



Panduan Pengguna

AWS Clean Rooms



AWS Clean Rooms: Panduan Pengguna

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Merek dagang dan tampilan dagang Amazon tidak boleh digunakan sehubungan dengan produk atau layanan apa pun yang bukan milik Amazon, dengan cara apa pun yang dapat menyebabkan kebingungan di antara pelanggan, atau dengan cara apa pun yang merendahkan atau mendiskreditkan Amazon. Semua merek dagang lain yang tidak dimiliki oleh Amazon merupakan hak milik masing-masing pemiliknya, yang mungkin atau mungkin tidak terafiliasi, terkait dengan, atau disponsori oleh Amazon.

Table of Contents

Apa itu AWS Clean Rooms?	1
Apakah Anda AWS Clean Rooms pengguna pertama kali?	2
Bagaimana cara AWS Clean Rooms kerja	2
Layanan terkait	2
AWS layanan	2
Layanan pihak ketiga	4
Mengakses AWS Clean Rooms	4
Harga untuk AWS Clean Rooms	5
Tagihan untuk AWS Clean Rooms	5
Aturan analisis	6
Jenis aturan analisis	7
Aturan analisis agregasi	9
Aturan analisis daftar	29
Aturan analisis kustom	38
Aturan analisis tabel pemetaan ID	45
AWS Clean Rooms Privasi Diferensial	55
Privasi diferensial	56
Bagaimana Privasi Diferensial bekerja AWS Clean Rooms	56
Kebijakan privasi diferensial	57
Kemampuan SQL	59
Kiat dan contoh kueri SQL	88
Batasan	89
AWS Clean Room ML	91
Terminologi AWS Clean Rooms	91
Cara AWS Clean Rooms MLbekerja dengan AWS model	92
Cara AWS Clean Rooms MLbekerja dengan model khusus	93
AWS model dalam Kamar Bersih	95
Model kustom di Kamar Bersih ML	103
Komputasi kriptografi	111
Pertimbangan-pertimbangan	112
Jenis file dan data yang didukung	115
Nama kolom	120
Jenis kolom	121
Parameter	123

Bendera opsional	128
Kueri dengan C3R	131
Pedoman	132
Menyiapkan AWS Clean Rooms	159
Mendaftar untuk AWS	159
Siapkan peran layanan untuk AWS Clean Rooms	159
Buat pengguna administrator	160
Buat peran IAM untuk anggota kolaborasi	161
Membuat peran layanan untuk membaca data dari Amazon S3	162
Membuat peran layanan untuk membaca data dari Amazon Athena	166
Buat peran layanan untuk membaca data dari Snowflake	169
Buat peran layanan untuk membaca kode dari bucket S3 (peran template PySpark analisis)	174
Buat peran layanan untuk menulis hasil PySpark pekerjaan	177
Buat peran layanan untuk menerima hasil	180
Mengatur peran layanan untuk AWS Clean Rooms ML	184
Siapkan peran layanan untuk pemodelan mirip	184
Siapkan peran layanan untuk pemodelan khusus	198
Kolaborasi dan keanggotaan	214
Menciptakan kolaborasi	215
Membuat kolaborasi untuk kueri	215
Membuat kolaborasi untuk pertanyaan dan pekerjaan	226
Membuat kolaborasi untuk pemodelan ML	238
Membuat keanggotaan dan bergabung dengan kolaborasi	247
Mengedit kolaborasi	254
Edit nama dan deskripsi kolaborasi	254
Perbarui mesin analitik kolaborasi	255
Matikan penyimpanan log	255
Mengedit setelan log kolaborasi	256
Edit tag kolaborasi	257
Edit tag keanggotaan	258
Menambahkan anggota baru	258
Edit kemampuan anggota yang ada	259
Edit pengaturan persetujuan otomatis kolaborasi	259
Mengedit tag tabel terkait	259
Edit tag templat analisis	259

Edit tag kebijakan privasi diferensial	260
Perubahan permintaan	261
Menambahkan anggota baru ke kolaborasi	261
Perbarui kemampuan anggota yang ada	262
Edit pengaturan persetujuan otomatis kolaborasi	263
Menghapus kolaborasi	264
Melihat kolaborasi	265
Mengundang anggota	265
Memantau anggota	266
Menambahkan anggota	266
Menghapus anggota	267
Meninggalkan Kolaborasi	268
Tabel data	270
Format data	271
Format data yang didukung untuk PySpark pekerjaan	271
Format data yang didukung untuk kueri SQL	271
Jenis data yang didukung	272
Jenis kompresi file untuk AWS Clean Rooms	273
Enkripsi sisi server untuk AWS Clean Rooms	273
Apache Iceberg tabel	274
Tipe data yang didukung untuk tabel Iceberg	275
Mempersiapkan tabel data	276
Mempersiapkan tabel data di Amazon S3	276
Mempersiapkan tabel data di Amazon Athena	279
Mempersiapkan tabel data di Snowflake	281
Mempersiapkan tabel data terenkripsi	283
Langkah 1: Selesaikan prasyarat	284
Langkah 2: Unduh klien enkripsi C3R	285
Langkah 3: (Opsional) Lihat perintah yang tersedia di klien enkripsi C3R	285
Langkah 4: Buat skema enkripsi untuk file tabular	286
Langkah 5: Buat kunci rahasia bersama	293
Langkah 6: Simpan kunci rahasia bersama dalam variabel lingkungan	295
Langkah 7: Enkripsi data	296
Langkah 8: Verifikasi enkripsi data	297
(Opsional) Buat skema (pengguna tingkat lanjut)	298
Mendekripsi tabel data	307

Tabel yang dikonfigurasi	309
Membuat tabel yang dikonfigurasi	310
Sumber data Amazon S3	310
Sumber data Amazon Athena	314
Sumber data kepingan salju	317
Menambahkan aturan analisis ke tabel yang dikonfigurasi	321
Menambahkan aturan analisis agregasi ke tabel (aliran terpandu)	322
Menambahkan aturan analisis daftar ke tabel (alur terpandu)	326
Menambahkan aturan analisis kustom ke tabel (alur terpandu)	328
Menambahkan aturan analisis ke tabel (editor JSON)	332
Langkah selanjutnya	334
Mengaitkan tabel terkonfigurasi ke kolaborasi	334
Anggaran akses data	335
Kaitkan tabel yang dikonfigurasi	335
Mengonfigurasi anggaran akses data	341
Melihat anggaran akses data	342
Menambahkan anggaran akses data ke tabel terkait yang ada	342
Mengedit anggaran akses data	344
Menghapus anggaran akses data	345
Menambahkan aturan analisis kolaborasi ke tabel yang dikonfigurasi	346
Mengkonfigurasi kebijakan privasi diferensial (opsional)	348
Melihat log penggunaan privasi diferensial	349
Mengedit kebijakan privasi diferensial	349
Menghapus kebijakan privasi diferensial	350
Melihat parameter privasi diferensial yang dihitung	351
Melihat tabel dan aturan analisis	352
Mengedit tabel yang dikonfigurasi	353
Mengedit tag tabel yang dikonfigurasi	353
Mengedit aturan analisis tabel yang dikonfigurasi	354
Menghapus aturan analisis tabel yang dikonfigurasi	355
Kolom tabel yang dikonfigurasi tidak diizinkan	355
Mengedit asosiasi tabel yang dikonfigurasi	359
Memutuskan tabel yang dikonfigurasi	359
Resolusi Entitas AWS di AWS Clean Rooms	361
Ruang nama ID	362
Membuat dan mengaitkan namespace ID baru	362

Mengaitkan namespace ID yang ada	365
Mengedit asosiasi namespace ID	367
Memutuskan asosiasi namespace ID	368
Tabel pemetaan ID	369
Membuat dan mengisi tabel pemetaan ID baru	370
Mengisi tabel pemetaan ID yang ada	378
Mengedit tabel pemetaan ID	379
Menghapus tabel pemetaan ID	379
Template analisis	381
Templat analisis SQL	381
Membuat templat analisis SQL	382
Meninjau template analisis SQL	385
PySpark template analisis	386
Keamanan	387
Batasan	387
Praktik terbaik	388
Membuat skrip pengguna	390
Bekerja dengan parameter dalam templat PySpark analisis	393
Membuat lingkungan virtual (opsional)	397
Menyimpan skrip pengguna dan lingkungan virtual di S3	398
Membuat template PySpark analisis	400
Meninjau template PySpark analisis	404
Templat analisis pemecahan masalah PySpark	407
Memecahkan masalah kode Anda	408
Pekerjaan template analisis tidak dimulai	408
Pekerjaan template analisis dimulai tetapi gagal selama pemrosesan	410
Pengaturan lingkungan virtual gagal	411
Analisis	413
Menjalankan kueri SQL	413
Prasyarat	414
Konfigurasi properti percikan untuk kueri SQL	415
Menanyakan tabel yang dikonfigurasi	416
Menanyakan tabel pemetaan ID	435
Kueri tabel yang dikonfigurasi menggunakan templat analisis	438
Menanyakan dengan pembuat analisis	453
Melihat dampak privasi diferensial	459

Melihat kueri terbaru	460
Melihat detail kueri	461
Menjalankan PySpark pekerjaan	461
Menjalankan pekerjaan menggunakan template analisis	462
Melihat pekerjaan terbaru	465
Melihat detail tugas	465
Hasil analisis	467
Menerima hasil kueri	468
Menerima hasil pekerjaan	469
Mengedit nilai default untuk pengaturan hasil kueri	470
Mengedit nilai default untuk pengaturan hasil pekerjaan	472
Menggunakan output kueri di lain Layanan AWS	472
Pemodelan khusus	474
Pembuatan dataset sintetis yang ditingkatkan privasi	475
Pertimbangan untuk pembuatan data sintetis	476
Membuat dan bergabung dengan kolaborasi	479
Menciptakan kolaborasi untuk pembelajaran mesin	479
Bergabung dengan kolaborasi	483
Menyumbang data pelatihan	485
Mengkonfigurasi algoritma model	491
Mengaitkan algoritma model yang dikonfigurasi	493
Membuat saluran input ML	498
Membuat model terlatih	518
Menggunakan pelatihan inkremental	521
Menggunakan pelatihan terdistribusi	525
Mengekspor artefak model	529
Menjalankan inferensi pada model terlatih	531
Pemodelan ML untuk penyedia data pelatihan	534
Mengimpor data pelatihan	535
Membuat model yang mirip	536
Mengkonfigurasi model yang mirip	537
Mengaitkan model mirip yang dikonfigurasi	538
Memperbarui model mirip yang dikonfigurasi	539
Pemodelan ML untuk penyedia data benih	541
Membuat segmen yang mirip	541
Mengekspor segmen yang mirip	556

Pemecahan masalah	558
Satu atau beberapa tabel yang direferensikan oleh kueri tidak dapat diakses oleh peran layanan terkait. table/role Pemilik harus memberikan akses peran layanan ke tabel.	558
Salah satu kumpulan data yang mendasarinya memiliki format file yang tidak didukung.	558
Hasil kueri tidak seperti yang diharapkan saat menggunakan Komputasi Kriptografi untuk Clean Rooms.	559
AWS Clean Rooms Spark SQL: Data partisi tidak ada	559
Keamanan	560
Perlindungan data	561
Enkripsi saat diam	562
Enkripsi saat bergerak	563
Mengkripsi data yang mendasarinya	563
Kebijakan kunci	563
Menggunakan Peran Terkait Layanan	570
Izin peran terkait layanan untuk AWS Clean Rooms	570
Membuat peran terkait layanan untuk AWS Clean Rooms	571
Mengedit peran terkait layanan untuk AWS Clean Rooms	571
Menghapus peran terkait layanan untuk AWS Clean Rooms	571
Wilayah yang Didukung untuk AWS Clean Rooms peran terkait layanan	572
Retensi data	572
Praktik terbaik	573
Praktik terbaik dengan AWS Clean Rooms	573
Praktik terbaik untuk menggunakan aturan analisis di AWS Clean Rooms	573
Identity and Access Management	575
Audiens	576
Mengautentikasi dengan identitas	576
Mengelola akses menggunakan kebijakan	578
Bagaimana AWS Clean Rooms bekerja dengan IAM	580
Contoh kebijakan berbasis identitas	586
AWS kebijakan terkelola	589
Pemecahan masalah	597
Pencegahan "confused deputy" lintas layanan	599
Perilaku IAM untuk AWS Clean Rooms MS	601
Perilaku IAM untuk Kamar Bersih Model Kustom	605
Validasi kepatuhan	607
Ketahanan	607

Keamanan infrastruktur	608
Keamanan jaringan	608
AWS PrivateLink	609
Pertimbangan-pertimbangan	609
Membuat sebuah titik akhir antarmuka	609
Memantau	611
Analisis masuk AWS Clean Rooms	612
Menerima kueri dan log pekerjaan	613
Tindakan yang disarankan untuk kueri dan log pekerjaan	614
Pemantauan terperinci dengan CloudWatch in AWS Clean Rooms	614
Tipe metrik	615
Dimensi metrik	616
Siapa yang dapat mengakses metrik	617
Penetapan harga	617
CloudTrail log	617
AWS Clean Rooms informasi di CloudTrail	618
Memahami entri file AWS Clean Rooms log	619
Contoh AWS Clean Rooms CloudTrail peristiwa	619
Mengintegrasikan ke dalam AWS Clean Rooms EDAs	622
AWS Clean Rooms acara	623
Acara perutean AWS Clean Rooms	627
Referensi detail acara	628
Penandaan Alokasi Biaya	638
Sumber Daya Taggable untuk Alokasi Biaya	638
AWS CloudFormation sumber daya	640
AWS Clean Rooms dan CloudFormation template	640
Pelajari lebih lanjut tentang CloudFormation	640
Kuota	642
AWS Clean Rooms kuota	642
AWS Clean Rooms batas parameter sumber daya	658
Kuota AWS Clean Rooms	658
Kuota pembatasan API Clean Rooms	691
Riwayat dokumen	698
Glosarium	714
Aturan analisis agregasi	714
Aturan analisis	714

Template analisis	714
Klien enkripsi C3R	715
Kolom Cleartext	715
Kolaborasi	715
Pencipta kolaborasi	715
Tabel yang dikonfigurasi	716
Aturan analisis kustom	716
Dekripsi	716
Privasi diferensial	717
Enkripsi	717
Kolom sidik jari	717
Metode alur kerja pemetaan ID	717
Tabel pemetaan ID	717
Aturan analisis tabel pemetaan ID	718
Alur kerja pemetaan ID	718
Ruang nama ID	718
Asosiasi namespace ID	718
Pekerjaan	718
Aturan analisis daftar	719
Model mirip	719
Segmen mirip	719
Anggota	719
Anggota yang dapat menanyakan	719
Anggota yang dapat menjalankan kueri dan pekerjaan	719
Anggota yang dapat menerima hasil	720
Anggota membayar biaya komputasi kueri	720
Anggota membayar biaya kueri dan komputasi pekerjaan	720
Keanggotaan	721
Kolom tertutup	721
Data benih	721
Mesin analitik percikan	721
Kueri	721
.....	dccxxiii

Apa itu AWS Clean Rooms?

AWS Clean Rooms membantu Anda dan mitra Anda menganalisis dan berkolaborasi dalam kumpulan data kolektif Anda untuk mendapatkan wawasan baru tanpa mengungkapkan data yang mendasarinya satu sama lain. AWS Clean Rooms adalah ruang kerja kolaborasi yang aman, tempat Anda membuat kamar bersih sendiri dalam hitungan menit, dan menganalisis kumpulan data kolektif Anda hanya dengan beberapa langkah. Anda memilih mitra dengan siapa Anda ingin berkolaborasi, memilih kumpulan data mereka, dan mengonfigurasi kontrol peningkatan privasi untuk mitra tersebut.

Dengan AWS Clean Rooms, Anda dapat berkolaborasi dengan ribuan perusahaan yang sudah menggunakan AWS. Kolaborasi tidak memerlukan pemindahan data dari AWS atau memuatnya ke penyedia layanan cloud lain. Saat Anda menjalankan kueri atau pekerjaan, AWS Clean Rooms membaca data dari lokasi asli data tersebut dan menerapkan aturan analisis bawaan untuk membantu Anda mempertahankan kendali atas data tersebut.

AWS Clean Rooms menyediakan kontrol akses data bawaan dan kontrol dukungan audit yang dapat Anda konfigurasi. Kontrol ini meliputi:

- [Aturan analisis](#) untuk membatasi kueri SQL dan memberikan kendala keluaran.
- [Komputasi Kriptografi Clean Rooms untuk](#) menjaga data terenkripsi, bahkan saat kueri diproses, untuk mematuhi kebijakan penanganan data yang ketat.
- [Log analisis](#) untuk meninjau kueri dan pekerjaan di AWS Clean Rooms dan membantu mendukung audit.
- [Privasi diferensial](#) untuk melindungi dari upaya identifikasi pengguna. AWS Clean Rooms Privasi Diferensial adalah kemampuan yang dikelola sepenuhnya yang melindungi privasi pengguna Anda dengan teknik yang didukung secara matematis dan kontrol intuitif yang dapat Anda terapkan dalam beberapa langkah.
- [AWS Clean Rooms](#) untuk memungkinkan dua pihak mengidentifikasi pengguna serupa dalam data mereka tanpa perlu berbagi data satu sama lain. Pihak pertama membuat dan mengonfigurasi model yang mirip dari data pelatihan mereka. Kemudian, data benih dibawa ke kolaborasi untuk membuat segmen mirip yang menyerupai data pelatihan.

Video berikut menjelaskan lebih lanjut tentang AWS Clean Rooms.

[AWS Clean Rooms](#)

Apakah Anda AWS Clean Rooms pengguna pertama kali?

Jika Anda adalah pengguna pertama kali AWS Clean Rooms, kami sarankan Anda mulai dengan membaca bagian berikut:

- [Bagaimana cara AWS Clean Rooms kerja](#)
- [Mengakses AWS Clean Rooms](#)
- [Menyiapkan AWS Clean Rooms](#)
- [AWS Clean Rooms Glosarium](#)

Bagaimana cara AWS Clean Rooms kerja

Di AWS Clean Rooms, Anda membuat kolaborasi dan menambahkan Akun AWS yang ingin Anda undang, atau membuat keanggotaan untuk bergabung dengan kolaborasi yang telah diundang. Anda kemudian menautkan sumber daya data yang diperlukan untuk kasus penggunaan Anda: tabel yang dikonfigurasi untuk data peristiwa, model yang dikonfigurasi untuk pemodelan ML, atau ruang nama ID untuk resolusi entitas. Anda memiliki opsi untuk membuat atau menyetujui templat analisis untuk menyetujui terlebih dahulu pada pertanyaan dan pekerjaan yang tepat yang ingin Anda izinkan dalam kolaborasi. Terakhir, Anda menganalisis data bersama dengan menjalankan kueri atau PySpark pekerjaan SQL pada tabel yang dikonfigurasi, melakukan resolusi entitas dalam tabel pemetaan ID, atau menggunakan pemodelan ML untuk menghasilkan segmen audiens yang mirip.

Diagram berikut menunjukkan cara AWS Clean Rooms kerja.

Layanan terkait

AWS layanan

Layanan AWS Berikut ini terkait dengan AWS Clean Rooms:

- Amazon Athena

Anggota kolaborasi dapat menyimpan data yang mereka bawa AWS Clean Rooms sebagai AWS Glue Data Catalog tampilan di Amazon Athena. Untuk informasi selengkapnya, lihat topik berikut:

Untuk informasi selengkapnya, lihat topik berikut:

[Mempersiapkan tabel data untuk kueri di AWS Clean Rooms](#)

[Membuat tabel yang dikonfigurasi - sumber data Amazon Athena](#)

[Apa itu Amazon Athena?](#) di Panduan Pengguna Amazon Athena

- CloudFormation

Buat sumber daya berikut di CloudFormation: kolaborasi, tabel yang dikonfigurasi, asosiasi tabel yang dikonfigurasi, dan keanggotaan

Untuk informasi selengkapnya, lihat [Menciptakan AWS Clean Rooms sumber daya dengan AWS CloudFormation](#).

- AWS CloudTrail

Gunakan AWS Clean Rooms dengan CloudTrail log untuk meningkatkan analisis Layanan AWS aktivitas Anda.

Untuk informasi selengkapnya, lihat [Pencatatan panggilan AWS Clean Rooms API menggunakan AWS CloudTrail](#).

- Resolusi Entitas AWS

Gunakan AWS Clean Rooms dengan Resolusi Entitas AWS untuk melakukan resolusi entitas.

Untuk informasi selengkapnya, lihat [Resolusi Entitas AWS di AWS Clean Rooms](#).

- AWS Glue

Anggota kolaborasi dapat membuat AWS Glue tabel dari data mereka di Amazon S3 untuk digunakan. AWS Clean Rooms

Untuk informasi selengkapnya, lihat topik berikut:

[Mempersiapkan tabel data untuk kueri di AWS Clean Rooms](#)

[Apa itu AWS Glue?](#) di Panduan Developer AWS Glue

- Amazon Simple Storage Service (Amazon S3)

Anggota kolaborasi dapat menyimpan data yang mereka bawa AWS Clean Rooms di Amazon S3.

Untuk informasi selengkapnya, lihat topik berikut:

[Mempersiapkan tabel data untuk kueri di AWS Clean Rooms](#)

[Membuat tabel yang dikonfigurasi - sumber data Amazon S3](#)

[Apa itu Amazon S3?](#) di Panduan Pengguna Layanan Penyimpanan Sederhana Amazon

- AWS Secrets Manager

Anggota kolaborasi dapat membuat rahasia untuk mengakses dan membaca data yang disimpan di Snowflake.

Untuk informasi selengkapnya, lihat topik berikut:

[Buat peran layanan untuk membaca data dari Snowflake](#)

[Mempersiapkan tabel data untuk kueri di AWS Clean Rooms](#)

[Apa itu AWS Secrets Manager?](#) dalam AWS Secrets Manager Panduan Pengguna

Layanan pihak ketiga

Layanan pihak ketiga berikut ini terkait dengan AWS Clean Rooms:

- Kepingan salju

Anggota kolaborasi dapat menyimpan data yang mereka bawa ke gudang AWS Clean Rooms Snowflake.

Untuk informasi selengkapnya, lihat topik berikut:

[Mempersiapkan tabel data untuk kueri di AWS Clean Rooms](#)

[Membuat tabel yang dikonfigurasi — Sumber data Snowflake](#)

Mengakses AWS Clean Rooms

Anda dapat mengakses AWS Clean Rooms melalui opsi berikut:

- Langsung melalui AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms/>.

- Secara terprogram melalui API. AWS Clean Rooms Untuk informasi lebih lanjut, lihat [Referensi API AWS Clean Rooms](#).

Harga untuk AWS Clean Rooms

Untuk informasi harga, lihat [Harga AWS Clean Rooms](#).

Note

Untuk anggota kolaborasi yang mengaitkan data yang disimpan di Snowflake, Anda akan dikenakan biaya oleh penyedia gudang data atau penyedia cloud masing-masing untuk keluar data dan komputasi setiap kali kueri dijalankan yang menggunakan data yang disimpan di lokasi tersebut.

Tagihan untuk AWS Clean Rooms

AWS Clean Rooms memberikan pencipta kolaborasi kemampuan untuk menunjuk anggota mana yang membayar untuk kueri atau biaya komputasi pekerjaan dalam kolaborasi.

Dalam kebanyakan kasus, [anggota yang dapat menanyakan](#) dan [anggota yang membayar biaya komputasi kueri](#) adalah sama. Namun, jika anggota yang dapat melakukan kueri dan anggota yang membayar biaya komputasi kueri berbeda, maka, ketika anggota yang dapat melakukan kueri menjalankan kueri terhadap sumber daya keanggotaan mereka sendiri, sumber daya keanggotaan anggota yang membayar biaya komputasi kueri akan ditagih.

Anggota yang membayar biaya komputasi kueri tidak melihat peristiwa apa pun untuk kueri yang dijalankan dalam riwayat CloudTrail Acara mereka karena pembayar bukanlah orang yang menjalankan kueri maupun pemilik sumber daya tempat kueri dijalankan. Namun, pembayar memang melihat biaya yang dihasilkan pada sumber daya keanggotaan mereka untuk semua kueri yang dijalankan oleh anggota yang dapat menjalankan kueri dalam kolaborasi.

Untuk informasi selengkapnya tentang cara membuat kolaborasi dan mengonfigurasi anggota yang membayar biaya komputasi kueri, lihat [Menciptakan kolaborasi](#).

Aturan analisis di AWS Clean Rooms

Sebagai bagian dari mengaktifkan tabel untuk digunakan AWS Clean Rooms untuk analisis kolaborasi, anggota kolaborasi harus mengonfigurasi aturan analisis.

Aturan analisis adalah kontrol peningkatan privasi yang disiapkan oleh setiap pemilik data pada tabel yang dikonfigurasi. Aturan analisis menentukan bagaimana tabel yang dikonfigurasi dapat dianalisis.

Aturan analisis adalah kontrol tingkat akun pada tabel yang dikonfigurasi (sumber daya tingkat akun) dan diberlakukan dalam kolaborasi apa pun di mana tabel yang dikonfigurasi dikaitkan. Jika tidak ada aturan analisis yang dikonfigurasi, tabel yang dikonfigurasi dapat dikaitkan dengan kolaborasi tetapi tidak dapat ditanyakan. Kueri hanya dapat mereferensikan tabel yang dikonfigurasi dengan jenis aturan analisis yang sama.

Untuk mengonfigurasi aturan analisis, pertama-tama Anda memilih jenis analisis dan kemudian menentukan aturan analisis. Untuk kedua langkah tersebut, Anda harus mempertimbangkan kasus penggunaan yang ingin Anda aktifkan dan bagaimana Anda ingin melindungi data yang mendasarinya.

AWS Clean Rooms memberlakukan kontrol yang lebih ketat di semua tabel yang dikonfigurasi yang direferensikan dalam kueri.

Contoh berikut menggambarkan kontrol restriktif.

Example Kontrol restriktif: Kendala keluaran

- Kolaborator A memiliki kendala keluaran pada kolom pengidentifikasi 100.
- Kolaborator B memiliki kendala keluaran pada kolom pengidentifikasi 150.

Kueri agregasi yang mereferensikan kedua tabel yang dikonfigurasi memerlukan setidaknya 150 nilai pengidentifikasi yang berbeda dalam baris keluaran agar dapat ditampilkan dalam output kueri. Output kueri tidak menunjukkan bahwa hasil dihapus karena kendala keluaran.

Example Kontrol restriktif: Template analisis tidak disetujui

- Kolaborator A telah mengizinkan templat analisis dengan kueri yang mereferensikan tabel yang dikonfigurasi dari Kolaborator A dan Kolaborator B dalam aturan analisis kustom mereka.
- Kolaborator B tidak mengizinkan template analisis.

Karena Collaborator B tidak mengizinkan templat analisis, anggota yang dapat melakukan kueri tidak dapat menjalankan templat analisis tersebut.

Jenis aturan analisis

Ada tiga jenis aturan analisis: [agregasi](#), [daftar](#), dan [kustom](#). Tabel berikut membandingkan jenis aturan analisis. Setiap jenis memiliki bagian terpisah yang menjelaskan menentukan aturan analisis.

Note

Ada jenis aturan analisis yang disebut aturan analisis tabel pemetaan ID. Namun, aturan analisis ini dikelola oleh AWS Clean Rooms dan tidak dapat dimodifikasi. Untuk informasi selengkapnya, lihat [Aturan analisis tabel pemetaan ID](#).

Bagian berikut menjelaskan kasus penggunaan dan kontrol yang didukung untuk setiap jenis aturan analisis.

Kasus penggunaan yang didukung

Tabel berikut menunjukkan ringkasan perbandingan kasus penggunaan yang didukung untuk setiap jenis aturan analisis.

Kasus penggunaan	Agregasi	Daftar	Kustom
Analisis yang didukung	Kueri yang menggabungkan statistik menggunakan fungsi COUNT, SUM, dan AVG sepanjang dimensi opsional	Kueri yang menampilkan daftar tingkat baris dari tumpang tindih antara beberapa tabel	Analisis kustom apa pun selama templat analisis atau pembuat analisis telah ditinjau dan diizinkan
Kasus penggunaan umum	Analisis segmen, pengukuran, atribusi	Pengayaan, pembangunan segmen	Atribusi sentuhan pertama, analisis inkremental, penemuan audiens

Kasus penggunaan	Agregasi	Daftar	Kustom
Konstruksi SQL	• Pernyataan JOIN: INNER JOIN • Fungsi agregat: COUNT/COUNT DISTINCT, SUM/ SUM DISTINCT, dan AVG • Fungsi skalar: Subset terbatas	• Pernyataan JOIN: INNER JOIN • Fungsi skalar: Tidak ada	Mayoritas fungsi SQL dan konstruksi SQL tersedia dengan perintah SELECT
Subquery dan ekspresi tabel umum () CTEs	Tidak	Tidak	Ya
Template analisis	Tidak	Tidak	Ya

Kontrol yang didukung

Tabel berikut menunjukkan ringkasan perbandingan tentang bagaimana setiap jenis aturan analisis melindungi data dasar Anda.

Pengendalian	Agregasi	Daftar	Kustom
Mekanisme kontrol	Kontrol bagaimana data dalam tabel dapat digunakan dalam kueri (Misalnya, izinkan COUNT dan SUM kolom hashed_email.)	Kontrol bagaimana data dalam tabel dapat digunakan dalam kueri (Misalnya, izinkan penggunaan kolom hashed_email hanya untuk bergabung.)	Kontrol kueri apa yang diizinkan untuk berjalan di atas meja (Misalnya, izinkan hanya kueri yang ditentukan dalam templat analisis "Kueri khusus 1".)
Teknik peningkatan privasi bawaan	• Pertandingan buta • Diperlukan agregasi	• Pertandingan buta	• Privasi diferensial

Pengendalian	Agregasi	Daftar	Kustom
	- Ambang agregasi min $\geq$ 2 Struktur kueri yang telah ditentukan sebelumnya	- Diperlukan tumpang tindih - Struktur kueri yang telah ditentukan sebelumnya - Analisis tambahan yang diizinkan	Kolom keluaran yang tidak diizinkan
Tinjau kueri sebelum dapat dijalankan	Tidak	Tidak	Ya, menggunakan templat analisis

Untuk informasi selengkapnya tentang aturan analisis yang tersedia AWS Clean Rooms, lihat topik berikut.

- [Aturan analisis agregasi](#)
- [Aturan analisis daftar](#)
- [Aturan analisis kustom di AWS Clean Rooms](#)

Aturan analisis agregasi

Dalam AWS Clean Rooms, aturan analisis agregasi menghasilkan statistik agregat menggunakan fungsi COUNT, SUM, and/or AVG sepanjang dimensi opsional. Ketika aturan analisis agregasi ditambahkan ke tabel yang dikonfigurasi, ini memungkinkan anggota yang dapat melakukan kueri untuk menjalankan kueri pada tabel yang dikonfigurasi.

Aturan analisis agregasi mendukung penggunaan kasus seperti perencanaan kampanye, jangkauan media, pengukuran frekuensi, dan atribusi.

Struktur kueri dan sintaks yang didukung didefinisikan dalam [Struktur kueri agregasi dan sintaks](#).

Parameter aturan analisis, yang didefinisikan dalam [Aturan analisis agregasi - kontrol kueri](#), termasuk kontrol kueri dan kontrol hasil kueri. Kontrol kuerinya mencakup kemampuan untuk mengharuskan tabel yang dikonfigurasi digabungkan ke setidaknya satu tabel yang dikonfigurasi yang dimiliki oleh anggota yang dapat melakukan kueri, baik secara langsung maupun transitif. Persyaratan ini

memungkinkan Anda untuk memastikan bahwa kueri dijalankan di persimpangan (INNERJOIN) tabel Anda dan mereka.

Struktur kueri agregasi dan sintaks

Kueri pada tabel yang memiliki aturan analisis agregasi harus mematuhi sintaks berikut.

```
--select_aggregate_function_expression
SELECT
aggregation_function(column_name) [[AS] column_alias ] [, ...]

--select_grouping_column_expression
[, {column_name|scalar_function(arguments)} [[AS] column_alias ]][, ...]

--table_expression
FROM table_name [[AS] table_alias ]
  [[INNER] JOIN table_name [[AS] table_alias] ON join_condition] [...]

--where_expression
[WHERE where_condition]

--group_by_expression
[GROUP BY {column_name|scalar_function(arguments)}, ...]]

--having_expression
[HAVING having_condition]

--order_by_expression
[ORDER BY {column_name|scalar_function(arguments)} [{ASC|DESC}]] [,...]]
```

Tabel berikut menjelaskan setiap ekspresi yang tercantum dalam sintaks sebelumnya.

Ekspresi	Definisi	Contoh
<i>select_aggregate_function_expression</i>	Daftar dipisahkan koma yang berisi ekspresi berikut: select_aggregation_function_expression	SELECT SUM(PRICE), user_segment

Ekspresi	Definisi	Contoh
	<code>select_aggregate_expression</code>  Note Harus ada setidaknya a satu <code>select_aggregation_function_expression</code> <code>diselect_aggregate_expression</code> .	
<i><code>select_aggregation_function_expression</code></i>	Satu atau lebih fungsi agregasi yang didukung diterapkan pada satu atau beberapa kolom. Hanya kolom yang diizinkan sebagai argumen fungsi agregasi.  Note Harus ada setidaknya a satu <code>select_aggregation_function_expression</code> <code>diselect_aggregate_expression</code> .	<code>AVG(PRICE)</code> <code>COUNT(DISTINCT user_id)</code>

Ekspresi	Definisi	Contoh
<i><code>select_grouping_column_expression</code></i>	Ekspresi yang dapat berisi ekspresi apa pun menggunakan berikut ini: • Nama kolom tabel • Fungsi skalar yang didukung • String literal • Literal numerik  Note <code>select_aggregate_expression</code> dapat alias kolom dengan atau tanpa AS parameter. Untuk informasi selengkapnya, lihat Referensi AWS Clean Rooms SQL.	<code>TRUNC(timestampColumn)</code> <code>UPPER(campaignName)</code>

Ekspresi	Definisi	Contoh
<i>table_expression</i>	Sebuah tabel, atau gabungan tabel, menghubungkan menggabungkan ekspresi kondisional dengan <code>join_condition</code> . <code>join_condition</code> mengembalikan Boolean. <code>table_expression</code> Dukungan: • <code>JOIN</code> Tipe tertentu (<code>INNER JOIN</code>) • Kondisi perbandingan kesetaraan dalam a <code>join_condition</code> () = • Operator logis (<code>AND</code>, <code>OR</code>).	<pre>FROM consumer_table INNER JOIN provider_ table ON consumer_table.ide ntifier1 = provider_ table.identifier1 AND consumer_table .identifier2 = provider_table.ide ntifier2</pre>

Ekspresi	Definisi	Contoh
<i>where_expression</i>	Ekspresi kondisional yang mengembalikan Boolean. Ini mungkin terdiri dari yang berikut: • Nama kolom tabel • Fungsi skalar yang didukung • Operator matematika • String literal • Literal numerik Kondisi perbandingan yang didukung adalah (=, >, <, <=, >=, <>, !=, NOT, IN, NOT IN, LIKE, IS NULL, IS NOT NULL). Operator logis yang didukung adalah (AND, OR). where_expression itu opsional.	<pre>WHERE where_condition WHERE price > 100 WHERE TRUNC(timestampColumn) = '1/1/2022' WHERE timestampColumn = timestampColumn2 - 14</pre>
<i>group_by_expression</i>	Daftar ekspresi yang dipisahkan koma yang cocok dengan persyaratan untuk select_grouping_column_expression	<pre>GROUP BY TRUNC(timestampColumn), UPPER(campaignName), segment</pre>

Ekspresi	Definisi	Contoh
<i>having_expression</i>	Ekspresi kondisional yang mengembalikan Boolean. Mereka memiliki fungsi agregasi yang didukung diterapkan ke satu kolom (misalnya, SUM(price)) dan dibandingkan dengan literal numerik. Kondisi yang didukung adalah (=, >, <, <=, >=, <>, !=). Operator logis yang didukung adalah (AND, OR). having_expression itu opsional.	HAVING SUM(SALES) > 500

Ekspresi	Definisi	Contoh
<i>order_by_expression</i>	Daftar ekspresi yang dipisahkan koma yang kompatibel dengan persyaratan yang sama yang didefinisikan dalam <code>select_aggregate_expression</code> didefinisikan sebelumnya. <code>order_by_expression</code> itu opsional.  Note <code>order_by_expression</code> izin ASC dan DESC parameter. Untuk informasi selengkapnya, lihat parameter ASC DESC di Referensi AWS Clean Rooms SQL.	ORDER BY SUM(SALES), UPPER(campaignName)

Untuk struktur kueri agregasi dan sintaks, perhatikan hal berikut:

- Perintah SQL selain SELECT tidak didukung.
- Sub-kueri dan ekspresi tabel umum (misalnya, WITH) tidak didukung.
- Operator yang menggabungkan beberapa kueri (misalnya, UNION) tidak didukung.
- TOP, LIMIT, dan OFFSET parameter tidak didukung.

Aturan analisis agregasi - kontrol kueri

Dengan kontrol kueri agregasi, Anda dapat mengontrol bagaimana kolom dalam tabel Anda digunakan untuk menanyakan tabel. Misalnya, Anda dapat mengontrol kolom mana yang digunakan untuk bergabung, kolom mana yang dapat dihitung, atau kolom mana yang dapat digunakan dalam WHERE pernyataan.

Bagian berikut menjelaskan setiap kontrol.

Topik

- [Kontrol agregasi](#)
- [Bergabunglah dengan kontrol](#)
- [Kontrol dimensi](#)
- [Fungsi skalar](#)

Kontrol agregasi

Dengan menggunakan kontrol agregasi, Anda dapat menentukan fungsi agregasi mana yang akan diizinkan, dan kolom apa yang harus diterapkan. Fungsi agregasi dapat digunakan dalam SELECT, HAVING, dan ORDER BY ekspresi.

Kontrol	Definisi	Penggunaan
aggregateColumns	Kolom kolom tabel dikonfigurasi yang Anda izinkan untuk digunakan dalam fungsi agregasi.	aggregateColumns dapat digunakan di dalam fungsi agregasi di SELECT, HAVING, dan ORDER BY ekspresi. Beberapa juga aggregate Columns dapat dikategorikan sebagai joinColumn (didefinisikan nanti). Diberikan tidak aggregate Column dapat juga dikategorikan sebagai dimension Column (didefinisikan nanti).

Kontrol	Definisi	Penggunaan
<code>function</code>	Fungsi COUNT, SUM, dan AVG yang Anda izinkan untuk digunakan di atas. <code>aggregateColumns</code>	<code>function</code> dapat diterapkan pada <code>aggregateColumns</code> yang terkait dengannya.

Bergabunglah dengan kontrol

Sebuah JOIN klausa digunakan untuk menggabungkan baris dari dua atau lebih tabel, berdasarkan kolom terkait di antara mereka.

Anda dapat menggunakan kontrol Gabung untuk mengontrol bagaimana tabel Anda dapat digabungkan ke tabel lain di `table_expression`. AWS Clean Rooms hanya mendukung INNERJOIN. INNERJOIN pernyataan hanya dapat menggunakan kolom yang telah dikategorikan secara eksplisit sebagai `joinColumn` aturan analisis Anda, tunduk pada kontrol yang Anda tentukan.

INNERJOIN harus beroperasi pada `joinColumn` dari tabel yang dikonfigurasi dan `joinColumn` dari tabel lain yang dikonfigurasi dalam kolaborasi. Anda memutuskan kolom mana dari tabel Anda dapat digunakan sebagai `joinColumn`.

Setiap kondisi kecocokan dalam ON klausa diperlukan untuk menggunakan kondisi perbandingan kesetaraan (=) antara dua kolom.

Beberapa kondisi pertandingan dalam suatu ON klausa dapat berupa:

- Dikombinasikan menggunakan operator AND logis
- Dipisahkan menggunakan operator OR logis

Note

Semua kondisi JOIN pertandingan harus cocok dengan satu baris dari setiap sisi JOIN. Semua kondisional yang dihubungkan oleh OR atau operator AND logis harus mematuhi persyaratan ini juga.

Berikut ini adalah contoh dari query dengan operator AND logis.

```
SELECT some_col, other_col
FROM table1
  JOIN table2
  ON table1.id = table2.id AND table1.name = table2.name
```

Berikut ini adalah contoh dari query dengan operator OR logis.

```
SELECT some_col, other_col
FROM table1
  JOIN table2
  ON table1.id = table2.id OR table1.name = table2.name
```

Kontrol	Definisi	Penggunaan
joinColumns	Kolom (jika ada) yang ingin Anda izinkan anggota yang dapat kueri untuk digunakan dalam INNER JOIN pernyataan.	Spesifik juga joinColumn dapat dikategorikan sebagai aggregateColumn (lihat Kontrol agregasi). Kolom yang sama tidak dapat digunakan sebagai joinColumn dan dimensionColumns (lihat nanti). Kecuali itu juga telah dikategorikan sebagai aggregateColumn, a tidak joinColumn dapat digunakan di bagian lain dari kueri selain. INNER JOIN
joinRequired	Kontrol apakah Anda memerlukan tabel INNER JOIN yang dikonfigurasi dari anggota yang dapat melakukan kueri.	Jika Anda mengaktifkan parameter ini, INNER JOIN diperlukan. Jika Anda tidak mengaktifkan parameter ini, INNER JOIN adalah opsional.

Kontrol	Definisi	Penggunaan
		Dengan asumsi Anda mengaktifkan parameter ini, anggota yang dapat melakukan kueri diminta untuk menyertakan tabel yang mereka miliki di INNERJOIN . Mereka harus meja JOIN Anda dengan mereka, baik secara langsung atau transitif (yaitu, menggabungkan meja mereka ke meja lain, yang dengan sendirinya bergabung dengan meja Anda).

Berikut ini adalah contoh transitivitas.

```
ON
my_table.identifer = third_party_table.identifier
....
ON
third_party_table.identifier = member_who_can_query_table.id
```

Note

Anggota yang dapat melakukan query juga dapat menggunakan `joinRequired` parameter. Dalam hal ini, kueri harus menggabungkan tabel mereka dengan setidaknya satu tabel lainnya.

Kontrol dimensi

Kontrol dimensi mengontrol kolom di mana kolom agregasi dapat disaring, dikelompokkan, atau digabungkan.

Kontrol	Definisi	Penggunaan
<code>dimensionColumns</code>	Kolom (jika ada) yang Anda izinkan anggota yang dapat kueri untuk digunakan <code>SELECT</code> , <code>WHERE</code> , <code>GROUPBY</code> , dan <code>ORDERBY</code> .	A <code>dimensionColumn</code> dapat digunakan dalam <code>SELECT (select_grouping_column_expression), WHERE, GROUPBY, dan ORDERBY</code>. Kolom yang sama tidak bisa berupa <code>adimensionColumn</code>, <code>ajoinColumn</code>, and/or <code>anaggregateColumn</code>.

Fungsi skalar

Fungsi skalar mengontrol fungsi skalar mana yang dapat digunakan pada kolom dimensi.

Kontrol	Definisi	Penggunaan
<code>scalarFunctions</code>	Fungsi skalar yang dapat digunakan <code>dimensionColumns</code> dalam kueri.	Menentukan fungsi skalar (jika ada) yang Anda izinkan (misalnya, <code>CAST</code>) untuk diterapkan pada <code>dimensionColumns</code>. Fungsi skalar tidak dapat digunakan di atas fungsi lain atau di dalam fungsi lainnya. Argumen fungsi skalar dapat berupa kolom, literal string, atau literal numerik.

Fungsi skalar berikut didukung:

- Fungsi matematika - ABS, CEILING, FLOOR, LOG, LN, ROUND, SQRT
- Fungsi pemformatan tipe data - CAST, CONVERT, TO_CHAR, TO_DATE, TO_NUMBER, TO_TIMESTAMP
- Fungsi string - BAWAH, ATAS, TRIM, RTRIM, SUBSTRING
 - Untuk RTRIM, set karakter khusus untuk dipangkas tidak diperbolehkan.
- Ekspresi bersyarat - COALESCE
- Fungsi tanggal - EXTRACT, GETDATE, CURRENT_DATE, DATEADD
- Fungsi lainnya - TRUNC

Untuk detail selengkapnya, lihat [Referensi AWS Clean Rooms SQL](#).

Aturan analisis agregasi - kontrol hasil kueri

Dengan kontrol hasil kueri agregasi, Anda dapat mengontrol hasil mana yang dikembalikan dengan menentukan satu atau beberapa kondisi yang harus dipenuhi oleh setiap baris keluaran agar dapat dikembalikan. AWS Clean Rooms mendukung kendala agregasi dalam bentuk. `COUNT (DISTINCT column) >= X` Formulir ini mengharuskan setiap baris menggabungkan setidaknya X nilai pilihan yang berbeda dari tabel Anda yang dikonfigurasi (misalnya, jumlah minimum `user_id` nilai yang berbeda). Ambang batas minimum ini secara otomatis diberlakukan, bahkan jika kueri yang dikirimkan itu sendiri tidak menggunakan kolom yang ditentukan. Mereka diberlakukan secara kolektif di setiap tabel yang dikonfigurasi dalam kueri dari tabel yang dikonfigurasi dari setiap anggota dalam kolaborasi.

Setiap tabel yang dikonfigurasi harus memiliki setidaknya satu batasan agregasi dalam aturan analisisnya. Pemilik tabel yang dikonfigurasi dapat menambahkan beberapa `columnName` dan terkait minimum dan diberlakukan secara kolektif.

Kendala agregasi

Batasan agregasi mengontrol baris mana dalam hasil kueri yang dikembalikan. Untuk dikembalikan, baris harus memenuhi jumlah minimum yang ditentukan dari nilai berbeda di setiap kolom yang ditentukan dalam batasan agregasi. Persyaratan ini berlaku bahkan jika kolom tidak disebutkan secara eksplisit dalam kueri atau di bagian lain dari aturan analisis.

Kontrol	Definisi	Penggunaan
columnName	aggregateColumn Yang digunakan dalam kondisi bahwa setiap baris output harus memenuhi.	Dapat berupa kolom apa pun di tabel yang dikonfigurasi.
minimum	Jumlah minimum nilai berbeda untuk yang terkait aggregate Column yang harus dimiliki baris keluaran (misalnya , COUNT DISTINCT) agar dapat dikembalikan dalam hasil kueri.	minimumHarus minimal nilai 2.

Struktur aturan analisis agregasi

Contoh berikut menunjukkan struktur yang telah ditetapkan untuk aturan analisis agregasi.

Dalam contoh berikut, *MyTable* mengacu pada tabel data Anda. Anda dapat mengganti masing-masing *user input placeholder* dengan informasi Anda sendiri.

```
{
  "aggregateColumns": [
    {
      "columnNames": [MyTable column names], "function": [Allowed Agg Functions]
    },
  ],
  "joinRequired": ["QUERY_RUNNER"],
  "joinColumns": [MyTable column names],
  "dimensionColumns": [MyTable column names],
  "scalarFunctions": [Allowed Scalar functions],
  "outputConstraints": [
    {
      "columnName": [MyTable column names], "minimum": [Numeric value]
    },
  ]
}
```

Aturan analisis agregasi - contoh

Contoh berikut menunjukkan bagaimana dua perusahaan dapat berkolaborasi dalam AWS Clean Rooms menggunakan analisis agregasi.

Perusahaan A memiliki data pelanggan dan penjualan. Perusahaan A tertarik untuk memahami aktivitas pengembalian produk. Perusahaan B adalah salah satu pengecer Perusahaan A dan memiliki data pengembalian. Perusahaan B juga memiliki atribut segmen pada pelanggan yang berguna bagi Perusahaan A (misalnya, membeli produk terkait, menggunakan layanan pelanggan dari pengecer). Perusahaan B tidak ingin memberikan data pengembalian pelanggan tingkat baris dan informasi atribut. Perusahaan B hanya ingin mengaktifkan serangkaian kueri untuk Perusahaan A untuk mendapatkan statistik agregat tentang pelanggan yang tumpang tindih pada ambang agregasi minimum.

Perusahaan A dan Perusahaan B memutuskan untuk berkolaborasi sehingga Perusahaan A dapat memahami aktivitas pengembalian produk dan memberikan produk yang lebih baik di Perusahaan B dan saluran lainnya.

Untuk membuat kolaborasi dan menjalankan analisis agregasi, perusahaan melakukan hal berikut:

1. Perusahaan A menciptakan kolaborasi dan menciptakan keanggotaan. Kolaborasi ini menjadikan Perusahaan B sebagai anggota lain dalam kolaborasi tersebut. Perusahaan A memungkinkan pencatatan kueri dalam kolaborasi, dan memungkinkan pencatatan kueri di akun mereka.
2. Perusahaan B menciptakan keanggotaan dalam kolaborasi. Ini memungkinkan pencatatan kueri di akunnya.
3. Perusahaan A membuat tabel penjualan yang dikonfigurasi.
4. Perusahaan A menambahkan aturan analisis agregasi berikut ke tabel yang dikonfigurasi penjualan.

```
{
  "aggregateColumns": [
    {
      "columnNames": [
        "identifier"
      ],
      "function": "COUNT_DISTINCT"
    },
    {
      "columnNames": [
        "purchases"
      ]
    }
  ]
}
```

```
    ],
    "function": "AVG"
  },
  {
    "columnNames": [
      "purchases"
    ],
    "function": "SUM"
  }
],
"joinColumns": [
  "hashedemail"
],
"dimensionColumns": [
  "demoseg",
  "purchasedate",
  "productline"
],
"scalarFunctions": [
  "CAST",
  "COALESCE",
  "TRUNC"
],
"outputConstraints": [
  {
    "columnName": "hashedemail",
    "minimum": 2,
    "type": "COUNT_DISTINCT"
  },
]
}
```

`aggregateColumnsPerusahaan A` ingin menghitung jumlah pelanggan unik dalam tumpang tindih antara data penjualan dan data pengembalian. Perusahaan A juga ingin menjumlahkan jumlah yang `purchases` dibuat untuk dibandingkan dengan `jumlahreturns`.

`joinColumns`—Perusahaan A ingin menggunakan `identifier` untuk mencocokkan pelanggan dari data penjualan ke pelanggan dari data pengembalian. Ini akan membantu perusahaan A `match` kembali ke pembelian yang tepat. Ini juga membantu segmen Perusahaan A tumpang tindih pelanggan.

dimensionColumns— Perusahaan A menggunakan **dimensionColumns** untuk memfilter berdasarkan produk tertentu, membandingkan pembelian dan pengembalian selama periode waktu tertentu, memastikan tanggal pengembalian setelah tanggal produk, dan membantu segmen pelanggan yang tumpang tindih.

scalarFunctions— Perusahaan A memilih fungsi CAST skalar untuk membantu memperbarui format tipe data jika diperlukan berdasarkan tabel yang dikonfigurasi Perusahaan A terkait dengan kolaborasi. Ini juga menambahkan fungsi skalar untuk membantu memformat kolom jika diperlukan.

outputConstraints— Perusahaan A menetapkan batasan output minimum. Tidak perlu membatasi hasil karena analisis diizinkan untuk melihat data tingkat baris dari tabel penjualan mereka

 Note

Perusahaan A tidak termasuk **joinRequired** dalam aturan analisis. Ini memberikan fleksibilitas bagi analisis mereka untuk menanyakan tabel penjualan saja.

5. Perusahaan B membuat tabel yang dikonfigurasi pengembalian.
6. Perusahaan B menambahkan aturan analisis agregasi berikut ke tabel pengembalian yang dikonfigurasi.

```
{
  "aggregateColumns": [
    {
      "columnNames": [
        "identifier"
      ],
      "function": "COUNT_DISTINCT"
    },
    {
      "columnNames": [
        "returns"
      ],
      "function": "AVG"
    },
    {
      "columnNames": [
        "returns"
      ]
    }
  ]
}
```

```

    ],
    "function": "SUM"
  }
],
"joinColumns": [
  "hashedemail"
],
"joinRequired": [
  "QUERY_RUNNER"
],
"dimensionColumns": [
  "state",
  "popularpurchases",
  "customerserviceuser",
  "productline",
  "returndate"
],
"scalarFunctions": [
  "CAST",
  "LOWER",
  "UPPER",
  "TRUNC"
],
"outputConstraints": [
  {
    "columnName": "hashedemail",
    "minimum": 100,
    "type": "COUNT_DISTINCT"
  },
  {
    "columnName": "producttype",
    "minimum": 2,
    "type": "COUNT_DISTINCT"
  }
]
}

```

aggregateColumns— Perusahaan B memungkinkan Perusahaan A untuk menjumlahkan **returns** untuk dibandingkan dengan jumlah pembelian. Mereka memiliki setidaknya satu kolom agregat karena mereka mengaktifkan kueri agregat.

joinColumns— Perusahaan B memungkinkan Perusahaan A untuk bergabung **identifier** untuk mencocokkan pelanggan dari data pengembalian ke pelanggan dari data penjualan.

`identifierdata` sangat sensitif dan memilikinya sebagai `joinColumn` memastikan bahwa data tidak akan pernah dikeluarkan dalam kueri.

`joinRequired`— Perusahaan B membutuhkan kueri pada data pengembalian agar tumpang tindih dengan data penjualan. Mereka tidak ingin mengaktifkan Perusahaan A untuk menanyakan semua individu dalam kumpulan data mereka. Mereka juga menyetujui pembatasan itu dalam perjanjian kolaborasi mereka.

`dimensionColumns`— Perusahaan B memungkinkan Perusahaan A untuk memfilter dan mengelompokkan berdasarkan `statepopularpurchases`, dan `customerserviceuser` yang merupakan atribut unik yang dapat membantu membuat analisis untuk Perusahaan A. Perusahaan B memungkinkan Perusahaan A untuk menggunakan `returndate` untuk menyaring output pada `returndate` yang terjadi setelahnya `purchasedate`. Dengan penyaringan ini, output lebih akurat untuk mengevaluasi dampak perubahan produk.

`scalarFunctions`— Perusahaan B memungkinkan hal-hal berikut:

- `TRUNC` untuk tanggal
- `LOWER` dan `UPPER` jika `producttype` dimasukkan dalam format yang berbeda dalam data mereka
- `CAST` jika Perusahaan A perlu mengonversi tipe data dalam penjualan agar sama dengan tipe data dalam pengembalian

Perusahaan A tidak mengaktifkan fungsi skalar lainnya karena mereka tidak percaya bahwa mereka diperlukan untuk kueri.

`outputConstraints`Perusahaan B menetapkan batasan output minimum `hashedemail` untuk membantu mengurangi kemampuan untuk mengidentifikasi kembali pelanggan. Ini juga menambahkan kendala keluaran minimum `producttype` untuk mengurangi kemampuan mengidentifikasi kembali produk tertentu yang dikembalikan. Jenis produk tertentu bisa lebih dominan berdasarkan dimensi output (misalnya, `state`). Kendala output mereka akan selalu diberlakukan terlepas dari kendala output yang ditambahkan oleh Perusahaan A ke data mereka.

7. Perusahaan A menciptakan asosiasi tabel penjualan untuk kolaborasi.
8. Perusahaan B menciptakan asosiasi tabel pengembalian untuk kolaborasi.
9. Perusahaan A menjalankan kueri, seperti contoh berikut, untuk lebih memahami jumlah pengembalian di Perusahaan B dibandingkan dengan total pembelian berdasarkan lokasi pada tahun 2022.

```
SELECT
  companyB.state,
  SUM(companyB.returns),
  COUNT(DISTINCT companyA.hashemail)
FROM
  sales companyA
  INNER JOIN returns companyB ON companyA.identifier = companyB.identifier
WHERE
  companyA.purchasedate BETWEEN '2022-01-01' AND '2022-12-31' AND
  TRUNC(companyB. returndate) > companyA.purchasedate
GROUP BY
  companyB.state;
```

10 Perusahan A dan Perusahan B meninjau log kueri. Perusahan B memverifikasi bahwa kueri sejalan dengan apa yang disepakati dalam perjanjian kolaborasi.

Memecahkan masalah aturan analisis agregasi

Gunakan informasi di sini untuk membantu Anda mendiagnosis dan memperbaiki masalah umum saat Anda bekerja dengan aturan analisis agregasi.

Masalah

- [Kueri saya tidak mengembalikan hasil apa pun](#)

Kueri saya tidak mengembalikan hasil apa pun

Hal ini dapat terjadi ketika tidak ada hasil yang cocok atau ketika hasil yang cocok tidak memenuhi satu atau lebih ambang agregasi minimum.

Untuk informasi selengkapnya tentang ambang agregasi minimum, lihat [Aturan analisis agregasi - contoh](#)

Aturan analisis daftar

Dalam AWS Clean Rooms, aturan analisis daftar menampilkan daftar tingkat baris tumpang tindih antara tabel yang dikonfigurasi yang ditambahkan dan tabel yang dikonfigurasi dari anggota yang dapat melakukan kueri. Anggota yang dapat melakukan kueri menjalankan kueri yang menyertakan aturan analisis daftar.

Jenis aturan analisis daftar mendukung penggunaan kasus seperti pengayaan dan pembangunan audiens.

Untuk informasi selengkapnya tentang struktur kueri dan sintaks yang telah ditentukan untuk aturan analisis ini, lihat. [Daftar aturan analisis struktur yang telah ditentukan](#)

Parameter aturan analisis daftar, didefinisikan dalam [Aturan analisis daftar - kontrol kueri](#), memiliki kontrol kueri. Kontrol kuerinya mencakup kemampuan untuk memilih kolom yang dapat dicantumkan dalam output. Kueri diperlukan untuk memiliki setidaknya satu gabungan dengan tabel yang dikonfigurasi dari anggota yang dapat melakukan kueri, baik secara langsung maupun transitif.

Tidak ada kontrol hasil kueri seperti yang ada untuk [aturan analisis Agregasi](#).

Kueri daftar hanya dapat menggunakan operator matematika. Mereka tidak dapat menggunakan fungsi lain (seperti agregasi atau skalar).

Topik

- [Daftar struktur kueri dan sintaks](#)
- [Aturan analisis daftar - kontrol kueri](#)
- [Daftar aturan analisis struktur yang telah ditentukan](#)
- [Aturan analisis daftar - contoh](#)

Daftar struktur kueri dan sintaks

Kueri pada tabel yang memiliki aturan analisis daftar harus mematuhi sintaks berikut.

```
--select_list_expression
SELECT DISTINCT column_name [[AS] column_alias ] [, ...]

--table_expression
FROM table_name [[AS] table_alias ]
    [[INNER] JOIN table_name [[AS] table_alias] ON join_condition] [...]

--where_expression
[WHERE where_condition]

--limit_expression
[LIMIT number]
```

Tabel berikut menjelaskan setiap ekspresi yang tercantum dalam sintaks sebelumnya.

Ekspresi	Definisi	Contoh
<i>select_list_expression</i>	Daftar dipisahkan koma yang berisi setidaknya satu nama kolom tabel. Diperlukan DISTINCT parameter.  Note Kolom alias <code>select_list_expression</code> kaleng dengan atau tanpa AS parameter. Untuk informasi selengkapnya, lihat Referensi AWS Clean Rooms SQL.	SELECT DISTINCT segment
<i>table_expression</i>	Sebuah tabel, atau gabungan tabel, dengan <code>join_condition</code> untuk menghubungkannya <code>join_condition</code>. <code>join_condition</code> mengembalikan Boolean. <code>table_expression</code> Dukungan: Jenis JOIN tertentu (INNERJOIN)	<pre>FROM consumer_table INNER JOIN provider_table ON consumer_table.identifier1 = provider_table.identifier1 AND consumer_table.identifier2 = provider_table.identifier2</pre>

Ekspresi	Definisi	Contoh
	- Kondisi perbandingan kesetaraan dalam a <code>join_condition () =</code> - Operator logis (AND,OR).	
<i>where_expression</i>	Ekspresi kondisional yang mengembalikan Boolean. Ini dapat terdiri dari yang berikut: - Nama kolom tabel - Operator matematika - String literal - Literal numerik Kondisi perbandingan yang didukung adalah (=, >, <, <=, >=, <>, !=, NOT, IN, NOT IN, LIKE, IS NULL, IS NOT NULL). Operator logika yang didukung adalah (AND, OR). <code>where_expression</code> itu opsional.	<pre>WHERE state + '_' + city = 'NY_NYC'</pre> <pre>WHERE timestampColumn = timestampColumn2 - 14</pre>
<i>limit_expression</i>	Ekspresi ini harus mengambil bilangan bulat positif. <code>limit_expression</code> itu opsional.	<code>LIMIT 100</code>

Untuk struktur kueri daftar dan sintaks, perhatikan hal berikut:

- Perintah SQL selain SELECT tidak didukung.

- Subkueri dan ekspresi tabel umum (misalnya, WITH) tidak didukung
- BYKlausul HAVING GROUPBY,, dan ORDER tidak didukung
- Parameter OFFSET tidak didukung

Aturan analisis daftar - kontrol kueri

Dengan kontrol kueri daftar, Anda dapat mengontrol bagaimana kolom dalam tabel Anda digunakan untuk menanyakan tabel. Misalnya, Anda dapat mengontrol kolom mana yang digunakan untuk bergabung, atau kolom mana yang dapat digunakan dalam pernyataan dan WHERE klausa SELECT.

Bagian berikut menjelaskan setiap kontrol.

Topik

- [Bergabunglah dengan kontrol](#)
- [Kontrol daftar](#)

Bergabunglah dengan kontrol

Dengan kontrol Gabung, Anda dapat mengontrol bagaimana tabel Anda dapat digabungkan ke tabel lain di table_expression. AWS Clean Rooms hanya mendukung INNER JOIN. Dalam aturan analisis daftar, setidaknya satu INNER JOIN diperlukan dan anggota yang dapat melakukan kueri diminta untuk menyertakan tabel yang mereka miliki di INNER JOIN. Ini berarti mereka harus menggabungkan meja Anda dengan meja mereka, baik secara langsung maupun transitif.

Berikut ini adalah contoh transitivitas.

```
ON
my_table.identifer = third_party_table.identifier
....
ON
third_party_table.identifier = member_who_can_query_table.id
```

INNERPernyataan JOIN hanya dapat menggunakan kolom yang secara eksplisit dikategorikan sebagai aturan analisis joinColumn Anda.

INNERJOIN harus beroperasi pada joinColumn dari tabel yang dikonfigurasi dan joinColumn dari tabel lain yang dikonfigurasi dalam kolaborasi. Anda memutuskan kolom mana dari tabel Anda dapat digunakan sebagai joinColumn.

Setiap kondisi kecocokan dalam ON klausa diperlukan untuk menggunakan kondisi perbandingan kesetaraan (=) antara dua kolom.

Beberapa kondisi pertandingan dalam suatu ON klausa dapat berupa:

- Dikombinasikan menggunakan operator AND logis
- Dipisahkan menggunakan operator OR logis

 Note

Semua kondisi JOIN pertandingan harus cocok dengan satu baris dari setiap sisi JOIN. Semua kondisional yang dihubungkan oleh OR atau operator AND logis harus mematuhi persyaratan ini juga.

Berikut ini adalah contoh dari query dengan operator AND logis.

```
SELECT some_col, other_col
FROM table1
  JOIN table2
    ON table1.id = table2.id AND table1.name = table2.name
```

Berikut ini adalah contoh dari query dengan operator OR logis.

```
SELECT some_col, other_col
FROM table1
  JOIN table2
    ON table1.id = table2.id OR table1.name = table2.name
```

Kontrol	Definisi	Penggunaan
<code>joinColumns</code>	Kolom yang ingin Anda izinkan anggota yang dapat kueri untuk digunakan dalam pernyataan INNER JOIN.	Kolom yang sama tidak dapat dikategorikan sebagai a <code>joinColumn</code> dan <code>listColumn</code> (lihat Kontrol daftar).

Kontrol	Definisi	Penggunaan
		<code>joinColumn</code> tidak dapat digunakan di bagian lain dari kueri selain INNER JOIN.

Kontrol daftar

Kontrol daftar mengontrol kolom yang dapat dicantumkan dalam output kueri (yaitu, digunakan dalam pernyataan SELECT) atau digunakan untuk memfilter hasil (yaitu, digunakan dalam WHERE pernyataan).

Kontrol	Definisi	Penggunaan
<code>listColumns</code>	Kolom yang Anda izinkan anggota yang dapat kueri untuk digunakan dalam SELECT dan WHERE	A <code>listColumn</code> dapat digunakan di SELECT dan WHERE. Kolom yang sama tidak dapat digunakan sebagai a <code>listColumn</code> dan <code>joinColumn</code> .

Daftar aturan analisis struktur yang telah ditentukan

Contoh berikut mencakup struktur yang telah ditentukan yang menunjukkan bagaimana Anda menyelesaikan aturan analisis daftar.

Dalam contoh berikut, *MyTable* mengacu pada tabel data Anda. Anda dapat mengganti masing-masing *user input placeholder* dengan informasi Anda sendiri.

```
{
  "joinColumns": [MyTable column name(s)],
  "listColumns": [MyTable column name(s)],
}
```

Aturan analisis daftar - contoh

Contoh berikut menunjukkan bagaimana dua perusahaan dapat berkolaborasi dalam AWS Clean Rooms menggunakan analisis daftar.

Perusahaan A memiliki data manajemen hubungan pelanggan (CRM). Perusahaan A ingin mendapatkan data segmen tambahan pada pelanggannya untuk mempelajari lebih lanjut tentang pelanggan mereka dan berpotensi menggunakan atribut sebagai masukan ke dalam analisis lain. Perusahaan B memiliki data segmen yang terdiri dari atribut segmen unik yang mereka buat berdasarkan data pihak pertama mereka. Perusahaan B ingin memberikan atribut segmen unik kepada Perusahaan A hanya pada pelanggan yang tumpang tindih antara data mereka dan data Perusahaan A.

Perusahaan memutuskan untuk berkolaborasi sehingga Perusahaan A dapat memperkaya data yang tumpang tindih. Perusahaan A adalah anggota yang dapat menanyakan, dan Perusahaan B adalah kontributor.

Untuk membuat kolaborasi dan menjalankan analisis daftar secara kolaborasi, perusahaan melakukan hal berikut:

1. Perusahaan A menciptakan kolaborasi dan menciptakan keanggotaan. Kolaborasi ini memiliki Perusahaan B sebagai anggota lain dalam kolaborasi tersebut. Perusahaan A memungkinkan pencatatan kueri dalam kolaborasi, dan memungkinkan pencatatan kueri di akunnya.
2. Perusahaan B menciptakan keanggotaan dalam kolaborasi. Ini memungkinkan pencatatan kueri di akunnya.
3. Perusahaan A membuat tabel yang dikonfigurasi CRM
4. Perusahaan A menambahkan aturan analisis ke tabel yang dikonfigurasi pelanggan, seperti yang ditunjukkan pada contoh berikut.

```
{
  "joinColumns": [
    "identifier1",
    "identifier2"
  ],
  "listColumns": [
    "internalid",
    "segment1",
    "segment2",
    "customercategory"
  ]
}
```

}

`joinColumns`—Perusahaan A ingin menggunakan `hashedemail` and/or `thirdpartyid` (diperoleh dari vendor identitas) untuk mencocokkan pelanggan dari data CRM ke pelanggan dari data segmen. Ini akan membantu memastikan Perusahaan A mencocokkan data yang diperkaya untuk pelanggan yang tepat. Mereka memiliki dua `JoinColumns` untuk berpotensi meningkatkan tingkat kecocokan analisis.

`listColumns`—Perusahaan A menggunakan `listColumns` untuk mendapatkan kolom yang diperkaya di samping yang `internalid` mereka gunakan dalam sistem mereka sendiri. Mereka menambahkan `segment1`, `segment2`, dan `customercategory` berpotensi membatasi pengayaan ke segmen tertentu dengan menggunakannya dalam filter.

5. Perusahaan B membuat tabel yang dikonfigurasi segmen.
6. Perusahaan B menambahkan aturan analisis ke tabel yang dikonfigurasi segmen.

```
{
  "joinColumns": [
    "identifier2"
  ],
  "listColumns": [
    "segment3",
    "segment4"
  ]
}
```

`joinColumns`—Perusahaan B memungkinkan Perusahaan A untuk bergabung `identifier2` untuk mencocokkan pelanggan dari data segmen ke data CRM. Perusahaan A dan Perusahaan B bekerja dengan vendor identitas untuk mendapatkan `identifier2` mana yang cocok untuk kolaborasi ini. Mereka tidak menambahkan yang lain `joinColumns` karena mereka percaya `identifier2` memberikan tingkat kecocokan tertinggi dan paling akurat dan pengidentifikasi lain tidak diperlukan untuk kueri.

`listColumns`—Perusahaan B memungkinkan Perusahaan A untuk memperkaya data mereka dengan `segment3` dan `segment4` atribut yang merupakan atribut unik yang telah mereka buat, kumpulkan, dan selaraskan (dengan pelanggan A) untuk menjadi bagian dari pengayaan data. Mereka ingin Perusahaan A mendapatkan segmen ini untuk tumpang tindih pada tingkat baris karena ini adalah kolaborasi pengayaan data.

7. Perusahaan A menciptakan asosiasi tabel CRM untuk kolaborasi.

8. Perusahaan B menciptakan asosiasi tabel segmen untuk kolaborasi.
9. Perusahaan A menjalankan kueri, seperti yang berikut untuk memperkaya data pelanggan yang tumpang tindih.

```
SELECT companyA.internalid, companyB.segment3, companyB.segment4
INNER JOIN returns companyB
  ON companyA.identifier2 = companyB.identifier2
WHERE companyA.customercategory > 'xxx'
```

10. Perusahaan A dan Perusahaan B meninjau log kueri. Perusahaan B memverifikasi bahwa kueri sejalan dengan apa yang disepakati dalam perjanjian kolaborasi.

Aturan analisis kustom di AWS Clean Rooms

Dalam AWS Clean Rooms, aturan analisis kustom adalah jenis aturan analisis baru yang memungkinkan kueri kustom dijalankan pada tabel yang dikonfigurasi. [Kueri SQL khusus masih dibatasi untuk hanya memiliki SELECT perintah tetapi dapat menggunakan lebih banyak konstruksi SQL daripada agregasi dan kueri daftar \(misalnya, fungsi jendela, OUTER JOIN CTEs, atau subquery; lihat Referensi SQL untuk daftar lengkap\).](#) [AWS Clean Rooms Kueri SQL kustom tidak harus mengikuti struktur kueri seperti agregasi dan kueri daftar.](#)

Aturan analisis kustom mendukung kasus penggunaan yang lebih maju daripada yang dapat didukung oleh aturan agregasi dan analisis daftar seperti analisis atribusi khusus, perbandingan, analisis inkrementalitas, dan penemuan audiens. Ini merupakan tambahan dari superset kasus penggunaan yang didukung oleh agregasi dan aturan analisis daftar.

Aturan analisis kustom juga mendukung privasi diferensial. Privasi diferensial adalah kerangka kerja yang ketat secara matematis untuk perlindungan privasi data. Untuk informasi selengkapnya, lihat [AWS Clean Rooms Privasi Diferensial](#). Saat Anda membuat templat analisis, Privasi AWS Clean Rooms Diferensial memeriksa templat untuk menentukan apakah templat tersebut kompatibel dengan struktur kueri tujuan umum untuk AWS Clean Rooms Privasi Diferensial. Validasi ini memastikan bahwa Anda tidak membuat templat analisis yang tidak diizinkan dengan tabel yang dilindungi privasi diferensial.

Untuk mengonfigurasi aturan analisis kustom, pemilik data dapat memilih untuk mengizinkan kueri khusus tertentu, yang disimpan dalam [templat analisis](#), untuk dijalankan pada tabel yang dikonfigurasi. Pemilik data meninjau templat analisis sebelum menambahkannya ke kontrol analisis yang diizinkan dalam aturan analisis khusus. Template analisis tersedia dan hanya terlihat dalam

kolaborasi di mana mereka dibuat (bahkan jika tabel dikaitkan dengan kolaborasi lain) dan hanya dapat dijalankan oleh anggota yang dapat melakukan kueri dalam kolaborasi itu.

Atau, anggota dapat memilih untuk mengizinkan anggota lain (penyedia kueri) untuk membuat kueri tanpa ulasan. Anggota menambahkan akun penyedia kueri yang dikendalikan oleh penyedia kueri yang diizinkan dalam aturan analisis kustom. Jika penyedia kueri adalah anggota yang dapat melakukan kueri, mereka dapat menjalankan kueri apa pun secara langsung pada tabel yang dikonfigurasi. Penyedia kueri juga dapat membuat kueri dengan [membuat templat analisis](#). Setiap kueri yang telah dibuat oleh penyedia kueri secara otomatis diizinkan untuk berjalan di atas meja di semua kolaborasi di mana Akun AWS ada dan tabel terkait.

Pemilik data hanya dapat mengizinkan templat analisis atau akun untuk membuat kueri, bukan keduanya. Jika pemilik data membiarkannya kosong, anggota yang dapat melakukan kueri tidak dapat menjalankan kueri pada tabel yang dikonfigurasi.

Topik

- [Aturan analisis kustom struktur yang telah ditentukan](#)
- [Contoh aturan analisis kustom](#)
- [Aturan analisis khusus dengan privasi diferensial](#)

Aturan analisis kustom struktur yang telah ditentukan

Contoh berikut mencakup struktur yang telah ditentukan yang menunjukkan kepada Anda cara menyelesaikan aturan analisis kustom dengan privasi diferensial diaktifkan. `userIdentifier`Nilai adalah kolom yang secara unik mengidentifikasi pengguna Anda, seperti `user_id`. Bila Anda memiliki dua atau lebih tabel dengan privasi diferensial diaktifkan dalam kolaborasi, AWS Clean Rooms Anda harus mengonfigurasi kolom yang sama dengan kolom pengenalan pengguna di kedua aturan analisis untuk mempertahankan definisi pengguna yang konsisten di seluruh tabel.

```
{
  "allowedAnalyses": ["ANY_QUERY"] | string[],
  "allowedAnalysisProviders": [],
  "differentialPrivacy": {
    "columns": [
      {
        "name": "userIdentifier"
      }
    ]
  }
}
```

```
}
```

Anda dapat:

- Tambahkan template analisis ARNs ke kontrol analisis yang diizinkan. Dalam hal ini, `allowedAnalysisProviders` kontrol tidak termasuk.

```
{  
  allowedAnalyses: string[]  
}
```

- Tambahkan anggota Akun AWS IDs ke `allowedAnalysisProviders` kontrol. Dalam hal ini, Anda `ANY_QUERY` menambah `allowedAnalyses` kontrol.

```
{  
  allowedAnalyses: ["ANY_QUERY"],  
  allowedAnalysisProviders: string[]  
}
```

Contoh aturan analisis kustom

Contoh berikut menunjukkan bagaimana dua perusahaan dapat berkolaborasi dalam AWS Clean Rooms menggunakan aturan analisis kustom.

Perusahaan A memiliki data pelanggan dan penjualan. Perusahaan A tertarik untuk memahami peningkatan penjualan kampanye iklan di situs Perusahaan B. Perusahaan B memiliki data pemirsa dan atribut segmen yang berguna bagi Perusahaan (misalnya, perangkat yang mereka gunakan saat melihat iklan).

Perusahaan A memiliki kueri inkrementalitas tertentu yang ingin mereka jalankan dalam kolaborasi.

Untuk membuat kolaborasi dan menjalankan analisis kustom dalam kolaborasi, perusahaan melakukan hal berikut:

1. Perusahaan A menciptakan kolaborasi dan menciptakan keanggotaan. Kolaborasi ini memiliki Perusahaan B sebagai anggota lain dalam kolaborasi tersebut. Perusahaan A memungkinkan pencatatan kueri dalam kolaborasi, dan memungkinkan pencatatan kueri di akunnya.
2. Perusahaan B menciptakan keanggotaan dalam kolaborasi. Ini memungkinkan pencatatan kueri di akunnya.

3. Perusahaan A membuat tabel yang dikonfigurasi CRM
4. Perusahaan A menambahkan aturan analisis kustom kosong ke tabel penjualan yang dikonfigurasi.
5. Perusahaan A mengaitkan tabel penjualan yang dikonfigurasi untuk kolaborasi.
6. Perusahaan B membuat tabel yang dikonfigurasi pemirsa.
7. Perusahaan B menambahkan aturan analisis kustom kosong ke tabel yang dikonfigurasi pemirsa.
8. Perusahaan B mengaitkan tabel yang dikonfigurasi pemirsa dengan kolaborasi.
9. Perusahaan A melihat tabel penjualan dan tabel pemirsa yang terkait dengan kolaborasi dan membuat templat analisis, menambahkan kueri inkrementalitas dan parameter untuk bulan kampanye.

```
{
  "analysisParameters": [
    {
      "defaultValue": ""
      "type": "DATE"
      "name": "campaign_month"
    }
  ],
  "description": "Monthly incrementality query using sales and viewership data"
  "format": "SQL"
  "name": "Incrementality analysis"
  "source":
    "WITH labeleddata AS
    (
      SELECT hashedemail, deviceid, purchases, unitprice, purchasedate,
      CASE
        WHEN testvalue IN ('value1', 'value2', 'value3') THEN 0
        ELSE 1
      END AS testgroup
      FROM viewershipdata
    )
    SELECT labeleddata.purchases, provider.impressions
    FROM labeleddata
    INNER JOIN salesdata
      ON labeleddata.hashedemail = provider.hashedemail
    WHERE MONTH(labeleddata.purchasedate) > :campaignmonth
    AND testgroup = :group
    "
}
```

10 Perusahaan A menambahkan akun mereka (misalnya, 444455556666) ke kontrol penyedia analisis yang diizinkan dalam aturan analisis khusus. Mereka menggunakan kontrol penyedia analisis yang diizinkan karena mereka ingin mengizinkan kueri apa pun yang mereka buat berjalan di tabel yang dikonfigurasi penjualan mereka.

```
{
  "allowedAnalyses": [
    "ANY_QUERY"
  ],
  "allowedAnalysisProviders": [
    "444455556666"
  ]
}
```

11 Perusahaan B melihat template analisis yang dibuat dalam kolaborasi dan meninjau isinya termasuk string kueri dan parameter.

12 Perusahaan B menentukan bahwa templat analisis mencapai kasus penggunaan inkrementalitas dan memenuhi persyaratan privasi mereka tentang bagaimana tabel yang dikonfigurasi pemirsa mereka dapat ditanyakan.

13 Perusahaan B menambahkan templat analisis ARN ke kontrol analisis yang diizinkan dalam aturan analisis khusus dari tabel pemirsa. Mereka menggunakan kontrol analisis yang diizinkan karena mereka hanya ingin mengizinkan kueri inkrementalitas berjalan pada tabel yang dikonfigurasi pemirsa mereka.

```
{
  "allowedAnalyses": [
    "arn:aws:cleanrooms:us-east-1:111122223333:membership/41327cc4-bbf0-43f1-b70c-a160dddceb08/analysistemplate/1ff1bf9d-781c-418d-a6ac-2b80c09d6292"
  ]
}
```

14 Perusahaan A menjalankan template analisis dan menggunakan nilai parameter 05-01-2023.

Aturan analisis khusus dengan privasi diferensial

Pada tahun AWS Clean Rooms, aturan analisis kustom mendukung privasi diferensial. Privasi diferensial adalah kerangka kerja yang ketat secara matematis untuk perlindungan privasi data yang membantu Anda melindungi data Anda dari upaya identifikasi ulang.

Privasi diferensial mendukung analisis agregat seperti perencanaan kampanye iklan, post-ad-campaign pengukuran, perbandingan dalam konsorsium lembaga keuangan, dan pengujian A/B untuk penelitian kesehatan.

Struktur kueri dan sintaks yang didukung didefinisikan dalam [Struktur kueri dan sintaks](#).

Aturan analisis khusus dengan contoh privasi diferensial

 Note

AWS Clean Rooms Privasi Diferensial hanya tersedia untuk kolaborasi di mana data disimpan di Amazon S3.

Pertimbangkan [contoh aturan analisis kustom](#) yang disajikan di bagian sebelumnya. Contoh ini menunjukkan bagaimana Anda dapat menggunakan privasi diferensial untuk melindungi data Anda dari upaya identifikasi ulang sambil memungkinkan mitra Anda mempelajari wawasan penting bisnis dari data Anda. Asumsikan bahwa Perusahaan B, yang memiliki data pemirsa, ingin melindungi data mereka menggunakan privasi diferensial. Untuk menyelesaikan pengaturan privasi diferensial, Perusahaan B menyelesaikan langkah-langkah berikut:

1. Perusahaan B mengaktifkan privasi diferensial sambil menambahkan aturan analisis kustom ke tabel yang dikonfigurasi pemirsa. Perusahaan B memilih `viewershipdata.hashemail` sebagai kolom pengenalan pengguna.
2. Perusahaan B [menambahkan kebijakan privasi diferensial](#) dalam kolaborasi untuk membuat tabel data pemirsa mereka tersedia untuk kueri. Perusahaan B memilih kebijakan default untuk menyelesaikan penyiapan dengan cepat.

Perusahaan A, yang ingin memahami peningkatan penjualan kampanye iklan di situs Perusahaan B, menjalankan templat analisis. Karena kueri kompatibel dengan [struktur kueri](#) tujuan umum Privasi AWS Clean Rooms Diferensial, kueri berjalan dengan sukses.

Struktur kueri dan sintaks

Kueri yang berisi setidaknya satu tabel yang mengaktifkan privasi diferensial harus mematuhi sintaks berikut.

```
query_statement:  
  [cte, ...] final_select
```

```

cte:
    WITH sub_query AS (
        inner_select
        [ UNION | INTERSECT | UNION_ALL | EXCEPT/MINUS ]
        [ inner_select ]
    )

inner_select:
    SELECT [user_id_column, ] expression [, ...]
    FROM table_reference [, ...]
    [ WHERE condition ]
    [ GROUP BY user_id_column[, expression] [, ...] ]
    [ HAVING condition ]

final_select:
    SELECT [expression, ...] | COUNT | COUNT_DISTINCT | SUM | AVG | STDDEV
    FROM table_reference [, ...]
    [ WHERE condition ]
    [ GROUP BY expression [, ...] ]
    [ HAVING COUNT | COUNT_DISTINCT | SUM | AVG | STDDEV | condition ]
    [ ORDER BY column_list ASC | DESC ]
    [ OFFSET literal ]
    [ LIMIT literal ]

expression:
    column_name [, ...] | expression AS alias | aggregation_functions |
window_functions_on_user_id | scalar_function | CASE | column_name math_expression [,
expression]

window_functions_on_user_id:
    function () OVER (PARTITION BY user_id_column, [column_name] [ORDER BY column_list
ASC|DESC])

```

Note

Untuk struktur dan sintaks kueri privasi diferensial, perhatikan hal-hal berikut:

- Sub-kueri tidak didukung.
- Common Table Expressions (CTEs) harus memancarkan kolom pengenalan pengguna jika tabel atau CTE melibatkan data yang dilindungi oleh privasi diferensial. Filter, pengelompokan, dan agregasi harus dilakukan di tingkat pengguna.

- `Final_select` memungkinkan fungsi agregat `COUNT DISTINCT`, `COUNT`, `SUM`, `AVG`, dan `STDDEV`.

Untuk detail selengkapnya tentang kata kunci SQL yang didukung untuk privasi diferensial, lihat [Kemampuan SQL dari AWS Clean Rooms Differential Privacy](#)

Aturan analisis tabel pemetaan ID

Dalam AWS Clean Rooms, aturan analisis tabel pemetaan ID bukanlah aturan analisis mandiri. Jenis aturan analisis ini dikelola oleh AWS Clean Rooms dan digunakan untuk menggabungkan data identitas yang berbeda untuk memfasilitasi kueri. Ini secara otomatis ditambahkan ke tabel pemetaan ID dan tidak dapat diedit. Ini mewarisi perilaku aturan analisis lain dalam kolaborasi — selama aturan analisis tersebut homogen.

Aturan analisis tabel pemetaan ID memberlakukan keamanan pada tabel pemetaan ID. Ini membatasi anggota kolaborasi dari langsung memilih atau memeriksa populasi yang tidak tumpang tindih antara kumpulan data kedua anggota menggunakan tabel pemetaan ID. Aturan analisis tabel pemetaan ID digunakan untuk melindungi data sensitif dalam tabel pemetaan ID saat digunakan dalam kueri dengan aturan analisis lainnya secara implisit.

Dengan aturan analisis tabel pemetaan ID, AWS Clean Rooms memberlakukan tumpang tindih di kedua sisi tabel pemetaan ID di SQL yang diperluas. Ini memungkinkan Anda untuk melakukan tugas-tugas berikut:

- Gunakan tumpang tindih tabel pemetaan ID dalam JOIN pernyataan.

AWS Clean Rooms memungkinkan `INNER`, `LEFT`, atau `RIGHT` bergabung pada tabel pemetaan ID jika menghormati tumpang tindih. Untuk melindungi informasi pemetaan sensitif, tabel pemetaan ID harus selalu berada di sisi "inner" dari operasi apa pun JOIN. Misalnya, JOIN operasi berikut ini valid:

- `table LEFT JOIN id_mapping_table`
- `id_mapping_table RIGHT JOIN table`
- `table INNER JOIN id_mapping_table`

JOIN operasi berikut tidak valid:

- `id_mapping_table LEFT JOIN table`
- `table RIGHT JOIN id_mapping_table`

Ini mencegah eksposur catatan pemetaan yang tidak memiliki kecocokan yang sesuai dalam kumpulan data Anda. Mengizinkan operasi semacam itu berpotensi mengungkapkan informasi sensitif tentang pemetaan data anggota kolaborasi lainnya.

- Gunakan kolom tabel pemetaan dalam JOIN pernyataan.

Anda tidak dapat menggunakan kolom tabel pemetaan dalam pernyataan berikut: SELECT,,, WHERE HAVINGGROUP BY, atau ORDER BY (kecuali perlindungan diubah pada asosiasi namespace ID sumber atau asosiasi namespace ID target).

- Dalam SQL diperluas, AWS Clean Rooms juga mendukung, implisit OUTER JOINJOIN, dan CROSS. JOIN Penggabungan ini tidak dapat memenuhi persyaratan tumpang tindih. Sebagai gantinya, AWS Clean Rooms gunakan `requireOverlap` untuk menentukan kolom mana yang harus digabungkan.

Struktur kueri dan sintaks yang didukung didefinisikan dalam. [Struktur kueri tabel pemetaan ID dan sintaks](#)

Parameter aturan analisis, didefinisikan dalam[Kontrol kueri aturan analisis tabel pemetaan ID](#), termasuk kontrol kueri dan kontrol hasil kueri. Kontrol kuerinya mencakup kemampuan untuk memerlukan tumpang tindih tabel pemetaan ID dalam JOIN pernyataan (yaitu,`requireOverlap`).

Topik

- [Struktur kueri tabel pemetaan ID dan sintaks](#)
- [Kontrol kueri aturan analisis tabel pemetaan ID](#)
- [Aturan analisis tabel pemetaan ID struktur yang telah ditentukan](#)
- [Aturan analisis tabel pemetaan ID - contoh](#)

Struktur kueri tabel pemetaan ID dan sintaks

Kueri pada tabel yang memiliki aturan analisis tabel pemetaan ID harus mematuhi sintaks berikut.

```
--select_list_expression
SELECT
provider.data_col, consumer.data_col

--table_expression
FROM provider
```

```
JOIN idMappingTable idmt ON provider.id = idmt.sourceId

JOIN consumer ON consumer.id = idmt.targetId
```

Tabel kolaborasi

Tabel berikut mewakili tabel dikonfigurasi yang ada dalam AWS Clean Rooms kolaborasi. Kolom id di tabel `cr_drivers_license` dan `cr_insurance` mewakili kolom yang cocok dengan tabel pemetaan ID.

`cr_drivers_license`

id	driver_name	state_of_registrasi
1	Eduard	TX
2	Dana	MA
3	Gweneth	IL

`cr_asuransi`

id	pemegang kebijakan_email	policy_number
a	eduardo@internal.company.com	17f9d04e-f5be-4426-bdc4-250ed59c6529
b	gwen@internal.company.com	3f0092db-2316-48a8-8d44-09cf8f6e6c64
c	rosa@internal.company.com	d7692e84-3d3c-47b8-b46d-a0d5345f0601

Tabel pemetaan ID

Tabel berikut merupakan tabel pemetaan ID yang ada yang cocok pada tabel `cr_drivers_license` dan `cr_insurance`. Tidak semua entri akan dimiliki IDs untuk kedua tabel kolaborasi.

cr_drivers_license_id	cr_insurance_id
1	a
2	null
3	b
null	c

Aturan analisis tabel pemetaan ID hanya memungkinkan kueri berjalan pada kumpulan data yang tumpang tindih, yang akan terlihat sebagai berikut:

cr_driver s_license_id	cr_insura nce_id	driver_name	state_of_ registrasi	pemegang kebijakan _email	policy_nu mber
1	a	Eduard	TX	eduardo@i nternal.c ompany.com	17f9d04e- f5be-4426 -bdc4-250 ed59c6529
3	b	Gweneth	IL	gwen@inte rnal.comp any.com	3f0092db- 2316-48a8 -8d44-09c f8f6e6c64

Kueri contoh

Contoh berikut menunjukkan lokasi yang valid untuk bergabung dengan tabel pemetaan ID:

```
-- Single ID mapping table
SELECT
  [ select_items ]FROM
  cr_drivers_license cr_dl
  [ INNER | LEFT ] JOIN cr_identity_mapping_table idmt ON idmt.cr_drivers_license_id
= cr_dl.id
```

```

    [ INNER | RIGHT ] JOIN cr_insurance cr_in          ON idmt.cr_insurance_id
    = cr_in.id
;
-- Single ID mapping table (Subquery)
SELECT
    [ select_items ]FROM (
    SELECT
        [ select_items ]
    FROM
        cr_drivers_license cr_dl
        [ INNER | LEFT ] JOIN cr_identity_mapping_table idmt ON
idmt.cr_drivers_license_id = cr_dl.id
        [ INNER | RIGHT ] JOIN cr_insurance cr_in          ON idmt.cr_insurance_id
        = cr_in.id
    )
;
-- Single ID mapping table (CTE)
WITH
    matched_ids AS (
        SELECT
            [ select_items ]
        FROM
            cr_drivers_license cr_dl
            [ INNER | LEFT ] JOIN cr_identity_mapping_table idmt ON
idmt.cr_drivers_license_id = cr_dl.id
            [ INNER | RIGHT ] JOIN cr_insurance cr_in          ON
idmt.cr_insurance_id      = cr_in.id
        )SELECT
        [ select_items ]FROM
        matched_ids
;

```

Pertimbangan-pertimbangan

Untuk struktur kueri tabel pemetaan ID dan sintaks, perhatikan hal berikut:

- Anda tidak dapat mengeditnya.
- Ini diterapkan ke tabel pemetaan ID secara default.
- Ini menggunakan asosiasi namespace ID sumber dan target di dalam kolaborasi.
- Tabel pemetaan ID dikonfigurasi secara default untuk memberikan perlindungan default untuk kolom yang berasal dari namespace ID. Anda dapat memodifikasi konfigurasi ini sehingga kolom

yang berasal dari namespace ID (salah satu `sourceID` atau `targetID`) dapat diizinkan di mana saja dalam kueri. Untuk informasi selengkapnya, lihat [Ruang nama ID di AWS Clean Rooms](#).

- Aturan analisis tabel pemetaan ID mewarisi batasan SQL dari aturan analisis lainnya dalam kolaborasi.

Kontrol kueri aturan analisis tabel pemetaan ID

Dengan kontrol kueri tabel pemetaan ID, AWS Clean Rooms mengontrol bagaimana kolom dalam tabel Anda digunakan untuk menanyakan tabel. Misalnya, ia mengontrol kolom mana yang digunakan untuk bergabung, dan kolom mana yang memerlukan tumpang tindih.

Aturan analisis tabel pemetaan ID juga mencakup fungsionalitas yang memungkinkan Anda mengizinkan `sourceID`, `targetID`, atau keduanya, diproyeksikan tanpa memerlukan JOIN.

Tabel berikut menjelaskan setiap kontrol.

Kontrol	Definisi	Penggunaan
<code>joinColumns</code>	Kolom yang dapat digunakan anggota yang dapat kueri dalam pernyataan INNER JOIN.	Anda tidak dapat menggunakan <code>joinColumns</code> di bagian lain dari kueri selain INNER JOIN. Untuk informasi selengkapnya, lihat Bergabunglah dengan kontrol .
<code>dimensionColumns</code>	Kolom (jika ada) yang dapat digunakan anggota yang dapat kueri dalam pernyataan SELECT dan GROUP BY.	A <code>dimensionColumn</code> dapat digunakan di SELECT dan GROUPBY. A <code>dimensionColumn</code> dapat muncul sebagai <code>joinKeys</code> . Anda hanya dapat menggunakan <code>dimensionColumns</code> dalam klausa JOIN jika Anda menentukannya dalam tanda kurung.

Kontrol	Definisi	Penggunaan
<code>queryConstraints:RequireOverlap</code>	Kolom dalam tabel pemetaan ID yang harus digabungkan sehingga kueri dapat berjalan.	Kolom ini harus digunakan untuk BERGABUNG dengan tabel Pemetaan ID dan tabel kolaborasi.

Aturan analisis tabel pemetaan ID struktur yang telah ditentukan

Struktur yang ditentukan sebelumnya untuk aturan analisis tabel pemetaan ID dilengkapi dengan perlindungan default yang diterapkan pada `sourceID` `targetID`. Ini berarti bahwa kolom dengan perlindungan yang diterapkan harus digunakan dalam kueri.

Anda dapat mengonfigurasi aturan analisis tabel pemetaan ID dengan cara berikut:

- Keduanya `sourceID` dan `targetID` dilindungi

Dalam konfigurasi ini, `sourceID` dan keduanya tidak `targetID` dapat diproyeksikan. `sourceID` dan `targetID` harus digunakan dalam JOIN ketika tabel pemetaan ID direferensikan.

- Hanya `targetID` dilindungi

Dalam konfigurasi ini, tidak `targetID` dapat diproyeksikan. `targetID` harus digunakan dalam JOIN ketika tabel pemetaan ID direferensikan. Ini `sourceID` dapat digunakan dalam kueri.

- Hanya `sourceID` dilindungi

Dalam konfigurasi ini, tidak `sourceID` dapat diproyeksikan. `sourceID` harus digunakan dalam JOIN ketika tabel pemetaan ID direferensikan. Ini `targetID` dapat digunakan dalam kueri.

- Tidak ada `sourceID` atau `targetID` dilindungi

Dalam konfigurasi ini, tabel pemetaan ID tidak tunduk pada penegakan khusus apa pun yang dapat digunakan dalam kueri.

Contoh berikut menunjukkan struktur yang telah ditentukan sebelumnya untuk aturan analisis tabel pemetaan ID dengan perlindungan default yang diterapkan pada `sourceID` `targetID`. Dalam contoh ini, aturan analisis tabel pemetaan ID hanya mengizinkan INNER JOIN pada `sourceID` kolom dan `targetID` kolom.

```
{
  "joinColumns": [
    "source_id",
    "target_id"
  ],
  "queryConstraints": [
    {
      "requireOverlap": {
        "columns": [
          "source_id",
          "target_id"
        ]
      }
    }
  ],
  "dimensionColumns": [] // columns that can be used in SELECT and JOIN
}
```

Contoh berikut menunjukkan struktur yang telah ditentukan sebelumnya untuk aturan analisis tabel pemetaan ID dengan perlindungan yang diterapkan pada `targetID`. Dalam contoh ini, aturan analisis tabel pemetaan ID hanya mengizinkan INNER JOIN pada `sourceID` kolom.

```
{
  "joinColumns": [
    "source_id",
    "target_id"
  ],
  "queryConstraints": [
    {
      "requireOverlap": {
        "columns": [
          "target_id"
        ]
      }
    }
  ],
  "dimensionColumns": [
    "source_id"
  ]
}
```

Contoh berikut menunjukkan struktur yang telah ditentukan sebelumnya untuk aturan analisis tabel pemetaan ID dengan perlindungan yang diterapkan pada `sourceID`. Dalam contoh ini, aturan analisis tabel pemetaan ID hanya mengizinkan INNER JOIN pada `targetID` kolom.

```
{
  "joinColumns": [
    "source_id",
    "target_id"
  ],
  "queryConstraints": [
    {
      "requireOverlap": {
        "columns": [
          "source_id"
        ]
      }
    }
  ],
  "dimensionColumns": [
    "target_id"
  ]
}
```

Contoh berikut menunjukkan struktur yang ditentukan sebelumnya untuk aturan analisis tabel pemetaan ID tanpa perlindungan yang diterapkan pada `sourceID` `targetID`. Dalam contoh ini, aturan analisis tabel pemetaan ID memungkinkan INNER JOIN pada `sourceID` kolom dan `targetID` kolom.

```
{
  "joinColumns": [
    "source_id",
    "target_id"
  ],
  "queryConstraints": [
    {
      "requireOverlap": {
        "columns": []
      }
    }
  ],
  "dimensionColumns": [
    "source_id",
    "target_id"
  ]
}
```



```
"target_id"  
]  
}
```

Aturan analisis tabel pemetaan ID - contoh

Daripada menulis pernyataan air terjun panjang yang merujuk Informasi Identifikasi Pribadi (PII), misalnya, perusahaan dapat menggunakan aturan analisis tabel pemetaan ID untuk menggunakan transcoding multi-pihak. LiveRamp Contoh berikut menunjukkan bagaimana Anda dapat berkolaborasi dalam AWS Clean Rooms menggunakan aturan analisis tabel pemetaan ID.

Perusahaan A adalah pengiklan yang memiliki data pelanggan dan penjualan, yang akan digunakan sebagai sumber. Perusahaan A juga melakukan transcoding atas nama para pihak dalam kolaborasi, dan membawa LiveRamp kredensialnya.

Perusahaan B adalah penerbit yang memiliki data peristiwa, yang akan digunakan sebagai target.

Note

Baik Perusahaan A atau Perusahaan B dapat memberikan kredensi LiveRamp transcoding dan melakukan transcoding.

Untuk membuat kolaborasi yang memungkinkan analisis tabel pemetaan ID bekerja sama, perusahaan melakukan hal berikut:

1. Perusahaan A menciptakan kolaborasi dan menciptakan keanggotaan. Ini menambahkan Perusahaan B, yang juga menciptakan keanggotaan dalam kolaborasi.
2. Perusahaan A mengaitkan sumber namespace ID yang ada atau membuat yang baru dalam Resolusi Entitas AWS menggunakan konsol. AWS Clean Rooms

Perusahaan A membuat tabel yang dikonfigurasi dengan data penjualannya dan kolom yang dikunci `sourceId` pada tabel pemetaan ID.

Sumber namespace ID menyediakan data untuk transkode.

3. Perusahaan B mengaitkan target namespace ID yang ada atau membuat yang baru dalam Resolusi Entitas AWS menggunakan konsol. AWS Clean Rooms

Perusahaan B membuat tabel yang dikonfigurasi dengan data peristiwa mereka dan kolom yang dikunci ke `targetId` dalam tabel pemetaan ID.

Target namespace ID tidak menyediakan data untuk ditranskode, hanya metadata di sekitar konfigurasi. LiveRamp

- Perusahaan A menemukan dua ruang nama ID yang terkait dengan kolaborasi dan membuat serta mengisi tabel pemetaan ID.
- Perusahaan A menjalankan kueri di dua kumpulan data dengan bergabung pada tabel pemetaan ID.

```
--- this would be valid for Custom or List
SELECT provider.data_col, consumer.data_col
FROM provider
  JOIN idMappingTable-123123123123-myMappingWFName idmt
    ON provider.id = idmt.sourceId
  JOIN consumer
    ON consumer.id = idmt.targetId
```

AWS Clean Rooms Privasi Diferensial

AWS Clean Rooms Privasi Diferensial membantu Anda melindungi privasi pengguna Anda dengan teknik yang didukung secara matematis yang diimplementasikan dengan kontrol intuitif dalam beberapa klik. Sebagai kemampuan yang dikelola sepenuhnya, tidak diperlukan pengalaman privasi diferensial sebelumnya untuk membantu Anda mencegah identifikasi ulang pengguna Anda. AWS Clean Rooms secara otomatis menambahkan jumlah noise yang dikalibrasi dengan hati-hati ke hasil kueri saat runtime untuk membantu melindungi data tingkat individu Anda.

AWS Clean Rooms Privasi Diferensial mendukung berbagai kueri analitis dan cocok untuk berbagai kasus penggunaan, di mana sejumlah kecil kesalahan dalam hasil kueri tidak akan membahayakan kegunaan analisis Anda. Dengan itu, mitra Anda dapat menghasilkan wawasan penting bisnis tentang kampanye iklan, keputusan investasi, penelitian klinis, dan banyak lagi, semuanya tanpa memerlukan pengaturan tambahan dari mitra Anda.

AWS Clean Rooms Privasi Diferensial melindungi dari overflow atau kesalahan cast tidak valid yang menggunakan fungsi skalar atau simbol operator matematika dengan cara yang berbahaya.

Untuk informasi selengkapnya tentang Privasi AWS Clean Rooms Diferensial, lihat topik berikut.

Topik

- [Privasi diferensial](#)

- [Bagaimana Privasi Diferensial bekerja AWS Clean Rooms](#)
- [Kebijakan privasi diferensial](#)
- [Kemampuan SQL dari AWS Clean Rooms Differential Privacy](#)
- [Kiat dan contoh kueri Privasi Diferensial](#)
- [Batasan Privasi AWS Clean Rooms Diferensial](#)

Privasi diferensial

Privasi diferensial hanya memungkinkan wawasan agregat dan mengaburkan kontribusi data individu dalam wawasan tersebut. Privasi diferensial melindungi data kolaborasi dari anggota yang dapat menerima hasil belajar tentang individu tertentu. Tanpa privasi diferensial, anggota yang dapat menerima hasil dapat mencoba menyimpulkan data pengguna individu dengan menambahkan atau menghapus catatan tentang individu dan mengamati perbedaan dalam hasil kueri.

Ketika privasi diferensial diaktifkan, jumlah noise tertentu ditambahkan ke hasil kueri untuk mengaburkan kontribusi pengguna individu. Jika anggota yang dapat menerima hasil mencoba mengamati perbedaan hasil kueri setelah menghapus catatan tentang individu dari kumpulan data mereka, variabilitas dalam hasil kueri membantu mencegah identifikasi data individu. AWS Clean Rooms Privasi Diferensial menggunakan [SampCert](#) sampler, implementasi sampler yang terbukti benar yang dikembangkan oleh AWS.

Bagaimana Privasi Diferensial bekerja AWS Clean Rooms

Alur kerja untuk mengaktifkan privasi diferensial AWS Clean Rooms memerlukan langkah-langkah tambahan berikut saat [menyelesaikan alur kerja untuk](#): AWS Clean Rooms

1. Anda mengaktifkan privasi diferensial saat menambahkan [aturan analisis khusus](#).
2. [Anda mengonfigurasi kebijakan privasi diferensial untuk kolaborasi](#) agar tabel data Anda dilindungi dengan privasi diferensial yang tersedia untuk kueri.

Setelah Anda menyelesaikan langkah-langkah ini, anggota yang dapat melakukan kueri dapat mulai menjalankan kueri pada data yang dilindungi privasi diferensial. AWS Clean Rooms mengembalikan hasil yang sesuai dengan kebijakan privasi diferensial. AWS Clean Rooms Privasi Diferensial melacak perkiraan jumlah kueri yang tersisa yang dapat Anda jalankan, mirip dengan pengukur gas di mobil yang menunjukkan tingkat bahan bakar mobil saat ini. Jumlah kueri yang dapat dijalankan

oleh anggota yang dapat melakukan kueri dibatasi oleh anggaran Privasi dan Kebisingan yang ditambahkan per parameter kueri yang diatur dalam [Kebijakan privasi diferensial](#).

Pertimbangan-pertimbangan

Saat menggunakan privasi diferensial di AWS Clean Rooms, pertimbangkan hal berikut:

- Anggota yang dapat menerima hasil tidak dapat menggunakan privasi diferensial. Mereka akan mengonfigurasi aturan analisis khusus dengan privasi diferensial dimatikan untuk tabel yang dikonfigurasi.
- Anggota yang dapat melakukan kueri tidak dapat menggabungkan tabel dari dua atau lebih penyedia data ketika keduanya mengaktifkan privasi diferensial.

Kebijakan privasi diferensial

Kebijakan privasi diferensial mengontrol berapa banyak fungsi agregasi yang diizinkan oleh anggota yang dapat kueri untuk dijalankan dalam suatu kolaborasi. Anggaran Privasi mendefinisikan sumber daya umum dan terbatas yang diterapkan semua tabel dalam kolaborasi. Kebisingan yang ditambahkan per kueri mengatur tingkat di mana anggaran privasi habis.

Kebijakan privasi diferensial diperlukan untuk membuat tabel yang dilindungi privasi diferensial Anda tersedia untuk pertanyaan. Ini adalah langkah satu kali dalam kolaborasi dan mencakup dua input:

- Anggaran privasi — Dikukur dalam hal epsilon, anggaran privasi mengontrol tingkat perlindungan privasi. Ini adalah sumber daya umum dan terbatas yang diterapkan untuk semua tabel Anda yang dilindungi dengan privasi diferensial dalam kolaborasi, karena tujuannya adalah untuk menjaga privasi pengguna Anda yang informasinya dapat hadir dalam beberapa tabel.

Anggaran Privasi dikonsumsi setiap kali kueri dijalankan di tabel Anda. Ketika anggaran privasi sepenuhnya habis, anggota kolaborasi yang dapat melakukan kueri tidak dapat menjalankan kueri tambahan hingga ditingkatkan atau di-refresh. Dengan menetapkan anggaran privasi yang lebih besar, anggota yang dapat menerima hasil dapat mengurangi ketidakpastian mereka tentang individu dalam data. Pilih anggaran privasi yang menyeimbangkan persyaratan kolaborasi Anda dengan kebutuhan privasi Anda dan setelah berkonsultasi dengan pengambil keputusan bisnis.

Anda dapat memilih Segarkan anggaran privasi setiap bulan untuk secara otomatis membuat anggaran privasi baru setiap bulan kalender, jika Anda berencana untuk secara teratur membawa data baru ke dalam kolaborasi. Memilih opsi ini memungkinkan jumlah informasi yang sewenang-

wenang untuk diungkapkan tentang baris data ketika berulang kali ditanyakan di seluruh penyegaran. Hindari memilih ini jika baris yang sama akan berulang kali ditanyakan antara penyegaran anggaran privasi.

- Kebisingan yang ditambahkan per kueri diukur dalam hal jumlah pengguna yang kontribusinya ingin Anda kaburkan. Nilai ini mengatur tingkat di mana anggaran privasi habis. Nilai noise yang lebih besar mengurangi tingkat di mana anggaran privasi habis, dan karenanya memungkinkan lebih banyak kueri untuk dijalankan pada data Anda. Namun, ini harus diimbangi dengan merilis wawasan data yang kurang akurat. Pertimbangkan akurasi yang diinginkan untuk wawasan kolaborasi saat menetapkan nilai ini.

Anda dapat menggunakan kebijakan privasi diferensial default untuk menyelesaikan pengaturan dengan cepat atau menyesuaikan kebijakan privasi diferensial Anda sesuai kasus penggunaan Anda. AWS Clean Rooms Privasi Diferensial menyediakan kontrol intuitif untuk mengonfigurasi kebijakan. AWS Clean Rooms Privasi Diferensial memungkinkan Anda melihat pratinjau utilitas dalam hal jumlah agregasi yang mungkin di semua kueri pada data Anda dan memperkirakan berapa banyak kueri yang dapat dijalankan dalam kolaborasi data.

Anda dapat menggunakan contoh interaktif untuk memahami bagaimana nilai yang berbeda dari anggaran Privasi dan Kebisingan yang ditambahkan per kueri akan memengaruhi hasil untuk berbagai jenis kueri SQL. Secara umum, Anda perlu menyeimbangkan kebutuhan privasi Anda dengan jumlah pertanyaan yang ingin Anda izinkan dan keakuratan pertanyaan tersebut. Anggaran Privasi yang lebih kecil atau Noise yang lebih besar yang ditambahkan per kueri dapat melindungi privasi pengguna dengan lebih baik, tetapi memberikan wawasan yang kurang berarti bagi mitra kolaborasi Anda.

Jika Anda meningkatkan anggaran Privasi sambil menjaga parameter Noise yang ditambahkan per kueri tetap sama, anggota yang dapat melakukan kueri dapat menjalankan lebih banyak agregasi pada tabel Anda dalam kolaborasi. Anda dapat meningkatkan anggaran Privasi kapan saja selama kolaborasi. Jika Anda mengurangi anggaran Privasi sambil menjaga parameter Noise yang ditambahkan per kueri tetap sama, anggota yang dapat melakukan kueri dapat menjalankan agregasi yang lebih sedikit. Anda tidak dapat mengurangi anggaran Privasi setelah anggota yang dapat melakukan kueri mulai menganalisis data Anda.

Jika Anda meningkatkan Noise yang ditambahkan per kueri sambil menjaga input anggaran Privasi tetap sama, anggota yang dapat melakukan kueri dapat menjalankan lebih banyak agregasi pada tabel Anda dalam kolaborasi. Jika Anda mengurangi Noise yang ditambahkan per kueri sambil menjaga input anggaran Privasi tetap sama, anggota yang dapat melakukan kueri dapat menjalankan

agregasi yang lebih sedikit. Anda dapat menambah atau mengurangi Noise yang ditambahkan per kueri kapan saja selama kolaborasi.

Kebijakan privasi diferensial dikelola oleh tindakan API templat anggaran privasi.

Kemampuan SQL dari AWS Clean Rooms Differential Privacy

AWS Clean Rooms Privasi Diferensial menggunakan struktur kueri tujuan umum untuk mendukung kueri SQL yang kompleks. Template analisis kustom divalidasi terhadap struktur ini untuk memastikan bahwa mereka dapat berjalan pada tabel yang dilindungi oleh privasi diferensial. Tabel berikut menunjukkan fungsi mana yang didukung. Untuk informasi selengkapnya, lihat [Struktur kueri dan sintaks](#).

Kategori

SELECT

JOIN

GROUP BY

di

mesin

analitik

Spark

Fungsi

ANY_VALUE

PERCENTILE

AVG

COUNT

PERCENTILE

COUNT

DISTINCT

STDDEV

AVG

COUNT

COUNT

DISTINCT

MAX

Kategori
SELECT
SUM
(COUNT)
di
mesin
analitik
Spark
menghasilkan
MEDIAN
data
fungsi
dengan
MIN
catatan
• Fungsi
tingkat
PERCENTIL
pengguna
E-CONT
Anda
• Fungsi
harus
STDDEV_SAMP
menulis
MP
ekspresi
dan
SELECT
STDDEV_POP
pada
P
mereka
• Fungsi
yang
SUM
CIES
dan
menggunakan
SUM
an
SELECT
DISTINCT
fungsi
MARCSAMP
dan.
format
MAR_POP

Kemampuan

SELECT

SQL

(CTE)

di

mesin

analitik

Spark

DENGAN

dengan

DENGAN

kalau

CTE

menggunakan

an

tabel

yang

dilindungi

i

privasi

diferensi

al

harus

menghasil

kan

data

dengan

catatan

tingkat

pengguna.

Anda

harus

menulis

ekspresi

SELECT

Kemampuan
SELECT
SQL
(CTEs)
di
mesin
analitik
Spark
pada
mereka
yang
CTEs
menggunak
an
`SELECT
userIdent
ifierColu
mn...`
format.

Kondisi
SELECT
SQL
(CTE)
di
mesin
analitik
Spark

Subquery
dapat
• MEMILIKI
• memiliki
• BERGABUNG
subquery
• Kondisi
yang
BERGABUNG
tidak
• FROM
mereferen
sikan
WHERE
hubungan
privasi
diferensi
al
dalam
konstruks
i
ini.
Anda
dapat
memiliki
subquery
yang
mereferen
sikan
hubungan
privasi
diferensi
al

Kemampuan
SELECT
SQL
(CTE)
di
mesin
analitik
Spark
dalam
klausa
FROM
dan
JOIN
saja.

Kemampuan SQL

SELECT

SQL

(CTE)

di

mesin

analitik

Spark

BERGABUNG

DENGAN

KIRI

syarat

bahwa

BERGABUNG

hanya

• KIRI

fungsi

SEMI

JOIN

BERGABUNG

yang

• KIRI

equi-

ANTI

join

BERGABUNG

pada

• BERGABUNG

kolom

DENGAN

pengenal

BENAR

pengguna

• BERGABUNG

yang

di

• PENYISIR

dan

BERGABUN

wajib

•

ATAU

menanyakan

operator

• CROSS

dua

JOIN

atau

lebih

tabel

dengan

privasi

Konfirmasi bahwa pemilik tabel telah mengonfigurasi kolom pengenalan pengguna yang sama di semua tabel. Pastikan bahwa kondisi equi-join wajib sudah benar. Konfirmasi bahwa pemilik tabel telah mengonfigurasi kolom pengenalan pengguna yang sama di semua tabel.

Ketika

SELECT

SQL

(CTE)

di

mesin

analitik

Spark

sehingga

definisi

pengguna

tetap

konsisten

di

seluruh

tabel.

Fungsi

CROSS

JOIN

tidak

didukung

saat

menggabun

gkan

dua

atau

lebih

relasi

dengan

privasi

diferensi

al

diaktifka

n.

Kemampuan
SELECT
SQL
(CTE)
di
mesin
analitik
Spark
UNION
UNION
ALL,
INTERSECT
”
KECUALI
||
MINUS
(ini
adalah
sinonim))

Kategori
SELECT
SQL
(CTE)
di
mesin
analitik
Spark

Fungsi
jendela
dengan
• Fungsi
kondisi
jendela
bahwa
AVG
kolom
• Fungsi
pengenal
jendela
di COUNT
klasifikasi
jendela
fungsi
jendela
• Fungsi
diperlukan
jendela
n DENSE_RANK
saat
Anda
• Fungsi
menanyakan
jendela
n FIRST_VALUE
hubungan
UE
dengan
• Fungsi
privasi
jendela
diferensi
al LAG
• Fungsi
diaktifkan
n.jendela
LAST_VALUE
E

Kemampuan

SELECT

SQL

(CTE)

di

mesin

analitik

Spark

- Fungsi
jendela
LEAD
- Fungsi
jendela
MAX
- Fungsi
jendela
MEDIAN
- Fungsi
jendela
MIN
- Fungsi
jendela
NTH_VALUE
- Fungsi
jendela
STDDEV_SAMP
dan
STDDEV_POP
(STDDEV_SAMP
dan
STDDEV

Kategori

SELECT

SUM

(CTE)

di

mesin

analitik

Spark

adalah

sinonim)

- Fungsi
jendela
SUM
- Fungsi
jendela
VAR_SAMP
dan
VAR_POP
(VAR_SAMP
dan
VARIANCE
adalah
sinonim)

Fungsi

peringkat

- Fungsi
jendela
DENSE_RANK
- Fungsi
jendela
NTILE

Kemampuan

SELECT

SQL

(CTE)

di

mesin

analitik

Spark

- Fungsi
jendela
PERCENT_R
ANK
- Fungsi
jendela
RANK
- Fungsi
jendela
ROW_NUMBE
R

Kategori

SELECT

JOIN

GROUP

di

mesin

analitik

Spark

Ekspresi

COALESCE

CASE

- Ekspresi
COALESCE
- Fungsi
TERBESAR
dan
PALING
KECIL
- Fungsi
NVL
dan
COALESCE
- NVL2
fungsi
- Fungsi
NULLIF

Ketika
SELECT
SQL
(CTE)
di
mesin
analitik
Spark
Kondisi
tidak
Ingan
dapat
• kondisi
digunakan
logis
karena
• Kondisi
mereka
pencocoka
memerluka
n
n pola
subquery.
• ANTARA
Semua
kondisi
yang
rentang
lain
• Kondisi
didukung.
nol

Kemampuan

SELECT

SQL

(CTE) yang

di

mesin

analitik

Spark

Fungsi

terdapat

w dan

aktu

dalam

transaksi

- Operator penggabungan
- Fungsi
ADD_MONTHS
- Fungsi
CONVERT_TIMEZONE
- Fungsi
CURRENT_DATE
- Fungsi
DATEADD
- Fungsi
DATEDIFF
- fungsi
DATE_PART

Kemampuan

SELECT

SQL

(CTE)

di

mesin

analitik

Spark

- Fungsi
DATE_TRUNC
- Fungsi
EKSTRAK
- Fungsi
TO_TIMESTAMP
- Bagian
tanggal
untuk
fungsi
tanggal
atau
stempel
waktu

Kategori

SELECT

SQL

(CTE)

di

mesin

analitik

Spark

Sampai

dukungan

(penggabungan)

- Fungsi
BTRIM
- Fungsi
CHAR_LENGTH
- Fungsi
CHARACTER_LENGTH
- Fungsi
CONCAT
- Fungsi
KIRI
dan
KANAN
- Fungsi
LEN
- Fungsi
PANJANG
- Fungsi
LOWER
- Fungsi
LPAD

Kemampuan

SELECT

SQL

(CTE)

di

mesin

analitik

Spark

dan

RPAD

- Fungsi
LTRIM
- Fungsi
POSISI
- Fungsi
REGEXP_CO
UNT
- Fungsi
REGEXP_IN
STR
- Fungsi
REGEXP_RE
PLACE
- Fungsi
REGEXP_SU
BSTR
- Fungsi
REPEAT
- GANTI
fungsi
- Fungsi
REVERSE
- Fungsi
RTRIM

Kemampuan

SELECT

SQL

(CTE)

di

mesin

analitik

Spark

- Fungsi
SPLIT_PAR
T
- Fungsi
SUBSTRING
- FUNGSI
TRANSLATE
- Fungsi
TRIM
- Fungsi
UPPER

Fungsi

TO_CHAR

an

tip

• Fungsi

data

TO_DATE

- TO_NUMBER
- String
format
datetime
- String
format
numerik

Kategori

SELECT

SQL

(CTE)

di

mesin

analitik

Spark

SAES_ENCRY

dukungan

- AES_DEKRI
PSI
- MENGKODEK
AN
- MEMECAHKA
N
KODE
- MD5
fungsi
- SHA1
fungsi
- SHA2
fungsi
- XX_
HASH64

Semua

dukungan

matematik

%,

dan

@

Kategori

SELECT

JOIN

(CTE)

di

mesin

analitik

Spark

Fungsi

analitik

a Fungsi

ACOS

- Fungsi

ASIN

- Fungsi

ATAN

- ATAN2

fungsi

- Fungsi

CBRT

- Fungsi

CEILING

(atau

CEIL)

- Fungsi

COS

- Fungsi

COT

- Fungsi

DERAJAT

- Fungsi

LTRIM

Ketika

SELECT

SQL

(CTE)

di

mesin

analitik

Spark

- Fungsi
EXP
- Fungsi
FLOOR
- Fungsi
LN
- Fungsi
LOG
- Fungsi
MOD
- Fungsi
PI
- Fungsi
POWER
- Fungsi
RADIANS
- fungsi
RANDOM
- Fungsi
ROUND
- Fungsi
SIGN
- Fungsi
SIN

Kategori

SELECT

SQL

(CTE)

di

mesin

analitik

Spark

- Fungsi

SQRT

- Fungsi

TRUNC

SQL

UNBASE64

- SIHIR

- HLL_SKETC

H_AGG,

- HLL_SKETC

H_ESTIMAT

E

- HLL_UNION

- HLL_UNION

_AGG

SQL

didukung

GET_JSON_

OBJEK

Kategori

SELECT

JOIN

(CTE)

di

mesin

analitik

Spark

ARRAY_CON

ARRAYING

- ARRAY_DISTINCT
- ARRAY_KEYQUALI
- ARRAY_BER POTONGAN
- ARRAY_JOIN
- ARRAY_REMOVE
- ARRAY_SORT
- ARRAY_UNION

GROUP

GROUPING

any,

GROUP,

KUBUS

Ketika

SELECT

SQL

(CTE)

di

mesin

analitik

Spark

ORDER

dengan

syarat

bahwa

klausa

ORDER

BY

hanya

didukung

dalam

klausa

partisi

fungsi

jendela

saat

menanyakan

n

tabel

dengan

privasi

diferensi

al

diaktifka

n.

Kemampuan

SELECT

SQL

(CTEs)

di

mesin

analitik

Spark

BASE, ~~BASE~~

SECRET, ~~SECRET~~

dalam

CTEs

menggunak

an

tabel

yang

dilindung

i

privasi

diferensi

al

Dilindungi

tabel

dan

kolom

Dilindungi

matematik

a

pada

fungsi

agregat

Kategori

SELECT

SUM

(COUNT)

di

mesin

analitik

Spark

Didukung

skalar

dalam

fungsi

agregat

Alternatif umum untuk konstruksi SQL yang tidak didukung

Kategori	Konstruksi SQL	Alternatif
Fungsi jendela	• LISTAGG • PERSENTILE_CONT • PERCENTILE_DISC	Anda dapat menggunakan fungsi agregat setara dengan GROUP BY.
Simbol operator matematika	• \$ kolom 2 • \$ kolom 2 • \$ kolom ^ 2	• CBRT • SQRT • DAYA (\$ kolom, 2)
Fungsi skalar	• SYSDATE • \$ kolom: :integer • mengkonversi (jenis, \$ kolom)	• CURRENT_DATE • CAST \$ kolom AS integer • CAST \$ kolom tipe AS
Literal	INTERVAL '1 DETIK'	INTERVAL '1' DETIK
Pembatasan baris	TOP n	BATAS n

Kategori	Konstruksi SQL	Alternatif
Join	• MENGGUNAKAN • ALAMI	Klausa ON harus secara eksplisit berisi kriteria gabungan.

Kiat dan contoh kueri Privasi Diferensial

AWS Clean Rooms Privasi Diferensial menggunakan [struktur kueri tujuan umum](#) untuk mendukung berbagai macam konstruksi SQL seperti Common Table Expressions (CTEs) untuk persiapan data dan fungsi agregat yang umum digunakan seperti, atau. COUNT SUM Untuk mengaburkan kontribusi pengguna yang mungkin dalam data Anda dengan menambahkan noise ke hasil kueri agregat saat run-time, Privasi AWS Clean Rooms Diferensial mengharuskan fungsi agregat di final dijalankan pada data tingkat pengguna. SELECT statement

Contoh berikut menggunakan dua tabel bernama `socialco_impressions` dan `socialco_users` dari penerbit media yang ingin melindungi data menggunakan privasi diferensial saat berkolaborasi dengan merek atletik dengan data. `athletic_brand_sales` Penerbit media telah mengonfigurasi `user_id` kolom sebagai kolom pengenalan pengguna sambil mengaktifkan privasi diferensial. AWS Clean Rooms Pengiklan tidak memerlukan perlindungan privasi diferensial dan ingin menjalankan kueri menggunakan CTEs data gabungan. Karena CTE mereka menggunakan tabel yang dilindungi privasi diferensial, pengiklan menyertakan kolom pengenalan pengguna dari tabel yang dilindungi tersebut dalam daftar kolom CTE dan bergabung dengan tabel yang dilindungi pada kolom pengenalan pengguna.

```
WITH matches_table AS(
  SELECT si.user_id, si.campaign_id, s.sale_id, s.sale_price
  FROM socialco_impressions si
  JOIN socialco_users su
    ON su.user_id = si.user_id
  JOIN athletic_brand_sales s
    ON s.emailsha256 = su.emailsha256
  WHERE s.timestamp > si.timestamp

  UNION ALL

  SELECT si.user_id, si.campaign_id, s.sale_id, s.sale_price
  FROM socialco_impressions si
  JOIN socialco_users su
```

```

        ON su.user_id = si.user_id
    JOIN athletic_brand_sales s
        ON s.phonesha256 = su.phonesha256
    WHERE s.timestamp > si.timestamp
)

SELECT COUNT (DISTINCT user_id) as unique_users
FROM matches_table
GROUP BY campaign_id
ORDER BY COUNT (DISTINCT user_id) DESC
LIMIT 5

```

Demikian pula, jika Anda ingin menjalankan fungsi jendela pada tabel data yang dilindungi privasi diferensial, Anda harus menyertakan kolom pengenal pengguna dalam klausa. `PARTITION BY`

```

ROW_NUMBER() OVER (PARTITION BY conversion_id, user_id ORDER BY match_type, match_age)
AS row

```

Batasan Privasi AWS Clean Rooms Diferensial

AWS Clean Rooms Privasi Diferensial tidak membahas situasi berikut:

1. AWS Clean Rooms Privasi Diferensial hanya mendukung kueri dengan tabel yang didukung Amazon AWS Glue S3. Itu tidak mendukung kueri dengan tabel Snowflake atau Amazon Athena.
2. AWS Clean Rooms Privasi Diferensial tidak mengatasi serangan waktu. Misalnya, serangan ini dimungkinkan dalam skenario di mana pengguna individu menyumbangkan sejumlah besar baris dan menambahkan atau menghapus pengguna ini secara signifikan mengubah waktu perhitungan kueri.
3. AWS Clean Rooms Privasi Diferensial tidak menjamin privasi diferensial ketika kueri SQL dapat mengakibatkan overflow atau kesalahan cast tidak valid pada waktu berjalan karena penggunaan konstruksi SQL tertentu.

Tabel berikut adalah daftar beberapa, tetapi tidak semua, konstruksi SQL yang dapat menghasilkan kesalahan run-time dan harus diverifikasi dalam template analisis. Sebaiknya Anda menyetujui templat analisis yang meminimalkan kemungkinan kesalahan waktu proses tersebut dan meninjau log kueri secara berkala untuk menentukan apakah kueri sesuai dengan perjanjian kolaborasi.

Konstruksi SQL berikut rentan terhadap kesalahan overflow:

Kategori	Konstruksi SQL rentan terhadap kesalahan overflow di mesin analitik Spark SQL
Fungsi agregat	• AVG • JUMLAH/SUM_DISTINCT
Fungsi pemformatan tipe data	• TO_TIMESTAMP • TO_DATE
Fungsi tanggal dan waktu	• ADD_MONTHS • DATEADD • DATEDIFF
Fungsi matematika	• +, -, *, / • DAYA
Fungsi string	• • CONCAT • ULANGI
Fungsi jendela	• AVG • JUMLAH

4. Fungsi pemformatan tipe data CAST rentan terhadap kesalahan cast yang tidak valid.

Anda dapat mengonfigurasi [CloudWatch untuk membuat filter metrik untuk grup log](#) dan kemudian [membuat CloudWatch alarm](#) pada filter metrik tersebut untuk menerima peringatan jika terjadi kemungkinan kesalahan overflow atau cast.

Secara khusus, Anda harus memantau kode kesalahan `CastError`, `OverflowError`, `ConversionError`. Kehadiran kode kesalahan ini menunjukkan potensi serangan saluran samping, tetapi mungkin menunjukkan kueri SQL yang salah.

Lihat informasi yang lebih lengkap di [Analisis masuk AWS Clean Rooms](#).

AWS Clean Room ML

AWS Clean Rooms ML memungkinkan dua pihak atau lebih untuk menjalankan model pembelajaran mesin pada data mereka tanpa perlu berbagi data satu sama lain. Layanan ini menyediakan kontrol peningkatan privasi yang memungkinkan pemilik data untuk menjaga data dan IP model mereka dengan aman. Anda dapat menggunakan model yang AWS tulis atau membawa model kustom Anda sendiri.

Untuk penjelasan lebih rinci tentang cara kerjanya, lihat [Lowongan kerja lintas akun](#).

Untuk informasi selengkapnya tentang kemampuan model Clean Rooms ML, lihat topik berikut.

Topik

- [Terminologi AWS Clean Rooms](#)
- [Cara AWS Clean Rooms ML bekerja dengan AWS model](#)
- [Cara AWS Clean Rooms ML bekerja dengan model khusus](#)
- [AWS model dalam Kamar Bersih](#)
- [Model kustom di Kamar Bersih ML](#)

Terminologi AWS Clean Rooms

Penting untuk memahami terminologi berikut saat menggunakan Clean Rooms ML:

- **Penyedia data pelatihan** — Pihak yang menyumbangkan data pelatihan, membuat dan mengonfigurasi model yang mirip, dan kemudian mengaitkan model yang mirip dengan kolaborasi.
- **Penyedia data benih** — Pihak yang menyumbangkan data benih, menghasilkan segmen yang mirip, dan mengeksport segmen mirip mereka.
- **Data pelatihan** — Data penyedia data pelatihan, yang digunakan untuk menghasilkan model yang mirip. Data pelatihan digunakan untuk mengukur kesamaan dalam perilaku pengguna.

Data pelatihan harus berisi ID pengguna, ID item, dan kolom stempel waktu. Secara opsional, data pelatihan dapat berisi interaksi lain sebagai fitur numerik atau kategoris. Contoh interaksi adalah daftar video yang ditonton, item yang dibeli, atau artikel yang dibaca.

- **Data benih** — Data penyedia data benih, yang digunakan untuk membuat segmen yang mirip. Data benih dapat diberikan secara langsung atau dapat berasal dari hasil AWS Clean Rooms kueri. Output segmen mirip adalah sekumpulan pengguna dari data pelatihan yang paling mirip dengan pengguna benih.

- **Model Lookalike** — Model pembelajaran mesin dari data pelatihan yang digunakan untuk menemukan pengguna serupa di kumpulan data lain.

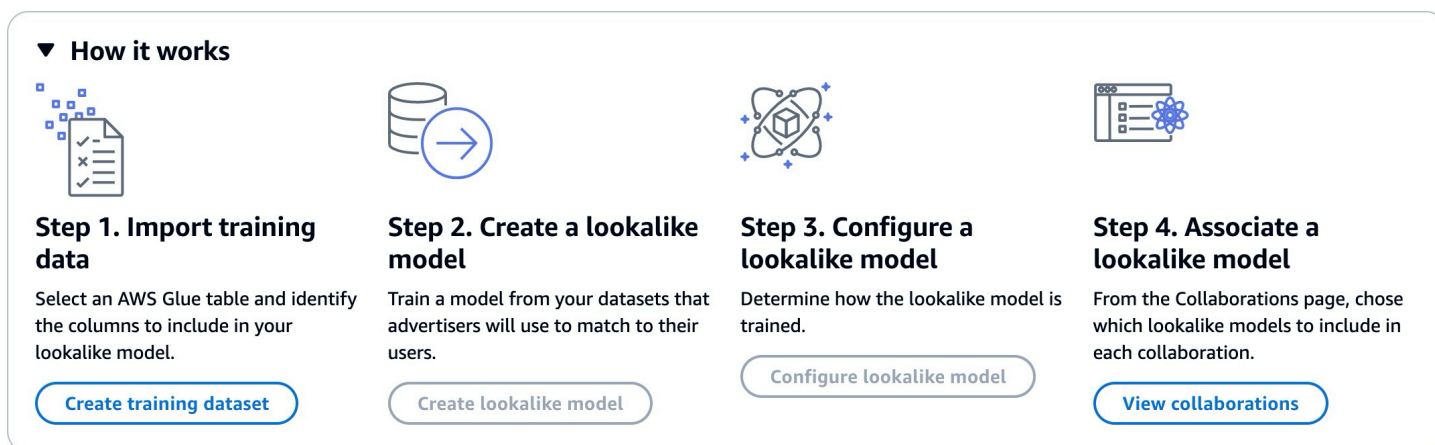
Saat menggunakan API, istilah model audiens digunakan secara setara dengan model yang mirip. Misalnya, Anda menggunakan [CreateAudienceModel](#) API untuk membuat model yang mirip.

- **Segmen mirip** — Subset dari data pelatihan yang paling mirip dengan data benih.

Saat menggunakan API, Anda membuat segmen mirip dengan API. [StartAudienceGenerationJob](#)

Data penyedia data pelatihan tidak pernah dibagikan dengan penyedia data benih dan data penyedia data benih tidak pernah dibagikan dengan penyedia data pelatihan. Output segmen yang mirip dibagikan dengan penyedia data pelatihan, tetapi tidak pernah penyedia data benih.

Cara AWS Clean Rooms ML bekerja dengan AWS model



Bekerja dengan model yang mirip mengharuskan dua pihak, penyedia data pelatihan dan penyedia data benih, bekerja secara berurutan AWS Clean Rooms untuk membawa data mereka ke dalam kolaborasi. Ini adalah alur kerja yang harus diselesaikan oleh penyedia data pelatihan terlebih dahulu:

1. Data penyedia data pelatihan harus disimpan dalam tabel katalog AWS Glue data interaksi item pengguna. Minimal, data pelatihan harus berisi kolom ID pengguna, kolom ID interaksi, dan kolom stempel waktu.
2. Penyedia data pelatihan mendaftarkan data pelatihan dengan AWS Clean Rooms.
3. Penyedia data pelatihan membuat model mirip yang dapat dibagikan dengan beberapa penyedia data benih. Model mirip adalah jaringan saraf dalam yang dapat memakan waktu hingga 24 jam

untuk dilatih. Ini tidak dilatih ulang secara otomatis dan kami sarankan Anda melatih ulang model setiap minggu.

4. Penyedia data pelatihan mengonfigurasi model yang mirip, termasuk apakah akan berbagi metrik relevansi dan lokasi Amazon S3 dari segmen keluaran. Penyedia data pelatihan dapat membuat beberapa model mirip yang dikonfigurasi dari satu model mirip.
5. Penyedia data pelatihan mengaitkan model audiens yang dikonfigurasi dengan kolaborasi yang dibagikan dengan penyedia data benih.

Ini adalah alur kerja yang harus diselesaikan oleh penyedia data seed selanjutnya:

1. Data penyedia data seed dapat disimpan dalam bucket Amazon S3 atau dapat berasal dari hasil kueri.
2. Penyedia data benih membuka kolaborasi yang mereka bagikan dengan penyedia data pelatihan.
3. Penyedia data seed membuat segmen mirip dari tab Clean Rooms di halaman kolaborasi.
4. Penyedia data benih dapat mengevaluasi metrik relevansi, jika dibagikan, dan mengekspor segmen yang mirip untuk digunakan di luar. AWS Clean Rooms

Cara AWS Clean Rooms MLbekerja dengan model khusus

Dengan Clean Rooms, anggota kolaborasi dapat menggunakan algoritma model kustom dockerized yang disimpan di Amazon ECR untuk bersama-sama menganalisis data mereka. Untuk melakukan ini, penyedia model harus membuat gambar dan menyimpannya di Amazon ECR. Ikuti langkah-langkah di [Amazon Elastic Container Registry User Guide](#) untuk membuat repositori pribadi yang akan berisi model HTML kustom.

Setiap anggota kolaborasi dapat menjadi penyedia model, asalkan mereka memiliki izin yang benar. Semua anggota kolaborasi dapat menyumbangkan data pelatihan, data inferensi, atau keduanya ke model. Untuk tujuan panduan ini, data yang berkontribusi anggota disebut sebagai penyedia data. Anggota yang menciptakan kolaborasi adalah pembuat kolaborasi, dan anggota ini dapat berupa penyedia model, salah satu penyedia data, atau keduanya.

Pada level tertinggi, berikut adalah langkah-langkah yang harus diselesaikan untuk melakukan pemodelan ML kustom:

1. Pembuat kolaborasi menciptakan kolaborasi dan menugaskan setiap anggota kemampuan anggota dan konfigurasi pembayaran yang tepat. Pembuat kolaborasi harus menetapkan kemampuan anggota untuk menerima keluaran model atau menerima hasil inferensi kepada

anggota yang sesuai dalam langkah ini karena tidak dapat diperbarui setelah kolaborasi dibuat. Untuk informasi selengkapnya, lihat [Membuat dan bergabung dengan kolaborasi di AWS Clean Rooms ML](#).

2. Penyedia model mengonfigurasi dan mengaitkan model HTML kontainerisasi mereka ke kolaborasi dan memastikan batasan privasi ditetapkan untuk data yang diekspor. Untuk informasi selengkapnya, lihat [Mengonfigurasi algoritme model di AWS Clean Rooms ML](#).
3. Penyedia data menyumbangkan data mereka untuk kolaborasi dan memastikan kebutuhan privasi mereka ditentukan. Penyedia data harus mengizinkan model untuk mengakses data mereka. Untuk informasi selengkapnya, lihat [Menyumbangkan data pelatihan di AWS Clean Rooms MS](#) dan [Mengaitkan algoritma model yang dikonfigurasi di AWS Clean Rooms MS](#).
4. Anggota kolaborasi membuat konfigurasi ML, yang menentukan ke mana artefak model atau hasil inferensi diekspor.
5. Anggota kolaborasi membuat saluran input ML yang memberikan masukan ke wadah pelatihan atau wadah inferensi. Saluran input ML adalah kueri yang mendefinisikan data yang akan digunakan dalam konteks algoritma model.
6. Anggota kolaborasi memanggil pelatihan model menggunakan saluran input ML dan algoritma model yang dikonfigurasi. Untuk informasi selengkapnya, lihat [Membuat model terlatih di AWS Clean Rooms ML](#).
7. (Opsional) Pelatih model memanggil pekerjaan ekspor model dan artefak model dikirim ke penerima hasil model. Hanya anggota dengan konfigurasi ML yang valid dan kemampuan anggota untuk menerima keluaran model yang dapat menerima artefak model. Untuk informasi selengkapnya, lihat [Mengekspor artefak model dari AWS Clean Rooms](#).
8. (Opsional) Seorang anggota kolaborasi memanggil inferensi model menggunakan saluran input ML, ARN model terlatih, dan algoritma model yang dikonfigurasi inferensi. Hasil inferensi dikirim ke penerima keluaran inferensi. Hanya anggota dengan konfigurasi ML yang valid dan kemampuan anggota untuk menerima output inferensi yang dapat menerima hasil inferensi.

Berikut langkah-langkah yang harus diselesaikan oleh penyedia model:

1. Buat image docker Amazon ECR yang kompatibel dengan SageMaker AI. Clean Rooms ML hanya mendukung gambar docker yang kompatibel dengan SageMaker AI.
2. Setelah Anda membuat gambar docker yang kompatibel dengan SageMaker AI, dorong gambar ke Amazon ECR. Ikuti petunjuk di [Amazon Elastic Container Registry User Guide](#) untuk membuat gambar pelatihan kontainer.
3. Konfigurasi algoritma model untuk digunakan di Clean Rooms ML.

- a. Berikan tautan repositori Amazon ECR dan argumen apa pun yang diperlukan untuk mengonfigurasi algoritma model.
- b. Menyediakan peran akses layanan yang memungkinkan Clean Rooms MS mengakses repositori Amazon ECR.
- c. Kaitkan algoritma model yang dikonfigurasi dengan kolaborasi. Ini termasuk menyediakan kebijakan privasi yang mendefinisikan kontrol untuk log kontainer, log kegagalan, CloudWatch metrik, dan batasan tentang berapa banyak data yang dapat diekspor dari hasil penampung.

Berikut adalah langkah-langkah yang harus diselesaikan oleh penyedia data untuk berkolaborasi dengan model HTML kustom:

1. Konfigurasi AWS Glue tabel yang ada dengan aturan analisis kustom. Hal ini memungkinkan serangkaian kueri tertentu yang telah disetujui sebelumnya atau akun yang telah disetujui sebelumnya untuk menggunakan data Anda.
2. Kaitkan tabel yang dikonfigurasi dengan kolaborasi dan berikan peran akses layanan yang dapat mengakses AWS Glue tabel Anda.
3. [Tambahkan aturan analisis kolaborasi](#) ke tabel yang memungkinkan asosiasi algoritma model yang dikonfigurasi untuk mengakses tabel yang dikonfigurasi.
4. Setelah model dan data dikaitkan dan dikonfigurasi di Clean Rooms, anggota dengan kemampuan untuk menjalankan kueri menyediakan kueri SQL dan memilih algoritma model yang akan digunakan.

Setelah pelatihan model selesai, anggota tersebut memulai ekspor artefak pelatihan model atau hasil inferensi. Artefak atau hasil ini dikirim ke anggota dengan kemampuan untuk menerima keluaran model terlatih. Penerima hasil harus mengonfigurasi mereka `MachineLearningConfiguration` sebelum mereka dapat menerima output model.

AWS model dalam Kamar Bersih

AWS Clean Rooms ML menyediakan metode pelestarian privasi bagi dua pihak untuk mengidentifikasi pengguna serupa dalam data mereka tanpa perlu berbagi data satu sama lain. Pihak pertama membawa data pelatihan AWS Clean Rooms sehingga mereka dapat membuat dan mengonfigurasi model yang mirip dan mengaitkannya dengan kolaborasi. Kemudian, data benih dibawa ke kolaborasi untuk membuat segmen mirip yang menyerupai data pelatihan.

Untuk penjelasan lebih rinci tentang cara kerjanya, lihat [Lowongan kerja lintas akun](#).

Topik berikut memberikan informasi tentang cara membuat dan mengkonfigurasi AWS model di Clean Rooms ML.

Topik

- [Perlindungan privasi AWS Clean Rooms](#)
- [Persyaratan data pelatihan untuk Clean Rooms](#)
- [Persyaratan data benih untuk Kamar Bersih ML](#)
- [Metrik evaluasi model AWS Clean Rooms](#)

Perlindungan privasi AWS Clean Rooms

Clean Rooms ML dirancang untuk mengurangi risiko serangan inferensi keanggotaan di mana penyedia data pelatihan dapat mempelajari siapa yang ada dalam data benih dan penyedia data benih dapat mempelajari siapa yang ada dalam data pelatihan. Beberapa langkah diambil untuk mencegah serangan ini.

Pertama, penyedia data benih tidak secara langsung mengamati output Clean Rooms ML dan penyedia data pelatihan tidak akan pernah dapat mengamati data benih. Penyedia data benih dapat memilih untuk memasukkan data benih di segmen output.

Selanjutnya, model mirip dibuat dari sampel acak data pelatihan. Sampel ini mencakup sejumlah besar pengguna yang tidak cocok dengan audiens benih. Proses ini membuat lebih sulit untuk menentukan apakah pengguna tidak ada dalam data, yang merupakan jalan lain untuk inferensi keanggotaan.

Selanjutnya, beberapa pelanggan benih dapat digunakan untuk setiap parameter pelatihan model mirip spesifik benih. Ini membatasi seberapa banyak model yang dapat disesuaikan, dan dengan demikian berapa banyak yang dapat disimpulkan tentang pengguna. Sebagai hasilnya, kami merekomendasikan bahwa ukuran minimum data benih adalah 500 pengguna.

Akhirnya, metrik tingkat pengguna tidak pernah diberikan kepada penyedia data pelatihan, yang menghilangkan jalan lain untuk serangan inferensi keanggotaan.

Persyaratan data pelatihan untuk Clean Rooms

Agar berhasil membuat model yang mirip, data pelatihan Anda harus memenuhi persyaratan berikut:

- Data pelatihan harus dalam format Parquet, CSV, atau JSON.

Note

Data Paket terkompresi Zstandard (ZSTD) tidak didukung.

- Data pelatihan Anda harus dikatalogkan. AWS Glue Untuk informasi selengkapnya, lihat [Memulai Katalog Data AWS Glue](#) di Panduan AWS Glue Pengembang. Sebaiknya gunakan AWS Glue crawler untuk membuat tabel Anda karena skema disimpulkan secara otomatis.
- Bucket Amazon S3 yang berisi data pelatihan dan data benih berada di AWS wilayah yang sama dengan sumber daya Clean Rooms Anda yang lain.
- Data pelatihan harus berisi setidaknya 100.000 pengguna unik IDs dengan setidaknya dua interaksi item masing-masing.
- Data pelatihan harus berisi setidaknya 1 juta catatan.
- Skema yang ditentukan dalam [CreateTrainingDataset](#) tindakan harus sejajar dengan skema yang ditentukan saat AWS Glue tabel dibuat.
- Bidang wajib, sebagaimana didefinisikan dalam tabel yang disediakan, didefinisikan dalam [CreateTrainingDataset](#) tindakan.

Jenis bidang	Jenis data yang didukung	Diperlukan	Deskripsi
USER_ID	string, int, kecil	Ya	Pengidentifikasi unik untuk setiap pengguna dalam kumpulan data. Ini harus menjadi nilai Informasi Identifikasi Non-Pribadi

Jenis bidang	Jenis data yang didukung	Diperlukan	Deskripsi
			di (PII). Ini mungkin mengenal hash atau ID pelanggan.
ITEM_ID	string, int, kecil	Ya	Pengidentifikasi unik untuk setiap item yang berinteraksi dengan pengguna.
TIMESTAMP	bigint, int, stempel waktu	Ya	Waktu ketika pengguna berinteraksi dengan item. Nilai harus dalam waktu epoch Unix dalam format detik.

Jenis bidang	Jenis data yang didukung	Diperlukan	Deskripsi
CATEGORICAL_FEATURE	string, int, float, bigint, ganda, boolean, array	Tidak	Menangkap data kategoris yang terkait dengan pengguna atau item. Ini dapat mencakup hal-hal seperti jenis acara (seperti klik atau pembelian), demografi pengguna (kelompok usia, jenis kelamin - anonim), lokasi pengguna (kota, negara - anonim), kategori barang (seperti pakaian atau elektronik),

Jenis bidang	Jenis data yang didukung	Diperlukan	Deskripsi
			atau merek barang.
NUMERICAL_FEATURE	ganda, mengapung, int, kecil	Tidak	Menangkap data numerik yang terkait dengan pengguna atau item. Ini dapat mencakup hal-hal seperti riwayat pembelian pengguna (jumlah total yang dihabiskan), harga item, berapa kali item dikunjungi, atau peringkat pengguna untuk item.

- Secara opsional, Anda dapat menyediakan hingga 10 fitur kategoris atau numerik total.

Berikut adalah contoh kumpulan data pelatihan yang valid dalam format CSV

```
USER_ID, ITEM_ID, TIMESTAMP, EVENT_TYPE(CATEGORICAL FEATURE), EVENT_VALUE (NUMERICAL FEATURE)
```

```
196,242,881250949,click,15
```

```
186,302,891717742,click,13
```

```
22,377,878887116,click,10
```

```
244,51,880606923,click,20
```

```
166,346,886397596,click,10
```

Persyaratan data benih untuk Kamar Bersih ML

Data benih untuk model yang mirip dapat datang langsung dari bucket Amazon S3 atau dari hasil kueri SQL.

Data benih yang diberikan secara langsung harus memenuhi persyaratan berikut:

- Data benih harus dalam format baris JSON dengan daftar pengguna IDs.
- Ukuran benih harus antara 25 dan 500.000 pengguna IDs unik.
- Jumlah minimum pengguna seed harus sesuai dengan nilai ukuran benih pencocokan minimum yang ditentukan saat Anda membuat model audiens yang dikonfigurasi.

Berikut ini adalah contoh kumpulan data pelatihan yang valid dalam format CSV

```
{"user_id": "abc"}  
{"user_id": "def"}  
{"user_id": "ghijkl"}  
{"user_id": "123"}  
{"user_id": "456"}  
{"user_id": "7890"}
```

Metrik evaluasi model AWS Clean Rooms

Clean Rooms ML menghitung skor recall dan relevansi untuk menentukan seberapa baik kinerja model Anda. Recall membandingkan kesamaan antara data mirip dan data pelatihan. Skor relevansi digunakan untuk memutuskan seberapa besar audiens seharusnya, bukan apakah model tersebut berkinerja baik.

Ingat adalah ukuran yang tidak bias tentang seberapa mirip segmen yang mirip dengan data pelatihan. Recall adalah persentase pengguna yang paling mirip (secara default, 20% paling mirip)

dari sampel data pelatihan yang disertakan dalam audiens benih oleh pekerjaan pembuatan audiens. Nilai berkisar dari 0-1, nilai yang lebih besar menunjukkan audiens yang lebih baik. Nilai recall kira-kira sama dengan persentase bin maksimum menunjukkan bahwa model audiens setara dengan pemilihan acak.

Kami menganggap ini sebagai metrik evaluasi yang lebih baik daripada akurasi, presisi, dan skor F1 karena Clean Rooms ML tidak secara akurat memberi label pengguna negatif sejati saat membangun modelnya.

Skor relevansi tingkat segmen adalah ukuran kesamaan dengan nilai mulai dari -1 (paling tidak mirip) hingga 1 (paling mirip). Clean Rooms ML menghitung serangkaian skor relevansi untuk berbagai ukuran segmen untuk membantu Anda menentukan ukuran segmen terbaik untuk data Anda. Skor relevansi menurun secara monoton seiring bertambahnya ukuran segmen, sehingga seiring bertambahnya ukuran segmen, hal itu bisa kurang mirip dengan data benih. Ketika skor relevansi tingkat segmen mencapai 0, model memprediksi bahwa semua pengguna di segmen mirip berasal dari distribusi yang sama dengan data benih. Meningkatkan ukuran output kemungkinan akan menyertakan pengguna di segmen mirip yang tidak berasal dari distribusi yang sama dengan data benih.

Skor relevansi dinormalisasi dalam satu kampanye dan tidak boleh digunakan untuk membandingkan di seluruh kampanye. Skor relevansi tidak boleh digunakan sebagai bukti bersumber tunggal untuk hasil bisnis apa pun karena dipengaruhi oleh beberapa faktor kompleks selain relevansi, seperti kualitas inventaris, jenis inventaris, waktu iklan, dan sebagainya.

Skor relevansi tidak boleh digunakan untuk menilai kualitas benih, melainkan jika dapat ditingkatkan atau diturunkan. Pertimbangkan contoh berikut:

- Semua skor positif — Ini menunjukkan bahwa ada lebih banyak pengguna keluaran yang diprediksi serupa daripada yang termasuk dalam segmen mirip. Ini umum untuk data benih yang merupakan bagian dari pasar besar, seperti semua orang yang telah membeli pasta gigi dalam sebulan terakhir. Kami merekomendasikan untuk melihat data benih yang lebih kecil, seperti semua orang yang telah membeli pasta gigi lebih dari sekali dalam sebulan terakhir.
- Semua skor negatif atau negatif untuk ukuran segmen mirip yang Anda inginkan — Ini menunjukkan bahwa Clean Rooms MS memprediksi tidak ada cukup pengguna serupa dalam ukuran segmen mirip yang diinginkan. Ini bisa jadi karena data benih terlalu spesifik atau pasarnya terlalu kecil. Kami merekomendasikan untuk menerapkan lebih sedikit filter ke data benih atau memperluas pasar. Misalnya, jika data benih asli adalah pelanggan yang membeli kereta dorong dan kursi mobil, Anda dapat memperluas pasar ke pelanggan yang membeli beberapa produk bayi.

Penyedia data pelatihan menentukan apakah skor relevansi diekspos dan keranjang tempat skor relevansi dihitung.

Model kustom di Kamar Bersih ML

Dengan Clean Rooms, anggota kolaborasi dapat menggunakan algoritma model kustom dockerized yang disimpan di Amazon ECR untuk bersama-sama menganalisis data mereka. Untuk melakukan ini, penyedia model harus membuat gambar dan menyimpannya di Amazon ECR. Ikuti langkah-langkah di [Amazon Elastic Container Registry User Guide](#) untuk membuat repositori pribadi yang akan berisi model HTML kustom.

Setiap anggota kolaborasi dapat menjadi penyedia model, asalkan mereka memiliki izin yang benar. Semua anggota kolaborasi dapat menyumbangkan data ke model. Untuk tujuan panduan ini, data yang berkontribusi anggota disebut sebagai penyedia data. Anggota yang menciptakan kolaborasi adalah pembuat kolaborasi, dan anggota ini dapat berupa penyedia model, salah satu penyedia data, atau keduanya.

Topik berikut menjelaskan informasi yang diperlukan untuk membuat model ML kustom

Topik

- [Prasyarat pemodelan HTML khusus](#)
- [Pedoman penulisan model untuk wadah pelatihan](#)
- [Pedoman pembuatan model untuk wadah inferensi](#)
- [Menerima log dan metrik model](#)

Prasyarat pemodelan HTML khusus

Sebelum Anda dapat melakukan pemodelan HTML kustom, Anda harus mempertimbangkan hal berikut:

- Tentukan apakah pelatihan model dan inferensi pada model yang dilatih akan dilakukan dalam kolaborasi.
- Tentukan peran yang akan dilakukan setiap anggota kolaborasi dan berikan mereka kemampuan yang sesuai.
 - Tetapkan CAN_QUERY kemampuan kepada anggota yang akan melatih model dan menjalankan inferensi pada model yang dilatih.
 - Tetapkan CAN_RECEIVE_RESULTS untuk setidaknya satu anggota kolaborasi.

- Tetapkan `CAN_RECEIVE_MODEL_OUTPUT` atau `CAN_RECEIVE_INFERENCE_OUTPUT` kemampuan kepada anggota yang akan menerima ekspor model terlatih atau output inferensi, masing-masing. Anda dapat memilih untuk menggunakan kedua kemampuan jika diperlukan oleh kasus penggunaan Anda.
- Tentukan ukuran maksimum artefak model terlatih atau hasil inferensi yang akan Anda izinkan untuk diekspor.
- Kami menyarankan agar semua pengguna memiliki `CleanroomsFullAccess` dan `CleanroomsMLFullAccess` kebijakan yang melekat pada peran mereka. Menggunakan model ML khusus memerlukan penggunaan AWS Clean Rooms AWS Clean Rooms dan AWS Clean Rooms SDKs.
- Pertimbangkan informasi berikut tentang peran IAM.
 - Semua penyedia data harus memiliki peran akses layanan yang memungkinkan AWS Clean Rooms untuk membaca data dari AWS Glue katalog dan tabel mereka, dan lokasi Amazon S3 yang mendasarinya. Peran ini mirip dengan yang diperlukan untuk query SQL. Ini memungkinkan Anda untuk menggunakan `CreateConfiguredTableAssociation` tindakan. Untuk informasi selengkapnya, lihat [Membuat peran layanan untuk membuat asosiasi tabel yang dikonfigurasi](#).
 - Semua anggota yang ingin menerima metrik harus memiliki peran akses layanan yang memungkinkan mereka menulis CloudWatch metrik dan log. Peran ini digunakan oleh Clean Rooms ML untuk menulis semua metrik model dan log ke anggota Akun AWS selama pelatihan model dan inferensi. Kami juga menyediakan kontrol privasi untuk menentukan anggota mana yang memiliki akses ke metrik dan log. Ini memungkinkan Anda untuk menggunakan `CreateMLConfiguration` tindakan. Untuk informasi selengkapnya, lihat [Buat peran layanan untuk pemodelan ML kustom - Konfigurasi ML](#).

Anggota yang menerima hasil harus menyediakan peran akses layanan dengan izin untuk menulis ke bucket Amazon S3 mereka. Peran ini memungkinkan Clean Rooms ML untuk mengeksport hasil (artefak model terlatih atau hasil inferensi) ke bucket Amazon S3. Ini memungkinkan Anda untuk menggunakan `CreateMLConfiguration` tindakan. Untuk informasi selengkapnya, lihat [Buat peran layanan untuk pemodelan ML kustom - Konfigurasi ML](#).

- Penyedia model harus menyediakan peran akses layanan dengan izin untuk membaca repositori dan gambar Amazon ECR mereka. Ini memungkinkan Anda untuk menggunakan `CreateConfigureModelAlgorithm` tindakan. Untuk informasi selengkapnya, lihat [Buat peran layanan untuk menyediakan model ML kustom](#).

- Anggota yang membuat `MLInputChannel` untuk menghasilkan kumpulan data untuk pelatihan atau inferensi harus menyediakan peran akses layanan yang memungkinkan Clean Rooms ML untuk menjalankan kueri SQL di AWS Clean Rooms. Ini memungkinkan Anda untuk menggunakan `CreateTrainedModel` dan `StartTrainedModelInferenceJob` tindakan. Untuk informasi selengkapnya, lihat [Membuat peran layanan untuk menanyakan kumpulan data](#).
- Penulis model harus mengikuti [Pedoman penulisan model untuk wadah pelatihan](#) dan [Pedoman pembuatan model untuk wadah inferensi](#) untuk memastikan input dan output model dikonfigurasi seperti yang diharapkan oleh AWS Clean Rooms.

Pedoman penulisan model untuk wadah pelatihan

Bagian ini merinci pedoman yang harus diikuti oleh penyedia model saat membuat algoritme model ML khusus untuk Clean Rooms.

- Gunakan gambar dasar kontainer yang didukung pelatihan SageMaker AI yang sesuai, seperti yang dijelaskan dalam Panduan [Pengembang SageMaker AI](#). Kode berikut memungkinkan Anda untuk menarik gambar dasar kontainer yang didukung dari titik akhir SageMaker AI publik.

```
ecr_registry_endpoint='763104351884.dkr.ecr.$REGION.amazonaws.com'  
base_image='pytorch-training:2.3.0-cpu-py311-ubuntu20.04-sagemaker'  
aws ecr get-login-password --region $REGION | docker login --username AWS --password-  
stdin $ecr_registry_endpoint  
docker pull $ecr_registry_endpoint/$base_image
```

- Saat membuat model secara lokal, pastikan hal berikut sehingga Anda dapat menguji model Anda secara lokal, pada instance pengembangan, pada Pelatihan SageMaker AI di Akun AWS, dan di Clean Rooms MS.
- Kami merekomendasikan menulis skrip pelatihan yang mengakses properti yang berguna tentang lingkungan pelatihan melalui berbagai variabel lingkungan. Clean Rooms MS menggunakan argumen berikut untuk menjalankan pelatihan pada kode model Anda: `SM_MODEL_DIR`, `SM_OUTPUT_DIR`, `SM_CHANNEL_TRAIN`, dan `FILE_FORMAT`. Default ini digunakan oleh Clean Rooms untuk melatih model ML Anda di lingkungan eksekusinya sendiri dengan data dari semua pihak.
- Clean Rooms ML membuat saluran input pelatihan Anda tersedia melalui `/opt/ml/input/data/channel-name` direktori di wadah docker. Setiap saluran input ML dipetakan berdasarkan yang sesuai yang `channel_name` disediakan dalam `CreateTrainedModel` permintaan.

```

parser = argparse.ArgumentParser()# Data, model, and output directories

parser.add_argument('--model_dir', type=str, default=os.environ.get('SM_MODEL_DIR',
"/opt/ml/model"))
parser.add_argument('--output_dir', type=str,
default=os.environ.get('SM_OUTPUT_DIR', "/opt/ml/output/data"))
parser.add_argument('--train_dir', type=str,
default=os.environ.get('SM_CHANNEL_TRAIN', "/opt/ml/input/data/train"))
parser.add_argument('--train_file_format', type=str,
default=os.environ.get('FILE_FORMAT', "csv"))

```

- Pastikan Anda dapat menghasilkan dataset sintetis atau pengujian berdasarkan skema kolaborator yang akan digunakan dalam kode model Anda.
- Pastikan Anda dapat menjalankan pekerjaan pelatihan SageMaker AI sendiri Akun AWS sebelum mengaitkan algoritma model dengan AWS Clean Rooms kolaborasi.

Kode berikut berisi contoh file Docker yang kompatibel dengan pengujian lokal, pengujian lingkungan Pelatihan SageMaker AI, dan Clean Rooms

```

FROM 763104351884.dkr.ecr.us-west-2.amazonaws.com/pytorch-training:2.3.0-cpu-
py311-ubuntu20.04-sagemaker
MAINTAINER $author_name

ENV PYTHONDONTWRITEBYTECODE=1 \
    PYTHONUNBUFFERED=1 \
    LD_LIBRARY_PATH="${LD_LIBRARY_PATH}:/usr/local/lib"

ENV PATH="/opt/ml/code:${PATH}"

# this environment variable is used by the SageMaker PyTorch container to determine
our user code directory
ENV SAGEMAKER_SUBMIT_DIRECTORY /opt/ml/code

# copy the training script inside the container
COPY train.py /opt/ml/code/train.py
# define train.py as the script entry point
ENV SAGEMAKER_PROGRAM train.py
ENTRYPOINT ["python", "/opt/ml/code/train.py"]

```

- Untuk memantau kegagalan kontainer dengan sebaik-baiknya, kami sarankan untuk mengeksport log dan debugging untuk alasan kegagalan. Sebagai `GetTrainedModel` tanggapan, Clean Rooms ML mengembalikan 1024 karakter pertama dari file ini di bawah `StatusDetails`.
- Setelah Anda menyelesaikan perubahan model apa pun dan Anda siap untuk mengujinya di lingkungan SageMaker AI, jalankan perintah berikut dalam urutan yang disediakan.

```
export ACCOUNT_ID=xxx
export REPO_NAME=xxx
export REPO_TAG=xxx
export REGION=xxx

docker build -t $ACCOUNT_ID.dkr.ecr.us-west-2.amazonaws.com/$REPO_NAME:$REPO_TAG

# Sign into AWS $ACCOUNT_ID/ Run aws configure
# Check the account and make sure it is the correct role/credentials
aws sts get-caller-identity
aws ecr create-repository --repository-name $REPO_NAME --region $REGION
aws ecr describe-repositories --repository-name $REPO_NAME --region $REGION

# Authenticate Docker
aws ecr get-login-password --region $REGION | docker login --username AWS --password-
stdin $ACCOUNT_ID.dkr.ecr.$REGION.amazonaws.com

# Push To ECR Images
docker push $ACCOUNT_ID.dkr.ecr.$REGION.amazonaws.com$REPO_NAME:$REPO_TAG

# Create Sagemaker Training job
# Configure the training_job.json with
# 1. TrainingImage
# 2. Input DataConfig
# 3. Output DataConfig
aws sagemaker create-training-job --cli-input-json file://training_job.json --region
$REGION
```

Setelah pekerjaan SageMaker AI selesai dan Anda puas dengan algoritme model Anda, Anda dapat mendaftarkan Amazon ECR Registry dengan AWS Clean Rooms ML. Gunakan `CreateConfiguredModelAlgorithm` tindakan untuk mendaftarkan algoritma model dan `CreateConfiguredModelAlgorithmAssociation` mengaitkannya dengan kolaborasi.

Pedoman pembuatan model untuk wadah inferensi

Bagian ini merinci pedoman yang harus diikuti oleh penyedia model saat membuat algoritme inferensi untuk Clean Rooms ML.

- Gunakan gambar dasar kontainer yang didukung inferensi SageMaker AI yang sesuai, seperti yang dijelaskan dalam Panduan Pengembang [SageMaker AI](#). Kode berikut memungkinkan Anda untuk menarik gambar dasar kontainer yang didukung dari titik akhir SageMaker AI publik.

```
ecr_registry_endpoint='763104351884.dkr.ecr.$REGION.amazonaws.com'  
base_image='pytorch-inference:2.3.0-cpu-py311-ubuntu20.04-sagemaker'  
aws ecr get-login-password --region $REGION | docker login --username AWS --password-  
stdin $ecr_registry_endpoint  
docker pull $ecr_registry_endpoint/$base_image
```

- Saat membuat model secara lokal, pastikan hal berikut sehingga Anda dapat menguji model Anda secara lokal, pada instance pengembangan, pada Transformasi SageMaker Batch AI di Akun AWS, dan di Clean Rooms MS.
 - Clean Rooms ML membuat artefak model Anda dari inferensi tersedia untuk digunakan oleh kode inferensi Anda melalui /opt/ml/model direktori di wadah docker.
 - Clean Rooms ML membagi input demi baris, menggunakan strategi MultiRecord batch, dan menambahkan karakter baris baru di akhir setiap catatan yang diubah.
 - Pastikan Anda dapat menghasilkan kumpulan data inferensi sintetis atau pengujian berdasarkan skema kolaborator yang akan digunakan dalam kode model Anda.
 - Pastikan Anda dapat menjalankan pekerjaan transformasi batch SageMaker AI sendiri Akun AWS sebelum mengaitkan algoritme model dengan AWS Clean Rooms kolaborasi.

Kode berikut berisi contoh file Docker yang kompatibel dengan pengujian lokal, pengujian lingkungan transformasi SageMaker AI, dan Clean Rooms

```
FROM 763104351884.dkr.ecr.us-east-1.amazonaws.com/pytorch-inference:1.12.1-cpu-  
py38-ubuntu20.04-sagemaker  
  
ENV PYTHONUNBUFFERED=1  
  
COPY serve.py /opt/ml/code/serve.py  
COPY inference_handler.py /opt/ml/code/inference_handler.py  
COPY handler_service.py /opt/ml/code/handler_service.py  
COPY model.py /opt/ml/code/model.py
```

```
RUN chmod +x /opt/ml/code/serve.py
```

```
ENTRYPOINT ["/opt/ml/code/serve.py"]
```

- Setelah Anda menyelesaikan perubahan model apa pun dan Anda siap untuk mengujinya di lingkungan SageMaker AI, jalankan perintah berikut dalam urutan yang disediakan.

```
export ACCOUNT_ID=xxx
export REPO_NAME=xxx
export REPO_TAG=xxx
export REGION=xxx

docker build -t $ACCOUNT_ID.dkr.ecr.us-west-2.amazonaws.com/$REPO_NAME:$REPO_TAG

# Sign into AWS $ACCOUNT_ID/ Run aws configure
# Check the account and make sure it is the correct role/credentials
aws sts get-caller-identity
aws ecr create-repository --repository-name $REPO_NAME --region $REGION
aws ecr describe-repositories --repository-name $REPO_NAME --region $REGION

# Authenticate Docker
aws ecr get-login-password --region $REGION | docker login --username AWS --password-stdin $ACCOUNT_ID.dkr.ecr.$REGION.amazonaws.com

# Push To ECR Repository
docker push $ACCOUNT_ID.dkr.ecr.$REGION.amazonaws.com$REPO_NAME:$REPO_TAG

# Create Sagemaker Model
# Configure the create_model.json with
# 1. Primary container -
#    a. ModelDataUrl - S3 Uri of the model.tar from your training job
aws sagemaker create-model --cli-input-json file://create_model.json --region $REGION

# Create Sagemaker Transform Job
# Configure the transform_job.json with
# 1. Model created in the step above
# 2. MultiRecord batch strategy
# 3. Line SplitType for TransformInput
# 4. AssembleWith Line for TransformOutput
aws sagemaker create-transform-job --cli-input-json file://transform_job.json --region $REGION
```


Setelah pekerjaan SageMaker AI selesai dan Anda puas dengan transformasi batch Anda, Anda dapat mendaftarkan Amazon ECR Registry dengan AWS Clean Rooms ML. Gunakan `CreateConfiguredModelAlgorithm` tindakan untuk mendaftarkan algoritma model dan `CreateConfiguredModelAlgorithmAssociation` mengaitkannya dengan kolaborasi.

Menerima log dan metrik model

Untuk menerima log dan metrik dari pelatihan atau inferensi model kustom, anggota harus telah [membuat Konfigurasi ML](#) dengan peran valid yang memberikan CloudWatch izin yang diperlukan (lihat [Membuat peran layanan untuk pemodelan HTML kustom - Konfigurasi ML](#)).

Metrik sistem

Metrik sistem untuk pelatihan dan inferensi, seperti CPU dan pemanfaatan memori, dipublikasikan ke semua anggota dalam kolaborasi dengan Konfigurasi ML yang valid. Metrik ini dapat dilihat saat pekerjaan berlangsung melalui CloudWatch Metrik di `/aws/cleanroomsm1/TrainedModels` atau ruang `/aws/cleanroomsm1/TrainedModelInferenceJobs` nama, masing-masing.

Log model

Akses ke log model disediakan oleh kebijakan konfigurasi privasi dari setiap algoritma model yang dikonfigurasi. Penulis model menetapkan kebijakan konfigurasi privasi saat mengaitkan algoritma model yang dikonfigurasi (baik melalui konsol atau `CreateConfiguredModelAlgorithmAssociation` API) ke kolaborasi. Menyetel kebijakan konfigurasi privasi mengontrol anggota mana yang dapat menerima log model.

Selain itu, pembuat model dapat mengatur pola filter dalam kebijakan konfigurasi privasi untuk memfilter peristiwa log. Semua log yang dikirim oleh wadah model ke `stdout` atau `stderr` dan yang cocok dengan pola filter (jika disetel), dikirim ke Amazon CloudWatch Logs. Log model tersedia dalam grup CloudWatch log `/aws/cleanroomsm1/TrainedModels` atau `/aws/cleanroomsm1/TrainedModelInferenceJobs`, masing-masing.

Metrik yang ditentukan khusus

Saat Anda mengonfigurasi algoritme model (baik melalui konsol atau `CreateConfiguredModelAlgorithm` API), pembuat model dapat memberikan nama metrik dan pernyataan regex tertentu untuk dicari di log keluaran. Ini dapat dilihat saat pekerjaan berlangsung melalui CloudWatch Metrik di namespace. `/aws/cleanroomsm1/TrainedModels` Saat mengaitkan algoritme model yang dikonfigurasi, pembuat model dapat menetapkan tingkat

kebisingan opsional dalam konfigurasi privasi metrik untuk menghindari keluaran data mentah sambil tetap memberikan visibilitas ke tren metrik khusus. Jika tingkat kebisingan diatur, metrik dipublikasikan di akhir pekerjaan daripada secara real time.

Komputasi Kriptografi untuk Clean Rooms

Komputasi Kriptografi untuk Clean Rooms (C3R) adalah kemampuan AWS Clean Rooms yang dapat digunakan selain aturan [analisis](#). Dengan C3R, organisasi dapat menyatukan data sensitif untuk memperoleh wawasan baru dari analitik data sementara secara kriptografis membatasi apa yang dapat dipelajari oleh pihak mana pun dalam prosesnya. C3R dapat digunakan oleh dua pihak atau lebih yang ingin berkolaborasi dengan data sensitif mereka tetapi diharuskan hanya menggunakan data terenkripsi di cloud.

Klien enkripsi C3R adalah alat enkripsi sisi klien yang dapat Anda gunakan untuk [mengenkripsi data](#) Anda untuk digunakan. AWS Clean Rooms Saat Anda menggunakan klien enkripsi C3R, data tetap dilindungi secara kriptografis saat digunakan dalam kolaborasi. AWS Clean Rooms Seperti halnya AWS Clean Rooms kolaborasi reguler, data input adalah tabel database relasional, dan komputasi dinyatakan sebagai query SQL. Namun, C3R hanya mendukung subset terbatas dari kueri SQL pada data terenkripsi.

Secara khusus, C3R mendukung SQL JOIN and SELECT pernyataan tentang data yang dilindungi secara kriptografi. Setiap kolom dalam tabel input dapat digunakan tepat di salah satu jenis pernyataan SQL berikut:

- Kolom yang dilindungi secara kriptografi untuk digunakan JOIN Pernyataan disebut fingerprint kolom.
- Kolom yang dilindungi secara kriptografi untuk digunakan SELECT Pernyataan disebut sealed kolom.
- Kolom yang tidak dilindungi secara kriptografi untuk digunakan JOIN atau SELECT Pernyataan disebut cleartext kolom.

Dalam beberapa kasus, GROUP BY pernyataan didukung pada fingerprint kolom. Untuk informasi selengkapnya, lihat [Fingerprintkolom](#). Saat ini, C3R tidak mendukung penggunaan konstruksi SQL lainnya pada data terenkripsi, seperti WHERE klausa atau fungsi agregat seperti SUM and AVERAGE, bahkan jika mereka diizinkan oleh aturan analisis yang relevan.

C3R dirancang untuk melindungi data dalam sel individual tabel. Menggunakan konfigurasi default untuk C3R, data dasar yang disediakan pelanggan kepada pihak ketiga melalui kolaborasi tetap

dienkripsi saat konten sedang digunakan di dalamnya. AWS Clean Rooms C3R menggunakan enkripsi AES-GCM standar industri untuk semua sealed kolom dan fungsi pseudorandom standar industri, yang dikenal sebagai Hash Based Message Authentication Code (HMAC), untuk melindungi fingerprint kolom.

Meskipun C3R mengenkripsi data dalam tabel Anda, informasi berikut mungkin masih dapat disimpulkan:

- Informasi tentang tabel itu sendiri, termasuk jumlah kolom, nama kolom, dan jumlah baris dalam tabel Anda.
- Seperti kebanyakan bentuk enkripsi standar, C3R tidak mencoba menyembunyikan panjang nilai terenkripsi. C3R memang menawarkan kemampuan untuk memasukkan nilai terenkripsi untuk menyembunyikan panjang teks yang tepat. Namun, batas atas pada panjang cleartext di setiap kolom masih bisa diungkapkan ke pihak lain.
- Informasi tingkat pencatatan, seperti ketika baris tertentu ditambahkan ke tabel C3R terenkripsi.

Untuk informasi selengkapnya tentang C3R, lihat topik berikut.

Topik

- [Pertimbangan saat menggunakan Komputasi Kriptografi untuk Clean Rooms](#)
- [Jenis file dan data yang didukung dalam Komputasi Kriptografi untuk Clean Rooms](#)
- [Nama kolom dalam Komputasi Kriptografi untuk Clean Rooms](#)
- [Jenis kolom dalam Komputasi Kriptografi untuk Clean Rooms](#)
- [Parameter komputasi kriptografi](#)
- [Bendera opsional dalam Komputasi Kriptografi untuk Clean Rooms](#)
- [Kueri dengan Komputasi Kriptografi untuk Clean Rooms](#)
- [Pedoman untuk klien enkripsi C3R](#)

Pertimbangan saat menggunakan Komputasi Kriptografi untuk Clean Rooms

Cryptographic Computing for Clean Rooms (C3R) berupaya memaksimalkan perlindungan data. Namun, beberapa kasus penggunaan mungkin mendapat manfaat dari tingkat perlindungan data yang lebih rendah dengan imbalan fungsionalitas tambahan. Anda dapat membuat pengorbanan khusus ini dengan memodifikasi C3R dari konfigurasi yang paling aman. Sebagai pelanggan, Anda

harus menyadari pengorbanan ini dan menentukan apakah mereka sesuai untuk kasus penggunaan Anda. Pengorbanan untuk dipertimbangkan meliputi yang berikut:

Topik

- [Mengizinkan data campuran cleartext dan terenkripsi dalam tabel Anda](#)
- [Mengizinkan nilai berulang dalam fingerprint kolom](#)
- [Melonggarkan pembatasan tentang bagaimana fingerprint kolom diberi nama](#)
- [Menentukan bagaimana NULL nilai direpresentasikan](#)

Untuk informasi selengkapnya tentang cara mengatur parameter untuk skenario ini, lihat [Parameter komputasi kriptografi](#).

Mengizinkan data campuran cleartext dan terenkripsi dalam tabel Anda

Memiliki semua data dienkripsi sisi klien memberikan perlindungan data maksimum. Namun, ini membatasi jenis kueri tertentu (misalnya, fungsi SUM agregat). Risiko mengizinkan cleartext data adalah layak bahwa siapa pun yang memiliki akses ke tabel terenkripsi dapat menyimpulkan beberapa informasi tentang nilai terenkripsi. Ini dapat dilakukan dengan melakukan analisis statistik pada cleartext dan data terkait.

Misalnya, bayangkan Anda memiliki kolom `City` dan `State`. `City` kolom adalah cleartext dan `State` kolom dienkripsi. Ketika Anda melihat nilai `Chicago` di `City` kolom, itu membantu Anda menentukan dengan probabilitas tinggi bahwa `State` itu `Illinois`. Sebaliknya, jika satu kolom `City` dan kolom lainnya `EmailAddress`, a cleartext `City` tidak mungkin mengungkapkan apa pun tentang terenkripsi `EmailAddress`.

Untuk informasi selengkapnya tentang parameter untuk skenario ini, lihat [izinkan parameter cleartext kolom](#).

Mengizinkan nilai berulang dalam fingerprint kolom

Untuk pendekatan yang paling aman, kami berasumsi bahwa fingerprint kolom apa pun berisi persis satu instance variabel. Tidak ada item yang dapat diulang dalam fingerprint kolom. Klien enkripsi C3R memetakan cleartext nilai-nilai ini menjadi nilai unik yang tidak dapat dibedakan dari nilai acak. Oleh karena itu, tidak mungkin untuk menyimpulkan informasi tentang cleartext dari nilai-nilai acak ini.

Risiko nilai berulang dalam fingerprint kolom adalah bahwa nilai berulang akan menghasilkan nilai yang tampak acak berulang. Dengan demikian, siapa pun yang memiliki akses ke tabel terenkripsi

dapat, secara teori, melakukan analisis statistik fingerprint kolom yang mungkin mengungkapkan informasi tentang cleartext nilai.

Sekali lagi, misalkan fingerprint kolomnya `State`, dan setiap baris tabel sesuai dengan rumah tangga AS. Dengan melakukan analisis frekuensi, seseorang dapat menyimpulkan keadaan mana California dan mana Wyoming dengan probabilitas tinggi. Kesimpulan ini dimungkinkan karena California memiliki lebih banyak penduduk daripada Wyoming. Sebaliknya, katakanlah fingerprint kolom berada pada pengidentifikasi rumah tangga dan setiap rumah tangga muncul dalam database antara 1 dan 4 kali dalam database jutaan entri. Tidak mungkin analisis frekuensi akan mengungkapkan informasi yang berguna.

Untuk informasi selengkapnya tentang parameter untuk skenario ini, lihat [Izinkan parameter duplikat](#).

Melonggarkan pembatasan tentang bagaimana fingerprint kolom diberi nama

Secara default, kami berasumsi bahwa ketika dua tabel digabungkan menggunakan kolom terenkripsi, fingerprint kolom tersebut memiliki nama yang sama di setiap tabel. Alasan teknis untuk hasil ini adalah bahwa, secara default, kami memperoleh kunci kriptografi yang berbeda untuk mengenkripsi setiap kolom. fingerprint Kunci itu berasal dari kombinasi kunci rahasia bersama untuk kolaborasi dan nama kolom. Jika kami mencoba menggabungkan dua kolom dengan nama kolom yang berbeda, kami memperoleh kunci yang berbeda dan kami tidak dapat menghitung gabungan yang valid.

Untuk mengatasi masalah ini, Anda dapat menonaktifkan fitur yang memperoleh kunci dari setiap nama kolom. Kemudian, klien enkripsi C3R menggunakan kunci turunan tunggal untuk semua fingerprint kolom. Risikonya adalah bahwa jenis lain dari analisis frekuensi dapat dilakukan yang mungkin mengungkapkan informasi.

Mari kita gunakan `State` contoh `City` dan lagi. Jika kita memperoleh nilai acak yang sama untuk setiap fingerprint kolom (dengan tidak memasukkan nama kolom). New York memiliki nilai acak yang sama di `State` kolom `City` dan. New York adalah salah satu dari beberapa kota di AS di mana `City` namanya sama dengan `State` namanya. Sebaliknya, jika kumpulan data Anda memiliki nilai yang sama sekali berbeda di setiap kolom, tidak ada informasi yang bocor.

Untuk informasi selengkapnya tentang parameter untuk skenario ini, lihat [Izinkan JOIN kolom dengan parameter nama yang berbeda](#).

Menentukan bagaimana NULL nilai direpresentasikan

Opsi yang tersedia untuk Anda adalah apakah akan memproses nilai kriptografi (enkripsi dan HMAC) seperti NULL nilai lainnya. Jika Anda tidak memproses NULL nilai seperti nilai lainnya, informasi mungkin akan terungkap.

Misalnya, anggaplah bahwa NULL di Middle Name kolom di cleartext menunjukkan orang tanpa nama tengah. Jika Anda tidak mengenkripsi nilai-nilai tersebut, Anda membocorkan baris mana dalam tabel terenkripsi yang digunakan untuk orang tanpa nama tengah. Informasi itu mungkin menjadi sinyal pengenalan bagi beberapa orang di beberapa populasi. Tetapi jika Anda memproses NULL nilai secara kriptografi, kueri SQL tertentu bertindak berbeda. Misalnya, GROUP BY klausa tidak akan mengelompokkan fingerprint NULL nilai dalam fingerprint kolom bersama-sama.

Untuk informasi selengkapnya tentang parameter untuk skenario ini, lihat [Pertahankan parameter NULL nilai](#).

Jenis file dan data yang didukung dalam Komputasi Kriptografi untuk Clean Rooms

Klien enkripsi C3R mengenali jenis file berikut:

- Berkas CSV
- Parquetberkas

Anda dapat menggunakan `--fileFormat` bendera di klien enkripsi C3R untuk menentukan format file secara eksplisit. Ketika ditentukan secara eksplisit, format file tidak ditentukan oleh ekstensi file.

Topik

- [Berkas CSV](#)
- [Parquetberkas](#)
- [Mengekripsi nilai non-string](#)

Berkas CSV

File dengan ekstensi.csv diasumsikan berformat CSV dan berisi teks yang dikodekan UTF-8. Klien enkripsi C3R memperlakukan semua nilai sebagai string.

Properti yang didukung dalam file.csv

Klien enkripsi C3R mensyaratkan bahwa file.csv memiliki properti berikut:

- Mungkin atau mungkin tidak berisi baris header awal yang secara unik menamai setiap kolom.
- Dibatasi koma. (Saat ini, pembatas khusus tidak didukung.)
- Teks yang dikodekan UTF-8.

Pemangkasan ruang putih dari entri .csv

Spasi putih depan dan belakang dipangkas dari entri .csv.

NULLPengkodean khusus untuk file.csv

File.csv dapat menggunakan pengkodean khususNULL.

Dengan klien enkripsi C3R, Anda dapat menentukan pengkodean kustom untuk NULL entri dalam data input dengan menggunakan bendera. `--csvInputNULLValue=<csv-input-null>` Klien enkripsi C3R dapat menggunakan pengkodean khusus dalam file keluaran yang dihasilkan untuk entri NULL dengan menggunakan bendera. `--csvOutputNULLValue=<csv-output-null>`

Note

NULLEntri dianggap kurang konten, khususnya dalam konteks format tabel yang lebih kaya seperti tabel SQL. Meskipun .csv tidak secara eksplisit mendukung karakterisasi ini karena alasan historis, itu adalah konvensi umum untuk mempertimbangkan entri kosong yang hanya berisi spasi putih. NULL Oleh karena itu, itulah perilaku default klien enkripsi C3R dan dapat disesuaikan sesuai kebutuhan.

Bagaimana entri .csv ditafsirkan oleh C3R

Tabel berikut memberikan contoh bagaimana entri .csv disusun (cleartextcleartextuntuk kejelasan) berdasarkan nilai (jika ada) yang disediakan untuk dan flag. `--csvInputNULLValue=<csv-input-null>` `--csvOutputNULLValue=<csv-output-null>` Memimpin dan membuntuti ruang putih di luar tanda kutip dipangkas sebelum C3R menafsirkan makna nilai apa pun.

<csv-input-null>	<csv-output-null>	Masukan entri	Entri keluaran
Tidak ada	Tidak ada	,AnyProduct,	,AnyProduct,
Tidak ada	Tidak ada	, AnyProduct ,	,AnyProduct,
Tidak ada	Tidak ada	,"AnyProduct",	,AnyProduct,
Tidak ada	Tidak ada	, "AnyProdu ct" ,	,AnyProduct,
Tidak ada	Tidak ada	,,	,,
Tidak ada	Tidak ada	, ,	,,
Tidak ada	Tidak ada	, "",	,,
Tidak ada	Tidak ada	, " ",	, " ",
Tidak ada	Tidak ada	, " " ,	, " ",
"AnyProduct"	"NULL"	,AnyProduct,	,NULL,
"AnyProduct"	"NULL"	, AnyProduct ,	,NULL,
"AnyProduct"	"NULL"	,"AnyProduct",	,NULL,
"AnyProduct"	"NULL"	, "AnyProdu ct" ,	,NULL,
Tidak ada	"NULL"	,,	,NULL,
Tidak ada	"NULL"	, ,	,NULL,
Tidak ada	"NULL"	, "",	,NULL,
Tidak ada	"NULL"	, " ",	, " ",
Tidak ada	"NULL"	, " " ,	, " ",
""	"NULL"	,,	,NULL,

<csv-input-null>	<csv-output-null>	Masukan entri	Entri keluaran
""	"NULL"	, ,	,NULL,
""	"NULL"	, "",	, "",
""	"NULL"	, " ",	, " ",
""	"NULL"	, " " ,	, " " ,
"\""	"NULL"	, ,	, ,
"\""	"NULL"	, ,	, ,
"\""	"NULL"	, "",	,NULL,
"\""	"NULL"	, " ",	, " ",
"\""	"NULL"	, " " ,	, " " ,

File CSV tanpa header

File sumber.csv tidak perlu memiliki header di baris pertama yang secara unik memberi nama setiap kolom. Namun, file.csv tanpa baris header memerlukan skema enkripsi posisi. Skema enkripsi posisi diperlukan alih-alih skema pemetaan khas yang digunakan untuk file.csv dengan baris header dan file. Parquet

Skema enkripsi posisi menentukan kolom keluaran berdasarkan posisi, bukan dengan nama. Skema enkripsi yang dipetakan memetakan nama kolom sumber untuk menargetkan nama kolom. Untuk informasi lebih lanjut, termasuk diskusi rinci dan contoh dari kedua format skema, lihat [Skema tabel yang dipetakan dan posisi](#).

Parquetberkas

File dengan .parquet ekstensi diasumsikan dalam Apache Parquet format.

Tipe Parquet data yang didukung

Klien enkripsi C3R dapat memproses data non-kompleks (yaitu tipe primitif) dalam Parquet file yang mewakili tipe data yang didukung oleh. AWS Clean Rooms

Namun, hanya kolom string yang dapat digunakan untuk sealed kolom.

Tipe data Parquet berikut didukung:

- Binarytipe primitif dengan anotasi logis berikut:
 - Tidak ada jika `--parquetBinaryAsString` diatur (tipe STRING data)
 - `Decimal(scale, precision)`(tipe DECIMAL data)
 - `String`(tipe STRING data)
- Booleantipe data primitif tanpa anotasi logis (tipe BOOLEAN data)
- Doubletipe data primitif tanpa anotasi logis (tipe DOUBLE data)
- `Fixed_Len_Binary_Array`tipe primitif dengan anotasi `Decimal(scale, precision)` logis (tipe DECIMAL data)
- Floattipe data primitif tanpa anotasi logis (tipe FLOAT data)
- Int32tipe primitif dengan anotasi logis berikut:
 - Tidak ada (tipe INT data)
 - `Date`(tipe DATE data)
 - `Decimal(scale, precision)`(tipe DECIMAL data)
 - `Int(16, true)`(tipe SMALLINT data)
 - `Int(32, true)`(tipe INT data)
- Int64tipe data primitif dengan anotasi logis berikut:
 - Tidak ada (tipe BIGINT data)
 - `Decimal(scale, precision)`(tipe DECIMAL data)
 - `Int(64, true)`(tipe BIGINT data)
 - `Timestamp(isUTCAdjusted, TimeUnit.MILLIS)`(tipe TIMESTAMP data)
 - `Timestamp(isUTCAdjusted, TimeUnit.MICROS)`(tipe TIMESTAMP data)
 - `Timestamp(isUTCAdjusted, TimeUnit.NANOS)`(tipe TIMESTAMP data)

Mengenkripsi nilai non-string

Saat ini, hanya nilai string yang didukung untuk sealed kolom.

Untuk file.csv, klien enkripsi C3R memperlakukan semua nilai sebagai teks yang dikodekan UTF-8 dan tidak berusaha untuk menafsirkannya secara berbeda sebelum enkripsi.

Untuk kolom sidik jari, jenis dikelompokkan ke dalam kelas ekivalensi. Kelas kesetaraan adalah sekumpulan tipe data yang dapat dibandingkan secara jelas untuk kesetaraan melalui tipe data yang representatif.

Kelas kesetaraan memungkinkan sidik jari identik untuk ditetapkan ke nilai semantik yang sama terlepas dari representasi aslinya. Namun, nilai yang sama dalam dua kelas ekivalensi tidak akan menghasilkan kolom sidik jari yang sama.

Misalnya, INTEGRAL nilai 42 akan diberikan sidik jari yang sama terlepas dari apakah itu awalnya SMALLINT, INT, atau BIGINT. Juga, INTEGRAL nilai tidak 0 akan pernah cocok dengan BOOLEAN nilai FALSE (yang diwakili oleh nilai 0).

Kelas kesetaraan berikut dan tipe AWS Clean Rooms data yang sesuai didukung oleh kolom sidik jari:

Kelas kesetaraan	Tipe AWS Clean Rooms data yang didukung
BOOLEAN	BOOLEAN
DATE	DATE
INTEGRAL	BIGINT, INT, SMALLINT
STRING	CHAR, STRING, VARCHAR

Nama kolom dalam Komputasi Kriptografi untuk Clean Rooms

Secara default, nama-nama kolom penting dalam Komputasi Kriptografi untuk Clean Rooms.

Jika nilai Izinkan JOIN kolom dengan parameter nama yang berbeda salah, nama kolom digunakan selama enkripsi fingerprint kolom. Untuk alasan ini, secara default, kolaborator harus berkoordinasi terlebih dahulu dan menggunakan nama kolom target yang sama untuk data yang akan menggunakan JOIN pernyataan dalam kueri. Secara default, kolom yang dienkripsi JOIN dengan nama berbeda tidak berhasil JOIN pada nilai apa pun.

Jika nilai Izinkan JOIN kolom dengan parameter nama yang berbeda benar, JOIN pernyataan di seluruh kolom dienkripsi sebagai fingerprint kolom berhasil. Mengenkripsi data dengan parameter ini mungkin memungkinkan beberapa inferensi nilai. cleartext Misalnya, jika baris memiliki nilai Kode

Otentikasi Pesan (HMAC) berbasis Hash yang sama di City kolom dan State kolom, nilainya mungkin. New York

Normalisasi nama header kolom

Nama header kolom dinormalisasi oleh klien enkripsi C3R. Setiap spasi putih depan dan belakang dihapus, dan nama kolom dibuat huruf kecil untuk output yang diubah.

Normalisasi diterapkan sebelum semua perhitungan, perhitungan, atau operasi lain yang mungkin dapat dipengaruhi oleh nama kolom. File keluaran yang dipancarkan hanya berisi nama yang dinormalisasi.

Jenis kolom dalam Komputasi Kriptografi untuk Clean Rooms

Topik ini memberikan informasi tentang jenis kolom dalam Komputasi Kriptografi untuk Clean Rooms.

Topik

- [Fingerprintkolom](#)
- [Kolom tertutup](#)
- [Cleartextkolom](#)

Fingerprintkolom

Fingerprintkolom adalah kolom yang dilindungi secara kriptografis untuk digunakan dalam JOIN pernyataan.

Data dari fingerprint kolom tidak dapat didekripsi. Hanya data dari kolom tertutup yang dapat didekripsi.

Fingerprintkolom hanya boleh digunakan dalam klausa dan fungsi SQL berikut:

- JOIN (INNER, OUTER, LEFT, RIGHT, or FULL) terhadap fingerprint kolom lain:
 - Jika nilai `allowJoinsOnColumnsWithDifferentNames` parameter diatur ke `false`, kedua fingerprint kolom juga JOIN harus memiliki nama yang sama.
- SELECT COUNT()
- SELECT COUNT(DISTINCT)
- GROUP BY (Hanya gunakan jika kolaborasi telah menetapkan nilai `preserveNulls` parameter ke `true`.)

Kueri yang melanggar batasan ini mungkin menghasilkan hasil yang salah.

Kolom tertutup

Kolom tertutup adalah kolom yang dilindungi secara kriptografis untuk digunakan dalam SELECT pernyataan.

Kolom tertutup hanya boleh digunakan dalam klausa dan fungsi SQL berikut:

- SELECT
- SELECT ... AS
- SELECT COUNT()

Note

SELECT COUNT(DISTINCT) tidak didukung.

Kueri yang melanggar batasan ini mungkin menghasilkan hasil yang salah.

Padding data untuk sealed kolom sebelum enkripsi

Saat Anda menentukan bahwa kolom harus berupa sealed kolom, C3R menanyakan padding jenis apa yang harus dipilih. Padding data sebelum enkripsi adalah opsional. Tanpa padding (tipe padnone), panjang data terenkripsi menunjukkan ukuran. cleartext Dalam beberapa keadaan, ukuran cleartext dapat mengekspos plaintext. Dengan padding (tipe pad fixed atau max), semua nilai pertama-tama diempuk ke ukuran umum dan kemudian dienkripsi. Dengan padding, panjang data terenkripsi tidak memberikan informasi tentang cleartext panjang aslinya, selain memberikan batas atas pada ukurannya.

Jika Anda ingin padding untuk kolom dan panjang byte maksimal data di kolom itu diketahui, gunakan fixed padding. Gunakan length nilai yang setidaknya sebesar panjang byte dari nilai terpanjang di kolom itu.

Note

Terjadi kesalahan dan enkripsi gagal jika nilai lebih panjang dari yang disediakan length.

Jika Anda ingin padding untuk kolom dan panjang byte maksimal data di kolom itu tidak diketahui, gunakan max padding. Mode padding ini membungkus semua data dengan panjang nilai terpanjang ditambah length byte tambahan.

Note

Anda mungkin ingin mengenkripsi data dalam batch, atau memperbarui tabel Anda dengan data baru secara berkala. Ketahuilah bahwa max padding akan memasukkan entri ke panjang (plus length byte) dari entri plaintext terpanjang dalam batch tertentu. Ini berarti bahwa panjang ciphertext dapat bervariasi dari batch ke batch. Oleh karena itu, jika Anda mengetahui panjang byte maksimum untuk kolom, maka Anda harus menggunakan fixed sebagai gantinya. max

Cleartextkolom

Cleartextkolom adalah kolom yang tidak dilindungi secara kriptografis untuk digunakan dalam JOIN atau SELECT pernyataan.

Cleartextkolom dapat digunakan di bagian manapun dari query SQL.

Parameter komputasi kriptografi

[Parameter komputasi kriptografi tersedia untuk kolaborasi menggunakan Cryptographic Computing for Clean Rooms \(C3R\) saat membuat kolaborasi.](#) Anda dapat membuat kolaborasi menggunakan AWS Clean Rooms konsol atau operasi CreateCollaboration API. Di konsol, Anda dapat mengatur nilai untuk parameter dalam parameter komputasi kriptografi setelah Anda mengaktifkan opsi komputasi kriptografi Support. Untuk informasi selengkapnya, lihat topik berikut.

Topik

- [Izinkan parameter cleartext kolom](#)
- [Izinkan parameter duplikat](#)
- [Izinkan JOIN kolom dengan parameter nama yang berbeda](#)
- [Pertahankan parameter NULL nilai](#)

Izinkan parameter cleartext kolom

Di konsol, Anda dapat mengatur parameter Izinkan cleartext kolom saat [membuat kolaborasi](#) untuk menentukan apakah cleartext data diizinkan dalam tabel dengan data terenkripsi.

Tabel berikut menjelaskan nilai-nilai untuk parameter Izinkan cleartext kolom.

Nilai parameter	Deskripsi
Tidak	Cleartextkolom tidak diizinkan dalam tabel terenkripsi. Semua data dilindungi secara kriptografi.
Ya	Cleartextkolom diperbolehkan dalam tabel terenkripsi. Cleartextkolom tidak dilindungi secara kriptografis dan disertakan sebagai. cleartext Anda harus mencatat apa yang mungkin diungkapkan oleh cleartext data baris Anda tentang data lain dalam tabel. Untuk menjalankan SUM atau AVG pada kolom tertentu, kolom harus masukcleartext.

Menggunakan operasi CreateCollaboration API, untuk dataEncryptionMetadata parameter, Anda dapat mengatur nilai allowCleartext ke true atau false. Untuk informasi selengkapnya tentang operasi API, lihat [Referensi AWS Clean Rooms API](#).

Cleartextkolom sesuai dengan kolom yang diklasifikasikan sebagai cleartext dalam skema khusus tabel. Data dalam kolom ini tidak dienkripsi dan dapat digunakan dengan cara apa pun. Cleartextkolom dapat berguna jika data tidak sensitif and/or jika lebih banyak fleksibilitas diperlukan daripada sealed kolom atau fingerprint kolom terenkripsi memungkinkan.

Izinkan parameter duplikat

Di konsol, Anda dapat mengatur parameter Izinkan duplikat saat [membuat kolaborasi untuk menentukan apakah kolom yang](#) dienkripsi untuk JOIN kueri dapat berisi duplikat non-nilai. NULL

⚠ Important

Parameter Izinkan duplikat, [JOINIzinkan kolom dengan nama yang berbeda](#), dan [Pertahankan NULL nilai](#) memiliki efek terpisah tetapi terkait.

Tabel berikut menjelaskan nilai untuk parameter Izinkan duplikat.

Nilai parameter	Deskripsi
Tidak	Nilai berulang tidak diperbolehkan dalam fingerprint kolom. Semua nilai dalam satu fingerprint kolom harus unik.
Ya	Nilai berulang diperbolehkan dalam fingerprint kolom. Jika Anda perlu menggabungkan kolom dengan nilai berulang, atur nilai ini ke Ya. Ketika diatur ke Ya, pola frekuensi yang muncul dalam fingerprint kolom dalam tabel C3R atau hasil mungkin menyiratkan beberapa informasi tambahan tentang struktur data. cleartext

Menggunakan operasi CreateCollaboration API, untuk dataEncryptionMetadata parameter Anda dapat mengatur nilai allowDuplicates ke true atau false. Untuk informasi selengkapnya tentang operasi API, lihat [Referensi AWS Clean Rooms API](#).

Secara default, jika data terenkripsi harus digunakan dalam JOIN kueri, klien enkripsi C3R mengharuskan kolom tersebut tidak memiliki nilai duplikat. Persyaratan ini merupakan upaya untuk meningkatkan perlindungan data. Perilaku ini dapat membantu memastikan bahwa pola berulang dalam data tidak dapat diamati. Namun, jika Anda ingin bekerja dengan data terenkripsi dalam JOIN kueri dan tidak peduli tentang nilai duplikat, parameter Izinkan duplikat dapat menonaktifkan pemeriksaan konservatif ini.

Izinkan JOIN kolom dengan parameter nama yang berbeda

Di konsol, Anda dapat mengatur parameter Izinkan JOIN kolom dengan nama yang berbeda saat [membuat kolaborasi](#) untuk menentukan apakah JOIN pernyataan antara kolom dengan nama berbeda didukung.

Untuk informasi selengkapnya, lihat [Normalisasi nama header kolom](#)

Tabel berikut menjelaskan nilai untuk Izinkan JOIN kolom dengan parameter nama yang berbeda.

Nilai parameter	Deskripsi
Tidak	Gabungan fingerprint kolom dengan nama berbeda tidak didukung. JOIN pernyataan hanya memberikan hasil yang akurat pada kolom yang memiliki nama yang sama.  Important Nilai No memberikan peningkatan keamanan informasi tetapi mengharuskan peserta kolaborasi untuk menyetujui sebelumnya tentang nama kolom. Jika dua kolom memiliki nama yang berbeda ketika dienkripsi sebagai fingerprint kolom dan Izinkan JOIN kolom dengan nama yang berbeda diatur ke Tidak, JOIN pernyataan pada kolom tersebut tidak menghasilkan hasil. Ini karena tidak ada nilai pasca-enkripsi yang dibagi di antara mereka.
Ya	Gabungan fingerprint kolom dengan nama berbeda didukung. Untuk fleksibilitas tambahan, pengguna dapat mengatur nilai ini ke Ya, yang memungkinkan JOIN pernyataan pada kolom terlepas dari nama kolom mereka. Jika disetel ke Ya, klien enkripsi C3R tidak mempertimbangkan nama kolom saat melindungi fingerprint kolom. Akibatnya, nilai umum di fingerprint kolom yang berbeda dapat diamati dalam tabel C3R. Misalnya, jika baris memiliki JOIN nilai terenkripsi yang sama di City kolom dan State kolom, mungkin masuk akal untuk menyimpulkan bahwa nilainya adalah New York

Menggunakan operasi `CreateCollaboration` API, untuk `dataEncryptionMetadata` parameter, Anda dapat mengatur nilai `allowJoinsOnColumnsWithDifferentNames` ke `true` atau `false`. Untuk informasi selengkapnya tentang operasi API, lihat [Referensi AWS Clean Rooms API](#).

Secara default, enkripsi fingerprint kolom dipengaruhi oleh kolom `targetHeader` untuk itu, diatur [Langkah 4: Buat skema enkripsi untuk file tabular](#). Oleh karena itu, cleartext nilai yang sama memiliki representasi terenkripsi yang berbeda di setiap fingerprint kolom yang berbeda yang dienkripsi untuknya.

Parameter ini dapat berguna untuk mencegah inferensi cleartext nilai dalam beberapa kasus. Misalnya, melihat nilai terenkripsi yang sama di fingerprint kolom `City` dan `State` dapat digunakan untuk menyimpulkan nilainya secara wajar. New York Namun, penggunaan parameter ini memerlukan koordinasi tambahan terlebih dahulu, sehingga semua kolom yang akan digabungkan dalam kueri memiliki nama bersama.

Anda dapat menggunakan Izinkan JOIN kolom dengan parameter nama yang berbeda untuk melonggarkan batasan ini. Ketika nilai parameter disetel ke `Yes`, ini memungkinkan kolom apa pun yang dienkripsi JOIN untuk digunakan bersama terlepas dari nama.

Pertahankan parameter NULL nilai

Di konsol, Anda dapat mengatur parameter Pertahankan NULL nilai saat [membuat kolaborasi](#) untuk menunjukkan bahwa tidak ada nilai yang ada untuk kolom tersebut.

Tabel berikut menjelaskan nilai untuk parameter `Preserve NULL values`.

Nilai parameter	Deskripsi
Tidak	NULL nilai-nilai tidak dipertahankan. NULL nilai tidak muncul seperti NULL dalam tabel terenkripsi. NULL nilai muncul sebagai nilai acak unik dalam tabel C3R.
Ya	NULL nilai-nilai dipertahankan. NULL nilai muncul seperti NULL dalam tabel terenkripsi. Jika Anda memerlukan semantik SQL NULL nilai, Anda dapat mengatur nilai ini ke Ya. Akibatnya, NULL entri muncul seperti NULL pada tabel C3R, terlepas dari apakah kolom dienkripsi dan terlepas dari pengaturan parameter untuk Izinkan duplikat.

Menggunakan operasi `CreateCollaboration` API, untuk `dataEncryptionMetadata` parameter, Anda dapat mengatur nilai `preserveNulls` ke `true` atau `false`. Untuk informasi selengkapnya tentang operasi API, lihat [Referensi AWS Clean Rooms API](#).

Saat parameter `Pertahankan NULL` nilai disetel ke `Tidak` untuk kolaborasi:

1. `NULL` entri dalam `cleartext` kolom tidak berubah.
2. `NULL` entri dalam `fingerprint` kolom terenkripsi dienkripsi sebagai nilai acak untuk menyembunyikan isinya. Bergabung di kolom terenkripsi dengan `NULL` entri di `cleartext` kolom tidak menghasilkan kecocokan apa pun untuk entri mana pun. `NULL` Tidak ada kecocokan yang dibuat karena mereka masing-masing menerima konten acak unik mereka sendiri.
3. `NULL` entri dalam `sealed` kolom terenkripsi dienkripsi.

Ketika nilai parameter `Pertahankan NULL` nilai disetel ke `Ya` untuk kolaborasi, `NULL` entri dari semua kolom tetap sebagai `NULL` terlepas dari apakah kolom dienkripsi.

Parameter `Preserve NULL values` berguna dalam skenario seperti pengayaan data, di mana Anda ingin berbagi kekurangan informasi yang dinyatakan sebagai `NULL`. Parameter `Preserve NULL values` juga berguna dalam format `fingerprint` atau `HMAC` jika Anda memiliki `NULL` nilai di kolom yang Anda inginkan `JOIN` atau `GROUP BY`.

Jika nilai parameter `Izinkan duplikat` dan `Pertahankan NULL` nilai diatur ke `Tidak`, memiliki lebih dari satu `NULL` entri dalam `fingerprint` kolom menghasilkan kesalahan dan menghentikan enkripsi. Jika nilai salah satu parameter disetel ke `Ya`, tidak ada kesalahan seperti itu terjadi.

Bendera opsional dalam Komputasi Kriptografi untuk Clean Rooms

Bagian berikut menjelaskan flag opsional yang dapat Anda atur saat Anda [mengenkripsi data menggunakan klien enkripsi](#) C3R untuk kustomisasi dan pengujian file tabular.

Topik

- [--csvInputNULLValuebendera](#)
- [--csvOutputNULLValuebendera](#)
- [--enableStackTracesbendera](#)
- [--dryRunbendera](#)
- [--tempDirbendera](#)

--csvInputNULLValuebendera

Anda dapat menggunakan --csvInputNULLValue bendera untuk menentukan pengkodean kustom untuk NULL entri dalam data input saat Anda [mengenkripsi data menggunakan klien enkripsi C3R](#).

Tabel berikut merangkum penggunaan dan parameter flag ini.

Penggunaan	Parameter
Tidak wajib. Pengguna dapat menentukan pengkodean khusus untuk NULL entri dalam data input.	Pengkodean NULL nilai yang ditentukan pengguna dalam file CSV input

NULLEntri adalah entri yang dianggap kurang konten, khususnya dalam konteks format tabel yang lebih kaya seperti tabel SQL. Meskipun .csv tidak secara eksplisit mendukung karakterisasi ini karena alasan historis, itu adalah konvensi umum untuk mempertimbangkan entri kosong yang hanya berisi spasi putih. NULL Oleh karena itu, itulah perilaku default klien enkripsi C3R dan dapat disesuaikan sesuai kebutuhan.

--csvOutputNULLValuebendera

Anda dapat menggunakan --csvOutputNULLValue bendera untuk menentukan pengkodean kustom untuk NULL entri dalam data keluaran saat Anda [mengenkripsi data menggunakan klien enkripsi C3R](#).

Tabel berikut merangkum penggunaan dan parameter flag ini.

Penggunaan	Parameter
Tidak wajib. Pengguna dapat menentukan pengkodean khusus dalam file keluaran yang dihasilkan untuk NULL entri.	Pengkodean NULL nilai yang ditentukan pengguna dalam file CSV keluaran

NULLEntri adalah entri yang dianggap kurang konten, khususnya dalam konteks format tabel yang lebih kaya seperti tabel SQL. Meskipun .csv tidak secara eksplisit mendukung karakterisasi ini karena alasan historis, itu adalah konvensi umum untuk mempertimbangkan entri kosong yang hanya berisi

spasi putih. NULL Oleh karena itu, itulah perilaku default klien enkripsi C3R dan dapat disesuaikan sesuai kebutuhan.

--enableStackTracesbendera

Saat Anda [mengenkripsi data](#) menggunakan klien enkripsi C3R, gunakan --enableStackTraces tanda untuk memberikan informasi kontekstual tambahan untuk pelaporan kesalahan saat C3R menemukan kesalahan.

AWS tidak mengumpulkan kesalahan. Jika Anda mengalami kesalahan, gunakan jejak tumpukan untuk memecahkan masalah kesalahan sendiri atau mengirim jejak tumpukan Dukungan untuk mendapatkan bantuan.

Tabel berikut merangkum penggunaan dan parameter flag ini.

Penggunaan	Parameter
Tidak wajib. Digunakan untuk memberikan informasi kontekstual tambahan untuk pelaporan kesalahan ketika klien enkripsi C3R mengalami kesalahan.	Tidak ada

--dryRunbendera

[Enkripsi](#) dan [dekripsi perintah klien enkripsi](#) C3R menyertakan bendera opsional. --dryRun Bendera mengambil semua argumen yang disediakan pengguna dan memeriksa validitas dan konsistensi.

Anda dapat menggunakan --dryRun bendera untuk memeriksa apakah file skema Anda valid dan konsisten dengan file input yang sesuai.

Tabel berikut merangkum penggunaan dan parameter flag ini.

Penggunaan	Parameter
Tidak wajib. Menyebabkan klien enkripsi C3R mengurai parameter dan memeriksa file, tetapi tidak melakukan enkripsi atau dekripsi.	Tidak ada

--tempDir bendera

Anda mungkin ingin menggunakan direktori sementara karena file terenkripsi terkadang bisa lebih besar dari file yang tidak dienkripsi, tergantung pada pengaturannya. Kumpulan data juga harus dienkripsi per kolaborasi agar berfungsi dengan benar.

Saat Anda [mengenkripsi data](#) menggunakan C3R, gunakan --tempDir bendera untuk menentukan lokasi di mana file sementara dapat dibuat saat memproses input.

Tabel berikut merangkum penggunaan dan parameter flag ini.

Penggunaan	Parameter
Pengguna dapat menentukan lokasi di mana file sementara dapat dibuat saat memproses input.	Default ke direktori sementara sistem.

Kueri dengan Komputasi Kriptografi untuk Clean Rooms

Topik ini memberikan informasi tentang menulis kueri yang menggunakan tabel data yang telah dienkripsi menggunakan Cryptographic Computing untuk. Clean Rooms

Topik

- [Kueri yang bercabang di NULL](#)
- [Memetakan satu kolom sumber ke beberapa kolom target](#)
- [Menggunakan data yang sama untuk keduanya JOIN dan SELECT kueri](#)

Kueri yang bercabang di NULL

Untuk memiliki cabang kueri pada NULL pernyataan berarti menggunakan sintaks seperti IF x IS NULL THEN 0 ELSE 1.

Kueri selalu dapat bercabang pada NULL pernyataan di cleartext kolom.

Kueri dapat bercabang pada NULL pernyataan di sealed kolom dan fingerprint kolom hanya jika nilai parameter Preserve NULL values (preserveNulls) diatur ke. true

Kueri yang melanggar batasan ini mungkin menghasilkan hasil yang salah.

Memetakan satu kolom sumber ke beberapa kolom target

Satu kolom sumber dapat dipetakan ke beberapa kolom target. Misalnya, Anda mungkin ingin keduanya JOIN dan SELECT pada kolom.

Untuk informasi selengkapnya, lihat [Menggunakan data yang sama untuk keduanya JOIN dan SELECT kueri](#).

Menggunakan data yang sama untuk keduanya JOIN dan SELECT kueri

Jika data dalam kolom tidak sensitif, itu dapat muncul di kolom cleartext target, yang memungkinkannya digunakan untuk tujuan apa pun.

Jika data dalam kolom sensitif dan harus digunakan untuk keduanya JOIN dan SELECT kueri, petakan kolom sumber itu ke dua kolom target dalam file output. Satu kolom dienkripsi dengan type sebagai fingerprint kolom, dan satu kolom dienkripsi dengan kolom type sebagai tertutup. Pembuatan skema interaktif dari klien enkripsi C3R menyarankan sufiks header dan. `_fingerprint` `_sealed` Sufiks header ini dapat menjadi konvensi yang berguna untuk membedakan kolom tersebut dengan cepat.

Pedoman untuk klien enkripsi C3R

Klien enkripsi C3R adalah alat yang memungkinkan organisasi untuk menyatukan data sensitif untuk mendapatkan wawasan baru dari analisis data. Alat ini secara kriptografis membatasi apa yang dapat dipelajari oleh pihak mana pun dan AWS dalam prosesnya. Meskipun ini sangat penting, proses pengamanan data secara kriptografis dapat menambah overhead yang signifikan baik dalam hal sumber daya komputasi maupun penyimpanan. Oleh karena itu, penting untuk memahami pengorbanan menggunakan setiap pengaturan dan cara mengoptimalkan pengaturan sambil tetap mempertahankan jaminan kriptografi yang diinginkan. Topik ini berfokus pada implikasi kinerja dari pengaturan yang berbeda dalam klien dan skema enkripsi C3R.

Semua pengaturan enkripsi klien enkripsi C3R memberikan jaminan kriptografi yang berbeda. Pengaturan tingkat kolaborasi paling aman secara default. Mengaktifkan fungsionalitas tambahan sambil membuat kolaborasi melemahkan jaminan privasi, memungkinkan aktivitas seperti analisis frekuensi dilakukan pada ciphertext. Untuk informasi lebih lanjut tentang bagaimana pengaturan ini digunakan dan apa implikasinya, lihat [the section called “Komputasi kriptografi”](#).

Topik

- [Implikasi kinerja untuk jenis kolom](#)

- [Memecahkan masalah peningkatan ukuran ciphertext yang tidak terduga](#)

Implikasi kinerja untuk jenis kolom

C3R menggunakan tiga jenis kolom: cleartext, fingerprint, dan sealed. Masing-masing jenis kolom ini memberikan jaminan kriptografi yang berbeda dan memiliki tujuan penggunaan yang berbeda. Pada bagian berikut, implikasi kinerja dari jenis kolom dibahas dan dampak kinerja dari setiap pengaturan.

Topik

- [Cleartextkolom](#)
- [Fingerprintkolom](#)
- [Sealedkolom](#)

Cleartextkolom

Cleartextkolom tidak diubah dari format aslinya dan tidak diproses secara kriptografi dengan cara apa pun. Jenis kolom ini tidak dapat dikonfigurasi dan tidak memengaruhi kinerja penyimpanan atau komputasi.

Fingerprintkolom

Fingerprintkolom dimaksudkan untuk digunakan untuk menggabungkan data di beberapa tabel. Untuk tujuan ini, ukuran ciphertext yang dihasilkan harus selalu sama. Namun, kolom ini dipengaruhi oleh pengaturan tingkat kolaborasi. Fingerprintkolom mungkin memiliki berbagai tingkat dampak pada ukuran file output tergantung pada yang cleartext terkandung dalam input.

Topik

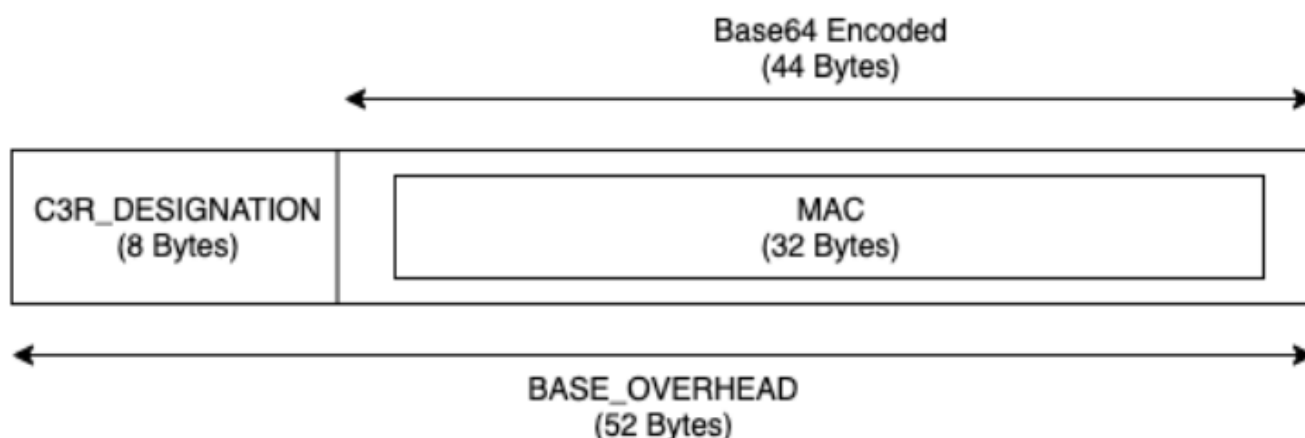
- [Basis overhead untuk kolom fingerprint](#)
- [Pengaturan kolaborasi untuk fingerprint kolom](#)
- [Contoh data untuk fingerprint kolom](#)
- [Kolom pemecahan masalah fingerprint](#)

Basis overhead untuk kolom fingerprint

Ada overhead dasar untuk fingerprint kolom. Overhead ini konstan dan menggantikan ukuran cleartext byte.

Data dalam fingerprint kolom diproses secara kriptografis melalui fungsi Kode Otentikasi Pesan berbasis Hash (HMAC), yang mengubah data menjadi kode otentikasi pesan 32 byte (MAC). Data ini kemudian diproses melalui encoder base64, menambahkan sekitar 33 persen ke ukuran byte. Ini pra-penandaan dengan penunjukan C3R 8 byte untuk menunjuk jenis kolom yang dimiliki data dan versi klien yang menghasilkannya. Hasil akhirnya adalah 52 byte. Hasil ini kemudian dikalikan dengan jumlah baris untuk mendapatkan total overhead basis (gunakan jumlah total null non-nilai jika `preserveNulls` disetel ke `true`).

Gambar berikut menunjukkan bagaimana $BASE\_OVERHEAD = C3R\_DESIGNATION + (MAC * 1.33)$



Output ciphertext dalam fingerprint kolom akan selalu 52 byte. Ini bisa menjadi penurunan penyimpanan yang signifikan jika cleartext data input rata-rata lebih dari 52 byte (misalnya, alamat jalan lengkap). Ini bisa menjadi peningkatan penyimpanan yang signifikan jika cleartext data input rata-rata kurang dari 52 byte (misalnya, usia pelanggan).

Pengaturan kolaborasi untuk fingerprint kolom

Setelan **preserveNulls**

Ketika pengaturan `preserveNulls` tingkat kolaborasi `false` (default), setiap null nilai diganti dengan 32 byte acak yang unik dan diproses seolah-olah tidak null. Hasilnya adalah bahwa setiap null nilai sekarang 52 byte. Ini dapat menambahkan persyaratan penyimpanan yang signifikan untuk tabel yang berisi data yang sangat jarang dibandingkan dengan saat pengaturan ini `true` dan null nilai dilewatkan sebagai null.

Jika Anda tidak memerlukan jaminan privasi dari pengaturan ini dan lebih memilih untuk mempertahankan null nilai dalam kumpulan data Anda, aktifkan `preserveNulls` pengaturan pada saat kolaborasi dibuat. `preserveNulls` Pengaturan tidak dapat diubah setelah kolaborasi dibuat.

Contoh data untuk fingerprint kolom

Berikut ini adalah contoh kumpulan data input dan output untuk fingerprint kolom dengan pengaturan untuk mereproduksi. Pengaturan tingkat kolaborasi lainnya menyukai `allowCleartext` dan `allowDuplicates` tidak memengaruhi hasil dan dapat disetel sebagai `true` atau `false` jika mencoba mereproduksi secara lokal.

Contoh rahasia bersama: `wJa1rXUtnFEMI/K7MDENG/bPxRfiCYEXAMPLEKEY`

Contoh ID kolaborasi: `a1b2c3d4-5678-90ab-cdef-EXAMPLE11111`

`allowJoinsOnColumnsWithDifferentNames`: Pengaturan `True` ini tidak memengaruhi kinerja atau persyaratan penyimpanan. Namun, pengaturan ini membuat pilihan nama kolom tidak relevan saat mereproduksi nilai yang ditunjukkan dalam tabel berikut.

Contoh 1

Input	<code>null</code>
<code>preserveNulls</code>	<code>TRUE</code>
Output	<code>null</code>
Deterministik	<code>Yes</code>
Byte masukan	<code>0</code>
Byte keluaran	<code>0</code>

Contoh 2

Input	<code>null</code>
<code>preserveNulls</code>	<code>FALSE</code>
Output	<code>01: hmac: 31kFjthvV3IUu6mMvFc1a +XAHwgw/E1m0q4p3Yg25kk=</code>
Deterministik	<code>No</code>
Byte masukan	<code>0</code>

Byte keluaran	52

Contoh 3

Input	empty string
preserveNulls	-
Output	01:hmac:oKTgi3Gba+eUb3JteSz 2EMgXUkF1WgM77UP0Ydw5kPQ=
Deterministik	Yes
Byte masukan	0
Byte keluaran	52

Contoh 4

Input	abcdefghijklmnopqrstuvwxyz
preserveNulls	-
Output	01:hmac:kU/IqwG7FMmzzshr0B9 scomE0UJUEE7j9keTctp1Gww=
Deterministik	Yes
Byte masukan	26
Byte keluaran	52

Contoh 5

Input	abcdefghijklmnopqrstuvwxyzA BCDEFGHIJKLMNOPQRSTUVWXYZ01 23456789

<code>preserveNulls</code>	-
Output	01: hmac: ks3htnQbw2vdhCRFF6J NzW5LMndJaHG57uvE26mBtSs=
Deterministik	Yes
Byte masukan	62
Byte keluaran	52

Kolom pemecahan masalah fingerprint

Mengapa ciphertext di fingerprint kolom saya beberapa kali lebih besar dari ukuran cleartext yang masuk ke dalamnya?

Ciphertext dalam fingerprint kolom selalu 52 byte panjangnya. Jika data input Anda kecil (misalnya, usia pelanggan), itu akan menunjukkan peningkatan ukuran yang signifikan. Ini juga bisa terjadi jika `preserveNulls` pengaturan diatur ke `false`.

Mengapa ciphertext di fingerprint kolom saya beberapa kali lebih kecil dari ukuran cleartext yang masuk ke dalamnya?

Ciphertext dalam fingerprint kolom selalu 52 byte panjangnya. Jika data input Anda besar (misalnya, alamat jalan lengkap pelanggan), itu akan menunjukkan penurunan ukuran yang signifikan.

Bagaimana saya tahu jika saya membutuhkan jaminan kriptografi yang disediakan oleh?

`preserveNulls`

Sayangnya, jawabannya adalah itu tergantung. Minimal, [the section called “Parameter”](#) harus ditinjau untuk bagaimana `preserveNulls` pengaturan melindungi data Anda. Namun, kami menyarankan Anda untuk mereferensikan persyaratan penanganan data organisasi Anda dan kontrak apa pun yang berlaku untuk kolaborasi masing-masing.

Mengapa saya harus mengeluarkan biaya overhead base64?

Untuk memungkinkan kompatibilitas dengan format file tabular seperti CSV, pengkodean base64 diperlukan. Meskipun beberapa format file seperti Parquet mungkin mendukung representasi biner data, penting bahwa semua peserta dalam kolaborasi mewakili data dengan cara yang sama untuk memastikan hasil kueri yang tepat.

Sealedkolom

Sealedkolom dimaksudkan untuk digunakan untuk mentransfer data antara anggota kolaborasi. Ciphertext dalam kolom ini bersifat non-deterministik dan memiliki dampak signifikan pada kinerja dan penyimpanan berdasarkan bagaimana kolom dikonfigurasi. Kolom ini dapat dikonfigurasi secara individual dan seringkali memiliki dampak terbesar pada kinerja klien enkripsi C3R dan ukuran file keluaran yang dihasilkan.

Topik

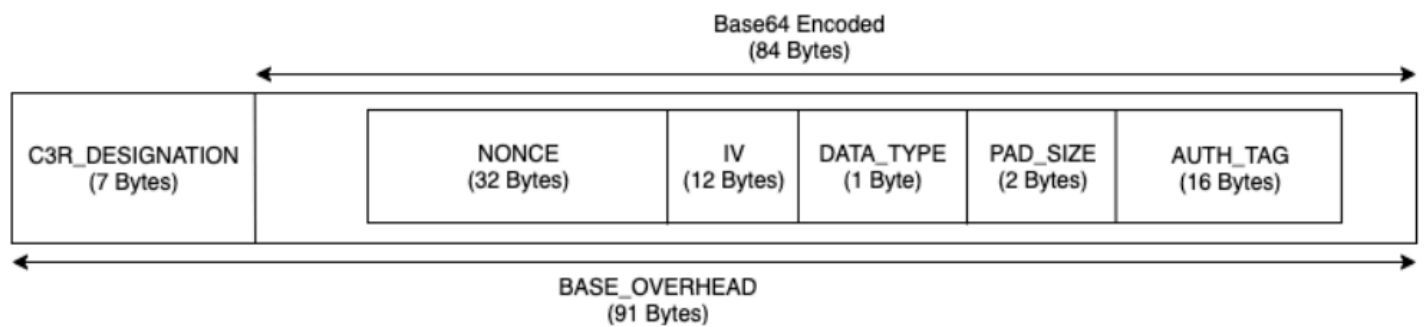
- [Basis overhead untuk kolom sealed](#)
- [Pengaturan kolaborasi untuk sealed kolom](#)
- [sealedKolom pengaturan skema: jenis padding](#)
- [Contoh data untuk sealed kolom](#)
- [Kolom pemecahan masalah sealed](#)

Basis overhead untuk kolom sealed

Ada overhead dasar untuk sealed kolom. Overhead ini konstan dan selain ukuran byte cleartext dan padding (jika ada).

Sebelum enkripsi apa pun, data dalam sealed kolom pra-pended dengan karakter 1 byte yang menunjuk jenis data apa yang terkandung. Jika padding dipilih, data kemudian empuk dan ditambahkan dengan 2 byte yang menyatakan ukuran pad. Setelah byte ini ditambahkan, data diproses secara kriptografi dengan menggunakan AES-GCM dan disimpan dengan IV (12 byte), (32 byte), dan nonce (16 byte). Auth Tag Data ini kemudian diproses melalui encoder base64, menambahkan sekitar 33 persen ke ukuran byte. Data pra-penandaan dengan penunjukan C3R 7 byte untuk menentukan jenis kolom apa yang dimiliki data dan versi klien yang digunakan untuk memproduksinya. Hasilnya adalah overhead basis akhir 91 byte. Hasil ini kemudian dapat dikalikan dengan jumlah baris untuk mendapatkan total overhead basis (gunakan jumlah total nilai non-null jika `preserveNulls` disetel ke `true`).

Gambar berikut menunjukkan bagaimana $BASE\_OVERHEAD = C3R\_DESIGNATION + ((NONCE + IV + DATA\_TYPE + PAD\_SIZE + AUTH\_TAG) * 1.33)$



Pengaturan kolaborasi untuk sealed kolom

Setelan **preserveNulls**

Ketika pengaturan `preserveNulls` tingkat kolaborasi `false` (default), setiap `null` nilai unik, acak 32 byte dan diproses seolah-olah tidak. `null` Hasilnya adalah bahwa setiap `null` nilai sekarang 91 byte (lebih jika empuk). Ini dapat menambahkan persyaratan penyimpanan yang signifikan untuk tabel yang berisi data yang sangat jarang dibandingkan dengan saat pengaturan ini `true` dan `null` nilai dilewatkan sebagai `null`.

Jika Anda tidak memerlukan jaminan privasi dari pengaturan ini dan lebih memilih untuk mempertahankan `null` nilai dalam kumpulan data Anda, aktifkan `preserveNulls` pengaturan pada saat kolaborasi dibuat. `preserveNulls` Pengaturan tidak dapat diubah setelah kolaborasi dibuat.

sealedKolom pengaturan skema: jenis padding

Topik

- [Jenis pad none](#)
- [Jenis pad fixed](#)
- [Jenis pad max](#)

Jenis pad **none**

Memilih jenis pad `none` tidak menambahkan padding apa pun ke cleartext dan tidak menambahkan overhead tambahan ke overhead dasar yang dijelaskan sebelumnya. Tidak ada padding yang menghasilkan ukuran output yang paling hemat ruang. Namun, itu tidak memberikan jaminan privasi yang sama dengan tipe `fixed` dan `max` padding. Ini karena ukuran yang mendasarinya cleartext dapat dilihat dari ukuran ciphertext.

Jenis pad **fixed**

Memilih jenis pad **fixed** adalah ukuran pelestarian privasi untuk menyembunyikan panjang data yang terkandung dalam kolom. Hal ini dilakukan dengan padding semua cleartext ke yang disediakan `pad_length` sebelum dienkripsi. Setiap data yang melebihi ukuran itu menyebabkan klien enkripsi C3R gagal.

Mengingat bahwa padding ditambahkan ke cleartext sebelum dienkripsi, AES-GCM memiliki pemetaan 1-ke-1 untuk byte ciphertext. cleartext Pengkodean base64 akan menambah 33 persen. Overhead penyimpanan tambahan dari padding dapat dihitung dengan mengurangi panjang rata-rata dari cleartext dari nilai `pad_length` dan mengalikannya dengan 1,33. Hasilnya adalah overhead rata-rata padding per record. Hasil ini kemudian dapat dikalikan dengan jumlah baris untuk mendapatkan overhead padding total (gunakan jumlah total `null` non-nilai jika `preserveNulls` diatur ke). `true`

$$PADDING\_OVERHEAD = (PAD\_LENGTH - AVG\_CLEARTEXT\_LENGTH) * 1.33 * ROW\_COUNT$$

Kami menyarankan Anda memilih minimum `pad_length` yang mencakup nilai terbesar dalam kolom. Misalnya, jika nilai terbesar adalah 50 byte, a `pad_length` dari 50 sudah cukup. Nilai yang lebih besar dari itu hanya akan menambah overhead penyimpanan tambahan.

Padding tetap tidak menambahkan overhead komputasi yang signifikan.

Jenis pad **max**

Memilih jenis pad **max** adalah ukuran pelestarian privasi untuk menyembunyikan panjang data yang terkandung dalam kolom. Hal ini dilakukan dengan padding semua cleartext ke nilai terbesar di kolom ditambah tambahan `pad_length` sebelum dienkripsi. Umumnya, max padding memberikan jaminan yang sama dengan **fixed** padding untuk satu kumpulan data sementara memungkinkan untuk tidak mengetahui nilai terbesar cleartext di kolom. Namun, max padding mungkin tidak memberikan jaminan privasi yang sama seperti **fixed** padding di seluruh pembaruan karena nilai terbesar dalam kumpulan data individu mungkin berbeda.

Kami menyarankan Anda memilih tambahan `pad_length` 0 saat menggunakan max padding. Panjang ini bantalan semua nilai menjadi ukuran yang sama dengan nilai terbesar di kolom. Nilai yang lebih besar dari itu hanya akan menambah overhead penyimpanan tambahan.

Jika cleartext nilai terbesar diketahui untuk kolom tertentu, kami sarankan Anda menggunakan jenis **fixed** pad sebagai gantinya. Menggunakan **fixed** padding menciptakan konsistensi di seluruh kumpulan data yang diperbarui. Menggunakan max padding menghasilkan setiap subset data yang diempuk ke nilai terbesar yang ada di subset.

Contoh data untuk sealed kolom

Berikut ini adalah contoh kumpulan data input dan output untuk sealed kolom dengan pengaturan untuk mereproduksi. Pengaturan tingkat kolaborasi lainnya seperti `allowClearText`, `allowJoinsOnColumnsWithDifferentNames`, dan `allowDuplicates` tidak memengaruhi hasil dan dapat disetel sebagai `true` atau `false` jika mencoba mereproduksi secara lokal. Meskipun ini adalah pengaturan dasar untuk mereproduksi, sealed kolom tidak deterministik dan nilai akan berubah setiap saat. Tujuannya adalah untuk menunjukkan byte dalam dibandingkan dengan byte keluar. Contoh `pad_length` nilai dipilih dengan sengaja. Mereka menunjukkan bahwa `fixed` padding menghasilkan nilai yang sama dengan `max` padding dengan `pad_length` pengaturan minimum yang disarankan atau ketika padding tambahan diinginkan.

Contoh rahasia bersama: `wJa1rXUtnFEMI/K7MDENG/bPxRfiCYEXAMPLEKEY`

Contoh ID kolaborasi: `a1b2c3d4-5678-90ab-cdef-EXAMPLE11111`

Topik

- [Jenis pad none](#)
- [Jenis pad fixed \(Contoh 1\)](#)
- [Jenis pad fixed \(Contoh 2\)](#)
- [Jenis pad max \(Contoh 1\)](#)
- [Jenis pad max \(Contoh 2\)](#)

Jenis pad **none**

Contoh 1

Input	<code>null</code>
<code>preserveNulls</code>	<code>TRUE</code>
Output	<code>null</code>
Deterministik	<code>Yes</code>
Byte masukan	<code>0</code>
Byte keluaran	<code>0</code>

Contoh 2

Input	null
preserveNulls	FALSE
Output	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmlmDEyMzQ1Njc40TBqfRY Z98t5KU6aWfssGSPbNIJfG3iXmu 6cbCUrizuV
Deterministik	No
Byte masukan	0
Byte keluaran	91

Contoh 3

Input	empty string
preserveNulls	-
Output	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmlmDEyMzQ1Njc40TBqfRY Z98t5KU6aWfstGSPeM6qR8DWC2P B2GM1X41YK
Deterministik	No
Byte masukan	0
Byte keluaran	91

Contoh 4

Input	abcdefghijklmnopqrstuvwxyz
preserveNulls	-

Output	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfstEE1GKEPiRzyh0 h7t60mWMLTWCv02ckr6pkx9sGL5 VLDQeHzh6DmPpyWNuI=
Deterministik	No
Byte masukan	26
Byte keluaran	127

Contoh 5

Input	abcdefghijklmnopqrstuvwxyza BCDEFGHIJKLMNOPQRSTUVWXYZ01 23456789
preserveNulls	-
Output	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfstEE1GKEPiRzyh0 h7t60mWMLTWCv02ckr6plwtH/8t RFnn2rF91bcB9G4+n8GiRfJNmqdP4/ Q0Q3cXb/pbvPcnnohrHIGSX54ua+1/ JfcVjc=
Deterministik	No
Byte masukan	62
Byte keluaran	175

Jenis pad **fixed** (Contoh 1)

Dalam contoh ini, pad_length adalah 62 dan masukan terbesar adalah 62 byte.

Contoh 1

Input	<code>null</code>
<code>preserveNulls</code>	<code>TRUE</code>
Output	<code>null</code>
Deterministik	<code>Yes</code>
Byte masukan	<code>0</code>
Byte keluaran	<code>0</code>

Contoh 2

Input	<code>null</code>
<code>preserveNulls</code>	<code>FALSE</code>
Output	<code>01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc40TBqfRYZ98t5KU6aWfssGSNWfMRp7nSb7SMX2s3JKLOhK1+7r75Tk+Mx9jy48Fcg1y0PvBqRSZ7oqy1V3UKfYTLEZb/hCz7oaIneVsrcoNpATs0GzbnLkor4L+/aSuA=</code>
Deterministik	<code>No</code>
Byte masukan	<code>0</code>
Byte keluaran	<code>175</code>

Contoh 3

Input	<code>empty string</code>
<code>preserveNulls</code>	<code>-</code>

Output	<pre> 01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4 0TBqfRY Z98t5KU6aWfstGSNWfMRp7nSb7S MX2s3JKL0hK1+7r75Tk+Mx9jy48 Fcg1y0PvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcolB53l07VZp A60wkuXu29CA= </pre>
Deterministik	No
Byte masukan	0
Byte keluaran	175

Contoh 4

Input	abcdefghijklmnopqrstuvwxyz
preserveNulls	-
Output	<pre> 01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4 0TBqfRY Z98t5KU6aWfstEE1GKEPiRzyh0 h7t60mWMLTWCv02ckr6pkx9jy48 Fcg1y0PvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcutBAc0+Mb9t uU2KIIHH31AWg= </pre>
Deterministik	No
Byte masukan	26
Byte keluaran	175

Contoh 5

Input	abcdefghijklmnopqrstuvwxyzA BCDEFGHIJKLMNOPQRSTUVWXYZ01 23456789
preserveNulls	-
Output	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfstEE1GKEPiRzyh0 h7t60mWMLTWcV02ckr6plwtH/8t RFnn2rF91bcB9G4+n8GiRfJNmqdP4/ Q0Q3cXb/pbvPcnnohrHIGSX54ua+1/ JfcVjc=
Deterministik	No
Byte masukan	62
Byte keluaran	175

Jenis pad **fixed** (Contoh 2)

Dalam contoh ini, pad_length adalah 162 dan masukan terbesar adalah 62 byte.

Contoh 1

Input	null
preserveNulls	TRUE
Output	null
Deterministik	Yes
Byte masukan	0
Byte keluaran	0

Contoh 2

Input	null
preserveNulls	FALSE
Output	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc40TBqfRY Z98t5KU6aWfssGSNWfMRp7nSb7S MX2s3JKL0hK1+7r75Tk+Mx9jy48 Fcgl1y0PvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B8000Gp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwv/xAySX+xcntotL703aBTBb
Deterministik	No
Byte masukan	0
Byte keluaran	307

Contoh 3

Input	empty string
preserveNulls	-
Output	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc40TBqfRY Z98t5KU6aWfstGSNWfMRp7nSb7S MX2s3JKL0hK1+7r75Tk+Mx9jy48 Fcgl1y0PvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B8000Gp

	pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwv84lVaT9Yd+6oQx65/+gdVT
Deterministik	No
Byte masukan	0
Byte keluaran	307

Contoh 4

Input	abcdefghijklmnopqrstuvwxyz
preserveNulls	-
Output	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfstEE1GKEPiRzyh0 h7t60mWMLTWCv02ckr6pkx9jy48 Fcgl1y0PvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B8000Gp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwX5Hn1+Wyf06ks3QMaRDGSf
Deterministik	No
Byte masukan	26
Byte keluaran	307

Contoh 5

Input	abcdefghijklmnopqrstuvwxyzA BCDEFGHIJKLMNOPQRSTUVWXYZ01 23456789
preserveNulls	-
Output	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfstEE1GKEPiRzyh0 h7t60mWMLTWcV02ckr6plwtH/8t RFnn2rF91bcB9G4+n8GiRfJNmqd P4/Q0Q3cXb/pbvPcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvTkSSVc2B8000Gp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrnwjkJXQZ0gPdeFX9Yr/8a1V5i
Deterministik	No
Byte masukan	62
Byte keluaran	307

Jenis pad **max** (Contoh 1)

Dalam contoh ini, pad_length adalah 0 dan masukan terbesar adalah 62 byte.

Contoh 1

Input	null
preserveNulls	TRUE
Output	null
Deterministik	Yes

Masukan Byte	0
Byte Keluaran	0

Contoh 2

Input	null
preserveNulls	FALSE
Output	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc40TBqfRYZ98t5KU6aWfssGSNWfMRp7nSb7SMX2s3JKL0hK1+7r75Tk+Mx9jy48Fcg1y0PvBqRSZ7oqy1V3UKfYTL EZb/hCz7oaIneVsrcoNpATs0GzbnLkor4L+/aSuA=
Deterministik	No
Byte masukan	0
Byte keluaran	175

Contoh 3

Input	empty string
preserveNulls	-
Output	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc40TBqfRYZ98t5KU6aWfstGSNWfMRp7nSb7SMX2s3JKL0hK1+7r75Tk+Mx9jy48Fcg1y0PvBqRSZ7oqy1V3UKfYTL EZb/hCz7oaIneVsrcoLB53l07VZpA60wkuXu29CA=

Deterministik	No
Byte masukan	0
Byte keluaran	175

Contoh 4

Input	abcdefghijklmnopqrstuvwxy
preserveNulls	-
Output	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc40TBqfRYZ98t5KU6aWfstEE1GKEPiRzyh0h7t60mWMLTWCv02ckr6pkx9jy48Fcg1y0PvBqRSZ7oqy1V3UKfYTLEZb/hCz7oaIneVsrcutBAC0+Mb9tuU2KIIHH31AWg=
Deterministik	No
Byte masukan	26
Byte keluaran	175

Contoh 5

Input	abcdefghijklmnopqrstuvwxyA BCDEFGHIJKLMNOPQRSTUVWXYZ01 23456789
preserveNulls	-
Output	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc40TBqfRYZ98t5KU6aWfstEE1GKEPiRzyh0h7t60mWMLTWCv02ckr6plwtH/8t

	RFnn2rF91bcB9G4+n8GiRfJNmqdP4/ Q0Q3cXb/pbvPcnnohrHIGSX54ua+1/ JfcVjc=
Deterministik	No
Byte masukan	62
Byte keluaran	175

Jenis pad **max** (Contoh 2)

Dalam contoh ini, `pad_length` adalah 100 dan masukan terbesar adalah 62 byte.

Contoh 1

Input	null
<code>preserveNulls</code>	TRUE
Output	null
Deterministik	Yes
Byte masukan	0
Byte keluaran	0

Contoh 2

Input	null
<code>preserveNulls</code>	FALSE
Output	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmlMDEyMzQ1Njc40TBqfRY Z98t5KU6aWfssGSNWfMRp7nSb7S MX2s3JKL0hK1+7r75Tk+Mx9jy48 Fcg1y0PvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z

	NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B8000Gp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwv/xAySX+xcntotL703aBTBb
Deterministik	No
Byte masukan	0
Byte keluaran	307

Contoh 3

Input	empty string
preserveNulls	-
Output	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmn1MDEyMzQ1Njc40TBqfRY Z98t5KU6aWfstGSNWfMRp7nSb7S MX2s3JKL0hK1+7r75Tk+Mx9jy48 Fcgl1y0PvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B8000Gp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwv841VaT9Yd+6oQx65/+gdVT
Deterministik	No
Byte masukan	0
Byte keluaran	307

Contoh 4

Input	abcdefghijklmnopqrstuvwxyz
preserveNulls	-
Output	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfstEE1GKEPiRzyh0 h7t60mWMLTWcV02ckr6pkx9jy48 Fcgl1y0PvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B8000Gp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmtX5Hn1+Wyf06ks3QMaRDGSf
Deterministik	No
Byte masukan	26
Byte keluaran	307

Contoh 5

Input	abcdefghijklmnopqrstuvwxyzA BCDEFGHIJKLMNOPQRSTUVWXYZ01 23456789
preserveNulls	-
Output	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfstEE1GKEPiRzyh0 h7t60mWMLTWcV02ckr6plwtH/8t RFnn2rF91bcB9G4+n8GiRfJNmqd P4/Q0Q3cXb/pbvPcnkB0xbLWD7z

```
NdAqQGR0rXoSESdW0I0vpNoGcBf
v4cJbG0A3h1DvtkSSVc2B8000Gp
pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn
+8o4WtG/ClipNcjDXvXVtK4vfCohcCA6
uwrmwjkJXQZ0gPdeFX9Yr/8a1V5i
```

Deterministik	No
Byte masukan	62
Byte keluaran	307

Kolom pemecahan masalah sealed

Mengapa ciphertext di sealed kolom saya beberapa kali lebih besar dari ukuran cleartext yang masuk ke dalamnya?

Ini tergantung pada beberapa faktor. Untuk satu, ciphertext dalam Cleartext kolom selalu setidaknya 91 byte panjangnya. Jika data input Anda kecil (misalnya, usia pelanggan), itu akan menunjukkan peningkatan ukuran yang signifikan. Kedua, jika `preserveNulls` disetel ke `false` dan data input Anda berisi banyak `null` nilai, masing-masing `null` nilai tersebut akan berubah menjadi 91 byte ciphertext. Akhirnya, jika Anda menggunakan padding, menurut definisi byte ditambahkan ke cleartext data sebelum dienkripsi.

Sebagian besar data saya di sealed kolom sangat kecil, dan saya perlu menggunakan padding. Bisakah saya menghapus nilai besar dan memprosesnya secara terpisah untuk menghemat ruang?

Kami tidak menyarankan Anda menghapus nilai besar dan memprosesnya secara terpisah. Melakukan hal itu mengubah jaminan privasi yang disediakan oleh klien enkripsi C3R. Sebagai model ancaman, asumsikan bahwa pengamat dapat melihat kedua kumpulan data terenkripsi. Jika pengamat melihat bahwa satu subset data memiliki kolom yang dilapisi secara signifikan lebih atau kurang dari subset lain, mereka dapat membuat kesimpulan tentang ukuran data di setiap subset. Misalnya, asumsikan `fullName` kolom diempuk dengan total 40 byte dalam satu file dan diempuk hingga 800 byte di file lain. Seorang pengamat mungkin dapat berasumsi bahwa satu kumpulan data berisi nama terpanjang di dunia (747 byte).

Apakah saya perlu memberikan padding tambahan saat menggunakan tipe **max** padding?

Tidak. Saat menggunakan `max padding`, kami merekomendasikan bahwa `pad_length`, juga dikenal sebagai padding tambahan di luar nilai terbesar di kolom, diatur ke 0.

Bisakah saya memilih yang besar **`pad_length`** saat menggunakan **`fixed`** padding untuk menghindari kekhawatiran jika nilai terbesar akan cocok?

Ya, tetapi panjang `pad` yang besar tidak efisien dan menggunakan lebih banyak penyimpanan dari yang diperlukan. Kami menyarankan Anda untuk memeriksa untuk melihat seberapa besar nilai terbesar dan menetapkan `pad_length` ke nilai itu.

Bagaimana saya tahu jika saya membutuhkan jaminan kriptografi yang disediakan oleh?

`preserveNulls`

Sayangnya, jawabannya adalah itu tergantung. Minimal, [Komputasi Kriptografi untuk Clean Rooms](#) harus ditinjau untuk bagaimana `preserveNulls` pengaturan melindungi data Anda. Namun, kami menyarankan Anda untuk mereferensikan persyaratan penanganan data organisasi Anda dan kontrak apa pun yang berlaku untuk kolaborasi masing-masing.

Mengapa saya harus mengeluarkan biaya overhead `base64`?

Untuk memungkinkan kompatibilitas dengan format file tabular seperti CSV, pengkodean `base64` diperlukan. Meskipun beberapa format file seperti Parquet mungkin mendukung representasi biner data, penting bahwa semua peserta dalam kolaborasi mewakili data dengan cara yang sama untuk memastikan hasil kueri yang tepat.

Memecahkan masalah peningkatan ukuran ciphertext yang tidak terduga

Katakanlah Anda mengenkripsi data Anda, dan ukuran data yang dihasilkan sangat besar. Langkah-langkah berikut dapat membantu Anda mengidentifikasi di mana peningkatan ukuran terjadi dan apa, jika ada, tindakan yang dapat Anda ambil.

Mengidentifikasi di mana peningkatan ukuran terjadi

Sebelum Anda dapat memecahkan masalah mengapa data terenkripsi Anda secara signifikan lebih besar daripada `cleartext` data Anda, Anda harus terlebih dahulu mengidentifikasi di mana peningkatan ukurannya. `Cleartext` kolom dapat diabaikan dengan aman karena tidak berubah. Lihatlah yang tersisa `fingerprint` dan `sealed` kolom, dan pilih salah satu yang tampak signifikan.

Mengidentifikasi alasan peningkatan ukuran terjadi

`fingerprint` Kolom atau `sealed` kolom mungkin berkontribusi pada peningkatan ukuran.

Topik

- [Apakah peningkatan ukuran berasal dari fingerprint kolom?](#)
- [Apakah peningkatan ukuran berasal dari sealed kolom?](#)

Apakah peningkatan ukuran berasal dari fingerprint kolom?

Jika kolom yang paling berkontribusi terhadap peningkatan penyimpanan adalah fingerprint kolom, ini mungkin karena cleartext datanya kecil (misalnya, usia pelanggan). Setiap fingerprint ciphertext yang dihasilkan panjangnya 52 byte. Sayangnya, tidak ada yang bisa dilakukan tentang masalah ini column-by-column atas dasar. Untuk informasi selengkapnya, lihat [Basis overhead untuk kolom fingerprint](#) detail tentang kolom ini, termasuk dampaknya terhadap persyaratan penyimpanan.

Kemungkinan penyebab lain dari peningkatan ukuran dalam fingerprint kolom adalah pengaturan kolaborasi, `preserveNulls`. Jika pengaturan kolaborasi untuk `preserveNulls` dinonaktifkan (pengaturan default), semua `null` nilai dalam fingerprint kolom akan menjadi 52 byte ciphertext. Tidak ada yang bisa dilakukan untuk ini dalam kolaborasi saat ini. `preserveNulls` pengaturan diatur pada saat kolaborasi dibuat dan semua kolaborator harus menggunakan pengaturan yang sama untuk memastikan hasil kueri yang benar. Untuk informasi selengkapnya tentang `preserveNulls` pengaturan dan bagaimana pengaktifannya memengaruhi jaminan privasi data Anda, lihat [the section called “Komputasi kriptografi”](#)

Apakah peningkatan ukuran berasal dari sealed kolom?

Jika kolom yang paling berkontribusi terhadap peningkatan penyimpanan adalah sealed kolom, ada beberapa detail yang dapat berkontribusi pada peningkatan ukuran.

Jika cleartext data kecil (misalnya, usia pelanggan), setiap sealed ciphertext yang dihasilkan setidaknya 91 byte panjangnya. Sayangnya, tidak ada yang bisa dilakukan tentang masalah ini. Untuk informasi selengkapnya, lihat [Basis overhead untuk kolom sealed](#) detail tentang kolom ini, termasuk dampaknya terhadap persyaratan penyimpanan.

Penyebab utama kedua untuk peningkatan penyimpanan sealed kolom adalah padding. Padding menambahkan byte ekstra ke cleartext sebelum dienkripsi untuk menyembunyikan ukuran nilai individual dalam kumpulan data. Kami menyarankan Anda mengatur padding ke nilai minimum yang mungkin untuk kumpulan data Anda. Minimal, `pad_length` untuk `fixed` padding harus diatur untuk mencakup nilai terbesar yang mungkin di kolom. Pengaturan yang lebih tinggi dari itu tidak menambahkan jaminan privasi tambahan. Misalnya, jika Anda tahu nilai terbesar yang mungkin dalam kolom bisa 50 byte, kami sarankan Anda menyetel `pad_length` ke 50 byte. Namun, jika

sealed kolom menggunakan max padding, kami sarankan Anda mengatur `pad_length` ke 0 byte. Ini karena max padding mengacu pada padding tambahan di luar nilai terbesar di kolom.

Kemungkinan penyebab akhir dari peningkatan ukuran dalam sealed kolom adalah pengaturan kolaborasi, `preserveNulls`. Jika pengaturan kolaborasi untuk `preserveNulls` dinonaktifkan (pengaturan default), semua `null` nilai dalam sealed kolom akan menjadi 91 byte ciphertext. Tidak ada yang bisa dilakukan untuk ini dalam kolaborasi saat ini. `preserveNulls` Pengaturan diatur pada saat kolaborasi dibuat, dan semua kolaborator harus menggunakan pengaturan yang sama untuk memastikan hasil kueri yang benar. Untuk informasi selengkapnya tentang pengaturan ini dan bagaimana cara mengaktifkannya memengaruhi jaminan privasi data Anda, lihat. [the section called “Komputasi kriptografi”](#)

Menyiapkan AWS Clean Rooms

Topik berikut menjelaskan cara mengaturnya AWS Clean Rooms.

Topik

- [Mendaftar untuk AWS](#)
- [Siapkan peran layanan untuk AWS Clean Rooms](#)
- [Mengatur peran layanan untuk AWS Clean Rooms ML](#)

Mendaftar untuk AWS

Sebelum Anda dapat menggunakan AWS Clean Rooms, atau apa pun Layanan AWS, Anda harus mendaftar AWS dengan Akun AWS.

Jika Anda tidak memiliki Akun AWS, selesaikan langkah-langkah berikut untuk membuatnya.

Untuk mendaftar untuk Akun AWS

1. Buka <https://portal.aws.amazon.com/billing/pendaftaran>.
2. Ikuti petunjuk online.

Selama prosedur pendaftaran, Anda akan menerima panggilan telepon dengan kode verifikasi yang akan Anda masukkan pada keypad telepon.

3. Ketika Anda mendaftar untuk Akun AWS, pengguna Akun AWS root dibuat. Pengguna root memiliki akses ke semua Layanan AWS dan sumber daya di akun. Sebagai praktik terbaik keamanan, [tetapkan akses administratif ke pengguna administratif](#), dan hanya gunakan pengguna root untuk melakukan [tugas-tugas yang memerlukan akses pengguna root](#).

Siapkan peran layanan untuk AWS Clean Rooms

Bagian berikut menjelaskan peran yang diperlukan untuk melakukan setiap tugas.

Topik

- [Buat pengguna administrator](#)
- [Buat peran IAM untuk anggota kolaborasi](#)
- [Membuat peran layanan untuk membaca data dari Amazon S3](#)

- [Membuat peran layanan untuk membaca data dari Amazon Athena](#)
- [Buat peran layanan untuk membaca data dari Snowflake](#)
- [Buat peran layanan untuk membaca kode dari bucket S3 \(peran template PySpark analisis\)](#)
- [Buat peran layanan untuk menulis hasil PySpark pekerjaan](#)
- [Buat peran layanan untuk menerima hasil](#)

Buat pengguna administrator

Untuk menggunakannya AWS Clean Rooms, Anda perlu membuat pengguna administrator untuk diri sendiri dan menambahkan pengguna administrator ke grup administrator.

Untuk membuat pengguna administrator, pilih salah satu opsi berikut.

Pilih salah satu cara untuk mengelola administrator Anda	Untuk	Oleh	Anda juga bisa
Di Pusat Identitas IAM (Direkomendasikan)	Gunakan kredensi jangka pendek untuk mengakses. AWS Ini sejalan dengan praktik terbaik keamanan. Untuk informasi tentang praktik terbaik, lihat Praktik terbaik keamanan di IAM di Panduan Pengguna IAM.	Mengikuti petunjuk di Memulai di Panduan AWS IAM Identity Center Pengguna.	Konfigurasi akses terprogram dengan Mengonfigurasi AWS CLI yang akan digunakan AWS IAM Identity Center dalam AWS Command Line Interface Panduan Pengguna.

Pilih salah satu cara untuk mengelola administrator Anda	Untuk	Oleh	Anda juga bisa
Di IAM (Tidak direkomendasikan)	Gunakan kredensi jangka panjang untuk mengakses. AWS	Mengikuti petunjuk di Buat pengguna IAM untuk akses darurat di Panduan Pengguna IAM.	Konfigurasi akses terprogram dengan Mengelola kunci akses untuk pengguna IAM di Panduan Pengguna IAM .

Buat peran IAM untuk anggota kolaborasi

Anggota adalah AWS pelanggan yang merupakan peserta dalam kolaborasi.

Untuk membuat peran IAM untuk anggota kolaborasi

- Ikuti [Membuat peran untuk mendelegasikan izin ke prosedur pengguna IAM di Panduan Pengguna](#).AWS Identity and Access Management
- Untuk langkah Buat kebijakan, pilih tab JSON di editor Kebijakan, lalu tambahkan kebijakan tergantung pada kemampuan yang diberikan kepada anggota kolaborasi.

AWS Clean Rooms menawarkan kebijakan terkelola berikut berdasarkan kasus penggunaan umum.

Jika Anda ingin...	Kemudian gunakan...
Lihat sumber daya dan metadata	AWS kebijakan terkelola: AWSCleanRoomsReadOnlyAccess

Jika Anda ingin...	Kemudian gunakan...
Kueri	AWS kebijakan terkelola: AWSCleanRoomsFullAccess
Kueri dan jalankan pekerjaan	AWS kebijakan terkelola: AWSCleanRoomsFullAccess
Kueri dan terima hasil	AWS kebijakan terkelola: AWSCleanRoomsFullAccess
Kelola sumber daya kolaborasi tetapi jangan kueri	AWS kebijakan terkelola: AWSCleanRoomsFullAccessNoQuerying

Untuk informasi tentang berbagai kebijakan terkelola yang ditawarkan oleh AWS Clean Rooms, lihat [AWS kebijakan terkelola untuk AWS Clean Rooms](#),

Membuat peran layanan untuk membaca data dari Amazon S3

AWS Clean Rooms menggunakan peran layanan untuk membaca data dari Amazon S3.

Ada dua cara untuk membuat peran layanan ini.

- Jika Anda memiliki izin IAM yang diperlukan untuk membuat peran layanan, gunakan AWS Clean Rooms konsol untuk membuat peran layanan.
- Jika Anda tidak memiliki `iam:CreateRole`, `iam:CreatePolicy` dan `iam:AttachRolePolicy` izin atau ingin membuat peran IAM secara manual, lakukan salah satu hal berikut:
 - Gunakan prosedur berikut untuk membuat peran layanan menggunakan kebijakan kepercayaan khusus.
 - Minta administrator Anda untuk membuat peran layanan menggunakan prosedur berikut.

Note

Anda atau administrator IAM Anda harus mengikuti prosedur ini hanya jika Anda tidak memiliki izin yang diperlukan untuk membuat peran layanan menggunakan konsol. AWS Clean Rooms

Untuk membuat peran layanan untuk membaca data dari Amazon S3 menggunakan kebijakan kepercayaan khusus

1. Buat peran menggunakan kebijakan kepercayaan khusus. Untuk informasi selengkapnya, lihat prosedur [Membuat peran menggunakan kebijakan kepercayaan kustom \(konsol\)](#) di Panduan AWS Identity and Access Management Pengguna.
2. Gunakan kebijakan kepercayaan kustom berikut sesuai dengan prosedur [Membuat peran menggunakan kebijakan kepercayaan kustom \(konsol\)](#).

 Note

Jika Anda ingin membantu memastikan bahwa peran tersebut hanya digunakan dalam konteks keanggotaan kolaborasi tertentu, Anda dapat meringkas kebijakan kepercayaan lebih lanjut. Untuk informasi selengkapnya, lihat [Pencegahan "confused deputy" lintas layanan](#).

JSON

```
{  
  
  "Version": "2012-10-17",  
  "Statement": [  
    {  
      "Sid": "RoleTrustPolicyForCleanRoomsService",  
      "Effect": "Allow",  
      "Principal": {  
        "Service": "cleanrooms.amazonaws.com"  
      },  
      "Action": "sts:AssumeRole"  
    }  
  ]  
}
```

3. Gunakan kebijakan izin berikut sesuai dengan prosedur [Membuat peran menggunakan kebijakan kepercayaan kustom \(konsol\)](#).

Note

Kebijakan contoh berikut mendukung izin yang diperlukan untuk membaca AWS Glue metadata dan data Amazon S3 yang sesuai. Namun, Anda mungkin perlu mengubah kebijakan ini tergantung pada cara Anda mengatur data Amazon S3 Anda. Misalnya, jika Anda telah menyiapkan kunci KMS khusus untuk data Amazon S3, Anda mungkin perlu mengubah kebijakan ini dengan izin AWS Key Management Service tambahan AWS KMS().

AWS Glue Sumber daya Anda dan sumber daya Amazon S3 yang mendasarinya harus AWS Region sama dengan kolaborasi. AWS Clean Rooms

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "NecessaryGluePermissions",
      "Effect": "Allow",
      "Action": [
        "glue:GetDatabase",
        "glue:GetDatabases",
        "glue:GetTable",
        "glue:GetTables",
        "glue:GetPartition",
        "glue:GetPartitions",
        "glue:BatchGetPartition"
      ],
      "Resource": [
        "arn:aws:glue:us-east-1:111122223333:database/databaseName",
        "arn:aws:glue:us-east-1:111122223333:table/databaseName/tableName",
        "arn:aws:glue:us-east-1:111122223333:catalog"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
```

```

        "glue:GetSchema",
        "glue:GetSchemaVersion"
    ],
    "Resource": [
        "*"
    ]
},
{
    "Sid": "NecessaryS3BucketPermissions",
    "Effect": "Allow",
    "Action": [
        "s3:GetBucketLocation",
        "s3:ListBucket"
    ],
    "Resource": [
        "arn:aws:s3:::bucket"
    ],
    "Condition": {
        "StringEquals": {
            "s3:ResourceAccount": [
                "444455556666"
            ]
        }
    }
},
{
    "Sid": "NecessaryS3ObjectPermissions",
    "Effect": "Allow",
    "Action": [
        "s3:GetObject"
    ],
    "Resource": [
        "arn:aws:s3:::bucket/prefix/*"
    ],
    "Condition": {
        "StringEquals": {
            "s3:ResourceAccount": [
                "444455556666"
            ]
        }
    }
}
]

```


}

Note

Kebijakan ini mereferensikan dua hal yang berbeda Akun AWS IDs untuk mendukung AWS Clean Rooms kolaborasi dimana metadata katalog data dan penyimpanan data aktual dikelola oleh berbagai pihak:

- 111122223333 - Ini adalah akun yang memiliki sumber daya Katalog AWS Glue Data (database, tabel, dan katalog). Pernyataan pertama memberikan izin untuk mengakses skema tabel, informasi partisi, dan metadata dari katalog akun ini. AWS Glue
- 444455556666 - Ini adalah akun yang memiliki bucket Amazon S3 yang berisi file data aktual. Izin Amazon S3 (pernyataan 3 dan 4) dibatasi untuk bucket yang dimiliki oleh akun ini melalui kondisi tersebut. `s3:ResourceAccount`

Konfigurasi ini mendukung arsitektur data perusahaan umum di mana satu tim mengelola katalog data dan definisi skema sementara tim lain memiliki infrastruktur penyimpanan data yang mendasarinya. `s3:ResourceAccount` Kondisi ini memberikan lapisan keamanan tambahan dengan memastikan operasi Amazon S3 hanya berfungsi pada bucket yang dimiliki oleh akun yang ditunjuk.

4. Ganti masing-masing *placeholder* dengan informasi Anda sendiri.
5. Terus ikuti prosedur [Membuat peran menggunakan kebijakan kepercayaan khusus \(konsol\)](#) untuk membuat peran.

Membuat peran layanan untuk membaca data dari Amazon Athena

AWS Clean Rooms menggunakan peran layanan untuk membaca data dari Amazon Athena.

Untuk membuat peran layanan untuk membaca data dari Athena menggunakan kebijakan kepercayaan khusus

1. Buat peran menggunakan kebijakan kepercayaan khusus. Untuk informasi selengkapnya, lihat prosedur [Membuat peran menggunakan kebijakan kepercayaan kustom \(konsol\)](#) di Panduan AWS Identity and Access Management Pengguna.

- Gunakan kebijakan kepercayaan kustom berikut sesuai dengan prosedur [Membuat peran menggunakan kebijakan kepercayaan kustom \(konsol\)](#).

 Note

Jika Anda ingin membantu memastikan bahwa peran tersebut hanya digunakan dalam konteks keanggotaan kolaborasi tertentu, Anda dapat meringkas kebijakan kepercayaan lebih lanjut. Untuk informasi selengkapnya, lihat [Pencegahan "confused deputy" lintas layanan](#).

JSON

```
{  
  
  "Version": "2012-10-17",  
  "Statement": [  
    {  
      "Sid": "RoleTrustPolicyForCleanRoomsService",  
      "Effect": "Allow",  
      "Principal": {  
        "Service": "cleanrooms.amazonaws.com"  
      },  
      "Action": "sts:AssumeRole"  
    }  
  ]  
}
```

- Gunakan kebijakan izin berikut sesuai dengan prosedur [Membuat peran menggunakan kebijakan kepercayaan kustom \(konsol\)](#).

 Note

Kebijakan contoh berikut mendukung izin yang diperlukan untuk membaca AWS Glue metadata dan data Athena yang terkait. Namun, Anda mungkin perlu mengubah kebijakan ini tergantung pada cara Anda mengatur data Amazon S3 Anda. Misalnya, jika Anda telah menyiapkan kunci KMS khusus untuk data Amazon S3, Anda mungkin perlu mengubah kebijakan ini dengan izin tambahan. AWS KMS

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "athena:GetWorkGroup",
        "athena:GetTableMetadata",
        "athena:GetDataCatalog",
        "athena:StartQueryExecution",
        "athena:GetQueryExecution",
        "athena:GetQueryResults"
      ],
      "Resource": [
        "arn:aws:athena:region:accountId:workgroup/workgroup",
        "arn:aws:athena:region:accountId:datacatalog/federatedCatalogName"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "glue:GetDatabase",
        "glue:GetTable",
        "glue:GetCatalog"
      ],
      "Resource": [
        "arn:aws:glue:region:accountId:catalog",
        "arn:aws:glue:region:accountId:catalog/federatedCatalogName",
        "arn:aws:glue:region:accountId:database/federatedCatalogName/databaseName",
        "arn:aws:glue:region:accountId:table/federatedCatalogName/databaseName/tableName"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "s3:GetObject",
        "s3:GetBucketLocation",
        "s3:AbortMultipartUpload",
        "s3:ListBucket",
        "s3:PutObject",

```

```

        "s3:ListMultipartUploadParts"
    ],
    "Resource": [
        "arn:aws:s3:::athenaResultsBucket",
        "arn:aws:s3:::athenaResultsBucket/*"
    ],
    "Condition": {
        "StringEquals": {
            "aws:ResourceAccount": "accountId"
        }
    }
},
{
    "Effect": "Allow",
    "Action": "lakeformation:GetDataAccess",
    "Resource": "*"
}
]
}

```

4. Ganti masing-masing *placeholder* dengan informasi Anda sendiri.
5. Terus ikuti prosedur [Membuat peran menggunakan kebijakan kepercayaan khusus \(konsol\)](#) untuk membuat peran.

Siapkan izin Lake Formation

Jika Anda meminta sumber daya yang dilindungi dengan izin Lake Formation, peran layanan harus memiliki izin akses Pilih dan Jelaskan pada izin table/view/catalog dan Jelaskan pada database. AWS Glue

Untuk informasi lebih lanjut, lihat:

- [Gunakan Athena untuk menanyakan data yang terdaftar AWS Lake Formation](#) di Panduan Pengguna Amazon Athena
- [Izin orientasi ke Lake Formation](#) di Panduan Pengembang AWS Lake Formation

Buat peran layanan untuk membaca data dari Snowflake

AWS Clean Rooms menggunakan peran layanan untuk mengambil kredensial Anda untuk Snowflake untuk membaca data Anda dari sumber ini.

Ada dua cara untuk membuat peran layanan ini:

- Jika Anda memiliki izin IAM yang diperlukan untuk membuat peran layanan, gunakan AWS Clean Rooms konsol untuk membuat peran layanan.
- Jika Anda tidak memiliki `iam:CreateRole`, `iam:CreatePolicy` dan `iam:AttachRolePolicy` izin atau ingin membuat peran IAM secara manual, lakukan salah satu hal berikut:
 - Gunakan prosedur berikut untuk membuat peran layanan menggunakan kebijakan kepercayaan khusus.
 - Minta administrator Anda untuk membuat peran layanan menggunakan prosedur berikut.

 Note

Anda atau administrator IAM Anda harus mengikuti prosedur ini hanya jika Anda tidak memiliki izin yang diperlukan untuk membuat peran layanan menggunakan konsol. AWS Clean Rooms

Untuk membuat peran layanan untuk membaca data dari Snowflake menggunakan kebijakan kepercayaan khusus

1. Buat peran menggunakan kebijakan kepercayaan khusus. Untuk informasi selengkapnya, lihat prosedur [Membuat peran menggunakan kebijakan kepercayaan kustom \(konsol\)](#) di Panduan AWS Identity and Access Management Pengguna.
2. Gunakan kebijakan kepercayaan kustom berikut sesuai dengan prosedur [Membuat peran menggunakan kebijakan kepercayaan kustom \(konsol\)](#).

 Note

Jika Anda ingin membantu memastikan bahwa peran tersebut hanya digunakan dalam konteks keanggotaan kolaborasi tertentu, Anda dapat meringkas kebijakan kepercayaan lebih lanjut. Untuk informasi selengkapnya, lihat [Pencegahan "confused deputy" lintas layanan](#).

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowIfSourceArnMatches",
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "ForAnyValue:ArnEquals": {
          "aws:SourceArn": [
            "arn:aws:cleanrooms:us-  
east-1:111122223333:membership/membershipId",
            "arn:aws:cleanrooms:us-  
east-1:444455556666:membership/queryRunnerMembershipId"
          ]
        }
      }
    }
  ]
}
```

 Note

Kebijakan kepercayaan ini mereferensikan dua hal yang berbeda Akun AWS IDs untuk mendukung AWS Clean Rooms kolaborasi di mana tanggung jawab eksekusi kueri didistribusikan ke berbagai pihak:

- 111122223333 - Ini adalah akun yang berisi keanggotaan yang berpartisipasi dalam kolaborasi. Keanggotaan ini dapat memiliki tabel data, aturan analisis, atau sumber kolaborasi lainnya yang memerlukan akses peran.
- 444455556666 - Ini adalah akun yang berisi keanggotaan yang bertanggung jawab untuk menjalankan query ("query runner"). Keanggotaan ini mengeksekusi kueri yang

dilindungi dan perlu mengambil peran ini untuk mengakses sumber daya komputasi dan data yang diperlukan.

Konfigurasi ini memungkinkan skenario di mana satu pihak menyediakan templat data atau analisis sementara pihak lain menjalankan kueri yang sebenarnya. Kedua peran memerlukan izin yang berbeda tetapi saling melengkapi melalui peran eksekusi yang sama. `aws:SourceArnKondisi` ini memastikan bahwa hanya AWS Clean Rooms operasi yang berasal dari dua keanggotaan spesifik ini yang dapat mengambil peran, menjaga keamanan sambil mendukung pelaksanaan pekerjaan terdistribusi dan alur kerja manajemen hasil.

- Gunakan salah satu kebijakan izin berikut sesuai dengan prosedur [Membuat peran menggunakan kebijakan kepercayaan kustom \(konsol\)](#).

Kebijakan izin untuk rahasia yang dienkripsi dengan kunci KMS milik pelanggan

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": "secretsmanager:GetSecretValue",
      "Resource": "arn:aws:secretsmanager:us-east-1:111122223333:secret:secretIdentifier",
      "Effect": "Allow"
    },
    {
      "Sid": "AllowDecryptViaSecretsManagerForKey",
      "Action": "kms:Decrypt",
      "Resource": "arn:aws:kms:us-east-1:444455556666:key/keyIdentifier",
      "Effect": "Allow",
      "Condition": {
        "StringEquals": {
          "kms:ViaService": "secretsmanager.us-east-1.amazonaws.com",
          "kms:EncryptionContext:SecretARN":
            "arn:aws:secretsmanager:us-east-1:111122223333:secret:secretIdentifier"
        }
      }
    }
  ]
}
```

```

    }
  }
}

```

Note

Kebijakan ini mereferensikan dua hal yang berbeda Akun AWS IDs untuk mendukung skenario manajemen rahasia lintas akun:

- 111122223333 - Ini adalah akun yang memiliki dan menyimpan rahasia di. Pernyataan pertama memberikan izin untuk mengambil nilai rahasia dari akun ini.
- 444455556666 - Ini adalah akun yang memiliki kunci yang digunakan untuk mengenkripsi rahasia. AWS KMS Pernyataan kedua memberikan izin untuk mendekripsi rahasia menggunakan AWS KMS kunci dari akun ini.

Konfigurasi ini umum di lingkungan perusahaan di mana:

- Rahasia dikelola secara terpusat dalam satu akun (Akun 1)
- Kunci enkripsi dikelola oleh akun keamanan atau layanan bersama terpisah (Akun 2)
- Kebijakan AWS KMS utama di Akun 2 juga harus mengizinkan layanan di Akun 1 untuk menggunakan kunci untuk encryption/decryption operasi

`kms:EncryptionContext:SecretARN` Kondisi ini memastikan bahwa AWS KMS kunci hanya dapat digunakan untuk mendekripsi rahasia khusus ini, memberikan lapisan keamanan tambahan untuk akses lintas akun.

Kebijakan izin untuk rahasia yang dienkripsi dengan Kunci yang dikelola AWS

JSON

```

{
  "Version": "2012-10-17",
  "Statement": [
    {

```



```

        "Action": "secretsmanager:GetSecretValue",
        "Resource": "arn:aws:secretsmanager:us-
east-1:111122223333:secret:secretIdentifier",
        "Effect": "Allow"
    }
]
}

```

4. Ganti masing-masing *placeholder* dengan informasi Anda sendiri.
5. Terus ikuti prosedur [Membuat peran menggunakan kebijakan kepercayaan khusus \(konsol\)](#) untuk membuat peran.

Buat peran layanan untuk membaca kode dari bucket S3 (peran template PySpark analisis)

AWS Clean Rooms menggunakan peran layanan untuk membaca kode dari bucket S3 yang ditentukan anggota kolaborasi saat menggunakan templat PySpark analisis.

Untuk membuat peran layanan untuk membaca kode dari bucket S3

1. Buat peran menggunakan kebijakan kepercayaan khusus. Untuk informasi selengkapnya, lihat prosedur [Membuat peran menggunakan kebijakan kepercayaan kustom \(konsol\)](#) di Panduan AWS Identity and Access Management Pengguna.
2. Gunakan kebijakan kepercayaan kustom berikut sesuai dengan prosedur [Membuat peran menggunakan kebijakan kepercayaan kustom \(konsol\)](#).

JSON

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "ForAnyValue:ArnEquals": {

```

```

        "aws:SourceArn": [
            "arn:aws:cleanrooms:us-
east-1:111122223333:membership/jobRunnerMembershipId",
            "arn:aws:cleanrooms:us-
east-1:444455556666:membership/analysisTemplateOwnerMembershipId"
        ]
    }
}
]
}

```

Note

Kebijakan kepercayaan ini mereferensikan dua hal berbeda Akun AWS IDs untuk mendukung skenario AWS Clean Rooms kolaborasi multi-pihak:

- 111122223333 - Ini adalah akun yang berisi keanggotaan yang bertanggung jawab untuk menjalankan kueri (“pelari pekerjaan”). Keanggotaan ini menjalankan pekerjaan analisis dan perlu mengambil peran ini untuk mengakses sumber daya yang diperlukan.
- 444455556666 - Ini adalah akun yang memiliki template analisis dan keanggotaan terkait (“pemilik template analisis”). Keanggotaan ini mendefinisikan kueri apa yang dapat dijalankan dan juga perlu mengambil peran ini untuk mengelola dan menjalankan analisis.

Konfigurasi ini khas dalam AWS Clean Rooms kolaborasi di mana banyak pihak berpartisipasi dalam kolaborasi yang sama, masing-masing dengan milik mereka sendiri Akun AWS dan keanggotaan. Baik pelaksana kueri dan pemilik templat analisis memerlukan akses ke sumber daya bersama. `aws:SourceArnKondisi` ini memastikan bahwa hanya AWS Clean Rooms operasi yang berasal dari dua keanggotaan spesifik ini yang dapat mengambil peran, memberikan kontrol akses yang tepat untuk kolaborasi multi-pihak.

3. Gunakan kebijakan izin berikut sesuai dengan prosedur [Membuat peran menggunakan kebijakan kepercayaan kustom \(konsol\)](#).

Note

Contoh kebijakan berikut mendukung izin yang diperlukan untuk membaca kode Anda dari Amazon S3. Namun, Anda mungkin perlu mengubah kebijakan ini tergantung pada cara Anda menyiapkan data S3.

Sumber daya Amazon S3 Anda harus AWS Region sama dengan kolaborasi. AWS Clean Rooms

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:GetObject",
        "s3:GetObjectVersion"
      ],
      "Resource": ["arn:aws:s3:::s3Path"],
      "Condition": {
        "StringEquals": {
          "s3:ResourceAccount": [
            "s3BucketOwnerAccountId"
          ]
        }
      }
    }
  ]
}
```

4. Ganti masing-masing *placeholder* dengan informasi Anda sendiri:

- *s3Path*— Lokasi bucket S3 kode Anda.
- *s3BucketOwnerAccountId*— Akun AWS ID pemilik bucket S3.
- *region*- Nama AWS Region. Misalnya, **us-east-1**.

- *jobRunnerAccountId*— Akun AWS ID anggota yang dapat menjalankan kueri dan pekerjaan.
 - *jobRunnerMembershipId*— ID Keanggotaan anggota yang dapat menanyakan dan menjalankan pekerjaan. ID Keanggotaan dapat ditemukan di tab Detail kolaborasi. Ini memastikan bahwa AWS Clean Rooms mengasumsikan peran hanya ketika anggota ini menjalankan analisis dalam kolaborasi ini.
 - *analysisTemplateAccountId*— Akun AWS ID dari template analisis.
 - *analysisTemplateOwnerMembershipId*— ID Keanggotaan anggota yang memiliki templat analisis. ID Keanggotaan dapat ditemukan di tab Detail kolaborasi.
5. Terus ikuti prosedur [Membuat peran menggunakan kebijakan kepercayaan khusus \(konsol\)](#) untuk membuat peran.

Buat peran layanan untuk menulis hasil PySpark pekerjaan

AWS Clean Rooms menggunakan peran layanan untuk menulis hasil PySpark pekerjaan ke bucket S3 tertentu.

Untuk membuat peran layanan untuk menulis hasil PySpark pekerjaan

1. Buat peran menggunakan kebijakan kepercayaan khusus. Untuk informasi selengkapnya, lihat prosedur [Membuat peran menggunakan kebijakan kepercayaan kustom \(konsol\)](#) di Panduan AWS Identity and Access Management Pengguna.
2. Gunakan kebijakan kepercayaan kustom berikut sesuai dengan prosedur [Membuat peran menggunakan kebijakan kepercayaan kustom \(konsol\)](#).

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
```

```

    "ForAnyValue:ArnEquals": {
      "aws:SourceArn": [
        "arn:aws:cleanrooms:us-
east-1:111122223333:membership/jobRunnerMembershipId",
        "arn:aws:cleanrooms:us-
east-1:444455556666:membership/rrMembershipId"
      ]
    }
  }
}
]
}

```

Note

Kebijakan kepercayaan ini merujuk dua hal yang berbeda Akun AWS IDs untuk mendukung AWS Clean Rooms kolaborasi dengan peran operasional yang berbeda:

- 111122223333 - Ini adalah akun yang berisi keanggotaan yang bertanggung jawab untuk menjalankan pekerjaan analisis (“pelari pekerjaan”). Keanggotaan ini mengeksekusi beban kerja komputasi dan perlu mengambil peran ini untuk mengakses sumber daya pemrosesan.
- 444455556666 - Ini adalah akun yang berisi keanggotaan dengan tanggung jawab penerima hasil (RR). Keanggotaan ini berwenang untuk menerima dan mengakses output pekerjaan analisis, dan membutuhkan akses peran untuk menulis hasil ke lokasi yang ditentukan.

Konfigurasi ini memungkinkan AWS Clean Rooms skenario di mana satu pihak menjalankan analisis komputasi sementara pihak lain menerima dan mengelola hasilnya. Kedua peran memerlukan izin yang berbeda tetapi saling melengkapi melalui peran eksekusi yang sama. `aws:SourceArn` kondisi ini memastikan bahwa hanya AWS Clean Rooms operasi yang berasal dari dua keanggotaan spesifik ini yang dapat mengambil peran, menjaga keamanan sambil mendukung pelaksanaan pekerjaan terdistribusi dan alur kerja manajemen hasil.

3. Gunakan kebijakan izin berikut sesuai dengan prosedur [Membuat peran menggunakan kebijakan kepercayaan kustom \(konsol\)](#).

 Note

Kebijakan contoh berikut mendukung izin yang diperlukan untuk menulis ke Amazon S3. Namun, Anda mungkin perlu mengubah kebijakan ini tergantung pada cara Anda mengatur S3.

Sumber daya Amazon S3 Anda harus AWS Region sama dengan kolaborasi. AWS Clean Rooms

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:PutObject"
      ],
      "Resource": "arn:aws:s3:::bucket/optionalPrefix/*",
      "Condition": {
        "StringEquals": {
          "s3:ResourceAccount": [
            "s3BucketOwnerAccountId"
          ]
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "s3:GetBucketLocation",
        "s3:ListBucket"
      ],
      "Resource": "arn:aws:s3:::bucket",
      "Condition": {
        "StringEquals": {
          "s3:ResourceAccount": [
            "s3BucketOwnerAccountId"
          ]
        }
      }
    }
  ]
}
```

```

    }
  }
}

```

4. Ganti masing-masing *placeholder* dengan informasi Anda sendiri:
 - *region*- Nama AWS Region. Misalnya, **us-east-1**.
 - *jobRunnerAccountId*— Akun AWS ID tempat bucket S3 berada.
 - *jobRunnerMembershipId*— ID Keanggotaan anggota yang dapat menanyakan dan menjalankan pekerjaan. ID Keanggotaan dapat ditemukan di tab Detail kolaborasi. Ini memastikan bahwa AWS Clean Rooms mengasumsikan peran hanya ketika anggota ini menjalankan analisis dalam kolaborasi ini.
 - *rrAccountId*— Akun AWS ID tempat bucket S3 berada.
 - *rrMembershipId*— ID Keanggotaan anggota yang dapat menerima hasil. ID Keanggotaan dapat ditemukan di tab Detail kolaborasi. Ini memastikan bahwa AWS Clean Rooms mengasumsikan peran hanya ketika anggota ini menjalankan analisis dalam kolaborasi ini.
 - *bucket*— Nama dan lokasi bucket S3.
 - *optionalPrefix*- Awalan opsional jika Anda ingin menyimpan hasil Anda di bawah awalan S3 tertentu.
 - *s3BucketOwnerAccountId*— Akun AWS ID pemilik bucket S3.
5. Terus ikuti prosedur [Membuat peran menggunakan kebijakan kepercayaan khusus \(konsol\)](#) untuk membuat peran.

Buat peran layanan untuk menerima hasil

Note

Jika Anda adalah anggota yang hanya dapat menerima hasil (di konsol, kemampuan anggota Anda hanya Terima hasil), ikuti prosedur ini.

Jika Anda adalah anggota yang dapat menanyakan dan menerima hasil (di konsol, Kemampuan anggota Anda adalah hasil Kueri dan Terima), Anda dapat melewati prosedur ini.

Untuk anggota kolaborasi yang hanya dapat menerima hasil, AWS Clean Rooms gunakan peran layanan untuk menulis hasil data kueri dalam kolaborasi ke bucket S3 yang ditentukan.

Ada dua cara untuk membuat peran layanan ini:

- Jika Anda memiliki izin IAM yang diperlukan untuk membuat peran layanan, gunakan AWS Clean Rooms konsol untuk membuat peran layanan.
- Jika Anda tidak memiliki `iam:CreateRole`, `iam:CreatePolicy` dan `iam:AttachRolePolicy` izin atau ingin membuat peran IAM secara manual, lakukan salah satu hal berikut:
 - Gunakan prosedur berikut untuk membuat peran layanan menggunakan kebijakan kepercayaan khusus.
 - Minta administrator Anda untuk membuat peran layanan menggunakan prosedur berikut.

Note

Anda atau administrator IAM Anda harus mengikuti prosedur ini hanya jika Anda tidak memiliki izin yang diperlukan untuk membuat peran layanan menggunakan konsol. AWS Clean Rooms

Untuk membuat peran layanan untuk menerima hasil menggunakan kebijakan kepercayaan khusus

1. Buat peran menggunakan kebijakan kepercayaan khusus. Untuk informasi selengkapnya, lihat prosedur [Membuat peran menggunakan kebijakan kepercayaan kustom \(konsol\)](#) di Panduan AWS Identity and Access Management Pengguna.
2. Gunakan kebijakan kepercayaan kustom berikut sesuai dengan prosedur [Membuat peran menggunakan kebijakan kepercayaan kustom \(konsol\)](#).

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowIfExternalIdMatches",
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms.amazonaws.com"
```



```

    },
    "Action": "sts:AssumeRole",
    "Condition": {
        "ArnLike": {
            "sts:ExternalId":
"arn:aws*:region*:dbuser:*/a1b2c3d4-5678-90ab-cdef-EXAMPLEaaaa*"
        }
    },
    },
    {
        "Sid": "AllowIfSourceArnMatches",
        "Effect": "Allow",
        "Principal": {
            "Service": "cleanrooms.amazonaws.com"
        },
        "Action": "sts:AssumeRole",
        "Condition": {
            "ForAnyValue:ArnEquals": {
                "aws:SourceArn": [
                    "arn:aws:cleanrooms:us-east-1:555555555555:membership/a1b2c3d4-5678-90ab-cdef-EXAMPLEaaaa"
                ]
            }
        }
    }
}
]
}

```

- Gunakan kebijakan izin berikut sesuai dengan prosedur [Membuat peran menggunakan kebijakan kepercayaan kustom \(konsol\)](#).

Note

Kebijakan contoh berikut mendukung izin yang diperlukan untuk membaca AWS Glue metadata dan data Amazon S3 yang sesuai. Namun, Anda mungkin perlu mengubah kebijakan ini tergantung pada cara Anda menyiapkan data S3.

AWS Glue Sumber daya Anda dan sumber daya Amazon S3 yang mendasarinya harus AWS Region sama dengan kolaborasi. AWS Clean Rooms

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:GetBucketLocation",
        "s3:ListBucket"
      ],
      "Resource": [
        "arn:aws:s3:::bucket_name"
      ],
      "Condition": {
        "StringEquals": {
          "aws:ResourceAccount": "accountId"
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "s3:PutObject"
      ],
      "Resource": [
        "arn:aws:s3:::bucket_name/optional_key_prefix/*"
      ],
      "Condition": {
        "StringEquals": {
          "aws:ResourceAccount": "accountId"
        }
      }
    }
  ]
}
```

4. Ganti masing-masing *placeholder* dengan informasi Anda sendiri:

- *region*- Nama AWS Region. Misalnya, **us-east-1**.

- *a1b2c3d4-5678-90ab-cdef-EXAMPLEaaaaa*— ID Keanggotaan anggota yang dapat melakukan query. ID Keanggotaan dapat ditemukan di tab Detail kolaborasi. Ini memastikan bahwa AWS Clean Rooms mengasumsikan peran hanya ketika anggota ini menjalankan analisis dalam kolaborasi ini.
- *arn:aws:cleanrooms:us-east-1:555555555555:membership/a1b2c3d4-5678-90ab-cdef-EXAMPLEaaaaa*— ARN Keanggotaan tunggal dari anggota yang dapat menanyakan. ARN Keanggotaan dapat ditemukan di tab Detail kolaborasi. AWS Clean Rooms Ini memastikan mengasumsikan peran hanya ketika anggota ini menjalankan analisis dalam kolaborasi ini.
- *bucket_name*— Nama Sumber Daya Amazon (ARN) dari ember S3. Nama Sumber Daya Amazon (ARN) dapat ditemukan di tab Properties bucket di Amazon S3.
- *accountId*— Akun AWS ID tempat bucket S3 berada.

bucket_name/optional_key_prefix— Nama Sumber Daya Amazon (ARN) dari tujuan hasil di Amazon S3. Nama Sumber Daya Amazon (ARN) dapat ditemukan di tab Properties bucket di Amazon S3.

5. Terus ikuti prosedur [Membuat peran menggunakan kebijakan kepercayaan khusus \(konsol\)](#) untuk membuat peran.

Mengatur peran layanan untuk AWS Clean Rooms ML

Peran yang diperlukan untuk melakukan pemodelan mirip berbeda dari yang dibutuhkan untuk menggunakan model khusus. Bagian berikut menjelaskan peran yang diperlukan untuk melakukan setiap tugas.

Topik

- [Siapkan peran layanan untuk pemodelan mirip](#)
- [Siapkan peran layanan untuk pemodelan khusus](#)

Siapkan peran layanan untuk pemodelan mirip

Topik

- [Membuat peran layanan untuk membaca data pelatihan](#)
- [Buat peran layanan untuk menulis segmen yang mirip](#)

- [Buat peran layanan untuk membaca data benih](#)

Membuat peran layanan untuk membaca data pelatihan

AWS Clean Rooms menggunakan peran layanan untuk membaca data pelatihan. Anda dapat membuat peran ini menggunakan konsol jika Anda memiliki izin IAM yang diperlukan. Jika Anda tidak memiliki `CreateRole` izin, minta administrator Anda untuk membuat peran layanan.

Untuk membuat peran layanan untuk melatih kumpulan data

1. Masuk ke konsol IAM (<https://console.aws.amazon.com/iam/>) dengan akun administrator Anda.
2. Di bagian Manajemen akses, pilih Kebijakan.
3. Pilih Buat kebijakan.
4. Di editor Kebijakan, pilih tab JSON, lalu salin dan tempel kebijakan berikut.

Note

Kebijakan contoh berikut mendukung izin yang diperlukan untuk membaca AWS Glue metadata dan data Amazon S3 yang sesuai. Namun, Anda mungkin perlu mengubah kebijakan ini tergantung pada cara Anda menyiapkan data S3. Kebijakan ini tidak menyertakan kunci KMS untuk mendekripsi data.

AWS Glue Sumber daya Anda dan sumber daya Amazon S3 yang mendasarinya harus AWS Region sama dengan kolaborasi. AWS Clean Rooms

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "glue:GetDatabase",
        "glue:GetDatabases",
        "glue:GetTable",
        "glue:GetTables",
        "glue:GetPartitions",
```

```

        "glue:GetPartition",
        "glue:BatchGetPartition",
        "glue:GetUserDefinedFunctions"
    ],
    "Resource": [
        "arn:aws:glue:us-east-1:111122223333:database/databases",
        "arn:aws:glue:us-east-1:111122223333:table/databases/tables",
        "arn:aws:glue:us-east-1:111122223333:catalog",
        "arn:aws:glue:us-east-1:111122223333:database/default"
    ]
},
{
    "Effect": "Allow",
    "Action": [
        "glue:CreateDatabase"
    ],
    "Resource": [
        "arn:aws:glue:us-east-1:111122223333:database/default"
    ]
},
{
    "Effect": "Allow",
    "Action": [
        "s3:ListBucket",
        "s3:GetBucketLocation"
    ],
    "Resource": [
        "arn:aws:s3::bucket"
    ],
    "Condition": {
        "StringEquals": {
            "s3:ResourceAccount": [
                "111122223333"
            ]
        }
    }
},
{
    "Effect": "Allow",
    "Action": [
        "s3:GetObject"
    ],
    "Resource": [
        "arn:aws:s3::bucketFolders/*"
    ]
}

```

```

    ],
    "Condition": {
      "StringEquals": {
        "s3:ResourceAccount": [
          "111122223333"
        ]
      }
    }
  }
]
}

```

Jika Anda perlu menggunakan kunci KMS untuk mendekripsi data, tambahkan AWS KMS pernyataan ini ke template sebelumnya:

```

{
  "Effect": "Allow",
  "Action": [
    "kms:Decrypt",
  ],
  "Resource": [
    "arn:aws:kms:region:accountId:key/keyId"
  ],
  "Condition": {
    "ArnLike": {
      "kms:EncryptionContext:aws:s3:arn":
        "arn:aws:s3:::bucketFolders*"
    }
  }
}

```

5. Ganti masing-masing *placeholder* dengan informasi Anda sendiri:

- *region*— Nama AWS Region. Misalnya, **us-east-1**.
- *accountId*— Akun AWS ID tempat bucket S3 berada.
- *database/databases*, *table/databases/tablescatalog*, dan *database/default* — Lokasi data pelatihan yang AWS Clean Rooms perlu diakses.

- **bucket**— Nama Sumber Daya Amazon (ARN) dari ember S3. Nama Sumber Daya Amazon (ARN) dapat ditemukan di tab Properties bucket di Amazon S3.
- **bucketFolders**— Nama folder tertentu di bucket S3 yang AWS Clean Rooms perlu diakses.

6. Pilih Berikutnya.
7. Untuk meninjau dan membuat, masukkan nama Kebijakan dan Deskripsi, dan tinjau Ringkasan.
8. Pilih Buat kebijakan.

Anda telah membuat kebijakan untuk AWS Clean Rooms.

9. Di bawah Manajemen akses, pilih Peran.

Dengan Peran, Anda dapat membuat kredensi jangka pendek, yang direkomendasikan untuk meningkatkan keamanan. Anda juga dapat memilih Pengguna untuk membuat kredensi jangka panjang.

10. Pilih Buat peran.
11. Di wizard Buat peran, untuk jenis entitas Tepercaya, pilih Kebijakan kepercayaan khusus.
12. Salin dan tempel kebijakan kepercayaan khusus berikut ke editor JSON.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowAssumeRole",
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms-ml.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "StringEqualsIfExists": {
          "aws:SourceAccount": ["111122223333"]
        },
        "ArnLikeIfExists": {
          "aws:SourceArn": "arn:aws:cleanrooms-ml:us-east-1:111122223333:training-dataset/*"
        }
      }
    }
  ]
}
```

```
}
]
}
```

SourceAccountItu selalu milikmu Akun AWS. Ini SourceArn dapat dibatasi pada kumpulan data pelatihan tertentu, tetapi hanya setelah kumpulan data itu dibuat. Karena Anda belum mengetahui kumpulan data pelatihan ARN, wildcard ditentukan di sini.

accountId adalah ID Akun AWS yang berisi data pelatihan.

13. Pilih Berikutnya dan di bawah Tambahkan izin, masukkan nama kebijakan yang baru saja Anda buat. (Anda mungkin perlu memuat ulang halaman.)
14. Pilih kotak centang di samping nama kebijakan yang Anda buat, lalu pilih Berikutnya.
15. Untuk Nama, tinjau, dan buat, masukkan nama Peran dan Deskripsi.

Note

Nama Peran harus cocok dengan pola dalam passRole izin yang diberikan kepada anggota yang dapat melakukan kueri dan menerima hasil dan peran anggota.

- a. Tinjau Pilih entitas tepercaya, dan edit jika perlu.
- b. Tinjau izin di Tambahkan izin, dan edit jika perlu.
- c. Tinjau Tag, dan tambahkan tag jika perlu.
- d. Pilih Buat peran.

Anda telah membuat peran layanan untuk AWS Clean Rooms.

Buat peran layanan untuk menulis segmen yang mirip

AWS Clean Rooms menggunakan peran layanan untuk menulis segmen yang mirip ke ember. Anda dapat membuat peran ini menggunakan konsol jika Anda memiliki izin IAM yang diperlukan. Jika Anda tidak memiliki CreateRole izin, minta administrator Anda untuk membuat peran layanan.

Untuk membuat peran layanan untuk menulis segmen mirip

1. Masuk ke konsol IAM (<https://console.aws.amazon.com/iam/>) dengan akun administrator Anda.

2. Di bagian Manajemen akses, pilih Kebijakan.
3. Pilih Buat kebijakan.
4. Di editor Kebijakan, pilih tab JSON, lalu salin dan tempel kebijakan berikut.

Note

Kebijakan contoh berikut mendukung izin yang diperlukan untuk membaca AWS Glue metadata dan data Amazon S3 yang sesuai. Namun, Anda mungkin perlu mengubah kebijakan ini tergantung pada cara Anda mengatur data Amazon S3 Anda. Kebijakan ini tidak menyertakan kunci KMS untuk mendekripsi data.

AWS Glue Sumber daya Anda dan sumber daya Amazon S3 yang mendasarinya harus AWS Region sama dengan kolaborasi. AWS Clean Rooms

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:ListBucket",
        "s3:GetBucketLocation"
      ],
      "Resource": [
        "arn:aws:s3:::buckets"
      ],
      "Condition": {
        "StringEquals": {
          "s3:ResourceAccount": [
            "accountId"
          ]
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "s3:PutObject"
      ]
    }
  ]
}
```

```

    ],
    "Resource": [
      "arn:aws:s3:::bucketFolders/*"
    ],
    "Condition": {
      "StringEquals": {
        "s3:ResourceAccount": [
          "accountId"
        ]
      }
    }
  }
}

```

Jika Anda perlu menggunakan kunci KMS untuk mengenkripsi data, tambahkan AWS KMS pernyataan ini ke template:

```

{
  "Effect": "Allow",
  "Action": [
    "kms:Encrypt",
    "kms:GenerateDataKey*",
    "kms:ReEncrypt*",
  ],
  "Resource": [
    "arn:aws:kms:region:accountId:key/keyId"
  ],
  "Condition": {
    "ArnLike": {
      "kms:EncryptionContext:aws:s3:arn":
        "arn:aws:s3:::bucketFolders*"
    }
  }
}

```

5. Ganti masing-masing *placeholder* dengan informasi Anda sendiri:

- *buckets*— Nama Sumber Daya Amazon (ARN) dari ember S3. Nama Sumber Daya Amazon (ARN) dapat ditemukan di tab Properties bucket di Amazon S3.

- *accountId*— Akun AWS ID tempat bucket S3 berada.
- *bucketFolders*— Nama folder tertentu di bucket S3 yang AWS Clean Rooms perlu diakses.
- *region*— Nama AWS Region. Misalnya, **us-east-1**.
- *keyId*— Kunci KMS diperlukan untuk mengenkripsi data Anda.

6. Pilih Berikutnya.

7. Untuk meninjau dan membuat, masukkan nama Kebijakan dan Deskripsi, dan tinjau Ringkasan.

8. Pilih Buat kebijakan.

Anda telah membuat kebijakan untuk AWS Clean Rooms.

9. Di bawah Manajemen akses, pilih Peran.

Dengan Peran, Anda dapat membuat kredensi jangka pendek, yang direkomendasikan untuk meningkatkan keamanan. Anda juga dapat memilih Pengguna untuk membuat kredensi jangka panjang.

10. Pilih Buat peran.

11. Di wizard Buat peran, untuk jenis entitas Tepercaya, pilih Kebijakan kepercayaan khusus.

12. Salin dan tempel kebijakan kepercayaan khusus berikut ke editor JSON.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowAssumeRole",
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms-ml.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "StringEqualsIfExists": {
          "aws:SourceAccount": ["111122223333"]
        },
        "ArnLikeIfExists": {
          "aws:SourceArn": "arn:aws:cleanrooms-ml:us-east-1:111122223333:configured-audience-model/*"
        }
      }
    }
  ]
}
```

```
}  
}  
}  
]  
}
```

SourceAccountItu selalu milikmu Akun AWS. Ini SourceArn dapat dibatasi pada kumpulan data pelatihan tertentu, tetapi hanya setelah kumpulan data itu dibuat. Karena Anda belum mengetahui kumpulan data pelatihan ARN, wildcard ditentukan di sini.

13. Pilih Berikutnya.
14. Pilih kotak centang di samping nama kebijakan yang Anda buat, lalu pilih Berikutnya.
15. Untuk Nama, tinjau, dan buat, masukkan nama Peran dan Deskripsi.

 Note

Nama Peran harus cocok dengan pola dalam passRole izin yang diberikan kepada anggota yang dapat melakukan kueri dan menerima hasil dan peran anggota.

- a. Tinjau Pilih entitas tepercaya, dan edit jika perlu.
- b. Tinjau izin di Tambahkan izin, dan edit jika perlu.
- c. Tinjau Tag, dan tambahkan tag jika perlu.
- d. Pilih Buat peran.

Anda telah membuat peran layanan untuk AWS Clean Rooms.

Buat peran layanan untuk membaca data benih

AWS Clean Rooms menggunakan peran layanan untuk membaca data benih. Anda dapat membuat peran ini menggunakan konsol jika Anda memiliki izin IAM yang diperlukan. Jika Anda tidak memiliki CreateRole izin, minta administrator Anda untuk membuat peran layanan.

Untuk membuat peran layanan untuk membaca data benih yang disimpan dalam bucket S3.

1. Masuk ke konsol IAM (<https://console.aws.amazon.com/iam/>) dengan akun administrator Anda.
2. Di bagian Manajemen akses, pilih Kebijakan.

3. Pilih Buat kebijakan.
4. Di editor Kebijakan, pilih tab JSON, lalu salin dan tempel salah satu kebijakan berikut.

 Note

Kebijakan contoh berikut mendukung izin yang diperlukan untuk membaca AWS Glue metadata dan data Amazon S3 yang sesuai. Namun, Anda mungkin perlu mengubah kebijakan ini tergantung pada cara Anda mengatur data Amazon S3 Anda. Kebijakan ini tidak menyertakan kunci KMS untuk mendekripsi data.

AWS Glue Sumber daya Anda dan sumber daya Amazon S3 yang mendasarinya harus AWS Region sama dengan kolaborasi. AWS Clean Rooms

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:ListBucket"
      ],
      "Resource": [
        "arn:aws:s3:::buckets"
      ],
      "Condition": {
        "StringEquals": {
          "s3:ResourceAccount": [
            "accountId"
          ]
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "s3:GetObject"
      ],
      "Resource": [
```

```

        "arn:aws:s3:::bucketFolders/*"
    ],
    "Condition": {
        "StringEquals": {
            "s3:ResourceAccount": [
                "accountId"
            ]
        }
    }
}
]
}

```

Note

Contoh kebijakan berikut mendukung izin yang diperlukan untuk membaca hasil query SQL dan menggunakannya sebagai data input. Namun, Anda mungkin perlu mengubah kebijakan ini tergantung pada bagaimana kueri Anda terstruktur. Kebijakan ini tidak menyertakan kunci KMS untuk mendekripsi data.

JSON

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowCleanRoomsStartQuery",
      "Effect": "Allow",
      "Action": [
        "cleanrooms:GetCollaborationAnalysisTemplate",
        "cleanrooms:GetSchema",
        "cleanrooms:StartProtectedQuery"
      ],
      "Resource": "*"
    },
    {
      "Sid": "AllowCleanRoomsGetAndUpdateQuery",
      "Effect": "Allow",
      "Action": [

```

```

        "cleanrooms:GetProtectedQuery",
        "cleanrooms:UpdateProtectedQuery"
    ],
    "Resource": [
        "arn:aws:cleanrooms:us-east-1:111122223333:membership/queryRunnerMembershipId"
    ]
}

```

Jika Anda perlu menggunakan kunci KMS untuk mendekripsi data, tambahkan AWS KMS pernyataan ini ke template:

```

{
    "Effect": "Allow",
    "Action": [
        "kms:Decrypt",
        "kms:DescribeKey"
    ],
    "Resource": [
        "arn:aws:kms:region:accountId:key/keyId"
    ],
    "Condition": {
        "ArnLike": {
            "kms:EncryptionContext:aws:s3:arn":
                "arn:aws:s3:::bucketFolders*"
        }
    }
}

```

5. Ganti masing-masing *placeholder* dengan informasi Anda sendiri:

- *buckets*— Nama Sumber Daya Amazon (ARN) dari ember S3. Nama Sumber Daya Amazon (ARN) dapat ditemukan di tab Properties bucket di Amazon S3.
- *accountId*— Akun AWS ID tempat bucket S3 berada.
- *bucketFolders*— Nama folder tertentu di bucket S3 yang AWS Clean Rooms perlu diakses.
- *region*— Nama AWS Region. Misalnya, **us-east-1**.

- *queryRunnerAccountId*— Akun AWS ID akun yang akan menjalankan kueri.
- *queryRunnerMembershipId*— ID Keanggotaan anggota yang dapat melakukan query. ID Keanggotaan dapat ditemukan di tab Detail kolaborasi. Ini memastikan bahwa AWS Clean Rooms mengasumsikan peran hanya ketika anggota ini menjalankan analisis dalam kolaborasi ini.
- *keyId*— Kunci KMS diperlukan untuk mengenkripsi data Anda.

6. Pilih Berikutnya.

7. Untuk meninjau dan membuat, masukkan nama Kebijakan dan Deskripsi, dan tinjau Ringkasan.

8. Pilih Buat kebijakan.

Anda telah membuat kebijakan untuk AWS Clean Rooms.

9. Di bawah Manajemen akses, pilih Peran.

Dengan Peran, Anda dapat membuat kredensi jangka pendek, yang direkomendasikan untuk meningkatkan keamanan. Anda juga dapat memilih Pengguna untuk membuat kredensi jangka panjang.

10. Pilih Buat peran.

11. Di wizard Buat peran, untuk jenis entitas Tepercaya, pilih Kebijakan kepercayaan khusus.

12. Salin dan tempel kebijakan kepercayaan khusus berikut ke editor JSON.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowAssumeRole",
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms-ml.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "StringEqualsIfExists": {
          "aws:SourceAccount": ["111122223333"]
        }
      }
    }
  ]
}
```



```
    "ArnLikeIfExists": {  
      "aws:SourceArn": "arn:aws:cleanrooms-ml:us-  
east-1:111122223333:audience-generation-job/*"  
    }  
  }  
}
```

SourceAccountItu selalu milikmu Akun AWS. Ini SourceArn dapat dibatasi pada kumpulan data pelatihan tertentu, tetapi hanya setelah kumpulan data itu dibuat. Karena Anda belum mengetahui kumpulan data pelatihan ARN, wildcard ditentukan di sini.

13. Pilih Berikutnya.
14. Pilih kotak centang di samping nama kebijakan yang Anda buat, lalu pilih Berikutnya.
15. Untuk Nama, tinjau, dan buat, masukkan nama Peran dan Deskripsi.

 Note

Nama Peran harus cocok dengan pola dalam passRole izin yang diberikan kepada anggota yang dapat melakukan kueri dan menerima hasil dan peran anggota.

- a. Tinjau Pilih entitas tepercaya, dan edit jika perlu.
- b. Tinjau izin di Tambahkan izin, dan edit jika perlu.
- c. Tinjau Tag, dan tambahkan tag jika perlu.
- d. Pilih Buat peran.

Anda telah membuat peran layanan untuk AWS Clean Rooms.

Siapkan peran layanan untuk pemodelan khusus

Topik

- [Buat peran layanan untuk pemodelan ML kustom - Konfigurasi ML](#)
- [Buat peran layanan untuk menyediakan model ML kustom](#)
- [Membuat peran layanan untuk menanyakan kumpulan data](#)

- [Membuat peran layanan untuk membuat asosiasi tabel yang dikonfigurasi](#)

Buat peran layanan untuk pemodelan ML kustom - Konfigurasi ML

AWS Clean Rooms menggunakan peran layanan untuk mengontrol siapa yang dapat membuat konfigurasi HTML kustom. Anda dapat membuat peran ini menggunakan konsol jika Anda memiliki izin IAM yang diperlukan. Jika Anda tidak memiliki `CreateRole` izin, minta administrator Anda untuk membuat peran layanan.

Peran ini memungkinkan Anda untuk menggunakan `MLConfiguration` tindakan [Put](#).

Untuk membuat peran layanan untuk memungkinkan pembuatan konfigurasi HTML kustom

1. Masuk ke konsol IAM (<https://console.aws.amazon.com/iam/>) dengan akun administrator Anda.
2. Di bagian Manajemen akses, pilih Kebijakan.
3. Pilih Buat kebijakan.
4. Di editor Kebijakan, pilih tab JSON, lalu salin dan tempel kebijakan berikut.

Note

Contoh kebijakan berikut mendukung izin yang diperlukan untuk mengakses dan menulis data ke bucket S3 dan untuk mempublikasikan CloudWatch metrik. Namun, Anda mungkin perlu mengubah kebijakan ini tergantung pada cara Anda mengatur data Amazon S3 Anda. Kebijakan ini tidak menyertakan kunci KMS untuk mendekripsi data. Sumber daya Amazon S3 Anda harus AWS Region sama dengan kolaborasi. AWS Clean Rooms

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowS3ObjectWriteForExport",
      "Effect": "Allow",
      "Action": [
        "s3:PutObject"
      ]
    }
  ]
}
```

```

    ],
    "Resource": [
        "arn:aws:s3:::bucket/*"
    ],
    "Condition": {
        "StringEquals": {
            "s3:ResourceAccount": [
                "111122223333"
            ]
        }
    }
},
{
    "Sid": "AllowS3KMSEncryptForExport",
    "Effect": "Allow",
    "Action": [
        "kms:Encrypt",
        "kms:GenerateDataKey*"
    ],
    "Resource": [
        "arn:aws:kms:us-east-1:111122223333:key/keyId"
    ],
    "Condition": {
        "StringLike": {
            "kms:EncryptionContext:aws:s3:arn":
"arn:aws:s3:::bucket*"
        }
    }
},
{
    "Sid": "AllowCloudWatchMetricsPublishingForTrainingJobs",
    "Action": "cloudwatch:PutMetricData",
    "Resource": "*",
    "Effect": "Allow",
    "Condition": {
        "StringLike": {
            "cloudwatch:namespace": "/aws/cleanroomsml/*"
        }
    }
},
{
    "Sid": "AllowCloudWatchLogsPublishingForTrainingOrInferenceJobs",
    "Effect": "Allow",
    "Action": [

```

```

        "logs:CreateLogGroup",
        "logs:CreateLogStream",
        "logs:DescribeLogStreams",
        "logs:PutLogEvents"
    ],
    "Resource": [
        "arn:aws:logs:us-east-1:111122223333:log-group:/aws/
cleanroomsm1/*"
    ]
}
]
}

```

5. Ganti masing-masing *placeholder* dengan informasi Anda sendiri:
 - *bucket*— Nama Sumber Daya Amazon (ARN) dari ember S3. Nama Sumber Daya Amazon (ARN) dapat ditemukan di tab Properties bucket di Amazon S3.
 - *region*- Nama AWS Region. Misalnya, **us-east-1**.
 - *accountId*— Akun AWS ID tempat bucket S3 berada.
 - *keyId*— Kunci KMS diperlukan untuk mengenkripsi data Anda.
 6. Pilih Berikutnya.
 7. Untuk meninjau dan membuat, masukkan nama Kebijakan dan Deskripsi, dan tinjau Ringkasan.
 8. Pilih Buat kebijakan.
- Anda telah membuat kebijakan untuk AWS Clean Rooms.
9. Di bawah Manajemen akses, pilih Peran.
- Dengan Peran, Anda dapat membuat kredensi jangka pendek, yang direkomendasikan untuk meningkatkan keamanan. Anda juga dapat memilih Pengguna untuk membuat kredensi jangka panjang.
10. Pilih Buat peran.
 11. Di wizard Buat peran, untuk jenis entitas Tepercaya, pilih Kebijakan kepercayaan khusus.
 12. Salin dan tempel kebijakan kepercayaan khusus berikut ke editor JSON.

JSON

```

{
  "Version": "2012-10-17",

```

```

    "Statement": [
      {
        "Effect": "Allow",
        "Principal": {
          "Service": "cleanrooms-ml.amazonaws.com"
        },
        "Action": "sts:AssumeRole",
        "Condition": {
          "StringEquals": {
            "aws:SourceAccount": "111122223333"
          },
          "ArnLike": {
            "aws:SourceArn": "arn:aws:cleanrooms:us-
east-1:111122223333:membership/membershipID"
          }
        }
      }
    ]
  }
}

```

SourceAccount itu selalu milikmu Akun AWS. Ini SourceArn dapat dibatasi pada kumpulan data pelatihan tertentu, tetapi hanya setelah kumpulan data itu dibuat. Karena Anda belum mengetahui kumpulan data pelatihan ARN, wildcard ditentukan di sini.

13. Pilih Berikutnya.
14. Pilih kotak centang di samping nama kebijakan yang Anda buat, lalu pilih Berikutnya.
15. Untuk Nama, tinjau, dan buat, masukkan nama Peran dan Deskripsi.

Note

Nama Peran harus cocok dengan pola dalam passRole izin yang diberikan kepada anggota yang dapat melakukan kueri dan menerima hasil dan peran anggota.

- a. Tinjau Pilih entitas tepercaya, dan edit jika perlu.
- b. Tinjau izin di Tambahkan izin, dan edit jika perlu.
- c. Tinjau Tag, dan tambahkan tag jika perlu.
- d. Pilih Buat peran.

Anda telah membuat peran layanan untuk AWS Clean Rooms.

Buat peran layanan untuk menyediakan model ML kustom

AWS Clean Rooms menggunakan peran layanan untuk mengontrol siapa yang dapat membuat algoritma model HTML kustom. Anda dapat membuat peran ini menggunakan konsol jika Anda memiliki izin IAM yang diperlukan. Jika Anda tidak memiliki `CreateRole` izin, minta administrator Anda untuk membuat peran layanan.

Peran ini memungkinkan Anda untuk menggunakan [CreateConfiguredModelAlgorithm](#) tindakan.

Untuk membuat peran layanan agar anggota dapat menyediakan model ML kustom

1. Masuk ke konsol IAM (<https://console.aws.amazon.com/iam/>) dengan akun administrator Anda.
2. Di bagian Manajemen akses, pilih Kebijakan.
3. Pilih Buat kebijakan.
4. Di editor Kebijakan, pilih tab JSON, lalu salin dan tempel kebijakan berikut.

Note

Contoh kebijakan berikut mendukung izin yang diperlukan untuk mengambil image docker yang berisi algoritma model. Namun, Anda mungkin perlu mengubah kebijakan ini tergantung pada cara Anda mengatur data Amazon S3 Anda. Kebijakan ini tidak menyertakan kunci KMS untuk mendekripsi data.

Sumber daya Amazon S3 Anda harus AWS Region sama dengan kolaborasi. AWS Clean Rooms

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowECRIImageDownloadForTrainingAndInferenceJobs",
      "Effect": "Allow",
      "Action": [
        "ecr:BatchGetImage",
        "ecr:BatchCheckLayerAvailability",
```

```

        "ecr:GetDownloadUrlForLayer"
      ],
      "Resource": "arn:aws:ecr:us-east-1:111122223333:repository/repoName"
    }
  ]
}

```

5. Ganti masing-masing *placeholder* dengan informasi Anda sendiri:
 - *region*— Nama AWS Region. Misalnya, **us-east-1**.
 - *accountId*— Akun AWS ID tempat bucket S3 berada.
 - *repoName*— Nama repositori yang berisi data Anda.
6. Pilih Berikutnya.
7. Untuk meninjau dan membuat, masukkan nama Kebijakan dan Deskripsi, dan tinjau Ringkasan.
8. Pilih Buat kebijakan.

Anda telah membuat kebijakan untuk AWS Clean Rooms.

9. Di bawah Manajemen akses, pilih Peran.

Dengan Peran, Anda dapat membuat kredensi jangka pendek, yang direkomendasikan untuk meningkatkan keamanan. Anda juga dapat memilih Pengguna untuk membuat kredensi jangka panjang.

10. Pilih Buat peran.
11. Di wizard Buat peran, untuk jenis entitas Tepercaya, pilih Kebijakan kepercayaan khusus.
12. Salin dan tempel kebijakan kepercayaan khusus berikut ke editor JSON.

JSON

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms-ml.amazonaws.com"
      },
      "Action": "sts:AssumeRole"
    }
  ]
}

```

```
}
]
}
```

SourceAccountItu selalu milik Anda Akun AWS . SourceArn Dapat dibatasi pada kumpulan data pelatihan tertentu, tetapi hanya setelah kumpulan data itu dibuat. Karena Anda belum mengetahui kumpulan data pelatihan ARN, wildcard ditentukan di sini.

13. Pilih Berikutnya.
14. Pilih kotak centang di samping nama kebijakan yang Anda buat, lalu pilih Berikutnya.
15. Untuk Nama, tinjau, dan buat, masukkan nama Peran dan Deskripsi.

Note

Nama Peran harus cocok dengan pola dalam passRole izin yang diberikan kepada anggota yang dapat melakukan kueri dan menerima hasil dan peran anggota.

- a. Tinjau Pilih entitas tepercaya, dan edit jika perlu.
- b. Tinjau izin di Tambahkan izin, dan edit jika perlu.
- c. Tinjau Tag, dan tambahkan tag jika perlu.
- d. Pilih Buat peran.

Anda telah membuat peran layanan untuk AWS Clean Rooms.

Membuat peran layanan untuk menanyakan kumpulan data

AWS Clean Rooms menggunakan peran layanan untuk mengontrol siapa yang dapat menanyakan kumpulan data yang akan digunakan untuk pemodelan HTML kustom. Anda dapat membuat peran ini menggunakan konsol jika Anda memiliki izin IAM yang diperlukan. Jika Anda tidak memiliki CreateRole izin, minta administrator Anda untuk membuat peran layanan.

Peran ini memungkinkan Anda untuk menggunakan tindakan [Buat MLInput Saluran](#).

Untuk membuat peran layanan untuk memungkinkan anggota melakukan kueri kumpulan data

1. Masuk ke konsol IAM (<https://console.aws.amazon.com/iam/>) dengan akun administrator Anda.
2. Di bagian Manajemen akses, pilih Kebijakan.

3. Pilih Buat kebijakan.
4. Di editor Kebijakan, pilih tab JSON, lalu salin dan tempel kebijakan berikut.

 Note

Contoh kebijakan berikut mendukung izin yang diperlukan untuk menanyakan kumpulan data yang akan digunakan untuk pemodelan HTML kustom. Namun, Anda mungkin perlu mengubah kebijakan ini tergantung pada cara Anda mengatur data Amazon S3 Anda. Kebijakan ini tidak menyertakan kunci KMS untuk mendekripsi data. Sumber daya Amazon S3 Anda harus AWS Region sama dengan kolaborasi. AWS Clean Rooms

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowCleanRoomsStartQueryForMLInputChannel",
      "Effect": "Allow",
      "Action": "cleanrooms:StartProtectedQuery",
      "Resource": "*"
    },
    {
      "Sid": "AllowCleanroomsGetSchemaAndGetAnalysisTemplateForMLInputChannel",
      "Effect": "Allow",
      "Action": [
        "cleanrooms:GetSchema",
        "cleanrooms:GetCollaborationAnalysisTemplate"
      ],
      "Resource": "*"
    },
    {
      "Sid": "AllowCleanRoomsGetAndUpdateQueryForMLInputChannel",
      "Effect": "Allow",
      "Action": [
        "cleanrooms:GetProtectedQuery",
        "cleanrooms:UpdateProtectedQuery"
      ]
    }
  ]
}
```

```

        ],
        "Resource": [
            "arn:aws:cleanrooms:us-east-1:111122223333:membership/queryRunnerMembershipId"
        ]
    }
]
}

```

5. Ganti masing-masing *placeholder* dengan informasi Anda sendiri:
 - *region*- Nama AWS Region. Misalnya, **us-east-1**.
 - *queryRunnerAccountId*— Akun AWS ID akun yang akan menjalankan kueri.
 - *queryRunnerMembershipId*— ID Keanggotaan anggota yang dapat melakukan query. ID Keanggotaan dapat ditemukan di tab Detail kolaborasi. Ini memastikan bahwa AWS Clean Rooms mengasumsikan peran hanya ketika anggota ini menjalankan analisis dalam kolaborasi ini.
6. Pilih Berikutnya.
7. Untuk meninjau dan membuat, masukkan nama Kebijakan dan Deskripsi, dan tinjau Ringkasan.
8. Pilih Buat kebijakan.

Anda telah membuat kebijakan untuk AWS Clean Rooms.

9. Di bawah Manajemen akses, pilih Peran.

Dengan Peran, Anda dapat membuat kredensi jangka pendek, yang direkomendasikan untuk meningkatkan keamanan. Anda juga dapat memilih Pengguna untuk membuat kredensi jangka panjang.

10. Pilih Buat peran.
11. Di wizard Buat peran, untuk jenis entitas Tepercaya, pilih Kebijakan kepercayaan khusus.
12. Salin dan tempel kebijakan kepercayaan khusus berikut ke editor JSON.

JSON

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",

```

```
    "Principal": {  
      "Service": "cleanrooms-ml.amazonaws.com"  
    },  
    "Action": "sts:AssumeRole"  
  }  
]  
}
```

SourceAccountItu selalu milik Anda Akun AWS . SourceArn Dapat dibatasi pada kumpulan data pelatihan tertentu, tetapi hanya setelah kumpulan data itu dibuat. Karena Anda belum mengetahui kumpulan data pelatihan ARN, wildcard ditentukan di sini.

13. Pilih Berikutnya.
14. Pilih kotak centang di samping nama kebijakan yang Anda buat, lalu pilih Berikutnya.
15. Untuk Nama, tinjau, dan buat, masukkan nama Peran dan Deskripsi.

 Note

Nama Peran harus cocok dengan pola dalam passRole izin yang diberikan kepada anggota yang dapat melakukan kueri dan menerima hasil dan peran anggota.

- a. Tinjau Pilih entitas tepercaya, dan edit jika perlu.
- b. Tinjau izin di Tambahkan izin, dan edit jika perlu.
- c. Tinjau Tag, dan tambahkan tag jika perlu.
- d. Pilih Buat peran.

Anda telah membuat peran layanan untuk AWS Clean Rooms.

Membuat peran layanan untuk membuat asosiasi tabel yang dikonfigurasi

AWS Clean Rooms menggunakan peran layanan untuk mengontrol siapa yang dapat membuat asosiasi tabel yang dikonfigurasi. Anda dapat membuat peran ini menggunakan konsol jika Anda memiliki izin IAM yang diperlukan. Jika Anda tidak memiliki CreateRole izin, minta administrator Anda untuk membuat peran layanan.

Peran ini memungkinkan Anda untuk menggunakan CreateConfiguredTableAssociation tindakan.

Untuk membuat peran layanan untuk memungkinkan pembuatan asosiasi tabel yang dikonfigurasi

1. Masuk ke konsol IAM (<https://console.aws.amazon.com/iam/>) dengan akun administrator Anda.
2. Di bagian Manajemen akses, pilih Kebijakan.
3. Pilih Buat kebijakan.
4. Di editor Kebijakan, pilih tab JSON, lalu salin dan tempel kebijakan berikut.

 Note

Contoh kebijakan berikut mendukung pembuatan asosiasi tabel yang dikonfigurasi. Namun, Anda mungkin perlu mengubah kebijakan ini tergantung pada cara Anda mengatur data Amazon S3 Anda. Kebijakan ini tidak menyertakan kunci KMS untuk mendekripsi data.

Sumber daya Amazon S3 Anda harus AWS Region sama dengan kolaborasi. AWS Clean Rooms

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "kms:Decrypt",
        "kms:DescribeKey"
      ],
      "Resource": "arn:aws:kms:us-east-1:111122223333:key/KMS-key-ID",
      "Effect": "Allow"
    },
    {
      "Action": [
        "s3:ListBucket",
        "s3:GetBucketLocation"
      ],
      "Resource": "arn:aws:s3:::bucket-name",
      "Effect": "Allow"
    },
    {
```

```

        "Action": "s3:GetObject",
        "Resource": "arn:aws:s3:::bucket-name/*",
        "Effect": "Allow"
    },
    {
        "Action": [
            "glue:GetDatabase",
            "glue:GetDatabases",
            "glue:GetTable",
            "glue:GetTables",
            "glue:GetPartitions",
            "glue:GetPartition",
            "glue:BatchGetPartition"
        ],
        "Resource": [
            "arn:aws:glue:us-east-1:111122223333:catalog",
            "arn:aws:glue:us-east-1:111122223333:database/Glue database name",
            "arn:aws:glue:us-east-1:111122223333:table/Glue database name/Glue table name"
        ],
        "Effect": "Allow"
    },
    {
        "Action": [
            "glue:GetSchema",
            "glue:GetSchemaVersion"
        ],
        "Resource": "*",
        "Effect": "Allow"
    }
]
}

```

Ganti Sumber Daya Placeholder ARNs

Saat menggunakan kebijakan ini, Anda harus mengganti pengenalan sumber daya placeholder dengan sumber daya yang sebenarnya ARNs :

- AWS KMS Sumber Daya Kunci: Ganti *KMS-key-ID* dengan ID AWS KMS kunci aktual yang mengenkripsi data Amazon S3 Anda. Kuncinya harus berada di akun yang sama (111122223333) yang memiliki sumber daya katalog. AWS Glue

- Sumber Daya Bucket Amazon S3: Ganti *bucket-name* dengan nama sebenarnya bucket Amazon S3 yang berisi AWS Glue data tabel Anda. Perhatikan bahwa bucket Amazon S3 ARNs tidak menyertakan akun IDs karena nama bucket unik secara global.
- AWS Glue Sumber Daya: Ganti placeholder berikut dengan nama sumber daya Anda yang sebenarnya:
 - *Glue database name*- Nama AWS Glue database Anda
 - *Glue table name*- Nama AWS Glue meja Anda

Semua AWS Glue sumber daya (katalog, database, dan tabel) harus sama Akun AWS (111122223333) untuk memastikan izin akses yang konsisten. Akun ini harus sama dengan yang memiliki AWS KMS kunci yang digunakan untuk enkripsi data, menciptakan batas keamanan terpadu untuk sumber daya data Anda. AWS Clean Rooms

5. Ganti masing-masing *placeholder* dengan informasi Anda sendiri:
 - *KMS key used to encrypt the Amazon S3 data*— Kunci KMS yang digunakan untuk mengenkripsi data Amazon S3. Untuk mendekripsi data, Anda perlu memberikan kunci KMS yang sama yang digunakan untuk mengenkripsi data.
 - *Amazon S3 bucket of AWS Glue table*— Nama bucket Amazon S3 yang berisi AWS Glue tabel yang berisi data Anda.
 - *region*- Nama AWS Region. Misalnya, **us-east-1**.
 - *accountId*— Akun AWS ID akun yang memiliki data.
 - *AWS Glue database name*— Nama AWS Glue database yang berisi data Anda.
 - *AWS Glue table name*— Nama AWS Glue tabel yang berisi data Anda.
6. Pilih Berikutnya.
7. Untuk meninjau dan membuat, masukkan nama Kebijakan dan Deskripsi, dan tinjau Ringkasan.
8. Pilih Buat kebijakan.

Anda telah membuat kebijakan untuk AWS Clean Rooms.

9. Di bawah Manajemen akses, pilih Peran.

Dengan Peran, Anda dapat membuat kredensi jangka pendek, yang direkomendasikan untuk meningkatkan keamanan. Anda juga dapat memilih Pengguna untuk membuat kredensi jangka panjang.

10. Pilih Buat peran.
11. Di wizard Buat peran, untuk jenis entitas Tepercaya, pilih Kebijakan kepercayaan khusus.
12. Salin dan tempel kebijakan kepercayaan khusus berikut ke editor JSON.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms-ml.amazonaws.com"
      },
      "Action": "sts:AssumeRole"
    }
  ]
}
```

SourceAccountItu selalu milik Anda Akun AWS . SourceArn Dapat dibatasi pada kumpulan data pelatihan tertentu, tetapi hanya setelah kumpulan data itu dibuat. Karena Anda belum mengetahui kumpulan data pelatihan ARN, wildcard ditentukan di sini.

13. Pilih Berikutnya.
14. Pilih kotak centang di samping nama kebijakan yang Anda buat, lalu pilih Berikutnya.
15. Untuk Nama, tinjau, dan buat, masukkan nama Peran dan Deskripsi.

 Note

Nama Peran harus cocok dengan pola dalam passRole izin yang diberikan kepada anggota yang dapat melakukan kueri dan menerima hasil dan peran anggota.

- a. Tinjau Pilih entitas tepercaya, dan edit jika perlu.
- b. Tinjau izin di Tambahkan izin, dan edit jika perlu.
- c. Tinjau Tag, dan tambahkan tag jika perlu.
- d. Pilih Buat peran.

Anda telah membuat peran layanan untuk AWS Clean Rooms.

Kolaborasi dan keanggotaan di AWS Clean Rooms

Kolaborasi adalah batas logis yang aman AWS Clean Rooms di mana anggota dapat melakukan analisis pada tabel yang dikonfigurasi.

Setiap anggota AWS Clean Rooms dapat membuat kolaborasi.

Pembuat kolaborasi dapat menunjuk satu anggota untuk menganalisis tabel yang dikonfigurasi dan menerima hasil. Namun, pembuat kolaborasi mungkin ingin mencegah anggota yang dapat menjalankan analisis memiliki akses ke hasil kueri. Dalam hal ini, pembuat kolaborasi dapat menunjuk satu [anggota yang dapat meminta](#) atau [satu anggota yang dapat menjalankan kueri dan pekerjaan dan anggota lain yang dapat menerima](#) hasil.

Dalam kebanyakan kasus, anggota yang dapat menanyakan atau anggota yang dapat menanyakan dan menjalankan pekerjaan juga merupakan [anggota yang membayar biaya komputasi](#). Namun, pembuat kolaborasi dapat mengonfigurasi anggota yang berbeda agar bertanggung jawab membayar biaya komputasi kueri.

Untuk informasi tentang cara membuat kolaborasi menggunakan AWS SDKs, lihat [Referensi AWS Clean Rooms API](#).

Topik

- [Menciptakan kolaborasi](#)
- [Membuat keanggotaan dan bergabung dengan kolaborasi](#)
- [Mengedit kolaborasi](#)
- [Ubah permintaan di AWS Clean Rooms](#)
- [Menghapus kolaborasi](#)
- [Melihat kolaborasi](#)
- [Mengundang anggota untuk berkolaborasi](#)
- [Memantau anggota](#)
- [Menambahkan anggota ke kolaborasi](#)
- [Menghapus anggota dari kolaborasi](#)
- [Meninggalkan Kolaborasi](#)

Menciptakan kolaborasi

Ada tiga cara untuk membuat kolaborasi di AWS Clean Rooms.

Bentuk paling dasar adalah [kolaborasi untuk kueri](#). Kolaborasi ini berfokus pada analisis query SQL dan mempertahankan struktur sederhana dengan dua peran utama: satu anggota yang dapat menjalankan query dan yang lain yang dapat menerima hasil. Pengaturan kolaborasi dasar ini berfungsi dengan baik untuk tugas analisis data sederhana.

Bentuk kedua, [kolaborasi untuk pertanyaan dan pekerjaan](#), memperluas fungsionalitas dengan menggabungkan kueri dan pekerjaan SQL. PySpark Pengaturan kolaborasi ini mempertahankan struktur peran dasar yang sama tetapi memperluas izin untuk menyertakan eksekusi pekerjaan. Persyaratan penting adalah bahwa anggota yang membuat templat PySpark analisis juga harus menjadi orang yang menerima hasil, memastikan akuntabilitas yang jelas dalam proses analisis.

Bentuk ketiga, [kolaborasi untuk pemodelan ML](#), dibangun untuk alur kerja pembelajaran mesin. Pengaturan kolaborasi ini menambahkan dua peran lagi: satu untuk pengguna yang membutuhkan hasil dari model terlatih, dan satu lagi untuk pengguna yang membutuhkan hasil dari menggunakan model tersebut untuk membuat prediksi. Pengaturan kolaborasi ini membantu anggota kolaborasi bekerja sama dalam proyek data yang kompleks sambil menjaga peran dan izin semua orang tetap jelas.

Topik berikut menjelaskan cara membuat kolaborasi untuk kueri, pekerjaan, dan pemodelan ML.

Topik

- [Membuat kolaborasi untuk kueri](#)
- [Membuat kolaborasi untuk pertanyaan dan pekerjaan](#)
- [Membuat kolaborasi untuk pemodelan ML](#)

Membuat kolaborasi untuk kueri

Dalam prosedur ini, Anda sebagai [pembuat kolaborasi](#) melakukan tugas-tugas berikut:

- [Buat kolaborasi](#).
- Undang satu atau lebih [anggota](#) ke [kolaborasi](#).
- Tetapkan kemampuan untuk anggota, seperti [anggota yang dapat meminta](#) dan [anggota yang dapat menerima hasil](#).

Jika pembuat kolaborasi juga anggota yang dapat menerima hasil, mereka menentukan tujuan dan format hasil. Mereka juga menyediakan peran layanan Amazon Resource Name (ARN) untuk menulis hasil ke tujuan hasil.

- Konfigurasi [anggota mana yang bertanggung jawab untuk membayar biaya komputasi dalam kolaborasi](#).

Sebelum Anda mulai, pastikan Anda telah menyelesaikan prasyarat berikut:

- Anda memiliki nama dan Akun AWS ID untuk setiap anggota yang ingin Anda undang ke kolaborasi.
- Anda memiliki izin untuk membagikan nama dan Akun AWS ID untuk setiap anggota dengan semua anggota kolaborasi.

 Note

Anda tidak dapat menambahkan lebih banyak anggota setelah Anda membuat kolaborasi.

Untuk informasi tentang cara membuat kolaborasi menggunakan AWS SDKs, lihat [Referensi AWS Clean Rooms API](#).

Untuk membuat kolaborasi untuk kueri

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Akun AWS yang akan berfungsi sebagai pembuat kolaborasi.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Di sudut kanan atas, pilih Buat kolaborasi.
4. Untuk Langkah 1: Tentukan kolaborasi, lakukan hal berikut:
 - a. Untuk Detail, masukkan Nama dan Deskripsi kolaborasi.

Informasi ini akan terlihat oleh anggota kolaborasi yang diundang untuk berpartisipasi dalam kolaborasi. Nama dan Deskripsi membantu mereka memahami apa yang dimaksud dengan kolaborasi.

- b. Untuk Anggota:

- i. Untuk Anggota 1: Anda, masukkan nama tampilan Anggota sesuai keinginan untuk kolaborasi.

 Note

Akun AWS ID Anda disertakan secara otomatis untuk Akun AWS ID Anggota.

- ii. Untuk Anggota 2, masukkan nama tampilan Anggota dan Akun AWS ID Anggota untuk anggota yang ingin Anda undang ke kolaborasi.

Nama tampilan Anggota dan Akun AWS ID Anggota akan terlihat oleh semua orang yang diundang ke kolaborasi. Setelah Anda memasukkan dan menyimpan nilai untuk bidang ini, Anda tidak dapat mengeditnya.

 Note

Anda harus memberi tahu anggota kolaborasi bahwa Akun AWS ID Anggota dan nama tampilan Anggota mereka akan terlihat oleh semua kolaborator yang diundang dan aktif dalam kolaborasi.

- iii. Jika Anda ingin menambahkan anggota lain, pilih Tambahkan anggota lain. Kemudian masukkan nama tampilan Anggota dan Akun AWS ID Anggota untuk setiap anggota yang dapat menyumbangkan data yang ingin Anda undang ke kolaborasi.
- c. Jika Anda ingin mengaktifkan pencatatan Analisis, pilih kotak centang Aktifkan pencatatan analisis.
 - Pilih kotak centang Log dari kueri di bawah Jenis log yang didukung.

Anda akan menerima log yang dihasilkan dari kueri SQL di akun Amazon CloudWatch Logs Anda.

- d. Jika Anda ingin mengaktifkan Pemantauan terperinci, pilih kotak centang Aktifkan pemantauan terperinci.

Pelari analisis dan pembayar yang dikonfigurasi dapat memilih untuk mengaktifkan metrik terperinci saat mereka membuat keanggotaan mereka. Jika diaktifkan, metrik pemantauan terperinci akan dipublikasikan CloudWatch untuk pemantauan operasional kolaborasi, termasuk kinerja kueri dan pemanfaatan sumber daya. Metrik ini akan tersedia untuk pelari analisis dan pembayar yang dikonfigurasi di masing-masing s. Akun AWS

Untuk informasi selengkapnya tentang CloudWatch harga, lihat [CloudWatch Harga](#).

- e. Di bawah Wilayah hasil kueri yang diizinkan, pilih satu atau Wilayah AWS beberapa tempat Anda ingin mengirim hasil kueri.

Secara default, hanya Wilayah saat ini (seperti N. Virginia us-east-1) yang dipilih.

 Important

Saat Anda mengaktifkan pengiriman hasil kueri lintas wilayah, hasil Anda dapat diproses dan disimpan di luar Wilayah sumber.

Untuk informasi selengkapnya tentang Wilayah, lihat [Wilayah dan Titik Akhir](#) di. Referensi Umum AWS

- f. (Opsional) Kelola akses ke data Anda dengan Persetujuan permintaan perubahan otomatis dengan mengonfigurasi pengaturan mana yang dapat diubah secara otomatis tanpa persetujuan manual untuk permintaan perubahan. Secara default, beberapa pengaturan hanya dapat diubah dengan mengirimkan permintaan perubahan yang harus disetujui oleh semua anggota sebelum dapat diterapkan.
 - Berikan kemampuan anggota — Pilih kemampuan yang dapat diberikan kepada anggota kolaborasi tanpa persetujuan manual. Anggota selalu dapat menyumbangkan data.
 - Pilih kemampuan:
 - Kontribusikan data (selalu diaktifkan)
 - Menerima hasil
 - Persetujuan otomatis anggota baru dengan kemampuan ini — Jika diizinkan, setiap anggota yang ditambahkan dengan kemampuan yang dipilih di atas akan langsung bergabung dengan kolaborasi. Anggota yang ditambahkan dengan kemampuan lain masih memerlukan persetujuan manual untuk bergabung.
 - Kemampuan yang dapat dicabut secara otomatis - Pilih kemampuan yang dapat dicabut tanpa persetujuan manual. Anggota selalu dapat menyumbangkan data.
 - Pilih kemampuan:
 - Kontribusikan data (selalu diaktifkan)
 - Menerima hasil

Jika Anda memilih opsi ini, Anda dapat melacak semua modifikasi konfigurasi kolaborasi melalui riwayat Ubah permintaan, yang terletak di tab Detail di halaman detail kolaborasi.

- g. (Opsional) Jika Anda ingin mengaktifkan kemampuan komputasi kriptografi, pilih kotak centang Aktifkan komputasi kriptografi.

- i. Pilih parameter cakupan kriptografi berikut:

- Izinkan plaintext kolom

Pilih Tidak jika Anda memerlukan tabel terenkripsi sepenuhnya.

Pilih Ya jika Anda ingin cleartext kolom diizinkan dalam tabel terenkripsi.

Untuk menjalankan SUM atau AVG pada kolom tertentu, kolom harus masukcleartext.

- Pertahankan NULL nilai

Pilih Tidak jika Anda tidak ingin mempertahankan NULL nilai. NULLnilai tidak akan muncul seperti NULL dalam tabel terenkripsi.

Pilih Ya jika Anda ingin mempertahankan NULL nilai. NULLnilai akan muncul seperti NULL dalam tabel terenkripsi.

- ii. Pilih parameter Sidik Jari berikut:

- Izinkan duplikat

Pilih Tidak jika Anda tidak ingin entri duplikat diizinkan dalam kolom. fingerprint

Pilih Ya jika Anda ingin entri duplikat diizinkan dalam kolom. fingerprint

- Izinkan JOIN kolom dengan nama yang berbeda

Pilih Tidak jika Anda tidak ingin bergabung dengan fingerprint kolom dengan nama yang berbeda.

Pilih Ya jika Anda ingin bergabung dengan fingerprint kolom dengan nama yang berbeda.

Untuk informasi lebih lanjut tentang parameter komputasi kriptografi, lihat [Parameter komputasi kriptografi](#).

Untuk informasi selengkapnya tentang cara mengenkripsi data Anda untuk digunakan AWS Clean Rooms, lihat [Mempersiapkan tabel data terenkripsi dengan Cryptographic Computing untuk Clean Rooms](#).

 Note

Verifikasi konfigurasi ini dengan hati-hati sebelum menyelesaikan langkah berikutnya. Setelah membuat kolaborasi, Anda hanya dapat mengedit nama kolaborasi, deskripsi, dan apakah log disimpan di Amazon CloudWatch Logs.

- h. Jika Anda ingin mengaktifkan Tag untuk sumber kolaborasi, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.
 - i. Pilih Berikutnya.
5. Untuk Langkah 2: Tentukan kemampuan anggota, untuk Analisis menggunakan kueri dan pekerjaan, di bawah Jenis analisis yang didukung biarkan kotak centang Kueri dipilih dan lakukan tindakan yang disarankan, berdasarkan tujuan Anda.

Tujuan Anda	Tindakan yang disarankan
Kueri data dalam kolaborasi dan terima hasilnya	1. Pilih diri Anda sebagai anggota yang dapat Menjalankan kueri. 2. Pilih diri Anda sebagai anggota yang dapat Menerima hasil dari analisis dari daftar dropdown.
Kueri data dalam kolaborasi dan tetapkan anggota yang berbeda untuk menerima hasil	1. Pilih diri Anda sebagai anggota yang dapat Menjalankan kueri. 2. Pilih anggota yang dapat Menerima hasil dari analisis dari daftar dropdown.
Menerima hasil kueri dalam kolaborasi dan menetapkan anggota yang berbeda untuk kueri data	1. Pilih anggota yang dapat Jalankan kueri dari daftar dropdown.

Tujuan Anda	Tindakan yang disarankan
	2. Pilih diri Anda sebagai anggota yang dapat Menerima hasil dari analisis dari daftar dropdown.
Buat dan kelola kolaborasi, tetapkan anggota yang berbeda untuk menanyakan data, dan menetapkan anggota lain untuk menerima hasil	1. Pilih anggota yang dapat Jalankan kueri dari daftar dropdown. 2. Pilih anggota yang dapat Menerima hasil dari analisis dari daftar dropdown.

- a. Jika Anda menggunakan Clean Rooms ML, untuk pemodelan ML menggunakan alur kerja yang dibuat khusus,
 - i. (Opsional) Pilih anggota yang dapat Menerima output dari model terlatih dari daftar dropdown.
 - ii. (Opsional) Pilih anggota yang dapat Menerima output dari inferensi model dari daftar dropdown.
 - b. Lihat kemampuan anggota di bawah resolusi ID menggunakan Resolusi Entitas AWS.
 - c. Pilih Berikutnya.
6. Untuk Langkah 3: Konfigurasi pembayaran, untuk Analisis menggunakan kueri, lakukan salah satu tindakan berikut berdasarkan tujuan Anda.

Tujuan Anda	Tindakan yang disarankan
Tetapkan anggota yang dapat Menjalankan kueri untuk menjadi anggota yang membayar biaya komputasi kueri	1. Untuk Analisis menggunakan kueri, pilih anggota yang akan Membayar kueri agar sama dengan anggota yang dapat Menjalankan kueri. 2. Pilih Berikutnya.
Tetapkan anggota yang berbeda untuk membayar biaya komputasi kueri	1. Untuk Analisis menggunakan kueri, pilih diri Anda sebagai anggota yang akan Membayar kueri. 2. Pilih Berikutnya.

Untuk pemodelan ML menggunakan alur kerja yang dibuat khusus, Pencipta model mirip yang dikonfigurasi adalah anggota yang akan Membayar untuk pemodelan mirip.

Untuk resolusi ID dengan Resolusi Entitas AWS, Pembuat tabel pemetaan ID adalah anggota yang akan membayar tabel pemetaan ID.

7. Untuk Langkah 4: Konfigurasikan keanggotaan, pilih salah satu opsi berikut:

Yes, join by creating membership now

1. Untuk pengaturan Hasil default, untuk pengaturan hasil Kueri, jika Anda adalah anggota yang dapat Menerima hasil,
 - a. Untuk tujuan Hasil di Amazon S3, masukkan tujuan Amazon S3 atau pilih Jelajahi S3 untuk memilih bucket S3.
 - b. Untuk format hasil kueri, pilih CSV atau PARQUET.
 - c. (Hanya percikan) Untuk file Hasil, pilih Multiple atau Single.
 - d. (Opsional) Untuk akses Layanan, jika Anda ingin mengirimkan kueri yang memakan waktu hingga 24 jam ke tujuan S3 Anda, pilih kotak centang Tambahkan peran layanan untuk mendukung kueri yang membutuhkan waktu hingga 24 jam untuk diselesaikan.

Kueri besar yang membutuhkan waktu hingga 24 jam untuk diselesaikan akan dikirimkan ke tujuan S3 Anda.

Jika Anda tidak memilih kotak centang, hanya kueri yang selesai dalam waktu 12 jam yang akan dikirimkan ke lokasi S3 Anda.

- e. Tentukan izin akses Layanan dengan memilih Buat dan gunakan peran layanan baru atau Gunakan peran layanan yang ada.

Jika Anda memilih untuk...	Lalu...
Membuat dan menggunakan peran layanan baru	- AWS Clean Rooms membuat peran layanan dengan kebijakan yang diperlukan untuk tabel ini. - Nama peran Layanan default adalah <code>cleanrooms-result-receiver-<timestamp></code>

Jika Anda memilih untuk...	Lalu...
	Anda harus memiliki izin untuk membuat peran dan melampirkan kebijakan.
Gunakan peran layanan yang ada	- Pilih nama peran layanan yang ada dari daftar tarik-turun. Daftar peran ditampilkan jika Anda memiliki izin untuk membuat daftar peran. Jika Anda tidak memiliki izin untuk membuat daftar peran, Anda dapat memasukkan Nama Sumber Daya Amazon (ARN) peran yang ingin Anda gunakan. - Lihat peran layanan dengan memilih tautan eksternal Lihat di IAM. Jika tidak ada peran layanan yang ada, opsi untuk Menggunakan peran layanan yang ada tidak tersedia. Secara default, AWS Clean Rooms tidak mencoba memperbarui kebijakan peran yang ada untuk menambahkan izin yang diperlukan.

 Note

- AWS Clean Rooms memerlukan izin untuk melakukan kueri sesuai dengan aturan analisis. Untuk informasi selengkapnya tentang izin AWS Clean Rooms, lihat [AWS kebijakan terkelola untuk AWS Clean Rooms](#).
- Jika peran tidak memiliki izin yang memadai AWS Clean Rooms, Anda akan menerima pesan galat yang menyatakan bahwa peran tersebut tidak memiliki

izin yang memadai untuk peran tersebut. AWS Clean Rooms Kebijakan peran harus ditambahkan sebelum melanjutkan.

- Jika Anda tidak dapat mengubah kebijakan peran, Anda akan menerima pesan galat yang menyatakan bahwa AWS Clean Rooms tidak dapat menemukan kebijakan untuk peran layanan.

2. Untuk pengaturan Log, pilih salah satu opsi berikut untuk penyimpanan Log di Amazon CloudWatch Logs:

 Note

Bagian Pengaturan Log muncul jika Anda memilih untuk mengaktifkan pencatatan Kueri.

- a. Pilih Aktifkan dan log kueri yang relevan dengan Anda akan disimpan di akun Amazon CloudWatch Logs Anda.

Setiap anggota hanya dapat menerima log untuk kueri yang mereka mulai atau yang berisi data mereka.

Anggota yang dapat menerima hasil juga menerima log untuk semua kueri yang dijalankan dalam kolaborasi, meskipun data mereka tidak diakses dalam kueri.

Di bawah Jenis log yang didukung, kotak centang Log kueri diaktifkan secara default.

 Note

Setelah Anda mengaktifkan pencatatan Kueri, penyimpanan log dapat diatur beberapa menit dan mulai menerima log di Amazon CloudWatch Logs. Selama periode singkat ini, anggota yang dapat melakukan kueri mungkin menjalankan kueri yang sebenarnya tidak mengirim log.

- b. Pilih Matikan dan log kueri yang relevan dengan Anda tidak akan disimpan di akun Amazon CloudWatch Logs Anda.
3. Jika Anda ingin mengaktifkan Tag untuk sumber daya keanggotaan, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.

4. Jika Anda adalah anggota yang membayar untuk komputasi Kueri, tunjukkan penerimaan Anda dengan memilih kotak centang Saya setuju untuk membayar biaya komputasi dalam kolaborasi ini.

 Note

Anda harus memilih kotak centang ini untuk melanjutkan.

Untuk informasi selengkapnya tentang cara penghitungan harga, lihat [Harga untuk AWS Clean Rooms](#).

Jika Anda adalah [anggota yang membayar biaya komputasi kueri](#) tetapi bukan [anggota yang dapat melakukan kueri](#), disarankan agar Anda menggunakan AWS Budgets untuk mengonfigurasi anggaran AWS Clean Rooms dan menerima pemberitahuan setelah anggaran maksimum tercapai. Untuk informasi selengkapnya tentang menyiapkan anggaran, lihat [Mengelola biaya Anda AWS Budgets](#) di Panduan AWS Cost Management Pengguna. Untuk informasi selengkapnya tentang mengatur notifikasi, lihat [Membuat topik Amazon SNS untuk pemberitahuan anggaran](#) di AWS Cost Management Panduan Pengguna. Jika anggaran maksimum telah tercapai, Anda dapat menghubungi anggota yang dapat menjalankan kueri atau [meninggalkan kolaborasi](#). Jika Anda meninggalkan kolaborasi, tidak ada lagi kueri yang diizinkan untuk dijalankan, dan oleh karena itu Anda tidak akan lagi ditagih untuk biaya komputasi kueri.

5. Pilih Berikutnya.

Kolaborasi dan keanggotaan Anda dibuat.

Status Anda dalam kolaborasi aktif.

No, I will create a membership later

1. Pilih Berikutnya.

Hanya kolaborasi yang dibuat.

Status Anda dalam kolaborasi tidak aktif.

8. Untuk Langkah 5: Tinjau dan buat, lakukan hal berikut:

- a. Tinjau pilihan yang Anda buat untuk langkah-langkah sebelumnya dan edit jika perlu.

b. Pilih salah satu opsi.

Jika Anda memilih untuk...	Kemudian pilih...
Buat keanggotaan dengan kolaborasi (Ya, bergabung dengan membuat keanggotaan sekarang)	Buat kolaborasi dan keanggotaan
Buat kolaborasi, dan bukan untuk membuat keanggotaan saat ini (Tidak, saya akan membuat keanggotaan nanti)	Buat kolaborasi

Setelah kolaborasi Anda berhasil dibuat, Anda dapat melihat halaman detail kolaborasi di bawah Kolaborasi.

Anda sekarang siap untuk:

- [Siapkan tabel data Anda untuk dianalisis AWS Clean Rooms](#). (Opsional jika Anda ingin menganalisis data peristiwa Anda sendiri atau jika Anda ingin menanyakan data identitas.)
- [Kaitkan tabel yang dikonfigurasi dengan kolaborasi Anda](#). (Opsional jika Anda ingin menganalisis data acara Anda sendiri.)
- [Tambahkan aturan analisis untuk tabel yang dikonfigurasi](#). (Opsional jika Anda ingin menganalisis data acara Anda sendiri.)
- [Buat keanggotaan dan bergabunglah dengan kolaborasi](#). (Opsional jika Anda sudah membuat keanggotaan.)
- [Undang anggota untuk bergabung dengan kolaborasi](#).

Membuat kolaborasi untuk pertanyaan dan pekerjaan

Dalam prosedur ini, Anda sebagai [pembuat kolaborasi](#) melakukan tugas-tugas berikut:

- [Buat kolaborasi](#).
- Undang satu atau lebih [anggota](#) ke [kolaborasi](#).
- Tetapkan kemampuan untuk anggota, seperti [anggota yang dapat menjalankan pertanyaan dan pekerjaan dan anggota yang dapat menerima hasil](#).

Jika pembuat kolaborasi juga anggota yang dapat menerima hasil, mereka menentukan tujuan dan format hasil. Mereka juga menyediakan peran layanan Amazon Resource Name (ARN) untuk menulis hasil ke tujuan hasil.

- Konfigurasi [anggota mana yang bertanggung jawab untuk membayar biaya kueri dan komputasi pekerjaan dalam kolaborasi](#).

Sebelum Anda mulai, pastikan Anda telah menyelesaikan prasyarat berikut:

- Anda memiliki nama dan Akun AWS ID untuk setiap anggota yang ingin Anda undang ke kolaborasi.
- Anda memiliki izin untuk membagikan nama dan Akun AWS ID untuk setiap anggota dengan semua anggota kolaborasi.

 Note

Anda tidak dapat menambahkan lebih banyak anggota setelah Anda membuat kolaborasi.

Untuk informasi tentang cara membuat kolaborasi menggunakan AWS SDKs, lihat [Referensi AWS Clean Rooms API](#).

Untuk membuat kolaborasi untuk pertanyaan dan pekerjaan

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Akun AWS yang akan berfungsi sebagai pembuat kolaborasi.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Di sudut kanan atas, pilih Buat kolaborasi.
4. Untuk Langkah 1: Tentukan kolaborasi, lakukan hal berikut:
 - a. Untuk Detail, masukkan Nama dan Deskripsi kolaborasi.

Informasi ini akan terlihat oleh anggota kolaborasi yang diundang untuk berpartisipasi dalam kolaborasi. Nama dan Deskripsi membantu mereka memahami apa yang dimaksud dengan kolaborasi.

- b. Untuk Anggota:

- i. Untuk Anggota 1: Anda, masukkan nama tampilan Anggota sesuai keinginan untuk kolaborasi.

 Note

Akun AWS ID Anda disertakan secara otomatis untuk Akun AWS ID Anggota.

- ii. Untuk Anggota 2, masukkan nama tampilan Anggota dan Akun AWS ID Anggota untuk anggota yang ingin Anda undang ke kolaborasi.

Nama tampilan Anggota dan Akun AWS ID Anggota akan terlihat oleh semua orang yang diundang ke kolaborasi. Setelah Anda memasukkan dan menyimpan nilai untuk bidang ini, Anda tidak dapat mengeditnya.

 Note

Anda harus memberi tahu anggota kolaborasi bahwa Akun AWS ID Anggota dan nama tampilan Anggota mereka akan terlihat oleh semua kolaborator yang diundang dan aktif dalam kolaborasi.

- iii. Jika Anda ingin menambahkan anggota lain, pilih Tambahkan anggota lain. Kemudian masukkan nama tampilan Anggota dan Akun AWS ID Anggota untuk setiap anggota yang dapat menyumbangkan data yang ingin Anda undang ke kolaborasi.
- c. Jika Anda ingin mengaktifkan Pencatatan analisis, pilih kotak centang Aktifkan pencatatan analisis, lalu pilih Jenis log yang didukung.
 - Jika Anda ingin menerima log yang dihasilkan dari kueri SQL, pilih kotak centang Log dari kueri.
 - Jika Anda ingin menerima log yang dihasilkan dari pekerjaan menggunakan PySpark, pilih kotak centang Log dari pekerjaan.
- d. Jika Anda ingin mengaktifkan Pemantauan terperinci, pilih kotak centang Aktifkan pemantauan terperinci.

Pelari analisis dan pembayar yang dikonfigurasi dapat memilih untuk mengaktifkan metrik terperinci saat mereka membuat keanggotaan mereka. Jika diaktifkan, metrik pemantauan terperinci akan dipublikasikan CloudWatch untuk pemantauan operasional kolaborasi,

termasuk kinerja kueri dan pemanfaatan sumber daya. Metrik ini akan tersedia untuk pelari analisis dan pembayar yang dikonfigurasi di masing-masing s. Akun AWS

Untuk informasi selengkapnya tentang CloudWatch harga, lihat [CloudWatch Harga](#).

- e. Di bawah Wilayah hasil kueri yang diizinkan, pilih satu atau Wilayah AWS beberapa tempat Anda ingin mengirim hasil kueri.

Secara default, hanya Wilayah saat ini (seperti N. Virginia us-east-1) yang dipilih.

 Important

Saat Anda mengaktifkan pengiriman hasil kueri lintas wilayah, hasil Anda dapat diproses dan disimpan di luar Wilayah sumber.

Untuk informasi selengkapnya tentang Wilayah, lihat [Wilayah dan Titik Akhir](#) di. Referensi Umum AWS

- f. (Opsional) Kelola akses ke data Anda dengan Persetujuan permintaan perubahan otomatis dengan mengonfigurasi pengaturan mana yang dapat diubah secara otomatis tanpa persetujuan manual untuk permintaan perubahan. Secara default, beberapa pengaturan hanya dapat diubah dengan mengirimkan permintaan perubahan yang harus disetujui oleh semua anggota sebelum dapat diterapkan.
 - Berikan kemampuan anggota — Pilih kemampuan yang dapat diberikan kepada anggota kolaborasi tanpa persetujuan manual. Anggota selalu dapat menyumbangkan data.
 - Pilih kemampuan:
 - Kontribusikan data (selalu diaktifkan)
 - Menerima hasil
 - Persetujuan otomatis anggota baru dengan kemampuan ini — Jika diizinkan, setiap anggota yang ditambahkan dengan kemampuan yang dipilih di atas akan langsung bergabung dengan kolaborasi. Anggota yang ditambahkan dengan kemampuan lain masih memerlukan persetujuan manual untuk bergabung.
 - Kemampuan yang dapat dicabut secara otomatis - Pilih kemampuan yang dapat dicabut tanpa persetujuan manual. Anggota selalu dapat menyumbangkan data.
 - Pilih kemampuan:
 - Kontribusikan data (selalu diaktifkan)

- Menerima hasil

Jika Anda memilih opsi ini, Anda dapat melacak semua modifikasi konfigurasi kolaborasi melalui riwayat Ubah permintaan, yang terletak di tab Detail di halaman detail kolaborasi.

- g. (Opsional) Jika Anda ingin mengaktifkan kemampuan komputasi kriptografi, pilih kotak centang Aktifkan komputasi kriptografi.

- i. Pilih parameter cakupan kriptografi berikut:

- Izinkan plaintext kolom

Pilih Tidak jika Anda memerlukan tabel terenkripsi sepenuhnya.

Pilih Ya jika Anda ingin cleartext kolom diizinkan dalam tabel terenkripsi.

Untuk menjalankan SUM atau AVG pada kolom tertentu, kolom harus masukcleartext.

- Pertahankan NULL nilai

Pilih Tidak jika Anda tidak ingin mempertahankan NULL nilai. NULLnilai tidak akan muncul seperti NULL dalam tabel terenkripsi.

Pilih Ya jika Anda ingin mempertahankan NULL nilai. NULLnilai akan muncul seperti NULL dalam tabel terenkripsi.

- ii. Pilih parameter Sidik Jari berikut:

- Izinkan duplikat

Pilih Tidak jika Anda tidak ingin entri duplikat diizinkan dalam kolom. fingerprint

Pilih Ya jika Anda ingin entri duplikat diizinkan dalam kolom. fingerprint

- Izinkan JOIN kolom dengan nama yang berbeda

Pilih Tidak jika Anda tidak ingin bergabung dengan fingerprint kolom dengan nama yang berbeda.

Pilih Ya jika Anda ingin bergabung dengan fingerprint kolom dengan nama yang berbeda.

Untuk informasi lebih lanjut tentang parameter komputasi kriptografi, lihat [Parameter komputasi kriptografi](#).

Untuk informasi selengkapnya tentang cara mengenkripsi data Anda untuk digunakan AWS Clean Rooms, lihat [Mempersiapkan tabel data terenkripsi dengan Cryptographic Computing untuk Clean Rooms](#).

 Note

Verifikasi konfigurasi ini dengan hati-hati sebelum menyelesaikan langkah berikutnya. Setelah membuat kolaborasi, Anda hanya dapat mengedit nama kolaborasi, deskripsi, dan apakah log disimpan di Amazon CloudWatch Logs.

- h. Jika Anda ingin mengaktifkan Tag untuk sumber kolaborasi, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.
 - i. Pilih Berikutnya.
5. Untuk Langkah 2: Tentukan kemampuan anggota, lakukan hal berikut:
- a. Untuk Analisis menggunakan kueri dan pekerjaan, di bawah Jenis analisis yang didukung, pilih kotak centang Pekerjaan.

Kotak centang Kueri dipilih secara default.

- i. Pilih anggota yang dapat Menjalankan kueri dan pekerjaan dari daftar dropdown.
- ii. Pilih anggota yang dapat Menerima hasil dari analisis dari daftar dropdown.

 Note

Anggota yang membuat template PySpark analisis juga harus menjadi anggota yang menerima hasil.

- b. Jika Anda menggunakan Clean Rooms ML, untuk pemodelan ML menggunakan alur kerja yang dibuat khusus,
 - i. (Opsional) Pilih anggota yang dapat Menerima output dari model terlatih dari daftar dropdown.

- ii. (Opsional) Pilih anggota yang dapat Menerima output dari inferensi model dari daftar dropdown.
 - c. Lihat kemampuan anggota di bawah resolusi ID menggunakan Resolusi Entitas AWS.
 - d. Pilih Berikutnya.
6. Untuk Langkah 3: Konfigurasi pembayaran,
- a. Untuk Analisis menggunakan kueri dan pekerjaan, pilih anggota yang akan Membayar untuk pertanyaan dan pekerjaan.

Anda dapat menetapkan anggota yang dapat Menjalankan kueri dan pekerjaan untuk menjadi anggota yang membayar kueri dan biaya perhitungan pekerjaan.

Anda dapat menetapkan anggota yang berbeda untuk membayar kueri dan biaya perhitungan pekerjaan.

- b. Untuk pemodelan ML menggunakan alur kerja yang dibuat khusus, Pencipta model mirip yang dikonfigurasi adalah anggota yang akan Membayar untuk pemodelan mirip.
 - c. Untuk resolusi ID dengan Resolusi Entitas AWS, Pembuat tabel pemetaan ID adalah anggota yang akan membayar tabel pemetaan ID.
 - d. Pilih Berikutnya.
7. Untuk Langkah 4: Konfigurasi keanggotaan, pilih salah satu opsi berikut:

Yes, join by creating membership now

- 1. Untuk pengaturan Hasil default, untuk pengaturan hasil Kueri, jika Anda adalah anggota yang dapat Menerima hasil,
 - a. Pilih kotak centang Setel pengaturan default untuk kueri. Untuk tujuan Hasil di Amazon S3, masukkan tujuan Amazon S3 atau pilih Jelajahi S3 untuk memilih bucket S3.
 - b. Untuk format hasil kueri, pilih CSV atau PARQUET.
 - c. (Hanya percikan) Untuk file Hasil, pilih Multiple atau Single.
 - d. (Opsional) Untuk akses Layanan, jika Anda ingin mengirimkan kueri yang memakan waktu hingga 24 jam ke tujuan S3 Anda, pilih kotak centang Tambahkan peran layanan untuk mendukung kueri yang membutuhkan waktu hingga 24 jam untuk diselesaikan.

Kueri besar yang membutuhkan waktu hingga 24 jam untuk diselesaikan akan dikirimkan ke tujuan S3 Anda.

Jika Anda tidak memilih kotak centang, hanya kueri yang selesai dalam waktu 12 jam yang akan dikirimkan ke lokasi S3 Anda.

- e. Tentukan izin akses Layanan dengan memilih Buat dan gunakan peran layanan baru atau Gunakan peran layanan yang ada.

Jika Anda memilih untuk...	Lalu...
Membuat dan menggunakan peran layanan baru	• AWS Clean Rooms membuat peran layanan dengan kebijakan yang diperlukan untuk tabel ini. • Nama peran Layanan default adalah <code>cleanrooms-result-receiver-<timestamp></code> • Anda harus memiliki izin untuk membuat peran dan melampirkan kebijakan.

Jika Anda memilih untuk...	Lalu...
Gunakan peran layanan yang ada	i. Pilih nama peran layanan yang ada dari daftar tarik-turun. Daftar peran ditampilkan jika Anda memiliki izin untuk membuat daftar peran. Jika Anda tidak memiliki izin untuk membuat daftar peran, Anda dapat memasukkan Nama Sumber Daya Amazon (ARN) peran yang ingin Anda gunakan. ii. Lihat peran layanan dengan memilih tautan eksternal Lihat di IAM. Jika tidak ada peran layanan yang ada, opsi untuk Menggunakan peran layanan yang ada tidak tersedia. Secara default, AWS Clean Rooms tidak mencoba memperbarui kebijakan peran yang ada untuk menambahkan izin yang diperlukan.

Note

- AWS Clean Rooms memerlukan izin untuk melakukan kueri sesuai dengan aturan analisis. Untuk informasi selengkapnya tentang izin AWS Clean Rooms, lihat [AWS kebijakan terkelola untuk AWS Clean Rooms](#).
- Jika peran tidak memiliki izin yang memadai AWS Clean Rooms, Anda akan menerima pesan galat yang menyatakan bahwa peran tersebut tidak memiliki izin yang memadai untuk peran tersebut. AWS Clean Rooms Kebijakan peran harus ditambahkan sebelum melanjutkan.

- Jika Anda tidak dapat mengubah kebijakan peran, Anda akan menerima pesan galat yang menyatakan bahwa AWS Clean Rooms tidak dapat menemukan kebijakan untuk peran layanan.

2. Untuk hasil Job,

Example

Misalnya: **s3://bucket/prefix**

- a. Pilih kotak centang Setel pengaturan default untuk pekerjaan, lalu tentukan tujuan Hasil di Amazon S3 dengan memasukkan tujuan S3 atau pilih Jelajahi S3 untuk memilih dari daftar bucket S3 yang tersedia.
 - b. Tentukan izin akses Layanan dengan memilih nama peran Layanan yang ada dari daftar tarik-turun.
- ## 3. Untuk pengaturan Log, pilih salah satu opsi berikut untuk penyimpanan Log di Amazon CloudWatch Logs:

Note

Bagian Pengaturan Log muncul jika Anda memilih untuk mengaktifkan pencatatan Kueri.

- a. Pilih Aktifkan dan log kueri yang relevan dengan Anda akan disimpan di akun Amazon CloudWatch Logs Anda.

Setiap anggota hanya dapat menerima log untuk kueri yang mereka mulai atau yang berisi data mereka.

Anggota yang dapat menerima hasil juga menerima log untuk semua kueri yang dijalankan dalam kolaborasi, meskipun data mereka tidak diakses dalam kueri.

Di bawah Jenis log yang didukung, pilih dari jenis log yang dipilih pembuat kolaborasi untuk didukung:

Di bawah Jenis log yang didukung, kotak centang Log Kueri dan log Job diaktifkan secara default.

 Note

Setelah Anda mengaktifkan pencatatan Analisis, penyimpanan log dapat diatur beberapa menit dan mulai menerima log di Amazon CloudWatch Logs. Selama periode singkat ini, anggota yang dapat melakukan kueri mungkin menjalankan kueri yang sebenarnya tidak mengirim log.

- b. Pilih Matikan dan log kueri yang relevan dengan Anda tidak akan disimpan di akun Amazon CloudWatch Logs Anda.
4. Jika Anda ingin mengaktifkan tag Keanggotaan untuk sumber daya keanggotaan, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.
5. Jika Anda adalah anggota yang membayar untuk perhitungan Query, atau Job compute, atau keduanya, tunjukkan penerimaan Anda dengan memilih kotak centang Saya setuju untuk membayar biaya komputasi dalam kolaborasi ini.

 Note

Anda harus memilih kotak centang ini untuk melanjutkan.
Untuk informasi selengkapnya tentang cara penghitungan harga, lihat [Harga untuk AWS Clean Rooms](#).

Jika Anda adalah [anggota yang membayar biaya komputasi kueri](#) tetapi bukan [anggota yang dapat melakukan kueri](#), disarankan agar Anda menggunakan AWS Budgets untuk mengonfigurasi anggaran AWS Clean Rooms dan menerima pemberitahuan setelah anggaran maksimum tercapai. Untuk informasi selengkapnya tentang menyiapkan anggaran, lihat [Mengelola biaya Anda AWS Budgets](#) di Panduan AWS Cost Management Pengguna. Untuk informasi selengkapnya tentang mengatur notifikasi, lihat [Membuat topik Amazon SNS untuk pemberitahuan anggaran](#) di AWS Cost Management Panduan Pengguna. Jika anggaran maksimum telah tercapai, Anda dapat menghubungi anggota yang dapat menjalankan kueri atau [meninggalkan kolaborasi](#). Jika Anda meninggalkan kolaborasi, tidak ada lagi kueri yang diizinkan untuk dijalankan, dan oleh karena itu Anda tidak akan lagi ditagih untuk biaya komputasi kueri.

6. Pilih Berikutnya.

Kolaborasi dan keanggotaan Anda dibuat.

Status Anda dalam kolaborasi aktif.

No, I will create a membership later

1. Pilih Berikutnya.

Hanya kolaborasi yang dibuat.

Status Anda dalam kolaborasi tidak aktif.

8. Untuk Langkah 5: Tinjau dan buat, lakukan hal berikut:

- a. Tinjau pilihan yang Anda buat untuk langkah-langkah sebelumnya dan edit jika perlu.
- b. Pilih salah satu opsi.

Jika Anda memilih untuk...	Kemudian pilih...
Buat keanggotaan dengan kolaborasi (Ya, bergabung dengan membuat keanggotaan sekarang)	Buat kolaborasi dan keanggotaan
Buat kolaborasi, dan bukan untuk membuat keanggotaan saat ini (Tidak, saya akan membuat keanggotaan nanti)	Buat kolaborasi

Setelah kolaborasi Anda berhasil dibuat, Anda dapat melihat halaman detail kolaborasi di bawah Kolaborasi.

Anda sekarang siap untuk:

- [Siapkan tabel data Anda untuk dianalisis AWS Clean Rooms](#). (Opsional jika Anda ingin menganalisis data peristiwa Anda sendiri atau jika Anda ingin menanyakan data identitas.)
- [Kaitkan tabel yang dikonfigurasi dengan kolaborasi Anda](#). (Opsional jika Anda ingin menganalisis data acara Anda sendiri.)
- [Tambahkan aturan analisis untuk tabel yang dikonfigurasi](#). (Opsional jika Anda ingin menganalisis data acara Anda sendiri.)

- [Buat keanggotaan dan bergabunglah dengan kolaborasi](#). (Opsional jika Anda sudah membuat keanggotaan.)
- [Undang anggota untuk bergabung dengan kolaborasi](#).

Membuat kolaborasi untuk pemodelan ML

Dalam prosedur ini, Anda sebagai [pembuat kolaborasi](#) melakukan tugas-tugas berikut:

- [Buat kolaborasi](#).
- Undang satu atau lebih [anggota](#) ke [kolaborasi](#).
- Menetapkan kemampuan untuk anggota, seperti
 - [Anggota yang dapat menanyakan](#)
 - [Anggota yang dapat menerima hasil](#)
 - Anggota yang dapat menerima output dari model terlatih
 - Anggota yang dapat mengeluarkan dari inferensi model

Jika pembuat kolaborasi juga anggota yang dapat menerima hasil, mereka menentukan tujuan dan format hasil. Mereka juga menyediakan peran layanan Amazon Resource Name (ARN) untuk menulis hasil ke tujuan hasil.

- Konfigurasi [anggota mana yang bertanggung jawab untuk membayar biaya komputasi, pelatihan model, dan biaya inferensi model dalam kolaborasi](#).

Sebelum Anda mulai, pastikan Anda telah menyelesaikan prasyarat berikut:

- Anda memiliki nama dan Akun AWS ID untuk setiap anggota yang ingin Anda undang ke kolaborasi.
- Anda memiliki izin untuk membagikan nama dan Akun AWS ID untuk setiap anggota dengan semua anggota kolaborasi.

Note

Anda tidak dapat menambahkan lebih banyak anggota setelah Anda membuat kolaborasi.

Untuk informasi tentang cara membuat kolaborasi menggunakan AWS SDKs, lihat [Referensi AWS Clean Rooms API](#).

Untuk membuat kolaborasi untuk pemodelan ML

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Akun AWS yang akan berfungsi sebagai pembuat kolaborasi.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Di sudut kanan atas, pilih Buat kolaborasi.
4. Untuk Langkah 1: Tentukan kolaborasi, lakukan hal berikut:

- a. Untuk Detail, masukkan Nama dan Deskripsi kolaborasi.

Informasi ini akan terlihat oleh anggota kolaborasi yang diundang untuk berpartisipasi dalam kolaborasi. Nama dan Deskripsi membantu mereka memahami apa yang dimaksud dengan kolaborasi.

- b. Untuk Anggota:

- i. Untuk Anggota 1: Anda, masukkan nama tampilan Anggota sesuai keinginan untuk kolaborasi.

 Note

Akun AWS ID Anda disertakan secara otomatis untuk Akun AWS ID Anggota.

- ii. Untuk Anggota 2, masukkan nama tampilan Anggota dan Akun AWS ID Anggota untuk anggota yang ingin Anda undang ke kolaborasi.

Nama tampilan Anggota dan Akun AWS ID Anggota akan terlihat oleh semua orang yang diundang ke kolaborasi. Setelah Anda memasukkan dan menyimpan nilai untuk bidang ini, Anda tidak dapat mengeditnya.

 Note

Anda harus memberi tahu anggota kolaborasi bahwa Akun AWS ID Anggota dan nama tampilan Anggota mereka akan terlihat oleh semua kolaborator yang diundang dan aktif dalam kolaborasi.

- iii. Jika Anda ingin menambahkan anggota lain, pilih Tambahkan anggota lain. Kemudian masukkan nama tampilan Anggota dan Akun AWS ID Anggota untuk setiap anggota yang dapat menyumbangkan data yang ingin Anda undang ke kolaborasi.
- c. Jika Anda ingin mengaktifkan Pencatatan analisis, pilih kotak centang Aktifkan pencatatan analisis, lalu di bawah Jenis log yang didukung, pilih Log dari kueri.
- d. Jika Anda ingin mengaktifkan Pemantauan terperinci, pilih kotak centang Aktifkan pemantauan terperinci.

Pelari analisis dan pembayar yang dikonfigurasi dapat memilih untuk mengaktifkan metrik terperinci saat mereka membuat keanggotaan mereka. Jika diaktifkan, metrik pemantauan terperinci akan dipublikasikan CloudWatch untuk pemantauan operasional kolaborasi, termasuk kinerja kueri dan pemanfaatan sumber daya. Metrik ini akan tersedia untuk pelari analisis dan pembayar yang dikonfigurasi di masing-masing s. Akun AWS

Untuk informasi selengkapnya tentang CloudWatch harga, lihat [CloudWatch Harga](#).

- e. Di bawah Wilayah hasil kueri yang diizinkan, pilih satu atau Wilayah AWS beberapa tempat Anda ingin mengirim hasil kueri.

Secara default, hanya Wilayah saat ini (seperti N. Virginia us-east-1) yang dipilih.

 Important

Saat Anda mengaktifkan pengiriman hasil kueri lintas wilayah, hasil Anda dapat diproses dan disimpan di luar Wilayah sumber.

Untuk informasi selengkapnya tentang Wilayah, lihat [Wilayah dan Titik Akhir](#) di. Referensi Umum AWS

- f. (Opsional) Kelola akses ke data Anda dengan Persetujuan permintaan perubahan otomatis dengan mengonfigurasi pengaturan mana yang dapat diubah secara otomatis tanpa persetujuan manual untuk permintaan perubahan. Secara default, beberapa pengaturan hanya dapat diubah dengan mengirimkan permintaan perubahan yang harus disetujui oleh semua anggota sebelum dapat diterapkan.
 - Berikan kemampuan anggota — Pilih kemampuan yang dapat diberikan kepada anggota kolaborasi tanpa persetujuan manual. Anggota selalu dapat menyumbangkan data.
 - Pilih kemampuan:

- Kontribusi data (selalu diaktifkan)
- Menerima hasil
- Persetujuan otomatis anggota baru dengan kemampuan ini — Jika diizinkan, setiap anggota yang ditambahkan dengan kemampuan yang dipilih di atas akan langsung bergabung dengan kolaborasi. Anggota yang ditambahkan dengan kemampuan lain masih memerlukan persetujuan manual untuk bergabung.
- Kemampuan yang dapat dicabut secara otomatis - Pilih kemampuan yang dapat dicabut tanpa persetujuan manual. Anggota selalu dapat menyumbangkan data.
 - Pilih kemampuan:
 - Kontribusi data (selalu diaktifkan)
 - Menerima hasil

Jika Anda memilih opsi ini, Anda dapat melacak semua modifikasi konfigurasi kolaborasi melalui riwayat Ubah permintaan, yang terletak di tab Detail di halaman detail kolaborasi.

g. (Opsional) Jika Anda ingin mengaktifkan kemampuan komputasi kriptografi, pilih kotak centang Aktifkan komputasi kriptografi.

i. Pilih parameter cakupan kriptografi berikut:

- Izinkan plaintext kolom

Pilih Tidak jika Anda memerlukan tabel terenkripsi sepenuhnya.

Pilih Ya jika Anda ingin cleartext kolom diizinkan dalam tabel terenkripsi.

Untuk menjalankan SUM atau AVG pada kolom tertentu, kolom harus masukcleartext.

- Pertahankan NULL nilai

Pilih Tidak jika Anda tidak ingin mempertahankan NULL nilai. NULLnilai tidak akan muncul seperti NULL dalam tabel terenkripsi.

Pilih Ya jika Anda ingin mempertahankan NULL nilai. NULLnilai akan muncul seperti NULL dalam tabel terenkripsi.

ii. Pilih parameter Sidik Jari berikut:

- Izinkan duplikat

Pilih Tidak jika Anda tidak ingin entri duplikat diizinkan dalam kolom. fingerprint

Pilih Ya jika Anda ingin entri duplikat diizinkan dalam kolom. fingerprint

- Izinkan JOIN kolom dengan nama yang berbeda

Pilih Tidak jika Anda tidak ingin bergabung dengan fingerprint kolom dengan nama yang berbeda.

Pilih Ya jika Anda ingin bergabung dengan fingerprint kolom dengan nama yang berbeda.

Untuk informasi lebih lanjut tentang parameter komputasi kriptografi, lihat [Parameter komputasi kriptografi](#).

Untuk informasi selengkapnya tentang cara mengenkripsi data Anda untuk digunakan AWS Clean Rooms, lihat [Mempersiapkan tabel data terenkripsi dengan Cryptographic Computing untuk Clean Rooms](#).

 Note

Verifikasi konfigurasi ini dengan hati-hati sebelum menyelesaikan langkah berikutnya. Setelah membuat kolaborasi, Anda hanya dapat mengedit nama kolaborasi, deskripsi, dan apakah log disimpan di Amazon CloudWatch Logs.

- h. Jika Anda ingin mengaktifkan Tag untuk sumber kolaborasi, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.
 - i. Pilih Berikutnya.
5. Untuk Langkah 2: Tentukan kemampuan anggota,
 - a. Untuk Analisis menggunakan kueri dan pekerjaan, di bawah Jenis analisis yang didukung, biarkan kotak centang Kueri dipilih.
 - b. Untuk kueri Jalankan, pilih anggota yang akan memulai pelatihan model
 - c. Untuk Menerima hasil dari analisis, pilih satu atau beberapa anggota yang akan menerima hasil kueri.
 - d. Untuk pemodelan ML menggunakan alur kerja yang dibuat khusus,

- i. Untuk Menerima output dari model terlatih, pilih anggota yang akan menerima hasil model terlatih, termasuk artefak model dan metrik.
 - ii. Untuk Menerima output dari inferensi model, pilih anggota yang akan menerima hasil inferensi model.
- e. Lihat kemampuan anggota di bawah resolusi ID menggunakan Resolusi Entitas AWS.
6. Untuk Langkah 3: Konfigurasi pembayaran,
 - a. Di bawah Analisis menggunakan kueri, untuk Bayar untuk kueri, lakukan salah satu tindakan berikut:
 - Agar anggota yang sama membayar dan menjalankan kueri, pilih anggota yang sama yang Anda pilih untuk kueri Jalankan.
 - Untuk meminta anggota yang berbeda membayar biaya kueri, pilih akun anggota Anda.
 - b. Untuk pemodelan ML menggunakan alur kerja yang dibuat khusus,
 - Pilih anggota yang akan membayar untuk pelatihan model.
 - c. Pilih anggota yang akan membayar untuk pekerjaan inferensi.
 - d. Untuk Pay for lookalike modeling, tidak diperlukan tindakan. Pencipta model mirip yang dikonfigurasi adalah anggota yang akan membayar untuk pemodelan mirip.
 - e. (Opsional) Pilih anggota yang akan Membayar untuk pembuatan data Sintetis.
 - f. Untuk resolusi ID dengan Resolusi Entitas AWS, tidak diperlukan tindakan. Pembuat tabel pemetaan ID adalah anggota yang akan membayar tabel pemetaan ID.
7. Pilih Berikutnya.
8. Untuk Langkah 4: Konfigurasi keanggotaan, di bawah keanggotaan Kolaborasi, pilih salah satu opsi berikut:

Yes, join by creating membership now

1. Untuk pengaturan Hasil default, untuk pengaturan hasil Kueri, jika Anda adalah anggota yang dapat Menerima hasil,
 - a. Pilih kotak centang Setel pengaturan default untuk kueri.
 - b. Untuk tujuan Hasil di Amazon S3, masukkan tujuan Amazon S3 atau pilih Jelajahi S3 untuk memilih bucket S3.
 - c. Untuk format hasil kueri, pilih CSV atau PARQUET.

- d. (Hanya percikan) Untuk file Hasil, pilih Multiple atau Single.
- e. (Opsional) Jika Anda ingin mengirimkan kueri yang memakan waktu hingga 24 jam ke tujuan S3 Anda, pilih kotak centang Tambahkan peran layanan untuk mendukung kueri yang membutuhkan waktu hingga 24 jam untuk diselesaikan.

Kueri besar yang membutuhkan waktu hingga 24 jam untuk diselesaikan akan dikirimkan ke tujuan S3 Anda.

Jika Anda tidak memilih kotak centang, hanya kueri yang selesai dalam waktu 12 jam yang akan dikirimkan ke lokasi S3 Anda.

- f. Tentukan izin akses Layanan dengan memilih Buat dan gunakan peran layanan baru atau Gunakan peran layanan yang ada.

Jika Anda memilih untuk...	Lalu...
Membuat dan menggunakan peran layanan baru	• AWS Clean Rooms membuat peran layanan dengan kebijakan yang diperlukan untuk tabel ini. • Nama peran Layanan default adalah <code>cleanrooms-result-receiver-<timestamp></code> • Anda harus memiliki izin untuk membuat peran dan melampirkan kebijakan.

Jika Anda memilih untuk...	Lalu...
Gunakan peran layanan yang ada	i. Pilih nama peran layanan yang ada dari daftar tarik-turun. Daftar peran ditampilkan jika Anda memiliki izin untuk membuat daftar peran. Jika Anda tidak memiliki izin untuk membuat daftar peran, Anda dapat memasukkan Nama Sumber Daya Amazon (ARN) peran yang ingin Anda gunakan. ii. Lihat peran layanan dengan memilih tautan eksternal Lihat di IAM. Jika tidak ada peran layanan yang ada, opsi untuk Menggunakan peran layanan yang ada tidak tersedia. Secara default, AWS Clean Rooms tidak mencoba memperbarui kebijakan peran yang ada untuk menambahkan izin yang diperlukan.

Note

- AWS Clean Rooms memerlukan izin untuk melakukan kueri sesuai dengan aturan analisis. Untuk informasi selengkapnya tentang izin AWS Clean Rooms, lihat [AWS kebijakan terkelola untuk AWS Clean Rooms](#).
- Jika peran tidak memiliki izin yang memadai AWS Clean Rooms, Anda akan menerima pesan galat yang menyatakan bahwa peran tersebut tidak memiliki izin yang memadai untuk peran tersebut. AWS Clean Rooms Kebijakan peran harus ditambahkan sebelum melanjutkan.

- Jika Anda tidak dapat mengubah kebijakan peran, Anda akan menerima pesan galat yang menyatakan bahwa AWS Clean Rooms tidak dapat menemukan kebijakan untuk peran layanan.

2. Untuk konfigurasi ML,

- a. Pilih kotak centang Buat konfigurasi ML, lalu tentukan tujuan keluaran Model di Amazon S3 dengan memasukkan tujuan S3 atau pilih Jelajahi S3 untuk memilih dari daftar bucket S3 yang tersedia.
 - b. Tentukan izin akses Layanan dengan memilih Membuat dan menggunakan peran layanan baru atau Menggunakan peran layanan yang ada.
 - c. Jika bucket S3 dienkripsi, pilih kotak centang Bucket tujuan dienkripsi dengan kunci KMS, lalu masukkan AWS KMS key atau pilih Buat untuk membuat kunci KMS baru. AWS KMS key
3. Jika Anda ingin mengaktifkan tag Keanggotaan untuk sumber daya keanggotaan, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.
4. Jika Anda adalah anggota yang membayar untuk komputasi Kueri, tunjukkan penerimaan Anda dengan memilih kotak centang Saya setuju untuk membayar biaya komputasi dalam kolaborasi ini.

 Note

Anda harus memilih kotak centang ini untuk melanjutkan.
Untuk informasi selengkapnya tentang cara penghitungan harga, lihat [Harga untuk AWS Clean Rooms](#).

Jika Anda adalah [anggota yang membayar biaya komputasi kueri](#) tetapi bukan [anggota yang dapat melakukan kueri](#), disarankan agar Anda menggunakan AWS Budgets untuk mengonfigurasi anggaran AWS Clean Rooms dan menerima pemberitahuan setelah anggaran maksimum tercapai. Untuk informasi selengkapnya tentang menyiapkan anggaran, lihat [Mengelola biaya Anda AWS Budgets](#) di Panduan AWS Cost Management Pengguna. Untuk informasi selengkapnya tentang mengatur notifikasi, lihat [Membuat topik Amazon SNS untuk pemberitahuan anggaran](#) di AWS Cost Management Panduan Pengguna. Jika anggaran maksimum telah tercapai, Anda dapat menghubungi anggota yang dapat menjalankan kueri atau [meninggalkan kolaborasi](#). Jika Anda meninggalkan

kolaborasi, tidak ada lagi kueri yang diizinkan untuk dijalankan, dan oleh karena itu Anda tidak akan lagi ditagih untuk biaya komputasi kueri.

5. Pilih Berikutnya.

Kolaborasi dan keanggotaan Anda dibuat.

Status Anda dalam kolaborasi aktif.

No, I will create a membership later

1. Pilih Berikutnya.

Hanya kolaborasi yang dibuat.

Status Anda dalam kolaborasi tidak aktif.

9. Untuk Langkah 5: Tinjau dan buat, lakukan hal berikut:

- Tinjau pilihan yang Anda buat untuk langkah-langkah sebelumnya dan edit jika perlu.
- Pilih salah satu opsi.

Jika Anda memilih untuk...	Kemudian pilih...
Buat keanggotaan dengan kolaborasi (Ya, bergabung dengan membuat keanggotaan sekarang)	Buat kolaborasi dan keanggotaan
Buat kolaborasi, dan bukan untuk membuat keanggotaan saat ini (Tidak, saya akan membuat keanggotaan nanti)	Buat kolaborasi

Membuat keanggotaan dan bergabung dengan kolaborasi

Keanggotaan adalah sumber daya yang dibuat saat anggota bergabung dengan kolaborasi. AWS Clean Rooms

Anda dapat bergabung dengan kolaborasi sebagai

- [anggota yang dapat menanyakan](#)

- [anggota yang dapat menjalankan kueri dan pekerjaan](#)
- [anggota yang dapat menerima hasil](#) dari query atau pekerjaan
- [anggota membayar biaya komputasi kueri](#)
- [anggota membayar untuk pertanyaan dan pekerjaan](#)

Semua anggota dapat menyumbangkan data.

Untuk informasi tentang cara membuat keanggotaan dan bergabung dengan kolaborasi menggunakan AWS SDKs, lihat [Referensi AWS Clean Rooms API](#).

Dalam prosedur ini, anggota yang diundang [bergabung dengan kolaborasi dengan membuat sumber daya keanggotaan](#).

Jika anggota yang diundang adalah anggota yang dapat menerima hasil, mereka menentukan tujuan dan format hasil. Mereka juga menyediakan peran layanan ARN untuk menulis ke tujuan hasil.

Jika anggota yang diundang adalah anggota yang bertanggung jawab untuk membayar biaya komputasi, mereka menerima tanggung jawab pembayaran mereka sebelum bergabung dengan kolaborasi.

Untuk membuat keanggotaan dan bergabung dengan kolaborasi

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan anggota Anda Akun AWS.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pada tab Tersedia untuk bergabung, untuk Kolaborasi yang tersedia untuk bergabung, pilih Nama kolaborasi.
4. Pada halaman detail kolaborasi, di bagian Ikhtisar, lihat detail kolaborasi, termasuk detail anggota Anda dan daftar anggota lainnya.

Verifikasi bahwa Akun AWS IDs untuk setiap anggota kolaborasi adalah orang-orang yang ingin Anda ikuti untuk kolaborasi.

5. Pilih Buat keanggotaan.
6. Pada halaman Buat keanggotaan, di Ikhtisar, lihat nama Kolaborasi, Deskripsi kolaborasi, Akun AWS ID pembuat Kolaborasi, detail anggota Anda, dan Akun AWS ID anggota yang akan Membayar kueri.

7. Jika pembuat kolaborasi telah memilih untuk mengaktifkan pencatatan Analisis, pilih salah satu opsi berikut untuk penyimpanan Log di Amazon CloudWatch Logs:

Jika Anda memilih...	Lalu...
Nyalakan	Log yang relevan untuk Anda disimpan di Amazon CloudWatch Logs. Setiap anggota hanya dapat menerima log untuk kueri yang mereka mulai atau yang berisi data mereka. Anggota yang dapat menerima hasil juga menerima log untuk semua analisis yang dijalankan dalam kolaborasi, bahkan jika data mereka tidak diakses dalam analisis. Di bawah Jenis log yang didukung, pilih dari jenis log yang dipilih pembuat kolaborasi untuk didukung: 1. Jika Anda ingin menerima log yang dihasilkan dari kueri SQL, pilih kotak centang Log dari kueri. 2. Jika Anda ingin menerima log yang dihasilkan dari pekerjaan menggunakan PySpark, pilih kotak centang Log dari pekerjaan.
Matikan	Log kueri yang relevan dengan Anda tidak disimpan di akun Amazon CloudWatch Logs Anda.

 Note

Setelah Anda mengaktifkan pencatatan Analisis, penyimpanan log dapat diatur beberapa menit dan mulai menerima log di Amazon CloudWatch Logs. Selama periode singkat ini,

anggota yang dapat melakukan kueri mungkin menjalankan kueri yang sebenarnya tidak mengirim log.

8. Jika pembuat kolaborasi mengaktifkan Pemantauan terperinci untuk kolaborasi ini, pilih apakah Anda ingin menerima metrik observabilitas terperinci di akun Anda CloudWatch .

Untuk pemantauan terperinci:

Opsi pemantauan terperinci

Opsi	Deskripsi
Nyalakan	AWS Clean Rooms akan mempublikasikan metrik pemantauan terperinci CloudWatch untuk kolaborasi ini di akun Anda. Anda dapat menggunakan metrik ini untuk pemantauan operasional, termasuk kinerja kueri dan pemanfaatan sumber daya. CloudWatchBiaya tambahan berlaku. Untuk informasi selengkapnya, silakan lihat Harga CloudWatch .
Matikan	Tidak ada metrik terperinci yang akan diekspor ke akun Anda CloudWatch . Anggota lain masih dapat melihat metrik pemantauan terperinci di akun mereka sendiri jika mereka telah mengaktifkan opsi ini.

 Note

Metrik pemantauan terperinci hanya tersedia untuk anggota yang dapat menjalankan kueri (pelari analisis) dan anggota yang dikonfigurasi sebagai pembayar untuk kolaborasi.

9. Jika kemampuan anggota Anda termasuk Menerima hasil, untuk pengaturan Hasil default:
 - a. Untuk hasil Kueri, pilih kotak centang Setel pengaturan default untuk kueri, lalu tentukan tujuan Hasil di Amazon S3 dengan memasukkan tujuan S3 atau pilih Jelajahi S3 untuk memilih dari daftar bucket S3 yang tersedia.

Example

Misalnya: **s3://bucket/prefix**

- i. Untuk format Hasil, pilih CSV atau PARQUET.
- ii. (Hanya percikan) Untuk file Hasil, pilih Multiple atau Single.
- iii. (Opsional) Untuk akses Layanan, jika Anda ingin mengirimkan kueri yang memakan waktu hingga 24 jam ke tujuan S3 Anda, pilih kotak centang Tambahkan peran layanan untuk mendukung kueri yang membutuhkan waktu hingga 24 jam untuk diselesaikan.

Kueri besar yang membutuhkan waktu hingga 24 jam untuk diselesaikan akan dikirimkan ke tujuan S3 Anda.

Jika Anda tidak memilih kotak centang, hanya kueri yang selesai dalam waktu 12 jam yang akan dikirimkan ke lokasi S3 Anda.

Note

Anda harus memilih peran layanan yang ada atau memiliki izin untuk membuat yang baru. Untuk informasi selengkapnya, lihat [Buat peran layanan untuk menerima hasil](#).

- iv. Tentukan izin akses Layanan dengan memilih Buat dan gunakan peran layanan baru atau Gunakan peran layanan yang ada.

Create and use a new service role

- AWS Clean Rooms membuat peran layanan dengan kebijakan yang diperlukan untuk tabel ini.
- Nama peran Layanan default adalah `cleanrooms-result-receiver-
<timestamp>`
- Anda harus memiliki izin untuk membuat peran dan melampirkan kebijakan.

Use an existing service role

1. Pilih nama peran layanan yang ada dari daftar tarik-turun.

Daftar peran ditampilkan jika Anda memiliki izin untuk membuat daftar peran.

Jika Anda tidak memiliki izin untuk membuat daftar peran, Anda dapat memasukkan Nama Sumber Daya Amazon (ARN) peran yang ingin Anda gunakan.

2. Lihat peran layanan dengan memilih tautan eksternal Lihat di IAM.

Jika tidak ada peran layanan yang ada, opsi untuk Menggunakan peran layanan yang ada tidak tersedia.

Secara default, AWS Clean Rooms tidak mencoba memperbarui kebijakan peran yang ada untuk menambahkan izin yang diperlukan.

 Note

- AWS Clean Rooms memerlukan izin untuk melakukan kueri sesuai dengan aturan analisis. Untuk informasi selengkapnya tentang izin AWS Clean Rooms, lihat [AWS kebijakan terkelola untuk AWS Clean Rooms](#).
- Jika peran tidak memiliki izin yang memadai AWS Clean Rooms, Anda akan menerima pesan galat yang menyatakan bahwa peran tersebut tidak memiliki izin yang memadai untuk peran tersebut. AWS Clean Rooms Kebijakan peran harus ditambahkan sebelum melanjutkan.
- Jika Anda tidak dapat mengubah kebijakan peran, Anda akan menerima pesan galat yang menyatakan bahwa AWS Clean Rooms tidak dapat menemukan kebijakan untuk peran layanan.

- b. Untuk hasil Job, pilih kotak centang Setel pengaturan default untuk pekerjaan, lalu tentukan tujuan Hasil di Amazon S3 dengan memasukkan tujuan S3 atau pilih Jelajahi S3 untuk memilih dari daftar bucket S3 yang tersedia.

Example

Misalnya: **s3://bucket/prefix**

- Tentukan izin akses Layanan dengan memilih nama peran Layanan yang ada dari daftar tarik-turun.

10. Jika Anda ingin mengaktifkan Tag untuk sumber daya keanggotaan, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.

11. Jika pembuat kolaborasi telah menunjuk Anda sebagai anggota yang akan Membayar kueri atau Bayar untuk pertanyaan dan pekerjaan, tunjukkan penerimaan Anda dengan memilih kotak centang Saya setuju untuk membayar biaya komputasi dalam kolaborasi ini.

 Note

Anda harus memilih kotak centang ini untuk melanjutkan.

Untuk informasi selengkapnya tentang cara penghitungan harga, lihat [Harga untuk AWS Clean Rooms](#).

Jika Anda adalah [anggota yang membayar biaya komputasi kueri](#) atau [anggota yang membayar pertanyaan dan biaya komputasi pekerjaan](#) tetapi bukan [anggota yang dapat menanyakan](#), disarankan agar Anda menggunakan AWS Budgets untuk mengonfigurasi anggaran AWS Clean Rooms dan menerima pemberitahuan setelah anggaran maksimum tercapai. Untuk informasi selengkapnya tentang menyiapkan anggaran, lihat [Mengelola biaya Anda AWS Budgets](#) di Panduan AWS Cost Management Pengguna. Untuk informasi selengkapnya tentang mengatur notifikasi, lihat [Membuat topik Amazon SNS untuk pemberitahuan anggaran](#) di AWS Cost Management Panduan Pengguna. Jika anggaran maksimum telah tercapai, Anda dapat menghubungi anggota yang dapat menjalankan kueri dan pekerjaan atau [meninggalkan kolaborasi](#). Jika Anda meninggalkan kolaborasi, tidak ada lagi kueri yang diizinkan untuk dijalankan, dan oleh karena itu Anda tidak akan lagi ditagih untuk biaya komputasi kueri.

12. Jika Anda yakin ingin membuat keanggotaan dan bergabung dengan kolaborasi, pilih Buat keanggotaan.

Anda diberi akses baca ke metadata kolaborasi. Ini termasuk informasi seperti nama tampilan dan deskripsi kolaborasi, di samping semua nama dan Akun AWS IDs anggota lainnya.

Anda sekarang siap untuk:

- [Siapkan tabel data Anda untuk ditanyakan](#). AWS Clean Rooms (Opsional jika Anda ingin menanyakan data acara Anda sendiri atau jika Anda ingin menanyakan data identitas.)
- [Kaitkan tabel yang dikonfigurasi ke kolaborasi Anda](#) - jika Anda ingin menanyakan data peristiwa.
- [Tambahkan aturan analisis untuk tabel yang dikonfigurasi](#) - jika Anda ingin menanyakan data peristiwa.

- [Buat dan kaitkan namespace ID baru](#) — jika Anda ingin membuat tabel pemetaan ID untuk menanyakan data identitas.

Untuk informasi tentang cara meninggalkan kolaborasi, lihat [Meninggalkan Kolaborasi](#).

Mengedit kolaborasi

Sebagai pembuat kolaborasi, Anda dapat mengedit berbagai bagian kolaborasi.

Untuk informasi tentang cara mengedit kolaborasi menggunakan AWS SDKs, lihat [Referensi API AWS Clean Rooms](#).

Topik

- [Edit nama dan deskripsi kolaborasi](#)
- [Perbarui mesin analitik kolaborasi](#)
- [Matikan penyimpanan log](#)
- [Mengedit setelan log kolaborasi](#)
- [Edit tag kolaborasi](#)
- [Edit tag keanggotaan](#)
- [Menambahkan anggota baru](#)
- [Edit kemampuan anggota yang ada](#)
- [Edit pengaturan persetujuan otomatis kolaborasi](#)
- [Mengedit tag tabel terkait](#)
- [Edit tag templat analisis](#)
- [Edit tag kebijakan privasi diferensial](#)

Edit nama dan deskripsi kolaborasi

Setelah Anda membuat kolaborasi, Anda hanya dapat mengedit nama dan deskripsi kolaborasi.

Untuk mengedit nama dan deskripsi kolaborasi

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang Anda buat.

4. Pada halaman detail kolaborasi, pilih Tindakan, lalu pilih Edit kolaborasi.
5. Pada halaman Edit kolaborasi, untuk Detail, edit Nama dan Deskripsi kolaborasi.
6. Pilih Simpan perubahan.

Perbarui mesin analitik kolaborasi

Setelah Anda membuat kolaborasi, Anda dapat mengubah mesin analitik dari AWS Clean Rooms SQL ke Spark.

Note

Mengubah mesin analitik dari AWS Clean Rooms SQL ke Spark dapat merusak alur kerja yang ada.

Untuk memperbarui mesin analitik kolaborasi

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang Anda buat.
4. Pada halaman detail kolaborasi, pilih Tindakan, lalu pilih Edit kolaborasi.
5. Pilih Simpan perubahan.

Matikan penyimpanan log

Jika Anda telah mengaktifkan pencatatan Analisis, Anda dapat mengedit apakah log analisis disimpan di akun Amazon CloudWatch Logs Anda.

Untuk mematikan penyimpanan log

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang mengaktifkan pencatatan analisis.
4. Pada halaman detail kolaborasi, pilih Tindakan, lalu pilih Matikan penyimpanan log.

 Note

Peringatan muncul, menunjukkan hal berikut:

- Kueri baru tidak akan lagi masuk ke CloudWatch akun Anda.
- Log yang ada akan dipertahankan sesuai dengan pengaturan retensi Anda saat ini.
- Jika Anda mengaktifkan kembali logging di masa depan, itu hanya akan berlaku untuk kueri yang dibuat setelah pengaktifan kembali.
- Perubahan ini hanya memengaruhi log Anda - pengaturan pencatatan anggota tim lainnya tetap tidak berubah.

5. Pilih Matikan.

Mengedit setelan log kolaborasi

Jika Anda telah mengaktifkan Pencatatan kueri, Anda dapat mengedit apakah log kueri disimpan di akun Amazon CloudWatch Logs Anda.

Untuk mengedit setelan log kolaborasi

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang Anda buat.
4. Pada halaman detail kolaborasi, lakukan salah satu hal berikut:
 - Pilih Tindakan, lalu pilih Pengaturan Edit log.
 - Pada tab Log, pilih Edit pengaturan log.
5. Pada modal pengaturan Edit log, untuk penyimpanan Log di Amazon CloudWatch Logs:
 - Jika Anda tidak ingin log yang relevan dengan Anda disimpan di akun Amazon CloudWatch Logs Anda, pilih Turn o.
 - Jika Anda ingin log yang relevan dengan Anda disimpan di akun Amazon CloudWatch Logs Anda, pilih Aktifkan.

Anda hanya dapat menerima log untuk kueri yang Anda mulai atau yang berisi data Anda.

Anggota yang dapat menerima hasil juga menerima log untuk semua kueri yang dijalankan dalam kolaborasi, meskipun data mereka tidak diakses dalam kueri.

1. Di bawah Jenis log yang didukung, pilih dari jenis log yang dipilih pembuat kolaborasi untuk didukung:

- Jika Anda ingin menerima log yang dihasilkan dari kueri SQL, pilih kotak centang Log dari kueri.
- Jika Anda ingin menerima log yang dihasilkan dari pekerjaan menggunakan PySpark, pilih kotak centang Log dari pekerjaan.

6. Pilih Simpan perubahan.

Note

Setelah Anda mengaktifkan logging, diperlukan beberapa menit untuk mengatur penyimpanan log dan mulai menerima log di Amazon CloudWatch Logs. Selama periode singkat ini, anggota yang dapat melakukan kueri mungkin menjalankan kueri yang sebenarnya tidak mengirim log.

Edit tag kolaborasi

Sebagai pembuat kolaborasi, setelah Anda membuat kolaborasi, Anda dapat mengelola tag pada sumber kolaborasi.

Untuk mengedit tag kolaborasi

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang Anda buat.
4. Pilih salah satu cara berikut:

Jika Anda...	Lalu...
Pencipta kolaborasi dan anggota kolaborasi	Pilih tab Detail.

Jika Anda...	Lalu...
Pencipta kolaborasi tetapi bukan anggota kolaborasi	Gulir ke bawah halaman ke bagian Tag.

5. Untuk detail Kolaborasi, pilih Kelola tag.
6. Pada halaman Kelola tag, Anda dapat melakukan hal berikut:
 - Untuk menghapus sebuah tag, pilih Hapus.
 - Untuk menambahkan tanda, pilih Tambahkan tanda baru.
 - Untuk menyimpan perubahan, pilih Simpan perubahan

Edit tag keanggotaan

Sebagai pembuat kolaborasi, setelah Anda membuat kolaborasi, Anda dapat mengelola tag pada sumber daya keanggotaan.

Untuk mengedit tag keanggotaan

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang Anda buat.
4. Pilih tab Detail.
5. Untuk detail Keanggotaan, pilih Kelola tag.
6. Pada halaman Kelola tag keanggotaan, Anda dapat melakukan hal berikut:
 - Untuk menghapus sebuah tag, pilih Hapus.
 - Untuk menambahkan tanda, pilih Tambahkan tanda baru.
 - Untuk menyimpan perubahan Anda, memilih Simpan perubahan.

Menambahkan anggota baru

Untuk informasi selengkapnya, lihat [Menambahkan anggota baru ke kolaborasi](#).

Edit kemampuan anggota yang ada

Untuk informasi selengkapnya, lihat [Perbarui kemampuan anggota yang ada](#).

Edit pengaturan persetujuan otomatis kolaborasi

Untuk informasi selengkapnya, lihat [Edit pengaturan persetujuan otomatis kolaborasi](#).

Mengedit tag tabel terkait

Sebagai pembuat kolaborasi, setelah Anda mengaitkan tabel ke kolaborasi, Anda dapat mengelola tag pada sumber daya tabel terkait.

Untuk mengedit tag tabel terkait

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang Anda buat.
4. Pilih tab Tabel.
5. Untuk Tabel yang terkait dengan Anda, pilih tabel.
6. Pada halaman detail tabel yang dikonfigurasi, untuk Tag, pilih Kelola tag.

Pada halaman Kelola tag, Anda dapat melakukan hal berikut:

- Untuk menghapus sebuah tag, pilih Hapus.
- Untuk menambahkan tanda, pilih Tambahkan tanda baru.
- Untuk menyimpan perubahan Anda, memilih Simpan perubahan.

Edit tag templat analisis

Sebagai pembuat kolaborasi, setelah Anda membuat kolaborasi, Anda dapat mengelola tag pada sumber daya templat analisis.

Untuk mengedit tag keanggotaan

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).

2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang Anda buat.
4. Pilih tab Template.
5. Pada bagian Template Analisis yang dibuat oleh Anda, pilih templat analisis.
6. Pada halaman detail tabel templat analisis, gulir ke bawah ke bagian Tag.
7. Pilih Kelola tanda.
8. Pada halaman Kelola tag, Anda dapat melakukan hal berikut:
 - Untuk menghapus sebuah tag, pilih Hapus.
 - Untuk menambahkan tanda, pilih Tambahkan tanda baru.
 - Untuk menyimpan perubahan Anda, memilih Simpan perubahan.

Edit tag kebijakan privasi diferensial

Sebagai pembuat kolaborasi, setelah Anda membuat kolaborasi, Anda dapat mengelola tag pada sumber daya templat analisis.

Untuk mengedit tag keanggotaan

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang berisi kebijakan privasi diferensial yang ingin Anda edit.
4. Pilih tab Tabel.
5. Pada tab Tabel, pilih Kelola tag.
6. Pada halaman Kelola tag, Anda dapat melakukan hal berikut:
 - Untuk menghapus sebuah tag, pilih Hapus.
 - Untuk menambahkan tanda, pilih Tambahkan tanda baru.
 - Untuk menyimpan perubahan Anda, memilih Simpan perubahan.

Ubah permintaan di AWS Clean Rooms

Permintaan perubahan memungkinkan Anda mengusulkan perubahan pada setelan kolaborasi yang ada untuk disetujui oleh anggota kolaborasi lainnya. Dengan permintaan perubahan, Anda dapat mengirimkan permintaan untuk menambahkan anggota baru, memperbarui kemampuan anggota yang ada, dan mengubah pengaturan persetujuan otomatis kolaborasi. Semua anggota kolaborasi harus menyetujui permintaan perubahan agar perubahan yang diusulkan berlaku.

Permintaan perubahan bersifat khusus untuk kolaborasi dan dapat dikirimkan oleh pembuat kolaborasi.

Anda dapat mengirimkan permintaan perubahan dengan cara berikut:

- Menambahkan anggota baru ke kolaborasi
- Perbarui kemampuan anggota yang ada
- Edit pengaturan persetujuan otomatis kolaborasi

Note

Anda harus menjadi pembuat kolaborasi untuk mengirimkan permintaan perubahan.

Menambahkan anggota baru ke kolaborasi

Untuk menambahkan anggota baru ke kolaborasi, Anda harus menjadi pencipta kolaborasi. Menambahkan anggota baru ke kolaborasi memerlukan persetujuan manual dari anggota kolaborasi yang ada, dan akan mengirimkan permintaan perubahan.

Anda dapat menambahkan anggota baru ke kolaborasi dengan mengikuti langkah-langkah berikut:

Untuk menambahkan anggota baru ke kolaborasi

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Roomskonsol](#) dengan anggota AndaAkun AWS.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi Anda untuk menavigasi ke detail kolaborasi Anda.
4. Pada halaman kolaborasi, navigasikan ke tab Anggota.
5. Dalam tabel Anggota, pilih Edit anggota.

6. Pilih Tambahkan anggota lain.
7. Masukkan informasi anggota baru:
 - Nama tampilan anggota
 - Akun AWSID Anggota
 - Tentukan apakah anggota dapat menerima hasil. Centang kotak untuk memberikan kemampuan anggota ini.
8. Pilih Simpan perubahan.
9. Konfirmasikan pengiriman permintaan perubahan Anda. Di modal konfirmasi, konfirmasikan perubahan dan pilih Kirim permintaan perubahan.

 Note

Jika jenis perubahan yang disetujui otomatis didukung dalam kolaborasi Anda, persetujuan manual atas permintaan perubahan mungkin tidak diperlukan. Anda dapat meninjau jenis perubahan mana yang tidak memerlukan permintaan perubahan di bagian “Ikhtisar” kolaborasi. Untuk informasi selengkapnya, lihat [the section called “Edit pengaturan persetujuan otomatis kolaborasi”](#).

Perbarui kemampuan anggota yang ada

Untuk memperbarui kemampuan anggota kolaborasi yang ada, Anda harus menjadi pencipta kolaborasi. Memperbarui kemampuan untuk anggota kolaborasi yang ada memerlukan persetujuan manual dari anggota yang ada, dan akan mengirimkan permintaan perubahan.

Kemampuan anggota yang dapat diperbarui adalah:

- Dapat menerima hasil

Anda dapat memperbarui kemampuan anggota untuk anggota kolaborasi yang ada dengan mengikuti langkah-langkah berikut:

Untuk memperbarui kemampuan anggota yang ada

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Roomskonsol](#) dengan anggota AndaAkun AWS.

2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi Anda untuk menavigasi ke detail kolaborasi Anda.
4. Pada halaman kolaborasi, navigasikan ke tab Anggota.
5. Dalam tabel Anggota, pilih Edit anggota.
6. Tentukan kemampuan anggota untuk berubah.
7. Pilih Simpan perubahan.
8. Konfirmasikan pengiriman permintaan perubahan Anda. Di modal konfirmasi, konfirmasikan perubahan dan pilih Kirim permintaan perubahan.

Note

Jika jenis perubahan yang disetujui otomatis didukung dalam kolaborasi Anda, persetujuan manual atas permintaan perubahan mungkin tidak diperlukan. Anda dapat meninjau jenis perubahan mana yang tidak memerlukan permintaan perubahan di bagian “Ikhtisar” kolaborasi. Untuk informasi selengkapnya, lihat [the section called “Edit pengaturan persetujuan otomatis kolaborasi”](#).

Edit pengaturan persetujuan otomatis kolaborasi

Untuk mengedit pengaturan kolaborasi untuk persetujuan otomatis, Anda harus menjadi pembuat kolaborasi dan mengirimkan permintaan perubahan untuk persetujuan oleh anggota kolaborasi lainnya.

Anda dapat mengedit pengaturan persetujuan otomatis untuk kolaborasi dengan mengikuti langkah-langkah berikut:

Untuk mengedit pengaturan persetujuan otomatis kolaborasi

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Roomskonsol](#) dengan anggota AndaAkun AWS.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi Anda untuk menavigasi ke detail kolaborasi Anda.
4. Pada halaman kolaborasi, pilih tombol Tindakan dan pilih Edit persetujuan otomatis.
5. Untuk memberikan kemampuan anggota kepada anggota yang ada tanpa menyetujui permintaan perubahan secara manual:

1. Arahkan ke bagian Grant member abilities.
2. Tentukan kemampuan anggota mana yang dapat diberikan secara otomatis.

 Note

Secara default, semua anggota kolaborasi dapat “Menyumbang data”.

3. (Opsional) Untuk memungkinkan anggota baru secara instan bergabung dengan kolaborasi dengan kemampuan yang ditentukan tanpa menyetujui permintaan perubahan secara manual, pilih Persetujuan otomatis anggota baru dengan kemampuan ini.
6. Untuk memungkinkan penghapusan kemampuan dari anggota yang ada tanpa permintaan perubahan:
 1. Arahkan ke bagian Kemampuan yang dapat dicabut secara otomatis.
 2. Tentukan kemampuan anggota mana yang dapat dihapus secara otomatis.
7. Pilih Simpan perubahan.
8. Konfirmasikan pengiriman permintaan perubahan Anda. Di modal konfirmasi, konfirmasikan perubahan dan pilih Kirim permintaan perubahan.

Menghapus kolaborasi

Sebagai pembuat kolaborasi, Anda dapat menghapus kolaborasi yang Anda buat.

 Note

Saat menghapus kolaborasi, Anda dan semua anggota tidak dapat menjalankan kueri, menerima hasil, atau menyumbangkan data. Setiap anggota kolaborasi terus memiliki akses ke data mereka sendiri sebagai bagian dari keanggotaan mereka.

Untuk menghapus kolaborasi

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang ingin Anda hapus.

4. Di bawah Tindakan, pilih Hapus kolaborasi.
5. Konfirmasikan penghapusan dan kemudian pilih Hapus.

Melihat kolaborasi

Sebagai pembuat kolaborasi, Anda dapat melihat semua kolaborasi yang Anda buat.

Untuk melihat kolaborasi

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pada halaman Kolaborasi, di bawah Terakhir digunakan, lihat 5 kolaborasi terakhir yang digunakan.
4. Pada tab Dengan keanggotaan aktif, lihat daftar Kolaborasi dengan keanggotaan aktif.

Anda dapat mengurutkan berdasarkan Nama, tanggal yang dibuat Keanggotaan, dan rincian anggota Anda.

Anda dapat menggunakan bilah Pencarian untuk mencari kolaborasi.

5. Pada tab Tersedia untuk bergabung, lihat daftar Kolaborasi yang tersedia untuk bergabung.
6. Pada tab Tidak lagi tersedia, lihat daftar kolaborasi yang dihapus dan Keanggotaan untuk kolaborasi yang tidak lagi tersedia (keanggotaan dihapus).

Mengundang anggota untuk berkolaborasi

Sebagai pembuat kolaborasi, setelah Anda membuat kolaborasi, Anda dapat mengirim tautan undangan ke anggota yang tercantum di tab Anggota.

Mengundang anggota untuk berkolaborasi

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang Anda buat.
4. Pilih tab Anggota.
5. Di tabel Anggota, pilih tombol Salin tautan undangan.

Tautan undangan disalin.

6. Rekatkan tautan undangan dalam metode komunikasi aman pilihan Anda dan kirimkan ke setiap anggota kolaborasi.

Memantau anggota

Sebagai pembuat kolaborasi, setelah Anda membuat kolaborasi, Anda dapat memantau status semua anggota di tab Anggota. Ini membantu memastikan kontrol akses yang tepat dan manajemen biaya di seluruh kolaborasi Anda.

Untuk memantau status anggota

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang Anda buat.
4. Pilih tab Anggota.
5. Dalam tabel Anggota, tinjau Status masing-masing anggota.
6. Dalam tabel Kemampuan Anggota, tinjau anggota mana yang dapat melakukan kueri, menerima hasil, menyumbangkan data, dan melakukan tugas lainnya.
7. Dalam tabel konfigurasi Pembayaran, tinjau anggota mana yang membayar untuk kueri, tabel pemetaan ID, dan pemodelan ML.

Menambahkan anggota ke kolaborasi

Prasyarat

- Akun AWS dengan izin untuk mengelola kolaborasi
- Akun AWSIDs Anggota yang ingin Anda tambahkan

Untuk menambahkan anggota ke kolaborasi

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang ingin Anda tambahkan anggota.

4. Pilih tab Anggota.
5. Pilih Edit anggota.
6. Pilih Tambahkan anggota lain dan masukkan informasi berikut:
 - Nama tampilan anggota
 - Akun AWSID Anggota
 - Tentukan apakah mereka dapat menerima hasil
7. Pilih Simpan perubahan.
8. Dalam modal konfirmasi, verifikasi bahwa informasi tersebut benar, dan pilih Kirim permintaan perubahan.
 - Jika permintaan perubahan memerlukan persetujuan anggota lain, semua anggota yang ada harus menyetujui permintaan perubahan sebelum anggota baru ditambahkan. Untuk informasi selengkapnya tentang permintaan perubahan, lihat [the section called “Perubahan permintaan”](#).
 - Jika jenis perubahan persetujuan otomatis didukung untuk anggota baru dengan kemampuan yang ditentukan, maka perubahan ini akan segera berlaku. Untuk informasi selengkapnya, lihat [the section called “Edit pengaturan persetujuan otomatis kolaborasi”](#).
9. Pada halaman detail kolaborasi, di bawah tab Anggota, verifikasi bahwa status Anggota anggota yang ditambahkan ditampilkan sebagai Diundang.

Setelah menyelesaikan langkah-langkah ini, anggota yang diundang dapat bergabung dengan kolaborasi. Untuk informasi selengkapnya tentang bergabung dengan kolaborasi, lihat [Membuat keanggotaan dan bergabung dengan kolaborasi](#)

Menghapus anggota dari kolaborasi

Note

Sebelum Anda mulai, perhatikan bahwa menghapus anggota:

- Menghapus semua kumpulan data terkait dari kolaborasi
- Jika [anggota membayar biaya komputasi kueri](#), tindakan ini menghentikan semua eksekusi kueri dalam kolaborasi.

Menghapus anggota juga menghapus semua kumpulan data terkait dari kolaborasi.

Prasyarat

- Anda harus menjadi pencipta kolaborasi
- Anda tidak dapat menghapus akun Anda sendiri

Untuk menghapus anggota dari kolaborasi

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang ingin Anda modifikasi.
4. Pilih tab Anggota.
5. Pilih tombol opsi di sebelah anggota yang ingin Anda hapus.
6. Pilih Hapus.
7. Di kotak dialog konfirmasi, ketik **confirm** untuk memverifikasi penghapusan.

Note

Setelah Anda menghapus anggota, semua kumpulan data yang terkait dengan akun mereka juga dihapus dari kolaborasi.

Important

Jika Anda menghapus anggota yang membayar biaya komputasi kueri, tidak ada kueri lebih lanjut yang dapat dijalankan dalam kolaborasi sampai Anda menunjuk anggota baru yang membayar.

Meninggalkan Kolaborasi

Sebagai anggota kolaborasi, Anda dapat meninggalkan kolaborasi dengan menghapus keanggotaan Anda. Jika Anda adalah pembuat kolaborasi, Anda hanya dapat meninggalkan kolaborasi dengan [menghapus kolaborasi](#).

 Note

Ketika Anda menghapus keanggotaan Anda, Anda meninggalkan kolaborasi dan tidak dapat bergabung kembali. Jika Anda adalah [anggota yang membayar biaya komputasi kueri](#) dan Anda menghapus keanggotaan Anda, tidak ada lagi kueri yang diizinkan untuk dijalankan.

Untuk meninggalkan kolaborasi

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Untuk Dengan keanggotaan aktif, pilih kolaborasi di mana Anda menjadi anggota.
4. Pilih Tindakan.
5. Pilih Hapus keanggotaan.
6. Di kotak dialog, konfirmasikan keputusan untuk meninggalkan kolaborasi dengan mengetikkan **confirm** kolom input teks, lalu pilih Kosongkan dan hapus keanggotaan.

Anda melihat pesan di konsol yang menunjukkan bahwa keanggotaan telah dihapus.

Pembuat kolaborasi melihat status Anggota sebagai Kiri.

Siapkan tabel data di AWS Clean Rooms

Note

Mempersiapkan tabel data dapat dilakukan sebelum atau setelah Anda bergabung dengan kolaborasi. Setelah tabel disiapkan, Anda dapat menggunakannya kembali di beberapa kolaborasi selama kebutuhan privasi Anda untuk tabel itu sama.

Sebagai anggota dalam kolaborasi, Anda harus menyiapkan tabel data Anda sebelum mereka dapat ditanyakan AWS Clean Rooms oleh anggota kolaborasi yang dapat melakukan kueri.

Tabel data yang Anda gunakan untuk kueri AWS Clean Rooms biasanya jenis tabel data yang sama yang Anda gunakan untuk aplikasi lain. Misalnya, jenis kumpulan data yang sama digunakan dengan Amazon Athena, Amazon EMR, Amazon Redshift Spectrum, dan Amazon Quick.

Anda dapat menanyakan data dalam format aslinya langsung dari salah satu sumber data berikut:

- Amazon Simple Storage Service (Amazon S3)
- Amazon Athena
- Kepingan salju

AWS Clean Rooms mengakses kumpulan data pada waktu proses kueri, memastikan bahwa anggota yang dapat melakukan kueri selalu mengakses sebagian besar up-to-date data. Setiap data yang sementara dibaca ke dalam AWS Clean Rooms kolaborasi akan dihapus setelah kueri selesai. Hasil kueri ditulis ke bucket Amazon S3 Anda.

Jika kasus penggunaan Anda melibatkan kueri data identitas, lihat [Resolusi Entitas AWS di AWS Clean Rooms](#).

Topik

- [Format data untuk AWS Clean Rooms](#)
- [Apache Iceberg tabel di AWS Clean Rooms](#)
- [Mempersiapkan tabel data untuk kueri di AWS Clean Rooms](#)
- [Mempersiapkan tabel data terenkripsi dengan Cryptographic Computing untuk Clean Rooms](#)

- [Mendekripsi tabel data dengan klien enkripsi C3R](#)

Format data untuk AWS Clean Rooms

Untuk menganalisis data, dataset harus dalam format yang AWS Clean Rooms mendukung.

Topik

- [Format data yang didukung untuk PySpark pekerjaan](#)
- [Format data yang didukung untuk kueri SQL](#)
- [Jenis data yang didukung](#)
- [Jenis kompresi file untuk AWS Clean Rooms](#)
- [Enkripsi sisi server untuk AWS Clean Rooms](#)

Format data yang didukung untuk PySpark pekerjaan

AWS Clean Rooms mendukung format terstruktur berikut untuk menjalankan PySpark pekerjaan.

- Parquet
- OpenCSV
- JSON

Format data yang didukung untuk kueri SQL

AWS Clean Rooms mendukung format terstruktur berikut untuk menjalankan kueri SQL.

- [Tabel Apache Iceberg](#)
- Parquet
- OpenCSV
- JSON

Note

timestampNilai dalam file teks harus dalam formatyyy-MM-dd HH:mm:ss.SSSSSS.
Misalnya:2017-05-01 11:30:59.000000.

Sebaiknya gunakan format file penyimpanan kolumnar, seperti Apache Parquet. Dengan format file penyimpanan kolumnar, Anda dapat meminimalkan pergerakan data dengan memilih hanya kolom yang Anda butuhkan. Untuk kinerja optimal, objek besar harus dibagi menjadi objek 100mb—1gb.

Jenis data yang didukung

AWS Clean Rooms mendukung tipe data berikut.

- ARRAY
- BIGINT
- BOOLEAN
- BYTE
- CHAR
- DATE
- DECIMAL
- FLOAT
- INTEGER
- INTERVAL
- LONG
- PETA
- REAL
- SHORT
- SMALLINT
- STRUCT
- TIME
- STAMP_LTZ
- TIMESTAMP_NTZ
- TINYINT
- VARCHAR

Untuk informasi selengkapnya, lihat [Tipe data](#) dalam Referensi AWS Clean Rooms SQL.

Jenis kompresi file untuk AWS Clean Rooms

Untuk mengurangi ruang penyimpanan, meningkatkan kinerja, dan meminimalkan biaya, kami sangat menyarankan Anda untuk mengompres kumpulan data Anda.

AWS Clean Rooms mengenali jenis kompresi file berdasarkan ekstensi file dan mendukung jenis kompresi dan ekstensi yang ditunjukkan pada tabel berikut.

Algoritma kompresi	Ekstensi file
GZIP	.gz
Bzip2	.bz2
Snappy	.snappy

Anda dapat menerapkan kompresi pada level yang berbeda. Paling umum, Anda mengompres seluruh file atau mengompres blok individual dalam file. Mengompresi format kolumnar pada tingkat file tidak menghasilkan manfaat kinerja.

Enkripsi sisi server untuk AWS Clean Rooms

Note

Enkripsi sisi server tidak menggantikan komputasi kriptografi untuk kasus penggunaan yang memerlukannya.

AWS Clean Rooms secara transparan mendekripsi kumpulan data yang dienkripsi menggunakan opsi enkripsi berikut:

- SSE-S3 - Enkripsi sisi server menggunakan kunci enkripsi AES-256 yang dikelola oleh Amazon S3
- SSE-KMS - Enkripsi sisi server dengan kunci yang dikelola oleh AWS Key Management Service

Untuk menggunakan SSE-S3, peran AWS Clean Rooms layanan yang digunakan untuk mengaitkan tabel yang dikonfigurasi ke kolaborasi harus memiliki izin dekripsi KMS. Untuk menggunakan SSE-KMS, kebijakan kunci KMS juga harus mengizinkan peran AWS Clean Rooms layanan untuk mendekripsi.

AWS Clean Rooms tidak mendukung enkripsi sisi klien Amazon S3. Untuk informasi selengkapnya tentang enkripsi sisi server, lihat [Melindungi data menggunakan enkripsi sisi server](#) di Panduan Pengguna Layanan Penyimpanan Sederhana Amazon.

Apache Iceberg tabel di AWS Clean Rooms

Apache Iceberg adalah format tabel sumber terbuka untuk data. AWS Clean Rooms dapat menggunakan statistik yang disimpan di Apache Iceberg metadata untuk mengoptimalkan rencana kueri dan mengurangi pemindaian file selama pemrosesan kueri ruang bersih. Untuk informasi lebih lanjut, lihat dokumentasi [Apache Iceberg](#).

Pertimbangkan hal berikut saat menggunakan AWS Clean Rooms dengan tabel Iceberg:

- Tabel Apache Iceberg untuk S3 - Apache Iceberg tabel harus didefinisikan AWS Glue Data Catalog berdasarkan [implementasi katalog lem open source](#).
- Tabel Apache Iceberg untuk Athena - [Untuk informasi lebih lanjut, lihat -iceberg.html https://docs.aws.amazon.com/athena/latest/ug/querying](#)
- Tabel Apache Iceberg untuk Snowflake - Untuk informasi [lebih lanjut, lihat panduan pengguna/ tabel-gunung es https://docs.snowflake.com/en/](#)
- Format file parquet - AWS Clean Rooms hanya mendukung tabel Iceberg dalam format file data Parquet.
- Kompresi GZIP dan Snappy — AWS Clean Rooms mendukung Parquet dengan GZIP dan Snappy kompresi
- Versi Iceberg - AWS Clean Rooms mendukung menjalankan kueri terhadap versi 1 dan versi 2 tabel Iceberg.
- Partisi - Anda tidak perlu menambahkan partisi secara manual untuk Apache Iceberg tabel di AWS Glue. AWS Clean Rooms mendeteksi partisi baru di Apache Iceberg tabel secara otomatis dan tidak diperlukan operasi manual untuk memperbarui partisi dalam definisi tabel. Partisi gunung es muncul sebagai kolom reguler dalam skema AWS Clean Rooms tabel dan tidak secara terpisah sebagai kunci partisi dalam skema tabel yang dikonfigurasi.
- Batasan
 - Hanya tabel Iceberg baru

Apache Iceberg tabel dikonversi dari Apache Parquet tabel tidak didukung.

- Pertanyaan perjalanan waktu

AWS Clean Rooms tidak mendukung kueri perjalanan waktu dengan Apache Iceberg tabel.

- Mesin Athena versi 2

Iceberg tabel yang dibuat dengan mesin Athena versi 2 tidak didukung.

- Format berkas

Avro dan format file Optimized Row Columnar (ORC) tidak didukung.

- Kompresi

Zstandard (Zstd) kompresi untuk Parquet tidak didukung.

Tipe data yang didukung untuk tabel Iceberg

AWS Clean Rooms bisa query Iceberg tabel yang berisi tipe data berikut:

- BOOLEAN
- DATE
- DECIMAL
- DOUBLE
- FLOAT
- INT
- LIST
- LONG
- MAP
- STRING
- STRUCT
- TIMESTAMP WITHOUT TIME ZONE

Untuk informasi selengkapnya tentang tipe data Gunung Es, lihat [Skema untuk Gunung Es di dokumentasi Apache Iceberg](#).

Mempersiapkan tabel data untuk kueri di AWS Clean Rooms

Jika kasus penggunaan Anda tidak mengharuskan Anda untuk membawa data Anda sendiri, Anda dapat melewati prosedur ini.

Jika kasus penggunaan Anda melibatkan kueri data identitas, lihat [Resolusi Entitas AWS di AWS Clean Rooms](#).

Untuk informasi selengkapnya tentang format data yang dapat Anda gunakan, lihat [Format data untuk AWS Clean Rooms](#).

Topik

- [Mempersiapkan tabel data di Amazon S3](#)
- [Mempersiapkan tabel data di Amazon Athena](#)
- [Mempersiapkan tabel data di Snowflake](#)

Mempersiapkan tabel data di Amazon S3

Anda dapat menganalisis tabel data yang telah dikatalogkan AWS Glue dan disimpan di Amazon S3. Jika tabel data Anda sudah dikatalogkan AWS Glue, lewati ke. [Membuat tabel yang dikonfigurasi di AWS Clean Rooms](#)

Important

Saat menyiapkan data Anda di Amazon S3 untuk digunakan AWS Clean Rooms, Anda harus memastikan bahwa lokasi AWS Glue tabel Anda sama persis dengan jalur direktori tempat file data Anda disimpan.

Misalnya: Jika data Anda disimpan dalam `s3://mybucket/folder/subfolder/data.parquet`, lokasi AWS Glue tabel Anda harus menunjuk ke `s3://mybucket/folder/subfolder/`. Mengatur lokasi tabel ke direktori induk (`s3://mybucket/folder/`) akan mengakibatkan tabel tampak kosong saat ditanyakan.

Mempersiapkan tabel data Anda di Amazon S3 melibatkan langkah-langkah berikut:

Topik

- [Langkah 1: Selesaikan prasyarat](#)

- [Langkah 2: \(Opsional\) Siapkan data Anda untuk komputasi kriptografi](#)
- [Langkah 3: Unggah tabel data Anda ke Amazon S3](#)
- [Langkah 4: Buat AWS Glue tabel](#)
- [Langkah 5: Langkah selanjutnya](#)

Langkah 1: Selesaikan prasyarat

Untuk menyiapkan tabel data Anda untuk digunakan AWS Clean Rooms, Anda harus menyelesaikan prasyarat berikut:

- Tabel data Anda disimpan sebagai salah satu [format data yang didukung untuk AWS Clean Rooms](#).
- Tabel data Anda dikatalogkan AWS Glue dan menggunakan [tipe data yang didukung](#) untuk AWS Clean Rooms
- Semua tabel data Anda disimpan di Amazon Simple Storage Service (Amazon S3) Simple Storage Service (Amazon S3) yang AWS Region sama saat kolaborasi dibuat.
- AWS Glue Data Catalog Harus berada di Wilayah yang sama dengan kolaborasi.
- AWS Glue Data Catalog Itu sama Akun AWS dengan keanggotaan.
- Bucket Amazon S3 tidak terdaftar. AWS Lake Formation

Langkah 2: (Opsional) Siapkan data Anda untuk komputasi kriptografi

(Opsional) Jika Anda menggunakan komputasi kriptografi dan tabel data Anda berisi informasi sensitif yang ingin Anda enkripsi, Anda harus mengenkripsi tabel data menggunakan klien enkripsi C3R.

Untuk mempersiapkan data Anda untuk komputasi kriptografi, ikuti prosedur di [Mempersiapkan tabel data terenkripsi dengan Cryptographic Computing untuk Clean Rooms](#).

Langkah 3: Unggah tabel data Anda ke Amazon S3

Note

Jika Anda bermaksud menggunakan tabel data terenkripsi dalam kolaborasi, Anda harus terlebih dahulu mengenkripsi data untuk komputasi kriptografi sebelum mengunggah tabel data Anda ke Amazon S3. Untuk informasi selengkapnya, lihat [Mempersiapkan tabel data terenkripsi dengan Cryptographic Computing untuk Clean Rooms](#).

Untuk mengunggah tabel data Anda ke Amazon S3

1. Masuk ke Konsol Manajemen AWS dan buka konsol Amazon S3 di <https://console.aws.amazon.com/s3/>
2. Pilih Bucket, lalu pilih bucket tempat Anda ingin menyimpan tabel data Anda.
3. Pilih Unggah, lalu ikuti petunjuknya.
4. Pilih tab Objek untuk melihat awalan tempat data Anda disimpan. Catat nama folder.

Anda dapat memilih folder untuk melihat data.

Langkah 4: Buat AWS Glue tabel

Jika Anda sudah memiliki tabel AWS Glue data, Anda dapat melewati langkah ini.

Pada langkah ini, Anda menyiapkan crawler yang meng-crawl semua file di bucket S3 dan membuat tabel. AWS Glue Untuk informasi selengkapnya, lihat [Mendefinisikan crawler AWS Glue di AWS Glue](#) Panduan Pengguna.

Untuk informasi selengkapnya tentang tipe AWS Glue Data Catalog data yang didukung, lihat [Jenis data yang didukung](#).

Note

AWS Clean Rooms saat ini tidak mendukung bucket S3 yang terdaftar. AWS Lake Formation

Prosedur berikut menjelaskan cara membuat AWS Glue tabel. Jika Anda ingin menggunakan AWS Glue Data Catalog objek terenkripsi dengan kunci AWS Key Management Service (AWS KMS), Anda perlu mengonfigurasi kebijakan izin kunci KMS untuk mengizinkan akses ke tabel terenkripsi tersebut. Untuk informasi selengkapnya, lihat [Menyiapkan enkripsi di AWS Glue](#) di Panduan AWS Glue Pengembang.

Untuk membuat AWS Glue tabel

1. Ikuti [Bekerja dengan crawler pada prosedur AWS Glue konsol](#) di Panduan AWS Glue Pengguna.
2. Buat catatan nama AWS Glue database dan nama AWS Glue tabel.

Langkah 5: Langkah selanjutnya

Sekarang setelah Anda menyiapkan tabel data di Amazon S3, Anda siap untuk:

- [Buat tabel yang dikonfigurasi](#)
- [Buat model ML](#)

Tabel dapat ditanyakan setelah:

- Pembuat kolaborasi telah membuat kolaborasi di AWS Clean Rooms. Untuk informasi selengkapnya, lihat [Menciptakan kolaborasi](#).
- Pembuat kolaborasi telah mengirimkan ID kolaborasi kepada Anda sebagai peserta dalam kolaborasi.

Mempersiapkan tabel data di Amazon Athena

Anda dapat menanyakan tabel data yang telah dibuat sebagai Tampilan AWS Glue Data Catalog (GDC) di Amazon Athena.

Tampilan GDC adalah tabel virtual, dibuat dari satu atau lebih AWS Glue tabel yang mendasarinya. Itu harus dibuat menggunakan Athena SQL di katalog Athena. `AwsGlueCatalog`

Mempersiapkan tabel data Anda di Amazon Athena melibatkan langkah-langkah berikut:

Topik

- [Langkah 1: Selesaikan prasyarat](#)
- [Langkah 2: \(Opsional\) Siapkan data Anda untuk komputasi kriptografi](#)
- [Langkah 3: Langkah selanjutnya](#)

Langkah 1: Selesaikan prasyarat

Untuk menyiapkan tabel data Anda untuk digunakan AWS Clean Rooms, Anda harus menyelesaikan prasyarat berikut:

- Tabel data Anda disimpan sebagai salah satu [format data yang didukung untuk AWS Clean Rooms](#).
- Tabel data Anda menggunakan [tipe data yang didukung untuk AWS Clean Rooms](#).

- Anda telah membuat Tampilan GDC di AWS Glue meja Anda menggunakan Athena SQL di katalog Athena. `AwsDataCatalog`

Tampilan akan muncul di:

- Konsol Athena (di bawah `AwsDataCatalog`) sebagai Tampilan: <https://console.aws.amazon.com/athena/>
- AWS Glue Konsol sebagai AWS Glue tabel: <https://console.aws.amazon.com/glue/>

Untuk informasi selengkapnya, lihat [Menggunakan tampilan Katalog Data di Athena](#) di Panduan Pengguna Amazon Athena.

Note

Anda memerlukan izin yang sesuai untuk membuat Tampilan di AWS Glue Athena dan. Selain itu, pastikan Anda memiliki akses ke tabel dasar yang direferensikan dalam definisi Tampilan Anda.

AWS Clean Rooms hanya mendukung Jenis AWS Glue Katalog untuk Athena, bukan Jenis Katalog Lambda atau Sarang.

- Tabel data atau Tampilan GDC Anda dikatalogkan AWS Glue dan terdaftar. AWS Lake Formation
- Anda telah membuat bucket keluaran terpisah di Amazon S3 untuk menerima hasil Athena.
- Anda telah menyiapkan peran layanan untuk membaca data dari Amazon Athena. Untuk informasi selengkapnya, lihat [Membuat peran layanan untuk membaca data dari Amazon Athena](#).
- Peran layanan memiliki izin akses Lake Formation Select dan Describe pada Tampilan atau tabel GDC.

Langkah 2: (Opsional) Siapkan data Anda untuk komputasi kriptografi

(Opsional) Jika Anda menggunakan komputasi kriptografi dan tabel data Anda berisi informasi sensitif yang ingin Anda enkripsi, Anda harus mengenkripsi tabel data menggunakan klien enkripsi C3R.

Untuk mempersiapkan data Anda untuk komputasi kriptografi, ikuti prosedur di [Mempersiapkan tabel data terenkripsi dengan Cryptographic Computing untuk Clean Rooms](#).

Langkah 3: Langkah selanjutnya

Sekarang setelah Anda menyiapkan tabel data di Amazon Athena, Anda siap untuk:

- [Buat tabel yang dikonfigurasi](#)
- [Buat model ML](#)

Tabel dapat ditanyakan setelah:

- Pembuat kolaborasi telah membuat kolaborasi di AWS Clean Rooms. Untuk informasi selengkapnya, lihat [Menciptakan kolaborasi](#).
- Pembuat kolaborasi telah mengirimkan ID kolaborasi kepada Anda sebagai peserta dalam kolaborasi.

Mempersiapkan tabel data di Snowflake

Anda dapat menanyakan tabel data yang telah disimpan di gudang data Snowflake.

Mempersiapkan tabel data Anda di Snowflake melibatkan langkah-langkah berikut:

Topik

- [Langkah 1: Selesaikan prasyarat](#)
- [Langkah 2: \(Opsional\) Siapkan data Anda untuk komputasi kriptografi](#)
- [Langkah 3: Buat AWS Secrets Manager rahasia](#)
- [Langkah 4: Langkah selanjutnya](#)

Langkah 1: Selesaikan prasyarat

Untuk menyiapkan tabel data Anda untuk digunakan AWS Clean Rooms, Anda harus menyelesaikan prasyarat berikut:

- Anda memiliki izin Akun AWS yang tepat yang diberikan untuk membaca tabel data Anda. Untuk informasi selengkapnya, lihat [Buat peran layanan untuk membaca data dari Snowflake](#).
- Tabel data Anda disimpan sebagai salah satu [format data yang didukung untuk AWS Clean Rooms](#).
- Tabel data Anda menggunakan [tipe data yang didukung untuk AWS Clean Rooms](#).
- Tabel data Anda disimpan di gudang Snowflake. Untuk informasi lebih lanjut, lihat dokumentasi [Snowflake](#).

- Anda telah menyiapkan pengguna Snowflake baru dengan hak istimewa hanya-baca ke tabel Snowflake yang akan Anda kaitkan dengan kolaborasi Anda.

Langkah 2: (Opsional) Siapkan data Anda untuk komputasi kriptografi

(Opsional) Jika Anda menggunakan komputasi kriptografi dan tabel data Anda berisi informasi sensitif yang ingin Anda enkripsi, Anda harus mengenkripsi tabel data menggunakan klien enkripsi C3R.

Untuk mempersiapkan data Anda untuk komputasi kriptografi, ikuti prosedur di [Mempersiapkan tabel data terenkripsi dengan Cryptographic Computing untuk Clean Rooms](#).

Langkah 3: Buat AWS Secrets Manager rahasia

Untuk terhubung ke Snowflake dari AWS Clean Rooms, Anda harus membuat dan menyimpan kredensi Snowflake Anda secara AWS Secrets Manager rahasia, lalu mengaitkan rahasia itu dengan tabel Snowflake di AWS Clean Rooms

Note

Kami menyarankan Anda membuat pengguna baru yang khusus untuk AWS Clean Rooms. Pengguna itu seharusnya hanya memiliki peran dengan izin Baca untuk data yang AWS Clean Rooms ingin Anda akses.

Untuk membuat AWS Secrets Manager rahasia

1. Di Snowflake, buat pengguna `snowflakeUser` dan atur otentikasi pasangan kunci.

Note

Pada November 2025, Snowflake akan bertransisi untuk hanya mendukung otentikasi pasangan kunci. Perubahan ini akan memengaruhi AWS Clean Rooms integrasi saat ini dengan Snowflake, yang menggunakan otentikasi nama pengguna dan kata sandi. Setelah tanggal ini, koneksi Snowflake AWS Clean Rooms akan memerlukan otentikasi pasangan kunci menggunakan kunci pribadi Snowflake Privacy Enhanced Mail (PEM).

2. Tentukan gudang Snowflake mana yang akan berinteraksi dengan pengguna ini, `snowflakeWarehouse` Entah mengaturnya sebagai `DEFAULT_WAREHOUSE` for `snowflakeUser` di Snowflake atau ingat untuk langkah selanjutnya.

3. Di [AWS Secrets Manager](#), buat rahasia menggunakan kredensial Snowflake Anda. Untuk membuat rahasia di Secrets Manager, ikuti tutorial yang tersedia di [Buat AWS Secrets Manager rahasia](#) di Panduan AWS Secrets Manager Pengguna. Setelah membuat rahasia, simpan nama Rahasia, `secretName` untuk langkah selanjutnya.
 - Saat memilih pasangan kunci/nilai, buat pasangan `snowflakeUser` dengan kunci. `sfUser`
 - Saat memilih pasangan kunci/nilai, buat pasangan untuk kunci pribadi Snowflake PEM Anda dengan kunci tersebut. `pem_private_key`
 - Saat memilih pasangan kunci/nilai, buat pasangan `snowflakeWarehouse` dengan kunci. `sfWarehouse`

Ini tidak diperlukan jika default diatur di Snowflake.

 - Saat memilih pasangan kunci/nilai, buat pasangan `snowflakeRole` dengan kunci. `sfRole`

Langkah 4: Langkah selanjutnya

Sekarang setelah Anda menyiapkan tabel data Anda di Snowflake, Anda siap untuk:

- [Buat tabel yang dikonfigurasi](#)
- [Buat model ML](#)

Tabel dapat ditanyakan setelah:

- Pembuat kolaborasi telah membuat kolaborasi di AWS Clean Rooms. Untuk informasi selengkapnya, lihat [Menciptakan kolaborasi](#).
- Pembuat kolaborasi telah mengirimkan ID kolaborasi kepada Anda sebagai peserta dalam kolaborasi.

Mempersiapkan tabel data terenkripsi dengan Cryptographic Computing untuk Clean Rooms

Cryptographic Computing for Clean Rooms (C3R) adalah kemampuan dalam AWS Clean Rooms Anda dapat menggunakan C3R untuk membatasi secara kriptografi apa yang dapat dipelajari oleh pihak mana pun dan AWS dalam kolaborasi. AWS Clean Rooms

Anda dapat mengenkripsi tabel data menggunakan klien enkripsi C3R, alat enkripsi sisi klien, sebelum mengunggah tabel data ke sumber data Anda: Amazon Simple Storage Service (Amazon S3), Amazon Athena, atau Snowflake.

Untuk informasi selengkapnya, lihat [Komputasi Kriptografi untuk Clean Rooms](#).

Mempersiapkan tabel data terenkripsi dengan C3R melibatkan langkah-langkah berikut:

Langkah-langkah

- [Langkah 1: Selesaikan prasyarat](#)
- [Langkah 2: Unduh klien enkripsi C3R](#)
- [Langkah 3: \(Opsional\) Lihat perintah yang tersedia di klien enkripsi C3R](#)
- [Langkah 4: Buat skema enkripsi untuk file tabular](#)
- [Langkah 5: Buat kunci rahasia bersama](#)
- [Langkah 6: Simpan kunci rahasia bersama dalam variabel lingkungan](#)
- [Langkah 7: Enkripsi data](#)
- [Langkah 8: Verifikasi enkripsi data](#)
- [\(Opsional\) Buat skema \(pengguna tingkat lanjut\)](#)

Langkah 1: Selesaikan prasyarat

Untuk menyiapkan tabel data Anda untuk digunakan dengan C3R, Anda harus menyelesaikan prasyarat berikut:

- Anda dapat mengakses Komputasi Kriptografi untuk Clean Rooms repositori di: GitHub

<https://github.com/aws/c3r>

- Anda telah menyiapkan AWS kredensial untuk menggunakan klien enkripsi C3R. Kredensial ini digunakan oleh klien enkripsi C3R untuk panggilan API hanya-baca untuk mengambil metadata kolaborasi. AWS Clean Rooms Untuk informasi selengkapnya, lihat [AWS CLI Mengonfigurasi Panduan AWS Command Line Interface Pengguna untuk Versi 2](#).
- Anda memiliki Java Runtime Environment (JRE) 11 atau lebih baru diinstal pada mesin Anda.
 - [Yang direkomendasikan Java Runtime Environment, Amazon Corretto 11 atau lebih tinggi, dapat diunduh dari/corretto. https://aws.amazon.com](#)

- The Java Development Kit (JDK) termasuk yang sesuai JRE dari versi yang sama. Namun, kemampuan tambahan tidak JDK diperlukan untuk menjalankan klien enkripsi Cryptographic Computing for Clean Rooms (C3R).
- File data tabular Anda (.csv) atau Parquet file (.parquet) disimpan secara lokal.
- Anda atau anggota lain dalam kolaborasi memiliki kemampuan untuk membuat kunci rahasia bersama. Untuk informasi selengkapnya, lihat [Langkah 5: Buat kunci rahasia bersama](#).
- Pencipta kolaborasi telah menciptakan kolaborasi AWS Clean Rooms dengan komputasi Cryptographic yang diaktifkan untuk kolaborasi. Untuk informasi selengkapnya, lihat [Menciptakan kolaborasi](#).
- Pembuat kolaborasi telah mengirimkan ID kolaborasi kepada Anda sebagai peserta dalam kolaborasi. Kolaborasi Amazon Resource Name (ARN) disertakan dalam undangan yang dikirim, yang berisi ID kolaborasi.

Langkah 2: Unduh klien enkripsi C3R

Untuk mengunduh klien enkripsi C3R dari GitHub

1. [Pergi ke Komputasi Kriptografi untuk Clean RoomsAWSGitHub repositori: c3r https://github.com/aws/](https://github.com/aws/)
2. Pilih dan unduh file.

Kode sumber, lisensi, dan materi terkait dapat dikloning atau diunduh sebagai file. zipfile dari halaman arahan GitHub repositori. (Lihat tombol Kode di kanan atas daftar konten repositori).

Klien enkripsi C3R terbaru yang ditandatangani Java Executable File (yaitu, aplikasi antarmuka baris perintah) ada di halaman Rilis repositori. GitHub

Paket klien enkripsi C3R untuk Apache Spark (`c3r-cli-spark`) adalah versi `c3r-cli` yang harus dikirimkan sebagai pekerjaan ke server Apache Spark yang sedang berjalan. Untuk informasi selengkapnya, lihat [Menjalankan C3R di Apache Spark](#).

Langkah 3: (Opsional) Lihat perintah yang tersedia di klien enkripsi C3R

Gunakan prosedur ini untuk membiasakan diri dengan perintah yang tersedia di klien enkripsi C3R.

Untuk melihat semua perintah yang tersedia di klien enkripsi C3R

1. Dari antarmuka baris perintah (CLI), navigasikan ke folder yang berisi file yang diunduh `c3r-cli.jar`.
2. Jalankan perintah berikut: `java -jar c3r-cli.jar`
3. Lihat daftar perintah dan opsi yang tersedia.

Langkah 4: Buat skema enkripsi untuk file tabular

Untuk mengenkripsi data, diperlukan skema enkripsi yang menjelaskan bagaimana data akan digunakan. Bagian ini menjelaskan bagaimana klien enkripsi C3R membantu dalam menghasilkan skema enkripsi untuk file CSV dengan baris header atau file. Parquet

Anda hanya perlu melakukan ini sekali per file. Setelah skema ada, dapat digunakan kembali untuk mengenkripsi file yang sama (atau file apa pun dengan nama kolom yang identik). Jika nama kolom atau skema enkripsi yang diinginkan berubah, Anda harus memperbarui file skema. Untuk informasi selengkapnya, lihat [\(Opsional\) Buat skema \(pengguna tingkat lanjut\)](#).

Important

Sangat penting bahwa semua pihak yang berkolaborasi menggunakan kunci rahasia bersama yang sama. Pihak yang berkolaborasi juga harus mengoordinasikan nama kolom agar sesuai jika mereka akan JOIN di-ed atau dibandingkan untuk kesetaraan dalam kueri. Jika tidak, kueri SQL mungkin menghasilkan hasil yang tidak terduga atau salah. Namun, ini tidak diperlukan jika pembuat kolaborasi mengaktifkan pengaturan `allowJoinsOnColumnsWithDifferentNames` enkripsi selama pembuatan kolaborasi. Untuk informasi selengkapnya tentang setelan yang relevan dengan enkripsi, lihat. [Parameter komputasi kriptografi](#)

Ketika dijalankan dalam mode skema, klien enkripsi C3R melewati kolom file input demi kolom, meminta Anda apakah dan bagaimana kolom itu harus diperlakukan. Jika file berisi banyak kolom yang tidak diinginkan untuk output terenkripsi, pembuatan skema interaktif mungkin menjadi membosankan karena Anda harus melewati setiap kolom yang tidak diinginkan. Untuk menghindari hal ini, Anda dapat menulis skema secara manual, atau membuat versi sederhana dari file input yang hanya menampilkan kolom yang diinginkan. Kemudian, generator skema interaktif dapat dijalankan pada file yang dikurangi itu. Klien enkripsi C3R mengeluarkan informasi tentang file skema dan

menanyakan bagaimana kolom sumber harus disertakan atau dienkripsi (jika ada) dalam output target.

Untuk setiap kolom sumber dalam file input, Anda diminta untuk:

1. Berapa banyak kolom target yang harus dihasilkan
2. Bagaimana setiap kolom target harus dienkripsi (jika ada)
3. Nama setiap kolom target
4. Bagaimana data harus diempuk sebelum enkripsi jika kolom dienkripsi sebagai kolom sealed

Note

Saat Anda mengenkripsi data untuk kolom yang telah dienkripsi sebagai sealed kolom, Anda harus menentukan data mana yang membutuhkan padding. Klien enkripsi C3R menyarankan padding default selama pembuatan skema yang membungkus semua entri dalam kolom dengan panjang yang sama.

Saat menentukan panjangnya `fixed`, perhatikan bahwa padding dalam byte, bukan bit.

Berikut ini adalah tabel keputusan untuk membuat skema.

Tabel keputusan skema

Keputusan	Jumlah kolom target dari kolom sumber <'name-of-column'>?	Jenis kolom target: [c]cleartext, [f]fingerprint, atau [s]sealed?	Nama judul kolom target <default 'name-of-column'>	Tambahkan akhiran <suffix>ke header untuk menunjukkan bagaimana itu dienkripsi, [y] ya atau [n] tidak <default 'yes'>	<'name-of-column_disegel'> tipe padding: [n] satu, [f] tetap, atau [m] maks <default 'max'>
Biarkan kolom tidak terenkripsi.	1	c	Tidak berlaku	Tidak berlaku	Tidak berlaku

Keputusan	Jumlah kolom target dari kolom sumber <'name-of-column'>?	Jenis kolom target: [c]cleartext, [f]fingerprint, atau [s]sealed?	Nama judul kolom target <default 'name-of-column'>	Tambahkan akhiran <suffix>ke header untuk menunjukkan bagaimana itu dienkripsi, [y] ya atau [n] tidak <default 'yes'>	<'name-of-column _disegel'> tipe padding: [n] satu, [f] tetap, atau [m] maks <default 'max'>
Enkripsi kolom sebagai fingerprint kolom.	1	f	Pilih default atau masukkan nama header baru.	Masukkan y untuk memilih default (<code>_fingerprint</code>) atau entern.	Tidak berlaku
Enkripsi kolom sebagai sealed kolom.	1	s	Pilih default atau masukkan nama header baru.	Masukkan y untuk memilih default (<code>_sealed</code>) atau entern.	Pilih jenis padding. Untuk informasi selengkapnya, lihat (Opsional) Buat skema (pengguna tingkat lanjut) .

Keputusan	Jumlah kolom target dari kolom sumber <'name-of-column'>?	Jenis kolom target: [c]cleartext, [f]fingerprint, atau [s]sealed?	Nama judul kolom target <default 'name-of-column'>	Tambahkan akhiran <suffix>ke header untuk menunjukkan bagaimana itu dienkripsi, [y] ya atau [n] tidak <default 'yes'>	<'name-of-column _disegel'> tipe padding: [n] satu, [f] tetap, atau [m] maks <default 'max'>
Enkripsi kolom sebagai keduanya fingerprint dansealed.	2	Masukkan kolom target pertama: f. Masukkan kolom target kedua: s.	Pilih header target untuk setiap kolom target.	Masuk y untuk memilih default atau masuk n .	Pilih jenis padding (hanya untuk sealed kolom). Untuk informasi selengkapnya, lihat (Opsional) Buat skema (pengguna tingkat lanjut) .

Berikut ini adalah dua contoh cara membuat skema enkripsi. Konten yang tepat dari interaksi Anda tergantung pada file input dan tanggapan yang Anda berikan.

Contoh

- [Contoh: Menghasilkan skema enkripsi untuk fingerprint kolom dan kolom cleartext](#)
- [Contoh: Menghasilkan skema enkripsi dengansealed,fingerprint, dan kolom cleartext](#)

Contoh: Menghasilkan skema enkripsi untuk fingerprint kolom dan kolom cleartext

Dalam contoh ini, untuk `ads.csv`, hanya ada dua kolom: `username` dan `ad_variant`. Untuk kolom ini, kami menginginkan yang berikut:

- Untuk `username` kolom yang akan dienkripsi sebagai kolom fingerprint
- Untuk `ad_variant` kolom menjadi `cleartext` kolom

Untuk menghasilkan skema enkripsi untuk fingerprint kolom dan kolom cleartext

1. (Opsional) Untuk memastikan `c3r-cli.jar` file dan file yang akan dienkripsi hadir:
 - a. Arahkan ke direktori yang diinginkan dan jalankan `ls` (jika menggunakan Mac atau Unix/Linux) atau `dir` jika menggunakan Windows).
 - b. Lihat daftar file data tabular (misalnya, `.csv`) dan pilih file untuk dienkripsi.

Dalam contoh ini, `ads.csv` adalah file yang ingin kita enkripsi.

2. Dari CLI, jalankan perintah berikut untuk membuat skema secara interaktif.

```
java -jar c3r-cli.jar schema ads.csv --interactive --output=ads.json
```

Note

- Kau bisa lari `java --jar PATH/T0/c3r-cli.jar`. Atau, jika Anda telah menambahkan `PATH/T0/c3r-cli.jar` ke variabel lingkungan `CLASSPATH` Anda, Anda juga dapat menjalankan nama kelas. Klien enkripsi C3R akan mencari di `CLASSPATH` untuk menemukannya (misalnya, `java com.amazon.psion.cli.Main`).
- `--interactive` Bendera memilih mode interaktif untuk mengembangkan skema. Ini memandu pengguna melalui wizard untuk membuat skema. Pengguna dengan keterampilan tingkat lanjut dapat membuat skema JSON mereka sendiri tanpa menggunakan wizard. Untuk informasi selengkapnya, lihat [\(Opsional\) Buat skema \(pengguna tingkat lanjut\)](#).
- `--output` Bendera menetapkan nama output. Jika Anda tidak menyertakan `--output` bendera, klien enkripsi C3R mencoba memilih nama keluaran default (seperti `<input>.out.csv` atau untuk skema, `<input>.json`).

3. Untuk `Number of target columns from source column 'username'?`, enter **1** dan kemudian tekan Enter.
4. Untuk `Target column type: [c]cleartext, [f]fingerprint, or [s]sealed?`, enter **f** dan kemudian tekan Enter.
5. Untuk `Target column headername <default 'username'>`, tekan Enter.

Nama default 'username' digunakan.

6. Untuk `Add suffix '_fingerprint' to header to indicate how it was encrypted, [y]es or [n]o <default 'yes'>`, enter **y** dan kemudian tekan Enter.

 Note

Mode interaktif menyarankan sufiks untuk ditambahkan ke header kolom terenkripsi (`_fingerprint` untuk kolom dan untuk fingerprint kolom). `_sealed` Sufiks mungkin berguna saat Anda melakukan tugas seperti mengunggah data Layanan AWS atau membuat kolaborasi. AWS Clean Rooms Sufiks ini dapat membantu menunjukkan apa yang dapat dilakukan dengan data terenkripsi di setiap kolom. Misalnya, hal-hal tidak akan berfungsi jika Anda mengenkripsi kolom sebagai sealed kolom (`_sealed`) dan mencoba JOIN melakukannya atau mencoba sebaliknya.

7. Untuk `Number of target columns from source column 'ad_variant'?`, enter **1** dan kemudian tekan Enter.
8. Untuk `Target column type: [c]cleartext, [f]fingerprint, or [s]sealed?`, enter **c** dan kemudian tekan Enter.
9. Untuk `Target column headername <default 'username'>`, tekan Enter.

Nama default 'ad_variant' digunakan.

Skema ditulis ke file baru bernama `aads.json`.

 Note

Anda dapat melihat skema dengan membukanya di editor teks apa pun, seperti Notepad on Windows atau TextEdit on macOS.

10. Anda sekarang siap untuk [mengenkripsi data](#).

Contoh: Menghasilkan skema enkripsi dengan `sealed`, `fingerprint`, dan kolom `cleartext`

Dalam contoh ini, untuk `sales.csv`, ada tiga kolom: `username`, `purchased`, dan `product`. Untuk kolom ini, kami menginginkan yang berikut:

- Untuk `product` kolom menjadi `sealed` kolom
- Untuk `username` kolom yang akan dienkripsi sebagai kolom `fingerprint`
- Untuk `purchased` kolom menjadi `cleartext` kolom

Untuk menghasilkan skema enkripsi dengan `sealed`, `fingerprint`, dan kolom `cleartext`

1. (Opsional) Untuk memastikan `c3r-cli.jar` file dan file yang akan dienkripsi hadir:
 - a. Arahkan ke direktori yang diinginkan dan jalankan `ls` (jika menggunakan Mac atau Unix/Linux) atau `dir` jika menggunakan Windows).
 - b. Lihat daftar file data tabular (`.csv`) dan pilih file untuk dienkripsi.

Dalam contoh ini, `sales.csv` adalah file yang ingin kita enkripsi.

2. Dari CLI, jalankan perintah berikut untuk membuat skema secara interaktif.

```
java -jar c3r-cli.jar schema sales.csv --interactive --  
output=sales.json
```

Note

- `--interactive` Bendera memilih mode interaktif untuk mengembangkan skema. Ini memandu pengguna melalui alur kerja terpandu untuk membuat skema.
- Jika Anda adalah pengguna tingkat lanjut, Anda dapat membuat skema JSON Anda sendiri tanpa menggunakan alur kerja yang dipandu. Untuk informasi selengkapnya, lihat [\(Opsional\) Buat skema \(pengguna tingkat lanjut\)](#).
- Untuk file `.csv` tanpa header kolom, lihat `--noHeaders` tanda untuk perintah skema yang tersedia di CLI.
- `--output` Bendera menetapkan nama output. Jika Anda tidak menyertakan `--output` bendera, klien enkripsi C3R mencoba memilih nama keluaran default (seperti `<input>.out` atau untuk skema, `<input>.json`)

3. Untuk `Number of target columns from source column 'username'?`, enter **1** dan kemudian tekan Enter.
4. Untuk `Target column type: [c]leartext, [f]ingerprint, or [s]ealed?`, enter **f** dan kemudian tekan Enter.
5. Untuk `Target column headername <default 'username'>`, tekan Enter.

Nama default 'username' digunakan.

6. Untuk `Add suffix '_fingerprint' to header to indicate how it was encrypted, [y]es or [n]o <default 'yes'>`, enter **y** dan kemudian tekan Enter.
7. Untuk `Number of target columns from source column 'purchased'?`, enter **1** dan kemudian tekan Enter.
8. Untuk `Target column type: [c]leartext, [f]ingerprint, or [s]ealed?`, enter **c** dan kemudian tekan Enter.
9. Untuk `Target column headername <default 'purchased'>`, tekan Enter.

Nama default 'purchased' digunakan.

10. Untuk `Number of target columns from source column 'product'?`, enter **1** dan kemudian tekan Enter.
11. Untuk `Target column type: [c]leartext, [f]ingerprint, or [s]ealed?`, enter **s** dan kemudian tekan Enter.
12. Untuk `Target column headername <default 'product'>`, tekan Enter.

Nama default 'product' digunakan.

13. Untuk `'product_sealed' padding type: [n]one, [f]ixed, or [m]ax <default 'max'>?`, tekan Enter untuk memilih default.
14. Untuk `Byte-length beyond max length to pad cleartext to in 'product_sealed' <default '0'>?` tekan Enter untuk memilih default.

Skema ditulis ke file baru bernama `sales.json`.

15. Anda sekarang siap untuk [mengkripsi data](#).

Langkah 5: Buat kunci rahasia bersama

Untuk mengkripsi tabel data, peserta kolaborasi harus menyetujui dan berbagi kunci rahasia bersama dengan aman.

Kunci rahasia bersama harus setidaknya 256-bit (32 byte). Anda dapat menentukan kunci yang lebih besar, tetapi itu tidak akan memberi Anda keamanan tambahan.

⚠ Important

Ingat, ID kunci dan kolaborasi yang digunakan untuk enkripsi dan dekripsi harus identik untuk semua peserta kolaborasi.

Bagian berikut memberikan contoh perintah konsol untuk menghasilkan kunci rahasia bersama yang disimpan seperti `secret.key` di direktori kerja terminal masing-masing saat ini.

Topik

- [Contoh: Pembuatan kunci menggunakan OpenSSL](#)
- [Contoh: Pembuatan kunci saat Windows menggunakan PowerShell](#)

Contoh: Pembuatan kunci menggunakan OpenSSL

Untuk pustaka kriptografi tujuan umum umum, jalankan perintah berikut untuk membuat kunci rahasia bersama.

```
openssl rand 32 > secret.key
```

Jika Anda menggunakan Windows dan belum OpenSSL menginstal, Anda dapat membuat kunci menggunakan contoh yang dijelaskan dalam [Contoh: Pembuatan kunci saat Windows menggunakan PowerShell](#).

Contoh: Pembuatan kunci saat Windows menggunakan PowerShell

Untuk PowerShell, aplikasi terminal tersedia Windows, jalankan perintah berikut untuk membuat kunci rahasia bersama.

```
$bs = New-Object Byte[](32);  
[Security.Cryptography.RandomNumberGenerator]::Create().GetBytes($bs); Set-  
Content 'secret.key' -Encoding Byte -Value $bs
```

Langkah 6: Simpan kunci rahasia bersama dalam variabel lingkungan

Variabel lingkungan adalah cara yang nyaman dan dapat diperluas bagi pengguna untuk menyediakan kunci rahasia dari berbagai toko utama seperti AWS Secrets Manager dan meneruskannya ke klien enkripsi C3R.

Klien enkripsi C3R dapat menggunakan kunci yang disimpan Layanan AWS jika Anda menggunakan AWS CLI untuk menyimpan kunci tersebut dalam variabel lingkungan yang relevan. Misalnya, klien enkripsi C3R dapat menggunakan kunci dari AWS Secrets Manager Untuk informasi selengkapnya, lihat [Membuat dan mengelola rahasia AWS Secrets Manager](#) di Panduan AWS Secrets Manager Pengguna.

Note

Namun, sebelum Anda menggunakan Layanan AWS seperti AWS Secrets Manager untuk menahan kunci C3R Anda, verifikasi bahwa kasus penggunaan Anda mengizinkannya. Kasus penggunaan tertentu mungkin mengharuskan kunci ditahan. AWS Ini untuk memastikan bahwa data terenkripsi dan kunci tidak pernah dipegang oleh pihak ketiga yang sama.

Satu-satunya persyaratan untuk kunci rahasia bersama adalah bahwa kunci rahasia bersama base64 dikodekan dan disimpan dalam variabel lingkungan. C3R_SHARED_SECRET

Bagian berikut menjelaskan perintah konsol untuk mengonversi secret . key file ke base64 dan menyimpannya sebagai variabel lingkungan. secret . keyFile bisa saja dihasilkan dari salah satu perintah yang tercantum di [Langkah 5: Buat kunci rahasia bersama](#) dan hanya merupakan sumber contoh.

Simpan kunci dalam variabel lingkungan saat Windows menggunakan PowerShell

Untuk mengonversi ke base64 dan mengatur variabel lingkungan saat Windows menggunakan PowerShell, jalankan perintah berikut.

```
$Bytes=[IO.File]::ReadAllBytes((Get-Location).ToString()+'\secret.key');  
$env:C3R_SHARED_SECRET=[Convert]::ToBase64String($Bytes)
```

Simpan kunci dalam variabel lingkungan pada Linux atau macOS

Untuk mengkonversi ke base64 dan mengatur variabel lingkungan pada Linux atau macOS, jalankan perintah berikut.

```
export C3R_SHARED_SECRET="$(cat secret.key | base64)"
```

Langkah 7: Enkripsi data

Untuk melakukan langkah ini, Anda harus memperoleh ID AWS Clean Rooms kolaborasi dan kunci rahasia bersama. Untuk informasi lebih lanjut, lihat [Prasyarat](#).

Dalam contoh berikut, kita menjalankan enkripsi pada `ads.csv`, menggunakan skema yang kita buat disebut `ads.json`.

Untuk mengenkripsi data

1. Simpan kunci rahasia bersama untuk kolaborasi di [Langkah 6: Simpan kunci rahasia bersama dalam variabel lingkungan](#).
2. Dari baris perintah, masukkan perintah berikut.

```
java -jar c3r-cli.jar encrypt <name of input .csv file> --schema=<name of schema .json file> --id=<collaboration id> --output=<name of output.csv file> <optional flags>
```

3. Untuk `<name of input .csv file>`, masukkan nama file input `.csv`.
4. Untuk `schema=`, masukkan nama file skema enkripsi `.json`.
5. Untuk `id=`, masukkan ID kolaborasi.
6. Untuk `output=`, masukkan nama file output (misalnya, `ads-output.csv`).
7. Sertakan salah satu bendera baris perintah yang dijelaskan dalam [Parameter komputasi kriptografi](#) dan [Bendera opsional dalam Komputasi Kriptografi untuk Clean Rooms](#).
8. Jalankan perintah .

Dalam contoh untuk `ads.csv`, kita menjalankan perintah berikut.

```
java -jar c3r-cli.jar encrypt ads.csv --schema=ads.json --id=123e4567-e89b-42d3-a456-556642440000 --output=ads-output.csv
```

Dalam contoh untuksales.csv, kita menjalankan perintah berikut.

```
java -jar c3r-cli.jar encrypt sales.csv --schema=sales.json --id=123e4567-e89b-42d3-a456-556642440000
```

 Note

Dalam contoh ini, kita tidak menentukan nama file output (--output=**sales-output.csv**). Akibatnya, nama file output default name-of-file.out.csv dihasilkan.

Anda sekarang siap untuk memverifikasi data terenkripsi.

Langkah 8: Verifikasi enkripsi data

Untuk memverifikasi bahwa data dienkripsi

1. Lihat file data terenkripsi (misalnya,sales-output.csv).
2. Verifikasi kolom berikut:
 - a. Kolom 1 - Terenkripsi (misalnya,username_fingerprint).

Untuk fingerprint kolom (HMAC), setelah awalan versi dan tipe (misalnya,01:hmac:), ada 44 karakter data yang disandikan base64.

- b. Kolom 2 - Tidak dienkripsi (misalnya,purchased).
- c. Kolom 3 - Terenkripsi (misalnya,product_sealed).

Untuk kolom enkripsi (SELECT), panjang padding cleartext plus setelah awalan versi dan tipe (misalnya,01:enc:) berbanding lurus dengan panjang yang dienkripsi. cleartext Artinya, panjang adalah ukuran input ditambah sekitar 33 persen overhead karena pengkodean.

Anda sekarang siap untuk:

1. [Unggah data terenkripsi ke S3.](#)
2. [Buat AWS Glue tabel.](#)
3. [Buat tabel yang dikonfigurasi di AWS Clean Rooms.](#)

Klien enkripsi C3R akan membuat file sementara yang tidak berisi data yang tidak terenkripsi (kecuali data itu juga tidak dienkripsi dalam output akhir). Namun, beberapa nilai terenkripsi mungkin tidak diempuk dengan benar. Kolom sidik jari mungkin berisi nilai duplikat, meskipun setelah `allowRepeatedFingerprintValue` kolaborasinya. `false` Masalah ini terjadi karena file sementara ditulis sebelum panjang padding yang tepat dan properti penghapusan duplikat diperiksa.

Jika klien enkripsi C3R gagal atau terputus selama enkripsi, mungkin berhenti setelah menulis file sementara tetapi sebelum memeriksa properti ini dan menghapus file sementara. Oleh karena itu, file-file sementara ini mungkin masih ada di disk. Jika ini masalahnya, konten dalam file-file ini tidak melindungi data plaintext ke tingkat yang sama dengan output. Secara khusus, file-file sementara ini mungkin mengungkapkan data teks biasa ke analisis statistik yang tidak akan bekerja melawan output akhir. Pengguna harus menghapus file-file ini (terutama SQLite database) untuk mencegah file-file ini jatuh ke tangan yang tidak sah.

(Opsional) Buat skema (pengguna tingkat lanjut)

Membuat skema secara manual adalah untuk pengguna tingkat lanjut.

Berikut ini adalah deskripsi format file skema JSON untuk file input dengan atau tanpa header kolom. Pengguna tingkat lanjut dapat langsung menulis atau memodifikasi skema jika diinginkan.

Note

Klien enkripsi C3R dapat membantu Anda dalam membuat skema melalui proses interaktif yang dijelaskan dalam [Contoh: Menghasilkan skema enkripsi dengansealed,fingerprint, dan kolom cleartext](#) atau melalui pembuatan templat rintisan.

Skema tabel yang dipetakan dan posisi

Bagian berikut menjelaskan dua jenis skema tabel:

- Skema tabel yang dipetakan - Skema ini digunakan untuk mengenkripsi file.csv dengan baris header dan file. Apache Parquet
- Skema tabel posisi - Skema ini digunakan untuk mengenkripsi file.csv tanpa baris header.

Klien enkripsi C3R dapat mengenkripsi file tabular untuk kolaborasi. Untuk melakukan ini, ia harus memiliki file skema yang sesuai yang menentukan bagaimana output terenkripsi harus diturunkan dari input.

Klien enkripsi C3R dapat membantu menghasilkan skema untuk INPUT file dengan menjalankan perintah skema klien enkripsi C3R di baris perintah. Contoh dari sebuah perintah adalah `java -jar c3r-cli.jar schema --interactive INPUT`.

Skema menentukan informasi berikut:

1. Kolom sumber mana yang memetakan kolom yang mengubah kolom dalam file output melalui nama tajuknya (skema yang dipetakan) atau posisi (skema posisi)
2. Kolom target mana yang akan tetap cleartext
3. Kolom target mana yang akan dienkripsi untuk kueri SELECT
4. Kolom target mana yang akan dienkripsi untuk kueri JOIN

Informasi ini dikodekan dalam file skema JSON khusus tabel, yang terdiri dari satu objek yang bidangnya `headerRow` adalah nilai Boolean. Nilai harus `true` untuk Parquet file dan file.csv dengan baris header, dan `false` sebaliknya.

Skema tabel yang dipetakan

Skema yang dipetakan memiliki bentuk sebagai berikut.

```
{
  "headerRow": true,
  "columns": [
    {
      "sourceHeader": STRING,
      "targetHeader": STRING,
      "type": TYPE,
      "pad": PAD
    },
    ...
  ]
}
```

Jika `headerRow` yaitu `true`, bidang berikutnya dalam objek adalah `columns`, yang berisi larik skema kolom yang memetakan header sumber ke header target (yaitu, objek JSON yang menjelaskan apa yang harus berisi kolom keluaran).

- `sourceHeader`— Nama STRING header dari kolom sumber tempat datanya berasal.

Note

Kolom sumber yang sama dapat digunakan untuk beberapa kolom target. Kolom dari file input yang tidak terdaftar sebagai di `sourceHeader` mana saja dalam skema tidak muncul di file output.

- `targetHeader`— Nama `STRING` header dari kolom yang sesuai dalam file output.

Note

Bidang ini opsional untuk skema yang dipetakan. Jika bidang ini dihilangkan, `sourceHeader` digunakan kembali untuk nama header dalam output. Entah `_fingerprint` atau `_sealed` ditambahkan jika kolom output masing-masing adalah fingerprint kolom atau sealed kolom.

- `type`— Kolom target dalam file output. `TYPE` Yaitu, salah satu `cleartextsealed`, atau `fingerprint` tergantung pada bagaimana kolom akan digunakan dalam kolaborasi.
- `pad`— Bidang objek skema kolom yang hanya ada saat ada. `TYPE` `sealed` Nilai yang sesuai dari `PAD` adalah objek yang menggambarkan bagaimana data harus empuk sebelum dienkripsi.

```
{
  "type": PAD_TYPE,
  "length": INT
}
```

Untuk menentukan padding pra-enkripsi, `type` dan `length` digunakan sebagai berikut:

- `PAD_TYPE`as `none` — Tidak ada padding yang akan diterapkan pada data kolom dan `length` bidang tidak berlaku (yaitu, dihilangkan).
- `PAD_TYPE`as `fixed` — Data kolom diempuk dengan byte `length` yang ditentukan.
- `PAD_TYPE`as `max` — Data kolom dipadatkan dengan ukuran panjang byte nilai terpanjang ditambah `length` byte tambahan.

Berikut ini adalah contoh skema yang dipetakan, dengan kolom masing-masing jenis.

```
{
  "headerRow": true,
```

```

"columns": [
  {
    "sourceHeader": "FullName",
    "targetHeader": "name",
    "type": "cleartext"
  },
  {
    "sourceHeader": "City",
    "targetHeader": "city_sealed",
    "type": "sealed",
    "pad": {
      "type": "max",
      "length": 16
    }
  },
  {
    "sourceHeader": "PhoneNumber",
    "targetHeader": "phone_number_fingerprint",
    "type": "fingerprint"
  },
  {
    "sourceHeader": "PhoneNumber",
    "targetHeader": "phone_number_sealed",
    "type": "sealed",
    "pad": {
      "type": "fixed",
      "length": 20
    }
  }
]
}

```

Sebagai contoh yang lebih kompleks, berikut ini adalah contoh file.csv dengan header.

```

FirstName,LastName,Address,City,State,PhoneNumber,Title,Level,Notes
Jorge,Souza,12345 Mills Rd,Anytown,SC,703-555-1234,CEO,10,
Paulo,Santos,0 Street,Anytown,MD,404-555-111,CIO,9,This is a really long note that
could really be a paragraph
Mateo,Jackson,1 Two St,Anytown,NY,304-555-1324,C00,9,""
Terry,Whitlock,4 N St,Anytown,VA,407-555-8888,EA,7,Secret notes
Diego,Ramirez,9 Hollows Rd,Anytown,VA,407-555-1222,SDE I,4,null
John,Doe,8 Hollows Rd,Anytown,VA,407-555-4321,SDE I,4,Jane's younger brother
Jane,Doe,8 Hollows Rd,Anytown,VA,407-555-4322,SDE II,5,John's older sister

```


Dalam contoh skema dipetakan berikut, kolom `FirstName` dan `LastName` kolom. `cleartext` StateKolom dienkripsi sebagai `fingerprint` kolom dan sebagai `sealed` kolom dengan padding. `none` Kolom yang tersisa dihilangkan.

```
{
  "headerRow": true,
  "columns": [
    {
      "sourceHeader": "FirstName",
      "targetHeader": "GivenName",
      "type": "cleartext"
    },
    {
      "sourceHeader": "LastName",
      "targetHeader": "Surname",
      "type": "cleartext"
    },
    {
      "sourceHeader": "State",
      "targetHeader": "State_Join",
      "type": "fingerprint"
    },
    {
      "sourceHeader": "State",
      "targetHeader": "State",
      "type": "sealed",
      "pad": {
        "type": "none"
      }
    }
  ]
}
```

Berikut ini adalah file.csv yang dihasilkan dari skema yang dipetakan.

```
givenname,surname,state_fingerprint,state
John,Doe,01:hmac:UK8s8Cn/WR2J0/To2dTxD73aDEe2ZUXeSHy3Tv
+1Mk=,01:enc:FQ3n3Ahv9BQQNWQGcugeHzHYzEZE1vapHa2Uu4SRgSATZ3q0bjPA4TcsHt
+B0kMKBcnHWI13BeGG/SBqmj7vKpI=
Paulo,Santos,01:hmac:CHF4eIrtTNgAooU9v4h9Qjc
+txBnMidQTjdjWuaDTTA=,01:enc:KZ5n5GtaXACco65AXk48BQ02durDNR2ULc4YxmMC8NaZZKKJiksU1IwFadAvV4iBQ1
Mateo,Jackson,01:hmac:iIRnjfNBzryusIJ1w35lgNzeY1RQ1bSfq6PDHW8Xrbk=,01:enc:mLKpS5HIOsgphdEsrzEd
eN9nB02gAbIygt40Fn4La1Yn9Xyj/XUWXlmn8zFe2T4kyDTD8kG0vpQEUGxAUFk=
```

```
Diego,Ramirez,01:hmac:UK8s8Cn/WR2J0/To2dTxD73aDEe2ZUXeSHy3Tv+1Mk=,01:enc:rmZhT98Zm
+IIGw1UTjMIJP4IrW/AAItBLMXcHvnYfRgmWP623VFQ6aUnhsb2MDqEw4G5Uwg5rKKZepUxx5uKbfk=
Jorge,Souza,01:hmac:3BxJdXiFFyZ8HBbYNqqEhBVqhN0d7s2ZiKUe7QiTy08=,01:enc:vVaqWC1VRbhvkf8gnuR7q0z
Terry,Whitlock01:hmac:UK8s8Cn/WR2J0/To2dTxD73aDEe2ZUXeSHy3Tv+1Mk=,01:enc:3c9VEwb0D0/
xbQjdGuccLvI7oZTBdPU+SyrJIyr2kudfAxbuMQ2uRdU/q7rbgyJjxZS8M2U35ILJf/1DgTyg7cM=
Jane,Doe,01:hmac:UK8s8Cn/WR2J0/To2dTxD73aDEe2ZUXeSHy3Tv+1Mk=,01:enc:9RWv46YLveykeNZ/
G0Nd1YFg+AVd0nu05hHyAYTQkPLHnyX+0/jbzD/g9ZT8GCgVE9aB5bV4ooJIXHGBVMXcjrQ=
```

Skema tabel posisi

Skema posisi memiliki bentuk sebagai berikut.

```
{
  "headerRow": false,
  "columns": [
    [
      {
        "targetHeader": STRING,
        "type": TYPE,
        "pad": PAD
      },
      {
        "targetHeader": STRING,
        "type": TYPE,
        "pad": PAD
      }
    ],
    [],
    ...
  ]
}
```

Jika `headerRow` ya `false`, bidang berikutnya dalam objek adalah `columns`, yang berisi array entri. Setiap entri itu sendiri merupakan larik skema kolom posisi nol atau lebih (tanpa `sourceHeader` bidang), yang merupakan objek JSON yang menjelaskan apa yang harus dikandung output.

- `sourceHeader`— Nama `STRING` header dari kolom sumber tempat datanya berasal.

Note

Bidang ini harus dihilangkan dalam skema posisi. Dalam skema posisi, kolom sumber disimpulkan oleh indeks kolom yang sesuai dalam file skema.

- `targetHeader`— Nama STRING header dari kolom yang sesuai dalam file output.

Note

Bidang ini diperlukan untuk skema posisi.

- `type`— Kolom target dalam file output. TYPE Yaitu, salah satu `cleartextsealed`, atau `fingerprint` tergantung pada bagaimana kolom akan digunakan dalam kolaborasi.
- `pad`- Bidang objek skema kolom yang hanya ada saat ada. TYPE `sealed` Nilai yang sesuai dari PAD adalah objek yang menggambarkan bagaimana data harus empuk sebelum dienkripsi.

```
{
  "type": PAD_TYPE,
  "length": INT
}
```

Untuk menentukan padding pra-enkripsi, `type` dan `length` digunakan sebagai berikut:

- `PAD_TYPE`as `none` — Tidak ada padding yang akan diterapkan pada data kolom dan `length` bidang tidak berlaku (yaitu, dihilangkan).
- `PAD_TYPE`as `fixed` — Data kolom diempuk dengan byte `length` yang ditentukan.
- `PAD_TYPE`as `max` — Data kolom dipadatkan dengan ukuran panjang byte nilai terpanjang ditambah `length` byte tambahan.

Note

`fixed` berguna jika Anda tahu sebelumnya batas atas pada ukuran byte data kolom. Kesalahan muncul jika ada data di kolom itu lebih panjang dari yang ditentukan `length`. `max` nyaman ketika ukuran yang tepat dari data input tidak diketahui karena berfungsi terlepas dari ukuran data. Namun, `max` membutuhkan waktu pemrosesan tambahan karena mengenkripsi data dua kali. `max` mengenkripsi data sekali ketika dibaca ke file sementara dan sekali setelah entri data terpanjang di kolom diketahui.

Juga, panjang nilai terpanjang tidak disimpan di antara pemanggilan klien. Jika Anda berencana untuk mengenkripsi data Anda dalam batch, atau mengenkripsi data baru secara berkala, ketahuilah bahwa panjang ciphertext yang dihasilkan dapat bervariasi antar batch.

Berikut ini adalah contoh skema posisi.

```
{
  "headerRow": false,
  "columns": [
    [
      {
        "targetHeader": "name",
        "type": "cleartext"
      }
    ],
    [
      {
        "targetHeader": "city_sealed",
        "type": "sealed",
        "pad": {
          "type": "max",
          "length": 16
        }
      }
    ],
    [
      {
        "targetHeader": "phone_number_fingerprint",
        "type": "fingerprint"
      },
      {
        "targetHeader": "phone_number_sealed",
        "type": "sealed",
        "pad": {
          "type": "fixed",
          "length": 20
        }
      }
    ]
  ]
}
```

```
}
```

Sebagai contoh kompleks, berikut ini adalah contoh file.csv jika tidak memiliki baris pertama dengan header.

```
Jorge,Souza,12345 Mills Rd,Anytown,SC, 703 -555 -1234,CEO, 10,
Paulo,Santos, 0 Street,Anytown,MD, 404-555-111,CIO, 9,This is a really long note that
could really be a paragraph
Mateo,Jackson, 1 Two St,Anytown,NY, 304-555-1324,C00, 9, ""
Terry,Whitlock, 4 N St,Anytown,VA, 407-555-8888,EA, 7,Secret notes
Diego,Ramirez, 9 Hollows Rd,Anytown,VA, 407-555-1222,SDE I, 4,null
John,Doe, 8 Hollows Rd,Anytown,VA, 407-555-4321,SDE I, 4,Jane's younger brother
Jane,Doe, 8 Hollows Rd,Anytown,VA, 407-555-4322,SDE II, 5,John's older sister
```

Skema posisi memiliki bentuk berikut.

```
{
  "headerRow": false,
  "columns": [
    [
      {
        "targetHeader": "GivenName",
        "type": "cleartext"
      }
    ],
    [
      {
        "targetHeader": "Surname",
        "type": "cleartext"
      }
    ],
    [],
    [],
    [
      {
        "targetHeader": "State_Join",
        "type": "fingerprint"
      },
      {
        "targetHeader": "State",
        "type": "sealed",
        "pad": {
          "type": "none"
        }
      }
    ]
  ]
}
```

```

    }
  }
],
[],
[],
[],
[]
]
}

```

Skema sebelumnya menghasilkan file output berikut dengan baris header yang berisi header target yang ditentukan.

```

givenname,surname,state_fingerprint,state
Mateo,Jackson,01: hmac:iIRnjfNBzryusIJ1w35lgNzeY1RQ1bSfq6PDHW8Xrbk=,01: enc:ENS6QD3cMV19vQEGfe9MN
Q8m/Y5SA89dJwKpT5rGPP8e36h6klwDoslpFzGvU0=
Jorge,Souza,01: hmac:3BxJdXiFFyZ8HBbYNqqEhBVqhN0d7s2ZiKUe7QiTy08=,01: enc:LKo0zirq2+
+XEIIIMNRjAsGMdyWUDwYaum0B+IFP+rUf1BNeZDJjtFe1Z+zbZfXQWwJy52Rt7HqvAb2WIK1oMmk=
Paulo,Santos,01: hmac:CHF4eIrtTNgAooU9v4h9Qjc
+txBnMidQTjdjWuaDTTA=,01: enc:MyQKyWxJ9kvK1xDQQtX1UNwv3F+yRBRr0xrUY/1BGg5KFg0n9pK+MZ7g
+ZNqZEPcPz4lht1u0t/wbTaqz0CLXFQ=
Jane,Doe,01: hmac:UK8s8Cn/WR2J0/To2dTxD73aDEe2ZUXeSHy3Tv+1Mk=,01: enc: Pd8sbITBfb0/
ttUB4svVsGoYkDfnDvgkvxzeCi0Yxq54rLSwccy1o3/B50C3cpkkn56dovCwzgmmPNwimCmYtb4=
Terry,Whitlock01: hmac:UK8s8Cn/WR2J0/To2dTxD73aDEe2ZUXeSHy3Tv
+1Mk=,01: enc:Qmtzu3B3GAXKh2KkRYTiEAaMopYedsSdF2e/
ADUiBQ9kv2CxKPzWyYTD3ztmKPMka19dHre5VhUHNp030+j1AQ8=
Diego,Ramirez,01: hmac:UK8s8Cn/WR2J0/To2dTxD73aDEe2ZUXeSHy3Tv+1Mk=,01: enc:ysdg
+GHKdeZrS/geBIoo0EPLHG68MsWpx1dh3xjb+fG5rmFmqUcJLNuuYBHhHA1xchM2WVeV1fmHkBX3mvZNVkc=
John,Doe,01: hmac:UK8s8Cn/WR2J0/To2dTxD73aDEe2ZUXeSHy3Tv+1Mk=,01: enc:9uX0wZu07kAPAx
+Hf6uvQownkWqFSKtWS7gQIJS5aXFquKWCK6yZN0X5Ea2N3bn03Uj1kh0agDWoiP9FRZGJA4=

```

Mendekripsi tabel data dengan klien enkripsi C3R

Ikuti prosedur ini untuk kolaborasi yang menggunakan Cryptographic Computing untuk Clean Rooms dan klien enkripsi C3R untuk mengenkripsi tabel data. Gunakan prosedur ini setelah Anda memiliki [data kueri dalam kolaborasi](#).

Kunci rahasia bersama dan ID kolaborasi diperlukan untuk prosedur ini.

Anggota yang dapat menerima hasil mendekripsi data menggunakan kunci rahasia bersama dan ID kolaborasi yang sama yang digunakan untuk mengenkripsi data untuk kolaborasi.

Note

AWS Clean Rooms kolaborasi sudah membatasi siapa yang dapat melakukan dan melihat hasil kueri. Untuk melakukan dekripsi, siapa pun yang memiliki akses ke hasil ini memerlukan kunci rahasia bersama dan ID kolaborasi yang sama yang digunakan untuk mengenkripsi data.

Untuk mendekripsi tabel data terenkripsi

1. (Opsional) [Lihat perintah yang tersedia di klien enkripsi C3R](#).
2. (Opsional) Arahkan ke direktori yang diinginkan dan jalankan `ls` (macOS) atau `dir` (Windows).
 - Verifikasi bahwa `c3r-cli.jar` file dan file data hasil kueri terenkripsi berada di direktori yang diinginkan.

Note

Jika hasil kueri diunduh dari antarmuka AWS Clean Rooms konsol, kemungkinan ada di folder Unduhan untuk akun pengguna Anda. (Misalnya, folder Unduhan di direktori pengguna Anda di Windows and macOS.) Kami menyarankan Anda memindahkan file hasil kueri ke folder yang sama dengan `c3r-cli.jar`.

3. Simpan kunci rahasia bersama dalam variabel `C3R_SHARED_SECRET` lingkungan. Untuk informasi selengkapnya, lihat [Langkah 6: Simpan kunci rahasia bersama dalam variabel lingkungan](#).
4. Dari AWS Command Line Interface (AWS CLI), jalankan perintah berikut.

```
java -jar c3r-cli.jar decrypt <name of input .csv file> --id=<collaboration id> --  
output=<output file name>
```

5. Ganti masing-masing *user input placeholder* dengan informasi Anda sendiri:
 - a. Untuk `id=`, masukkan ID kolaborasi.
 - b. Untuk `output=`, masukkan nama file output (misalnya, `results-decrypted.csv`).

Jika Anda tidak menentukan nama output, nama default ditampilkan di terminal.

- c. Lihat data yang didekripsi dalam file keluaran yang ditentukan menggunakan CSV pilihan Anda atau Parquet melihat aplikasi (seperti Microsoft Excel, editor teks, atau aplikasi lain).

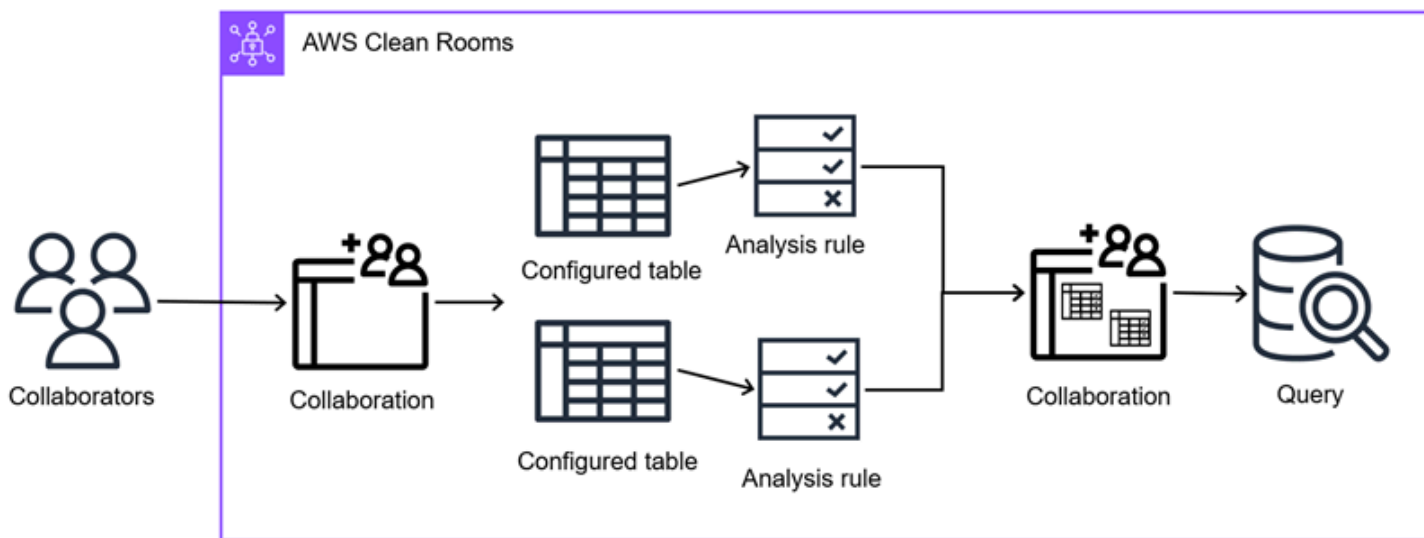
Tabel yang dikonfigurasi di AWS Clean Rooms

Tabel yang dikonfigurasi adalah referensi ke tabel yang ada di sumber data. Ini berisi aturan analisis yang menentukan bagaimana data dapat ditanyakan AWS Clean Rooms dan dapat menyertakan anggaran akses data untuk mengontrol penggunaan tabel. Tabel yang dikonfigurasi dapat dikaitkan dengan satu atau lebih kolaborasi.

Dengan AWS Clean Rooms, Anda dapat melakukan analisis agregasi pada data peristiwa, seperti jumlah pembelian dibandingkan dengan jumlah pembelian. Anda juga dapat melakukan analisis daftar pada data peristiwa, seperti memperkaya data pelanggan yang tumpang tindih dari data segmen ke data CRM. Anda juga dapat melakukan kueri khusus dan mengatur privasi diferensial pada data peristiwa, seperti data pemirsa dan atribut segmen. Untuk salah satu jenis analisis ini, Anda dapat mengatur anggaran akses data untuk memantau dan mengontrol berapa banyak data Anda diakses melalui kueri.

Pertama, Anda membuat kolaborasi AWS Clean Rooms dan menambahkan yang ingin Akun AWS Anda undang, atau bergabung dengan kolaborasi yang Anda undang dengan membuat keanggotaan. Selanjutnya, Anda dan anggota lain dalam kolaborasi membuat tabel yang dikonfigurasi. Anda berdua menambahkan aturan analisis ke tabel yang dikonfigurasi (agregasi, daftar, atau kustom) dan secara opsional mengatur anggaran akses data. Kemudian, Anda mengaitkan tabel yang dikonfigurasi dengan kolaborasi. Terakhir, anggota yang dapat melakukan kueri menjalankan kueri di dua tabel data, menggunakan anggaran akses data saat kueri dijalankan.

Diagram berikut merangkum cara bekerja dengan data peristiwa di AWS Clean Rooms.



- [Membuat tabel yang dikonfigurasi di AWS Clean Rooms](#)
- [Menambahkan aturan analisis ke tabel yang dikonfigurasi](#)
- [Mengaitkan tabel terkonfigurasi ke kolaborasi](#)
- [Mengonfigurasi anggaran akses data](#)
- [Menambahkan aturan analisis kolaborasi ke tabel yang dikonfigurasi](#)
- [Mengkonfigurasi kebijakan privasi diferensial \(opsional\)](#)
- [Melihat tabel dan aturan analisis](#)
- [Mengedit tabel yang dikonfigurasi](#)
- [Mengedit tag tabel yang dikonfigurasi](#)
- [Mengedit aturan analisis tabel yang dikonfigurasi](#)
- [Menghapus aturan analisis tabel yang dikonfigurasi](#)
- [Kolom tabel yang dikonfigurasi tidak diizinkan](#)
- [Mengedit asosiasi tabel yang dikonfigurasi](#)
- [Memutuskan tabel yang dikonfigurasi](#)

Membuat tabel yang dikonfigurasi di AWS Clean Rooms

Tabel yang dikonfigurasi adalah referensi ke tabel yang ada di sumber data. Ini berisi aturan analisis yang menentukan bagaimana data dapat ditanyakan. AWS Clean Rooms Tabel yang dikonfigurasi dapat dikaitkan dengan satu atau lebih kolaborasi.

Untuk informasi tentang cara membuat tabel yang dikonfigurasi menggunakan AWS SDKs, lihat [Referensi AWS Clean Rooms API](#).

Topik

- [Membuat tabel yang dikonfigurasi - sumber data Amazon S3](#)
- [Membuat tabel yang dikonfigurasi - sumber data Amazon Athena](#)
- [Membuat tabel yang dikonfigurasi — Sumber data Snowflake](#)

Membuat tabel yang dikonfigurasi - sumber data Amazon S3

Dalam prosedur ini, [anggota](#) melakukan tugas-tugas berikut:

- Mengkonfigurasi AWS Glue tabel yang ada untuk digunakan di AWS Clean Rooms. (Langkah ini dapat dilakukan sebelum atau sesudah bergabung dengan kolaborasi, kecuali menggunakan Cryptographic Computing untuk Clean Rooms.)

 Note

AWS Clean Rooms mendukung AWS Glue tabel. Untuk informasi selengkapnya tentang memasukkan data Anda AWS Glue, lihat [Langkah 3: Unggah tabel data Anda ke Amazon S3](#).

- Beri nama [tabel yang dikonfigurasi](#) dan pilih kolom mana yang akan digunakan dalam kolaborasi.

Prosedur berikut mengasumsikan bahwa:

- Anggota kolaborasi telah [mengunggah tabel data mereka ke Amazon S3](#) dan [membuat AWS Glue tabel](#).

 Note

Tujuan Hasil di Amazon S3 tidak dapat berada dalam bucket S3 yang sama dengan sumber data apa pun.

- (Opsional) Hanya untuk tabel data [terenkripsi](#), anggota kolaborasi telah [menyiapkan tabel data terenkripsi menggunakan klien enkripsi C3R](#).

Anda dapat menggunakan generasi statistik yang disediakan oleh AWS Glue untuk menghitung statistik tingkat kolom untuk tabel. AWS Glue Data Catalog Setelah AWS Glue menghasilkan statistik untuk tabel di Katalog Data, Amazon Redshift Spectrum secara otomatis menggunakan statistik tersebut untuk mengoptimalkan paket kueri. Untuk informasi selengkapnya tentang menggunakan statistik tingkat kolom komputasi AWS Glue, lihat [Mengoptimalkan performa kueri menggunakan statistik kolom](#) di Panduan Pengguna AWS Glue. Untuk informasi selengkapnya AWS Glue, lihat [Panduan Pengembang AWS Glue](#).

Untuk membuat tabel yang dikonfigurasi - sumber data Amazon S3

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Tabel.

3. Di sudut kanan atas, pilih Konfigurasi tabel baru.
4. Untuk Sumber data, di bawah sumber AWS data, pilih Amazon S3.
5. Di bawah tabel Amazon S3:
 - a. Pilih Wilayah tempat tabel S3 di-host.

Secara default, Wilayah saat ini (seperti N. Virginia us-east-1) dipilih.

 Warning

Jika sumber data Amazon S3 Anda berada di Wilayah yang berbeda dari lokasi pemrosesan Anda, pemrosesan data dapat terjadi sementara di luar Wilayah sumber. Sebelum melanjutkan, verifikasi bahwa pergerakan data lintas wilayah mematuhi persyaratan kedaulatan data Anda, kebijakan kepatuhan terhadap peraturan, dan standar tata kelola data.

Untuk informasi selengkapnya tentang Wilayah, lihat [Wilayah dan Titik Akhir](#) di Referensi Umum AWS

- b. Pilih Database dari daftar dropdown.
- c. Pilih Tabel yang ingin Anda konfigurasi dari daftar dropdown.

 Note

Untuk memverifikasi bahwa ini adalah tabel yang benar, lakukan salah satu dari yang berikut:

- Pilih Lihat di AWS Glue.
- Aktifkan Lihat skema dari AWS Glue untuk melihat skema.

 Important

Untuk AWS Glue tabel di mana data dalam format CSV, nama kolom dan urutan dalam skema Glue harus sama persis dengan data CSV. Jika tidak sejajar, daftar kolom yang diizinkan untuk tabel yang dikonfigurasi mungkin tidak diterapkan dengan benar.

6. Untuk Kolom dan metode analisis yang diizinkan dalam kolaborasi,
 - a. Untuk kolom mana yang ingin Anda izinkan dalam kolaborasi?
 - Pilih Semua kolom untuk memungkinkan semua kolom ditanyakan dalam kolaborasi.
 - Pilih Daftar kustom untuk mengizinkan satu atau beberapa kolom dari daftar tarik-turun. Tentukan kolom yang diizinkan untuk ditanyakan dalam kolaborasi.
 - b. Untuk metode analisis yang diizinkan,
 - i. Pilih Kueri langsung untuk memungkinkan kueri SQL dijalankan langsung di tabel ini
 - ii. Pilih Pekerjaan langsung untuk memungkinkan PySpark pekerjaan dijalankan langsung di tabel ini.

Example Contoh

Misalnya, jika Anda ingin mengizinkan anggota kolaborasi menjalankan kueri SQL langsung dan PySpark pekerjaan di semua kolom, pilih Semua kolom, Kueri langsung, dan Pekerjaan langsung.

7. Untuk detail tabel yang Dikonfigurasi,
 - a. Masukkan Nama untuk tabel yang dikonfigurasi.

Anda dapat menggunakan nama default atau mengganti nama tabel ini.
 - b. Masukkan Deskripsi tabel.

Deskripsi membantu membedakan antara tabel lain yang dikonfigurasi dengan nama yang mirip.
8. Jika Anda ingin mengaktifkan Tag untuk sumber daya tabel yang dikonfigurasi, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.
9. Pilih Konfigurasi tabel baru.

Sekarang setelah Anda membuat tabel yang dikonfigurasi, Anda siap untuk:

- [Tambahkan aturan analisis ke tabel yang dikonfigurasi](#)
- [Kaitkan tabel yang dikonfigurasi ke kolaborasi](#)

Membuat tabel yang dikonfigurasi - sumber data Amazon Athena

Opsi sumber data Amazon Athena memungkinkan Anda untuk menanyakan data yang disimpan di Amazon S3, dikatalogkan dalam katalog data atau katalog gabungan, AWS Glue dan akses yang dikontrol melalui AWS Lake Formation Kedua tabel dan AWS Glue Data Catalog Tampilan didukung. Tautan sumber daya Lake Formation dapat digunakan untuk berbagi tabel dan tampilan di seluruh Akun AWS dan di seberang Wilayah AWS ke akun AWS Clean Rooms anggota yang menggabungkannya ke AWS Clean Rooms kolaborasi.

Note

Hanya kumpulan data berbasis Amazon S3 yang dapat ditanyakan melalui integrasi sumber data Athena.

Dalam prosedur ini, [anggota](#) melakukan tugas-tugas berikut:

- Mengkonfigurasi tabel atau tampilan yang ada di AWS Glue Data Catalog untuk digunakan AWS Clean Rooms
- Beri nama [tabel yang dikonfigurasi](#) dan pilih kolom mana yang akan digunakan dalam kolaborasi.

Prosedur berikut mengasumsikan bahwa:

- Anggota kolaborasi telah membuat AWS Glue Data Catalog database dan tabel atau tampilan GDC.

Untuk membuat tabel yang dikonfigurasi — sumber data Athena

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Tabel.
3. Di sudut kanan atas, pilih Konfigurasi tabel baru.
4. Untuk Sumber data, di bawah sumber AWS data, pilih Amazon Athena.
5. Di bawah tabel Amazon Athena:
 - a. Pilih Wilayah tempat tabel Amazon Athena di-host.

Secara default, Wilayah saat ini (seperti N. Virginia us-east-1) dipilih.

 Warning

Jika sumber data Amazon Athena Anda berada di Wilayah yang berbeda dari lokasi pemrosesan Anda, pemrosesan data dapat terjadi sementara di luar Wilayah sumber. Sebelum melanjutkan, verifikasi bahwa pergerakan data lintas wilayah mematuhi persyaratan kedaulatan data Anda, kebijakan kepatuhan terhadap peraturan, dan standar tata kelola data.

Untuk informasi selengkapnya tentang Wilayah, lihat [Wilayah dan Titik Akhir](#) di Referensi Umum AWS

- b. Pilih Katalog dari daftar dropdown.

Secara default, Katalog AWS Glue Data dipilih.

- AWS Glue Katalog Data — Katalog default untuk tabel di AWS Glue.
- Katalog federasi - Tersedia jika Anda telah mengonfigurasi Federasi AWS Glue Katalog untuk terhubung ke katalog Apache Iceberg REST jarak jauh. Untuk informasi selengkapnya, lihat [Federasi katalog](#) di Panduan AWS Lake Formation Pengembang.

- c. Pilih Database dari daftar dropdown.
- d. Pilih Tabel yang ingin Anda konfigurasi dari daftar dropdown.

 Note

Untuk memverifikasi bahwa ini adalah tabel yang benar, lakukan salah satu dari yang berikut:

- Pilih Lihat di AWS Glue atau Lihat di AWS Lake Formation (tergantung pada jenis katalog Anda).
- Aktifkan Lihat skema dari AWS Glue untuk melihat skema.

6. Untuk konfigurasi Amazon Athena,

- a. Pilih Workgroup dari daftar dropdown.

- b. Untuk lokasi keluaran S3, pilih tindakan yang disarankan, berdasarkan salah satu skenario berikut.

Skenario	Tindakan yang disarankan
Workgroup Anda tidak memiliki lokasi output default.	Masukkan lokasi output S3 atau pilih Browse S3.
Workgroup memberlakukan lokasi output default Anda.	Lokasi output S3 dipilih secara otomatis dan Anda tidak dapat mengubahnya.
Workgroup Anda tidak menerapkan lokasi output default Anda.	Masukkan lokasi output S3 atau pilih Browse S3.

7. Untuk Kolom yang diizinkan dalam kolaborasi, pilih opsi berdasarkan tujuan Anda.

Tujuan Anda	Opsi yang disarankan
Izinkan semua kolom untuk digunakan AWS Clean Rooms (tunduk pada aturan analisis)	Semua kolom
Izinkan satu atau beberapa kolom dari daftar dropdown Tentukan kolom yang diizinkan	Daftar kustom

8. Untuk detail tabel yang Dikonfigurasi,

- a. Masukkan Nama untuk tabel yang dikonfigurasi.

Anda dapat menggunakan nama default atau mengganti nama tabel ini.

- b. Masukkan Deskripsi tabel.

Deskripsi membantu membedakan antara tabel lain yang dikonfigurasi dengan nama yang mirip.

- c. Jika Anda ingin mengaktifkan Tag untuk sumber daya tabel yang dikonfigurasi, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.

9. Pilih Konfigurasi tabel baru.

Sekarang setelah Anda membuat tabel yang dikonfigurasi, Anda siap untuk:

- [Tambahkan aturan analisis ke tabel yang dikonfigurasi](#)
- [Kaitkan tabel yang dikonfigurasi ke kolaborasi](#)

Membuat tabel yang dikonfigurasi — Sumber data Snowflake

Dalam prosedur ini, [anggota](#) melakukan tugas-tugas berikut:

- Mengkonfigurasi tabel Snowflake yang ada untuk digunakan di. AWS Clean Rooms(Langkah ini dapat dilakukan sebelum atau sesudah bergabung dengan kolaborasi, kecuali menggunakan Cryptographic Computing untukClean Rooms.)
- Beri nama [tabel yang dikonfigurasi](#) dan pilih kolom mana yang akan digunakan dalam kolaborasi.

Prosedur berikut mengasumsikan bahwa:

- Anggota kolaborasi telah mengunggah tabel data mereka ke Snowflake.
- (Opsional) Hanya untuk tabel data [terenkripsi](#), anggota kolaborasi telah [menyiapkan tabel data terenkripsi menggunakan klien enkripsi](#) C3R.

Untuk membuat tabel yang dikonfigurasi — Sumber data Snowflake

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Tabel.
3. Di sudut kanan atas, pilih Konfigurasi tabel baru.
4. Untuk sumber data, di bawah awan pihak ketiga dan sumber data, pilih Snowflake.
5. Tentukan kredensial Snowflake menggunakan ARN rahasia yang ada atau menyimpan rahasia baru untuk tabel ini.

Use existing secret ARN

1. Jika Anda memiliki ARN rahasia, masukkan di bidang Rahasia ARN.

Anda dapat mencari ARN rahasia Anda dengan memilih Go to. AWS Secrets Manager

2. Jika Anda memiliki rahasia yang ada dari tabel lain, pilih Impor rahasia ARN dari tabel yang ada.

 Note

Rahasia ARN bisa lintas akun.

Store a new secret for this table

1. Masukkan kredensi Snowflake berikut:
 - Nama pengguna Snowflake
 - Gudang kepingan salju
 - Peran kepingan salju
 - Kunci pribadi Snowflake Privacy Enhanced Mail (PEM)
2. Untuk enkripsi, lakukan salah satu hal berikut:
 - Untuk menggunakan Kunci yang dikelola AWS (default), biarkan kotak centang Sesuaikan pengaturan enkripsi dihapus.
 - Untuk menggunakan kustom AWS KMS key:
 - Pilih kotak centang Sesuaikan pengaturan enkripsi.
 - Untuk tombol KMS, masukkan kunci ARN atau pilih salah satu dari daftar.
3. Masukkan nama Rahasia untuk membantu Anda menemukan kredensialnya nanti.
6. Untuk detail tabel dan skema kepingan salju, masukkan detail secara manual atau impor detailnya secara otomatis.

Enter the details manually

1. Masukkan pengidentifikasi akun Snowflake.

Untuk informasi selengkapnya, lihat [Pengenal akun](#) di dokumentasi Snowflake.

Pengidentifikasi akun Anda harus dalam format yang digunakan untuk driver Snowflake. Anda perlu mengganti periode (.) dengan tanda hubung (-) sehingga pengenal diformat sebagai. **<orgname>-<account_name>**

2. Masukkan database Snowflake.

Untuk informasi selengkapnya, lihat [database Snowflake di dokumentasi](#) Snowflake.

3. Masukkan nama skema Snowflake.

4. Masukkan nama tabel Snowflake.

Untuk informasi lebih lanjut, lihat [Memahami Struktur Tabel Kepingan Salju dalam dokumentasi](#) Kepingan Salju.

5. Untuk Skema, masukkan nama Kolom dan pilih tipe Data dari daftar dropdown.

6. Pilih Tambahkan kolom untuk menambahkan lebih banyak kolom.

- Jika Anda memilih tipe data Object, tentukan skema Object.

Example Contoh skema objek

```
name STRING,  
location OBJECT(  
  x INT,  
  y INT,  
  metadata OBJECT(uuid STRING)  
),  
history ARRAY(TEXT)
```

- Jika Anda memilih tipe data Array, tentukan skema Array.

Example Contoh skema array

```
OBJECT(x INT, y INT)
```

- Jika Anda memilih tipe data Peta, tentukan skema Peta.

Example Contoh skema peta

```
STRING, OBJECT(x INT, y INT)
```

Automatically import the details

1. Ekspor tampilan KOLOM Anda dari Snowflake sebagai file CSV.

Untuk informasi selengkapnya tentang tampilan Snowflake COLUMNS, lihat tampilan [COLUMNS di dokumentasi](#) Snowflake.

2. Pilih Impor dari file untuk mengimpor file CSV dan menentukan informasi tambahan apa pun.

Nama database, nama skema, nama tabel, nama kolom dan tipe data secara otomatis diimpor.

- Jika Anda memilih tipe data Object, tentukan skema Object.
- Jika Anda memilih tipe data Array, tentukan skema Array.
- Jika Anda memilih tipe data Peta, tentukan skema Peta.

3. Masukkan pengidentifikasi akun Snowflake.

Untuk informasi selengkapnya, lihat [Pengenalan akun](#) di dokumentasi Snowflake.

Note

Hanya tabel S3 yang dikatalogkan AWS Glue dapat digunakan untuk mengambil skema tabel secara otomatis.

7. Untuk Kolom yang diizinkan dalam kolaborasi, pilih opsi berdasarkan tujuan Anda.

Tujuan Anda	Opsi yang disarankan
Izinkan semua kolom untuk digunakan AWS Clean Rooms (tunduk pada aturan analisis)	Semua kolom
Izinkan satu atau beberapa kolom dari daftar dropdown Tentukan kolom yang diizinkan	Daftar kustom

8. Untuk detail tabel yang Dikonfigurasi,

a. Masukkan Nama untuk tabel yang dikonfigurasi.

Anda dapat menggunakan nama default atau mengganti nama tabel ini.

b. Masukkan Deskripsi tabel.

Deskripsi membantu membedakan antara tabel lain yang dikonfigurasi dengan nama yang mirip.

c. Jika Anda ingin mengaktifkan Tag untuk sumber daya tabel yang dikonfigurasi, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.

9. Pilih Konfigurasi tabel baru.

Sekarang setelah Anda membuat tabel yang dikonfigurasi, Anda siap untuk:

- [Tambahkan aturan analisis ke tabel yang dikonfigurasi](#)
- [Kaitkan tabel yang dikonfigurasi ke kolaborasi](#)

Menambahkan aturan analisis ke tabel yang dikonfigurasi

Bagian berikut menjelaskan cara menambahkan aturan analisis ke tabel yang dikonfigurasi. Dengan menentukan aturan analisis, Anda dapat mengotorisasi anggota yang dapat melakukan kueri untuk menjalankan kueri yang cocok dengan aturan analisis tertentu yang didukung oleh AWS Clean Rooms

AWS Clean Rooms mendukung jenis aturan analisis berikut:

- [Aturan analisis agregasi](#)
- [Aturan analisis daftar](#)
- [Aturan analisis kustom di AWS Clean Rooms](#)

Hanya ada satu aturan analisis per tabel yang dikonfigurasi. Anda dapat mengonfigurasi aturan analisis kapan saja sebelum mengaitkan tabel yang dikonfigurasi dengan kolaborasi.

Important

Jika Anda menggunakan Komputasi Kriptografi untuk Clean Rooms dan memiliki tabel data terenkripsi dalam kolaborasi, aturan analisis yang Anda tambahkan ke tabel terkonfigurasi terenkripsi harus konsisten dengan cara data dienkripsi. Misalnya, jika Anda mengenkripsi data untuk SELECT (aturan analisis agregasi), Anda tidak boleh menambahkan aturan analisis untuk JOIN (aturan analisis daftar).

Topik

- [Menambahkan aturan analisis agregasi ke tabel \(aliran terpandu\)](#)
- [Menambahkan aturan analisis daftar ke tabel \(alur terpandu\)](#)
- [Menambahkan aturan analisis kustom ke tabel \(alur terpandu\)](#)
- [Menambahkan aturan analisis ke tabel \(editor JSON\)](#)

- [Langkah selanjutnya](#)

Menambahkan aturan analisis agregasi ke tabel (aliran terpandu)

Aturan analisis agregasi memungkinkan kueri yang mengumpulkan statistik tanpa mengungkapkan informasi tingkat baris menggunakan, COUNTSUM, dan berfungsi di sepanjang dimensi opsional. AVG

Prosedur ini menjelaskan proses penambahan aturan analisis agregasi ke tabel yang dikonfigurasi dengan menggunakan opsi Aliran terpandu di AWS Clean Rooms konsol.

Note

Tabel yang dikonfigurasi menggunakan sumber data non-S3 hanya mendukung aturan [analisis khusus](#).

Untuk menambahkan aturan analisis agregasi ke tabel (aliran terpandu)

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Tabel.
3. Pilih tabel yang dikonfigurasi.
4. Pada halaman detail tabel yang dikonfigurasi, pilih Konfigurasi aturan analisis.
5. Di bawah Langkah 1: Pilih jenis aturan analisis, di bawah Jenis aturan analisis, pilih Agregasi pilihan.
6. Di bawah Metode pembuatan, pilih Alur terpandu, lalu pilih Berikutnya.
7. Di bawah Langkah 2: Tentukan kontrol kueri, untuk fungsi Agregat:
 - a. Pilih fungsi Agregat dari dropdown:
 - MENGHITUNG
 - HITUNG BERBEDA
 - SUM
 - JUMLAH BERBEDA
 - AVG

- b. Pilih kolom mana yang dapat digunakan dalam fungsi Agregat dari dropdown Kolom.
- c. (Opsional) Pilih Tambahkan fungsi lain untuk menambahkan fungsi agregat lain dan mengaitkan satu atau beberapa kolom ke fungsi itu.

 Note

Setidaknya diperlukan satu fungsi agregat.

- d. (Opsional) Pilih Hapus untuk menghapus fungsi agregat.
8. Untuk kontrol Gabung,
- a. Pilih satu opsi untuk Izinkan tabel yang akan ditanyakan dengan sendirinya:

Jika Anda memilih...	Lalu...
Tidak, hanya tumpang tindih yang dapat ditanyakan	Tabel dapat ditanyakan hanya ketika bergabung ke tabel yang dimiliki oleh anggota yang dapat melakukan kueri.
Ya	Tabel dapat ditanyakan dengan sendirinya atau ketika bergabung ke tabel lain.

- b. Di bawah Tentukan kolom gabungan, pilih kolom yang ingin Anda izinkan untuk digunakan dalam INNER JOIN pernyataan.

Ini opsional jika Anda telah memilih Ya di langkah sebelumnya.

- c. Di bawah Tentukan operator yang diizinkan untuk pencocokan, pilih operator mana, jika ada, yang dapat digunakan untuk pencocokan pada beberapa kolom gabungan. Jika Anda memilih dua atau lebih JOIN kolom, salah satu operator ini diperlukan.

Jika Anda memilih...	Lalu...
DAN	Anda dapat memasukkan AND dalam kondisi INNER JOIN pertandingan untuk menggabungkan satu kolom ke kolom lain di antara tabel.

Jika Anda memilih...	Lalu...
ATAU	Anda dapat memasukkan OR dalam kondisi INNER JOIN kecocokan untuk menggabungkan beberapa kecocokan kolom antar tabel. Operator logis ini berguna untuk mendapatkan tingkat kecocokan yang lebih tinggi.

9. (Opsional) Untuk kontrol Dimensi, dalam menu tarik-turun Tentukan kolom dimensi, pilih kolom mana yang ingin Anda izinkan untuk digunakan dalam pernyataan SELECT, dan WHERE, GROUPBY, dan ORDER BY bagian dari kueri.

 Note

Fungsi agregat atau kolom gabungan tidak dapat digunakan sebagai kolom Dimensi.

10. Untuk fungsi Skalar, pilih satu opsi untuk Fungsi skalar mana yang ingin Anda izinkan?

Jika Anda memilih...	Lalu...
Semua saat ini didukung oleh AWS Clean Rooms	Anda mengizinkan semua fungsi skalar yang saat ini didukung oleh AWS Clean Rooms. Anda dapat memilih Lihat daftar untuk melihat seluruh daftar fungsi Skalar yang didukung. AWS Clean Rooms
Daftar kustom	Anda dapat menyesuaikan fungsi skalar mana yang akan diizinkan. Pilih satu atau beberapa opsi dari menu tarik-turun Tentukan fungsi skalar yang diizinkan.
Tidak ada	Anda tidak ingin mengizinkan fungsi skalar apa pun.

Untuk informasi selengkapnya, lihat [Fungsi skalar](#).

11. Pilih Berikutnya.
12. Di bawah Langkah 3: Tentukan kontrol hasil kueri, untuk kendala Agregasi:
 - a. Pilih daftar dropdown untuk setiap nama Kolom.
 - b. Pilih daftar dropdown untuk setiap Jumlah minimum nilai berbeda yang harus dipenuhi untuk setiap baris output yang akan dikembalikan, setelah COUNT DISTINCT fungsi diterapkan padanya.
 - c. Pilih Tambahkan kendala untuk menambahkan lebih banyak kendala agregasi.
 - d. (Opsional) Pilih Hapus untuk menghapus kendala agregasi.
13. Untuk analisis tambahan yang diterapkan pada output, pilih opsi berdasarkan tujuan Anda.

Tujuan Anda	Opsi yang disarankan
Izinkan hanya kueri langsung pada tabel ini. Tolak analisis tambahan agar tidak dijalankan pada hasil kueri. Tabel hanya dapat digunakan untuk query langsung.	Tidak diizinkan
Izinkan tetapi tidak memerlukan kueri langsung dan analisis tambahan pada tabel ini.	Diizinkan
Mengharuskan tabel hanya dapat digunakan dalam kueri langsung yang diproses dengan salah satu analisis tambahan yang diperlukan. Kueri langsung pada tabel ini harus diproses lebih lanjut sebelum dapat dikembalikan.	Diperlukan

14. Pilih Berikutnya.
15. Di bawah Langkah 4: Tinjau dan konfigurasi, tinjau pilihan yang telah Anda buat untuk langkah sebelumnya, edit jika perlu, lalu pilih Konfigurasi aturan analisis.

Anda melihat pesan konfirmasi bahwa Anda telah berhasil mengonfigurasi aturan analisis agregasi ke tabel.

Menambahkan aturan analisis daftar ke tabel (alur terpandu)

Aturan analisis daftar memungkinkan kueri yang menampilkan daftar tingkat baris tumpang tindih antara tabel terkait dan tabel anggota yang dapat melakukan kueri.

Prosedur ini menjelaskan proses menambahkan aturan analisis daftar ke tabel yang dikonfigurasi menggunakan opsi Aliran terpandu di AWS Clean Rooms konsol.

Note

Tabel yang dikonfigurasi menggunakan sumber data non-S3 hanya mendukung aturan [analisis khusus](#).

Untuk menambahkan aturan analisis daftar ke tabel (alur terpandu)

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Tabel.
3. Pilih tabel yang dikonfigurasi.
4. Pada halaman detail tabel yang dikonfigurasi, pilih Konfigurasi aturan analisis.
5. Di bawah Langkah 1: Pilih jenis aturan analisis, di bawah Jenis aturan analisis, pilih Daftar pilihan.
6. Di bawah Metode pembuatan, pilih Alur terpandu, lalu pilih Berikutnya.
7. Di bawah Langkah 2: Tentukan kontrol kueri, untuk kontrol Gabung:
 - a. Di bawah Tentukan kolom gabungan, pilih kolom yang ingin Anda izinkan untuk digunakan dalam INNER JOIN pernyataan.
 - b. Di bawah Tentukan operator yang diizinkan untuk pencocokan, pilih operator mana, jika ada, yang dapat digunakan untuk pencocokan pada beberapa kolom gabungan. Jika Anda memilih dua atau lebih JOIN kolom, salah satu operator ini diperlukan.

Jika Anda memilih...	Lalu...
DAN	Anda dapat memasukkan AND dalam kondisi INNER JOIN pertandingan untuk menggabungkan satu kolom ke kolom lain di antara tabel.
ATAU	Anda dapat memasukkan OR dalam kondisi INNER JOIN kecocokan untuk menggabungkan beberapa kecocokan kolom antar tabel. Operator logis ini berguna untuk mendapatkan tingkat kecocokan yang lebih tinggi.

8. (Opsional) Untuk kontrol Daftar, dalam menu tarik-turun Tentukan kolom daftar, pilih kolom mana yang ingin Anda izinkan untuk digunakan dalam output kueri (yaitu, digunakan dalam SELECT pernyataan), atau digunakan untuk memfilter hasil (yaitu, WHERE pernyataan).
9. Pilih Berikutnya.
10. Di bawah Langkah 3: Tentukan kontrol hasil kueri, untuk Analisis tambahan yang diterapkan pada output, pilih opsi berdasarkan tujuan Anda.

Tujuan Anda	Opsi yang disarankan
Izinkan hanya kueri langsung pada tabel ini. Tolak analisis tambahan agar tidak dijalankan pada hasil kueri. Tabel hanya dapat digunakan untuk query langsung.	Tidak diizinkan
Izinkan tetapi tidak memerlukan kueri langsung dan analisis tambahan pada tabel ini.	Diizinkan
Mengharuskan tabel hanya dapat digunakan dalam kueri langsung yang diproses dengan salah satu analisis tambahan yang diperlukan. Kueri langsung pada tabel ini harus	Diperlukan

Tujuan Anda	Opsi yang disarankan
diproses lebih lanjut sebelum dapat dikembalikan.	

- Di bawah Langkah 4: Tinjau dan konfigurasi, tinjau pilihan yang telah Anda buat untuk langkah sebelumnya, edit jika perlu, lalu pilih Konfigurasi aturan analisis.

Anda melihat pesan konfirmasi bahwa Anda telah berhasil mengonfigurasi aturan analisis daftar untuk tabel.

Menambahkan aturan analisis kustom ke tabel (alur terpandu)

Aturan analisis kustom memungkinkan kueri SQL kustom atau PySpark pekerjaan pada tabel dikonfigurasi. Aturan analisis kustom diperlukan jika Anda menggunakan:

- [Template analisis](#) untuk memungkinkan serangkaian kueri atau PySpark pekerjaan SQL yang telah disetujui sebelumnya atau kumpulan akun tertentu yang dapat memberikan kueri yang menggunakan data Anda.
- [AWS Clean Rooms Privasi Diferensial](#) untuk melindungi terhadap upaya identifikasi pengguna.
- Sumber data non-S3, seperti Amazon Athena atau Snowflake.

Prosedur ini menjelaskan proses penambahan aturan analisis kustom ke tabel yang dikonfigurasi menggunakan opsi Aliran terpandu di AWS Clean Rooms konsol.

Untuk menambahkan aturan analisis kustom ke tabel (alur terpandu)

- Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
- Di panel navigasi kiri, pilih Tabel.
- Pilih tabel yang dikonfigurasi.
- Pada halaman detail tabel yang dikonfigurasi, pilih Konfigurasi aturan analisis.
- Di bawah Langkah 1: Pilih jenis aturan analisis, di bawah Jenis aturan analisis, pilih Kustom pilihan.
- Di bawah Metode pembuatan, pilih Alur terpandu, lalu pilih Berikutnya.
- Di bawah Langkah 2: Tentukan kontrol analisis, untuk kontrol analisis langsung, pilih opsi berdasarkan tujuan Anda.

Tujuan Anda	Tindakan yang disarankan
Tinjau setiap analisis baru sebelum diizinkan untuk dijalankan pada tabel yang dikonfigurasi ini	1. Di bawah Template analisis diizinkan untuk dijalankan, pilih Tambahkan templat analisis. 2. Pilih Collaboration yang sesuai dan template Analisis dari daftar dropdown. 3. Pilih Berikutnya.
Izinkan kolaborator tertentu untuk menjalankan analisis apa pun dari jenis yang dipilih tanpa meninjau pada tabel ini	1. Di bawah jenis Analisis, a. Pilih kueri apa saja untuk mengizinkan kueri apa pun yang dibuat oleh yang Akun AWS Anda tentukan. b. Pilih Setiap kueri untuk mengizinkan pekerjaan apa pun yang dibuat oleh yang Akun AWS Anda tentukan. 2. Di bawah Akun AWS diizinkan untuk membuat analisis apa pun, pilih Tambah Akun AWS. 3. Masukkan Akun AWS atau pilih Akun AWS ID. dari daftar dropdown. 4. (Opsional) Pilih Tambahkan yang lain Akun AWS untuk menambahkan yang lain Akun AWS. 5. Pilih Berikutnya.

8. Di bawah Langkah 3: Tentukan kontrol hasil analisis,
 - a. Untuk kontrol hasil Job, perhatikan bahwa tidak ada kontrol hasil tambahan yang didukung.
 - b. Di bawah Kontrol hasil kueri, untuk Kolom yang tidak diizinkan dalam keluaran, pilih kolom yang ingin diizinkan dalam keluaran kueri, berdasarkan tujuan Anda.

Tujuan Anda	Tindakan yang disarankan
Izinkan semua kolom dikembalikan dalam output kueri	1. Pilih Tidak Ada 2. Lanjutkan ke Analisis tambahan yang diterapkan pada output.
Larang kolom tertentu dikembalikan dalam output kueri	1. Pilih daftar Kustom 2. Di bawah Tentukan kolom yang tidak diizinkan, pilih kolom yang ingin dihapus dari output kueri.

- c. Untuk Analisis tambahan yang diterapkan pada output, pilih apakah analisis tambahan dapat diterapkan ke output kueri, berdasarkan tujuan Anda.

Tujuan Anda	Opsi yang disarankan
• Izinkan hanya kueri langsung pada tabel ini. • Tolak analisis tambahan agar tidak dijalankan pada hasil kueri. • Tabel hanya dapat digunakan untuk query langsung.	Tidak diizinkan
Izinkan tetapi tidak memerlukan kueri langsung dan analisis tambahan pada tabel ini.	Izinkan
• Mengharuskan tabel hanya dapat digunakan dalam kueri langsung yang diproses dengan salah satu analisis tambahan yang diperlukan. • Kueri langsung pada tabel ini harus diproses lebih lanjut sebelum dapat dikembalikan.	Diperlukan

- d. Pilih Berikutnya.

9. (Opsional) Di bawah Langkah 4: Tetapkan privasi diferensial, tentukan apakah Anda ingin privasi diferensial diaktifkan atau dimatikan.

Privasi diferensial adalah teknik yang terbukti secara matematis untuk melindungi data Anda dari serangan identifikasi ulang.

 Note

AWS Clean Rooms Privasi Diferensial hanya tersedia untuk kolaborasi di mana data disimpan di Amazon S3.

Untuk privasi Diferensial, pilih apakah akan mengaktifkan atau menonaktifkan privasi diferensial, berdasarkan tujuan Anda.

Tujuan Anda	Tindakan yang disarankan
- Anda tidak memerlukan perlindungan terhadap upaya identifikasi ulang - Tabel Anda tidak memiliki data tingkat pengguna	- Pilih Matikan. - Pilih Berikutnya.

Tujuan Anda	Tindakan yang disarankan
- Anda memerlukan perlindungan terhadap upaya identifikasi ulang - Tabel Anda memiliki data tingkat pengguna	- Pilih Nyalakan. - Pilih kolom Pengenal pengguna yang berisi pengenal unik pengguna Anda, seperti <code>user_id</code> kolom, yang privasinya ingin Anda lindungi. Untuk mengaktifkan privasi diferensial untuk dua tabel atau lebih dalam kolaborasi, Anda harus mengonfigurasi kolom yang sama dengan kolom pengenal Pengguna di kedua aturan analisis untuk mempertahankan definisi pengguna yang konsisten di seluruh tabel. Jika terjadi kesalahan konfigurasi, anggota yang dapat melakukan kueri menerima pesan kesalahan bahwa ada dua kolom yang dapat dipilih untuk menghitung jumlah kontribusi pengguna (misalnya, jumlah tayangan iklan yang dibuat oleh pengguna) saat menjalankan kueri. - Pilih Berikutnya.

10. Di bawah Langkah 5: Tinjau dan konfigurasi, tinjau pilihan yang telah Anda buat untuk langkah sebelumnya, edit jika perlu, lalu pilih Konfigurasi aturan analisis.

Anda melihat pesan konfirmasi bahwa Anda telah berhasil mengonfigurasi aturan analisis kustom untuk tabel.

Menambahkan aturan analisis ke tabel (editor JSON)

Prosedur berikut menunjukkan cara menambahkan aturan analisis ke tabel menggunakan opsi editor JSON di AWS Clean Rooms konsol.

 Note

Tabel yang dikonfigurasi menggunakan sumber data non-S3 hanya mendukung aturan [analisis khusus](#).

Untuk menambahkan agregasi, daftar, atau aturan analisis kustom ke tabel (editor JSON)

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Tabel.
3. Pilih tabel yang dikonfigurasi.
4. Pada halaman detail tabel yang dikonfigurasi, pilih Konfigurasi aturan analisis.
5. Di bawah Langkah 1: Pilih jenis aturan analisis, di bawah Jenis aturan analisis, pilih opsi Agregasi, Daftar, atau Kustom.
6. Di bawah Metode pembuatan, pilih editor JSON, lalu pilih Berikutnya.
7. Di bawah Langkah 2: Tentukan kontrol, Anda dapat memilih untuk menyisipkan struktur kueri (Sisipkan template) atau menyisipkan file (Impor dari file).

Jika Anda memilih...	Lalu...
Sisipkan templat	1. Tentukan parameter untuk aturan analisis yang dipilih dalam Definisi aturan analisis. 2. Anda dapat menekan Ctrl+Spacebar untuk mengaktifkan pelengkapan otomatis. Untuk informasi selengkapnya tentang parameter aturan analisis agregasi, lihat Aturan analisis agregasi - kontrol kueri. Untuk informasi selengkapnya tentang parameter aturan analisis daftar, lihat Aturan analisis daftar - kontrol kueri.
Impor dari file	1. Pilih file JSON Anda dari drive lokal Anda. 2. Pilih Buka.

Jika Anda memilih...	Lalu...
	Definisi aturan Analisis menampilkan aturan analisis dari file yang diunggah.

8. Pilih Berikutnya.
9. Di bawah Langkah 3: Tinjau dan konfigurasi, tinjau pilihan yang telah Anda buat untuk langkah sebelumnya, edit jika perlu, lalu pilih Konfigurasi aturan analisis.

Anda menerima pesan konfirmasi bahwa Anda telah berhasil mengonfigurasi aturan analisis untuk tabel.

Langkah selanjutnya

Sekarang setelah Anda mengonfigurasi aturan analisis ke tabel yang dikonfigurasi, Anda siap untuk:

- [Kaitkan tabel yang dikonfigurasi ke kolaborasi](#)
- [Kueri tabel data](#) (sebagai anggota yang dapat melakukan kueri)

Mengaitkan tabel terkonfigurasi ke kolaborasi

Setelah Anda membuat tabel yang dikonfigurasi dan menambahkan aturan analisis ke dalamnya, Anda dapat mengaitkannya dengan kolaborasi dan memberikan AWS Clean Rooms peran layanan untuk mengakses AWS Glue tabel Anda.

Note

Peran layanan ini memiliki izin untuk tabel. Peran layanan hanya dapat diasumsikan oleh AWS Clean Rooms untuk menjalankan kueri yang diizinkan atas nama anggota yang dapat melakukan kueri. Tidak ada anggota kolaborasi (selain pemilik data) yang memiliki akses ke tabel yang mendasarinya dalam kolaborasi. Pemilik data dapat mengaktifkan privasi diferensial untuk membuat tabel mereka tersedia untuk kueri oleh anggota lain.

Anggaran akses data

Saat Anda mengaitkan tabel yang dikonfigurasi, Anda dapat menerapkan anggaran akses data. Anggaran akses data mengontrol berapa kali tabel dapat digunakan untuk kueri, pekerjaan, dan saluran input ML dalam suatu kolaborasi. Anggaran ini membantu organisasi mengelola pemanfaatan sumber daya dan mengendalikan biaya dengan membatasi penggunaan tabel.

Setiap kali tabel digunakan dalam kueri, pekerjaan, atau saluran input ML, anggaran untuk tabel tersebut dikurangi satu. Ketika anggaran mencapai nol, tabel tidak dapat digunakan dalam kueri SQL, pekerjaan Pyspark, atau sebagai bagian dari saluran input ML yang berasal dari tabel.

Anda dapat menetapkan anggaran per periode yang menyegarkan secara berkala, anggaran seumur hidup untuk penggunaan keseluruhan, atau keduanya. Secara default, penggunaan tabel tidak terbatas.

- Anggaran per periode — Alokasi terbaru yang membatasi berapa kali tabel ini dapat digunakan dalam periode waktu tertentu. Anda dapat mengatur periode menjadi harian, mingguan, atau bulanan. Anggaran ini dapat diatur untuk menyegarkan secara otomatis setiap hari, mingguan, atau bulanan.
- Anggaran seumur hidup — Alokasi berjalan yang membatasi jumlah total kali tabel ini dapat digunakan.

Kaitkan tabel yang dikonfigurasi

Topik berikut menjelaskan cara mengaitkan tabel yang dikonfigurasi dan menerapkan anggaran akses data ke kolaborasi menggunakan AWS Clean Rooms konsol.

Untuk informasi tentang cara mengaitkan tabel yang dikonfigurasi dengan kolaborasi menggunakan AWS SDKs, lihat [Referensi AWS Clean Rooms API](#).

Langkah 1: Selesaikan prasyarat

Untuk mengaitkan tabel yang dikonfigurasi, Anda harus menyelesaikan prasyarat berikut:

- AWS Glue Tabel yang menunjuk ke lokasi folder Amazon S3 (bukan satu file)
- Untuk tabel terenkripsi AWS Glue :
 - Peran layanan dengan izin untuk menggunakan AWS KMS kunci untuk mendekripsi tabel AWS Glue

- Untuk kumpulan data Amazon AWS KMS S3 yang dienkripsi: Peran layanan juga harus memiliki izin untuk menggunakan kunci untuk mendekripsi data Amazon S3 AWS KMS

Untuk informasi tentang mengonfigurasi enkripsi, lihat [Menyiapkan enkripsi AWS Glue di Panduan AWS Glue Pengembang](#).

Untuk memverifikasi lokasi AWS Glue tabel Anda:

1. Buka AWS Glue konsol di <https://console.aws.amazon.com/glue/>
2. Lihat detail tabel Anda dan konfirmasikan titik lokasi ke folder S3

Langkah 2: Kaitkan tabel yang dikonfigurasi

Untuk mengaitkan tabel yang dikonfigurasi

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Pilih metode untuk mengaitkan tabel Anda:
 - a. Dari halaman detail tabel yang dikonfigurasi:
 - i. Di panel navigasi kiri, pilih Tabel.
 - ii. Pilih tabel yang dikonfigurasi.
 - iii. Pada halaman detail tabel yang dikonfigurasi, pilih Kaitkan dengan kolaborasi.
 - iv. Untuk kotak dialog Tabel asosiasi ke kolaborasi, pilih Kolaborasi dari daftar dropdown.
 - b. Dari halaman detail kolaborasi:
 - i. Di panel navigasi kiri, pilih Kolaborasi.
 - ii. Pilih kolaborasi.
 - iii. Pada tab Tabel, pilih Tabel asosiasi.
3. Pada halaman tabel Associate, lakukan salah satu hal berikut:
 - Pilih tabel terkonfigurasi yang ada — Pilih nama tabel Konfigurasi yang ingin Anda kaitkan dengan kolaborasi dari daftar tarik-turun.
 - Konfigurasi tabel baru — Pilih Konfigurasi tabel baru dan ikuti petunjuk pada halaman Konfigurasi tabel baru.

- Lihat skema dan aturan analisis untuk tabel yang dikonfigurasi — Aktifkan Lihat skema dan aturan analisis.

4. Untuk detail asosiasi Tabel,

- a. Masukkan Nama untuk tabel terkait.

Anda dapat menggunakan nama default atau mengganti nama tabel ini.

- b. (Opsional) Masukkan Deskripsi tabel.

Deskripsi membantu menulis kueri.

5. Tentukan izin akses Layanan dengan memilih Buat dan gunakan peran layanan baru atau Gunakan peran layanan yang ada.

 Note

Jika Anda mengaitkan tabel yang dikonfigurasi yang didukung oleh Amazon Athena, pilih nama peran layanan yang ada dari daftar tarik-turun. Pastikan peran layanan memiliki IAM dan, jika diperlukan, izin Lake Formation ke kumpulan data.

Jika Anda memilih...	Lalu...
Membuat dan menggunakan peran layanan baru	• AWS Clean Rooms membuat peran layanan dengan kebijakan yang diperlukan untuk tabel ini. • Nama peran Layanan default adalah <code>cleanrooms-<timestamp></code> • Anda harus memiliki izin untuk membuat peran dan melampirkan kebijakan. • Jika data input Anda dienkripsi, Anda dapat memilih Data ini dienkripsi dengan kunci KMS dan kemudian masukkan AWS KMS key yang akan digunakan untuk mendekripsi input data Anda.

Jika Anda memilih...	Lalu...
Gunakan peran layanan yang ada	1. Pilih nama peran layanan yang ada dari daftar tarik-turun. Daftar peran ditampilkan jika Anda memiliki izin untuk membuat daftar peran. Jika Anda tidak memiliki izin untuk membuat daftar peran, Anda dapat memasukkan Nama Sumber Daya Amazon (ARN) peran yang ingin Anda gunakan. 2. Lihat peran layanan dengan memilih tautan eksternal View in IAM. Jika tidak ada peran layanan yang ada, opsi untuk Menggunakan peran layanan yang ada tidak tersedia. Secara default, AWS Clean Rooms tidak mencoba memperbarui kebijakan peran yang ada untuk menambahkan izin yang diperlukan. 3. (Opsional) Pilih kotak centang Tambahkan kebijakan yang telah dikonfigurasi sebelumnya dengan izin yang diperlukan untuk peran ini untuk menambahkan izin lampiran yang diperlukan ke peran. Anda harus memiliki izin untuk mengubah peran dan membuat kebijakan.

 Note

- AWS Clean Rooms memerlukan izin untuk melakukan kueri sesuai dengan aturan analisis. Untuk informasi selengkapnya tentang izin AWS Clean Rooms, lihat [AWS kebijakan terkelola untuk AWS Clean Rooms](#).
- Jika peran tidak memiliki izin yang memadai AWS Clean Rooms, Anda akan menerima pesan galat yang menyatakan bahwa peran tersebut tidak memiliki izin yang memadai untuk peran tersebut. AWS Clean Rooms Kebijakan peran harus ditambahkan sebelum melanjutkan.
- Jika Anda tidak dapat mengubah kebijakan peran, Anda akan menerima pesan galat yang menyatakan bahwa AWS Clean Rooms tidak dapat menemukan kebijakan untuk peran layanan.

6. Jika Anda ingin mengaktifkan tag asosiasi tabel yang dikonfigurasi untuk sumber daya asosiasi tabel yang dikonfigurasi, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.
7. Pilih Berikutnya.
8. Pada halaman aturan Konfigurasi analisis kolaborasi, pilih salah satu dari berikut ini:
 - Ya, buat aturan analisis kolaborasi sekarang — Kaitkan tabel Anda dengan kolaborasi ini dan buat aturan analisis kolaborasi
 - Tidak, saya akan membuat aturan analisis kolaborasi nanti — Kaitkan tabel Anda dengan kolaborasi ini saja. Anda dapat membuat aturan analisis kolaborasi nanti.
9. Jika Anda memilih Ya, buat aturan analisis kolaborasi sekarang, untuk pengiriman Hasil, pilih Anggota yang diizinkan untuk menerima hasil untuk output kueri dari daftar tarik-turun.
10. Pilih Berikutnya.
11. Pada halaman Tambah anggaran akses data, untuk konfigurasi anggaran akses data, pilih salah satu dari berikut ini:
 - Ya, tambahkan anggaran akses data sekarang — Kaitkan tabel Anda dengan kolaborasi ini dan tambahkan anggaran akses data. Anda dapat memilih anggaran periode, anggaran seumur hidup, atau keduanya.
 - Tidak, saya akan menambahkan anggaran akses data nanti — Kaitkan tabel Anda dengan kolaborasi ini saja. Anda dapat menambahkan anggaran akses data nanti.

Jika Anda memilih Tidak, saya akan menambahkan anggaran akses data nanti, lewati ke langkah 15.

12. Jika Anda memilih Ya, tambahkan anggaran akses data sekarang, pilih salah satu konfigurasi anggaran berikut:

Anggaran per periode saja	Anggaran seumur hidup saja	Baik per periode dan anggaran seumur hidup
1. Tinggalkan Tambah per periode anggaran dipilih. 2. Masukkan jumlah anggaran Per periode antara 1 dan 1.000.000. 3. Untuk Periode, pilih Harian, Mingguan, atau Bulanan. 4. (Opsional) Tinggalkan Segarkan anggaran mingguan secara otomatis dipilih untuk memperbarui alokasi. 5. Hapus Tambahkan anggaran seumur hidup.	1. Hapus Tambah per periode anggaran. 2. Pilih Tambahkan anggaran seumur hidup. 3. Masukkan jumlah anggaran seumur hidup antara 1 dan 1.000.000.	1. Tinggalkan Tambah per periode anggaran dipilih. 2. Masukkan jumlah anggaran Per periode antara 1 dan 1.000.000. 3. Untuk Periode, pilih Harian, Mingguan, atau Bulanan. 4. Tinggalkan Segarkan anggaran mingguan secara otomatis dipilih. 5. Pilih Tambahkan anggaran seumur hidup. 6. Masukkan jumlah anggaran seumur hidup antara 1 dan 1.000.000.

13. Tinjau pilihan Anda di bawah Ringkasan anggaran akses data.

Example Contoh

Misalnya, jika Anda telah memilih jumlah anggaran per periode sebesar 1.000, atur Periode ke Mingguan, biarkan kotak centang Segarkan anggaran mingguan secara otomatis dipilih, dan atur Anggaran seumur hidup menjadi 1.000.000, maka ringkasan anggaran Akses akan menampilkan pesan berikut: Setiap minggu, tabel ini dapat digunakan hingga 1.000 kali untuk menjalankan kueri atau pekerjaan. Anggaran ini diatur untuk disegarkan secara otomatis setiap hari Minggu pukul 00:00 UTC, dan akan terus disegarkan hingga tabel ini mencapai anggaran seumur hidup 1.000.000 penggunaan.

14. (Opsional) Jika Anda ingin mengaktifkan tag anggaran akses data untuk sumber daya anggaran akses, pilih Tambahkan tag baru dan masukkan pasangan Kunci dan Nilai.
15. Pilih Berikutnya.
16. Tinjau informasi di halaman Ulasan dan buat.
 - a. Jika Anda perlu mengedit bagian apa pun, pilih Edit.
 - b. Edit konfigurasi Anda, lalu pilih Berikutnya.
17. Pilih tabel Associate.

Langkah 3: Langkah selanjutnya

Sekarang setelah Anda mengaitkan tabel data yang dikonfigurasi ke kolaborasi, Anda siap untuk:

- [Tambahkan aturan analisis kolaborasi](#) ke tabel yang dikonfigurasi
- [Edit kolaborasi](#), jika Anda pembuat kolaborasi
- [Kueri tabel data](#) (sebagai anggota yang dapat melakukan kueri)

Mengonfigurasi anggaran akses data

Kolaborator dapat melihat, menambah, mengedit, dan menghapus anggaran akses data untuk menetapkan batas berapa kali tabel dapat digunakan dalam alur kerja. Gunakan anggaran ini untuk mengelola data dan biaya.

Setiap kali tabel ditanyakan atau pekerjaan dijalankan menggunakan saluran input ML yang berasal dari tabel, anggaran untuk tabel tersebut dikurangi satu. Ketika anggaran mencapai nol, tabel tidak dapat ditanyakan dan pekerjaan ML tidak dapat dijalankan menggunakan saluran input ML yang berasal dari tabel.

Anda dapat menetapkan anggaran per periode yang menyegarkan secara berkala, anggaran seumur hidup untuk penggunaan keseluruhan, atau keduanya. Secara default, penggunaan tabel tidak terbatas.

- Anggaran per periode — Alokasi terbaru yang membatasi berapa kali tabel ini dapat digunakan dalam periode waktu tertentu. Anda dapat mengatur periode menjadi harian, mingguan, atau bulanan. Anggaran ini dapat diatur untuk menyegarkan secara otomatis setiap hari, mingguan, atau bulanan.

- Anggaran seumur hidup — Alokasi berjalan yang membatasi jumlah total kali tabel ini dapat digunakan.

Topik

- [Melihat anggaran akses data](#)
- [Menambahkan anggaran akses data ke tabel terkait yang ada](#)
- [Mengedit anggaran akses data](#)
- [Menghapus anggaran akses data](#)

Melihat anggaran akses data

Anda dapat melihat anggaran akses data dari tab Tabel atau dari halaman detail tabel.

Untuk melihat anggaran akses data

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pilih tab Tabel.
5. Lakukan salah satu tindakan berikut:
 - Di bawah kolom Anggaran akses data yang tersisa, pilih anggaran untuk melihat detailnya.
 - Pilih tabel, dan pada halaman detail tabel, gulir ke bawah untuk melihat bagian Detail anggaran akses data.

Menambahkan anggaran akses data ke tabel terkait yang ada

Sebagai anggota kolaborasi, Anda dapat menambahkan anggaran akses data ke tabel terkait yang ada.

Untuk menambahkan anggaran akses data ke tabel terkait yang ada

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.

2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pilih tab Tabel.
5. Pilih tombol opsi di sebelah tabel yang ingin Anda tambahkan anggaran akses data.
6. Dari daftar dropdown Tindakan, di bawah Anggaran akses data, pilih Tambah (jika belum ada anggaran).
7. Pilih salah satu konfigurasi anggaran berikut:

Anggaran per periode saja	Anggaran seumur hidup saja	Baik per periode dan anggaran seumur hidup
1. Tinggalkan Tambah per periode anggaran dipilih. 2. Masukkan jumlah anggaran Per periode antara 1 dan 1.000.000. 3. Untuk Periode, pilih Harian, Mingguan, atau Bulanan. 4. (Opsional) Tinggalkan Segarkan anggaran mingguan secara otomatis dipilih untuk memperbarui alokasi. 5. Hapus Tambahkan anggaran seumur hidup.	1. Hapus Tambah per periode anggaran. 2. Pilih Tambahkan anggaran seumur hidup. 3. Masukkan jumlah anggaran seumur hidup antara 1 dan 1.000.000.	1. Tinggalkan Tambah per periode anggaran dipilih. 2. Masukkan jumlah anggaran Per periode antara 1 dan 1.000.000. 3. Untuk Periode, pilih Harian, Mingguan, atau Bulanan. 4. Tinggalkan Segarkan anggaran mingguan secara otomatis dipilih. 5. Pilih Tambahkan anggaran seumur hidup. 6. Masukkan jumlah anggaran seumur hidup antara 1 dan 1.000.000.

8. Tinjau pilihan Anda di bawah Ringkasan anggaran akses data.
9. Example Contoh

Misalnya, jika Anda telah memilih jumlah anggaran per periode sebesar 1.000, atur Periode ke Mingguan, biarkan kotak centang Segarkan anggaran mingguan secara otomatis dipilih, dan atur Anggaran seumur hidup menjadi 1.000.000, maka ringkasan anggaran Akses akan menampilkan pesan berikut: Setiap minggu, tabel ini dapat digunakan hingga 1.000 kali untuk menjalankan

kueri atau pekerjaan. Anggaran ini diatur untuk disegarkan secara otomatis setiap hari Minggu pukul 00:00 UTC, dan akan terus disegarkan hingga tabel ini mencapai anggaran seumur hidup 1.000.000 penggunaan.

10. (Opsional) Jika Anda ingin mengaktifkan tag anggaran akses data untuk sumber daya anggaran akses, pilih Tambahkan tag baru dan masukkan pasangan Kunci dan Nilai.
11. Pilih Tambahkan anggaran akses data.

Mengedit anggaran akses data

Sebagai anggota kolaborasi, Anda dapat mengedit anggaran akses data. Saat Anda mengedit anggaran akses data, itu mengatur ulang saldo anggaran saat ini.

Anda dapat mengedit anggaran akses data dari tab Tabel atau dari halaman detail tabel.

Tables tab

Untuk mengedit anggaran akses data dari tab Tabel

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pilih tab Tabel.
5. Pilih tombol opsi di sebelah tabel yang ingin Anda edit.
6. Dari daftar dropdown Tindakan, di bawah Anggaran akses data, pilih Edit.
7. Pada halaman Edit anggaran akses data, perbarui informasi anggaran Per periode atau anggaran seumur hidup.
8. Lihat ringkasan anggaran akses data untuk memverifikasi bahwa pengeditan yang Anda buat sudah benar.
9. Pilih Simpan perubahan.

Table details page

Untuk mengedit anggaran akses data dari halaman detail tabel

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pilih tab Tabel.
5. Pilih meja.
6. Pada halaman detail tabel, gulir ke bawah ke bagian Detail anggaran akses data.
7. Dari daftar dropdown Tindakan, pilih Edit.
8. Pada halaman Edit anggaran akses data, perbarui informasi anggaran Per periode atau anggaran seumur hidup.
9. Pilih Simpan perubahan.

Menghapus anggaran akses data

Anda dapat menghapus anggaran akses data dari tab Tabel atau dari halaman detail tabel.

Tables tab

Untuk menghapus anggaran akses data dari tab Tabel

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pilih tab Tabel.
5. Pilih tombol opsi di sebelah tabel yang ingin Anda hapus.
6. Dari daftar dropdown Tindakan, di bawah Anggaran akses data, pilih Hapus.

 Important

Anda tidak dapat membatalkan tindakan ini dan anggaran akses data Anda akan diatur ulang ke tak terbatas.

7. Jika Anda yakin ingin menghapus anggaran akses data, pilih Hapus.

Table details page

Untuk menghapus anggaran akses data dari halaman detail tabel

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pilih tab Tabel.
5. Pilih meja.
6. Pada halaman detail tabel, gulir ke bawah ke bagian Detail anggaran akses data.
7. Dari daftar dropdown Tindakan, pilih Hapus.

 Important

Anda tidak dapat membatalkan tindakan ini dan anggaran akses data Anda akan diatur ulang ke tak terbatas.

8. Jika Anda yakin ingin menghapus anggaran akses data, pilih Hapus.

Menambahkan aturan analisis kolaborasi ke tabel yang dikonfigurasi

Aturan analisis kolaborasi memungkinkan Anda menentukan kontrol yang spesifik untuk kolaborasi ini. Kontrol ini bekerja sama dengan aturan analisis tabel yang dikonfigurasi untuk menentukan bagaimana tabel ini dapat dianalisis dalam kolaborasi ini.

Anda menambahkan aturan analisis kolaborasi ke tabel yang dikonfigurasi setelah Anda [membuat tabel yang dikonfigurasi](#), [menambahkan aturan analisis](#), dan [mengaitkannya ke kolaborasi](#). Anda perlu menambahkan aturan analisis kolaborasi jika tabel dikonfigurasi untuk mendukung analisis langsung atau untuk memungkinkan analisis tambahan.

- Analisis langsung — Tabel dapat digunakan dalam kueri yang menganalisisnya secara langsung. Misalnya, dalam kueri yang mengeluarkan analisis pengukuran agregat atau daftar pengidentifikasi untuk aktivasi.
- Analisis tambahan — Tabel juga dapat digunakan sebagai masukan ke dalam analisis tambahan, selain kueri yang menganalisisnya secara langsung. Misalnya, tabel dapat digunakan dalam kueri yang merupakan benih untuk model ML yang mirip, atau saluran input ML untuk model HTML kustom.

Untuk menambahkan aturan analisis kolaborasi ke tabel

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pada tab Tabel, di bawah Tabel yang terkait dengan Anda, lihat tabel yang dikonfigurasi yang telah Anda kaitkan dengan kolaborasi.
 - Jika status analisis langsung atau status analisis tambahan memiliki status Siap, maka tabel siap untuk ditanyakan.
 - Jika Status analisis langsung atau Status analisis tambahan memiliki status Tidak siap, lalu pilih status, lalu pilih Konfigurasi di kotak dialog.
5. Pada halaman aturan Konfigurasi analisis kolaborasi, perluas Lihat aturan analisis tabel yang dikonfigurasi untuk melihat detailnya.
6. Untuk Analisis tambahan yang diizinkan, pilih opsi berdasarkan tujuan Anda.

Tujuan Anda	Opsi yang disarankan
Izinkan analisis tambahan di atas meja.	Apa saja
Izinkan hanya analisis tambahan di atas meja oleh anggota tertentu.	Setiap oleh anggota tertentu

Tujuan Anda	Opsi yang disarankan
Izinkan hanya analisis spesifik di atas meja.	Daftar kustom

- Untuk pengiriman Hasil, tentukan siapa yang dapat menerima hasil dari Anggota yang diizinkan menerima hasil untuk tarik-turun keluaran kueri.
- Pilih Konfigurasi aturan analisis.

Mengkonfigurasi kebijakan privasi diferensial (opsional)

Note

AWS Clean Rooms Privasi Diferensial hanya tersedia untuk kolaborasi di mana data disimpan di Amazon S3.

Prosedur ini menjelaskan proses konfigurasi kebijakan privasi diferensial dalam kolaborasi dengan menggunakan opsi Aliran terpandu di AWS Clean Rooms konsol. Ini adalah langkah satu kali untuk semua tabel dengan perlindungan privasi diferensial.

Untuk mengonfigurasi pengaturan privasi diferensial (aliran terpandu)

- Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
- Di panel navigasi kiri, pilih Kolaborasi.
- Pilih kolaborasi.
- Pada tab Tabel di halaman kolaborasi, pilih Konfigurasi kebijakan privasi diferensial.
- Pada halaman Konfigurasi kebijakan privasi diferensial, pilih nilai untuk properti berikut:
 - Anggaran privasi
 - Segarkan anggaran privasi setiap bulan
 - Kebisingan ditambahkan per kueri

Anda dapat menggunakan nilai default atau memasukkan nilai khusus yang mendukung kasus penggunaan spesifik Anda. Setelah memilih nilai untuk anggaran Privasi dan Noise ditambahkan

per kueri, Anda dapat melihat pratinjau utilitas yang dihasilkan dalam hal jumlah agregasi yang mungkin di semua kueri pada data Anda.

6. Pilih Konfigurasi

Anda akan melihat pesan konfirmasi bahwa Anda telah berhasil mengonfigurasi kebijakan privasi diferensial untuk kolaborasi tersebut.

Sekarang setelah Anda mengonfigurasi privasi diferensial, Anda siap untuk:

- [Kueri tabel data](#) (sebagai anggota yang dapat melakukan kueri)
- [Kolaborasi](#) (jika Anda pembuat kolaborasi)

Melihat log penggunaan privasi diferensial

Sebagai anggota kolaborasi yang melindungi data dengan privasi diferensial, setelah Anda membuat kolaborasi dengan privasi diferensial, Anda dapat memantau penggunaan anggaran privasi.

Untuk melihat berapa banyak agregasi yang dijalankan dan berapa banyak anggaran privasi yang digunakan

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pilih tab Tabel.
5. Pilih Lihat log penggunaan (teks biru).
6. Lihat detail penggunaan, termasuk anggaran privasi dan berapa banyak utilitas yang disediakan.

Mengedit kebijakan privasi diferensial

Kapan saja setelah mengonfigurasi kebijakan privasi diferensial, Anda dapat memperbaruinya untuk lebih mencerminkan kebutuhan privasi Anda.

Untuk mengedit kebijakan privasi diferensial

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pada tab Tabel pada halaman kolaborasi, di bawah Tabel yang terkait dengan Anda, pilih Edit.
5. Pada halaman Edit privasi diferensial, pilih nilai baru untuk properti berikut:
 - Anggaran privasi - Pindahkan bilah geser untuk menambah atau mengurangi anggaran kapan saja selama kolaborasi. Anda tidak dapat mengurangi anggaran setelah anggota yang dapat melakukan kueri mulai menanyakan data Anda. Jika anggaran Privasi meningkat, AWS Clean Rooms akan terus menggunakan anggaran yang ada sampai sepenuhnya dikonsumsi sebelum memanfaatkan anggaran privasi yang baru ditambahkan.
 - Noise ditambahkan per kueri - Pindahkan bilah penggeser untuk menambah atau mengurangi Noise yang ditambahkan per kueri kapan saja selama kolaborasi.

Note

Anda dapat memilih contoh Interaktif untuk mengeksplorasi bagaimana nilai yang berbeda dari anggaran Privasi dan Kebisingan yang ditambahkan per kueri memengaruhi jumlah fungsi agregat yang dapat Anda jalankan.

Anda tidak dapat mengubah nilai penyegaran anggaran Privasi. Untuk mengubah pilihan Anda, Anda harus menghapus kebijakan privasi diferensial dan membuat yang baru.

6. Pilih Simpan perubahan.

Anda melihat pesan konfirmasi bahwa Anda telah berhasil mengedit kebijakan privasi diferensial.

Menghapus kebijakan privasi diferensial

Anda dapat menghapus kebijakan privasi diferensial dari tab Tabel kolaborasi.

Untuk menghapus kebijakan privasi diferensial

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pada tab Tabel pada halaman kolaborasi, di samping Kebijakan privasi diferensial, pilih Hapus.
5. Jika Anda yakin ingin menghapus kebijakan privasi diferensial, pilih Hapus.

Setelah menghapus kebijakan privasi diferensial, Anda tidak dapat mengakses log penggunaan anggaran privasi dari kebijakan tersebut. Tabel dengan privasi diferensial diaktifkan tidak dapat ditanyakan jika kebijakan privasi diferensial dihapus.

Melihat parameter privasi diferensial yang dihitung

Untuk pengguna dengan keahlian dalam privasi diferensial, Anda dapat melihat parameter privasi diferensial yang dihitung dari tab Analisis kolaborasi.

Untuk melihat parameter privasi diferensial yang dihitung

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pada tab Analisis, di bagian Hasil, pilih Lihat parameter privasi diferensial yang dihitung.

Dalam tabel parameter privasi diferensial terhitung, Anda dapat melihat nilai sensitivitas fungsi agregat, yang didefinisikan sebagai jumlah maksimum yang dapat mengubah hasil fungsi jika catatan pengguna tunggal ditambahkan, dihapus, atau dimodifikasi. Daftar ini mencakup parameter privasi diferensial berikut:

- Batas kontribusi pengguna (UCL) adalah jumlah maksimum baris yang disumbangkan oleh pengguna dalam kueri SQL. Misalnya, jika Anda ingin menghitung jumlah tayangan yang cocok dalam kampanye tertentu di mana setiap pengguna dapat memiliki beberapa tayangan, Privasi AWS Clean Rooms Diferensial harus mengikat jumlah tayangan satu pengguna untuk memastikan bahwa perhitungan privasi diferensial akurat. Dengan kata lain, jika ada pengguna yang memiliki

lebih banyak tayangan daripada terikat, maka AWS Clean Rooms secara otomatis mengambil sampel acak seragam dari tayangan pengguna tersebut sesuai nilai UCL yang dihitung dan mengecualikan tayangan yang tersisa dari pengguna tersebut saat menjalankan kueri. Nilai UCL sama dengan 1 jika Anda menghitung jumlah pengguna unik. Ini karena menambahkan, menghapus, atau memodifikasi satu pengguna dapat mengubah jumlah pengguna yang berbeda paling banyak 1.

- Nilai minimum adalah batas bawah ekspresi yang digunakan dalam fungsi agregat seperti `sum()`. Misalnya, jika ekspresi adalah kolom yang dikenal sebagai `purchase_value`, nilai minimum adalah batas bawah kolom.
- Nilai maksimum adalah batas atas ekspresi yang digunakan dalam fungsi agregat seperti `sum()`. Misalnya, jika ekspresi adalah kolom yang dikenal sebagai `purchase_value`, nilai maksimum adalah batas atas kolom.

Dalam tabel parameter privasi diferensial terhitung, Anda dapat menggunakan parameter ini untuk lebih memahami jumlah total noise dalam hasil kueri. Misalnya, ketika Noise yang dikonfigurasi yang ditambahkan per kueri adalah 30 pengguna dan `COUNT DISTINCT (user_id)` kueri dijalankan, maka Privasi AWS Clean Rooms Diferensial menambahkan noise acak yang jatuh antara -30 dan 30 dengan probabilitas tinggi karena sensitivitas `COUNT DISTINCT` adalah 1. Dalam kasus `COUNT` kueri dengan konfigurasi yang sama, Privasi AWS Clean Rooms Diferensial menambahkan noise statistik yang diskalakan oleh batas kontribusi pengguna karena satu pengguna dapat menyumbangkan beberapa baris ke hasil kueri. Dalam kasus `SUM` kueri seperti `SUM (purchase_value)` di mana semua nilai kolom positif, total noise diskalakan oleh batas kontribusi pengguna dikalikan nilai maksimum. AWS Clean Rooms Privasi Diferensial secara otomatis menghitung parameter sensitivitas untuk melakukan penambahan noise pada waktu proses kueri dan menghabiskan anggaran privasi. Penipisan anggaran privasi diperlukan karena parameter sensitivitas bergantung pada data.

Melihat tabel dan aturan analisis

Untuk melihat tabel yang terkait dengan kolaborasi dan aturan analisis

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.

4. Pilih tab Tabel.
5. Pilih salah satu cara berikut:
 - a. Untuk melihat tabel yang terkait dalam kolaborasi, untuk Tabel yang terkait dengan Anda, pilih tabel (teks biru).
 - b. Untuk melihat tabel lain yang terkait dalam kolaborasi, untuk Tabel yang terkait dengan kolaborator, pilih tabel (teks biru).
6. Lihat detail tabel dan aturan analisis pada halaman detail tabel.

Mengedit tabel yang dikonfigurasi

Prasyarat:

- An Akun AWS dengan akses ke AWS Clean Rooms

Bagian berikut menjelaskan cara mengedit nama, deskripsi, dan detail konfigurasi tabel untuk Amazon S3, Amazon Athena, dan sumber data Snowflake.

Untuk informasi tentang cara mengedit tabel yang dikonfigurasi menggunakan AWS SDKs, lihat [Referensi AWS Clean Rooms API](#).

Untuk mengedit tabel yang dikonfigurasi

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Tabel.
3. Pilih tabel yang dikonfigurasi yang Anda buat.
4. Pada halaman detail tabel yang dikonfigurasi, pilih Edit.
5. Edit konfigurasi Anda.
6. Pilih Simpan perubahan.

Mengedit tag tabel yang dikonfigurasi

Sebagai anggota kolaborasi, setelah Anda membuat tabel yang dikonfigurasi, Anda dapat mengelola tag pada sumber daya tabel yang dikonfigurasi pada tab Tabel yang dikonfigurasi.

Untuk mengedit tag tabel yang dikonfigurasi

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Tabel.
3. Pilih tabel yang dikonfigurasi yang Anda buat.
4. Pada halaman detail tabel yang dikonfigurasi, gulir ke bawah ke bagian Tag.
5. Pilih Kelola tanda.
6. Pada halaman Kelola tag, Anda dapat melakukan hal berikut:
 - Untuk menghapus sebuah tag, pilih Hapus.
 - Untuk menambahkan tanda, pilih Tambahkan tanda baru.
 - Untuk menyimpan perubahan Anda, memilih Simpan perubahan.

Mengedit aturan analisis tabel yang dikonfigurasi

Untuk mengedit aturan analisis tabel yang dikonfigurasi

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Tabel.
3. Pilih tabel yang dikonfigurasi yang Anda buat.
4. Pada halaman detail tabel yang dikonfigurasi, gulir ke bawah ke aturan analisis agregasi, aturan analisis daftar, atau bagian Aturan analisis kustom. (Pilihan Anda tergantung pada jenis aturan analisis yang Anda pilih untuk tabel yang dikonfigurasi.)
5. Pilih Edit.
6. Pada halaman aturan Edit analisis, Anda dapat:
 - Ubah definisi aturan Analisis dengan:
 - Memodifikasi editor JSON.
 - Memilih Impor dari file untuk mengunggah definisi aturan analisis baru.
 - Pratinjau apa yang akan dilihat anggota dalam kolaborasi dengan memilih dari opsi berikut:
 - Tampilan tabel
 - JSON

- Contoh kueri

7. Pilih Simpan perubahan untuk menyimpan perubahan Anda.

Menghapus aturan analisis tabel yang dikonfigurasi

Warning

Tindakan ini tidak dapat dibatalkan dan berdampak pada semua sumber daya terkait.

Untuk menghapus aturan analisis tabel yang dikonfigurasi

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Tabel.
3. Pilih tabel yang dikonfigurasi yang Anda buat.
4. Pada halaman detail tabel yang dikonfigurasi, gulir ke bawah ke aturan analisis agregasi, aturan analisis daftar, atau bagian Aturan analisis kustom. (Pilihan Anda tergantung pada jenis aturan analisis yang Anda pilih untuk tabel yang dikonfigurasi.)
5. Pilih Hapus.
6. Jika Anda yakin ingin menghapus aturan analisis, pilih Hapus.

Kolom tabel yang dikonfigurasi tidak diizinkan

Konfigurasi kolom keluaran yang tidak diizinkan adalah kontrol dalam aturan analisis AWS Clean Rooms kustom yang memungkinkan Anda menentukan daftar kolom (jika ada) yang tidak diizinkan untuk diproyeksikan dalam hasil kueri. Kolom yang direferensikan dalam daftar ini dianggap “kolom keluaran yang tidak diizinkan”. Ini berarti bahwa referensi apa pun ke kolom tersebut melalui transformasi, aliasing, atau cara lain mungkin tidak ada dalam SELECT akhir (proyeksi) kueri.

Meskipun kemampuan melarang kolom diproyeksikan secara langsung dalam output, itu tidak sepenuhnya mencegah nilai yang mendasarinya disimpulkan secara tidak langsung melalui mekanisme lain. Kolom ini masih dapat digunakan dalam klausa proyeksi (seperti dalam subquery atau Common Table Expression (CTE)), selama mereka tidak direferensikan dalam proyeksi paling akhir.

Konfigurasi kolom keluaran yang tidak diizinkan memberi Anda fleksibilitas untuk menerapkan dan mengkodeifikasi kontrol pada tabel Anda dalam kombinasi dengan tinjauan tingkat templat analisis berdasarkan kasus penggunaan dan persyaratan privasi yang sesuai.

Untuk informasi selengkapnya tentang cara mengatur konfigurasi ini, lihat [Menambahkan aturan analisis kustom ke tabel \(alur terpandu\)](#).

Contoh

Contoh berikut menampilkan bagaimana kontrol kolom keluaran yang tidak diizinkan diterapkan.

- Anggota A bekerja sama dengan Anggota B.
- Anggota B adalah anggota yang dapat menjalankan kueri.
- Anggota A mendefinisikan pengguna tabel dengan umur kolom, jenis kelamin, email, dan nama. Usia kolom dan nama adalah kolom keluaran yang tidak diizinkan.
- Anggota B mendefinisikan hewan peliharaan tabel dengan kumpulan kolom usia, jenis kelamin, dan owner_name yang serupa. Namun, mereka tidak menetapkan batasan apa pun pada kolom keluaran, yang berarti bahwa semua kolom dalam tabel dapat diproyeksikan secara bebas dalam kueri.

Jika Anggota B menjalankan kueri berikut, kueri tersebut diblokir karena kolom keluaran yang tidak diizinkan tidak dapat diproyeksikan secara langsung:

```
SELECT
  age
FROM
  users
```

Jika Anggota B menjalankan kueri berikut, itu diblokir karena kolom keluaran yang tidak diizinkan tidak dapat diproyeksikan secara implisit melalui bintang proyek:

```
SELECT
  *
FROM
  users
```

Jika Anggota B menjalankan kueri berikut, kueri tersebut diblokir karena transformasi kolom keluaran yang tidak diizinkan tidak dapat diproyeksikan:

```
SELECT
  COUNT(age)
FROM
  users
```

Jika Anggota B menjalankan kueri berikut, kueri tersebut diblokir karena kolom keluaran yang tidak diizinkan tidak dapat direferensikan dalam proyeksi akhir menggunakan alias:

```
SELECT
  count_age
FROM
  (SELECT COUNT(age) AS count_age FROM users)
```

Jika Anggota B menjalankan kueri berikut, itu diblokir karena kolom terbatas yang diubah diproyeksikan dalam output:

```
SELECT
  CONCAT(name, email)
FROM
  users
```

Jika Anggota B menjalankan kueri berikut, itu diblokir karena kolom keluaran yang tidak diizinkan yang ditentukan dalam CTE tidak dapat direferensikan dalam proyeksi akhir:

```
WITH cte AS (
  SELECT
    age AS age_alias
  FROM
    users
)
SELECT age_alias FROM cte
```

Jika Anggota B menjalankan kueri berikut, itu diblokir karena kolom keluaran yang tidak diizinkan tidak dapat digunakan sebagai kunci pengurutan atau partisi dalam proyeksi akhir:


```
SELECT
  LISTAGG(gender) WITHIN GROUP (ORDER BY age) OVER (PARTITION BY age)
FROM
  users
```

Jika Anggota B menjalankan kueri berikut, itu berhasil karena kolom yang merupakan bagian dari kolom keluaran yang tidak diizinkan masih dapat digunakan di seluruh konstruksi lain dalam kueri, seperti dalam klausa gabungan atau filter.

```
SELECT
  u.name,
  p.gender,
  p.age
FROM
  users AS u
JOIN
  pets AS p
ON
  u.name = p.owner_name
```

Dalam skenario yang sama, Anggota B juga dapat menggunakan kolom nama di pengguna sebagai filter atau kunci sortir:

```
SELECT
  u.email,
  u.gender
FROM
  users AS u
WHERE
  u.name = 'Mike'
ORDER BY
  u.name
```

Selain itu, kolom keluaran yang tidak diizinkan dari pengguna dapat digunakan dalam proyeksi menengah seperti subquery dan CTEs, seperti:

```
WITH cte AS (
```

```
SELECT
    u.gender,
    u.id,
    u.first_name
FROM
    users AS u
)
SELECT
    first_name
FROM
    (SELECT cte.gender, cte.id, cte.first_name FROM cte)
```

Mengedit asosiasi tabel yang dikonfigurasi

Sebagai anggota kolaborasi, Anda dapat mengedit asosiasi tabel yang telah dikonfigurasi yang telah Anda buat.

Untuk mengedit asosiasi tabel yang dikonfigurasi

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pilih tab Tabel.
5. Untuk Tabel yang terkait dengan Anda, pilih tabel.
6. Pada halaman detail tabel, gulir ke bawah untuk melihat detail asosiasi Tabel.
7. Pilih Edit.
8. Pada halaman Edit asosiasi tabel yang dikonfigurasi, perbarui Deskripsi atau informasi akses Layanan.
9. Pilih Simpan perubahan.

Memutuskan tabel yang dikonfigurasi

Sebagai anggota kolaborasi, Anda dapat memisahkan tabel yang dikonfigurasi dari kolaborasi. Tindakan ini mencegah anggota yang dapat melakukan kueri dari menanyakan tabel.

Untuk memisahkan tabel yang dikonfigurasi

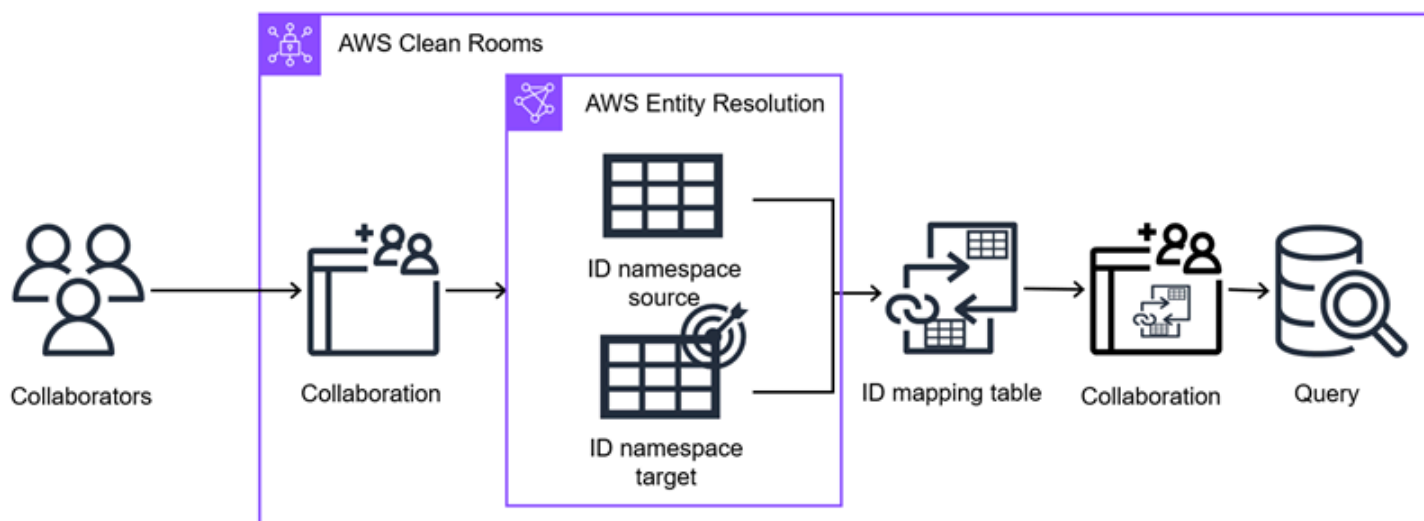
1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pilih tab Tabel.
5. Untuk Tabel yang terkait dengan Anda, pilih tombol opsi di sebelah tabel yang ingin Anda pisahkan.
6. Pilih Pisahkan.
7. Di kotak dialog, konfirmasi keputusan untuk memisahkan tabel yang dikonfigurasi dan mencegah anggota yang dapat melakukan kueri untuk menanyakan tabel dengan memilih Disassociate.

Resolusi Entitas AWS di AWS Clean Rooms

Dengan Resolusi Entitas AWS in AWS Clean Rooms, Anda dapat menerjemahkan data dari sumber ke target, mengisi tabel pemetaan ID dengan data yang diterjemahkan, dan menanyakan data.

Pertama, Anda membuat kolaborasi AWS Clean Rooms dan menambahkan yang ingin Akun AWS Anda undang, atau bergabung dengan kolaborasi yang Anda undang dengan membuat keanggotaan. Selanjutnya, Anda melakukan pemetaan ID pada dua tabel data. Anda melakukan ini dengan mengaitkan sumber namespace ID yang ada atau membuat yang baru di. Resolusi Entitas AWS Anggota kolaborasi lainnya mengaitkan target namespace ID yang ada atau membuat target namespace ID baru. Kemudian, Anda membuat dan mengisi tabel pemetaan ID dari dua ruang nama ID terkait. Terakhir, anggota yang dapat melakukan kueri menjalankan kueri di dua tabel data dengan bergabung pada tabel pemetaan ID.

Diagram berikut merangkum cara bekerja dengan Resolusi Entitas AWS in AWS Clean Rooms.



Note

Penyedia layanan transcoding yang saat ini didukung adalah LiveRamp, yang tersedia sebagai berikut Wilayah AWS: US East (Virginia N.), US East (Ohio), dan US West (Oregon).

Topik

- [Ruang nama ID di AWS Clean Rooms](#)
- [Tabel pemetaan ID di AWS Clean Rooms](#)

Ruang nama ID di AWS Clean Rooms

Namespace ID adalah pembungkus di sekitar tabel identitas Anda yang memungkinkan Anda memberikan metadata yang menjelaskan kumpulan data Anda dan cara menggunakannya dalam alur kerja pemetaan ID. Alur kerja pemetaan ID adalah pekerjaan pemrosesan data yang memetakan data dari sumber data input ke target data input berdasarkan metode pemetaan ID yang ditentukan. Ini menghasilkan tabel pemetaan ID.

Ada dua jenis ruang nama ID: Sumber dan Target. Sumber berisi konfigurasi untuk data sumber yang akan diproses dalam alur kerja pemetaan ID. Target berisi konfigurasi data target yang akan diselesaikan oleh semua sumber. Untuk menentukan data masukan yang ingin Anda selesaikan di dua Akun AWS, buat sumber namespace ID dan target namespace ID untuk menerjemahkan data Anda dari satu set (Sumber) ke set lainnya (Target).

Anda dapat membuat namespace ID baru atau Anda dapat mengaitkan namespace yang sudah ada. Untuk informasi selengkapnya tentang cara membuat namespace ID di Resolusi Entitas AWS, lihat [Mendefinisikan data input menggunakan namespace ID](#) di Panduan Pengguna.Resolusi Entitas AWS

Topik

- [Membuat dan mengaitkan namespace ID baru](#)
- [Mengaitkan namespace ID yang ada](#)
- [Mengedit asosiasi namespace ID](#)
- [Memutuskan asosiasi namespace ID](#)

Membuat dan mengaitkan namespace ID baru

Setiap anggota kolaborasi harus membuat dan mengaitkan Source namespace ID atau Target namespace ID sebelum membuat tabel pemetaan ID untuk menanyakan data identitas.

Jika Anda telah membuat namespace ID di Resolusi Entitas AWS, lewati ke. [Mengaitkan namespace ID yang ada](#)

Untuk membuat dan mengaitkan namespace ID baru

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms/>.
2. Di panel navigasi kiri, pilih Kolaborasi.

3. Pilih kolaborasi.
4. Pada tab Resolusi entitas, pilih Namespace ID Asosiasi.
5. Pada halaman namespace ID Associate, untuk data resolusi Entity, pilih Create ID namespace.

Resolusi Entitas AWS Konsol muncul di tab baru.

6. Ikuti petunjuk di halaman namespace Create ID di konsol. Resolusi Entitas AWS
 - a. Untuk Detail, masukkan nama namespace ID, Deskripsi, dan pilih jenis namespace ID (Sumber atau Target).
 - b. Untuk metode namespace ID, pilih metode berbasis Aturan untuk pencocokan berbasis aturan atau layanan Penyedia untuk transcoding pihak ketiga.
 - c. Tentukan tipe input Data, tergantung pada metode namespace ID yang Anda pilih.
 - d. Pilih Buat namespace ID.
7. Kembali ke AWS Clean Rooms konsol.
8. Pada halaman namespace ID Associate, untuk data resolusi Entity, pilih sumber namespace Resolusi Entitas AWS ID atau target yang ingin Anda kaitkan dengan kolaborasi dari daftar tarik-turun.
9. Untuk detail Asosiasi, ambil langkah-langkah berikut.
 - a. Masukkan Nama untuk namespace ID terkait.

Anda dapat menggunakan nama default atau mengganti nama namespace ID ini.
 - b. (Opsional) Masukkan Deskripsi namespace ID.

Deskripsi membantu menulis kueri.
10. Tentukan izin AWS Clean Rooms akses dengan memilih opsi dan kemudian mengambil tindakan yang disarankan.

Opsi	Tindakan yang disarankan
Izinkan AWS Clean Rooms untuk menambah dan mengelola kebijakan izin	AWS Clean Rooms membuat peran layanan dengan kebijakan yang diperlukan untuk asosiasi ini.
Tambahkan dan kelola izin secara manual	Lakukan salah satu tindakan berikut:

Opsi	Tindakan yang disarankan
	- Tinjau kebijakan Sumber Daya dan tambahkan izin yang diperlukan ke kebijakan. - Gunakan kebijakan yang ada dengan memilih Tambahkan pernyataan kebijakan. Anda harus memiliki izin untuk mengubah peran dan membuat kebijakan.  Note Jika Anda tidak dapat mengubah kebijakan peran, Anda akan menerima pesan galat yang menyatakan bahwa AWS Clean Rooms tidak dapat menemukan kebijakan untuk peran layanan.

11. (Opsional) Untuk konfigurasi tabel pemetaan ID Tingkat Lanjut, ubah perlindungan default untuk kolom yang berasal dari kartu nama ID.

Tabel pemetaan ID dikonfigurasi secara default untuk hanya mengizinkan INNER JOIN pada sourceID kolom dan targetID kolom. Anda dapat memodifikasi konfigurasi ini sehingga kolom yang berasal dari namespace ID ini (baik sourceID atau targetID) dapat diizinkan di mana saja dalam kueri.

Tujuan Anda	Opsi yang disarankan
Kategorikan kolom sebagai “kolom gabungan” dan hanya mengizinkannya dalam klausa INNER JOIN	Ya
Kategorikan kolom sebagai “kolom dimensi” dan izinkan di mana saja dalam kueri,	Tidak, izinkan di mana saja dalam kueri

Tujuan Anda	Opsi yang disarankan
termasuk JOIN klausa, SELECT, WHERE dan GROUP BY pernyataan kueri.	

12. (Opsional) Jika Anda ingin mengaktifkan Tag untuk sumber daya namespace ID, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.
13. Pilih Kaitkan.
14. Pada tab Resolusi entitas, di bawah tabel ruang nama ID Terkait, lihat namespace ID terkait dan verifikasi bahwa jenis namespace ID sudah benar (Sumber atau Target).

Setelah semua anggota dalam kolaborasi mengaitkan ruang nama ID mereka, Anda dapat [membuat tabel pemetaan ID](#) dan menanyakan data.

Mengaitkan namespace ID yang ada

Dalam prosedur ini, setiap anggota mengaitkan sumber namespace ID yang ada atau target namespace ID mereka dalam kolaborasi.

Untuk mengaitkan namespace ID yang ada

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms/>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pada tab Resolusi entitas, pilih Namespace ID Asosiasi.
5. Pada halaman namespace ID Associate, untuk data resolusi Entity, pilih sumber namespace Resolusi Entitas AWS ID atau target yang ingin Anda kaitkan dengan kolaborasi dari daftar tarik-turun.
6. Untuk detail Asosiasi, ambil langkah-langkah berikut.
 - a. Masukkan Nama untuk namespace ID terkait.

Anda dapat menggunakan nama default atau mengganti nama namespace ID ini.
 - b. (Opsional) Masukkan Deskripsi namespace ID.

Deskripsi membantu menulis kueri.

7. Tentukan izin AWS Clean Rooms akses dengan memilih opsi dan kemudian mengambil tindakan yang disarankan.

Opsi	Tindakan yang disarankan
Izinkan AWS Clean Rooms untuk menambah dan mengelola kebijakan izin	AWS Clean Rooms membuat peran layanan dengan kebijakan yang diperlukan untuk asosiasi ini.
Tambahkan dan kelola izin secara manual	Lakukan salah satu tindakan berikut: - Tinjau kebijakan Sumber Daya dan tambahkan izin yang diperlukan ke kebijakan. - Gunakan kebijakan yang ada dengan memilih Tambahkan pernyataan kebijakan. Anda harus memiliki izin untuk mengubah peran dan membuat kebijakan.  Note Jika Anda tidak dapat mengubah kebijakan peran, Anda akan menerima pesan galat yang menyatakan bahwa AWS Clean Rooms tidak dapat menemukan kebijakan untuk peran layanan.

8. (Opsional) Untuk konfigurasi tabel pemetaan ID Tingkat Lanjut, ubah perlindungan default untuk kolom yang berasal dari kartu nama ID.

Tabel pemetaan ID dikonfigurasi secara default untuk hanya mengizinkan INNER JOIN pada sourceID kolom dan targetID kolom. Anda dapat memodifikasi konfigurasi ini sehingga kolom yang berasal dari namespace ID ini (baik sourceID atau targetID) dapat diizinkan di mana saja dalam kueri.

Tujuan Anda	Opsi yang disarankan
Kategorikan kolom sebagai “kolom gabungan” dan hanya izinkan dalam klausa. INNER JOIN	Ya
Kategorikan kolom sebagai “kolom dimensi” dan izinkan di mana saja dalam kueri, termasuk JOIN klausa,, SELECTWHERE, dan GROUP BY pernyataan kueri.	Tidak, izinkan di mana saja dalam kueri

- (Opsional) Jika Anda ingin mengaktifkan Tag untuk sumber daya namespace ID, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.
- Pilih Kaitkan.
- Pada tab Resolusi entitas, di bawah tabel ruang nama ID Terkait, lihat namespace ID terkait dan verifikasi bahwa jenis namespace ID sudah benar (Sumber atau Target).

Setelah semua anggota dalam kolaborasi mengaitkan ruang nama ID mereka, Anda dapat [membuat tabel pemetaan ID](#) dan menanyakan data.

Mengedit asosiasi namespace ID

Sebagai anggota kolaborasi, Anda dapat mengedit asosiasi namespace ID yang telah Anda buat.

Untuk mengedit asosiasi namespace ID

- Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms/>.
- Di panel navigasi kiri, pilih Kolaborasi.
- Pilih kolaborasi.
- Pilih tab Resolusi entitas.
- Untuk ruang nama ID Terkait, pilih namespace ID.
- Pada halaman detail namespace ID, gulir ke bawah untuk melihat detail asosiasi namespace ID.
- Pilih Edit.
- Pada halaman asosiasi namespace Edit ID, edit salah satu dari berikut ini:

- a. Untuk detail Asosiasi, perbarui Nama atau Deskripsi.
- b. (Opsional) Untuk konfigurasi tabel pemetaan ID Tingkat Lanjut, ubah perlindungan default untuk kolom yang berasal dari kartu nama ID.

Tabel pemetaan ID dikonfigurasi secara default untuk hanya mengizinkan INNER JOIN pada `sourceID` kolom dan `targetID` kolom. Anda dapat memodifikasi konfigurasi ini sehingga kolom yang berasal dari namespace ID ini (baik `sourceID` atau `targetID`) dapat diizinkan di mana saja dalam kueri.

Tujuan Anda	Opsi yang disarankan
Kategorikan kolom sebagai “kolom gabungan” dan hanya mengizinkannya dalam klausa INNER JOIN	Ya
Kategorikan kolom sebagai “kolom dimensi” dan izinkan di mana saja dalam kueri, termasuk JOIN klausa, SELECT, WHERE dan GROUP BY pernyataan kueri.	Tidak, izinkan di mana saja dalam kueri

9. Pilih Simpan perubahan.

Memutuskan asosiasi namespace ID

Sebagai anggota kolaborasi, Anda dapat memisahkan namespace ID dari kolaborasi. Tindakan ini mencegah anggota yang dapat melakukan kueri dari menanyakan tabel.

Warning

Memutuskan asosiasi namespace ID dari kolaborasi menghapus data apa pun dari tabel pemetaan ID turunan, menjadikannya tidak dapat dikueri.

Misalnya, jika asosiasi namespace ID Anda digunakan sebagai SUMBER dalam tiga tabel pemetaan ID yang berbeda, maka semua data dari tabel pemetaan ID ini akan dihapus saat Anda memisahkan asosiasi namespace ID Anda.

Untuk memisahkan asosiasi namespace ID

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms/>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pilih tab Resolusi entitas.
5. Untuk ruang nama ID Terkait, pilih tombol opsi di sebelah namespace ID yang ingin Anda pisahkan.
6. Pilih Pisahkan.
7. Di kotak dialog, konfirmasi keputusan Anda untuk memutuskan sambungan namespace ID dengan memilih Disassociate. Tindakan ini mencegah setiap anggota yang dapat melakukan kueri mengakses tabel pemetaan ID.

Jika anggota kolaborasi menghapus salah satu ruang nama ID, Anda tidak dapat mengisi kembali tabel pemetaan ID jika sumber telah meninggalkan kolaborasi.

Meskipun tabel pemetaan ID telah diisi sebelumnya, memisahkan namespace ID berarti Anda tidak dapat lagi menjalankan kueri pada tabel itu.

Tabel pemetaan ID di AWS Clean Rooms

Tabel pemetaan ID adalah sumber daya AWS Clean Rooms yang memungkinkan pemetaan identitas multipihak dalam kolaborasi.

Sebelum membuat tabel pemetaan ID, Anda harus terlebih dahulu memiliki data sumber dan target yang dikonfigurasi sebagai ruang nama ID.

Setelah membuat tabel pemetaan ID, Anda menggunakan alur kerja pemetaan ID untuk menerjemahkan namespace ID sumber ke namespace ID target. Anda dapat melakukan ini menggunakan metode berbasis aturan, atau metode transcoding layanan penyedia.

Alur kerja pemetaan ID adalah pekerjaan pemrosesan data yang memetakan data dari sumber data input ke target data input berdasarkan metode alur kerja pemetaan ID yang ditentukan. Alur kerja ini mengisi tabel pemetaan ID.

 Note

Tabel pemetaan ID hanya dapat dibuat dari kumpulan data yang disimpan di Amazon S3 dan dirayapi ke dalam tabel. AWS Glue

Ada dua metode alur kerja pemetaan ID: pemetaan ID berbasis aturan atau pemetaan ID layanan penyedia:

- Pemetaan ID Berbasis aturan – Anda menggunakan aturan pencocokan untuk menerjemahkan data pihak pertama dari sumber ke target.
- Pemetaan ID layanan penyedia — Anda menggunakan layanan LiveRamp penyedia untuk menerjemahkan data pihak ketiga dari sumber ke target.

 Note

Penyedia layanan transcoding yang saat ini didukung adalah LiveRamp. Setiap anggota dalam kolaborasi yang memiliki langganan dengan LiveRamp through AWS Data Exchange dapat membuat tabel pemetaan ID. Jika Anda sudah berlangganan LiveRamp, tetapi tidak melalui AWS Data Exchange, hubungi LiveRamp untuk mendapatkan penawaran pribadi. Untuk informasi selengkapnya, lihat [Berlangganan layanan penyedia AWS Data Exchange di](#) Panduan Resolusi Entitas AWS Pengguna.

Topik

- [Membuat dan mengisi tabel pemetaan ID baru](#)
- [Mengisi tabel pemetaan ID yang ada](#)
- [Mengedit tabel pemetaan ID](#)
- [Menghapus tabel pemetaan ID](#)

Membuat dan mengisi tabel pemetaan ID baru

Prasyarat

Sebelum membuat tabel pemetaan ID, pastikan Anda memiliki:

- Sumber dan target namespace ID terkait

- Ruang nama ID dikonfigurasi untuk pemetaan ID berbasis aturan atau pemetaan ID layanan penyedia

Anda dapat membuat dua jenis tabel pemetaan ID:

- Berbasis aturan — Menggunakan aturan yang cocok untuk menerjemahkan data pihak pertama
- Layanan penyedia - Penggunaan LiveRamp untuk menerjemahkan Ramp IDs

Setelah Anda membuat tabel pemetaan ID, Anda dapat langsung mengisinya dengan menjalankan alur kerja pemetaan ID, atau Anda dapat menunggu untuk mengisi tabel nanti.

Setelah tabel pemetaan ID berhasil diisi, Anda kemudian dapat menjalankan kueri gabungan multi-tabel pada tabel pemetaan ID untuk bergabung `sourceId` dengan `targetId` dan menganalisis data.

Topik

- [Buat tabel pemetaan ID \(berbasis aturan\)](#)
- [Membuat tabel pemetaan ID \(layanan penyedia\)](#)

Buat tabel pemetaan ID (berbasis aturan)

Topik ini menjelaskan proses pembuatan tabel pemetaan ID yang menggunakan aturan pencocokan untuk menerjemahkan data pihak pertama dari sumber ke target.

Saat membuat tabel pemetaan ID berbasis aturan, Anda dapat memilih untuk memproses hanya catatan baru, diperbarui, atau dihapus dalam alur kerja dengan mengaktifkan pemrosesan tambahan.

Untuk membuat dan mengisi tabel pemetaan ID baru menggunakan metode berbasis aturan

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms/>.
2. Di panel navigasi, pilih Kolaborasi.
3. Pilih kolaborasi, lalu pilih tab Resolusi entitas.
4. Pilih Buat tabel pemetaan ID.
5. Di bawah pengaturan pemetaan ID, lakukan salah satu hal berikut:

- Untuk membuat alur kerja baru, tetap pilih alur kerja Pemetaan ID baru.
 - Untuk menggunakan alur kerja yang ada, kosongkan kotak centang dan pilih alur kerja dari daftar tarik-turun, lalu lewati ke langkah 9.
6. Di bawah Data identitas, lihat atau konfigurasi sumber dan target.
- Untuk pasangan namespace ID tunggal: Tinjau Sumber dan Target yang telah dipilih sebelumnya.
 - Untuk beberapa ruang nama ID: Pilih Sumber dan Target dari daftar dropdown.
7. (Opsional) Pilih Aktifkan pemrosesan tambahan untuk memproses hanya catatan baru, diperbarui, atau dihapus dalam alur kerja.

Resolusi Entitas AWS memproses hanya catatan baru, diperbarui, atau dihapus baik di namespace Sumber atau ID Target, daripada membuat ulang seluruh tabel pemetaan ID.

Jika Anda membiarkan opsi ini tidak dipilih, Resolusi Entitas AWS jalankan alur kerja pemetaan ID pemrosesan batch default pada tabel pemetaan ID.

8. Di bawah parameter Aturan, konfigurasi yang berikut ini:
- Kontrol aturan - Pilih apakah Target atau Sumber menyediakan aturan yang cocok.

Anda dapat melihat aturan dengan mengaktifkan Tampilkan aturan.

Kontrol aturan harus kompatibel antara sumber dan namespace ID target untuk digunakan dalam alur kerja pemetaan ID. Misalnya, jika namespace ID sumber membatasi aturan ke target tetapi namespace ID target membatasi aturan ke sumber, ini menghasilkan kesalahan.

- Jenis perbandingan secara otomatis diatur ke Beberapa bidang input.

Ini karena kedua peserta telah memilih opsi ini sebelumnya.

- Rekam pencocokan
 - Satu sumber ke satu target - Menyimpan satu catatan yang cocok per target
 - Banyak sumber untuk satu target - Menyimpan semua catatan yang cocok per target

 Note

Batasan yang ditentukan untuk ruang nama ID sumber dan target harus kompatibel.

9. Untuk detail pemetaan ID, konfigurasi hal berikut:

- a. Masukkan nama tabel pemetaan ID atau simpan nama default.
- b. (Opsional) Masukkan Deskripsi tabel pemetaan ID.

Deskripsi membantu menulis kueri.

10. Untuk AWS Clean Rooms akses, pilih salah satu:

- Izinkan AWS Clean Rooms untuk menambahkan dan mengelola kebijakan izin — Membuat peran layanan secara otomatis.
- Tambahkan dan kelola izin secara manual — Tinjau dan modifikasi kebijakan sumber daya atau pilih Tambahkan pernyataan kebijakan.

 Note

Jika Anda tidak dapat mengubah kebijakan peran, Anda akan menerima pesan galat yang menyatakan bahwa AWS Clean Rooms tidak dapat menemukan kebijakan untuk peran layanan.

11. Untuk Resolusi Entitas AWS akses, pilih salah satu:

Bagian ini hanya terlihat jika Anda membuat tabel pemetaan ID baru.

- Membuat dan menggunakan peran layanan baru
 - Nama peran Layanan default adalah `entityresolution-id-mapping-workflow-<timestamp>`
 - (Opsional) Untuk data terenkripsi, pilih Data ini dienkripsi oleh kunci KMS dan masukkan kunci.AWS KMS
- Gunakan peran layanan yang ada
 - Pilih nama peran layanan yang ada dari daftar tarik-turun atau masukkan ARN peran.

Daftar peran ditampilkan jika Anda memiliki izin untuk membuat daftar peran.

Lihat peran layanan dengan memilih tautan eksternal View in IAM.

Jika tidak ada peran layanan yang ada, opsi untuk Menggunakan peran layanan yang ada tidak tersedia.

Secara default, AWS Clean Rooms tidak mencoba memperbarui kebijakan peran yang ada untuk menambahkan izin yang diperlukan.

- (Opsional) Pilih Tambahkan kebijakan pra-konfigurasi dengan izin yang diperlukan untuk peran ini untuk melampirkan izin yang diperlukan ke peran.

Anda harus memiliki izin untuk mengubah peran dan membuat kebijakan.

12. (Opsional) Di bawah Pengaturan tambahan, konfigurasi:

a. Pengaturan tabel pemetaan ID

- Untuk mengaktifkan enkripsi khusus, pilih Sesuaikan pengaturan enkripsi dan masukkan AWS KMS kunci.

 Note

Kunci KMS ini perlu memberikan izin yang diperlukan untuk digunakan dalam menggunakan Resolusi Entitas AWS kebijakan `cleanrooms.amazonaws.com` kunci KMS. Untuk detail selengkapnya tentang izin yang diperlukan untuk bekerja dengan enkripsi dengan alur kerja pemetaan ID, lihat [Membuat peran pekerjaan alur kerja](#) di Panduan Pengguna. Resolusi Entitas AWS Resolusi Entitas AWS

- Untuk menambahkan tag, pilih Tambahkan tag baru dan masukkan pasangan nilai kunci

b. Pengaturan alur kerja pemetaan ID (hanya alur kerja baru):

- Untuk menggunakan nama yang berbeda, hapus Simpan nama dan deskripsi tabel pemetaan ID yang sama dan masukkan nilai baru.
- Untuk menambahkan tag, pilih Tambahkan tag baru dan masukkan pasangan nilai kunci

13. Pilih salah satu cara berikut:

- Buat tabel pemetaan ID - Membuat tabel kosong yang dapat Anda isi nanti () [Mengisi tabel pemetaan ID yang ada](#)
- Membuat dan mengisi tabel pemetaan ID - Membuat dan segera mengisi tabel (mungkin memakan waktu beberapa jam)

Proses alur kerja pemetaan ID dimulai. Selama proses ini, tabel pemetaan ID diisi dengan terjemahan. IDs Alur kerja pemetaan ID mungkin membutuhkan waktu beberapa jam untuk diproses.

Setelah tabel pemetaan ID berhasil diisi, Anda dapat [meminta tabel pemetaan ID](#) untuk bergabung sourceId dengan targetId dan menganalisis data.

Membuat tabel pemetaan ID (layanan penyedia)

Topik ini menjelaskan proses pembuatan tabel pemetaan ID yang menggunakan layanan penyedia (LiveRamp). Layanan LiveRamp penyedia menerjemahkan satu set Ramp sumber IDs ke yang lain menggunakan Ramp yang dipelihara atau diturunkan. IDs

Untuk membuat tabel pemetaan ID baru menggunakan metode layanan penyedia

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms/>.
2. Di panel navigasi, pilih Kolaborasi.
3. Pilih kolaborasi, lalu pilih tab Resolusi entitas.
4. Pilih Buat tabel pemetaan ID.
5. Di bawah pengaturan pemetaan ID, lakukan salah satu hal berikut:
 - Untuk membuat alur kerja baru, tetap pilih alur kerja Pemetaan ID baru.
 - Untuk menggunakan alur kerja yang ada, kosongkan kotak centang dan pilih alur kerja dari daftar tarik-turun, lalu lewati ke langkah 9.
6. Di bawah Data identitas, lihat atau konfigurasikan sumber dan target.
 - Untuk pasangan namespace ID tunggal: Tinjau Sumber dan Target yang telah dipilih sebelumnya.
 - Untuk beberapa ruang nama ID: Pilih Sumber dan Target dari daftar dropdown.
7. Di bawah Metode, verifikasi bahwa metode alur kerja pemetaan ID yang dipilih adalah LiveRamp transcoding.
8. Untuk LiveRamp konfigurasi, lakukan salah satu hal berikut:
 - Masukkan manajer LiveRamp ID ARN dan manajer LiveRamp rahasia ARN.
 - Pilih Impor dari alur kerja yang ada.
9. Untuk detail pemetaan ID, konfigurasikan hal berikut:
 - a. Masukkan nama tabel pemetaan ID atau simpan nama default.

- b. (Opsional) Masukkan Deskripsi tabel pemetaan ID.

Deskripsi membantu menulis kueri.

10. Untuk AWS Clean Rooms akses, pilih salah satu:

- Izinkan AWS Clean Rooms untuk menambahkan dan mengelola kebijakan izin — Membuat peran layanan secara otomatis.
- Tambahkan dan kelola izin secara manual — Tinjau dan modifikasi kebijakan sumber daya atau pilih Tambahkan pernyataan kebijakan.

 Note

Jika Anda tidak dapat mengubah kebijakan peran, Anda akan menerima pesan galat yang menyatakan bahwa AWS Clean Rooms tidak dapat menemukan kebijakan untuk peran layanan.

11. Untuk Resolusi Entitas AWS akses pilih salah satu:

Bagian ini hanya terlihat jika Anda membuat tabel pemetaan ID baru.

- Membuat dan menggunakan peran layanan baru
 - Nama peran Layanan default adalah `entityresolution-id-mapping-workflow-<timestamp>`
 - (Opsional) Untuk data terenkripsi, pilih Data ini dienkripsi oleh kunci KMS dan masukkan kunci.AWS KMS
- Gunakan peran layanan yang ada
 - Pilih nama peran layanan yang ada dari daftar tarik-turun atau masukkan ARN peran.

Daftar peran ditampilkan jika Anda memiliki izin untuk membuat daftar peran.

Lihat peran layanan dengan memilih tautan eksternal View in IAM.

Jika tidak ada peran layanan yang ada, opsi untuk Menggunakan peran layanan yang ada tidak tersedia.

Secara default, AWS Clean Rooms tidak mencoba memperbarui kebijakan peran yang ada untuk menambahkan izin yang diperlukan.

- (Opsional) Pilih Tambahkan kebijakan pra-konfigurasi dengan izin yang diperlukan untuk peran ini untuk melampirkan izin yang diperlukan ke peran.

Anda harus memiliki izin untuk mengubah peran dan membuat kebijakan.

12. (Opsional) Di bawah Pengaturan tambahan konfigurasi:

a. Pengaturan tabel pemetaan ID

- Untuk mengaktifkan enkripsi khusus, pilih Sesuaikan pengaturan enkripsi dan masukkan AWS KMS kunci.

 Note

Kunci KMS ini perlu memberikan izin yang diperlukan untuk digunakan dalam menggunakan Resolusi Entitas AWS kebijakan `cleanrooms.amazonaws.com` kunci KMS. Untuk detail selengkapnya tentang izin yang diperlukan untuk bekerja dengan enkripsi dengan alur kerja pemetaan ID, lihat [Membuat peran pekerjaan alur kerja](#) di Panduan Pengguna. Resolusi Entitas AWS Resolusi Entitas AWS

- Untuk menambahkan tag, pilih Tambahkan tag baru dan masukkan pasangan nilai kunci.

b. Pengaturan alur kerja pemetaan ID (hanya alur kerja baru):

- Untuk menggunakan nama yang berbeda, hapus Simpan nama dan deskripsi tabel pemetaan ID yang sama dan masukkan nilai baru.
- Untuk menambahkan tag, pilih Tambahkan tag baru dan masukkan pasangan nilai kunci.

13. Pilih salah satu cara berikut:

- Buat tabel pemetaan ID - Membuat tabel kosong yang dapat Anda isi nanti () [Mengisi tabel pemetaan ID yang ada](#)
- Membuat dan mengisi tabel pemetaan ID - Membuat dan segera mengisi tabel (mungkin memakan waktu beberapa jam)

Proses alur kerja pemetaan ID dimulai. Selama proses ini, tabel pemetaan ID diisi dengan terjemahan. IDs Alur kerja pemetaan ID mungkin membutuhkan waktu beberapa jam untuk diproses.

Setelah tabel pemetaan ID berhasil diisi, Anda dapat [meminta tabel pemetaan ID](#) untuk bergabung sourceId dengan targetId dan menganalisis data.

Mengisi tabel pemetaan ID yang ada

Saat data baru ditambahkan ke namespace ID, gunakan alur kerja ini. Jika Anda memilih untuk mengaktifkan pemrosesan tambahan saat [membuat tabel pemetaan ID](#), Anda memiliki kemampuan untuk memproses hanya catatan baru, diperbarui, atau dihapus di namespace Sumber atau ID Target, daripada membuat ulang seluruh tabel pemetaan ID.

Untuk mengisi tabel pemetaan ID yang ada

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms/>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Buka tab Resolusi entitas.
5. Di bawah bagian tabel pemetaan ID, pilih tabel pemetaan ID.
6. Jika Anda belum mengaktifkan pemrosesan inkremental untuk tabel pemetaan ID ini, pilih tabel pemetaan ID dan kemudian pilih Isi.
7. Jika Anda telah mengaktifkan pemrosesan inkremental untuk tabel pemetaan ID ini, pilih Isi dengan lalu pilih salah satu dari berikut ini:
 - Pemrosesan tambahan — Memproses hanya catatan baru, diperbarui, atau dihapus baik di namespace Sumber atau ID Target.

Direkomendasikan untuk pembaruan yang sering, pengoperasian harian, atau sinkronisasi data waktu nyata.
 - Pemrosesan Batch - Memproses seluruh tabel pemetaan ID.

Direkomendasikan untuk penyiapan awal, penyegaran penuh berkala, atau saat terjadi perubahan signifikan di ruang nama Source dan Target ID.
 - Hapus hanya pemrosesan - Memproses hanya menghapus catatan dari namespace ID Sumber dan memperbarui namespace ID Target yang sesuai.

Direkomendasikan untuk menyinkronkan penghapusan dengan cepat.

8. Proses alur kerja pemetaan ID dimulai.

Selama proses ini, tabel pemetaan ID diisi dengan transkode. IDs Alur kerja pemetaan ID mungkin membutuhkan waktu beberapa jam untuk diproses.

Setelah tabel pemetaan ID berhasil diisi, Anda dapat melakukan [kueri tabel pemetaan ID untuk bergabung dengan tabel](#) `sourceId`. `targetId`

Mengedit tabel pemetaan ID

Sebagai anggota kolaborasi, Anda dapat mengedit tabel pemetaan ID yang telah Anda buat.

Untuk mengedit tabel pemetaan ID

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pilih tab Resolusi entitas.
5. Untuk tabel pemetaan ID, pilih tabel.
6. Pada halaman detail tabel pemetaan ID, gulir ke bawah untuk melihat detail tabel pemetaan ID.
7. Pilih Edit.
8. Pada halaman tabel Edit ID pemetaan, perbarui Deskripsi atau informasi akses Layanan.
9. Pilih Simpan perubahan.

Menghapus tabel pemetaan ID

Sebagai anggota kolaborasi, Anda dapat menghapus tabel pemetaan ID yang telah Anda buat. Tindakan ini mencegah anggota yang dapat melakukan kueri dari menanyakan tabel.

Warning

Menghapus tabel pemetaan secara permanen akan menghapus data yang terisi.

Untuk menghapus tabel pemetaan ID

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pilih tab Resolusi entitas.
5. Untuk tabel pemetaan ID, pilih tabel.
6. Pada halaman detail tabel pemetaan ID, gulir ke bawah untuk melihat tabel pemetaan ID.
7. Pilih dan tabel pemetaan ID, lalu pilih Hapus.
8. Jika Anda yakin ingin menghapus tabel pemetaan ID, pilih Hapus.

Template analisis di AWS Clean Rooms

Template analisis bekerja dengan [Aturan analisis kustom di AWS Clean Rooms](#). Dengan template analisis, Anda dapat menentukan parameter untuk membantu Anda menggunakan kembali kueri yang sama. AWS Clean Rooms mendukung subset parameterisasi dengan nilai literal.

Template analisis khusus kolaborasi. Untuk setiap kolaborasi, anggota hanya dapat melihat kueri dalam kolaborasi tersebut. Jika Anda berencana untuk menggunakan privasi diferensial dalam sebuah kolaborasi, Anda harus memastikan bahwa templat analisis Anda kompatibel dengan [struktur kueri tujuan umum](#) Privasi Diferensial. AWS Clean Rooms

Anda dapat membuat template analisis dengan dua cara: menggunakan kode SQL atau menggunakan kode Python untuk Spark.

Topik

- [Templat analisis SQL](#)
- [PySpark template analisis](#)
- [Templat analisis pemecahan masalah PySpark](#)

Templat analisis SQL

Templat analisis SQL memungkinkan Anda untuk menanyakan dan menganalisis data di berbagai kumpulan data dalam kolaborasi. Anda dapat menggunakan template ini untuk melakukan berbagai jenis analisis, seperti mengidentifikasi tumpang tindih audiens dan menghitung metrik agregat.

Dengan template analisis SQL, Anda dapat:

- Tulis kueri SQL standar
- Tambahkan parameter untuk membuat kueri Anda dinamis
- Kontrol akses ke kolom dan tabel tertentu
- Tetapkan persyaratan agregasi untuk data sensitif
- Tentukan data input untuk pembuatan kumpulan data sintetis yang disempurnakan privasi untuk model pembelajaran mesin kustom (ML)

Topik

- [Membuat templat analisis SQL](#)

- [Meninjau template analisis SQL](#)

Membuat templat analisis SQL

Prasyarat

Sebelum Anda membuat template analisis SQL, Anda harus memiliki:

- AWS Clean Rooms Kolaborasi aktif
- Akses ke setidaknya satu tabel yang dikonfigurasi dalam kolaborasi

Untuk informasi tentang mengonfigurasi tabel di AWS Clean Rooms, lihat [Membuat tabel yang dikonfigurasi di AWS Clean Rooms](#).

- Izin untuk membuat templat analisis
- Pengetahuan dasar tentang sintaks kueri SQL

Prosedur berikut menjelaskan proses pembuatan template analisis SQL menggunakan [AWS Clean Rooms konsol](#).

Untuk informasi tentang cara membuat template analisis SQL menggunakan AWS SDKs, lihat [Referensi AWS Clean Rooms API](#).

Untuk membuat template analisis SQL

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Akun AWS yang akan berfungsi sebagai pembuat kolaborasi.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pada tab Template, buka bagian Analisis template yang dibuat oleh Anda.
5. Pilih Buat templat analisis.
6. Pada halaman template Buat analisis, untuk Detail,
 - a. Masukkan Nama untuk templat analisis.
 - b. (Opsional) Masukkan Deskripsi.
 - c. Untuk Format, biarkan opsi SQL dipilih.
7. Untuk Tabel, lihat tabel yang dikonfigurasi terkait dengan kolaborasi.

8. Untuk Definisi,

- a. Masukkan definisi untuk templat analisis.
- b. Pilih Impor dari untuk mengimpor definisi.
- c. (Opsional) Tentukan parameter di editor SQL dengan memasukkan titik dua (:) di depan nama parameter.

Contoh:

```
WHERE table1.date + :date_period > table1.date
```

9. Jika Anda menambahkan parameter sebelumnya, di bawah Parameter - opsional, untuk setiap nama Parameter, pilih nilai Jenis dan Default (opsional).
10. Untuk data Sintetis, jika Anda ingin menghasilkan data sintetis untuk pelatihan model, pilih kotak centang Memerlukan keluaran templat analisis menjadi sintetis.

Untuk informasi selengkapnya, lihat Pembuatan kumpulan data [sintetis yang disempurnakan privasi](#).

- a. Untuk klasifikasi Kolom, pilih Kolom dari daftar dropdown. Setidaknya lima kolom diperlukan.
 - i. Pilih Klasifikasi dari daftar dropdown. Ini mengidentifikasi tipe data untuk setiap kolom.

Jenis klasifikasi meliputi:

 - Numerik — Nilai numerik kontinu seperti pengukuran atau hitungan
 - Kategoris — Nilai atau kategori diskrit seperti label atau jenis
 - ii. Untuk menghapus kolom, pilih Hapus.
 - iii. Untuk menambahkan kolom lain, pilih Tambahkan kolom lain. Pilih Kolom dan Klasifikasi dari daftar dropdown.
 - iv. Untuk nilai Prediktif, pilih Kolom dari daftar dropdown. Ini adalah kolom yang digunakan model khusus untuk prediksi setelah dilatih pada kumpulan data sintetis.
- b. Pengaturan lanjutan memungkinkan Anda untuk mengatur tingkat Privasi dan ambang privasi. Sesuaikan pengaturan agar sesuai dengan kebutuhan Anda.
 - Untuk tingkat Privasi, masukkan nilai epsilon untuk menentukan seberapa banyak noise yang ditambahkan model sintetis untuk melindungi privasi dalam kumpulan data yang Anda hasilkan. Nilai harus antara 0,0001 dan 10.

- Nilai yang lebih rendah menambah lebih banyak noise, memberikan perlindungan privasi yang lebih kuat tetapi berpotensi mengurangi utilitas untuk model kustom hilir yang dilatih pada data ini.
- Nilai yang lebih tinggi menambahkan lebih sedikit derau, memberikan akurasi yang lebih besar tetapi berpotensi mengurangi perlindungan privasi.

Untuk ambang privasi, masukkan probabilitas tertinggi yang diizinkan bahwa serangan inferensi keanggotaan dapat mengidentifikasi anggota kumpulan data asli. Nilai harus antara 50,0 dan 100.

- Skor 50% menunjukkan bahwa serangan inferensi keanggotaan tidak berhasil membedakan anggota dari non-anggota lebih baik daripada tebakan acak.
- Tanpa batas privasi, masukkan 100%.

Nilai optimal bergantung pada kasus penggunaan spesifik dan persyaratan privasi Anda. Jika ambang privasi terlampaui, pembuatan saluran input ML gagal, dan Anda tidak dapat menggunakan kumpulan data sintetis untuk melatih model.

Warning

Pembuatan data sintetis melindungi dari menyimpulkan atribut individu apakah individu tertentu hadir dalam kumpulan data asli atau atribut pembelajaran individu tersebut hadir. Namun, itu tidak mencegah nilai literal dari kumpulan data asli, termasuk informasi yang dapat diidentifikasi secara pribadi (PII) muncul di kumpulan data sintetis.

Sebaiknya hindari nilai dalam kumpulan data input yang hanya terkait dengan satu subjek data karena ini dapat mengidentifikasi kembali subjek data. Misalnya, jika hanya satu pengguna yang tinggal dalam kode pos, keberadaan kode pos itu dalam kumpulan data sintetis akan mengonfirmasi bahwa pengguna berada dalam kumpulan data asli. Teknik seperti memotong nilai presisi tinggi atau mengganti katalog yang tidak biasa dengan yang lain dapat digunakan untuk mengurangi risiko ini. Transformasi ini dapat menjadi bagian dari query yang digunakan untuk membuat saluran input ML.

11. Jika Anda ingin mengaktifkan Tag untuk sumber daya, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.
12. Pilih Buat.

13. Anda sekarang siap memberi tahu anggota kolaborasi Anda bahwa mereka dapat [Meninjau template analisis](#). (Opsional jika Anda ingin menanyakan data Anda sendiri.)

Meninjau template analisis SQL

Setelah anggota kolaborasi membuat SQLanalysis template, Anda dapat meninjau dan menyetujuinya. Setelah template analisis dan disetujui, dapat digunakan dalam kueri di AWS Clean Rooms.

Note

Saat Anda membawa kode analisis Anda ke dalam kolaborasi, perhatikan hal-hal berikut:

- AWS Clean Rooms tidak memvalidasi atau menjamin perilaku kode analisis.
 - Jika Anda perlu memastikan perilaku tertentu, tinjau kode mitra kolaborasi Anda secara langsung atau bekerja dengan auditor pihak ketiga tepercaya untuk memeriksanya.
- Dalam model keamanan bersama:
 - Anda (pelanggan) bertanggung jawab atas keamanan kode yang berjalan di lingkungan.
 - AWS Clean Rooms bertanggung jawab atas keamanan lingkungan, memastikan bahwa
 - hanya kode yang disetujui yang berjalan
 - hanya tabel yang dikonfigurasi tertentu yang dapat diakses
 - satu-satunya tujuan output adalah bucket S3 penerima hasil.

Untuk meninjau template analisis SQL menggunakan konsol AWS Clean Rooms

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Akun AWS yang akan berfungsi sebagai pembuat kolaborasi.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pada tab Template, buka bagian Analisis template yang dibuat oleh anggota lain.
5. Pilih template analisis yang memiliki status Dapat menjalankan dari Tidak memerlukan ulasan Anda.
6. Pilih Tinjau.
7. Tinjau aturan analisis Ikhtisar, Definisi, dan Parameter (jika ada).

8. Tinjau tabel yang dikonfigurasi yang tercantum di bawah Tabel yang direferensikan dalam definisi.

Status di samping setiap tabel akan membaca Template tidak diperbolehkan.

9. Pilih meja.

Jika Anda	Kemudian pilih
Menyetujui template analisis	Izinkan template di atas meja. Konfirmasikan persetujuan Anda dengan memilih Izinkan.
Jangan menyetujui template analisis	Larang

Anda sekarang siap untuk query tabel dikonfigurasi menggunakan template analisis SQL. Untuk informasi selengkapnya, lihat [Menjalankan kueri SQL](#).

PySpark template analisis

PySpark template analisis memerlukan skrip pengguna Python dan lingkungan virtual opsional untuk menggunakan pustaka kustom dan sumber terbuka. File-file ini disebut artefak.

Sebelum Anda membuat template analisis, pertama-tama Anda membuat artefak dan kemudian menyimpan artefak dalam ember Amazon S3. AWS Clean Rooms menggunakan artefak ini saat menjalankan pekerjaan analisis. AWS Clean Rooms hanya mengakses artefak saat menjalankan pekerjaan.

Sebelum menjalankan kode apa pun pada templat PySpark analisis, AWS Clean Rooms validasi artefak dengan:

- Memeriksa versi objek S3 tertentu yang digunakan saat membuat template
- Memverifikasi hash SHA-256 dari artefak
- Gagal pekerjaan apa pun di mana artefak telah dimodifikasi atau dihapus

 Note

Ukuran maksimum semua artefak gabungan untuk templat PySpark analisis yang diberikan AWS Clean Rooms adalah 1 GB.

Keamanan untuk templat PySpark analisis

Untuk mempertahankan lingkungan komputasi yang aman, AWS Clean Rooms gunakan arsitektur komputasi dua tingkat untuk mengisolasi kode pengguna dari operasi sistem. Arsitektur ini didasarkan pada teknologi Amazon EMR Serverless Fine Grained Access Control, juga dikenal sebagai Membrane. Untuk informasi selengkapnya, lihat [Membrane — Kontrol akses data yang aman dan berkinerja di Apache Spark dengan adanya kode imperatif](#).

Komponen lingkungan komputasi dibagi menjadi ruang pengguna dan ruang sistem yang terpisah. Ruang pengguna mengeksekusi PySpark kode dalam template PySpark analisis. AWS Clean Rooms menggunakan ruang sistem untuk memungkinkan pekerjaan berjalan termasuk menggunakan peran layanan yang disediakan oleh pelanggan untuk membaca data untuk menjalankan pekerjaan dan menerapkan daftar izin kolom. Sebagai hasil dari arsitektur ini, PySpark kode pelanggan yang mempengaruhi ruang sistem, yang dapat mencakup sejumlah kecil Spark SQL dan PySpark DataFrames APIs, diblokir.

PySpark keterbatasan dalam AWS Clean Rooms

Ketika pelanggan mengirimkan template PySpark analisis yang disetujui, AWS Clean Rooms menjalankannya di lingkungan komputasi aman sendiri yang tidak dapat diakses oleh pelanggan. Lingkungan komputasi mengimplementasikan arsitektur komputasi dengan ruang pengguna dan ruang sistem untuk melestarikan lingkungan komputasi yang aman. Untuk informasi selengkapnya, lihat [Keamanan untuk templat PySpark analisis](#).

Pertimbangkan batasan berikut sebelum Anda menggunakannya PySpark AWS Clean Rooms.

Batasan

- Hanya DataFrame output yang didukung
- Sesi Spark Tunggal per eksekusi pekerjaan

Fitur yang tidak didukung

- Manajemen data
 - Format tabel gunung es
 - LakeFormation tabel terkelola
 - Kumpulan data terdistribusi tangguh (RDD)
 - Streaming percikan
 - Kontrol akses untuk kolom bersarang
- Fungsi dan ekstensi khusus
 - Fungsi tabel yang ditentukan pengguna () UDTFs
 - Sarang UDFs
 - Kelas khusus dalam fungsi yang ditentukan pengguna
 - Sumber data kustom
 - File JAR tambahan untuk:
 - Ekstensi percikan
 - Konektor
 - Konfigurasi metastore
- Pemantauan dan analisis
 - Penebangan percikan
 - Spark UI
 - ANALYZE TABLE perintah

Important

Keterbatasan ini ada untuk menjaga isolasi keamanan antara ruang pengguna dan sistem. Semua batasan berlaku terlepas dari konfigurasi kolaborasi. Pembaruan di masa mendatang dapat menambahkan dukungan untuk fitur tambahan berdasarkan evaluasi keamanan.

Praktik terbaik

Kami merekomendasikan praktik terbaik berikut saat membuat templat PySpark analisis.

- Rancang templat analisis Anda [PySpark keterbatasan dalam AWS Clean Rooms](#) dengan mempertimbangkan.
- Uji kode Anda di lingkungan pengembangan terlebih dahulu.
- Gunakan DataFrame operasi yang didukung secara eksklusif.
- Rencanakan struktur output Anda untuk bekerja dengan DataFrame keterbatasan.

Kami merekomendasikan praktik terbaik berikut untuk mengelola artefak

- Simpan semua artefak template PySpark analisis dalam bucket atau awalan S3 khusus.
- Gunakan penamaan versi yang jelas untuk versi artefak yang berbeda.
- Buat templat analisis baru saat pembaruan artefak diperlukan.
- Pertahankan inventaris templat mana yang menggunakan versi artefak mana.

Untuk informasi selengkapnya tentang cara menulis kode Spark, lihat berikut ini:

- [Contoh Apache Spark](#)
- [Tulis aplikasi Spark](#) di Panduan Rilis EMR Amazon
- [Tutorial: Menulis skrip AWS Glue untuk Spark](#) di AWS Glue Panduan Pengguna

Topik berikut menjelaskan cara membuat skrip dan pustaka pengguna Python sebelum membuat dan meninjau template analisis.

Topik

- [Membuat skrip pengguna](#)
- [Bekerja dengan parameter dalam templat PySpark analisis](#)
- [Membuat lingkungan virtual \(opsional\)](#)
- [Menyimpan skrip pengguna dan lingkungan virtual di S3](#)
- [Membuat template PySpark analisis](#)
- [Meninjau template PySpark analisis](#)

Membuat skrip pengguna

Script pengguna harus berisi fungsi entrypoint (dengan kata lain, handler). Anda dapat memberi nama file skrip pengguna Anda dengan nama file Python yang valid.

Prosedur berikut menjelaskan cara membuat skrip pengguna untuk menentukan fungsionalitas inti PySpark analisis Anda.

Prasyarat

- PySpark 1.0 (sesuai dengan Python 3.11 dan Spark 3.5.3)
- Kumpulan data di Amazon S3 hanya dapat dibaca sebagai asosiasi tabel yang dikonfigurasi dalam sesi Spark yang Anda tentukan.
- Kode Anda tidak dapat langsung memanggil Amazon S3 dan AWS Glue
- Kode Anda tidak dapat melakukan panggilan jaringan

Untuk membuat skrip pengguna

1. Buka editor teks atau Integrated Development Environment (IDE) pilihan Anda.

Anda dapat menggunakan editor teks atau IDE (seperti Visual Studio Code, PyCharm, atau Notepad ++) yang mendukung file Python.

2. Buat file Python baru dengan nama pilihan Anda (misalnya, **my_analysis.py**).
3. Mendefinisikan fungsi entrypoint yang menerima parameter objek konteks.

```
def entrypoint(context)
```

Parameter `context` objek adalah kamus yang menyediakan akses ke komponen Spark penting, tabel referensi, dan parameter analisis. Ini berisi:

Akses sesi percikan melalui `context['sparkSession']`

Tabel yang direferensikan melalui `context['referencedTables']`

Parameter analisis melalui `context['analysisParameters']` (jika parameter didefinisikan dalam template)

4. Tentukan hasil fungsi entrypoint:

```
return results
```

`results` harus mengembalikan objek yang berisi hasil kamus nama file ke output. `DataFrame`

Note

AWS Clean Rooms secara otomatis menulis `DataFrame` objek ke ember S3 dari penerima hasil.

5. Anda sekarang siap untuk:

- a. Simpan skrip pengguna ini di S3. Untuk informasi selengkapnya, lihat [Menyimpan skrip pengguna dan lingkungan virtual di S3](#).
- b. Buat lingkungan virtual opsional untuk mendukung pustaka tambahan apa pun yang diperlukan oleh skrip pengguna Anda. Untuk informasi selengkapnya, lihat [Membuat lingkungan virtual \(opsional\)](#).

Example Contoh 1

<caption>The following example demonstrates a generic user script for a PySpark analysis template.</caption>

```
# File name: my_analysis.py

def entrypoint(context):
    try:
        # Access Spark session
        spark = context['sparkSession']

        # Access input tables
        input_table1 = context['referencedTables']['table1_name']
        input_table2 = context['referencedTables']['table2_name']

        # Example data processing operations
        output_df1 = input_table1.select("column1", "column2")
        output_df2 = input_table2.join(input_table1, "join_key")
        output_df3 = input_table1.groupBy("category").count()

        # Return results - each key creates a separate output folder
        return {
```

```

        "results": {
            "output1": output_df1,          # Creates output1/ folder
            "output2": output_df2,          # Creates output2/ folder
            "analysis_summary": output_df3 # Creates analysis_summary/ folder
        }
    }

except Exception as e:
    print(f"Error in main function: {str(e)}")
    raise e

```

Struktur folder dari contoh ini adalah sebagai berikut:

```

analysis_results/
#
### output1/ # Basic selected columns
# ### part-000000.parquet
# ### _SUCCESS
#
### output2/ # Joined data
# ### part-000000.parquet
# ### _SUCCESS
#
### analysis_summary/ # Aggregated results
### part-000000.parquet
### _SUCCESS

```

Example Contoh 2

<caption>The following example demonstrates a more complex user script for a PySpark analysis template.</caption>

```

def entrypoint(context):
    try:
        # Get DataFrames from context
        emp_df = context['referencedTables']['employees']
        dept_df = context['referencedTables']['departments']

        # Apply Transformations
        emp_dept_df = emp_df.join(
            dept_df,
            on="dept_id",
            how="left"

```

```
        ).select(
            "emp_id",
            "name",
            "salary",
            "dept_name"
        )

    # Return Dataframes
    return {
        "results": {
            "outputTable": emp_dept_df
        }
    }

except Exception as e:
    print(f"Error in entrypoint function: {str(e)}")
    raise e
```

Bekerja dengan parameter dalam templat PySpark analisis

Parameter meningkatkan fleksibilitas templat PySpark analisis Anda dengan memungkinkan nilai yang berbeda diberikan pada waktu pengiriman pekerjaan. Parameter dapat diakses melalui objek konteks yang diteruskan ke fungsi entrypoint Anda.

Note

Parameter adalah string yang disediakan pengguna yang dapat berisi konten arbitrer.

- Tinjau kode untuk memastikan parameter ditangani dengan aman untuk mencegah perilaku tak terduga dalam analisis Anda.
- Desain penanganan parameter untuk bekerja dengan aman terlepas dari nilai parameter apa yang disediakan pada waktu pengiriman.

Mengakses parameter

Parameter tersedia di `context['analysisParameters']` kamus. Semua nilai parameter adalah string.

Example Mengakses parameter dengan aman

```
def entrypoint(context):
    # Access parameters from context
    parameters = context['analysisParameters']
    threshold = parameters['threshold']
    table_name = parameters['table_name']

    # Continue with analysis using parameters
    spark = context['sparkSession']
    input_df = context['referencedTables'][table_name]

    # Convert threshold value
    threshold_val = int(threshold)

    # Use parameter in DataFrame operation
    filtered_df = input_df.filter(input_df.amount > threshold_val)

    return {
        "results": {
            "output": filtered_df
        }
    }
```

Praktik terbaik keamanan parameter

Warning

Parameter adalah string yang disediakan pengguna yang dapat berisi konten arbitrer. Anda harus menangani parameter dengan aman untuk mencegah kerentanan keamanan dalam kode analisis Anda.

Pola penanganan parameter yang tidak aman untuk dihindari:

- Menjalankan parameter sebagai kode - Jangan pernah menggunakan `eval()` atau `exec()` pada nilai parameter

```
# UNSAFE - Don't do this
eval(parameters['expression']) # Can execute arbitrary code
```

- Interpolasi string SQL - Jangan pernah menggabungkan parameter langsung ke string SQL

```
# UNSAFE - Don't do this
sql = f"SELECT * FROM table WHERE column = '{parameters['value']}'" # SQL injection
risk
```

- Operasi jalur file yang tidak aman - Jangan pernah menggunakan parameter secara langsung dalam operasi sistem file tanpa validasi

```
# UNSAFE - Don't do this
file_path = f"/data/{parameters['filename']}" # Path traversal risk
```

Pola penanganan parameter aman:

- Gunakan parameter dalam DataFrame operasi — Spark DataFrames menangani nilai parameter dengan aman

```
# SAFE - Use parameters in DataFrame operations
threshold = int(parameters['threshold'])
filtered_df = input_df.filter(input_df.value > threshold)
```

- Validasi nilai parameter - Periksa apakah parameter memenuhi format yang diharapkan sebelum digunakan

```
# SAFE - Validate parameters before use
def validate_date(date_str):
    try:
        from datetime import datetime
        datetime.strptime(date_str, '%Y-%m-%d')
        return True
    except ValueError:
        return False

date_param = parameters['date_filter'] or '2024-01-01'
if not validate_date(date_param):
    raise ValueError(f"Invalid date format: {date_param}")
```

- Gunakan daftar yang diizinkan untuk nilai parameter - Jika memungkinkan, validasi parameter terhadap nilai baik yang diketahui

```
# SAFE - Use allowlists
allowed_columns = ['column1', 'column2', 'column3']
column_param = parameters['column_name']
if column_param not in allowed_columns:
    raise ValueError(f"Invalid column: {column_param}")
```

- Konversi tipe dengan penanganan kesalahan - Konversi parameter string ke tipe yang diharapkan dengan aman

```
# SAFE - Convert with error handling
try:
    batch_size = int(parameters['batch_size'] or '1000')
    if batch_size <= 0 or batch_size > 10000:
        raise ValueError(f"Batch size must be between 1 and 10000")
except ValueError as e:
    print(f"Invalid parameter: {e}")
    raise
```

Important

Ingat bahwa parameter melewati tinjauan kode saat pelari pekerjaan memberikan nilai yang berbeda. Rancang penanganan parameter Anda agar bekerja dengan aman terlepas dari nilai parameter apa yang disediakan.

Contoh parameter lengkap

Example Menggunakan parameter dengan aman dalam PySpark skrip

```
def entrypoint(context):
    try:
        # Access Spark session and tables
        spark = context['sparkSession']
        input_table = context['referencedTables']['sales_data']

        # Access parameters - fail fast if analysisParameters missing
        parameters = context['analysisParameters']

        # Validate and convert numeric parameter (handles empty strings with default)
        try:
```

```

        threshold = int(parameters['threshold'] or '100')
        if threshold <= 0:
            raise ValueError("Threshold must be positive")
    except (ValueError, TypeError) as e:
        print(f"Invalid threshold parameter: {e}")
        raise

    # Validate date parameter (handles empty strings with default)
    date_filter = parameters['start_date'] or '2024-01-01'
    from datetime import datetime
    try:
        datetime.strptime(date_filter, '%Y-%m-%d')
    except ValueError:
        raise ValueError(f"Invalid date format: {date_filter}")

    # Use parameters safely in DataFrame operations
    filtered_df = input_table.filter(
        (input_table.amount > threshold) &
        (input_table.date >= date_filter)
    )

    result_df = filtered_df.groupBy("category").agg(
        {"amount": "sum"}
    )

    return {
        "results": {
            "filtered_results": result_df
        }
    }

except Exception as e:
    print(f"Error in analysis: {str(e)}")
    raise

```

Membuat lingkungan virtual (opsional)

Jika Anda memiliki pustaka tambahan yang diperlukan oleh skrip pengguna, Anda memiliki opsi untuk membuat lingkungan virtual untuk menyimpan pustaka tersebut. Jika Anda tidak memerlukan pustaka tambahan, Anda dapat melewati langkah ini.

Saat bekerja dengan pustaka yang memiliki ekstensi asli, perlu diingat bahwa PySpark dalam AWS Clean Rooms beroperasi di Linux dengan ARM64 arsitektur.

Prosedur berikut menunjukkan cara membuat lingkungan virtual menggunakan perintah CLI dasar.

Untuk menciptakan lingkungan virtual

1. Buka terminal atau command prompt.
2. Tambahkan konten berikut:

```
# create and activate a python virtual environment
python3 -m venv pyspark_venvsource
source pyspark_venvsource/bin/activate

# install the python packages
pip3 install pandas # add packages here

# package the virtual environment into an archive
pip3 install venv-pack
venv-pack -f -o pyspark_venv.tar.gz

# optionally, remove the virtual environment directory
deactivate
rm -fr pyspark_venvsource
```

3. Anda sekarang siap untuk menyimpan lingkungan virtual ini di S3. Untuk informasi selengkapnya, lihat [Menyimpan skrip pengguna dan lingkungan virtual di S3](#).

Untuk informasi selengkapnya tentang bekerja dengan Docker dan Amazon ECR, lihat Panduan [Amazon ECRUser](#).

Menyimpan skrip pengguna dan lingkungan virtual di S3

Prosedur berikut menjelaskan cara menyimpan skrip pengguna dan lingkungan virtual opsional di Amazon S3. Selesaikan langkah ini sebelum membuat template PySpark analisis.

Important

Jangan memodifikasi atau menghapus artefak (skrip pengguna atau lingkungan virtual) setelah membuat templat analisis.

Melakukannya akan:

- Menyebabkan semua pekerjaan analisis future menggunakan template ini gagal.

- Memerlukan pembuatan template analisis baru dengan artefak baru.
- Tidak mempengaruhi pekerjaan analisis yang telah diselesaikan sebelumnya

Prasyarat

- An Akun AWS dengan izin yang sesuai
- File skrip pengguna (seperti `my_analysis.py`)
- (Opsional, jika ada) Paket lingkungan virtual (`.tar.gz` file)
- Akses untuk membuat atau memodifikasi peran IAM

Console

Untuk menyimpan skrip pengguna dan lingkungan virtual di S3 menggunakan konsol:

1. Masuk ke Konsol Manajemen AWS dan buka konsol Amazon S3 di <https://console.aws.amazon.com/s3/>
2. Buat bucket S3 baru atau gunakan yang sudah ada.
3. Aktifkan pembuatan versi untuk ember.
 - a. Pilih ember Anda.
 - b. Pilih Properti.
 - c. Di bagian Bucket Versioning, pilih Edit.
 - d. Pilih Aktifkan dan simpan perubahan.
4. Unggah artefak Anda dan aktifkan hash SHA-256.
 - a. Arahkan ke ember Anda.
 - b. Pilih Unggah.
 - c. Pilih Tambahkan file dan tambahkan file skrip pengguna Anda.
 - d. (Opsional, jika ada) Tambahkan file `.tar.gz` Anda.
 - e. Perluas Properti.
 - f. Di bawah Checksum, untuk fungsi Checksum, pilih. SHA256
 - g. Pilih Unggah.
5. Anda sekarang siap untuk membuat template PySpark analisis.

CLI

Untuk menyimpan skrip pengguna dan lingkungan virtual di S3 menggunakan: AWS CLI

1. Jalankan perintah berikut:

```
aws s3 cp --checksum-algorithm sha256 pyspark_venv.tar.gz s3://ARTIFACT-BUCKET/EXAMPLE-PREFIX/
```

2. Anda sekarang siap untuk membuat template PySpark analisis.

Note

Jika Anda perlu memperbarui skrip atau lingkungan virtual:

1. Unggah versi baru sebagai objek terpisah.
2. Buat template analisis baru menggunakan artefak baru.
3. Menghentikan template lama.
4. Simpan artefak asli di S3 jika template lama mungkin masih diperlukan.

Membuat template PySpark analisis

Note

Parameter adalah string yang disediakan pengguna yang dapat berisi konten arbitrer.

- Tinjau kode untuk memastikan parameter ditangani dengan aman untuk mencegah perilaku tak terduga dalam analisis Anda.
- Desain penanganan parameter untuk bekerja dengan aman terlepas dari nilai parameter apa yang disediakan pada waktu pengiriman.

Prasyarat

Sebelum Anda membuat template PySpark analisis, Anda harus memiliki:

- Keanggotaan dalam AWS Clean Rooms kolaborasi aktif

- Akses ke setidaknya satu tabel yang dikonfigurasi dalam kolaborasi aktif
- Izin untuk membuat templat analisis
- Skrip pengguna Python dan lingkungan virtual yang dibuat dan disimpan di S3
 - Bucket S3 mengaktifkan versi. Untuk informasi selengkapnya, lihat [Menggunakan pembuatan versi di bucket S3](#)
 - Bucket S3 dapat menghitung checksum SHA-256 untuk artefak yang diunggah. Untuk informasi selengkapnya, lihat [Menggunakan checksum](#)
- Izin untuk membaca kode dari bucket S3

Untuk informasi tentang membuat peran layanan yang diperlukan, lihat [Buat peran layanan untuk membaca kode dari bucket S3 \(peran template PySpark analisis\)](#).

Prosedur berikut menjelaskan proses pembuatan templat PySpark analisis menggunakan [AWS Clean Rooms konsol](#). Ini mengasumsikan bahwa Anda telah membuat skrip pengguna dan file lingkungan virtual dan menyimpan skrip pengguna dan file lingkungan virtual Anda dalam ember Amazon S3.

 Note

Anggota yang membuat template PySpark analisis juga harus menjadi anggota yang menerima hasil.

Untuk informasi tentang cara membuat template PySpark analisis menggunakan AWS SDKs, lihat [Referensi AWS Clean Rooms API](#).

Untuk membuat template PySpark analisis

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Akun AWS yang akan berfungsi sebagai pembuat kolaborasi.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pada tab Template, buka bagian Analisis template yang dibuat oleh Anda.
5. Pilih Buat templat analisis.
6. Pada halaman template Buat analisis, untuk Detail,

- a. Masukkan Nama untuk templat analisis.
 - b. (Opsional) Masukkan Deskripsi.
 - c. Untuk Format, pilih PySparkopsi.
7. Untuk Definisi,
- a. Tinjau Prasyarat dan pastikan setiap prasyarat terpenuhi sebelum melanjutkan.
 - b. Untuk file Entry point, masukkan bucket S3 atau pilih Browse S3.
 - c. (Opsional) Untuk file Libraries, masukkan bucket S3 atau pilih Browse S3.
8. Untuk Parameter - opsional, jika Anda ingin menambahkan parameter untuk membuat template analisis Anda dapat digunakan kembali:
- a. Pilih Add parameter (Tambahkan parameter).
 - b. Masukkan nama Parameter.

Nama parameter harus dimulai dengan huruf atau garis bawah, diikuti oleh karakter alfanumerik atau garis bawah.

- c. Untuk Type, STRING secara otomatis dipilih sebagai satu-satunya tipe yang didukung untuk template PySpark analisis.
- d. (Opsional) Masukkan nilai Default untuk parameter.

Jika Anda memberikan nilai default, job runner dapat menggunakan nilai ini saat menjalankan pekerjaan tanpa secara eksplisit memberikan nilai parameter.

- e. Untuk menambahkan lebih banyak parameter, pilih Tambahkan parameter lain dan ulangi langkah sebelumnya.

 Note

Anda dapat menentukan hingga 50 parameter per templat PySpark analisis. Setiap nilai parameter bisa sampai 1.000 karakter.

9. Untuk Tabel yang direferensikan dalam definisi,
- Jika semua tabel yang direferensikan dalam definisi telah dikaitkan dengan kolaborasi:
 - Biarkan kotak centang Semua tabel yang direferensikan dalam definisi telah dikaitkan dengan kolaborasi yang dipilih.

- Di bawah Tabel yang terkait dengan kolaborasi, pilih semua tabel terkait yang direferensikan dalam definisi.
- Jika semua tabel yang direferensikan dalam definisi belum dikaitkan dengan kolaborasi:
 - Kosongkan kotak centang Semua tabel yang direferensikan dalam definisi telah dikaitkan dengan kolaborasi.
 - Di bawah Tabel yang terkait dengan kolaborasi, pilih semua tabel terkait yang direferensikan dalam definisi.
 - Di bawah Tabel yang akan dikaitkan nanti, masukkan nama tabel.
 - Pilih Daftar tabel lain untuk daftar tabel lain.

10. Untuk konfigurasi pesan galat, pilih salah satu dari berikut ini:

- Pesan kesalahan dasar - mengembalikan pesan kesalahan dasar tanpa mengekspos data yang mendasarinya. Direkomendasikan untuk beban kerja produksi.
- Pesan kesalahan terperinci - mengembalikan pesan kesalahan terperinci untuk pemecahan masalah yang lebih cepat. Direkomendasikan dalam lingkungan pengembangan dan pengujian. Dapat mengekspos data sensitif, termasuk informasi identitas pribadi (PII).

 Note

Saat menggunakan Pesan kesalahan terperinci, semua anggota penyedia data harus menyetujui pengaturan ini untuk templat.

11. Tentukan izin akses Layanan dengan memilih nama peran Layanan yang ada dari daftar tarik-turun.

1. Daftar peran ditampilkan jika Anda memiliki izin untuk membuat daftar peran.

Jika Anda tidak memiliki izin untuk membuat daftar peran, Anda dapat memasukkan Nama Sumber Daya Amazon (ARN) peran yang ingin Anda gunakan.

2. Lihat peran layanan dengan memilih tautan eksternal View in IAM.

Jika tidak ada peran layanan yang ada, opsi untuk Menggunakan peran layanan yang ada tidak tersedia.

Secara default, AWS Clean Rooms tidak mencoba memperbarui kebijakan peran yang ada untuk menambahkan izin yang diperlukan.

 Note

- AWS Clean Rooms memerlukan izin untuk melakukan kueri sesuai dengan aturan analisis. Untuk informasi selengkapnya tentang izin AWS Clean Rooms, lihat [AWS kebijakan terkelola untuk AWS Clean Rooms](#).
- Jika peran tidak memiliki izin yang memadai AWS Clean Rooms, Anda akan menerima pesan galat yang menyatakan bahwa peran tersebut tidak memiliki izin yang memadai untuk peran tersebut. AWS Clean Rooms Kebijakan peran harus ditambahkan sebelum melanjutkan.
- Jika Anda tidak dapat mengubah kebijakan peran, Anda akan menerima pesan galat yang menyatakan bahwa AWS Clean Rooms tidak dapat menemukan kebijakan untuk peran layanan.

12. Jika Anda ingin mengaktifkan Tag untuk sumber daya tabel yang dikonfigurasi, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.
13. Pilih Buat.
14. Anda sekarang siap memberi tahu anggota kolaborasi Anda bahwa mereka dapat [Meninjau template analisis](#). (Opsional jika Anda ingin menanyakan data Anda sendiri.)

 Important

Jangan memodifikasi atau menghapus artefak (skrip pengguna atau lingkungan virtual) setelah membuat templat analisis.

Melakukannya akan:

- Menyebabkan semua pekerjaan analisis future menggunakan template ini gagal.
- Memerlukan pembuatan template analisis baru dengan artefak baru.
- Tidak mempengaruhi pekerjaan analisis yang telah diselesaikan sebelumnya.

Meninjau template PySpark analisis

Ketika anggota lain membuat templat analisis dalam kolaborasi Anda, Anda harus meninjau dan menyetujuinya sebelum dapat digunakan.

Prosedur berikut menunjukkan kepada Anda cara meninjau templat PySpark analisis, termasuk aturan, parameter, dan tabel yang direferensikan. Sebagai anggota kolaborasi, Anda akan menilai apakah template sesuai dengan perjanjian berbagi data dan persyaratan keamanan Anda.

Setelah templat analisis dan disetujui, dapat digunakan dalam pekerjaan di AWS Clean Rooms.

 Note

Saat Anda membawa kode analisis Anda ke dalam kolaborasi, perhatikan hal-hal berikut:

- AWS Clean Rooms tidak memvalidasi atau menjamin perilaku kode analisis.
 - Jika Anda perlu memastikan perilaku tertentu, tinjau kode mitra kolaborasi Anda secara langsung atau bekerja dengan auditor pihak ketiga tepercaya untuk memeriksanya.
- AWS Clean Rooms menjamin bahwa hash SHA-256 dari kode yang tercantum dalam template PySpark analisis cocok dengan kode yang berjalan di lingkungan analisis. PySpark
- AWS Clean Rooms tidak melakukan audit atau analisis keamanan pustaka tambahan yang Anda bawa ke lingkungan.
- Dalam model keamanan bersama:
 - Anda (pelanggan) bertanggung jawab atas keamanan kode yang berjalan di lingkungan.
 - Anda (pelanggan) bertanggung jawab untuk mengatur konfigurasi pesan kesalahan yang sesuai untuk lingkungan.
 - AWS Clean Rooms bertanggung jawab atas keamanan lingkungan, memastikan bahwa
 - hanya kode yang disetujui yang berjalan
 - hanya tabel yang dikonfigurasi tertentu yang dapat diakses
 - satu-satunya tujuan output adalah bucket S3 penerima hasil.

AWS Clean Rooms menghasilkan SHA-256 hash dari skrip pengguna dan lingkungan virtual untuk ulasan Anda. Namun, skrip pengguna dan pustaka yang sebenarnya tidak dapat diakses secara langsung di dalamnya AWS Clean Rooms.

Untuk memvalidasi bahwa skrip pengguna dan pustaka yang dibagikan sama dengan yang direferensikan dalam templat analisis, Anda dapat membuat hash SHA-256 dari file yang dibagikan dan membandingkannya dengan hash templat analisis yang dibuat oleh AWS Clean Rooms Hash dari kode yang dijalankan juga akan ada di log pekerjaan.

Prasyarat

- Sistem operasi Linux/Unix atau Subsistem Windows untuk Linux (WSL)
- File skrip pengguna yang ingin Anda hash
 - Minta pembuat templat analisis membagikan file melalui saluran aman.
- Hash template analisis yang dibuat oleh AWS Clean Rooms

Untuk meninjau template PySpark analisis menggunakan AWS Clean Rooms konsol

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Akun AWS yang akan berfungsi sebagai pembuat kolaborasi.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pada tab Template, buka bagian Analisis template yang dibuat oleh anggota lain.
5. Pilih template analisis yang memiliki status Dapat menjalankan dari Tidak memerlukan ulasan Anda.
6. Pilih Tinjau.
7. Tinjau aturan analisis Ikhtisar, Definisi, dan Parameter (jika ada).

Note

Parameter memungkinkan pelari analisis untuk mengirimkan nilai yang berbeda pada waktu pengiriman. Jika templat analisis mendukung parameter, tinjau bagaimana nilai parameter digunakan dalam kode mitra kolaborasi Anda untuk memastikannya memenuhi persyaratan Anda.

8. Validasi bahwa skrip dan pustaka pengguna bersama sama dengan yang direferensikan dalam templat analisis.
 - a. Buat hash SHA-256 dari file yang dibagikan dan bandingkan dengan hash template analisis yang dibuat oleh AWS Clean Rooms

Anda dapat menghasilkan hash dengan menavigasi ke direktori yang berisi file skrip pengguna Anda dan kemudian menjalankan perintah berikut:

```
sha256sum your_script_filename.py
```

Contoh output:

```
e3b0c44298fc1c149afb4c8996fb92427ae41e4649b934ca495991b7852b855 my_analysis.py
```

- b. Atau, Anda dapat menggunakan fitur checksum Amazon S3. Untuk informasi selengkapnya, lihat [Memeriksa integritas objek di Amazon S3](#) di Panduan Pengguna Amazon S3.
 - c. Alternatif lain adalah melihat hash dari kode yang dieksekusi di log pekerjaan.
9. Tinjau tabel yang dikonfigurasi yang tercantum di bawah Tabel yang direferensikan dalam definisi.

Status di samping setiap tabel akan membaca Template tidak diperbolehkan.

10. Pilih meja.
- a. Untuk menyetujui templat analisis, pilih Izinkan templat di atas meja. Konfirmasikan persetujuan Anda dengan memilih Izinkan.
 - b. Untuk menolak persetujuan, pilih Larang.

Jika Anda telah memilih untuk menyetujui template analisis, anggota yang dapat menjalankan pekerjaan sekarang dapat menjalankan PySpark pekerjaan pada tabel yang dikonfigurasi menggunakan templat PySpark analisis. Lihat informasi yang lebih lengkap di [Menjalankan PySpark pekerjaan](#).

Templat analisis pemecahan masalah PySpark

Saat menjalankan pekerjaan menggunakan templat PySpark analisis, Anda mungkin mengalami kegagalan selama inisialisasi atau eksekusi pekerjaan. Kegagalan ini biasanya berhubungan dengan konfigurasi skrip, izin akses data, atau pengaturan lingkungan.

Untuk informasi selengkapnya tentang PySpark batasan, lihat [PySpark keterbatasan dalam AWS Clean Rooms](#).

Topik

- [Memecahkan masalah kode Anda](#)
- [Pekerjaan template analisis tidak dimulai](#)
- [Pekerjaan template analisis dimulai tetapi gagal selama pemrosesan](#)
- [Pengaturan lingkungan virtual gagal](#)

Memecahkan masalah kode Anda

Untuk membantu Anda mengembangkan dan memecahkan masalah kode Anda, kami sarankan Anda mensimulasikan AWS Clean Rooms sendiri Akun AWS dengan mengaktifkan Pesan kesalahan terperinci dan menjalankan pekerjaan menggunakan data pengujian Anda sendiri.

Anda juga dapat mensimulasikan PySpark AWS Clean Rooms di Amazon EMR Tanpa Server dengan langkah-langkah berikut. Ini akan memiliki perbedaan kecil dengan PySpark AWS Clean Rooms tetapi sebagian besar mencakup bagaimana kode Anda dapat dijalankan.

Untuk mensimulasikan PySpark AWS Clean Rooms di EMR Tanpa Server

1. Buat kumpulan data di Amazon S3, katalogkan di AWS Glue Data Catalog, dan siapkan izin Lake Formation.
2. Daftarkan lokasi S3 dengan Lake Formation menggunakan peran khusus.
3. Buat instans Amazon EMR Studio jika Anda belum memilikinya (Amazon EMR Studio diperlukan untuk menggunakan Amazon EMR Tanpa Server).
4. Buat aplikasi EMR Tanpa Server
 - Pilih versi rilis emr-7.7.0.
 - Pilih ARM64 arsitektur.
 - Pilih Gunakan pengaturan khusus.
 - Nonaktifkan kapasitas pra-inisialisasi.
 - Jika Anda berencana untuk melakukan pekerjaan interaktif, pilih Endpoint interaktif > Aktifkan titik akhir untuk studio EMR.
 - Pilih Konfigurasi tambahan > Gunakan Lake Formation untuk kontrol akses berbutir halus.
 - Buat aplikasi.
5. Gunakan EMR-S baik melalui notebook EMR-Studio atau API. `StartJobRun`

Pekerjaan template analisis tidak dimulai

Penyebab umum

Pekerjaan template analisis dapat gagal segera saat startup karena tiga masalah konfigurasi utama:

- Penamaan skrip salah yang tidak cocok dengan format yang diperlukan

- Fungsi entrypoint yang hilang atau salah diformat dalam skrip pengguna
- Versi Python yang tidak kompatibel di lingkungan virtual

Resolusi

Untuk menyelesaikan:

1. Verifikasi skrip pengguna Anda:

- Periksa apakah skrip pengguna Anda memiliki nama file Python yang valid.

Nama file Python yang valid menggunakan huruf kecil, garis bawah untuk memisahkan kata, dan ekstensi.py.

2. Verifikasi fungsi entrypoint. Jika skrip pengguna Anda tidak memiliki fungsi entrypoint, tambahkan satu.

- a. Buka skrip pengguna Anda.
- b. Tambahkan fungsi entrypoint ini:

```
def entrypoint(context):  
    # Your analysis code here
```

- c. Pastikan nama fungsi dieja persis seperti `entrypoint`.
- d. Verifikasi fungsi menerima `context` parameter.

3. Periksa kompatibilitas versi Python:

- a. Verifikasi lingkungan virtual Anda menggunakan Python 3.9 atau 3.11.
- b. Untuk memeriksa versi Anda, jalankan: `python --version`
- c. Jika perlu, perbarui lingkungan virtual Anda:

```
conda create -n analysis-env python=3.9  
conda activate analysis-env
```

Pencegahan

- Gunakan kode awal template analisis yang disediakan yang mencakup struktur file yang benar.
- Siapkan lingkungan virtual khusus dengan Python 3.9 atau 3.11 untuk semua templat analisis.

- Uji template analisis Anda secara lokal menggunakan alat validasi templat sebelum mengirimkan pekerjaan.
- CI/CD Menerapkan pemeriksaan untuk memverifikasi penamaan skrip dan persyaratan fungsi titik masuk.

Pekerjaan template analisis dimulai tetapi gagal selama pemrosesan

Penyebab umum

Pekerjaan analisis dapat gagal selama eksekusi karena alasan keamanan dan pemformatan ini:

- Upaya akses langsung yang tidak sah ke AWS layanan seperti Amazon S3 atau AWS Glue
- Mengembalikan output dalam format yang salah yang tidak sesuai dengan DataFrame spesifikasi yang diperlukan
- Panggilan jaringan yang diblokir karena pembatasan keamanan di lingkungan eksekusi

Resolusi

Untuk menyelesaikan

1. Hapus akses AWS layanan langsung:
 - a. Cari kode Anda untuk impor dan panggilan AWS layanan langsung.
 - b. Ganti akses S3 langsung dengan metode sesi Spark yang disediakan.
 - c. Gunakan hanya tabel yang telah dikonfigurasi sebelumnya melalui antarmuka kolaborasi.
2. Format output dengan benar:
 - a. Verifikasi semua output adalah Spark DataFrames.
 - b. Perbarui pernyataan pengembalian Anda agar sesuai dengan format ini:

```
return {  
  "results": {  
    "output1": dataframe1  
  }  
}
```

- c. Hapus objek yang tidak DataFrame kembali.

3. Hapus panggilan jaringan:

- a. Identifikasi dan hapus panggilan API eksternal apa pun.
- b. Hapus urllib, permintaan, atau pustaka jaringan serupa.
- c. Hapus koneksi soket atau kode klien HTTP.

Pencegahan

- Gunakan linter kode yang disediakan untuk memeriksa AWS impor dan panggilan jaringan yang tidak sah.
- Uji pekerjaan di lingkungan pengembangan di mana pembatasan keamanan sesuai dengan produksi.
- Ikuti proses validasi skema keluaran sebelum menerapkan pekerjaan.
- Tinjau pedoman keamanan untuk pola akses layanan yang disetujui.

Pengaturan lingkungan virtual gagal

Penyebab umum

Kegagalan konfigurasi lingkungan virtual biasanya terjadi karena:

- Arsitektur CPU yang tidak cocok antara lingkungan pengembangan dan eksekusi
- Masalah pemformatan kode Python yang mencegah inisialisasi lingkungan yang tepat
- Konfigurasi gambar dasar salah dalam pengaturan wadah

Resolusi

Untuk menyelesaikan

1. Konfigurasi arsitektur yang benar:

- a. Periksa arsitektur Anda saat ini dengan `uname -m`.
- b. Perbarui Dockerfile Anda untuk menentukan: ARM64

```
FROM --platform=linux/arm64 public.ecr.aws/amazonlinux/amazonlinux:2023-minimal
```

- c. Membangun kembali wadah Anda dengan `docker build --platform=linux/arm64`.
2. Perbaiki lekukan Python:
 - a. Jalankan pemformat kode Python seperti `black` pada file kode Anda.
 - b. Verifikasi penggunaan spasi atau tab secara konsisten (bukan keduanya).
 - c. Periksa lekukan semua blok kode:

```
def my_function():  
    if condition:  
        do_something()  
    return result
```
 - d. Gunakan IDE dengan penyorotan lekukan Python.
3. Validasi konfigurasi lingkungan:
 - a. Jalankan `python -m py_compile your_script.py` untuk memeriksa kesalahan sintaks.
 - b. Uji lingkungan secara lokal sebelum penerapan.
 - c. Verifikasi semua dependensi terdaftar di `requirements.txt`

Pencegahan

- Gunakan Visual Studio Code atau PyCharm dengan plugin pemformatan Python
- Konfigurasi kait pra-komit untuk menjalankan pemformat kode secara otomatis
- Membangun dan menguji lingkungan secara lokal menggunakan gambar ARM64 dasar yang disediakan
- Terapkan pemeriksaan gaya kode otomatis di CI/CD pipeline Anda

Menganalisis data dalam kolaborasi

Di AWS Clean Rooms, Anda dapat menganalisis data dengan menjalankan kueri atau pekerjaan.

Query adalah metode untuk mengakses dan menganalisis tabel yang dikonfigurasi dalam kolaborasi, menggunakan serangkaian fungsi, kelas, dan variabel yang didukung. Bahasa kueri yang saat ini didukung AWS Clean Rooms adalah SQL. Ada 3 cara untuk menjalankan kueri AWS Clean Rooms: Tulis kode SQL, gunakan templat analisis SQL yang disetujui, atau gunakan UI pembuat Analisis.

Pekerjaan adalah metode untuk mengakses dan menganalisis tabel yang dikonfigurasi dalam kolaborasi menggunakan serangkaian fungsi, kelas, dan variabel yang didukung. Jenis pekerjaan yang saat ini didukung AWS Clean Rooms adalah PySpark. Ada satu cara untuk menjalankan pekerjaan AWS Clean Rooms: dengan menggunakan templat PySpark analisis yang disetujui.

Topik berikut menjelaskan cara menganalisis data AWS Clean Rooms dengan menjalankan kueri atau PySpark pekerjaan SQL.

Topik

- [Menjalankan kueri SQL](#)
- [Menjalankan PySpark pekerjaan](#)

Menjalankan kueri SQL

Note

Anda hanya dapat menjalankan kueri jika anggota yang bertanggung jawab membayar biaya komputasi kueri telah bergabung dengan kolaborasi sebagai anggota aktif.

Sebagai [anggota yang dapat melakukan query](#), Anda dapat menjalankan query SQL dengan:

- Membangun kueri SQL secara manual menggunakan editor kode SQL.
- Menggunakan [template analisis](#) SQL yang disetujui.
- Menggunakan UI pembuat Analisis untuk membuat kueri tanpa harus menulis kode SQL.

Ketika anggota yang dapat melakukan kueri menjalankan kueri SQL pada tabel dalam kolaborasi, AWS Clean Rooms mengasumsikan peran yang relevan untuk mengakses tabel atas nama mereka. AWS Clean Rooms menerapkan aturan analisis yang diperlukan untuk kueri input dan outputnya.

Aturan analisis dan kendala keluaran diberlakukan secara otomatis. AWS Clean Rooms hanya mengembalikan hasil yang sesuai dengan aturan analisis yang ditetapkan.

AWS Clean Rooms mendukung query SQL yang dapat berbeda dari mesin query lainnya. Untuk spesifikasi, lihat [Referensi AWS Clean Rooms SQL](#). Jika Anda ingin menjalankan kueri pada tabel data yang dilindungi dengan privasi diferensial, Anda harus memastikan bahwa kueri Anda kompatibel dengan struktur [kueri tujuan umum](#) Privasi Diferensial. AWS Clean Rooms

Note

Saat menggunakan [Cryptographic Computing untuk Clean Rooms](#), tidak semua operasi SQL menghasilkan hasil yang valid. Misalnya, Anda dapat melakukan COUNT pada kolom terenkripsi tetapi melakukan nomor terenkripsi menyebabkan kesalahan. SUM Selain itu, kueri mungkin juga menghasilkan hasil yang salah. Misalnya, kueri yang SUM disegel kolom menghasilkan kesalahan. Namun, GROUP BY kueri di atas kolom yang disegel tampaknya berhasil tetapi menghasilkan grup yang berbeda dari yang dihasilkan oleh GROUP BY kueri di atas cleartext.

[Anggota yang membayar biaya komputasi kueri](#) dibebankan untuk kueri yang dijalankan dalam kolaborasi.

Anggota yang dapat melakukan kueri dapat memilih beberapa [anggota yang dapat menerima hasil](#) untuk menerima hasil dari satu kueri. Untuk informasi selengkapnya, lihat [Menanyakan tabel yang dikonfigurasi menggunakan editor kode SQL](#). Untuk informasi umum tentang menerima hasil kueri, lihat [Menerima dan menggunakan hasil analisis](#).

Prasyarat

Sebelum Anda menjalankan query SQL, pastikan bahwa Anda memiliki yang berikut:

- Keanggotaan aktif dalam AWS Clean Rooms kolaborasi
- Akses ke setidaknya satu tabel yang dikonfigurasi dalam kolaborasi
- Konfirmasi bahwa anggota yang bertanggung jawab atas biaya komputasi kueri adalah anggota kolaborasi aktif

Untuk informasi tentang cara menanyakan data atau melihat kueri dengan memanggil operasi AWS Clean Rooms [StartProtectedQuery API](#) secara langsung atau menggunakan AWS SDKs, lihat [Referensi AWS Clean Rooms API](#).

Untuk informasi tentang pencatatan kueri, lihat [Analisis masuk AWS Clean Rooms](#).

 Note

Jika Anda menjalankan kueri pada tabel data [terenkripsi](#), hasil dari kolom terenkripsi dienkripsi.

Konfigurasi properti percikan untuk kueri SQL

AWS Clean Rooms memungkinkan Anda untuk secara opsional menyesuaikan perilaku runtime Spark dengan mengonfigurasi properti Spark yang didukung untuk kueri SQL. Properti ini memungkinkan Anda menyempurnakan kinerja, penggunaan memori, dan parameter eksekusi kueri. Dengan fitur ini, Anda memiliki kontrol yang lebih besar atas bagaimana kueri berbasis SPARK diproses, memungkinkan pengoptimalan berdasarkan persyaratan beban kerja spesifik Anda.

Sekarang Anda dapat menyesuaikan pengaturan seperti partisi acak, ambang batas gabungan siaran, dan parameter eksekusi kueri adaptif langsung dari konsol. AWS Clean Rooms Fitur ini sangat berguna untuk kueri kompleks atau kumpulan data besar di mana konfigurasi default mungkin tidak optimal. Dengan menyempurnakan properti Spark ini, Anda berpotensi meningkatkan kinerja kueri, mengurangi konsumsi sumber daya, dan mengelola penggunaan memori dengan lebih baik untuk analisis kolaborasi berbasis Spark Anda.

Untuk memanfaatkan fitur ini, Anda akan menemukan bagian properti Spark baru di antarmuka kueri. Anda dapat memilih dari daftar properti yang didukung dan menentukan nilai kustom. [Anda juga dapat mengonfigurasi properti Spark secara terprogram menggunakan API. StartProtectedQuery](#) Opsi konfigurasi lanjutan ini memberdayakan analis dan insinyur data untuk mengoptimalkan kueri mereka untuk meningkatkan efisiensi dan skalabilitas.

Untuk informasi selengkapnya tentang properti Spark, termasuk nilai default, lihat [Spark Properties dalam dokumentasi](#) Apache Spark.

Topik berikut menjelaskan cara kueri data dalam kolaborasi menggunakan AWS Clean Rooms konsol.

Topik

- [Menanyakan tabel yang dikonfigurasi menggunakan editor kode SQL](#)
- [Menanyakan tabel pemetaan ID menggunakan editor kode SQL](#)
- [Kueri tabel yang dikonfigurasi menggunakan template analisis SQL](#)
- [Menanyakan dengan pembuat analisis](#)
- [Melihat dampak privasi diferensial](#)
- [Melihat kueri terbaru](#)
- [Melihat detail kueri](#)

Menanyakan tabel yang dikonfigurasi menggunakan editor kode SQL

Sebagai anggota yang dapat melakukan kueri, Anda dapat membuat kueri secara manual dengan menulis kode SQL di editor kode SQL. Editor kode SQL terletak di bagian Analisis pada tab Analisis di AWS Clean Rooms konsol.

Editor kode SQL ditampilkan secara default. Jika Anda ingin menggunakan pembuat analisis untuk membuat kueri, lihat [Menanyakan dengan pembuat analisis](#).

Important

Jika Anda mulai menulis kueri SQL di editor kode dan kemudian mengaktifkan UI pembuat Analisis, kueri Anda tidak disimpan.

AWS Clean Rooms mendukung banyak perintah, fungsi, dan kondisi SQL. Untuk informasi selengkapnya, lihat [Referensi AWS Clean Rooms SQL](#).

Tip

Jika pemeliharaan terjadwal terjadi saat kueri sedang berjalan, kueri dihentikan dan digulung kembali. Anda harus memulai ulang kueri.

Untuk menanyakan tabel yang dikonfigurasi menggunakan editor kode SQL

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.

2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang memiliki status kemampuan anggota Anda dari Menjalankan kueri.
4. Pada tab Analisis, di bawah Tabel, lihat daftar tabel dan jenis aturan analisis terkait (Aturan analisis agregasi, Aturan analisis daftar, atau Aturan analisis kustom).

 Note

Jika Anda tidak melihat tabel yang Anda harapkan dalam daftar, mungkin karena alasan berikut:

- Tabel belum [dikaitkan](#).
- Tabel tidak memiliki [aturan analisis yang dikonfigurasi](#).

5. (Opsional) Untuk melihat skema tabel dan kontrol aturan analisis, perluas tabel dengan memilih ikon tanda plus (+).
6. Di bawah bagian Analisis, untuk mode Analisis, pilih Tulis kode SQL.

 Note

Bagian Analisis hanya ditampilkan jika anggota yang dapat menerima hasil dan anggota yang bertanggung jawab untuk membayar biaya komputasi kueri telah bergabung dengan kolaborasi sebagai anggota aktif.

7. Bangun kueri dengan mengetikkan kueri ke editor kode SQL.

Untuk informasi selengkapnya tentang perintah dan fungsi SQL yang didukung, lihat Referensi [AWS Clean Rooms SQL](#).

Anda juga dapat menggunakan opsi berikut untuk membuat kueri Anda.

Use an example query

Untuk menggunakan contoh kueri

1. Pilih tiga titik vertikal di sebelah tabel.
2. Di bawah Sisipkan di editor, pilih Contoh kueri.

 Note

Memasukkan kueri Contoh menambahkannya ke kueri yang sudah ada di editor.

Contoh kueri muncul. Semua tabel yang tercantum di bawah Tabel disertakan dalam kueri.

3. Edit nilai placeholder dalam kueri.

Insert column names or functions

Untuk menyisipkan nama kolom atau fungsi

1. Pilih tiga titik vertikal di sebelah kolom.
2. Di bawah Sisipkan di editor, pilih Nama kolom.
3. Untuk menyisipkan fungsi yang diizinkan secara manual pada kolom,
 - a. Pilih tiga titik vertikal di sebelah kolom.
 - b. Pilih Sisipkan di editor.
 - c. Pilih nama fungsi yang diizinkan (seperti INNER JOIN,, SUMSUM DISTINCT, atau COUNT).
4. Tekan Ctrl+Spasi untuk melihat skema tabel di editor kode.

 Note

Anggota yang dapat query dapat melihat dan menggunakan kolom partisi di setiap asosiasi tabel dikonfigurasi. Pastikan kolom partisi diberi label sebagai kolom partisi dalam tabel yang mendasari AWS Glue tabel yang dikonfigurasi.

5. Edit nilai placeholder dalam kueri.
8. Tentukan jenis Pekerja yang didukung dan Jumlah pekerja.

Anda dapat memilih jenis instans dan jumlah instance (pekerja) untuk menjalankan kueri SQL Anda.

Untuk CR.1X, Anda dapat memilih hingga 128 pekerja atau minimal 4 pekerja.

Untuk CR.4X, Anda dapat memilih hingga 32 pekerja atau minimal 4 pekerja.

Gunakan tabel berikut untuk menentukan jenis dan nomor atau pekerja yang Anda butuhkan untuk kasus penggunaan Anda.

Jenis pekerja	vCPU	Memori (GB)	Penyimpanan (GB)	Jumlah pekerja	Total Unit Pemrosesan Kamar Bersih (CRPU)
CR.1X (standar)	4	30	100	4	8
				128	256
CR.4X	16	120	400	4	32
				32	256

 Note

Jenis pekerja yang berbeda dan jumlah pekerja memiliki biaya terkait. Untuk mempelajari lebih lanjut tentang harga, lihat [AWS Clean Rooms harga](#).

9. Untuk Kirim hasil ke, tentukan siapa yang dapat menerima hasil.

 Note

Untuk menerima hasil, anggota kolaborasi harus dikonfigurasi sebagai penerima hasil dan harus menjadi peserta aktif dalam kolaborasi (Status: Aktif)

10. (Anggota yang hanya dapat melakukan kueri) Kotak centang Gunakan pengaturan hasil default Anda dipilih secara default. Tetap pilih ini jika Anda ingin mempertahankan pengaturan hasil default Anda.

Jika Anda ingin menentukan pengaturan hasil yang berbeda untuk kueri ini, kosongkan kotak centang Gunakan pengaturan hasil default Anda, lalu pilih yang berikut ini.

- a. Format hasil (CSV atau PARQUET)

- b. File hasil (Tunggal atau Ganda)
- c. Tujuan hasil di Amazon S3

Setiap anggota yang dapat menerima hasil dapat menentukan format Hasil, file Hasil, dan tujuan Hasil yang berbeda di Amazon S3.

11. Untuk menentukan properti Spark:

- a. Perluas properti Spark.
- b. Pilih Tambahkan properti Spark.
- c. Pada kotak dialog properti Spark, pilih nama Properti dari daftar dropdown dan masukkan Nilai.

Tabel berikut memberikan definisi untuk setiap properti.

Untuk informasi selengkapnya tentang properti Spark, lihat [Spark Properties](#) di dokumentasi Apache Spark.

Nama Properti	Deskripsi	nilai default
Spark.Task.MaxFailures	Mengontrol berapa kali berturut-turut tugas dapat gagal sebelum pekerjaan gagal. Membutuhkan nilai yang lebih besar dari atau sama dengan 1. Jumlah percobaan ulang yang diizinkan sama dengan nilai ini dikurangi 1. Hitungan kegagalan diatur ulang jika ada upaya yang berhasil. Kegagalan di berbagai tugas tidak menumpuk menuju batas ini.	4
spark.sql.file.maxPartitionBytes	Menetapkan jumlah maksimum byte untuk dikemas ke dalam satu partisi ketika membaca dari sumber berbasis file seperti Parquet, JSON, dan ORC.	128MB
Spark.hadoop.fs.s3.maxRetries	Menetapkan jumlah maksimum percobaan ulang untuk operasi file Amazon S3.	

Nama Properti	Deskripsi	nilai default
spark.network.timeout	Menetapkan batas waktu default untuk semua interaksi jaringan. Mengganti pengaturan batas waktu berikut jika tidak dikonfigurasi: • spark.storage.blockManagerHeartbeatTimeoutMs • spark.shuffle.io.connectionTimeout • Spark.rpc.askTimeout • Spark.rpc.lookupTimeout	120s
spark.rdd.compress	Menentukan apakah untuk mengompres partisi RDD serial menggunakan spark.io.compression.codec. Berlaku StorageLevel untuk MEMORY_ONLY_SER di Java dan Scala, atau .MEMORY_ONLY dengan Python. StorageLevel Mengurangi ruang penyimpanan tetapi membutuhkan waktu pemrosesan CPU tambahan.	SALAH
spark.shuffle.spill.compress	Menentukan apakah untuk mengompres data tumpahan shuffle menggunakan spark.io.compression.codec.	BETUL
spark.sql.adaptive.advisoryPartitionSizeInBytes	Menetapkan ukuran target dalam byte untuk partisi acak selama pengoptimalan adaptif ketika spark.sql.adaptive.enabled benar. Mengontrol ukuran partisi saat menggabungkan partisi kecil atau memisahkan partisi miring.	(nilai spark.sql.adaptive.shuffle.targetPostShuffleInputSize)

Nama Properti	Deskripsi	nilai default
spark.sql.adaptif.autoBroadcastJoinAmbang	Menetapkan ukuran tabel maksimum dalam byte untuk penyiaran ke node pekerja selama bergabung. Berlaku hanya dalam kerangka adaptif. Menggunakan nilai default yang sama seperti spark.sql. autoBroadcastJoinAmbang batas. Setel ke -1 untuk menonaktifkan penyiaran.	(tidak ada)
spark.sql.adaptive.coalescepartitions.enabled	Menentukan apakah akan menggabungkan partisi shuffle bersebelahan berdasarkan spark.sql.adaptive.advisoryPartitionSizeInBytes untuk mengoptimalkan ukuran tugas. Membutuhkan spark.sql.adaptive.enabled agar benar.	BETUL
Spark.sql.adaptive.CoalescePartitions.initialPartitionNum	Mendefinisikan jumlah awal partisi shuffle sebelum penggabungan. Memerlukan spark.sql.adaptive.enabled dan spark.sql.adaptive.CoalescePartitions.enabled untuk menjadi true. Default ke nilai spark.sql.shuffle.partitions.	(tidak ada)
Spark.sql.adaptive.CoalescePartitions.minPartitionSize	Menetapkan ukuran minimum untuk partisi shuffle yang digabungkan untuk mencegah partisi menjadi terlalu kecil selama optimasi adaptif.	1 MB

Nama Properti	Deskripsi	nilai default
spark.sql.adaptive.coalescepartition s.parallelismFirst	Menentukan apakah untuk menghitung ukuran partisi berdasarkan paralelisme cluster bukan spark.sql.adaptive.advisoryPartitionSizeInBytes selama penggabungan partisi. Menghasilkan ukuran partisi yang lebih kecil dari ukuran target yang dikonfigurasi untuk memaksimalkan paralelisme. Kami merekomendasikan pengaturan ini ke false pada cluster sibuk untuk meningkatkan pemanfaatan sumber daya dengan mencegah tugas-tugas kecil yang berlebihan.	BETUL
spark.sql.adaptive.enabled	Menentukan apakah akan mengaktifkan eksekusi kueri adaptif untuk mengoptimalkan kembali rencana kueri selama eksekusi kueri, berdasarkan statistik runtime yang akurat.	BETUL
spark.sql.adaptif.forceOptimizeSkewedBergabung	Menentukan apakah akan memaksa mengaktifkan OptimizeSkewedJoin bahkan jika memperkenalkan shuffle tambahan.	SALAH
spark.sql.adaptif.localShuffleReader.diaktifkan	Menentukan apakah akan menggunakan pembaca shuffle lokal ketika partisi shuffle tidak diperlukan, seperti setelah mengonversi dari gabungan sort-merge ke broadcast-hash bergabung. Membutuhkan spark.sql.adaptive.enabled agar benar.	BETUL

Nama Properti	Deskripsi	nilai default
spark.sql.adaptif. maxShuffledHashJoinLocalMapThreshold	Menetapkan ukuran partisi maksimum dalam byte untuk membangun peta hash lokal. Memprioritaskan gabungan hash yang diacak daripada gabungan sort-merge saat: • Nilai ini sama atau melebihi spark.sql.adaptive.advisoryPartitionSizeInBytes • Semua ukuran partisi berada dalam batas ini Mengesampingkan spark.sql.join.preferSortMergeBergabunglah dengan pengaturan.	0 byte
spark.sql.adaptif. optimizeSkewsInRebalancePartitions.d iaktifkan	Menentukan apakah akan mengoptimalkan partisi shuffle miring dengan membaginya menjadi partisi yang lebih kecil berdasarkan spark.sql.adaptive.advisoryPartitionSizeInBytes. Membutuhkan spark.sql.adaptive.enabled agar benar.	BETUL
spark.sql.adaptif. rebalancePartitionsSmallPartitionFactor	Mendefinisikan faktor ambang ukuran untuk menggabungkan partisi selama pemisahan. Partisi yang lebih kecil dari faktor ini dikalikan dengan spark.sql.adaptive.advisoryPartitionSizeInBytes digabung.	0,2
Spark.sql.adaptive. .skewjoin.enabled	Menentukan apakah untuk menangani data miring dalam bergabung diacak dengan memisahkan dan opsional mereplikasi partisi miring. Berlaku untuk gabungan hash sort-merge dan shuffled. Membutuhkan spark.sql.adaptive.enabled agar benar.	BETUL

Nama Properti	Deskripsi	nilai default
Spark.sql.adaptive.skewJoin.skewedPartitionFactor	Menentukan faktor ukuran yang menentukan kemiringan partisi. Partisi miring ketika ukurannya melebihi keduanya: - Faktor ini dikalikan dengan ukuran partisi median - Nilai spark.sql.adaptive.skewJoin.skewedPartitionThresholdInBytes	5
Spark.sql.adaptive.skewJoin.skewedPartitionThresholdInBytes	Menetapkan ambang ukuran dalam byte untuk mengidentifikasi partisi miring. Partisi miring ketika ukurannya melebihi keduanya: - Ambang batas ini - Ukuran partisi median dikalikan dengan spark.sql.adaptive.skewJoin.skewedPartitionFactor Kami merekomendasikan pengaturan nilai ini lebih besar dari spark.sql.adaptive.advisoryPartitionSizeInBytes.	256MB
spark.sql.autoBroadcastJoinAmbang	Menetapkan ukuran tabel maksimum dalam byte untuk penyiaran ke node pekerja selama bergabung. Setel ke -1 untuk menonaktifkan penyiaran.	10MB
Spark.Sql.BroadcastTimeout	Mengontrol periode batas waktu dalam hitungan detik untuk operasi siaran selama siaran bergabung.	300 detik
spark.sql.cbo.enabled	Menentukan apakah akan mengaktifkan optimasi berbasis biaya (CBO) untuk estimasi statistik rencana.	SALAH

Nama Properti	Deskripsi	nilai default
Spark.sql.cbo.join reorder.dp.star.filter	Menentukan apakah akan menerapkan heuristik filter star-join selama enumerasi bergabung berbasis biaya.	SALAH
Spark.sql.cbo.join reorder.dp.Threshold	Menetapkan jumlah maksimum node bergabung yang diizinkan dalam algoritma pemrograman dinamis.	12
Spark.sql.cbo.join reOrder.enabled	Menentukan apakah akan mengaktifkan pengurutan ulang bergabung dalam optimasi berbasis biaya (CBO).	SALAH
Spark.sql.cbo.plan stats.enabled	Menentukan apakah untuk mengambil jumlah baris dan statistik kolom dari katalog selama pembuatan rencana logis.	SALAH
spark.sql.cbo. starSchemaDetection	Menentukan apakah akan mengaktifkan pengurutan ulang gabungan berdasarkan deteksi skema bintang.	SALAH
spark.sql.file. maxPartitionNum	Menetapkan jumlah maksimum target partisi file split untuk sumber berbasis file (Parquet, JSON, dan ORC). Rescales partisi ketika hitungan awal melebihi nilai ini. Ini adalah target yang disarankan, bukan batas yang dijamin.	(tidak ada)
spark.sql.file. maxRecordsPerBerkas	Menetapkan jumlah maksimum catatan untuk menulis ke satu file. Tidak ada batasan yang berlaku saat disetel ke nol atau nilai negatif.	0

Nama Properti	Deskripsi	nilai default
spark.sql.file.minPartitionNum	Menetapkan jumlah minimum target partisi file terpisah untuk sumber berbasis file (Parquet, JSON, dan ORC). Default ke spark.sql. leafNodeDefaultParallelism. Ini adalah target yang disarankan, bukan batas yang dijamin.	(tidak ada)
spark.sql.inMemoryColumnarStorage.batchsize	Mengontrol ukuran batch untuk caching kolumnar. Meningkatkan ukuran meningkatkan pemanfaatan memori dan kompresi tetapi meningkatkan risiko out-of-memory kesalahan.	10000
spark.sql.inMemoryColumnarStorage.compression	Menentukan apakah akan secara otomatis memilih codec kompresi untuk kolom berdasarkan statistik data.	BETUL
spark.sql.inMemoryColumnarStorage.enableVectorizedReader	Menentukan apakah untuk mengaktifkan pembacaan vektor untuk caching kolumnar.	BETUL
spark.sql.legacy.allowHashOnMapType	Menentukan apakah untuk memungkinkan operasi hash pada struktur data tipe peta. Pengaturan lama ini mempertahankan kompatibilitas dengan penanganan jenis peta versi Spark yang lebih lama.	

Nama Properti	Deskripsi	nilai default
spark.sql.legacy.allowNegativeScaleOfDecimal	Menentukan apakah untuk memungkinkan nilai skala negatif dalam definisi tipe desimal. Setelan lama ini mempertahankan kompatibilitas dengan versi Spark lama yang mendukung skala desimal negatif.	
spark.sql.legacy.castComplexTypesToString.enabled	Menentukan apakah akan mengaktifkan perilaku warisan untuk casting tipe kompleks untuk string. Menjaga kompatibilitas dengan aturan konversi tipe versi Spark yang lebih lama.	
spark.sql.legacy.charVarcharAsString	Menentukan apakah untuk memperlakukan jenis CHAR dan VARCHAR sebagai tipe STRING. Pengaturan lama ini menyediakan kompatibilitas dengan penanganan tipe string versi Spark yang lebih lama.	
spark.sql.legacy.createEmptyCollectionUsingStringType	Menentukan apakah untuk membuat koleksi kosong menggunakan elemen tipe string. Setelan lama ini mempertahankan kompatibilitas dengan perilaku inisialisasi koleksi versi Spark yang lebih lama.	
spark.sql.legacy.exponentLiteralAsDecimal.enabled	Menentukan apakah untuk menafsirkan literal eksponensial sebagai tipe desimal. Pengaturan lama ini mempertahankan kompatibilitas dengan penanganan literal numerik versi Spark yang lebih lama.	
spark.sql.legacy.json.allowEmptyString.enabled	Menentukan apakah untuk mengizinkan string kosong dalam pengolahan JSON. Pengaturan lama ini mempertahankan kompatibilitas dengan perilaku penguraian JSON versi Spark yang lebih lama.	

Nama Properti	Deskripsi	nilai default
spark.sql.legacy.parquet.int96RebaseModelRead	Menentukan apakah akan menggunakan mode rebase INT96 stempel waktu lama saat membaca file Parquet. Pengaturan lama ini mempertahankan kompatibilitas dengan penanganan stempel waktu versi Spark yang lebih lama.	
spark.sql.legacy.timeParserPolicy	Mengontrol perilaku penguraian waktu untuk kompatibilitas mundur. Pengaturan lama ini menentukan bagaimana stempel waktu dan tanggal diuraikan dari string.	
Spark.sql.legacy.typeCoercion.datetimeToString.diaaktifkan	Menentukan apakah akan mengaktifkan perilaku pemaksaan tipe warisan saat mengonversi nilai datetime ke string. Menjaga kompatibilitas dengan aturan konversi datetime versi Spark yang lebih lama.	
spark.sql. maxSinglePartitionByte	Menetapkan ukuran partisi maksimum dalam byte. Perencana memperkenalkan operasi shuffle untuk partisi yang lebih besar untuk meningkatkan paralelisme.	128m
Spark.sql.metadata cache TTLSeconds	Mengontrol time-to-live (TTL) untuk cache metadata. Berlaku untuk metadata file partisi dan cache katalog sesi. Membutuhkan: • Nilai positif lebih besar dari nol • Spark.sql.CatalogImplementation diatur ke sarang • spark.sql.hive. filesourcePartitionFileCacheSize lebih besar dari nol • spark.sql.hive. manageFilesourcePartitions disetel ke true	-1000ms

Nama Properti	Deskripsi	nilai default
spark.sql.optimizer.collapseProjectAlwaysSebaris	Menentukan apakah untuk menciutkan proyeksi yang berdekatan dan ekspresi inline, bahkan ketika itu menyebabkan duplikasi.	SALAH
spark.sql.optimizer.dynamicPartitionPruning.diaktifkan	Menentukan apakah untuk menghasilkan predikat untuk kolom partisi yang digunakan sebagai kunci bergabung.	BETUL
spark.sql.optimizer.enableCsvExpressionOptimasi	Menentukan apakah akan mengoptimalkan ekspresi CSV dalam pengoptimal SQL dengan memangkas kolom yang tidak perlu dari operasi from_csv.	BETUL
spark.sql.optimizer.enableJsonExpressionOptimasi	Menentukan apakah untuk mengoptimalkan ekspresi JSON dalam pengoptimal SQL dengan: • Memangkas kolom yang tidak perlu dari operasi from_json • Menyederhanakan kombinasi from_json dan to_json • Mengoptimalkan operasi named_struct	BETUL
Spark.sql.optimizer.excludedrules	Mendefinisikan aturan pengoptimal untuk dinonaktifkan, diidentifikasi dengan nama aturan yang dipisahkan koma. Beberapa aturan tidak dapat dinonaktifkan karena diperlukan untuk kebenaran. Pengoptimal mencatat aturan mana yang berhasil dinonaktifkan.	(tidak ada)

Nama Properti	Deskripsi	nilai default
Spark.sql.Optimizer.Runtime.BloomFilter.applicationSideScanSizeThreshold	Menetapkan ukuran pemindaian agregat minimum dalam byte yang diperlukan untuk menyuntikkan filter Bloom di sisi aplikasi.	10GB
Spark.sql.Optimizer.Runtime.BloomFilter.creationSideThreshold	Mendefinisikan ambang batas ukuran maksimum untuk menyuntikkan filter Bloom di sisi pembuatan.	10MB
Spark.sql.Optimizer.Runtime.BloomFilter.enabled	Menentukan apakah akan menyisipkan filter Bloom untuk mengurangi data shuffle ketika satu sisi gabungan shuffle memiliki predikat selektif.	BETUL
Spark.sql.Optimizer.Runtime.BloomFilter.expectedNumItems	Mendefinisikan jumlah default item yang diharapkan dalam filter Bloom runtime.	1000000
Spark.sql.Optimizer.Runtime.BloomFilter.maxNumBits	Menetapkan jumlah maksimum bit yang diizinkan dalam filter Bloom runtime.	67108864
Spark.sql.Optimizer.Runtime.BloomFilter.maxNumItems	Menetapkan jumlah maksimum item yang diharapkan yang diizinkan dalam filter Bloom runtime.	4000000
spark.sql.optimizer.runtime.bloomfilter.number.threshold	Membatasi jumlah maksimum filter runtime non-DPP yang diizinkan per kueri untuk mencegah out-of-memory kesalahan pada driver.	10

Nama Properti	Deskripsi	nilai default
spark.sql.optimize r.runtime.bloomfilter.numbits	Mendefinisikan jumlah default bit yang digunakan dalam filter Bloom runtime.	8388608
spark.sql.optimize r.runtime.rowlevelOperationGroupFilter.enabled	Menentukan apakah akan mengaktifkan penyaringan grup runtime untuk operasi tingkat baris. Memungkinkan sumber data untuk: - Pangkas seluruh kelompok data (seperti file atau partisi) menggunakan filter sumber data - Jalankan kueri runtime untuk mengidentifikasi catatan yang cocok - Buang grup yang tidak perlu untuk menghindari penulisan ulang yang mahal Pembatasan: - Tidak semua ekspresi dapat dikonversi ke filter sumber data - Beberapa ekspresi memerlukan evaluasi Spark (seperti subquery)	BETUL
Spark.Sql.Optimizer.RuntimeFilter.semiJoinReduction.diaktifkan	Menentukan apakah akan menyisipkan semi-join untuk mengurangi data shuffle ketika salah satu sisi gabungan shuffle memiliki predikat selektif.	SALAH

Nama Properti	Deskripsi	nilai default
Spark.sql.parquet. AggregatePushdown	Menentukan apakah akan menekan agregat ke Parquet untuk optimasi. Mendukung: • MIN dan MAX untuk jenis boolean, integer, float, dan tanggal • COUNT untuk semua tipe data Melempar pengecualian jika statistik hilang dari footer file Parquet apa pun.	SALAH
spark.sql.parquet. columnarReaderBatchUkuran	Mengontrol jumlah baris di setiap batch pembaca vektor Parquet. Pilih nilai yang menyeimbangkan overhead kinerja dan penggunaan memori untuk mencegah out-of-memory kesalahan.	4096
Spark.sql.Session. Timezone	Mendefinisikan zona waktu sesi untuk menangani stempel waktu dalam literal string dan konversi objek Java. Menerima: • Berbasis wilayah IDs dalam area/city format (seperti Amerika/Los_Angeles) • Offset zona dalam format (+/-) HH, (+/-) HH: mm, atau (+/-) HH: mm: SS (seperti -08 atau + 01:00) • UTC atau Z sebagai alias untuk + 00:00	(nilai zona waktu lokal)
spark.sql.shuffle. partisi	Menetapkan jumlah default partisi untuk pengocokan data selama bergabung atau agregasi. Tidak dapat diubah antara restart kueri streaming terstruktur dari lokasi pos pemeriksaan yang sama.	200

Nama Properti	Deskripsi	nilai default
spark.sql.shuffledHashJoinFaktor	Mendefinisikan faktor perkalian yang digunakan untuk menentukan kelayakan bergabung hash shuffle. Gabungan hash shuffle dipilih ketika ukuran data sisi kecil dikalikan dengan faktor ini kurang dari ukuran data sisi besar.	3
spark.sql.sources.parallelPartitionDiscovery.ambang	Menetapkan jumlah maksimum jalur untuk daftar file sisi driver dengan sumber berbasis file (Parquet, JSON, dan ORC). Ketika terlampaui selama penemuan partisi, file terdaftar menggunakan pekerjaan terdistribusi Spark terpisah.	32
spark.sql.statistical.histogram.enabled	Menentukan apakah akan menghasilkan histogram equi-height selama perhitungan statistik kolom untuk meningkatkan akurasi estimasi. Membutuhkan pemindaian tabel tambahan di luar yang dibutuhkan untuk statistik kolom dasar.	SALAH

12. Pilih Jalankan.

Note

Anda tidak dapat menjalankan kueri jika anggota yang dapat menerima hasil belum mengonfigurasi setelan hasil kueri.

13. Lihat Hasilnya.

Untuk informasi selengkapnya, lihat [Menerima dan menggunakan hasil analisis](#).

14. Lanjutkan untuk menyesuaikan parameter dan jalankan kueri Anda lagi, atau pilih tombol + untuk memulai kueri baru di tab baru.

 Note

AWS Clean Rooms bertujuan untuk memberikan pesan kesalahan yang jelas. Jika pesan kesalahan tidak memiliki detail yang cukup untuk membantu Anda memecahkan masalah, hubungi tim akun. Berikan mereka deskripsi tentang bagaimana kesalahan terjadi dan pesan kesalahan (termasuk pengidentifikasi apa pun). Untuk informasi selengkapnya, lihat [Pemecahan masalah AWS Clean Rooms](#).

Menanyakan tabel pemetaan ID menggunakan editor kode SQL

Prosedur berikut menjelaskan cara menjalankan kueri gabungan multi-tabel pada tabel pemetaan ID untuk bergabung `sourceId` dengan `targetId`.

Sebelum Anda menanyakan tabel pemetaan ID, tabel pemetaan ID harus berhasil diisi.

Untuk menanyakan tabel pemetaan ID menggunakan editor kode SQL

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang memiliki status kemampuan anggota Anda dari Menjalankan kueri.
4. Pada tab Analisis, buka bagian Analisis.

 Note

Bagian Analisis hanya ditampilkan jika anggota yang dapat menerima hasil dan anggota yang bertanggung jawab untuk membayar biaya komputasi kueri telah bergabung dengan kolaborasi sebagai anggota aktif.

5. Pada tab Analisis, di bawah Tabel, lihat daftar tabel pemetaan ID (di bawah Dikelola oleh AWS Clean Rooms) dan jenis aturan analisis terkait (aturan analisis tabel pemetaan ID).

Note

Jika Anda tidak melihat tabel pemetaan ID yang Anda harapkan dalam daftar, mungkin karena tabel pemetaan ID belum berhasil diisi. Untuk informasi selengkapnya, lihat [Mengisi tabel pemetaan ID yang ada](#).

6. Bangun kueri dengan mengetikkan kueri ke editor kode SQL.

(Opsional) Jika Anda ingin menggunakan contoh kueri	(Opsional) Jika Anda ingin memasukkan nama tabel
- Pilih tiga titik vertikal di sebelah tabel. - Di bawah Sisipkan di editor, pilih Contoh pernyataan JOIN. Note Menyisipkan pernyataan Contoh JOIN menambahkan kueri yang sudah ada di editor. Contoh pernyataan JOIN muncul. - Edit nilai placeholder dalam kueri.	- Pilih tiga titik vertikal di sebelah kolom. - Di bawah Sisipkan di editor, pilih Nama tabel. - Edit nilai placeholder dalam kueri.

7. Tentukan jenis Pekerja yang didukung dan Jumlah pekerja.

Gunakan tabel berikut untuk menentukan jenis dan nomor atau pekerja yang Anda butuhkan untuk kasus penggunaan Anda.

Jenis pekerja	vCPU	Memori (GB)	Penyimpanan (GB)	Jumlah pekerja	Total Unit Pemrosesan Kamar Bersih (CRPU)
CR.1X (standar)	4	30	100	4	8
				128	256
CR.4X	16	120	400	4	32
				32	256

Note

Jenis pekerja yang berbeda dan jumlah pekerja memiliki biaya terkait. Untuk mempelajari lebih lanjut tentang harga, lihat [AWS Clean Rooms harga](#).

8. Pilih Jalankan.

Note

Anda tidak dapat menjalankan kueri jika anggota yang dapat menerima hasil belum mengonfigurasi setelan hasil kueri.

9. Lihat Hasilnya.

Untuk informasi selengkapnya, lihat [Menerima dan menggunakan hasil analisis](#).

10. Lanjutkan untuk menyesuaikan parameter dan jalankan kueri Anda lagi, atau pilih tombol + untuk memulai kueri baru di tab baru.

Note

AWS Clean Rooms bertujuan untuk memberikan pesan kesalahan yang jelas. Jika pesan kesalahan tidak memiliki detail yang cukup untuk membantu Anda memecahkan masalah,

hubungi tim akun. Berikan mereka deskripsi tentang bagaimana kesalahan terjadi dan pesan kesalahan (termasuk pengidentifikasi apa pun). Untuk informasi selengkapnya, lihat [Pemecahan masalah AWS Clean Rooms](#).

Kueri tabel yang dikonfigurasi menggunakan template analisis SQL

Prosedur ini menunjukkan cara menggunakan templat analisis di AWS Clean Rooms konsol untuk menanyakan tabel yang dikonfigurasi dengan aturan Analisis kustom.

Untuk menggunakan template analisis SQL untuk menanyakan tabel yang dikonfigurasi dengan aturan Analisis kustom

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang memiliki status kemampuan anggota Anda dari Menjalankan kueri.
4. Pada tab Analisis, di bawah bagian Tabel, lihat tabel dan jenis aturan analisis terkait (Aturan analisis kustom).

Note

Jika Anda tidak melihat tabel yang Anda harapkan dalam daftar, mungkin karena alasan berikut:

- Tabel belum [dikaitkan](#).
- Tabel tidak memiliki [aturan analisis yang dikonfigurasi](#).

5. Di bawah bagian Analisis, untuk mode Analisis, pilih Jalankan templat analisis dan kemudian pilih templat analisis dari daftar dropdown.
6. Parameter membentuk template analisis SQL akan secara otomatis terisi dalam Definisi.
7. Tentukan jenis Pekerja yang didukung dan Jumlah pekerja.

Gunakan tabel berikut untuk menentukan jenis dan nomor atau pekerja yang Anda butuhkan untuk kasus penggunaan Anda.

Jenis pekerja	vCPU	Memori (GB)	Penyimpanan (GB)	Jumlah pekerja	Total Unit Pemrosesan Kamar Bersih (CRPU)
CR.1X (standar)	4	30	100	4	8
				128	256
CR.4X	16	120	400	4	32
				32	256

 Note

Jenis pekerja yang berbeda dan jumlah pekerja memiliki biaya terkait. Untuk mempelajari lebih lanjut tentang harga, lihat [AWS Clean Rooms harga](#).

8. Tentukan properti Spark yang didukung.
 - a. Pilih Tambahkan properti Spark.
 - b. Pada kotak dialog properti Spark, pilih nama Properti dari daftar dropdown dan masukkan Nilai.

Tabel berikut memberikan definisi untuk setiap properti.

Untuk informasi selengkapnya tentang properti Spark, lihat [Spark Properties](#) di dokumentasi Apache Spark.

Nama Properti	Deskripsi	nilai default
Spark.Task.MaxFailures	Mengontrol berapa kali berturut-turut tugas dapat gagal sebelum pekerjaan gagal. Membutuhkan nilai yang lebih besar dari atau sama dengan 1. Jumlah percobaan	4

Nama Properti	Deskripsi	nilai default
	ulang yang diizinkan sama dengan nilai ini dikurangi 1. Hitungan kegagalan diatur ulang jika ada upaya yang berhasil. Kegagalan di berbagai tugas tidak menumpuk menuju batas ini.	
spark.sql.file.maxPartitionBytes	Menetapkan jumlah maksimum byte untuk dikemas ke dalam satu partisi ketika membaca dari sumber berbasis file seperti Parquet, JSON, dan ORC.	128MB
Spark.hadoop.fs.s3.maxRetries	Menetapkan jumlah maksimum percobaan ulang untuk operasi file Amazon S3.	
spark.network.timeout	Menetapkan batas waktu default untuk semua interaksi jaringan. Mengganti pengaturan batas waktu berikut jika tidak dikonfigurasi: • spark.storage.blockManagerHeartbeatTimeoutMs • spark.shuffle.io.connectionTimeout • Spark.rpc.askTimeout • Spark.rpc.lookupTimeout	120s
spark.rdd.compress	Menentukan apakah untuk mengompres partisi RDD serial menggunakan spark.io.compression.codec. Berlaku StorageLevel untuk MEMORY_ONLY_SER di Java dan Scala, atau .MEMORY_ONLY dengan Python. StorageLevel Mengurangi ruang penyimpanan tetapi membutuhkan waktu pemrosesan CPU tambahan.	SALAH

Nama Properti	Deskripsi	nilai default
<code>spark.shuffle.spill.compress</code>	Menentukan apakah untuk mengompres data tumpahan shuffle menggunakan <code>spark.io.compression.codec</code> .	BETUL
<code>spark.sql.adaptif.advisoryPartitionSizeInBytes</code>	Menetapkan ukuran target dalam byte untuk partisi acak selama pengoptimalan adaptif ketika <code>spark.sql.adaptive.enabled</code> benar. Mengontrol ukuran partisi saat menggabungkan partisi kecil atau memisahkan partisi miring.	(nilai <code>spark.sql.adaptive.shuffle.targetPostShuffleInputSize</code>)
<code>spark.sql.adaptif.autoBroadcastJoinAmbang</code>	Menetapkan ukuran tabel maksimum dalam byte untuk penyiaran ke node pekerja selama bergabung. Berlaku hanya dalam kerangka adaptif. Menggunakan nilai default yang sama seperti <code>spark.sql.autoBroadcastJoinAmbang</code> batas. Setel ke -1 untuk menonaktifkan penyiaran.	(tidak ada)
<code>spark.sql.adaptive.coalescepartitions.enabled</code>	Menentukan apakah akan menggabungkan partisi shuffle bersebelahan berdasarkan <code>spark.sql.adaptive.advisoryPartitionSizeInBytes</code> untuk mengoptimalkan ukuran tugas. Membutuhkan <code>spark.sql.adaptive.enabled</code> agar benar.	BETUL
<code>Spark.sql.adaptive.CoalescePartitions.initialPartitionNum</code>	Mendefinisikan jumlah awal partisi shuffle sebelum penggabungan. Memerlukan <code>spark.sql.adaptive.enabled</code> dan <code>spark.sql.adaptive.CoalescePartitions.enabled</code> untuk menjadi true. Default ke nilai <code>spark.sql.shuffle.partitions</code> .	(tidak ada)

Nama Properti	Deskripsi	nilai default
Spark.sql.adaptive .CoalescePartitions. minPartitionSize	Menetapkan ukuran minimum untuk partisi shuffle yang digabungkan untuk mencegah partisi menjadi terlalu kecil selama optimasi adaptif.	1 MB
spark.sql.adaptive .coalescepartition s.parallelismFirst	Menentukan apakah untuk menghitung ukuran partisi berdasarkan paralelisme cluster bukan spark.sql.adaptive. advisoryPartitionSizeInBytes selama penggabungan partisi. Menghasilkan ukuran partisi yang lebih kecil dari ukuran target yang dikonfigurasi untuk memaksimalkan paralelisme. Kami merekomendasikan pengaturan ini ke false pada cluster sibuk untuk meningkatkan pemanfaatan sumber daya dengan mencegah tugas-tugas kecil yang berlebihan.	BETUL
spark.sql.adaptive .enabled	Menentukan apakah akan mengaktifkan eksekusi kueri adaptif untuk mengoptimalkan kembali rencana kueri selama eksekusi kueri, berdasarkan statistik runtime yang akurat.	BETUL
spark.sql.adaptif. forceOptimizeSkewedBergabung	Menentukan apakah akan memaksa mengaktifkan OptimizeSkewedJoin bahkan jika memperkenalkan shuffle tambahan.	SALAH
spark.sql.adaptif. localShuffleReader .diaktifkan	Menentukan apakah akan menggunakan pembaca shuffle lokal ketika partisi shuffle tidak diperlukan, seperti setelah mengonversi dari gabungan sort-merge ke broadcast-hash bergabung. Membutuhkan spark.sql.adaptive.enabled agar benar.	BETUL

Nama Properti	Deskripsi	nilai default
spark.sql.adaptif. maxShuffledHashJoinLocalMapThreshold	Menetapkan ukuran partisi maksimum dalam byte untuk membangun peta hash lokal. Memprioritaskan gabungan hash yang diacak daripada gabungan sort-merge saat: • Nilai ini sama atau melebihi spark.sql.adaptive.advisoryPartitionSizeInBytes • Semua ukuran partisi berada dalam batas ini Mengesampingkan spark.sql.join.preferSortMergeBergabunglah dengan pengaturan.	0 byte
spark.sql.adaptif.optimizeSkewsInRebalancePartitions.diaaktifkan	Menentukan apakah akan mengoptimalkan partisi shuffle miring dengan membaginya menjadi partisi yang lebih kecil berdasarkan spark.sql.adaptive.advisoryPartitionSizeInBytes. Membutuhkan spark.sql.adaptive.enabled agar benar.	BETUL
spark.sql.adaptif.rebalancePartitionsSmallPartitionFactor	Mendefinisikan faktor ambang ukuran untuk menggabungkan partisi selama pemisahan. Partisi yang lebih kecil dari faktor ini dikalikan dengan spark.sql.adaptive.advisoryPartitionSizeInBytes digabung.	0.2
Spark.sql.adaptive.skewjoin.enabled	Menentukan apakah untuk menangani data miring dalam bergabung diacak dengan memisahkan dan opsional mereplikasi partisi miring. Berlaku untuk gabungan hash sort-merge dan shuffled. Membutuhkan spark.sql.adaptive.enabled agar benar.	BETUL

Nama Properti	Deskripsi	nilai default
Spark.sql.adaptive.skewJoin.skewedPartitionFactor	Menentukan faktor ukuran yang menentukan kemiringan partisi. Partisi miring ketika ukurannya melebihi keduanya: - Faktor ini dikalikan dengan ukuran partisi median - Nilai spark.sql.adaptive.skewJoin.skewedPartitionThresholdInBytes	5
Spark.sql.adaptive.skewJoin.skewedPartitionThresholdInBytes	Menetapkan ambang ukuran dalam byte untuk mengidentifikasi partisi miring. Partisi miring ketika ukurannya melebihi keduanya: - Ambang batas ini - Ukuran partisi median dikalikan dengan spark.sql.adaptive.skewJoin.skewedPartitionFactor Kami merekomendasikan pengaturan nilai ini lebih besar dari spark.sql.adaptive.advisoryPartitionSizeInBytes.	256MB
spark.sql.autoBroadcastJoinAmbang	Menetapkan ukuran tabel maksimum dalam byte untuk penyiaran ke node pekerja selama bergabung. Setel ke -1 untuk menonaktifkan penyiaran.	10MB
Spark.Sql.BroadcastTimeout	Mengontrol periode batas waktu dalam hitungan detik untuk operasi siaran selama siaran bergabung.	300 detik
spark.sql.cbo.enabled	Menentukan apakah akan mengaktifkan optimasi berbasis biaya (CBO) untuk estimasi statistik rencana.	SALAH

Nama Properti	Deskripsi	nilai default
Spark.sql.cbo.join reorder.dp.star.filter	Menentukan apakah akan menerapkan heuristik filter star-join selama enumerasi bergabung berbasis biaya.	SALAH
Spark.sql.cbo.join reorder.dp.Threshold	Menetapkan jumlah maksimum node bergabung yang diizinkan dalam algoritma pemrograman dinamis.	12
Spark.sql.cbo.join reOrder.enabled	Menentukan apakah akan mengaktifkan join reordering dalam optimasi berbasis biaya (CBO).	SALAH
Spark.sql.cbo.plan stats.enabled	Menentukan apakah untuk mengambil jumlah baris dan statistik kolom dari katalog selama pembuatan rencana logis.	SALAH
spark.sql.cbo. starSchemaDetection	Menentukan apakah akan mengaktifkan pengurutan ulang gabungan berdasarkan deteksi skema bintang.	SALAH
spark.sql.file. maxPartitionNum	Menetapkan jumlah maksimum target partisi file split untuk sumber berbasis file (Parquet, JSON, dan ORC). Rescales partisi ketika hitungan awal melebihi nilai ini. Ini adalah target yang disarankan, bukan batas yang dijamin.	(tidak ada)
spark.sql.file. maxRecordsPerBerkas	Menetapkan jumlah maksimum catatan untuk menulis ke satu file. Tidak ada batasan yang berlaku saat disetel ke nol atau nilai negatif.	0

Nama Properti	Deskripsi	nilai default
spark.sql.file.minPartitionNum	Menetapkan jumlah minimum target partisi file terpisah untuk sumber berbasis file (Parquet, JSON, dan ORC). Default ke spark.sql. leafNodeDefaultParallelism. Ini adalah target yang disarankan, bukan batas yang dijamin.	(tidak ada)
spark.sql.inMemoryColumnarStorage.batchsize	Mengontrol ukuran batch untuk caching kolumnar. Meningkatkan ukuran meningkatkan pemanfaatan memori dan kompresi tetapi meningkatkan risiko out-of-memory kesalahan.	10000
spark.sql.inMemoryColumnarStorage.compression	Menentukan apakah akan secara otomatis memilih codec kompresi untuk kolom berdasarkan statistik data.	BETUL
spark.sql.inMemoryColumnarStorage.enableVectorizedReader	Menentukan apakah untuk mengaktifkan pembacaan vektor untuk caching kolumnar.	BETUL
spark.sql.legacy.allowHashOnMapType	Menentukan apakah untuk memungkinkan operasi hash pada struktur data tipe peta. Pengaturan lama ini mempertahankan kompatibilitas dengan penanganan jenis peta versi Spark yang lebih lama.	

Nama Properti	Deskripsi	nilai default
spark.sql.legacy.allowNegativeScaleOfDecimal	Menentukan apakah untuk memungkinkan nilai skala negatif dalam definisi tipe desimal. Setelan lama ini mempertahankan kompatibilitas dengan versi Spark lama yang mendukung skala desimal negatif.	
spark.sql.legacy.castComplexTypesToString.enabled	Menentukan apakah akan mengaktifkan perilaku warisan untuk casting tipe kompleks untuk string. Menjaga kompatibilitas dengan aturan konversi tipe versi Spark yang lebih lama.	
spark.sql.legacy.charVarcharAsString	Menentukan apakah untuk memperlakukan jenis CHAR dan VARCHAR sebagai tipe STRING. Pengaturan lama ini menyediakan kompatibilitas dengan penanganan tipe string versi Spark yang lebih lama.	
spark.sql.legacy.createEmptyCollectionUsingStringType	Menentukan apakah untuk membuat koleksi kosong menggunakan elemen tipe string. Setelan lama ini mempertahankan kompatibilitas dengan perilaku inisialisasi koleksi versi Spark yang lebih lama.	
spark.sql.legacy.exponentLiteralAsDecimal.enabled	Menentukan apakah untuk menafsirkan literal eksponensial sebagai tipe desimal. Pengaturan lama ini mempertahankan kompatibilitas dengan penanganan literal numerik versi Spark yang lebih lama.	
spark.sql.legacy.json.allowEmptyString.enabled	Menentukan apakah untuk mengizinkan string kosong dalam pengolahan JSON. Pengaturan lama ini mempertahankan kompatibilitas dengan perilaku penguraian JSON versi Spark yang lebih lama.	

Nama Properti	Deskripsi	nilai default
spark.sql.legacy.parquet.int96RebaseModelRead	Menentukan apakah akan menggunakan mode rebase INT96 stempel waktu lama saat membaca file Parquet. Pengaturan lama ini mempertahankan kompatibilitas dengan penanganan stempel waktu versi Spark yang lebih lama.	
spark.sql.legacy.timeParserPolicy	Mengontrol perilaku penguraian waktu untuk kompatibilitas mundur. Pengaturan lama ini menentukan bagaimana stempel waktu dan tanggal diuraikan dari string.	
Spark.sql.legacy.typeCoercion.datetimeToString.default	Menentukan apakah akan mengaktifkan perilaku pemaksaan tipe warisan saat mengonversi nilai datetime ke string. Menjaga kompatibilitas dengan aturan konversi datetime versi Spark yang lebih lama.	
spark.sql. maxSinglePartitionByte	Menetapkan ukuran partisi maksimum dalam byte. Perencana memperkenalkan operasi shuffle untuk partisi yang lebih besar untuk meningkatkan paralelisme.	128m
Spark.sql.metadata cache TTLSeconds	Mengontrol time-to-live (TTL) untuk cache metadata. Berlaku untuk metadata file partisi dan cache katalog sesi. Membutuhkan: • Nilai positif lebih besar dari nol • Spark.sql.CatalogImplementation diatur ke sarang • spark.sql.hive. filesourcePartitionFileCacheSize lebih besar dari nol • spark.sql.hive. manageFilesourcePartitions disetel ke true	-1000ms

Nama Properti	Deskripsi	nilai default
spark.sql.optimizer.collapseProjectAlwaysSebaris	Menentukan apakah untuk menciutkan proyeksi yang berdekatan dan ekspresi inline, bahkan ketika itu menyebabkan duplikasi.	SALAH
spark.sql.optimizer.dynamicPartitionPruning.diaktifkan	Menentukan apakah untuk menghasilkan predikat untuk kolom partisi yang digunakan sebagai kunci bergabung.	BETUL
spark.sql.optimizer.enableCsvExpressionOptimasi	Menentukan apakah akan mengoptimalkan ekspresi CSV dalam pengoptimal SQL dengan memangkas kolom yang tidak perlu dari operasi from_csv.	BETUL
spark.sql.optimizer.enableJsonExpressionOptimasi	Menentukan apakah untuk mengoptimalkan ekspresi JSON dalam pengoptimal SQL dengan: • Memangkas kolom yang tidak perlu dari operasi from_json • Menyederhanakan kombinasi from_json dan to_json • Mengoptimalkan operasi named_struct	BETUL
Spark.sql.optimizer.excludedrules	Mendefinisikan aturan pengoptimal untuk dinonaktifkan, diidentifikasi dengan nama aturan yang dipisahkan koma. Beberapa aturan tidak dapat dinonaktifkan karena diperlukan untuk kebenaran. Pengoptimal mencatat aturan mana yang berhasil dinonaktifkan.	(tidak ada)

Nama Properti	Deskripsi	nilai default
Spark.sql.Optimizer.Runtime.BloomFilter.applicationSideScanSizeThreshold	Menetapkan ukuran pemindaian agregat minimum dalam byte yang diperlukan untuk menyuntikkan filter Bloom di sisi aplikasi.	10GB
Spark.sql.Optimizer.Runtime.BloomFilter.creationSideThreshold	Mendefinisikan ambang batas ukuran maksimum untuk menyuntikkan filter Bloom di sisi pembuatan.	10MB
Spark.sql.Optimizer.Runtime.BloomFilter.enabled	Menentukan apakah akan menyisipkan filter Bloom untuk mengurangi data shuffle ketika satu sisi gabungan shuffle memiliki predikat selektif.	BETUL
Spark.sql.Optimizer.Runtime.BloomFilter.expectedNumItems	Mendefinisikan jumlah default item yang diharapkan dalam filter Bloom runtime.	1000000
Spark.sql.Optimizer.Runtime.BloomFilter.maxNumBits	Menetapkan jumlah maksimum bit yang diizinkan dalam filter Bloom runtime.	67108864
Spark.sql.Optimizer.Runtime.BloomFilter.maxNumItems	Menetapkan jumlah maksimum item yang diharapkan yang diizinkan dalam filter Bloom runtime.	4000000
spark.sql.optimizer.runtime.bloomfilter.number.threshold	Membatasi jumlah maksimum filter runtime non-DPP yang diizinkan per kueri untuk mencegah out-of-memory kesalahan pada driver.	10

Nama Properti	Deskripsi	nilai default
spark.sql.optimize r.runtime.bloomfilter.numbits	Mendefinisikan jumlah default bit yang digunakan dalam filter Bloom runtime.	8388608
spark.sql.optimize r.runtime.rowlevelOperationGroupFilter.enabled	Menentukan apakah akan mengaktifkan penyaringan grup runtime untuk operasi tingkat baris. Memungkinkan sumber data untuk: - Pangkas seluruh kelompok data (seperti file atau partisi) menggunakan filter sumber data - Jalankan kueri runtime untuk mengidentifikasi catatan yang cocok - Buang grup yang tidak perlu untuk menghindari penulisan ulang yang mahal Pembatasan: - Tidak semua ekspresi dapat dikonversi ke filter sumber data - Beberapa ekspresi memerlukan evaluasi Spark (seperti subquery)	BETUL
Spark.Sql.Optimizer.RuntimeFilter.semiJoinReduction.diaktifkan	Menentukan apakah akan menyisipkan semi-join untuk mengurangi data shuffle ketika salah satu sisi gabungan shuffle memiliki predikat selektif.	SALAH

Nama Properti	Deskripsi	nilai default
Spark.sql.parquet. AggregatePushdown	Menentukan apakah akan menekan agregat ke Parquet untuk optimasi. Mendukung: • MIN dan MAX untuk jenis boolean, integer, float, dan tanggal • COUNT untuk semua tipe data Melempar pengecualian jika statistik hilang dari footer file Parquet apa pun.	SALAH
spark.sql.parquet. columnarReaderBatchSize	Mengontrol jumlah baris di setiap batch pembaca vektor Parquet. Pilih nilai yang menyeimbangkan overhead kinerja dan penggunaan memori untuk mencegah out-of-memory kesalahan.	4096
Spark.sql.Session. Timezone	Mendefinisikan zona waktu sesi untuk menangani stempel waktu dalam literal string dan konversi objek Java. Menerima: • Berbasis wilayah IDs dalam area/city format (seperti Amerika/Los_Angeles) • Offset zona dalam format (+/-) HH, (+/-) HH: mm, atau (+/-) HH: mm: SS (seperti -08 atau + 01:00) • UTC atau Z sebagai alias untuk + 00:00	(nilai zona waktu lokal)
spark.sql.shuffle. partitions	Menetapkan jumlah default partisi untuk pengocokan data selama bergabung atau agregasi. Tidak dapat diubah antara restart kueri streaming terstruktur dari lokasi pos pemeriksaan yang sama.	200

Nama Properti	Deskripsi	nilai default
spark.sql.shuffledHashJoinFaktor	Mendefinisikan faktor perkalian yang digunakan untuk menentukan kelayakan bergabung hash shuffle. Gabungan hash shuffle dipilih ketika ukuran data sisi kecil dikalikan dengan faktor ini kurang dari ukuran data sisi besar.	3
spark.sql.sources.parallelPartitionDiscovery.ambang	Menetapkan jumlah maksimum jalur untuk daftar file sisi driver dengan sumber berbasis file (Parquet, JSON, dan ORC). Ketika terlampaui selama penemuan partisi, file terdaftar menggunakan pekerjaan terdistribusi Spark terpisah.	32
spark.sql.statistical.histogram.enabled	Menentukan apakah akan menghasilkan histogram equi-height selama perhitungan statistik kolom untuk meningkatkan akurasi estimasi. Membutuhkan pemindaian tabel tambahan di luar yang dibutuhkan untuk statistik kolom dasar.	SALAH

9. Pilih Jalankan.

Note

Anda tidak dapat menjalankan kueri jika anggota yang dapat menerima hasil belum mengonfigurasi setelan hasil kueri.

10. Lanjutkan untuk menyesuaikan parameter dan jalankan kueri Anda lagi, atau pilih tombol + untuk memulai kueri baru di tab baru.

Menanyakan dengan pembuat analisis

Anda dapat menggunakan pembuat analisis untuk membuat kueri tanpa harus menulis kode SQL. Dengan pembuat analisis, Anda dapat membuat kueri untuk kolaborasi yang memiliki:

- Tabel tunggal yang menggunakan [aturan analisis agregasi](#) tanpa diperlukan JOIN
- Dua tabel (satu dari setiap anggota) yang keduanya menggunakan aturan [analisis agregasi](#)
- Dua tabel (satu dari setiap anggota) yang keduanya menggunakan [aturan analisis daftar](#)
- Dua tabel (satu dari setiap anggota) yang keduanya menggunakan aturan analisis agregasi dan dua tabel (satu dari setiap anggota) yang keduanya menggunakan aturan analisis daftar

Jika Anda ingin menulis kueri SQL secara manual, lihat. [Menanyakan tabel yang dikonfigurasi menggunakan editor kode SQL](#)

Pembuat analisis muncul sebagai opsi UI pembuat Analisis di bagian Analisis pada tab Analisis di AWS Clean Rooms konsol.

 Important

Jika Anda mengaktifkan UI pembuat Analisis, mulai membuat kueri di pembuat analisis, lalu matikan UI pembuat Analisis, kueri Anda tidak disimpan.

 Tip

Jika pemeliharaan terjadwal terjadi saat kueri sedang berjalan, kueri dihentikan dan digulung kembali. Anda harus memulai ulang kueri.

Topik berikut menjelaskan cara menggunakan pembuat analisis.

Topik

- [Gunakan pembuat analisis untuk menanyakan satu tabel \(agregasi\)](#)
- [Gunakan pembuat analisis untuk menanyakan dua tabel \(agregasi atau daftar\)](#)

Gunakan pembuat analisis untuk menanyakan satu tabel (agregasi)

Prosedur ini menunjukkan cara menggunakan UI pembuat Analisis di AWS Clean Rooms konsol untuk membuat kueri. Kueri adalah untuk kolaborasi yang memiliki satu tabel yang menggunakan [aturan analisis agregasi](#) tanpa JOIN diperlukan.

Untuk menggunakan pembuat analisis untuk menanyakan satu tabel

Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.

1. Di panel navigasi kiri, pilih Kolaborasi.
2. Pilih kolaborasi yang memiliki status kemampuan anggota Anda dari Menjalankan kueri.
3. Pada tab Analisis, di bawah Tabel, lihat tabel dan jenis aturan analisis terkait. (Jenis aturan analisis harus menjadi aturan analisis Agregasi.)

 Note

Jika Anda tidak melihat tabel yang Anda harapkan, mungkin karena alasan berikut:

- Tabel belum [dikaitkan](#).
- Tabel tidak memiliki [aturan analisis yang dikonfigurasi](#).

4. Di bawah bagian Analisis, aktifkan UI pembuat Analisis.
5. Bangun kueri.

Jika Anda ingin melihat semua metrik agregasi, lewati ke langkah 9.

- a. Untuk metrik Pilih, tinjau metrik agregat yang telah dipilih sebelumnya secara default dan hapus metrik apa pun jika diperlukan.
- b. (Opsional) Untuk Tambah segmen - opsional, pilih satu atau beberapa parameter.

 Note

Tambahkan segmen - opsional hanya ditampilkan jika dimensi ditentukan untuk tabel.

- c. (Opsional) Untuk Tambahkan filter — opsional, pilih Tambahkan filter, lalu pilih Parameter, operator, dan Nilai.

Untuk menambahkan lebih banyak filter, pilih Tambahkan filter lain.

Untuk menghapus filter, pilih Hapus.

 Note

ORDER BY tidak didukung untuk kueri agregasi.
Hanya AND operator yang didukung dalam filter.

- d. (Opsional) Untuk Tambahkan deskripsi — opsional, masukkan deskripsi untuk membantu mengidentifikasi kueri dalam daftar kueri.
6. Perluas kode SQL Pratinjau.
 - a. Lihat kode SQL yang dihasilkan dari pembuat analisis.
 - b. Untuk menyalin kode SQL, pilih Salin.
 - c. Untuk mengedit kode SQL, pilih Edit di editor kode SQL.
7. Pilih Jalankan.

 Note

Anda tidak dapat menjalankan kueri jika anggota yang dapat menerima hasil belum mengonfigurasi setelan hasil kueri.

8. Lanjutkan untuk menyesuaikan parameter dan jalankan kueri Anda lagi, atau pilih tombol + untuk memulai kueri baru di tab baru.

 Note

AWS Clean Rooms bertujuan untuk memberikan pesan kesalahan yang jelas. Jika pesan kesalahan tidak memiliki detail yang cukup untuk membantu Anda memecahkan masalah, hubungi tim akun. Berikan mereka deskripsi tentang bagaimana kesalahan terjadi dan pesan kesalahan (termasuk pengidentifikasi apa pun). Untuk informasi selengkapnya, lihat [Pemecahan masalah AWS Clean Rooms](#).

Gunakan pembuat analisis untuk menanyakan dua tabel (agregasi atau daftar)

Prosedur ini menjelaskan cara menggunakan pembuat analisis di AWS Clean Rooms konsol untuk membuat kueri untuk kolaborasi yang memiliki:

- Dua tabel (satu dari setiap anggota) yang keduanya menggunakan aturan [analisis agregasi](#)
- Dua tabel (satu dari setiap anggota) yang keduanya menggunakan [aturan analisis daftar](#)
- Dua tabel (satu dari setiap anggota) yang keduanya menggunakan aturan analisis agregasi dan dua tabel (satu dari setiap anggota) yang keduanya menggunakan aturan analisis daftar

Untuk menggunakan pembuat analisis untuk menanyakan dua tabel

Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.

1. Di panel navigasi kiri, pilih Kolaborasi.
2. Pilih kolaborasi yang memiliki status kemampuan anggota Anda Jalankan kueri..
3. Pada tab Analisis, di bawah Tabel, lihat dua tabel dan jenis aturan analisis terkait (Aturan analisis agregasi atau Aturan analisis daftar).

Note

Jika Anda tidak melihat tabel yang Anda harapkan dalam daftar, mungkin karena alasan berikut:

- Tabel belum [dikaitkan](#).
- Tabel tidak memiliki [aturan analisis yang dikonfigurasi](#).

4. Di bawah bagian Analisis, aktifkan UI pembuat Analisis.
5. Bangun kueri.

Jika kolaborasi berisi dua tabel yang menggunakan aturan analisis agregasi dan dua tabel yang menggunakan aturan analisis Daftar, pertama pilih Agregasi atau Daftar, lalu ikuti petunjuk berdasarkan aturan analisis yang dipilih.

Jika kedua tabel menggunakan aturan analisis agregasi

1. Untuk metrik Pilih, tinjau metrik agregat yang telah dipilih sebelumnya secara default dan hapus metrik apa pun jika diperlukan.

Jika kedua tabel menggunakan aturan analisis daftar

1. Untuk atribut Pilih, tinjau atribut daftar yang telah dipilih sebelumnya secara default dan hapus metrik apa pun jika diperlukan.

Jika kedua tabel menggunakan aturan analisis agregasi

2. Untuk catatan Pertandingan, pilih satu atau beberapa catatan.

Note

Saat menggunakan pembuat analisis, Anda hanya dapat mencocokkan pada satu pasang kolom.

3. (Opsional) Untuk Tambah segmen - opsional, pilih satu atau beberapa parameter.

Note

Tambahkan segmen - opsional hanya ditampilkan jika dimensi ditentukan untuk tabel.

4. (Opsional) Untuk Tambahkan filter — opsional, pilih Tambahkan filter, lalu pilih parameter, operator, dan nilai.

Untuk menambahkan lebih banyak filter, pilih Tambahkan filter lain.

Untuk menghapus filter, pilih Hapus.

Note

ORDER BY tidak didukung untuk kueri agregasi.
Hanya AND operator yang didukung dalam filter.

Jika kedua tabel menggunakan aturan analisis daftar

2. Untuk catatan Pertandingan, pilih satu atau beberapa catatan.

Note

Saat menggunakan pembuat analisis, Anda hanya dapat mencocokkan pada satu pasang kolom.

3. (Opsional) Untuk Tambahkan filter — opsional, pilih Tambahkan filter, lalu pilih parameter, operator, dan nilai.

Untuk menambahkan lebih banyak filter, pilih Tambahkan filter lain.

Untuk menghapus filter, pilih Hapus.

Note

LIMIT tidak didukung untuk kueri daftar.
Hanya AND operator yang didukung dalam filter.

4. (Opsional) Untuk Tambahkan deskripsi — opsional, masukkan deskripsi untuk membantu mengidentifikasi kueri dalam daftar kueri terbaru.

Jika kedua tabel menggunakan aturan analisis agregasi

Jika kedua tabel menggunakan aturan analisis daftar

5. (Opsional) Untuk Tambahkan deskripsi — opsional, masukkan deskripsi untuk membantu mengidentifikasi kueri dalam daftar kueri terbaru.

6. Perluas kode SQL Pratinjau.

- a. Lihat kode SQL yang dihasilkan dari pembuat analisis.
- b. Untuk menyalin kode SQL, pilih Salin.
- c. Untuk mengedit kode SQL, pilih Edit di editor kode SQL.

7. Pilih Jalankan.

Note

Anda tidak dapat menjalankan kueri jika anggota yang dapat menerima hasil belum mengonfigurasi setelan hasil kueri

8. Lanjutkan untuk menyesuaikan parameter dan jalankan kueri Anda lagi, atau pilih tombol + untuk memulai kueri baru di tab baru.

Note

AWS Clean Rooms bertujuan untuk memberikan pesan kesalahan yang jelas. Jika pesan kesalahan tidak memiliki detail yang cukup untuk membantu Anda memecahkan masalah, hubungi tim akun. Berikan mereka deskripsi tentang bagaimana kesalahan terjadi dan pesan kesalahan (termasuk pengidentifikasi apa pun). Untuk informasi selengkapnya, lihat [Pemecahan masalah AWS Clean Rooms](#).

Melihat dampak privasi diferensial

Secara umum, menulis dan menjalankan kueri tidak berubah ketika privasi diferensial diaktifkan. Namun, Anda tidak dapat menjalankan kueri jika tidak ada cukup anggaran privasi yang tersisa. Saat

Anda menjalankan kueri dan menggunakan anggaran privasi, Anda dapat melihat kira-kira berapa banyak agregasi yang dapat Anda jalankan dan bagaimana hal itu dapat memengaruhi kueri future.

Untuk melihat dampak privasi diferensial dalam sebuah kolaborasi

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang memiliki status detail anggota Anda dari Jalankan kueri.
4. Pada tab Analisis, di bawah Tabel, lihat anggaran privasi yang tersisa. Ini ditampilkan sebagai perkiraan jumlah fungsi agregasi yang tersisa dan Utilitas yang digunakan (diberikan sebagai persentase).

 Note

Perkiraan jumlah fungsi agregat yang tersisa dan persentase Utilitas yang digunakan hanya ditampilkan untuk anggota yang dapat melakukan query.

5. Pilih Lihat dampak untuk melihat seberapa banyak noise yang disuntikkan ke hasil dan kira-kira berapa banyak fungsi agregasi yang dapat Anda jalankan.

Melihat kueri terbaru

Anda dapat melihat kueri yang berjalan dalam 90 hari terakhir pada tab Analisis.

 Note

Jika satu-satunya kemampuan anggota Anda adalah Kontribusikan data, dan Anda bukan [anggota yang membayar biaya komputasi kueri](#), tab Analisis tidak akan muncul di konsol.

Untuk melihat kueri terbaru

Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.

1. Di panel navigasi kiri, pilih Kolaborasi.
2. Pilih kolaborasi.

3. Pada tab Analisis, di bawah Analisis, pilih Semua kueri dari menu tarik-turun, dan lihat kueri yang telah dijalankan dalam 90 hari terakhir.
4. Untuk mengurutkan kueri terbaru berdasarkan Status, pilih status dari daftar tarik-turun Semua status.

Statusnya adalah: Dikirim, Dimulai, Dibatalkan, Sukses, Gagal, dan Timed out.

Melihat detail kueri

Anda dapat melihat detail kueri sebagai anggota yang dapat menjalankan kueri atau sebagai anggota yang dapat menerima hasil.

Untuk melihat detail kueri

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pada tab Analisis, lakukan salah satu hal berikut:
 - Pilih tombol opsi untuk kueri tertentu yang ingin Anda lihat, lalu pilih Lihat detail.
 - Pilih ID kueri yang dilindungi.
5. Pada halaman Query details,
 - Jika Anda adalah anggota yang dapat menjalankan kueri, lihat detail Kueri, teks SQL, dan Hasil.

Anda melihat pesan yang mengonfirmasi bahwa hasil kueri dikirimkan ke anggota yang dapat menerima hasil.

 - Jika Anda adalah anggota yang dapat menerima hasil, lihat detail Kueri dan Hasil.

Menjalankan PySpark pekerjaan

Sebagai [anggota yang dapat melakukan kueri](#), Anda dapat menjalankan PySpark pekerjaan pada tabel yang dikonfigurasi dengan menggunakan [templat PySpark analisis](#) yang disetujui.

Prasyarat

Sebelum Anda menjalankan PySpark pekerjaan, Anda harus memiliki:

- Keanggotaan aktif dalam AWS Clean Rooms kolaborasi
- Akses ke setidaknya satu templat analisis dalam kolaborasi
- Akses ke setidaknya satu tabel yang dikonfigurasi dalam kolaborasi
- Izin untuk menulis hasil PySpark pekerjaan ke bucket S3 tertentu

Untuk informasi tentang membuat peran layanan yang diperlukan, lihat [Buat peran layanan untuk menulis hasil PySpark pekerjaan](#).

- Anggota yang bertanggung jawab untuk membayar biaya komputasi telah bergabung dengan kolaborasi sebagai anggota aktif

Untuk informasi tentang cara menanyakan data atau melihat kueri dengan memanggil operasi AWS Clean Rooms `StartProtectedJob` API secara langsung atau menggunakan AWS SDKs, lihat [Referensi AWS Clean Rooms API](#).

Untuk informasi tentang pencatatan pekerjaan, lihat [Analisis masuk AWS Clean Rooms](#).

Untuk informasi tentang menerima hasil pekerjaan, lihat [Menerima dan menggunakan hasil analisis](#).

Topik berikut menjelaskan cara menjalankan PySpark pekerjaan pada tabel yang dikonfigurasi dalam kolaborasi menggunakan AWS Clean Rooms konsol.

Topik

- [Menjalankan PySpark pekerjaan pada tabel yang dikonfigurasi menggunakan templat PySpark analisis](#)
- [Melihat pekerjaan terbaru](#)
- [Melihat detail tugas](#)

Menjalankan PySpark pekerjaan pada tabel yang dikonfigurasi menggunakan templat PySpark analisis

Prosedur ini menunjukkan cara menggunakan templat PySpark analisis di AWS Clean Rooms konsol untuk menganalisis tabel yang dikonfigurasi dengan aturan Analisis kustom.

Untuk menjalankan PySpark pekerjaan pada tabel yang dikonfigurasi menggunakan template PySpark analisis

Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.

1. Di panel navigasi kiri, pilih Kolaborasi.
2. Pilih kolaborasi yang memiliki status kemampuan anggota Anda dari Jalankan pekerjaan.
3. Pada tab Analisis, di bawah bagian Tabel, lihat tabel dan jenis aturan analisis terkait (Aturan analisis kustom).

 Note

Jika Anda tidak melihat tabel yang Anda harapkan dalam daftar, mungkin karena alasan berikut:

- Tabel belum [dikaitkan](#).
- Tabel tidak memiliki [aturan analisis yang dikonfigurasi](#).

4. Di bawah bagian Analisis, untuk mode Analisis, pilih Jalankan templat analisis.
5. Pilih template PySpark analisis dari daftar dropdown template Analisis.

Parameter dari template PySpark analisis akan secara otomatis terisi dalam Definisi.

6. Jika templat analisis memiliki parameter yang ditentukan, di bawah Parameter, berikan nilai untuk parameter:
 - a. Untuk setiap parameter, lihat nama Parameter dan nilai Default (jika dikonfigurasi).
 - b. Masukkan Nilai untuk setiap parameter yang ingin Anda timpa.

 Note

Jika Anda tidak memberikan nilai tetapi nilai default ada, nilai default akan digunakan.

⚠ Important

Nilai parameter dapat mencapai 1.000 karakter dan mendukung pengkodean UTF-8. Semua nilai parameter diperlakukan sebagai string dan diteruskan ke skrip pengguna Anda melalui objek konteks.

Pastikan skrip pengguna Anda memvalidasi dan menangani nilai parameter dengan aman. Untuk informasi selengkapnya tentang penanganan parameter aman, lihat [Bekerja dengan parameter dalam templat PySpark analisis](#).

7. Tentukan jenis Pekerja yang didukung dan Jumlah pekerja.

Gunakan tabel berikut untuk menentukan jenis dan nomor atau pekerja yang Anda butuhkan untuk kasus penggunaan Anda.

Jenis pekerja	vCPU	Memori (GB)	Penyimpanan (GB)	Jumlah pekerja	Total Unit Pemrosesan Kamar Bersih (CRPU)
CR.1X (standar)	4	30	100	4	8
				128	256
CR.4X	16	120	400	4	32
				32	256

ℹ Note

Jenis pekerja yang berbeda dan jumlah pekerja memiliki biaya terkait. Untuk mempelajari lebih lanjut tentang harga, lihat [AWS Clean Rooms harga](#).

8. Pilih Jalankan.

 Note

Anda tidak dapat menjalankan pekerjaan jika anggota yang dapat menerima hasil belum mengonfigurasi pengaturan hasil pekerjaan.

9. Lanjutkan untuk menyesuaikan parameter dan menjalankan pekerjaan Anda lagi, atau pilih tombol + untuk memulai pekerjaan baru di tab baru.

Melihat pekerjaan terbaru

Anda dapat melihat pekerjaan yang berjalan dalam 90 hari terakhir di tab Analisis.

 Note

Jika satu-satunya kemampuan anggota Anda adalah Kontribusikan data, dan Anda bukan [anggota yang membayar biaya komputasi pekerjaan](#), tab Analisis tidak akan muncul di konsol.

Untuk melihat lowongan terbaru

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pada tab Analisis, di bawah Analisis, pilih Semua pekerjaan dari menu tarik-turun, dan lihat pekerjaan yang telah dijalankan dalam 90 hari terakhir.
5. Untuk mengurutkan pekerjaan terbaru berdasarkan Status, pilih status dari daftar tarik-turun Semua status.

Statusnya adalah: Dikirim, Dimulai, Dibatalkan, Sukses, Gagal, dan Timed out.

Melihat detail tugas

Anda dapat melihat detail pekerjaan sebagai anggota yang dapat menjalankan pekerjaan atau sebagai anggota yang dapat menerima hasil.

Untuk melihat detail pekerjaan

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi.
4. Pada tab Analisis, di bawah Analisis, pilih Semua pekerjaan dari tarik-turun, lalu lakukan salah satu hal berikut:
 - Pilih tombol opsi untuk pekerjaan tertentu yang ingin Anda lihat, lalu pilih Lihat detail.
 - Pilih ID pekerjaan yang Dilindungi.
5. Pada halaman detail Job,
 - Jika Anda adalah anggota yang dapat menjalankan pekerjaan, lihat rincian Job, Job, dan Results.

Anda melihat pesan yang mengonfirmasi bahwa hasil pekerjaan dikirimkan ke anggota yang dapat menerima hasil.
 - Jika Anda adalah anggota yang dapat menerima hasil, lihat rincian dan Hasil Pekerjaan.

Menerima dan menggunakan hasil analisis

[Anggota yang dapat menerima hasil](#) dapat meninjau hasil kueri di AWS Clean Rooms konsol atau di bucket Amazon S3 yang mereka tentukan saat mereka bergabung dengan kolaborasi.

Note

[Hanya untuk tabel data terenkripsi, anggota yang dapat menerima hasil mendekripsi hasil kueri dengan menjalankan klien enkripsi C3R dalam mode dekripsi.](#)

Tujuan Hasil di Amazon S3 tidak dapat berada dalam bucket S3 yang sama dengan sumber data apa pun.

Topik berikut menjelaskan cara menerima hasil analisis menggunakan AWS Clean Rooms konsol.

Topik

- [Menerima hasil kueri](#)
- [Menerima hasil pekerjaan](#)
- [Mengedit nilai default untuk pengaturan hasil kueri](#)
- [Mengedit nilai default untuk pengaturan hasil pekerjaan](#)
- [Menggunakan output kueri di lain Layanan AWS](#)

Untuk informasi tentang cara menanyakan data atau melihat kueri dengan memanggil AWS Clean Rooms API secara langsung atau menggunakan AWS SDKs, lihat [Referensi AWS Clean Rooms API](#).

Untuk informasi tentang pencatatan kueri, lihat [Analisis masuk AWS Clean Rooms](#).

Note

Jika Anda menjalankan kueri pada tabel data terenkripsi, hasil dari kolom terenkripsi dienkripsi.

Menerima hasil kueri

Note

Tujuan Hasil di Amazon S3 tidak dapat berada dalam bucket S3 yang sama dengan sumber data apa pun.

Hasil kueri terletak di bagian default Pengaturan hasil pada tab Analisis di konsol. AWS Clean Rooms

Untuk menerima hasil kueri

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang memiliki status kemampuan anggota Anda dari Menerima hasil.
4. Untuk menerima hasil kueri langsung dari AWS Clean Rooms, pada tab Analisis, di bawah Analisis, pilih Semua kueri dari tarik-turun, lalu di bawah kolom ID kueri yang dilindungi, pilih kueri.
5. Pada halaman Query details, di bawah Results, lakukan salah satu hal berikut:

Jika Anda ingin...	Kemudian pilih...
Salin hasilnya.	Salin
Unduh hasilnya.	Unduh
	 Note Secara default, nama file yang diunduh adalah yang sesuai Query id yang ditampilkan saat kueri dijalankan. AWS Clean Rooms
Lihat hasilnya di Amazon S3.	Lihat di Amazon S3 Konsol Amazon S3 terbuka di tab terpisah.

6. Jika Anda menggunakan data terenkripsi, Anda sekarang dapat [mendekripsi tabel](#) data.

Untuk informasi selengkapnya, lihat [Mendekripsi tabel data dengan klien enkripsi C3R](#).

Menerima hasil pekerjaan

Note

Tujuan Hasil di Amazon S3 tidak dapat berada dalam bucket S3 yang sama dengan sumber data apa pun.

Hasil pekerjaan terletak di bagian default Pengaturan hasil pada tab Analisis di konsol. AWS Clean Rooms

Untuk menerima hasil pekerjaan

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang memiliki status kemampuan anggota Anda dari Menerima hasil.
4. Untuk menerima hasil pekerjaan langsung dari AWS Clean Rooms, pada tab Analisis, di bawah Analisis, pilih Semua pekerjaan dari menu tarik-turun, lalu di bawah kolom ID pekerjaan yang dilindungi, pilih pekerjaan.
5. Pada halaman Detail Job, di bawah Hasil, salin ID Job.

Kembali ke tab Analisis dan perluas default pengaturan Hasil.

Di bawah Tujuan hasil, pilih tautan untuk melihat hasil di Amazon S3.

Konsol Amazon S3 terbuka di tab terpisah.

Di Amazon S3, rekatkan ID Job di bilah Pencarian dan tekan enter.

Folder yang berisi hasil muncul. Pilih folder untuk melihat hasil pekerjaan.

Mengedit nilai default untuk pengaturan hasil kueri

Note

Tujuan Hasil di Amazon S3 tidak dapat berada dalam bucket S3 yang sama dengan sumber data apa pun.

Sebagai anggota yang dapat menerima hasil, Anda dapat mengedit nilai default untuk pengaturan hasil kueri di AWS Clean Rooms konsol.

Untuk mengedit nilai default untuk pengaturan hasil kueri

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang memiliki status kemampuan anggota Anda dari Menerima hasil.
4. Pada tab Analisis, di bawah Pengaturan hasil default, pilih Edit.
5. Pada halaman default Edit pengaturan hasil, ubah salah satu dari berikut ini, sesuai kebutuhan:
 - a. Di bawah Hasil kueri, ubah tujuan Hasil di Amazon S3, format Hasil, atau file Hasil.
 - b. (Opsional) Untuk akses Layanan, jika Anda ingin mengirimkan kueri yang memakan waktu hingga 24 jam ke tujuan S3 Anda, pilih kotak centang Tambahkan peran layanan untuk mendukung kueri yang membutuhkan waktu hingga 24 jam untuk diselesaikan.

Kueri besar yang membutuhkan waktu hingga 24 jam untuk diselesaikan akan dikirimkan ke tujuan S3 Anda.

Jika Anda tidak memilih kotak centang, hanya kueri yang selesai dalam waktu 12 jam yang akan dikirimkan ke lokasi S3 Anda.

- Tentukan izin akses Layanan dengan memilih Buat dan gunakan peran layanan baru atau Gunakan peran layanan yang ada.

Create and use a new service role

- AWS Clean Rooms membuat peran layanan dengan kebijakan yang diperlukan untuk tabel ini.

- Nama peran Layanan default adalah `cleanrooms-query-receiver-
<timestamp>`
- Anda harus memiliki izin untuk membuat peran dan melampirkan kebijakan.

Use an existing service role

1. Pilih nama peran layanan yang ada dari daftar tarik-turun.

Daftar peran ditampilkan jika Anda memiliki izin untuk membuat daftar peran.

Jika Anda tidak memiliki izin untuk membuat daftar peran, Anda dapat memasukkan Nama Sumber Daya Amazon (ARN) peran yang ingin Anda gunakan.

2. Lihat peran layanan dengan memilih tautan eksternal Lihat di IAM.

Jika tidak ada peran layanan yang ada, opsi untuk Menggunakan peran layanan yang ada tidak tersedia.

Secara default, AWS Clean Rooms tidak mencoba memperbarui kebijakan peran yang ada untuk menambahkan izin yang diperlukan.

Note

- AWS Clean Rooms memerlukan izin untuk melakukan kueri sesuai dengan aturan analisis. Untuk informasi selengkapnya tentang izin AWS Clean Rooms, lihat [AWS kebijakan terkelola untuk AWS Clean Rooms](#).
- Jika peran tidak memiliki izin yang memadai AWS Clean Rooms, Anda akan menerima pesan galat yang menyatakan bahwa peran tersebut tidak memiliki izin yang memadai untuk peran tersebut. AWS Clean Rooms Kebijakan peran harus ditambahkan sebelum melanjutkan.
- Jika Anda tidak dapat mengubah kebijakan peran, Anda akan menerima pesan galat yang menyatakan bahwa AWS Clean Rooms tidak dapat menemukan kebijakan untuk peran layanan.

6. Pilih Simpan perubahan.

7. Pengaturan hasil Kueri yang diperbarui muncul di halaman detail kolaborasi.

Mengedit nilai default untuk pengaturan hasil pekerjaan

Note

Tujuan Hasil di Amazon S3 tidak dapat berada dalam bucket S3 yang sama dengan sumber data apa pun.

Sebagai anggota yang dapat menerima hasil, Anda dapat mengedit nilai default untuk pengaturan hasil pekerjaan di AWS Clean Rooms konsol.

Untuk mengedit nilai default untuk pengaturan hasil pekerjaan

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pilih kolaborasi yang memiliki status kemampuan anggota Anda dari Menerima hasil.
4. Pada tab Analisis, di bawah Pengaturan hasil default, pilih Edit.
5. Pada halaman default Edit pengaturan hasil, ubah salah satu dari berikut ini, sesuai kebutuhan:
 - a. Di bawah hasil Job, ubah tujuan Hasil di Amazon S3.
 - b. Di bawah Akses layanan, ubah nama peran layanan yang ada.
6. Pilih Simpan perubahan.
7. Pengaturan hasil Job yang diperbarui muncul di halaman detail kolaborasi.

Menggunakan output kueri di lain Layanan AWS

Output kueri SQL dapat digunakan untuk data benih untuk model Clean Rooms MS. Untuk informasi selengkapnya, lihat [AWS Clean Rooms MS](#).

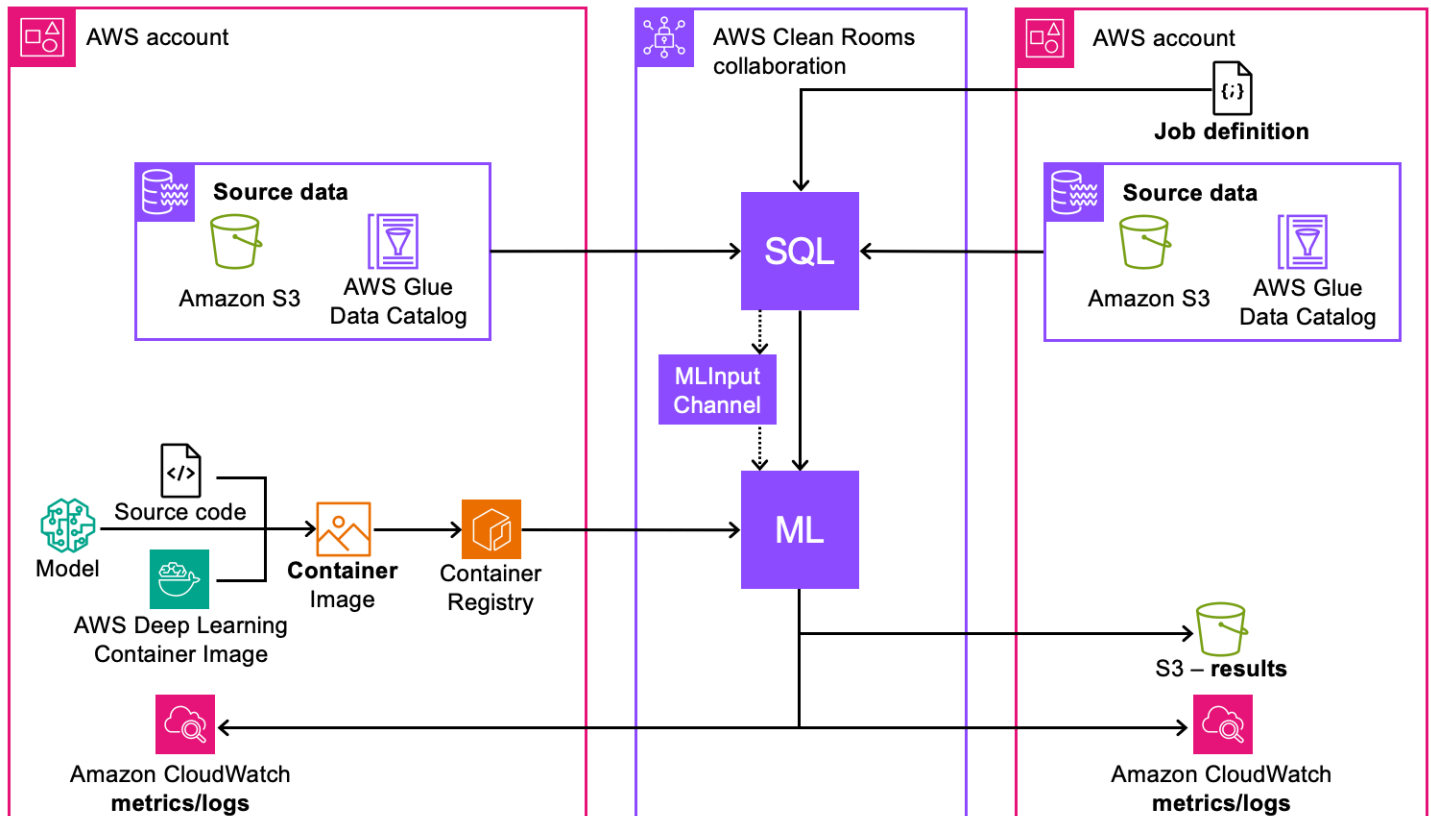
Output kueri dari AWS Clean Rooms tersedia di konsol (jika konsol digunakan untuk menjalankan kueri) dan diunduh di bucket Amazon S3 tertentu. Dari sana, Anda dapat menggunakan output kueri di tempat lain Layanan AWS, seperti Amazon Quick dan Amazon SageMaker AI, tergantung pada bagaimana layanan tersebut menggunakan data dari Amazon S3.

Untuk informasi selengkapnya tentang Amazon Quick, lihat [Dokumentasi Cepat Amazon](#).

Untuk informasi selengkapnya tentang Amazon SageMaker AI, lihat [Dokumentasi Amazon SageMaker AI](#).

Pemodelan kustom di AWS Clean Rooms

Dari sudut pandang teknis, diagram berikut menjelaskan cara kerja pemodelan ML kustom di AWS Clean Rooms ML.



Inilah cara kerja pemodelan ML kustom di Clean Rooms ML:

1. Konfigurasi Sumber Data

- Sumber data dapat disimpan dalam katalog Amazon S3, di, atau AWS Glue Data Catalog Snowflake
- AWS Glue Data Catalog digunakan untuk mengatur dan membuat katalog
- Data dari beberapa Akun AWS dapat digunakan dalam kolaborasi yang sama

2. Kueri SQL dan Pemrosesan Data

- Kueri SQL digunakan untuk mengakses dan memproses data sumber
- Kueri berjalan dalam batas-batas AWS Clean Rooms kolaborasi
- Umpan data yang diproses ke Saluran Input ML untuk pelatihan model

3. Pengembangan Model ML

- Kode sumber untuk model dapat dikembangkan menggunakan AWS Deep Learning Container Images
- Gambar wadah khusus harus dibuat dan disimpan di Amazon Elastic Container Registry

4. Komponen Infrastruktur

- Amazon Elastic Container Registry menyimpan dan mengelola wadah model ML
- Pemrosesan ML terjadi dalam lingkungan AWS Clean Rooms kolaborasi yang aman

5. Pemantauan dan Pencatatan

- Amazon CloudWatch menyediakan metrik dan log untuk kedua pihak yang berkolaborasi
- Pemantauan tersedia di seluruh Akun AWS yang terlibat dalam kolaborasi
- Metrik kinerja dan log operasional dapat diakses oleh pihak terkait

6. Manajemen Hasil

- Akses ke hasil dikontrol sesuai dengan izin kolaborasi

Sebelum Anda memulai, lihat [Prasyarat pemodelan HTML khusus](#) dan [Pedoman penulisan model untuk wadah pelatihan](#) untuk informasi lebih lanjut.

Pembuatan dataset sintetis yang ditingkatkan privasi

Dataset sintetis memiliki sifat statistik yang mirip dengan kumpulan data asli yang menjadi dasarnya, tetapi tidak berisi pengamatan dunia nyata yang ada dalam kumpulan data asli. Dengan menggunakan kumpulan data sintetis yang ditingkatkan privasi, Anda dapat membuka kasus penggunaan pelatihan model pembelajaran mesin (ML) baru yang sebelumnya dicegah oleh masalah privasi data. Saat membuat saluran input ML, Anda dapat menghasilkan data sintetis untuk melindungi informasi sensitif saat melatih model ML.

Saat membuat template dengan data sintetis, Anda harus:

- Memerlukan output template menjadi sintetis
- Klasifikasi kolom skema keluaran sebagai numerik atau kategoris
- Sesuaikan data sintetis berdasarkan kebutuhan organisasi
- Sesuaikan pengaturan privasi:
 - Tetapkan tingkat privasi (epsilon)
 - Konfigurasi ambang privasi

⚠ Warning

Pembuatan data sintetis melindungi dari menyimpulkan atribut individu apakah individu tertentu hadir dalam kumpulan data asli atau atribut pembelajaran individu tersebut hadir. Namun, itu tidak mencegah nilai literal dari kumpulan data asli, termasuk informasi yang dapat diidentifikasi secara pribadi (PII) muncul di kumpulan data sintetis.

Sebaiknya hindari nilai dalam kumpulan data input yang hanya terkait dengan satu subjek data karena ini dapat mengidentifikasi kembali subjek data. Misalnya, jika hanya satu pengguna yang tinggal dalam kode pos, keberadaan kode pos itu dalam kumpulan data sintetis akan mengonfirmasi bahwa pengguna berada dalam kumpulan data asli. Teknik seperti memotong nilai presisi tinggi atau mengganti katalog yang tidak biasa dengan yang lain dapat digunakan untuk mengurangi risiko ini. Transformasi ini dapat menjadi bagian dari query yang digunakan untuk membuat saluran input ML.

Untuk informasi selengkapnya tentang cara menghasilkan data sintetis untuk pelatihan model kustom, lihat [Membuat templat analisis SQL](#).

Template analisis dengan output sintetis hanya dapat digunakan untuk membuat saluran input ML. Untuk informasi selengkapnya, lihat [Membuat saluran input ML di AWS Clean Rooms](#).

Pertimbangan untuk pembuatan data sintetis

Dengan AWS Clean Rooms, anggota kolaborasi dapat membuat kumpulan data sintetis yang secara permanen menghapus identifikasi subjek kumpulan data asli dari kumpulan data kolektif mereka untuk melatih model pembelajaran mesin khusus. Saat membuat kolaborasi, Anda harus mengonfigurasi informasi pembayaran untuk menentukan siapa yang membayar untuk pembuatan data sintetis. Berikut adalah langkah-langkah tingkat tinggi untuk menghasilkan kumpulan data sintetis dan melatih model pembelajaran mesin khusus:

1. Anggota kolaborasi membuat templat analisis yang mencakup:
 - SQL diperlukan untuk mendefinisikan dataset yang akan disintesis.
 - Konfigurasi terkait privasi yang digunakan untuk memastikan data sintetis memenuhi persyaratan kepatuhan penyedia data.
2. Setelah semua penyedia data menyetujui template analisis, pelari kueri kolaborasi membuat saluran input machine learning (ML), menggunakan template.

3. Clean Rooms MS menghasilkan dataset sintetis dan memverifikasi bahwa itu memenuhi ambang privasi yang ditentukan dalam template analisis.
4. Jika semua ambang batas terpenuhi, saluran input ML diisi dengan kumpulan data sintetis.
5. Pelanggan kemudian dapat menggunakan saluran input ML ini untuk melatih model ML kustom yang terkait dengan kolaborasi.

Pertimbangan penting:

- Data sintetis yang dihasilkan di Clean Rooms ML tidak menghapus, menyunting, mengaburkan, atau membersihkan nilai individu apa pun, termasuk informasi identitas pribadi (PII) yang ditemukan dalam kumpulan data asli. Dataset sintetis dihasilkan oleh nilai pengambilan sampel, tetapi tidak seluruh catatan, dari kumpulan data asli.
- Jika kumpulan data asli berisi baris yang serupa, ada kemungkinan data sintetis berisi baris yang terlihat identik dengan baris dalam kumpulan data asli.

Persiapan dataset:

- Hindari kolom dengan distribusi kelas yang tidak seimbang secara signifikan. Ini sangat penting untuk nilai prediksi atau kolom “Y”. Ketidakseimbangan ekstrim mengurangi privasi keseluruhan kumpulan data sintetis.
- Clean Rooms ML tidak mendukung pembuatan data sintetis dari data deret waktu di mana menjaga korelasi di seluruh catatan sekuensial adalah penting.
- Clean Rooms ML tidak mendukung pembuatan data sintetis dari teks atau data tidak terstruktur.
- tipe data berikut didukung:

Nama

tipe

data

BIGINT

BOOLEAN

CHAR

DATE

Nama
tipe
data

DECIMAL

FLOAT

INTEGER

LONG

REAL

SHORT

SMALLINT

TIME

STAMP_LTZ

TIMESTAMP
_NTZ

TINYINT

VARCHAR

Pembatasan:

- Untuk pembuatan data sintetis, jumlah maksimum kolom prediktif adalah satu.
- Jika kolom target kategoris, jumlah maksimum kategori dalam kumpulan data asli adalah 100.
- Dalam kumpulan data asli, jumlah baris harus antara 1.500 dan 2,5 juta dan jumlah kolom maksimum adalah 1.000. Untuk nilai non-null di kolom target, jumlah baris minimum adalah 1.000.

Metrik privasi:

- Clean Rooms ML memberikan skor privasi yang mengukur seberapa terlindungi data sintetis yang dihasilkan terhadap serangan inferensi keanggotaan (MIAs). Layanan ini menyimpan 5% dari data asli dari proses sintesis untuk menghitung skor ini.
- Skor mendekati 50% dianggap baik; skor yang lebih tinggi menunjukkan perlindungan yang lebih sedikit terhadap MIAs. Skor secara signifikan di bawah 50% jarang terjadi dan mungkin karena non-representasi pola dari data asli dalam data yang disintesis.

Model kustom hilir:

- Data sintetis yang dihasilkan di Clean Rooms ML paling cocok untuk melatih model klasifikasi biner dan model klasifikasi multi-kelas hingga lima kelas.
- Model regresi pelatihan menggunakan data sintetis yang dihasilkan di Clean Rooms ML dapat menghasilkan akurasi model yang rendah, yang diukur dengan Root Mean Square Error (RMSE).

Membuat dan bergabung dengan kolaborasi di AWS Clean Rooms ML

Pencipta kolaborasi bertanggung jawab untuk menciptakan kolaborasi, mengundang anggota, dan menetapkan peran mereka. Anggota yang diundang bergabung dengan kolaborasi dan menentukan pengaturan hasil, pengaturan tujuan artefak model terlatih dan menerima tanggung jawab pembayaran, tergantung pada bagaimana kolaborasi diatur.

Menciptakan kolaborasi untuk pembelajaran mesin

Prosedur berikut menunjukkan cara membuat kolaborasi untuk pembelajaran mesin, mengundang satu atau lebih anggota, dan menetapkan anggota yang dapat memulai pelatihan model, menerima hasil, menerima hasil model yang terlatih, termasuk artefak dan metrik model, dan menerima hasil inferensi model. Pembuat kolaborasi juga menugaskan anggota yang akan membayar biaya komputasi kueri, pelatihan model, dan inferensi model.

Console

Untuk membuat kolaborasi untuk pembelajaran mesin (konsol)

1. [Buat kolaborasi dan undang satu atau lebih anggota untuk bergabung dengan kolaborasi](#)
2. Tetapkan kemampuan Anggota berikut untuk Analisis menggunakan kueri dan pekerjaan:

- Tetapkan kueri Jalankan ke anggota yang akan memulai pelatihan model.
 - Tetapkan Menerima hasil dari analisis kepada anggota yang akan menerima hasil kueri.
3. Tetapkan kemampuan Anggota berikut untuk pemodelan ML menggunakan alur kerja yang dibuat khusus:
- Tetapkan Output Terima dari model terlatih kepada anggota yang akan menerima hasil model terlatih, termasuk artefak dan metrik model.
 - Tetapkan Menerima output dari inferensi model ke anggota yang akan menerima hasil inferensi model.
4. Untuk Konfigurasi pembayaran, tentukan anggota yang akan membayar komputasi kueri, pelatihan model, inferensi model, dan biaya pembuatan kumpulan data sintetis.

Dataset sintetis memiliki sifat statistik yang mirip dengan kumpulan data asli yang menjadi dasarnya, tetapi tidak berisi pengamatan dunia nyata yang ada dalam kumpulan data asli. Dengan menggunakan kumpulan data sintetis yang disempurnakan privasi, Anda dapat membuka kasus penggunaan pelatihan model ML baru yang sebelumnya dicegah oleh masalah privasi data. Untuk informasi selengkapnya, lihat Pembuatan data [sintetis yang disempurnakan privasi](#).

Masing-masing biaya ini dapat diberikan kepada anggota yang sama atau berbeda. Jika anggota yang diundang adalah anggota yang bertanggung jawab untuk membayar biaya pembayaran, mereka harus menerima tanggung jawab pembayaran mereka sebelum bergabung dengan kolaborasi.

5. Untuk Konfigurasi keanggotaan, pembuat kolaborasi dapat memutuskan untuk bergabung dengan keanggotaan sekarang atau membuat keanggotaan nanti. Pembuat kolaborasi kemudian harus mengatur konfigurasi ML.
- a. Jika pembuat kolaborasi juga merupakan penerima hasil, mereka juga harus menentukan tujuan dan format hasil kueri dalam pengaturan Hasil default.
 - b. Konfigurasi ML menyediakan peran bagi Clean Rooms untuk memublikasikan metrik ke file Akun AWS. Jika pembuat kolaborasi juga menerima artefak model terlatih, mereka dapat menentukan bucket Amazon S3 yang digunakan untuk menerima hasil.
 - c. Di bagian Konfigurasi ML, pilih Buat konfigurasi ML lalu tentukan tujuan keluaran Model di Amazon S3 dan peran akses Layanan yang diperlukan untuk mengakses lokasi ini.

- d. Jika pembuat kolaborasi adalah anggota yang bertanggung jawab untuk membayar biaya pembayaran, mereka harus menerima tanggung jawab pembayaran mereka sebelum membuat kolaborasi.

API

Untuk membuat kolaborasi untuk machine learning (API)

1. [Buat kolaborasi dan undang satu atau lebih anggota untuk bergabung dengan kolaborasi](#)
2. Tetapkan peran berikut untuk anggota kolaborasi:
 - CAN_QUERY- ditugaskan untuk anggota yang akan memulai pelatihan model dan inferensi.
 - CAN_RECEIVE_MODEL_OUTPUT- ditugaskan untuk anggota yang akan menerima hasil model terlatih.
 - CAN_RECEIVE_INFERENCE_OUTPUT- ditugaskan untuk anggota yang akan menerima hasil inferensi model.

Jika pembuat kolaborasi juga merupakan penerima hasil, mereka juga harus menentukan tujuan dan format hasil kueri selama pembuatan kolaborasi. Mereka juga memberikan peran layanan Amazon Resource Name (ARN) untuk menulis hasil ke tujuan hasil kueri.

3. Tentukan anggota yang akan membayar untuk komputasi kueri, pelatihan model, dan biaya inferensi model. Masing-masing biaya ini dapat diberikan kepada anggota yang sama atau berbeda. Jika anggota yang diundang adalah anggota yang bertanggung jawab untuk membayar biaya pembayaran, mereka harus menerima tanggung jawab pembayaran mereka sebelum bergabung dengan kolaborasi.
4. Kode berikut membuat kolaborasi, mengundang anggota yang dapat menjalankan kueri dan menerima hasil, dan menentukan pembuat kolaborasi sebagai penerima artefak model.

```
import boto3
acr_client= boto3.client('cleanrooms')

collaboration = a_acr_client.create_collaboration(
    members=[
        {
            'accountId': 'invited_member_accountId',
            'memberAbilities': ["CAN_QUERY", "CAN_RECEIVE_RESULTS"],
            'displayName': 'member_display_name'
        }
    ],
```

```

    name='collaboration_name',
    description=collaboration_description,
    creatorMLMemberAbilities= {
        'customMLMemberAbilities':["CAN_RECEIVE_MODEL_OUTPUT",
"CAN_RECEIVE_INFERENCE_OUTPUT"],
    },
    creatorDisplayName='creator_display_name',
    queryLogStatus="ENABLED",
    analyticsEngine="SPARK",
    creatorPaymentConfiguration={
        "queryCompute": {
            "isResponsible": True
        },
        "machineLearning": {
            "modelTraining": {
                "isResponsible": True
            },
            "modelInference": {
                "isResponsible": True
            }
        }
    }
)

collaboration_id = collaboration['collaboration']['id']
print("collaborationId: {collaboration_id}")

member_membership = a_acr_client.create_membership(
    collaborationIdentifier = collaboration_id,
    queryLogStatus = 'ENABLED',
    paymentConfiguration={
        "queryCompute": {
            "isResponsible": True
        },
        "machineLearning": {
            "modelTraining": {
                "isResponsible": True
            },
            "modelInference": {
                "isResponsible": True
            }
        }
    }
)

```

)

5. Pembuat kolaborasi kemudian harus mengatur konfigurasi ML. Konfigurasi ML menyediakan peran bagi Clean Rooms untuk mempublikasikan metrik dan log ke file Akun AWS. Jika pembuat kolaborasi juga menerima hasil (artefak model atau hasil inferensi), mereka dapat menentukan bucket Amazon S3 yang digunakan untuk menerima hasil.

```
import boto3
acr_ml_client= boto3.client('cleanroomsm1')

acr_ml_client.put_ml_configuration(
    membershipIdentifier=membership_id,
    defaultOutputLocation={
        'roleArn': 'arn:aws:iam::account:role/roleName',
        'destination':{
            's3Destination':{
                's3Uri': "s3://bucketName/prefix"
            }
        }
    }
)
```

Bergabung dengan kolaborasi

Setelah pembuat kolaborasi menyelesaikan tugasnya, anggota yang diundang harus menyelesaikan tugasnya.

Console

Untuk membuat keanggotaan dan bergabung dengan kolaborasi (konsol)

1. Anggota yang diundang [membuat keanggotaan dan bergabung dengan kolaborasi](#).
2. Jika anggota yang diundang adalah anggota yang bertanggung jawab untuk membayar, termasuk perhitungan kueri, pelatihan model, dan biaya inferensi model, mereka harus menerima tanggung jawab pembayaran mereka sebelum bergabung dengan kolaborasi.
3. Anggota yang diundang menyiapkan konfigurasi ML, yang menyediakan peran bagi Clean Rooms untuk mempublikasikan metrik model ke file Akun AWS. Jika mereka juga anggota yang menerima artefak model terlatih, mereka harus memberikan ember Amazon S3 tempat artefak model terlatih disimpan.

API

Untuk membuat keanggotaan dan bergabung dengan kolaborasi (API)

1. Jika anggota yang diundang adalah anggota yang dapat menerima hasil, mereka menentukan tujuan dan format hasil kueri. Mereka juga menyediakan ARN peran layanan yang memungkinkan layanan untuk menulis ke tujuan hasil kueri

Jika anggota yang diundang adalah anggota yang bertanggung jawab untuk membayar, termasuk perhitungan kueri, pelatihan model, dan biaya inferensi model, mereka harus menerima tanggung jawab pembayaran mereka sebelum bergabung dengan kolaborasi.

Jika anggota yang diundang adalah anggota yang bertanggung jawab untuk membayar pelatihan model dan inferensi model untuk pemodelan khusus, mereka harus menerima tanggung jawab pembayaran mereka sebelum bergabung dengan kolaborasi.

Kode berikut membuat keanggotaan dengan pencatatan kueri diaktifkan.

```
import boto3
acr_client= boto3.client('cleanrooms')

acr_client.create_membership(
    membershipIdentifier='membership_id',
    queryLogStatus='ENABLED'
)
```

2. Anggota yang diundang menyiapkan konfigurasi ML, yang menyediakan peran bagi Clean Rooms untuk mempublikasikan metrik model ke file Akun AWS. Jika mereka juga anggota yang menerima artefak model terlatih, mereka harus menyediakan ember Amazon S3 tempat artefak model terlatih disimpan.

```
import boto3
acr_ml_client= boto3.client('cleanroomsm1')

acr_ml_client.put_ml_configuration(
    membershipIdentifier='membership_id',
    defaultOutputLocation={
        'roleArn':"arn:aws:iam::account:role/role_name",
        'destination':{
            's3Destination':{
                's3Uri':"s3://bucket_name/prefix"
            }
        }
    }
```

```
}  
}  
)
```

Menyumbangkan data pelatihan di AWS Clean Rooms MS

Setelah pembuat kolaborasi membuat kolaborasi dan anggota yang diundang bergabung, Anda siap untuk menyumbangkan data pelatihan untuk kolaborasi. Setiap anggota dapat menyumbangkan data pelatihan.

Console

Untuk menyumbangkan data pelatihan (konsol)

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Tabel.
3. Pada halaman Tabel, pilih Konfigurasi tabel baru.
4. Untuk Mengonfigurasi tabel baru, untuk Sumber data, pilih Amazon S3, Amazon Athena, atau Snowflake dan selesaikan langkah-langkah berikut, berdasarkan sumber data Anda:

Jika Anda menggunakan	Maka
Amazon S3	1. Pilih Database dari daftar dropdown, lalu pilih Tabel dari database. 2. Untuk Kolom yang diizinkan dalam kolaborasi, pilih Semua kolom atau Daftar kustom. 3. Untuk detail tabel yang Dikonfigurasi, berikan Nama dan Deskripsi opsional untuk tabel ini. 4. Jika Anda ingin melaporkan metrik model, masukkan Nama metrik dan pernyataan Regex yang akan mencari log keluaran untuk menemukan metrik.

Jika Anda menggunakan	Maka
	5. Jika Anda ingin mengaktifkan Tag untuk sumber daya tabel yang dikonfigurasi, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.
Amazon Athena	1. Pilih Database dari daftar dropdown, lalu pilih Tabel dari database. 2. Untuk Kolom yang diizinkan dalam kolaborasi, pilih Semua kolom atau Daftar kustom. 3. Untuk detail tabel yang Dikonfigurasi, berikan Nama dan Deskripsi opsional untuk tabel ini. 4. Jika Anda ingin melaporkan metrik model, masukkan Nama metrik dan pernyataan Regex yang akan mencari log keluaran untuk menemukan metrik. 5. Jika Anda ingin mengaktifkan Tag untuk sumber daya tabel yang dikonfigurasi, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.

Jika Anda menggunakan	Maka
Kepingan salju	1. Tentukan kredensial Snowflake menggunakan ARN rahasia yang ada atau menyimpan rahasia baru untuk tabel ini. 2. Untuk detail tabel dan skema kepingan salju, masukkan detail secara manual atau impor detailnya secara otomatis. 3. Untuk Skema, masukkan nama Kolom dan pilih tipe Data dari daftar dropdown. 4. Jika Anda ingin mengaktifkan Tag untuk sumber daya tabel yang dikonfigurasi, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.

5. Pilih Konfigurasi tabel baru.
6. Pada halaman detail tabel, pilih Konfigurasi aturan analisis untuk mengonfigurasi aturan analisis kustom untuk tabel ini. Aturan analisis kustom membatasi akses ke data Anda. Anda dapat mengizinkan serangkaian kueri pra-otorisasi tertentu pada data Anda atau mengizinkan serangkaian akun tertentu untuk menanyakan data Anda.
 - a. Untuk jenis aturan Analisis, pilih Kustom dan untuk metode Pembuatan, pilih Alur terpandu.
 - b. Pilih Berikutnya.
 - c. Untuk menentukan kontrol analisis, pilih antara Tinjau setiap analisis baru dan Izinkan analisis apa pun oleh kolaborator tertentu.
 - d. Pilih Berikutnya.
 - e. (Opsional) Untuk Tentukan kontrol hasil analisis, untuk Kolom yang tidak diizinkan dalam output tentukan apakah Anda ingin mengecualikan kolom apa pun dari output. Jika Anda memilih Tidak Ada, tidak ada kolom yang dikecualikan dari output. Jika Anda memilih Daftar kustom, Anda dapat menentukan kolom tertentu yang akan dihapus dari output.
 - f. Untuk Analisis tambahan yang diterapkan pada output, tentukan apakah Anda ingin mengizinkan, menolak, atau memerlukan analisis tambahan sebelum hasil dihasilkan.
 - g. Pilih Berikutnya.

- h. (Opsional) Untuk Mengatur privasi diferensial, pilih Matikan.
 - i. Pilih Berikutnya.
 - j. Tinjau informasi pada halaman Tinjau dan konfigurasi, lalu pilih Konfigurasi aturan analisis.
7. Dari halaman detail tabel, pilih Kaitkan dengan kolaborasi.
8. Di kotak dialog Tabel asosiasi, pilih kolaborasi yang ingin Anda kaitkan dengan tabel ini dan pilih Pilih kolaborasi.
9. Pada halaman tabel Rekanan, tinjau dan verifikasi informasi dalam detail asosiasi Tabel, Akses layanan, dan Tag. Pilih tabel Associate.
10. Di Tabel yang terkait dengan tabel Anda, pilih tombol radio di sebelah tabel yang baru saja Anda kaitkan. Dari menu Tindakan, pilih Konfigurasi dalam grup aturan Analisis kolaborasi.
11. Pada halaman aturan Konfigurasi analisis kolaborasi, untuk analisis tambahan yang Diizinkan, pilih apakah anggota kolaborasi atau anggota kolaborasi tertentu dapat melakukan analisis tambahan.

Untuk pengiriman Hasil, pilih anggota mana yang diizinkan untuk menerima hasil dari output kueri.

12. Pilih Konfigurasi aturan analisis.

API

Untuk menyumbangkan data pelatihan (API)

1. Konfigurasi AWS Glue tabel yang ada untuk digunakan AWS Clean Rooms dengan menyediakan tabel dan kolom yang dapat digunakan.

Jalankan kode berikut dengan parameter spesifik Anda.

```
import boto3
acr_client= boto3.client('cleanrooms')

acr_client.create_configured_table(
    name='configured_table_name',
    tableReference= {
        'glue': {
            'tableName': 'glue_table_name',
            'databaseName': 'glue_database_name'
```

```

    }
  },
  analysisMethod="DIRECT_QUERY",
  allowedColumns=["column1", "column2", "column3",...]
)

```

2. Konfigurasi aturan analisis khusus yang membatasi akses ke data Anda. Anda dapat mengizinkan serangkaian kueri pra-otorisasi tertentu pada data Anda atau mengizinkan serangkaian akun tertentu untuk menanyakan data Anda.

Jalankan kode berikut dengan parameter spesifik Anda.

```

import boto3
acr_client= boto3.client('cleanrooms')

acr_client.create_configured_table_analysis_rule(
    configuredTableIdentifier='configured_table_id',
    analysisRuleType='CUSTOM',
    analysisRulePolicy= {
        'v1': {
            'custom': {
                'allowedAnalyses': ['ANY_QUERY'],
                'allowedAnalysisProviders': ['query_runner_account'],
                'additionalAnalyses': "REQUIRED"
            }
        }
    }
)

```

Dalam contoh ini, akun tertentu diizinkan untuk menjalankan kueri apa pun pada data dan analisis tambahan diperlukan.

3. Kaitkan tabel yang dikonfigurasi ke kolaborasi dan berikan peran akses layanan ke AWS Glue tabel.

Jalankan kode berikut dengan parameter spesifik Anda.

```

import boto3
acr_client= boto3.client('cleanrooms')

acr_client.create_configured_table_association(
    name='configured_table_association_name',
    membershipIdentifier='membership_id',

```

```

configuredTableIdentifier='configured_table_id',
roleArn='arn:aws:iam::account:role/role_name'
)

```

Note

Peran layanan ini memiliki izin untuk tabel. Peran layanan hanya dapat diasumsikan oleh AWS Clean Rooms untuk menjalankan kueri yang diizinkan atas nama anggota yang dapat melakukan kueri. Tidak ada anggota kolaborasi (selain pemilik data) yang memiliki akses ke tabel yang mendasarinya dalam kolaborasi. Pemilik data dapat menonaktifkan privasi diferensial untuk membuat tabel mereka tersedia untuk kueri oleh anggota lain.

4. Terakhir, tambahkan aturan analisis ke asosiasi tabel yang dikonfigurasi.

Jalankan kode berikut dengan parameter spesifik Anda.

```

import boto3
acr_client= boto3.client('cleanrooms')

acr_client.create_configured_table_association_analysis_rule(

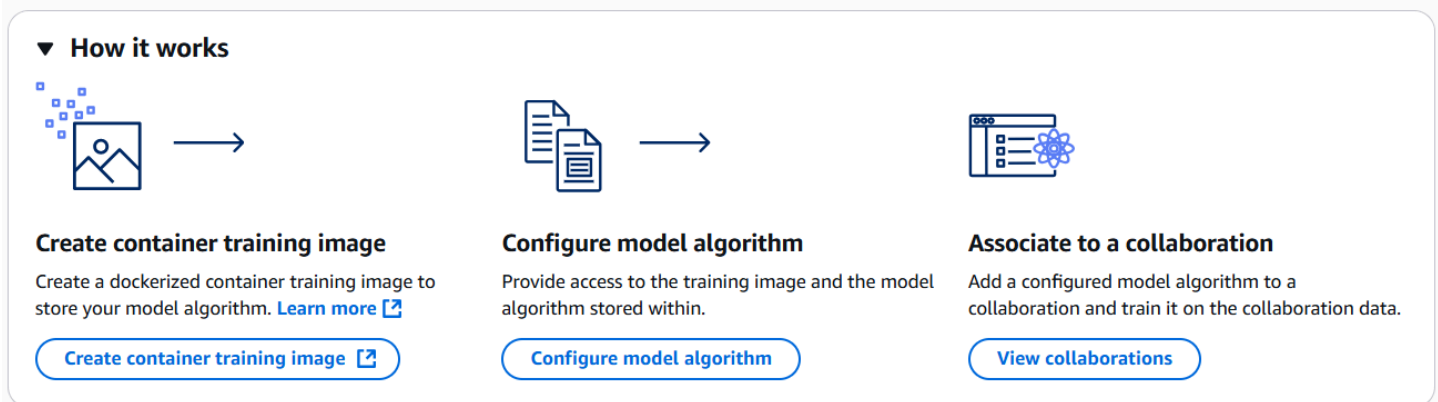
    configuredTableAssociationIdentifier='configured_table_association_identifier',
    membershipIdentifier='membership_id',
    configuredTableIdentifier='configured_table_id',
    analysisRuleType = 'CUSTOM',
    analysisRulePolicy= {
        'v1': {
            'custom': {
                'allowedAdditionalAnalyses':
                ['configured_model_algorithm_association_arns'],
                'allowedResultReceivers': ['query_runner_account']
            }
        }
    }
)

```

Mengonfigurasi algoritme model di AWS Clean Rooms ML

Setelah Anda [membuat gambar pelatihan kontainer](#), Anda harus mengonfigurasi algoritma model Anda. Mengonfigurasi algoritma model membuatnya tersedia untuk asosiasi ke kolaborasi.

Gambar berikut menunjukkan konfigurasi algoritma model sebagai langkah yang terjadi setelah Anda membuat image pelatihan kontainer dan sebelum Anda mengaitkannya dengan kolaborasi.



Console

Untuk mengkonfigurasi algoritma model ML kustom (konsol)

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih model Custom ML.
3. Pada halaman model Custom ML, pilih Configure model algorithm.
4. Pada halaman Konfigurasi algoritma model, untuk detail algoritma Model, masukkan Nama dan Deskripsi opsional.
5. Jika Anda ingin melakukan pelatihan model, untuk detail wadah ECR gambar Pelatihan,
 - a. Pilih kotak centang Tentukan URI gambar pelatihan.
 - b. Pilih Repositori yang memiliki model pelatihan, wadah inferensi, atau keduanya, dari daftar dropdown.
 - c. Pilih Gambar.
 - d. (Opsional) Masukkan Nilai untuk Entrypoints untuk mengakses gambar pelatihan.
 - e. (Opsional) Masukkan Nilai untuk Argumen.

6. (Opsional) Jika Anda ingin melaporkan metrik model, untuk metrik Pelatihan, masukkan Nama metrik dan pernyataan Regex yang akan mencari log keluaran untuk menemukan metrik.
7. Jika Anda ingin melakukan inferensi model, untuk detail wadah ECR gambar Inferensi,
 - a. Pilih kotak centang Tentukan URI gambar inferensi.
 - b. Pilih Repositori dari daftar dropdown.
 - c. Pilih Gambar.
8. Untuk akses Layanan, pilih nama peran layanan yang ada yang akan digunakan untuk mengakses tabel ini.
9. Untuk Enkripsi, pilih pengaturan Sesuaikan enkripsi untuk menentukan kunci KMS Anda sendiri dan informasi terkait. Jika tidak, Clean Rooms ML akan mengelola enkripsi.
10. Jika Anda ingin mengaktifkan Tag, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.
11. Pilih Konfigurasi algoritma model.

API

Untuk mengkonfigurasi algoritma model HTML kustom (API)

1. Buat image docker yang kompatibel dengan SageMaker AI. Clean Rooms ML hanya mendukung gambar docker yang kompatibel dengan SageMaker AI.
2. Setelah Anda membuat gambar docker yang kompatibel dengan SageMaker AI, gunakan Amazon ECR untuk membuat gambar pelatihan. Ikuti petunjuk di [Amazon Elastic Container Registry User Guide](#) untuk membuat gambar pelatihan kontainer.
3. Konfigurasi algoritma model untuk digunakan di Clean Rooms ML. Anda harus memberikan informasi berikut:
 - Tautan repositori Amazon ECR dan argumen tambahan untuk melatih model dan menjalankan inferensi. Clean Rooms ML mendukung menjalankan pekerjaan transformasi batch pada wadah inferensi.
 - Peran akses layanan yang memungkinkan Clean Rooms untuk mengakses repositori.
 - (Opsional) Wadah inferensi. Meskipun Anda dapat menyediakan ini dalam algoritma model terkonfigurasi terpisah, kami menyarankan Anda menyediakannya dalam langkah ini sehingga wadah pelatihan dan inferensi dikelola sebagai bagian dari sumber daya yang sama.

Jalankan kode berikut dengan parameter spesifik Anda.

```
import boto3
acr_ml_client= boto3.client('cleanroomsm1')

acr_ml_client.create_configured_model_algorithm(
    name='configured_model_algorithm_name',
    trainingContainerConfig={
        'imageUri': 'account.dkr.ecr.region.amazonaws.com/image_name:tag',
        'metricDefinitions': [
            {
                'name': 'custom_metric_name_1',
                'regex': 'custom_metric_regex_1'
            }
        ]
    },
    inferenceContainerConfig={
        'imageUri': 'account.dkr.ecr.region.amazonaws.com/image_name:tag',
    }
    roleArn='arn:aws:iam::account:role/role_name'
)
```

Mengaitkan algoritma model yang dikonfigurasi di AWS Clean Rooms MS

Setelah Anda mengonfigurasi algoritma model, Anda siap untuk mengaitkan algoritma model dengan kolaborasi. Mengaitkan algoritma model membuat algoritma model tersedia untuk semua anggota kolaborasi.

Gambar berikut menunjukkan mengaitkan algoritma model yang dikonfigurasi sebagai langkah terakhir, setelah membuat gambar pelatihan kontainer dan mengonfigurasi algoritma model.

▼ How it works



Create container training image

Create a dockerized container training image to store your model algorithm. [Learn more](#)

[Create container training image](#)



Configure model algorithm

Provide access to the training image and the model algorithm stored within.

[Configure model algorithm](#)



Associate to a collaboration

Add a configured model algorithm to a collaboration and train it on the collaboration data.

[View collaborations](#)

Console

Note

Setelah algoritma model dikaitkan, itu tidak dapat diedit. Untuk membuat perubahan, Anda dapat menghapus algoritma model terkait dan mengaitkan yang baru.

Untuk mengaitkan algoritma model ML kustom (konsol)

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih model Custom ML.
3. Pada halaman model Custom ML, pilih algoritma model yang dikonfigurasi yang ingin Anda kaitkan dengan kolaborasi, lalu pilih Associate to collaboration.
4. Di jendela algoritma model yang dikonfigurasi Associate, pilih Kolaborasi yang ingin Anda kaitkan.
5. Pilih Pilih kolaborasi.
6. Pada halaman algoritma model Associate, untuk detail asosiasi algoritma Model, masukkan Nama dan Deskripsi opsional.
7. Untuk algoritma Model, pilih algoritma model yang dikonfigurasi.
8. Untuk konfigurasi privasi ekspor model Terlatih,
 - a. Untuk mengeksport file model, pilih kotak centang File model.
 - b. Untuk mengeksport file output, pilih kotak centang File output.
 - c. Masukkan nilai ukuran Maks untuk data yang diekspor. Nilai harus antara 0,01 dan 10.

9. (Opsional) Jika Anda ingin mengirim log kesalahan lengkap atau ringkasan kesalahan yang lebih pendek kepada anggota, di bawah konfigurasi privasi pekerjaan inferensi model terlatih,
 - a. Di bawah Log penuh, pilih satu atau beberapa Akun IDs dari daftar dropdown.
 - b. (Opsional) Jika Anda ingin mengirim log yang cocok dengan pola filter, masukkan pola Filter.
 - c. (Opsional) Jika Anda ingin menambahkan akun lain dan pola filter opsional, pilih Tambahkan kebijakan log.
 - d. Di bawah Ringkasan kesalahan, pilih satu atau beberapa Akun IDs dari daftar dropdown.
 - e. (Opsional) Pilih satu atau beberapa Entitas yang akan disunting untuk menentukan entitas mana yang akan disunting dari log kesalahan atau ringkasan kesalahan.
 - PII — menyunting Informasi Identifikasi Pribadi
 - Angka - menyunting angka
 - Custom - redact berdasarkan pola redaksi kustom
 - i. Jika Anda memilih Custom pada langkah sebelumnya, masukkan pola redaksi kustom. Ini mencatat informasi yang cocok dengan pola ini.
 - ii. (Opsional) Jika Anda ingin menambahkan pola redaksi khusus lainnya, pilih Tambahkan pola kustom lainnya.
10. (Opsional) Jika Anda ingin mengonfigurasi metrik model terlatih, di bawah konfigurasi metrik model terlatih, pilih tingkat Kebisingan dari daftar tarik-turun.

Anda dapat memilih Tidak Ada, Rendah, Sedang dan Tinggi.
11. (Opsional) Jika Anda ingin mengatur ukuran artefak maksimum, di bawah Konfigurasi artefak, masukkan nilai ukuran artefak maks. Nilai harus antara 0,01 dan 10.
12. (Opsional) Jika Anda ingin mengaktifkan Tag, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.
13. Pilih Kaitkan.

API

Untuk mengaitkan algoritma model HTML kustom (API)

Jalankan kode berikut dengan parameter spesifik Anda.

Anda juga menyediakan kebijakan privasi yang menentukan siapa yang memiliki akses ke log yang berbeda, memungkinkan pelanggan untuk menentukan regex, dan berapa banyak data yang dapat diekspor dari output model pelatihan atau hasil inferensi.

 Note

Asosiasi algoritma model yang dikonfigurasi tidak dapat diubah.

```
import boto3
acr_ml_client= boto3.client('cleanroomsm1')

acr_ml_client.create_configured_model_algorithm_association(
    name='configured_model_algorithm_association_name',
    description='purpose of the association',
    configuredModelAlgorithmArn='arn:aws:cleanrooms-ml:region:account:membership/
membershipIdentifier/configured-model-algorithm/identifier',
    privacyConfiguration={
        "policies": {
            "trainedModelExports": {
                "filesToExport": ['files to export'],
                "containerLogs": [
                    {
                        "allowedAccountIds": ['member_account_id'],
                        "filterPattern": ['filter pattern'],
                        "logRedactionConfiguration": {
                            "entitiesToRedact": [
                                'ALL_PERSONALLY_IDENTIFIABLE_INFORMATION',
                                'NUMBERS',
                                'CUSTOM'
                            ],
                            "customEntityConfig": {
                                "customDataIdentifiers": [
                                    'custom_regex_1',
                                    'custom_regex_2'
                                ]
                            }
                        }
                    }
                ]
            },
            "containerMetrics": {
                "noiseLevel": 'noise value'
            }
        }
    }
)
```

```

    },
    "maxArtifactSize": {
      "unit": 'unit',
      "value": 'number'
    }
  },
  "trainedModelInferenceJobs": {
    "containerLogs": [
      {
        "allowedAccountIds": ['member_account_id'],
        "filterPattern": ['filter pattern'],
        "logRedactionConfiguration": {
          "entitiesToRedact": [
            'ALL_PERSONALLY_IDENTIFIABLE_INFORMATION',
            'NUMBERS',
            'CUSTOM'
          ],
          "customEntityConfig": {
            "customDataIdentifiers": [
              'custom_regex_1',
              'custom_regex_2'
            ]
          }
        }
      ]
    },
    "maxOutputSize": {
      "unit": 'unit',
      "value": 'number'
    }
  }
},
tags={
  'tag': 'tag'
}
)

```

Setelah algoritma model yang dikonfigurasi dikaitkan dengan kolaborasi, penyedia data pelatihan harus menambahkan aturan analisis kolaborasi ke tabel mereka. Aturan ini memungkinkan asosiasi algoritma model yang dikonfigurasi untuk mengakses tabel yang dikonfigurasi. Semua penyedia data pelatihan yang berkontribusi harus menjalankan kode berikut:

```
import boto3
acr_client= boto3.client('cleanrooms')

acr_client.create_configured_table_association_analysis_rule(
    membershipIdentifier= 'membership_id',
    configuredTableAssociationIdentifier= 'configured_table_association_id',
    analysisRuleType= 'CUSTOM',
    analysisRulePolicy = {
        'v1': {
            'custom': {
                'allowedAdditionalAnalyses': ['arn:aws:cleanrooms-
ml:region:*:membership/*/configured-model-algorithm-association/*'],
                'allowedResultReceivers': []
            }
        }
    }
)
```

Note

Karena asosiasi algoritma model yang dikonfigurasi tidak dapat diubah, kami merekomendasikan agar penyedia data pelatihan yang ingin mengizinkan model untuk digunakan untuk menggunakan kartu liar `allowedAdditionalAnalyses` selama beberapa iterasi pertama konfigurasi model kustom. Hal ini memungkinkan penyedia model untuk mengulangi kode mereka tanpa mengharuskan penyedia pelatihan lain untuk mengasosiasikan kembali sebelum melatih kode model mereka yang diperbarui dengan data.

Membuat saluran input ML di AWS Clean Rooms

Prasyarat:

- Sebuah Akun AWS dengan akses ke AWS Clean Rooms
- Kolaborasi yang disiapkan di AWS Clean Rooms tempat Anda ingin membuat saluran input ML
- Izin untuk menanyakan data dan membuat saluran input ML dalam kolaborasi.
- (Opsional) Algoritma model yang ada untuk diasosiasikan dengan saluran input ML, atau izin untuk membuat yang baru

- (Opsional) Tabel dengan aturan analisis yang dapat dijalankan untuk model yang Anda tentukan.
- (Opsional) Kueri SQL atau templat analisis yang ada untuk digunakan untuk menghasilkan kumpulan data
- (Opsional) Peran layanan yang ada dengan izin yang sesuai, atau izin untuk membuat peran layanan baru
- (Opsional) AWS KMS Kunci khusus jika Anda ingin menggunakan kunci enkripsi Anda sendiri
- Izin yang sesuai untuk membuat dan mengelola model ML dalam kolaborasi

Saluran input ML adalah kumpulan data yang dibuat dari kueri data tertentu. Anggota dengan kemampuan untuk query data dapat mempersiapkan data mereka untuk pelatihan dan inferensi dengan membuat saluran input ML. Membuat saluran input ML memungkinkan data tersebut digunakan dalam model pelatihan yang berbeda dalam kolaborasi yang sama. Anda harus membuat saluran input ML terpisah untuk pelatihan dan inferensi.

Untuk membuat saluran input ML, Anda harus menentukan query SQL yang digunakan untuk query data input dan membuat saluran input ML. Hasil kueri ini tidak pernah dibagikan dengan anggota mana pun dan tetap berada dalam batas-batas Clean Rooms. Referensi Amazon Resource Name (ARN) digunakan pada langkah selanjutnya untuk melatih model atau menjalankan inferensi.

Console

Untuk membuat saluran input ML (konsol)

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pada halaman Kolaborasi, pilih kolaborasi tempat Anda ingin membuat saluran input ML.
4. Setelah kolaborasi terbuka, pilih tab model ML.
5. Di bawah model Custom ML, di bagian saluran input ML, pilih Buat saluran input ML.
6. Pada halaman Create channel input ML, untuk detail saluran masukan ML, lakukan hal berikut:
 - a. Untuk Nama, masukkan nama unik untuk saluran Anda.
 - b. (Opsional) Untuk Deskripsi, masukkan deskripsi saluran Anda.
 - c. Untuk algoritma model terkait, pilih algoritma yang akan digunakan.

Pilih algoritma model Associate untuk menambahkan yang baru.

7. Untuk Dataset, pilih metode untuk menghasilkan kumpulan data pelatihan:

- Pilih kueri SQL untuk menggunakan hasil kueri SQL sebagai kumpulan data pelatihan.

Jika Anda memilih kueri SQL, masukkan kueri Anda di bidang kueri SQL.

(Opsional) Untuk mengimpor kueri yang baru saja Anda gunakan, pilih Impor dari kueri terbaru.

- Pilih template Analisis untuk menggunakan hasil template analisis sebagai kumpulan data pelatihan.

 Warning

Pembuatan data sintetis melindungi dari menyimpulkan atribut individu apakah individu tertentu hadir dalam kumpulan data asli atau atribut pembelajaran individu tersebut hadir. Namun, itu tidak mencegah nilai literal dari kumpulan data asli, termasuk informasi yang dapat diidentifikasi secara pribadi (PII) muncul di kumpulan data sintetis.

Sebaiknya hindari nilai dalam kumpulan data input yang hanya terkait dengan satu subjek data karena ini dapat mengidentifikasi kembali subjek data. Misalnya, jika hanya satu pengguna yang tinggal dalam kode pos, keberadaan kode pos itu dalam kumpulan data sintetis akan mengonfirmasi bahwa pengguna berada dalam kumpulan data asli. Teknik seperti memotong nilai presisi tinggi atau mengganti katalog yang tidak biasa dengan yang lain dapat digunakan untuk mengurangi risiko ini. Transformasi ini dapat menjadi bagian dari query yang digunakan untuk membuat saluran input ML.

- a. Jika tidak ada tabel yang terkait, pilih Tabel asosiasi untuk menambahkan tabel dengan aturan analisis yang dapat dijalankan untuk model yang ditentukan.
- b. Pilih Jenis pekerja yang akan digunakan saat membuat saluran data ini. Jenis pekerja default adalah CR.1X. Tentukan Jumlah pekerja yang akan digunakan. Nomor pekerja default adalah 16. Untuk menentukan properti Spark:
 - i. Perluas properti Spark.
 - ii. Pilih Tambahkan properti Spark.

- iii. Pada kotak dialog properti Spark, pilih nama Properti dari daftar dropdown dan masukkan Nilai.

Tabel berikut memberikan definisi untuk setiap properti.

Untuk informasi selengkapnya tentang properti Spark, lihat [Spark Properties](#) di dokumentasi Apache Spark.

Nama Properti	Deskripsi	nilai default
Spark.Task.MaxFailures	Mengontrol berapa kali berturut-turut tugas dapat gagal sebelum pekerjaan gagal. Membutuhkan nilai yang lebih besar dari atau sama dengan 1. Jumlah percobaan ulang yang diizinkan sama dengan nilai ini dikurangi 1. Hitungan kegagalan diatur ulang jika ada upaya yang berhasil. Kegagalan di berbagai tugas tidak menumpuk menuju batas ini.	4
spark.sql.file.maxPartitionBytes	Menetapkan jumlah maksimum byte untuk dikemas ke dalam satu partisi ketika membaca dari sumber berbasis file seperti Parquet, JSON, dan ORC.	128MB
Spark.hadoop.fs.s3.maxRetries	Menetapkan jumlah maksimum percobaan ulang untuk operasi file Amazon S3.	

Nama Properti	Deskripsi	nilai default
spark.network.timeout	Menetapkan batas waktu default untuk semua interaksi jaringan. Mengganti pengaturan batas waktu berikut jika tidak dikonfigurasi: - spark.storage.blockManagerHeartbeatTimeoutMs - spark.shuffle.io.connectionTimeout - Spark.rpc.askTimeout - Spark.rpc.lookupTimeout	120s
spark.rdd.compress	Menentukan apakah untuk mengompres partisi RDD serial menggunakan spark.io.compression.codec. Berlaku StorageLevel untuk MEMORY_ONLY_SER di Java dan Scala, atau MEMORY_ONLY dengan Python. StorageLevel Mengurangi ruang penyimpanan tetapi membutuhkan waktu pemrosesan CPU tambahan.	SALAH
spark.shuffle.spill.compress	Menentukan apakah untuk mengompres data tumpahan shuffle menggunakan spark.io.compression.codec.	BETUL
spark.sql.adaptive.advisoryPartitionSizeInBytes	Menetapkan ukuran target dalam byte untuk partisi acak selama pengoptimalan adaptif ketika spark.sql.adaptive.enabled benar. Mengontrol ukuran partisi saat menggabungkan partisi kecil atau memisahkan partisi miring.	(nilai spark.sql.adaptive.shuffle.targetPostShuffleInputSize)

Nama Properti	Deskripsi	nilai default
spark.sql.adaptif. autoBroad castJoinAmbang	Menetapkan ukuran tabel maksimum dalam byte untuk penyiaran ke node pekerja selama bergabung. Berlaku hanya dalam kerangka adaptif. Menggunakan nilai default yang sama seperti spark.sql. autoBroad castJoinAmbang batas. Setel ke -1 untuk menonaktifkan penyiaran.	(tidak ada)
spark.sql.adaptive .coalescepartition s.enabled	Menentukan apakah akan menggabungkan partisi shuffle bersebelahan berdasarkan spark.sql.adaptive.advisoryPartitionSizeInBytes untuk mengoptimalkan ukuran tugas. Membutuhkan spark.sql.adaptive.enabled agar benar.	BETUL
Spark.sql.adaptive .CoalescePartition s. initialPartitionNum	Mendefinisikan jumlah awal partisi shuffle sebelum penggabungan. Memerlukan spark.sql.adaptive.enabled dan spark.sql.adaptive.CoalescePartitions.enabled untuk menjadi true. Default ke nilai spark.sql.shuffle.partitions.	(tidak ada)
Spark.sql.adaptive .CoalescePartition s. minPartitionSize	Menetapkan ukuran minimum untuk partisi shuffle yang digabungkan untuk mencegah partisi menjadi terlalu kecil selama optimasi adaptif.	1 MB

Nama Properti	Deskripsi	nilai default
spark.sql.adaptive.coalescepartitions.parallelismFirst	Menentukan apakah untuk menghitung ukuran partisi berdasarkan paralelisme cluster bukan spark.sql.adaptive.advisoryPartitionSizeInBytes selama penggabungan partisi. Menghasilkan ukuran partisi yang lebih kecil dari ukuran target yang dikonfigurasi untuk memaksimalkan paralelisme. Kami merekomendasikan pengaturan ini ke false pada cluster sibuk untuk meningkatkan pemanfaatan sumber daya dengan mencegah tugas-tugas kecil yang berlebihan.	BETUL
spark.sql.adaptive.enabled	Menentukan apakah akan mengaktifkan eksekusi kueri adaptif untuk mengoptimalkan kembali rencana kueri selama eksekusi kueri, berdasarkan statistik runtime yang akurat.	BETUL
spark.sql.adaptive.forceOptimizeSkewedBergabung	Menentukan apakah akan memaksa mengaktifkan OptimizeSkewedJoin bahkan jika memperkenalkan shuffle tambahan.	SALAH
spark.sql.adaptive.localShuffleReader.diaktifkan	Menentukan apakah akan menggunakan pembaca shuffle lokal ketika partisi shuffle tidak diperlukan, seperti setelah mengonversi dari gabungan sort-merge ke broadcast-hash bergabung. Membutuhkan spark.sql.adaptive.enabled agar benar.	BETUL

Nama Properti	Deskripsi	nilai default
spark.sql.adaptif. maxShuffleHashJoinLocalMapThreshold	Menetapkan ukuran partisi maksimum dalam byte untuk membangun peta hash lokal. Memprioritaskan gabungan hash yang diacak daripada gabungan sort-merge saat: • Nilai ini sama atau melebihi spark.sql.adaptive.advisoryPartitionSizeInBytes • Semua ukuran partisi berada dalam batas ini Mengganti spark.sql.join.preferSortMergeBergabunglah dengan pengaturan.	0 byte
spark.sql.adaptif.optimizeShuffleRebalancePartitions.d iaktifkan	Menentukan apakah akan mengoptimalkan partisi shuffle miring dengan membaginya menjadi partisi yang lebih kecil berdasarkan spark.sql.adaptive.advisoryPartitionSizeInBytes. Membutuhkan spark.sql.adaptive.enabled agar benar.	BETUL
spark.sql.adaptif.rebalancePartitionsSmallPartitionFactor	Mendefinisikan faktor ambang ukuran untuk menggabungkan partisi selama pemisahan. Partisi yang lebih kecil dari faktor ini dikalikan dengan spark.sql.adaptive.advisoryPartitionSizeInBytes digabung.	0.2

Nama Properti	Deskripsi	nilai default
Spark.sql.adaptive.skewjoin.enabled	Menentukan apakah untuk menangani data miring dalam bergabung diacak dengan memisahkan dan opsional mereplikasi partisi miring. Berlaku untuk gabungan hash sort-merge dan shuffled. Membutuhkan spark.sql.adaptive.enabled agar benar.	BETUL
Spark.sql.adaptive.skewJoin.skewedPartitionFactor	Menentukan faktor ukuran yang menentukan kemiringan partisi. Partisi miring ketika ukurannya melebihi keduanya: - Faktor ini dikalikan dengan ukuran partisi median - Nilai spark.sql.adaptive.skewJoin.skewedPartitionThresholdInBytes	5
Spark.sql.adaptive.skewJoin.skewedPartitionThresholdInBytes	Menetapkan ambang ukuran dalam byte untuk mengidentifikasi partisi miring. Partisi miring ketika ukurannya melebihi keduanya: - Ambang batas ini - Ukuran partisi median dikalikan dengan spark.sql.adaptive.skewJoin.skewedPartitionFactor Kami merekomendasikan pengaturan nilai ini lebih besar dari spark.sql.adaptive.advisoryPartitionSizeInBytes.	256MB

Nama Properti	Deskripsi	nilai default
spark.sql. autoBroadcastJoinAmbang	Menetapkan ukuran tabel maksimum dalam byte untuk penyiaran ke node pekerja selama bergabung. Setel ke -1 untuk menonaktifkan penyiaran.	10MB
Spark.Sql. .BroadcastTimeout	Mengontrol periode batas waktu dalam hitungan detik untuk operasi siaran selama siaran bergabung.	300 detik
spark.sql.cbo.enabled	Menentukan apakah akan mengaktifkan optimasi berbasis biaya (CBO) untuk estimasi statistik rencana.	SALAH
Spark.sql.cbo.join reorder.dp.star.filter	Menentukan apakah akan menerapkan heuristik filter star-join selama enumerasi bergabung berbasis biaya.	SALAH
spark.sql.cbo.join reorder.dp.threshold	Menetapkan jumlah maksimum node bergabung yang diizinkan dalam algoritma pemrograman dinamis.	12
Spark.sql.cbo.join reOrder.enabled	Menentukan apakah akan mengaktifkan pengurutan ulang bergabung dalam optimasi berbasis biaya (CBO).	SALAH
Spark.sql.cbo.plan stats.enabled	Menentukan apakah untuk mengambil jumlah baris dan statistik kolom dari katalog selama pembuatan rencana logis.	SALAH
spark.sql.cbo. starSchemaDetection	Menentukan apakah akan mengaktifkan pengurutan ulang gabungan berdasarkan deteksi skema bintang.	SALAH

Nama Properti	Deskripsi	nilai default
spark.sql.file. maxPartitionNum	Menetapkan jumlah maksimum target partisi file split untuk sumber berbasis file (Parket, JSON, dan ORC). Rescales partisi ketika hitungan awal melebihi nilai ini. Ini adalah target yang disarankan, bukan batas yang dijamin.	(tidak ada)
spark.sql.file. maxRecord sPerBerkas	Menetapkan jumlah maksimum catatan untuk menulis ke satu file. Tidak ada batasan yang berlaku saat disetel ke nol atau nilai negatif.	0
spark.sql.file. minPartitionNum	Menetapkan jumlah minimum target partisi file terpisah untuk sumber berbasis file (Parket, JSON, dan ORC). Default ke spark.sql. leafNodeDefaultParalelisme. Ini adalah target yang disarankan, bukan batas yang dijamin.	(tidak ada)
spark.sql. inMemoryC olumnarSt orage.batchsize	Mengontrol ukuran batch untuk caching kolumnar. Meningkatkan ukuran meningkatkan pemanfaatan memori dan kompresi tetapi meningkatkan risiko out-of-memory kesalahan.	10000
spark.sql. inMemoryC olumnarpe nyimpanan .dikompresi	Menentukan apakah akan secara otomatis memilih codec kompresi untuk kolom berdasarkan statistik data.	BETUL

Nama Properti	Deskripsi	nilai default
spark.sql.inMemoryColumnarPe nyimpanan . enableVec torizedReader	Menentukan apakah untuk mengaktifkan pembacaan vektor untuk caching kolumnar.	BETUL
spark.sql.legacy. allowHash OnMapType	Menentukan apakah untuk memungkinkan operasi hash pada struktur data tipe peta. Pengaturan lama ini mempertahankan kompatibilitas dengan penanganan jenis peta versi Spark yang lebih lama.	
spark.sql.legacy. allowNeg ativeScale OfDecimal	Menentukan apakah untuk memungkinkan nilai skala negatif dalam definisi tipe desimal. Setelan lama ini mempertahankan kompatibilitas dengan versi Spark lama yang mendukung skala desimal negatif.	
spark.sql.legacy. castCompl exTypesTo String.diaktifkan	Menentukan apakah akan mengaktifkan perilaku warisan untuk casting tipe kompleks untuk string. Menjaga kompatibilitas dengan aturan konversi tipe versi Spark yang lebih lama.	
spark.sql.legacy. charVarcharAsTali	Menentukan apakah untuk memperlakukan jenis CHAR dan VARCHAR sebagai tipe STRING. Pengaturan lama ini menyediakan kompatibilitas dengan penanganan tipe string versi Spark yang lebih lama.	

Nama Properti	Deskripsi	nilai default
<code>spark.sql.legacy.createEmptyCollectionUsingStringType</code>	Menentukan apakah untuk membuat koleksi kosong menggunakan elemen tipe string. Setelan lama ini mempertahankan kompatibilitas dengan perilaku inisialisasi koleksi versi Spark yang lebih lama.	
<code>spark.sql.legacy.exponentLiteralAsDesimal</code> . Diaktifkan	Menentukan apakah untuk menafsirkan literal eksponensial sebagai tipe desimal. Pengaturan lama ini mempertahankan kompatibilitas dengan penanganan literal numerik versi Spark yang lebih lama.	
<code>spark.sql.legacy.json.allowEmptyString</code> .diaktifkan	Menentukan apakah untuk mengizinkan string kosong dalam pengolahan JSON. Pengaturan lama ini mempertahankan kompatibilitas dengan perilaku penguraian JSON versi Spark yang lebih lama.	
<code>spark.sql.legacy.parquet.int96RebaseModelRead</code>	Menentukan apakah akan menggunakan mode rebase INT96 stempel waktu lama saat membaca file Parquet. Pengaturan lama ini mempertahankan kompatibilitas dengan penanganan stempel waktu versi Spark yang lebih lama.	
<code>spark.sql.legacy.timeParserPolicy</code>	Mengontrol perilaku penguraian waktu untuk kompatibilitas mundur. Pengaturan lama ini menentukan bagaimana stempel waktu dan tanggal diuraikan dari string.	

Nama Properti	Deskripsi	nilai default
Spark.sql.legacy.typeCoercion.datetimeToString.diaaktifkan	Menentukan apakah akan mengaktifkan perilaku pemaksaan tipe warisan saat mengonversi nilai datetime ke string. Menjaga kompatibilitas dengan aturan konversi datetime versi Spark yang lebih lama.	
spark.sql.maxSinglePartitionByte	Menetapkan ukuran partisi maksimum dalam byte. Perencana memperkenalkan operasi shuffle untuk partisi yang lebih besar untuk meningkatkan paralelisme.	128m
Spark.sql.metadatacacheTTLSeconds	Mengontrol time-to-live (TTL) untuk cache metadata. Berlaku untuk metadata file partisi dan cache katalog sesi. Membutuhkan: • Nilai positif lebih besar dari nol • Spark.sql.CatalogImplementation diatur ke sarang • spark.sql.hive.filesourcePartitionFileCacheSize lebih besar dari nol • spark.sql.hive.manageFilesourcePartitions disetel ke true	-1000ms
spark.sql.optimizer.collapseProjectionAlwaysSebaris	Menentukan apakah untuk menciutkan proyeksi yang berdekatan dan ekspresi inline, bahkan ketika itu menyebabkan duplikasi.	SALAH
spark.sql.optimizer.dynamicPartitionPruning.diaaktifkan	Menentukan apakah untuk menghasilkan predikat untuk kolom partisi yang digunakan sebagai kunci bergabung.	BETUL

Nama Properti	Deskripsi	nilai default
spark.sql.optimize r.enableCsv Expressio nOptimasi	Menentukan apakah akan mengoptimalkan ekspresi CSV dalam pengoptimal SQL dengan memangkas kolom yang tidak perlu dari operasi from_csv.	BETUL
spark.sql.optimize r.enableJso nExpressi onOptimasi	Menentukan apakah untuk mengoptimalkan ekspresi JSON dalam pengoptimal SQL dengan: • Memangkas kolom yang tidak perlu dari operasi from_json • Menyederhanakan kombinasi from_json dan to_json • Mengoptimalkan operasi named_struct	BETUL
Spark.sql.optimize r.excludedrules	Mendefinisikan aturan pengoptimal untuk dinonaktifkan, diidentifikasi dengan nama aturan yang dipisahkan koma. Beberapa aturan tidak dapat dinonaktifkan karena diperlukan untuk kebenaran. Pengoptimal mencatat aturan mana yang berhasil dinonaktifkan.	(tidak ada)
Spark.sql.Optimize r.Runtime.BloomFil ter.applicati onSideSca nSizeThreshold	Menetapkan ukuran pemindaian agregat minimum dalam byte yang diperlukan untuk menyuntikkan filter Bloom di sisi aplikasi.	10GB
Spark.sql.Optimize r.Runtime.BloomFil ter.creationS ideThreshold	Mendefinisikan ambang batas ukuran maksimum untuk menyuntikkan filter Bloom di sisi pembuatan.	10MB

Nama Properti	Deskripsi	nilai default
Spark.sql.Optimize r.Runtime.BloomFilter.enabled	Menentukan apakah akan menyisipkan filter Bloom untuk mengurangi data shuffle ketika satu sisi gabungan shuffle memiliki predikat selektif.	BETUL
Spark.sql.Optimize r.Runtime.BloomFilter.expectedNumItems	Mendefinisikan jumlah default item yang diharapkan dalam filter Bloom runtime.	1000000
Spark.sql.Optimize r.Runtime.BloomFilter.maxNumBits	Menetapkan jumlah maksimum bit yang diizinkan dalam filter Bloom runtime.	67108864
Spark.sql.Optimize r.Runtime.BloomFilter.maxNumItems	Menetapkan jumlah maksimum item yang diharapkan yang diizinkan dalam filter Bloom runtime.	4000000
spark.sql.optimize r.runtime.bloomfilter.number.threshold	Membatasi jumlah maksimum filter runtime non-DPP yang diizinkan per kueri untuk mencegah out-of-memory kesalahan pada driver.	10
Spark.sql.Optimize r.Runtime.BloomFilter.Numbits	Mendefinisikan jumlah default bit yang digunakan dalam filter Bloom runtime.	8388608

Nama Properti	Deskripsi	nilai default
spark.sql .optimizer.runtime .rowlevelO perationGroupFilde r.enabled	Menentukan apakah akan mengaktifkan penyaringan grup runtime untuk operasi tingkat baris. Memungkinkan sumber data untuk: - Pangkas seluruh kelompok data (seperti file atau partisi) menggunakan filter sumber data - Jalankan kueri runtime untuk mengidentifikasi catatan yang cocok - Buang grup yang tidak perlu untuk menghindari penulisan ulang yang mahal Pembatasan: - Tidak semua ekspresi dapat dikonversi ke filter sumber data - Beberapa ekspresi memerlukan evaluasi Spark (seperti subquery)	BETUL
Spark.Sql.Optimize r.RuntimeFilter. semiJoinR eduction.diaktifkan	Menentukan apakah akan menyisipkan semi-join untuk mengurangi data shuffle ketika salah satu sisi gabungan shuffle memiliki predikat selektif.	SALAH

Nama Properti	Deskripsi	nilai default
Spark.sql.parquet. Aggregate Pushdown	Menentukan apakah akan menekan agregat ke Parquet untuk optimasi. Mendukung: • MIN dan MAX untuk jenis boolean, integer, float, dan tanggal • COUNT untuk semua tipe data Melempar pengecualian jika statistik hilang dari footer file Parquet apa pun.	SALAH
spark.sql.parquet. columnarReaderBatchUkuran	Mengontrol jumlah baris di setiap batch pembaca vektor Parquet. Pilih nilai yang menyeimbangkan overhead kinerja dan penggunaan memori untuk mencegah out-of-memory kesalahan.	4096
Spark.sql.Session. Timezone	Mendefinisikan zona waktu sesi untuk menangani stempel waktu dalam literal string dan konversi objek Java. Menerima: • Berbasis wilayah IDs dalam area/city format (seperti Amerika/Los_Angeles) • Offset zona dalam format (+/-) HH, (+/-) HH: mm, atau (+/-) HH: mm: SS (seperti -08 atau + 01:00) • UTC atau Z sebagai alias untuk + 00:00	(nilai zona waktu lokal)
spark.sql.shuffle. partisi	Menetapkan jumlah default partisi untuk pengocokan data selama bergabung atau agregasi. Tidak dapat diubah antara restart kueri streaming terstruktur dari lokasi pos pemeriksaan yang sama.	200

Nama Properti	Deskripsi	nilai default
spark.sql . shuffledH ashJoinFaktor	Mendefinisikan faktor perkalian yang digunakan untuk menentukan kelayakan bergabung hash shuffle. Gabungan hash shuffle dipilih ketika ukuran data sisi kecil dikalikan dengan faktor ini kurang dari ukuran data sisi besar.	3
spark.sql.sources. parallelPartitionD iscovery.ambang	Menetapkan jumlah maksimum jalur untuk daftar file sisi driver dengan sumber berbasis file (Parquet, JSON, dan ORC). Ketika terlampaui selama penemuan partisi, file terdaftar menggunakan pekerjaan terdistribusi Spark terpisah.	32
spark.sql.statisti cs.histogram.enabl ed	Menentukan apakah akan menghasilkan histogram equi-height selama perhitungan statistik kolom untuk meningkatkan akurasi estimasi. Membutuhkan pemindaian tabel tambahan di luar yang dibutuhkan untuk statistik kolom dasar.	SALAH

- c. Untuk penyimpanan data dalam beberapa hari, masukkan jumlah hari untuk menyimpan data.
- d. Untuk format Hasil, pilih CSV atau Parquet sebagai format data yang harus digunakan saluran input ML.
8. Untuk akses Layanan, pilih nama peran layanan yang ada yang akan digunakan untuk mengakses tabel ini atau pilih Buat dan gunakan peran layanan baru.
9. Untuk Enkripsi, pilih Enkripsi rahasia dengan kunci KMS khusus untuk menentukan kunci KMS Anda sendiri dan informasi terkait. Jika tidak, Clean Rooms ML akan mengelola enkripsi.
10. Pilih Buat saluran masukan ML.

Ini akan memakan waktu beberapa menit untuk membuat saluran input ML. Anda dapat melihat daftar saluran input ML pada tab model ML.

Note

Setelah saluran input ML dibuat, Anda tidak dapat mengeditnya.

API

Untuk membuat saluran input ML (API)

Jalankan kode berikut dengan parameter spesifik Anda:

```
import boto3
acr_client = boto3.client('cleanroomsm1')

acr_client.create_ml_input_channel(
    name="ml_input_channel_name",
    membershipIdentifier='membership_id',

    configuredModelAlgorithmAssociations=[configured_model_algorithm_association_arn],
    retentionInDays=1,
    inputChannel={
        "dataSource": {
            "protectedQueryInputParameters": {
                "sqlParameters": {
                    "queryString": "select * from table",
                    "computeConfiguration": {
                        "worker": {
                            "type": "CR.1X",
                            "number": 16,
                            "properties": {
                                "spark": {
                                    "spark configuration key": "spark configuration
value",
                                }
                            }
                        }
                    }
                },
                "resultFormat": "PARQUET"
            }
        },
        "roleArn": "arn:aws:iam::111122223333:role/role_name"
    }
}
```



```
)  
channel_arn = resp['ML Input Channel ARN']
```

Membuat model terlatih di AWS Clean Rooms ML

Prasyarat:

- Sebuah Akun AWS dengan akses ke AWS Clean Rooms
- Kolaborasi yang dibuat di AWS Clean Rooms
- Algoritma model yang dikonfigurasi terkait dengan kolaborasi
- Setidaknya satu saluran input ML yang dikonfigurasi
- Izin yang sesuai untuk membuat dan mengelola model ML dalam kolaborasi

Setelah Anda mengaitkan algoritma model yang dikonfigurasi ke kolaborasi, lalu membuat dan mengonfigurasi saluran input ML, Anda siap untuk membuat model terlatih. Model terlatih digunakan oleh anggota kolaborasi untuk bersama-sama menganalisis data mereka.

Anda dapat membuat model terlatih menggunakan prosedur berikut.

Atau, Anda dapat menggunakan pelatihan tambahan untuk meningkatkan model yang ada dengan data baru, atau pelatihan terdistribusi untuk melatih model di beberapa instance komputasi.

Topik

- [Menggunakan pelatihan tambahan di AWS Clean Rooms](#)
- [Menggunakan pelatihan terdistribusi di AWS Clean Rooms](#)

Console

Untuk membuat model terlatih (konsol)

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pada halaman Kolaborasi, pilih kolaborasi tempat Anda ingin membuat model terlatih.
4. Setelah kolaborasi terbuka, pilih tab model ML.

5. Di bawah model Custom MS, di bagian Model terlatih, pilih Buat model terlatih.
6. Pada halaman Buat model terlatih, untuk algoritma model terkait, tentukan algoritmanya.
7. Untuk detail model Terlatih, masukkan yang berikut ini:
 - a. Untuk Nama, masukkan nama unik untuk model dalam kolaborasi.
 - b. (Opsional) Untuk Deskripsi, masukkan deskripsi model yang dilatih.
 - c. Untuk mode input data Pelatihan, pilih salah satu dari berikut ini:
 - Pilih File jika Anda memiliki kumpulan data yang lebih kecil yang dapat muat pada volume penyimpanan ML dan Anda lebih memilih akses sistem file tradisional untuk skrip pelatihan Anda.
 - Pilih Pipe untuk kumpulan data besar untuk mengalirkan data langsung dari S3, menghindari kebutuhan untuk mengunduh semuanya ke disk, yang dapat meningkatkan kecepatan pelatihan dan mengurangi persyaratan penyimpanan.
 - Pilih FastFile apakah Anda ingin menggabungkan manfaat streaming dari S3 dengan akses sistem file, terutama untuk membaca data secara berurutan atau ketika berhadapan dengan lebih sedikit file untuk waktu startup yang lebih cepat.
8. Untuk detail saluran masukan ML, lakukan hal berikut:
 - a. Untuk saluran input ML, tentukan saluran input ML yang menyediakan data ke algoritma model.

Untuk menambahkan saluran lain, pilih Tambahkan saluran input ML lainnya. Anda dapat menambahkan hingga 19 saluran input ML tambahan.
 - b. Untuk nama Saluran, masukkan nama saluran input ML.
 - c. Untuk jenis distribusi data Amazon S3, pilih salah satu dari berikut ini:
 - Pilih Sepenuhnya direplikasi untuk memberikan setiap instance pelatihan dengan salinan lengkap kumpulan data Anda. Ini berfungsi paling baik ketika kumpulan data Anda cukup kecil untuk muat dalam memori atau ketika setiap instance membutuhkan akses ke semua data.
 - Pilih Sharded by S3 kunci untuk membagi dataset Anda di seluruh instans pelatihan berdasarkan kunci S3. Setiap instance menerima sekitar $1/n$ dari total objek S3, di mana 'n' adalah jumlah instance. Ini berfungsi paling baik untuk kumpulan data besar yang ingin Anda proses secara paralel.

 Note

Pertimbangkan ukuran set data dan persyaratan pelatihan Anda saat memilih jenis distribusi. Sepenuhnya direplikasi menyediakan akses data yang lengkap tetapi membutuhkan lebih banyak penyimpanan, sementara Sharded by S3 key memungkinkan pemrosesan terdistribusi kumpulan data besar.

9. Untuk Durasi pelatihan maksimum, pilih jumlah waktu maksimum yang Anda inginkan untuk melatih model Anda.
10. Untuk Hyperparameters, tentukan parameter spesifik algoritme apa pun dan nilai yang dimaksudkan. Hyperparameter khusus untuk model yang dilatih dan digunakan untuk menyempurnakan pelatihan model.
11. Untuk variabel Lingkungan, tentukan variabel spesifik algoritme apa pun dan nilai yang dimaksudkan. Variabel lingkungan diatur dalam wadah Docker.
12. Untuk Enkripsi, untuk menggunakan kustom AWS KMS key, pilih kotak centang Enkripsi rahasia dengan kunci KMS kustom.
13. Untuk konfigurasi Sumber Daya EC2, tentukan informasi tentang sumber daya komputasi yang digunakan untuk pelatihan model.
 - a. Untuk jenis Instance, pilih jenis instance yang ingin Anda jalankan.
 - b. Untuk hitungan Instance, masukkan jumlah instance.
 - c. Untuk ukuran Volume dalam GB, masukkan ukuran volume penyimpanan ML.
14. Pilih Buat model terlatih.

API

Untuk membuat model terlatih (API)

Anggota dengan kemampuan untuk melatih model memulai pelatihan dengan memilih saluran input ML dan algoritma model.

Jalankan kode berikut dengan parameter spesifik Anda:

```
import boto3
acr_ml_client= boto3.client('cleanroomsm1')
```

```
acr_ml_client.create_trained_model(  
    membershipIdentifier= 'membership_id',  
    configuredModelAlgorithmAssociationArn = 'arn:aws:cleanrooms-  
ml:region:account:membership/membershipIdentifier/configured-model-algorithm-  
association/identifier',  
    name='trained_model_name',  
    resourceConfig={  
        'instanceType': "ml.m5.xlarge",  
        'volumeSizeInGB': 1  
    },  
    dataChannels=[  
        {  
            "mlInputChannelArn": channel_arn_1,  
            "channelName": "channel_name"  
        },  
        {  
            "mlInputChannelArn": channel_arn_2,  
            "channelName": "channel_name"  
        }  
    ]  
)
```

Note

Setelah model terlatih dibuat, Anda tidak dapat mengeditnya. Untuk membuat perubahan, hapus model terlatih dan buat yang baru.

Menggunakan pelatihan tambahan di AWS Clean Rooms

Prasyarat:

- Sebuah Akun AWS dengan akses ke AWS Clean Rooms
- Model terlatih yang ada dalam kolaborasi
- Dataset baru atau yang diperbarui untuk pelatihan tambahan
- Izin yang sesuai untuk membuat dan mengelola model ML dalam kolaborasi
- Keakraban dengan hiperparameter dan konfigurasi model yang ada

Dengan pelatihan tambahan, Anda dapat menggunakan artefak model yang ada dan kumpulan data yang diperbarui untuk melatih model baru. Pelatihan tambahan menghemat waktu dan sumber daya.

Gunakan pelatihan tambahan untuk:

- Latih model baru menggunakan kumpulan data yang diperluas yang memiliki pola dasar yang tidak diperhitungkan dalam pelatihan sebelumnya.
- Latih beberapa varian model, baik dengan hiperparameter yang berbeda atau menggunakan kumpulan data yang berbeda.

Console

Untuk menjalankan pekerjaan pelatihan tambahan (konsol)

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pada halaman Kolaborasi, pilih kolaborasi di mana artefak model yang ingin Anda gunakan untuk pelatihan tambahan ada.
4. Setelah kolaborasi terbuka, pilih tab model ML.
5. Di bawah model Custom MS, di bagian Model terlatih, pilih tombol radio di sebelah model terlatih yang ingin Anda latih secara bertahap.
6. Pada halaman Ikhtisar, di bawah Versi,
 - a. Pilih tombol radio di sebelah model terlatih yang ingin Anda latih secara bertahap.
 - b. Pilih Kereta dari versi.
7. Pada halaman Buat model terlatih dari versi, untuk versi model Terlatih, pilih versinya.

Versi model dasar dipilih secara otomatis. Anda dapat mengubah versi ini jika ada versi lain.

8. Untuk detail model Terlatih, masukkan yang berikut ini:
 - a. Untuk Nama, masukkan nama unik untuk model dalam kolaborasi.
 - b. (Opsional) Untuk Deskripsi, masukkan deskripsi model yang dilatih.
 - c. Untuk mode input data Pelatihan, pilih salah satu dari berikut ini:

- Pilih File jika Anda memiliki kumpulan data yang lebih kecil yang dapat muat pada volume penyimpanan ML dan Anda lebih memilih akses sistem file tradisional untuk skrip pelatihan Anda.
 - Pilih Pipe untuk kumpulan data besar untuk mengalirkan data langsung dari S3, menghindari kebutuhan untuk mengunduh semuanya ke disk, yang dapat meningkatkan kecepatan pelatihan dan mengurangi persyaratan penyimpanan.
 - Pilih FastFile apakah Anda ingin menggabungkan manfaat streaming dari S3 dengan akses sistem file, terutama untuk membaca data secara berurutan atau ketika berhadapan dengan lebih sedikit file untuk waktu startup yang lebih cepat.
- d. Untuk nama saluran pelatihan tambahan, masukkan nama untuk saluran pelatihan tambahan

 Note

Jika Anda menentukan nama saluran pelatihan tambahan tanpa ID versi, sistem akan menggunakan model dasar untuk pelatihan tambahan.

9. Untuk detail saluran masukan ML, lakukan hal berikut:
- a. Untuk saluran input ML, tentukan saluran input ML yang menyediakan data ke algoritma model.

Untuk menambahkan saluran lain, pilih Tambahkan saluran input ML lainnya. Anda dapat menambahkan hingga 19 saluran input ML tambahan.
 - b. Untuk nama Saluran, masukkan nama saluran input ML.
 - c. Untuk jenis distribusi data Amazon S3, pilih salah satu dari berikut ini:
 - Pilih Sepenuhnya direplikasi untuk memberikan setiap instance pelatihan dengan salinan lengkap kumpulan data Anda. Ini berfungsi paling baik ketika kumpulan data Anda cukup kecil untuk muat dalam memori atau ketika setiap instance membutuhkan akses ke semua data.
 - Pilih Sharded by S3 kunci untuk membagi dataset Anda di seluruh instans pelatihan berdasarkan kunci S3. Setiap instance menerima sekitar $1/n$ dari total objek S3, di mana 'n' adalah jumlah instance. Ini berfungsi paling baik untuk kumpulan data besar yang ingin Anda proses secara paralel.

 Note

Pertimbangkan ukuran set data dan persyaratan pelatihan Anda saat memilih jenis distribusi. Sepenuhnya direplikasi menyediakan akses data yang lengkap tetapi membutuhkan lebih banyak penyimpanan, sementara Sharded by S3 key memungkinkan pemrosesan terdistribusi kumpulan data besar.

10. Untuk Durasi pelatihan maksimum, pilih jumlah waktu maksimum yang Anda inginkan untuk melatih model Anda.
11. Untuk Hyperparameters, tentukan parameter spesifik algoritme apa pun dan nilai yang dimaksudkan. Hyperparameter khusus untuk model yang dilatih dan digunakan untuk menyempurnakan pelatihan model.
12. Untuk variabel Lingkungan, tentukan variabel spesifik algoritme apa pun dan nilai yang dimaksudkan. Variabel lingkungan diatur dalam wadah Docker.
13. Untuk Enkripsi, untuk menggunakan kustom AWS KMS key, pilih kotak centang Enkripsi rahasia dengan kunci KMS kustom.
14. Untuk konfigurasi Sumber Daya EC2, tentukan informasi tentang sumber daya komputasi yang digunakan untuk pelatihan model.
 - a. Untuk jenis Instance, pilih jenis instance yang ingin Anda jalankan.
 - b. Untuk hitungan Instance, masukkan jumlah instance.
 - c. Untuk ukuran Volume dalam GB, masukkan ukuran volume penyimpanan ML.
15. Pilih Buat model terlatih dari versi.

API

Untuk menjalankan pekerjaan pelatihan tambahan (API)

Jalankan kode berikut dengan parameter spesifik Anda:

```
import boto3
acr_ml_client= boto3.client('cleanroomsm1')

acr_ml_client.create_trained_model(
    membershipIdentifier= 'membership_id',
```

```

    configuredModelAlgorithmAssociationArn = 'arn:aws:cleanrooms-
ml:region:account:membership/membershipIdentifier/configured-model-algorithm-
association/identifier',
    name='trained_model_name',
    resourceConfig={
        'instanceType': 'ml.m5.xlarge',
        'volumeSizeInGB': 1
    },
    incrementalTrainingDataChannels=[
        {
            'trainedModelArn': trained_model_arn,
            'channelName': 'channel_name'
        },
    ]
    dataChannels=[
        {
            'mlInputChannelArn': channel_arn_1,
            'channelName': 'channel_name'
        },
        {
            'mlInputChannelArn': channel_arn_2,
            'channelName': 'channel_name'
        }
    ]
)

```

Note

Batas: Maksimum total 20 saluran (termasuk keduanya dataChannels dan incrementalTrainingDataChannels).

Note

Setelah model terlatih dibuat, Anda tidak dapat mengeditnya. Untuk membuat perubahan, hapus model terlatih dan buat yang baru.

Menggunakan pelatihan terdistribusi di AWS Clean Rooms

Prasyarat:

- Sebuah Akun AWS dengan akses ke AWS Clean Rooms
- Kolaborasi yang dibuat di AWS Clean Rooms
- Algoritma model yang dikonfigurasi yang mendukung pelatihan terdistribusi
- Dataset besar yang cocok untuk pemrosesan terdistribusi
- Izin yang sesuai untuk membuat dan mengelola model ML dalam kolaborasi
- Kuota Amazon EC2 yang memadai untuk menjalankan beberapa instans untuk pelatihan terdistribusi

Pelatihan terdistribusi memanfaatkan kekuatan banyak node komputasi yang bekerja secara paralel untuk memproses sejumlah besar data dan memperbarui parameter model secara efisien.

Untuk informasi selengkapnya tentang pelatihan [terdistribusi](#), lihat [Konsep pelatihan terdistribusi](#) di Panduan Pengembang Amazon SageMaker AI.

Console

Untuk menjalankan pekerjaan pelatihan terdistribusi (konsol)

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pada halaman Kolaborasi, pilih kolaborasi tempat Anda ingin membuat model terlatih.
4. Setelah kolaborasi terbuka, pilih tab model ML.
5. Di bawah model Custom MS, di bagian Model terlatih, pilih Buat model terlatih.
6. Pada halaman Buat model terlatih, untuk algoritma model Terkait, tentukan algoritme
7. Untuk detail model Terlatih, masukkan yang berikut ini:
 - a. Untuk Nama, masukkan nama unik untuk model dalam kolaborasi.
 - b. (Opsional) Untuk Deskripsi, masukkan deskripsi model yang dilatih.
 - c. Untuk mode input data Pelatihan, pilih salah satu dari berikut ini:
 - Pilih File jika Anda memiliki kumpulan data yang lebih kecil yang dapat muat pada volume penyimpanan ML dan Anda lebih memilih akses sistem file tradisional untuk skrip pelatihan Anda.

- Pilih Pipe untuk kumpulan data besar untuk mengalirkan data langsung dari S3, menghindari kebutuhan untuk mengunduh semuanya ke disk, yang dapat meningkatkan kecepatan pelatihan dan mengurangi persyaratan penyimpanan.
- Pilih FastFile apakah Anda ingin menggabungkan manfaat streaming dari S3 dengan akses sistem file, terutama untuk membaca data secara berurutan atau ketika berhadapan dengan lebih sedikit file untuk waktu startup yang lebih cepat.

8. Untuk detail saluran masukan ML, lakukan hal berikut:

- a. Untuk saluran input ML, tentukan saluran input ML yang menyediakan data ke algoritma model.

Untuk menambahkan saluran lain, pilih Tambahkan saluran input ML lainnya. Anda dapat menambahkan hingga 19 saluran input ML tambahan.

- b. Untuk nama Saluran, masukkan nama saluran input ML.
- c. Untuk jenis distribusi data Amazon S3, pilih salah satu dari berikut ini:
- Pilih Sepenuhnya direplikasi untuk memberikan setiap instance pelatihan dengan salinan lengkap kumpulan data Anda. Ini berfungsi paling baik ketika kumpulan data Anda cukup kecil untuk muat dalam memori atau ketika setiap instance membutuhkan akses ke semua data.
 - Pilih Sharded by S3 kunci untuk membagi dataset Anda di seluruh instans pelatihan berdasarkan kunci S3. Setiap instance menerima sekitar $1/n$ dari total objek S3, di mana 'n' adalah jumlah instance. Ini berfungsi paling baik untuk kumpulan data besar yang ingin Anda proses secara paralel.

 Note

Pertimbangkan ukuran set data dan persyaratan pelatihan Anda saat memilih jenis distribusi. Sepenuhnya direplikasi menyediakan akses data yang lengkap tetapi membutuhkan lebih banyak penyimpanan, sementara Sharded by S3 key memungkinkan pemrosesan terdistribusi kumpulan data besar.

9. Untuk Durasi pelatihan maksimum, pilih jumlah waktu maksimum yang Anda inginkan untuk melatih model Anda.

10. Untuk Hyperparameters, tentukan parameter spesifik algoritme apa pun dan nilai yang dimaksudkan. Hyperparameter khusus untuk model yang dilatih dan digunakan untuk menyempurnakan pelatihan model.
11. Untuk variabel Lingkungan, tentukan variabel spesifik algoritme apa pun dan nilai yang dimaksudkan. Variabel lingkungan diatur dalam wadah Docker.
12. Untuk Enkripsi, untuk menggunakan kustom AWS KMS key, pilih kotak centang Enkripsi rahasia dengan kunci KMS kustom.
13. Untuk konfigurasi Sumber Daya EC2, tentukan informasi tentang sumber daya komputasi yang digunakan untuk pelatihan model.
 - a. Untuk jenis Instance, pilih jenis instance yang ingin Anda jalankan.

Jenis instans yang didukung untuk pelatihan terdistribusi adalah:

 - ml.m5.4xlarge
 - ml.m5.12xlarge
 - ml.m5.2xlarge
 - ml.g5.12xbesar
 - ml.g5.24xbesar
 - b. Untuk hitungan Instance, masukkan jumlah instance.
 - c. Untuk ukuran Volume dalam GB, masukkan ukuran volume penyimpanan ML.
14. Pilih Buat model terlatih.

API

Untuk menjalankan pekerjaan pelatihan terdistribusi (API)

Jalankan kode berikut dengan parameter spesifik Anda:

```
import boto3
acr_ml_client= boto3.client('cleanroomsm1')

acr_ml_client.create_trained_model(
    membershipIdentifier= 'membership_id',
    configuredModelAlgorithmAssociationArn = 'arn:aws:cleanrooms-
ml:region:account:membership/membershipIdentifier/configured-model-algorithm-
association/identifier',
    name='trained_model_name',
```

```
trainingInputMode: "File",
resourceConfig={
  'instanceCount': "3"
  'instanceType': "ml.m5.xlarge",
  'volumeSizeInGB': 3
},
dataChannels=[
  {
    "mlInputChannelArn": channel_arn_1,
    "channelName": "channel_name",
    "S3DataDistributionType": "FullyReplicated"
  }
]
)
```

Note

Setelah model terlatih dibuat, Anda tidak dapat mengeditnya. Untuk membuat perubahan, hapus model terlatih dan buat yang baru.

Mengekspor artefak model dari AWS Clean Rooms

Tugas ini bersifat opsional dan harus diselesaikan ketika Anda telah menetapkan kemampuan CAN_RECEIVE_MODEL_OUTPUT anggota untuk anggota kolaborasi.

Setelah pelatihan model selesai, anggota yang melatih model dapat memulai ekspor artefak model. Anggota yang melatih model memilih siapa yang akan menerima artefak model, asalkan anggota dapat menerima hasil dan konfigurasi ML yang valid.

Console

Untuk mengkonfigurasi algoritma model ML kustom (konsol)

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pada halaman Kolaborasi, pilih kolaborasi yang berisi model kustom yang ingin Anda ekspor.

4. Setelah kolaborasi terbuka, pilih tab Model ML, lalu pilih model Anda dari tabel model terlatih khusus
5. Pada halaman detail model terlatih khusus, klik Ekspor keluaran model.
6. Untuk keluaran model Ekspor, untuk detail keluaran model Ekspor, masukkan Nama dan Deskripsi opsional.

Pilih anggota mana yang akan menerima artefak model dalam keluaran Model yang diekspor ke anggota daftar drop-down kolaborasi.

7. Pilih Ekspor.

Hasilnya diekspor ke jalur berikut di lokasi Amazon S3 yang ditentukan dalam konfigurasi ML: `yourSpecifiedS3Path/collaborationIdentifier/trainedModelName/callerAccountId/jobName` Hanya File yang akan diekspor, hingga ukuran file maksimum yang ditentukan, yang Anda pilih saat mengaitkan algoritma model yang dikonfigurasi yang diekspor.

API

Untuk mengkonfigurasi algoritma model HTML kustom (API)

Memulai ekspor model dengan menjalankan kode berikut:

```
import boto3
acr_ml_client= boto3.client('cleanroomsm1')

acr_ml_client.start_trained_model_export_job(
    membershipIdentifier='membership_id',
    trainedModelArn='arn:aws:cleanrooms-ml:region:account:membership/
membershipIdentifier/trained-model/identifier',
    outputConfiguration={
        'member': {
            'accountId': 'model_output_receiver_account'
        }
    },
    name='export_job_name'
)
```

Hasilnya diekspor ke jalur berikut di lokasi Amazon S3 yang ditentukan dalam konfigurasi ML: `yourSpecifiedS3Path/collaborationIdentifier/trainedModelName/`

`callerAccountId/jobName` Hanya `filesToExport`, hingga yang `maxSize` ditentukan, yang Anda pilih saat mengaitkan algoritma model yang dikonfigurasi yang diekspor.

Menjalankan inferensi pada model terlatih di AWS Clean Rooms

MS

Anggota dengan kemampuan untuk menjalankan kueri juga dapat memulai pekerjaan inferensi setelah pekerjaan pelatihan selesai. Mereka memilih kumpulan data inferensi yang ingin mereka jalankan inferensi dan mereferensikan keluaran model terlatih yang ingin mereka jalankan dengan wadah inferensi.

Anggota yang akan menerima output inferensi harus diberikan kemampuan `CAN_RECEIVE_INFERENCE_OUTPUT` anggota.

Console

Untuk membuat pekerjaan inferensi model (konsol)

1. Masuk ke Konsol Manajemen AWS dan buka AWS Clean Rooms konsol di <https://console.aws.amazon.com/cleanrooms>.
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pada halaman Kolaborasi, pilih kolaborasi yang berisi model kustom tempat Anda ingin membuat pekerjaan inferensi.
4. Setelah kolaborasi terbuka, pilih tab Model ML, lalu pilih model Anda dari tabel model terlatih khusus.
5. Pada halaman detail model terlatih khusus, klik Mulai pekerjaan inferensi.
6. Untuk memulai pekerjaan inferensi, untuk detail pekerjaan Inferensi, masukkan Nama dan Deskripsi opsional.

Masukkan informasi berikut:

- Algoritma model terkait — Algoritma model terkait yang digunakan selama pekerjaan inferensi.
- Detail saluran input ML — Saluran input ML yang akan menyediakan data untuk pekerjaan inferensi ini.

- Transform Resources — Instance komputasi yang digunakan untuk melakukan fungsi transformasi dari pekerjaan inferensi.
- Konfigurasi keluaran — Siapa yang akan menerima output pekerjaan inferensi dan jenis output MIME.
- Enkripsi - Pilih pengaturan enkripsi Sesuaikan untuk menentukan kunci KMS Anda sendiri dan informasi terkait. Jika tidak, Clean Rooms ML akan mengelola enkripsi.
- Ubah detail pekerjaan — Muatan maksimum pekerjaan inferensi, dalam MB.
- Variabel lingkungan — Setiap variabel lingkungan yang diperlukan untuk mengakses gambar kontainer pekerjaan inferensi.

7. Pilih Mulai pekerjaan inferensi.

Hasilnya diekspor ke jalur berikut di lokasi Amazon S3 yang ditentukan dalam konfigurasi ML: `yourSpecifiedS3Path/collaborationIdentifier/trainedModelName/callerAccountId/jobName`

API

Untuk membuat pekerjaan inferensi model (API)

Memulai pekerjaan inferensi dengan menjalankan kode berikut:

```
import boto3
acr_ml_client= boto3.client('cleanroomsm1')

acr_ml_client.start_trained_model_inference_job(
    name="inference_job",
    membershipIdentifier='membership_id',
    trainedModelArn='arn:aws:cleanrooms-ml:region:account:membership/
membershipIdentifier/trained-model/identifier',

    dataSource={
        "mlInputChannelArn": 'channel_arn_3'
    },
    resourceConfig={'instanceType': 'ml.m5.xlarge'},
    outputConfiguration={
        'accept': 'text/csv',
        'members': [
            {
                "accountId": 'member_account_id'
```

```
}  
]  
}  
)
```

Hasilnya diekspor ke jalur berikut di lokasi Amazon S3 yang ditentukan dalam konfigurasi ML: `yourSpecifiedS3Path/collaborationIdentifier/trainedModelName/callerAccountId/jobName`

Buat model AWS Clean Rooms AWS sebagai penyedia data pelatihan

Model mirip adalah model data penyedia data pelatihan yang memungkinkan penyedia data benih untuk membuat segmen serupa dari data penyedia data pelatihan yang paling mirip dengan data benih mereka. Untuk membuat model mirip yang dapat digunakan dalam kolaborasi, Anda harus mengimpor data pelatihan Anda, membuat model mirip, mengonfigurasi model yang mirip, dan kemudian mengaitkannya dengan kolaborasi.

Bekerja dengan model yang mirip mengharuskan dua pihak, penyedia data pelatihan dan penyedia data benih, bekerja secara berurutan AWS Clean Rooms untuk membawa data mereka ke dalam kolaborasi. Ini adalah alur kerja yang harus diselesaikan oleh penyedia data pelatihan terlebih dahulu:

1. Data penyedia data pelatihan harus disimpan dalam tabel katalog AWS Glue data interaksi item pengguna. Minimal, data pelatihan harus berisi kolom ID pengguna, kolom ID interaksi, dan kolom stempel waktu.
2. Penyedia data pelatihan mendaftarkan data pelatihan dengan AWS Clean Rooms.
3. Penyedia data pelatihan membuat model mirip yang dapat dibagikan dengan beberapa penyedia data benih. Model mirip adalah jaringan saraf dalam yang dapat memakan waktu hingga 24 jam untuk dilatih. Ini tidak dilatih ulang secara otomatis dan kami sarankan Anda melatih ulang model setiap minggu.
4. Penyedia data pelatihan mengonfigurasi model yang mirip, termasuk apakah akan berbagi metrik relevansi dan lokasi Amazon S3 dari segmen keluaran. Penyedia data pelatihan dapat membuat beberapa model mirip yang dikonfigurasi dari satu model mirip.
5. Penyedia data pelatihan mengaitkan model audiens yang dikonfigurasi dengan kolaborasi yang dibagikan dengan penyedia data benih.

Setelah penyedia data pelatihan selesai membuat model ML, [penyedia data benih dapat membuat dan mengeksplor segmen yang mirip](#).

Topik

- [Mengimpor data pelatihan](#)
- [Membuat model yang mirip](#)

- [Mengkonfigurasi model yang mirip](#)
- [Megaikan model mirip yang dikonfigurasi](#)
- [Memperbarui model mirip yang dikonfigurasi](#)

Mengimpor data pelatihan

Note

Anda hanya dapat menyediakan kumpulan data pelatihan untuk digunakan dalam model mirip Clean Rooms ML. yang memiliki data yang disimpan di Amazon S3. Namun, Anda dapat menyediakan data benih untuk model mirip menggunakan SQL yang berjalan di seluruh data yang disimpan dalam sumber data yang didukung.

Sebelum Anda membuat model mirip, Anda harus menentukan AWS Glue tabel yang berisi data pelatihan. Clean Rooms ML tidak menyimpan salinan data ini, hanya metadata yang memungkinkannya mengakses data.

Untuk mengimpor data pelatihan di AWS Clean Rooms

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih model AWS ML.
3. Pada tab Kumpulan data Pelatihan, pilih Buat kumpulan data pelatihan.
4. Pada halaman Buat kumpulan data pelatihan, untuk detail kumpulan data Pelatihan, masukkan Nama dan Deskripsi opsional.
5. Pilih sumber data Pelatihan dengan memilih Database dan Tabel yang ingin Anda konfigurasi dari daftar dropdown.

Note

Untuk memverifikasi bahwa ini adalah tabel yang benar, lakukan salah satu dari yang berikut:

- Pilih Lihat di AWS Glue.
- Aktifkan Lihat skema untuk melihat skema.

6. Untuk detail Pelatihan, pilih kolom Pengenal pengguna, kolom pengenal item, dan kolom Timestamp dari daftar tarik-turun. Data pelatihan harus berisi tiga kolom ini. Anda juga dapat memilih kolom lain yang ingin Anda sertakan dalam data pelatihan.

Data di kolom Timestamp harus dalam waktu epoch Unix dalam format detik.

7. (Opsional) Jika Anda memiliki kolom Tambahan untuk dilatih, pilih nama Kolom dan Ketik dari daftar dropdown.
8. Dalam akses Layanan, Anda harus menentukan peran layanan yang dapat mengakses data Anda dan memberikan kunci KMS jika data Anda dienkripsi. Pilih Buat dan gunakan peran layanan baru dan Clean Rooms akan secara otomatis membuat peran layanan dan menambahkan kebijakan izin yang diperlukan. Pilih Gunakan peran layanan yang ada dan masukkan di bidang Nama peran layanan jika Anda memiliki peran layanan tertentu yang ingin Anda gunakan.

Jika data Anda dienkripsi, masukkan kunci KMS Anda di AWS KMS keybidang, atau klik Buat AWS KMS key untuk menghasilkan kunci KMS baru.

9. Jika Anda ingin mengaktifkan Tag untuk kumpulan data pelatihan, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.
10. Pilih Buat kumpulan data pelatihan.

Untuk tindakan API terkait, lihat [CreateTrainingDataset](#).

Membuat model yang mirip

Setelah Anda membuat kumpulan data pelatihan, Anda siap untuk membuat model yang mirip. Anda dapat membuat banyak model mirip dari satu kumpulan data pelatihan.

Anda harus membuat database default di AWS Glue Data Catalog atau menyertakan `glue:createDatabase` izin dalam peran yang disediakan.

Untuk membuat model yang mirip di AWS Clean Rooms

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih model AWS ML.
3. Pada tab Lookalike models, pilih Create lookalike model.

4. Pada halaman Buat model mirip, untuk detail model Lookalike, masukkan Nama dan Deskripsi opsional.
 - a. Pilih kumpulan data Pelatihan yang ingin Anda modelkan dari daftar tarik-turun.

 Note

Untuk memverifikasi bahwa ini adalah kumpulan data pelatihan yang benar, aktifkan Tampilkan detail kumpulan data pelatihan untuk melihat detailnya.

Untuk membuat kumpulan data pelatihan baru, pilih Buat kumpulan data pelatihan.

- b. (Opsional) Masukkan jendela Pelatihan.
5. Jika Anda ingin mengaktifkan pengaturan enkripsi khusus untuk model yang mirip, pilih Sesuaikan pengaturan enkripsi dan kemudian masukkan kunci KMS.
6. Jika Anda ingin mengaktifkan Tag untuk model mirip, pilih Tambahkan tag baru dan kemudian masukkan pasangan Kunci dan Nilai.
7. Pilih Buat model yang mirip.

 Note

Pelatihan model dapat memakan waktu beberapa jam hingga 2 hari.

Untuk tindakan API terkait, lihat [CreateAudienceModel](#).

Mengkonfigurasi model yang mirip

Setelah Anda membuat model yang mirip, Anda siap mengonfigurasinya untuk digunakan dalam kolaborasi. Anda dapat membuat beberapa model mirip yang dikonfigurasi dari satu model mirip.

Untuk mengonfigurasi model yang mirip di AWS Clean Rooms

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih model AWS ML.
3. Pada tab Configured lookalike models, pilih Configure lookalike model.

4. Pada halaman Configure lookalike model, untuk detail model mirip Configured, masukkan Nama dan Deskripsi opsional.
 - a. Pilih model Lookalike yang ingin Anda konfigurasikan dari daftar dropdown.
-  **Note**

Untuk memverifikasi bahwa ini adalah model mirip yang benar, aktifkan Tampilkan detail model yang mirip untuk melihat detailnya.

Untuk membuat model mirip baru, pilih Buat model mirip mirip.
- b. Pilih ukuran benih pencocokan minimum yang Anda inginkan. Ini adalah jumlah minimum pengguna dalam data penyedia data benih yang tumpang tindih dengan pengguna dalam data pelatihan. Nilai ini harus lebih besar dari 0.
 5. Agar Metrik dapat dibagikan dengan anggota lain, pilih apakah Anda ingin penyedia data benih dalam kolaborasi Anda menerima metrik model, termasuk skor relevansi.
 6. Untuk lokasi tujuan segmen Lookalike, masukkan bucket Amazon S3 tempat segmen mirip mirip diekspor. Bucket ini harus terletak di wilayah yang sama dengan sumber daya Anda yang lain.
 7. Untuk akses Layanan, pilih nama peran layanan yang ada yang akan digunakan untuk mengakses tabel ini.
 8. Untuk konfigurasi ukuran nampun lanjutan, tentukan jenis ukuran Audiens sebagai nomor Absolut atau Persentase.
 9. Jika Anda ingin mengaktifkan Tag untuk sumber daya tabel yang dikonfigurasi, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.
 10. Pilih Konfigurasikan model yang mirip.

Untuk tindakan API terkait, lihat [CreateConfiguredAudienceModel](#).

Mengaitkan model mirip yang dikonfigurasi

Setelah Anda mengonfigurasi model yang mirip, Anda dapat mengaitkannya dengan kolaborasi.

Untuk mengaitkan model mirip yang dikonfigurasi di AWS Clean Rooms

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).

2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pada tab Dengan keanggotaan aktif, pilih kolaborasi.
4. Pada tab model ML, di bawah R model mirip, pilih Associate eady-to-use lookalike model.
5. Pada halaman model mirip yang dikonfigurasi Associate, untuk detail asosiasi model mirip yang Dikonfigurasi:
 - a. Masukkan Nama untuk model audiens yang dikonfigurasi terkait.
 - b. Masukkan Deskripsi tabel.

Deskripsi membantu membedakan antara model audiens terkonfigurasi terkait lainnya dengan nama yang mirip.

6. Untuk model mirip yang dikonfigurasi, pilih model mirip yang dikonfigurasi dari daftar tarik-turun.
7. Pilih Kaitkan.

Untuk tindakan API terkait, lihat [CreateConfiguredAudienceModelAssociation](#).

Memperbarui model mirip yang dikonfigurasi

Setelah mengaitkan model mirip yang dikonfigurasi, Anda dapat memperbaruinya untuk mengubah informasi seperti nama, metrik yang akan dibagikan, atau menampilkan lokasi Amazon S3.

Untuk memperbarui model mirip yang dikonfigurasi terkait di AWS Clean Rooms

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih model AWS ML.
3. Pada tab Configured lookalike models, di bawah model mirip, pilih model Ready-to-use mirip yang dikonfigurasi dan pilih Edit.
4. Pada halaman Edit, untuk detail asosiasi model mirip yang Dikonfigurasi:
 - a. Perbarui Nama dan Deskripsi opsional.
 - b. Pilih model Lookalike yang ingin Anda konfigurasi dari daftar dropdown.
 - c. Pilih ukuran benih pencocokan minimum yang Anda inginkan. Ini adalah jumlah minimum pengguna dalam data penyedia data benih yang tumpang tindih dengan pengguna dalam data pelatihan. Nilai ini harus lebih besar dari 0.

5. Agar Metrik dapat dibagikan dengan anggota lain, pilih apakah Anda ingin penyedia data benih dalam kolaborasi Anda menerima metrik model, termasuk skor relevansi.
6. Untuk lokasi tujuan segmen Lookalike, masukkan bucket Amazon S3 tempat segmen mirip diekspor. Bucket ini harus terletak di wilayah yang sama dengan sumber daya Anda yang lain.
7. Untuk akses Layanan, pilih nama peran layanan yang ada yang akan digunakan untuk mengakses tabel ini.
8. Untuk konfigurasi ukuran tempat sampah lanjutan, pilih cara Anda ingin mengonfigurasi ukuran tempat sampah audiens.
9. Pilih Simpan perubahan.

Untuk tindakan API terkait, lihat [UpdateConfiguredAudienceModel](#).

Membuat model AWS Clean Rooms AWS sebagai penyedia data benih

Setelah penyedia data pelatihan selesai membuat model ML, penyedia data benih dapat membuat dan mengeksport segmen yang mirip. Segmen mirip adalah bagian dari data pelatihan yang paling mirip dengan data benih.

Ini adalah alur kerja yang harus diselesaikan oleh penyedia data benih:

1. Data penyedia data seed dapat disimpan dalam bucket Amazon S3 atau dapat berasal dari hasil kueri.
2. Penyedia data benih membuka kolaborasi yang mereka bagikan dengan penyedia data pelatihan.
3. Penyedia data seed membuat segmen mirip dari tab Clean Rooms di halaman kolaborasi.
4. Penyedia data benih dapat mengevaluasi metrik relevansi, jika dibagikan, dan mengeksport segmen yang mirip untuk digunakan di luar. AWS Clean Rooms

Topik

- [Membuat segmen yang mirip](#)
- [Mengeksport segmen yang mirip](#)

Membuat segmen yang mirip

Note

Anda hanya dapat menyediakan kumpulan data pelatihan untuk digunakan dalam model mirip Clean Rooms ML. yang memiliki data yang disimpan di Amazon S3. Namun, Anda dapat menyediakan data benih untuk model mirip menggunakan SQL yang berjalan di seluruh data yang disimpan dalam sumber data yang didukung.

Segmen mirip adalah bagian dari data pelatihan yang paling mirip dengan data benih.

Untuk membuat segmen mirip di AWS Clean Rooms

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pada tab Dengan keanggotaan aktif, pilih kolaborasi.
4. Pada tab Model ML, pilih Create lookalike segment.
5. Pada halaman Buat segmen mirip mirip, untuk model mirip mirip yang dikonfigurasi terkait, pilih model mirip mirip yang dikonfigurasi terkait untuk digunakan untuk segmen mirip ini.
6. Untuk detail segmen Lookalike masukkan Nama dan Deskripsi opsional.
7. Untuk profil Seed, pilih metode Seed Anda dengan memilih opsi dan kemudian mengambil tindakan yang disarankan.

Opsi	Tindakan yang disarankan
Jalur Amazon S3	1. Pilih lokasi Amazon S3. 2. (Opsional) Pilih Sertakan profil benih dalam output.
Kueri SQL	Tulis kueri SQL dan gunakan hasilnya sebagai data benih.
Template analisis	Pilih template analisis dari daftar dropdown dan gunakan hasil yang dibuat oleh template analisis.

8. Pilih jenis Worker yang akan digunakan saat membuat sumber data ini. Jenis pekerja default adalah CR.1X. Tentukan Jumlah pekerja yang akan digunakan. Standarnya adalah pekerja nomor 16. Untuk menentukan properti Spark:
 - a. Perluas properti Spark.
 - b. Pilih Tambahkan properti Spark.
 - c. Pada kotak dialog properti Spark, pilih nama Properti dari daftar dropdown dan masukkan Nilai.

Tabel berikut memberikan definisi untuk setiap properti.

Untuk informasi selengkapnya tentang properti Spark, lihat [Spark Properties](#) di dokumentasi Apache Spark.

Nama Properti	Deskripsi	nilai default
Spark.Task.MaxFailures	Mengontrol berapa kali berturut-turut tugas dapat gagal sebelum pekerjaan gagal. Membutuhkan nilai yang lebih besar dari atau sama dengan 1. Jumlah percobaan ulang yang diizinkan sama dengan nilai ini dikurangi 1. Hitungan kegagalan diatur ulang jika ada upaya yang berhasil. Kegagalan di berbagai tugas tidak menumpuk menuju batas ini.	4
spark.sql.file.maxPartitionBytes	Menetapkan jumlah maksimum byte untuk dikemas ke dalam satu partisi ketika membaca dari sumber berbasis file seperti Parquet, JSON, dan ORC.	128MB
Spark.hadoop.fs.s3.maxRetries	Menetapkan jumlah maksimum percobaan ulang untuk operasi file Amazon S3.	
spark.network.timeout	Menetapkan batas waktu default untuk semua interaksi jaringan. Mengganti pengaturan batas waktu berikut jika tidak dikonfigurasi: - spark.storage.blockManagerHeartbeatTimeoutMs - spark.shuffle.io.connectionTimeout - Spark.rpc.askTimeout - Spark.rpc.lookupTimeout	120s

Nama Properti	Deskripsi	nilai default
spark.rdd.compress	Menentukan apakah untuk mengompres partisi RDD serial menggunakan spark.io.compression.codec. Berlaku StorageLevel untuk MEMORY_ONLY_SER di Java dan Scala, atau .MEMORY_ONLY dengan Python. StorageLevel Mengurangi ruang penyimpanan tetapi membutuhkan waktu pemrosesan CPU tambahan.	SALAH
spark.shuffle.spill.compress	Menentukan apakah untuk mengompres data tumpahan shuffle menggunakan spark.io.compression.codec.	BETUL
spark.sql.adaptif.advisoryPartitionSizeInBytes	Menetapkan ukuran target dalam byte untuk partisi acak selama pengoptimalan adaptif ketika spark.sql.adaptive.enabled benar. Mengontrol ukuran partisi saat menggabungkan partisi kecil atau memisahkan partisi miring.	(nilai spark.sql.adaptive.shuffle.targetPostShuffleInputSize)
spark.sql.adaptif.autoBroadcastJoinAmbang	Menetapkan ukuran tabel maksimum dalam byte untuk penyiaran ke node pekerja selama bergabung. Berlaku hanya dalam kerangka adaptif. Menggunakan nilai default yang sama seperti spark.sql.autoBroadcastJoinAmbang batas. Setel ke -1 untuk menonaktifkan penyiaran.	(tidak ada)
spark.sql.adaptive.coalescePartitions.enabled	Menentukan apakah akan menggabungkan partisi shuffle bersebelahan berdasarkan spark.sql.adaptive.advisoryPartitionSizeInBytes untuk mengoptimalkan ukuran tugas. Membutuhkan spark.sql.adaptive.enabled agar benar.	BETUL

Nama Properti	Deskripsi	nilai default
Spark.sql.adaptive. .CoalescePartitions. initialPartitionNum	Mendefinisikan jumlah awal partisi shuffle sebelum penggabungan. Memerlukan spark.sql.adaptive.enabled dan spark.sql.adaptive.CoalescePartitions.enabled menjadi true. Default ke nilai spark.sql.shuffle.partitions.	(tidak ada)
Spark.sql.adaptive. .CoalescePartitions. minPartitionSize	Menetapkan ukuran minimum untuk partisi shuffle yang digabungkan untuk mencegah partisi menjadi terlalu kecil selama optimasi adaptif.	1 MB
spark.sql.adaptive. .coalescepartition s.parallelismFirst	Menentukan apakah untuk menghitung ukuran partisi berdasarkan paralelisme cluster bukan spark.sql.adaptive.advisoryPartitionSizeInBytes selama penggabungan partisi. Menghasilkan ukuran partisi yang lebih kecil dari ukuran target yang dikonfigurasi untuk memaksimalkan paralelisme. Kami merekomendasikan pengaturan ini ke false pada cluster sibuk untuk meningkatkan pemanfaatan sumber daya dengan mencegah tugas-tugas kecil yang berlebihan.	BETUL
spark.sql.adaptive. .enabled	Menentukan apakah akan mengaktifkan eksekusi kueri adaptif untuk mengoptimalkan kembali rencana kueri selama eksekusi kueri, berdasarkan statistik runtime yang akurat.	BETUL
spark.sql.adaptif. forceOptimizeSkewed dBergabungh	Menentukan apakah akan memaksa mengaktifkan OptimizeSkewedJoin bahkan jika memperkenalkan shuffle tambahan.	SALAH

Nama Properti	Deskripsi	nilai default
spark.sql.adaptif.localShuffleReader.diaktifkan	Menentukan apakah akan menggunakan pembaca shuffle lokal ketika partisi shuffle tidak diperlukan, seperti setelah mengonversi dari gabungan sort-merge ke broadcast-hash bergabung. Membutuhkan spark.sql.adaptive.enabled agar benar.	BETUL
spark.sql.adaptif.maxShuffledHashJoinLocalMapThreshold	Menetapkan ukuran partisi maksimum dalam byte untuk membangun peta hash lokal. Memprioritaskan gabungan hash yang diacak daripada gabungan sort-merge saat: • Nilai ini sama atau melebihi spark.sql.adaptive.advisoryPartitionSizeInBytes • Semua ukuran partisi berada dalam batas ini Mengesampingkan spark.sql.join.preferSortMergeBergabunglah dengan pengaturan.	0 byte
spark.sql.adaptif.optimizeSkewsInRebalancePartitions.diaaktifkan	Menentukan apakah akan mengoptimalkan partisi shuffle miring dengan membaginya menjadi partisi yang lebih kecil berdasarkan spark.sql.adaptive.advisoryPartitionSizeInBytes. Membutuhkan spark.sql.adaptive.enabled agar benar.	BETUL
spark.sql.adaptif.rebalancePartitionsSmallPartitionFactor	Mendefinisikan faktor ambang ukuran untuk menggabungkan partisi selama pemisahan. Partisi yang lebih kecil dari faktor ini dikalikan dengan spark.sql.adaptive.advisoryPartitionSizeInBytes digabung.	0,2

Nama Properti	Deskripsi	nilai default
Spark.sql.adaptive.skewjoin.enabled	Menentukan apakah untuk menangani data miring dalam bergabung diacak dengan memisahkan dan opsional mereplikasi partisi miring. Berlaku untuk gabungan hash sort-merge dan shuffled. Membutuhkan spark.sql.adaptive.enabled agar benar.	BETUL
Spark.sql.adaptive.skewJoin.skewedPartitionFactor	Menentukan faktor ukuran yang menentukan kemiringan partisi. Partisi miring ketika ukurannya melebihi keduanya: - Faktor ini dikalikan dengan ukuran partisi median - Nilai spark.sql.adaptive.skewJoin.skewedPartitionThresholdInBytes	5
Spark.sql.adaptive.skewJoin.skewedPartitionThresholdInBytes	Menetapkan ambang ukuran dalam byte untuk mengidentifikasi partisi miring. Partisi miring ketika ukurannya melebihi keduanya: - Ambang batas ini - Ukuran partisi median dikalikan dengan spark.sql.adaptive.skewJoin.skewedPartitionFactor Kami merekomendasikan pengaturan nilai ini lebih besar dari spark.sql.adaptive.advisoryPartitionSizeInBytes.	256MB
spark.sql.autoBroadcastJoinAmbang	Menetapkan ukuran tabel maksimum dalam byte untuk penyiaran ke node pekerja selama bergabung. Setel ke -1 untuk menonaktifkan penyiaran.	10MB

Nama Properti	Deskripsi	nilai default
Spark.Sql.BroadcastTimeout	Mengontrol periode batas waktu dalam hitungan detik untuk operasi siaran selama siaran bergabung.	300 detik
spark.sql.cbo.enabled	Menentukan apakah akan mengaktifkan optimasi berbasis biaya (CBO) untuk estimasi statistik rencana.	SALAH
Spark.sql.cbo.joinreorder.dp.star.filter	Menentukan apakah akan menerapkan heuristik filter star-join selama enumerasi bergabung berbasis biaya.	SALAH
Spark.sql.cbo.joinreorder.dp.Threshold	Menetapkan jumlah maksimum node bergabung yang diizinkan dalam algoritma pemrograman dinamis.	12
Spark.sql.cbo.joinreOrder.enabled	Menentukan apakah akan mengaktifkan pengurutan ulang bergabung dalam optimasi berbasis biaya (CBO).	SALAH
Spark.sql.cbo.planstats.enabled	Menentukan apakah untuk mengambil jumlah baris dan statistik kolom dari katalog selama pembuatan rencana logis.	SALAH
spark.sql.cbo.starSchemaDetection	Menentukan apakah akan mengaktifkan pengurutan ulang gabungan berdasarkan deteksi skema bintang.	SALAH
spark.sql.file.maxPartitionNum	Menetapkan jumlah maksimum target partisi file split untuk sumber berbasis file (Parquet, JSON, dan ORC). Rescales partisi ketika hitungan awal melebihi nilai ini. Ini adalah target yang disarankan, bukan batas yang dijamin.	(tidak ada)

Nama Properti	Deskripsi	nilai default
spark.sql.file.maxRecord sPerBerkas	Menetapkan jumlah maksimum catatan untuk menulis ke satu file. Tidak ada batasan yang berlaku saat disetel ke nol atau nilai negatif.	0
spark.sql.file.minPartitionNum	Menetapkan jumlah minimum target partisi file terpisah untuk sumber berbasis file (Parquet, JSON, dan ORC). Default ke spark.sql. leafNodeDefaultParallelisme. Ini adalah target yang disarankan, bukan batas yang dijamin.	(tidak ada)
spark.sql.inMemoryColumnarStorage.batchsize	Mengontrol ukuran batch untuk caching kolumnar. Meningkatkan ukuran meningkatkan pemanfaatan memori dan kompresi tetapi meningkatkan risiko out-of-memory kesalahan.	10000
spark.sql.inMemoryColumnarpe nyimpanan.dikompresi	Menentukan apakah akan secara otomatis memilih codec kompresi untuk kolom berdasarkan statistik data.	BETUL
spark.sql.inMemoryColumnarPe nyimpanan. enableVectorizedReader	Menentukan apakah untuk mengaktifkan pembacaan vektor untuk caching kolumnar.	BETUL
spark.sql.legacy.allowHash OnMapType	Menentukan apakah untuk memungkinkan operasi hash pada struktur data tipe peta. Pengaturan lama ini mempertahankan kompatibilitas dengan penanganan jenis peta versi Spark yang lebih lama.	

Nama Properti	Deskripsi	nilai default
spark.sql.legacy.allowNegativeScaleOfDecimal	Menentukan apakah untuk memungkinkan nilai skala negatif dalam definisi tipe desimal. Setelan lama ini mempertahankan kompatibilitas dengan versi Spark lama yang mendukung skala desimal negatif.	
spark.sql.legacy.castComplexTypesToString.diaktifkan	Menentukan apakah akan mengaktifkan perilaku warisan untuk casting tipe kompleks untuk string. Menjaga kompatibilitas dengan aturan konversi tipe versi Spark yang lebih lama.	
spark.sql.legacy.charVarcharAsTali	Menentukan apakah untuk memperlakukan jenis CHAR dan VARCHAR sebagai tipe STRING. Pengaturan lama ini menyediakan kompatibilitas dengan penanganan tipe string versi Spark yang lebih lama.	
spark.sql.legacy.createEmptyCollectionUsingStringType	Menentukan apakah untuk membuat koleksi kosong menggunakan elemen tipe string. Setelan lama ini mempertahankan kompatibilitas dengan perilaku inisialisasi koleksi versi Spark yang lebih lama.	
spark.sql.legacy.exponentLiteralAsDesimal. Diaktifkan	Menentukan apakah untuk menafsirkan literal eksponensial sebagai tipe desimal. Pengaturan lama ini mempertahankan kompatibilitas dengan penanganan literal numerik versi Spark yang lebih lama.	
spark.sql.legacy.json.allowEmptyString.diaktifkan	Menentukan apakah untuk mengizinkan string kosong dalam pengolahan JSON. Pengaturan lama ini mempertahankan kompatibilitas dengan perilaku penguraian JSON versi Spark yang lebih lama.	

Nama Properti	Deskripsi	nilai default
spark.sql.legacy.parquet.int96RebaseModelRead	Menentukan apakah akan menggunakan mode rebase INT96 stempel waktu lama saat membaca file Parquet. Pengaturan lama ini mempertahankan kompatibilitas dengan penanganan stempel waktu versi Spark yang lebih lama.	
spark.sql.legacy.timeParserPolicy	Mengontrol perilaku penguraian waktu untuk kompatibilitas mundur. Pengaturan lama ini menentukan bagaimana stempel waktu dan tanggal diuraikan dari string.	
Spark.sql.legacy.typeCoercion.datetimeToString.diaaktifkan	Menentukan apakah akan mengaktifkan perilaku pemaksaan tipe warisan saat mengonversi nilai datetime ke string. Menjaga kompatibilitas dengan aturan konversi datetime versi Spark yang lebih lama.	
spark.sql. maxSinglePartitionByte	Menetapkan ukuran partisi maksimum dalam byte. Perencana memperkenalkan operasi shuffle untuk partisi yang lebih besar untuk meningkatkan paralelisme.	128m
Spark.sql.metadata cache TTLSeconds	Mengontrol time-to-live (TTL) untuk cache metadata. Berlaku untuk metadata file partisi dan cache katalog sesi. Membutuhkan: • Nilai positif lebih besar dari nol • Spark.sql.CatalogImplementation diatur ke sarang • spark.sql.hive. filesourcePartitionFileCacheSize lebih besar dari nol • spark.sql.hive. manageFilesourcePartitions disetel ke true	-1000ms

Nama Properti	Deskripsi	nilai default
spark.sql.optimizer.collapseProjectAlwaysSebaris	Menentukan apakah untuk menciutkan proyeksi yang berdekatan dan ekspresi inline, bahkan ketika itu menyebabkan duplikasi.	SALAH
spark.sql.optimizer.dynamicPartitionPruning.diaktifkan	Menentukan apakah untuk menghasilkan predikat untuk kolom partisi yang digunakan sebagai kunci bergabung.	BETUL
spark.sql.optimizer.enableCsvExpressionOptimasi	Menentukan apakah akan mengoptimalkan ekspresi CSV dalam pengoptimal SQL dengan memangkas kolom yang tidak perlu dari operasi from_csv.	BETUL
spark.sql.optimizer.enableJsonExpressionOptimasi	Menentukan apakah untuk mengoptimalkan ekspresi JSON dalam pengoptimal SQL dengan: • Memangkas kolom yang tidak perlu dari operasi from_json • Menyederhanakan kombinasi from_json dan to_json • Mengoptimalkan operasi named_struct	BETUL
Spark.sql.optimizer.excludedrules	Mendefinisikan aturan pengoptimal untuk dinonaktifkan, diidentifikasi dengan nama aturan yang dipisahkan koma. Beberapa aturan tidak dapat dinonaktifkan karena diperlukan untuk kebenaran. Pengoptimal mencatat aturan mana yang berhasil dinonaktifkan.	(tidak ada)

Nama Properti	Deskripsi	nilai default
Spark.sql.Optimizer.Runtime.BloomFilter.applicationSideScanSizeThreshold	Menetapkan ukuran pemindaian agregat minimum dalam byte yang diperlukan untuk menyuntikkan filter Bloom di sisi aplikasi.	10GB
Spark.sql.Optimizer.Runtime.BloomFilter.creationSideThreshold	Mendefinisikan ambang batas ukuran maksimum untuk menyuntikkan filter Bloom di sisi pembuatan.	10MB
Spark.sql.Optimizer.Runtime.BloomFilter.enabled	Menentukan apakah akan menyisipkan filter Bloom untuk mengurangi data shuffle ketika satu sisi gabungan shuffle memiliki predikat selektif.	BETUL
Spark.sql.Optimizer.Runtime.BloomFilter.expectedNumItems	Mendefinisikan jumlah default item yang diharapkan dalam filter Bloom runtime.	1000000
Spark.sql.Optimizer.Runtime.BloomFilter.maxNumBits	Menetapkan jumlah maksimum bit yang diizinkan dalam filter Bloom runtime.	67108864
Spark.sql.Optimizer.Runtime.BloomFilter.maxNumItems	Menetapkan jumlah maksimum item yang diharapkan yang diizinkan dalam filter Bloom runtime.	4000000
spark.sql.optimizer.runtime.bloomfilter.number.threshold	Membatasi jumlah maksimum filter runtime non-DPP yang diizinkan per kueri untuk mencegah out-of-memory kesalahan pada driver.	10

Nama Properti	Deskripsi	nilai default
spark.sql.optimize r.runtime.bloomfilter.numbits	Mendefinisikan jumlah default bit yang digunakan dalam filter Bloom runtime.	8388608
spark.sql.optimize r.runtime.rowlevelOperationGroupFilter.enabled	Menentukan apakah akan mengaktifkan penyaringan grup runtime untuk operasi tingkat baris. Memungkinkan sumber data untuk: - Pangkas seluruh kelompok data (seperti file atau partisi) menggunakan filter sumber data - Jalankan kueri runtime untuk mengidentifikasi catatan yang cocok - Buang grup yang tidak perlu untuk menghindari penulisan ulang yang mahal Pembatasan: - Tidak semua ekspresi dapat dikonversi ke filter sumber data - Beberapa ekspresi memerlukan evaluasi Spark (seperti subquery)	BETUL
Spark.Sql.Optimizer.RuntimeFilter.semiJoinReduction.diaktifkan	Menentukan apakah akan menyisipkan semi-join untuk mengurangi data shuffle ketika salah satu sisi gabungan shuffle memiliki predikat selektif.	SALAH

Nama Properti	Deskripsi	nilai default
Spark.sql.parquet. AggregatePushdown	Menentukan apakah akan menekan agregat ke Parquet untuk optimasi. Mendukung: • MIN dan MAX untuk jenis boolean, integer, float, dan tanggal • COUNT untuk semua tipe data Melempar pengecualian jika statistik hilang dari footer file Parquet.	SALAH
spark.sql.parquet. columnarReaderBatchSize	Mengontrol jumlah baris di setiap batch pembaca vektor Parquet. Pilih nilai yang menyeimbangkan overhead kinerja dan penggunaan memori untuk mencegah out-of-memory kesalahan.	4096
Spark.sql.Session. Timezone	Mendefinisikan zona waktu sesi untuk menangani stempel waktu dalam literal string dan konversi objek Java. Menerima: • Berbasis wilayah IDs dalam area/city format (seperti Amerika/Los_Angeles) • Offset zona dalam format (+/-) HH, (+/-) HH: mm, atau (+/-) HH: mm: SS (seperti -08 atau + 01:00) • UTC atau Z sebagai alias untuk + 00:00	(nilai zona waktu lokal)
spark.sql.shuffle. partitions	Menetapkan jumlah default partisi untuk pengocokan data selama bergabung atau agregasi. Tidak dapat diubah antara restart kueri streaming terstruktur dari lokasi pos pemeriksaan yang sama.	200

Nama Properti	Deskripsi	nilai default
spark.sql.shuffledHashJoinFaktor	Mendefinisikan faktor perkalian yang digunakan untuk menentukan kelayakan bergabung hash shuffle. Gabungan hash shuffle dipilih ketika ukuran data sisi kecil dikalikan dengan faktor ini kurang dari ukuran data sisi besar.	3
spark.sql.sources.parallelPartitionDiscovery.ambang	Menetapkan jumlah maksimum jalur untuk daftar file sisi driver dengan sumber berbasis file (Parquet, JSON, dan ORC). Ketika terlampaui selama penemuan partisi, file terdaftar menggunakan pekerjaan terdistribusi Spark terpisah.	32
spark.sql.statistical.histogram.enabled	Menentukan apakah akan menghasilkan histogram equi-height selama perhitungan statistik kolom untuk meningkatkan akurasi estimasi. Membutuhkan pemindaian tabel tambahan di luar yang dibutuhkan untuk statistik kolom dasar.	SALAH

9. Untuk akses Layanan, pilih nama peran layanan yang ada yang akan digunakan untuk mengakses tabel ini.
10. Jika Anda ingin mengaktifkan Tag untuk kumpulan data pelatihan, pilih Tambahkan tag baru lalu masukkan pasangan Kunci dan Nilai.
11. Pilih Buat segmen mirip.

Untuk tindakan API terkait, lihat [StartAudienceGenerationJob](#).

Mengekspor segmen yang mirip

Setelah membuat segmen yang mirip, Anda dapat mengekspor data tersebut ke bucket Amazon S3.

Untuk mengekspor segmen yang mirip di AWS Clean Rooms

1. Masuk ke Konsol Manajemen AWS dan buka [AWS Clean Rooms konsol](#) dengan Anda Akun AWS (jika Anda belum melakukannya).
2. Di panel navigasi kiri, pilih Kolaborasi.
3. Pada tab Dengan keanggotaan aktif, pilih kolaborasi.
4. Pada tab Model ML, pilih segmen yang mirip dan pilih Ekspor.
5. Untuk model mirip Ekspor, untuk detail model mirip Ekspor masukkan Nama dan Deskripsi opsional.
6. Untuk ukuran Segmen, pilih ukuran yang Anda inginkan untuk segmen yang diekspor.
7. Pilih Ekspor.

Untuk tindakan API terkait, lihat [StartAudienceExportJob](#).

Pemecahan masalah AWS Clean Rooms

Bagian ini menjelaskan beberapa masalah umum yang mungkin timbul saat menggunakan AWS Clean Rooms dan cara memperbaikinya.

Masalah

- [Satu atau beberapa tabel yang direferensikan oleh kueri tidak dapat diakses oleh peran layanan terkait. table/role Pemilik harus memberikan akses peran layanan ke tabel.](#)
- [Salah satu kumpulan data yang mendasarinya memiliki format file yang tidak didukung.](#)
- [Hasil kueri tidak seperti yang diharapkan saat menggunakan Komputasi Kriptografi untuk Clean Rooms.](#)
- [AWS Clean Rooms Spark SQL: Data partisi tidak ada](#)

Satu atau beberapa tabel yang direferensikan oleh kueri tidak dapat diakses oleh peran layanan terkait. table/role Pemilik harus memberikan akses peran layanan ke tabel.

- Verifikasi bahwa izin untuk peran layanan disiapkan sesuai kebutuhan. Untuk informasi lebih lanjut, lihat [Menyiapkan AWS Clean Rooms](#).

Salah satu kumpulan data yang mendasarinya memiliki format file yang tidak didukung.

- Pastikan kumpulan data Anda berada dalam salah satu format file yang didukung:
 - Parquet
 - RCFile
 - TextFile
 - SequenceFile
 - RegexSerde
 - OpenCSV
 - AVRO

- JSON

Untuk informasi selengkapnya, lihat [Format data untuk AWS Clean Rooms](#).

Hasil kueri tidak seperti yang diharapkan saat menggunakan Komputasi Kriptografi untuk Clean Rooms.

Jika Anda menggunakan Cryptographic Computing for Clean Rooms (C3R), verifikasi bahwa kueri Anda menggunakan kolom terenkripsi dengan benar:

- `sealedKolom` hanya digunakan dalam SELECT klausa.
- `fingerprintKolom` hanya digunakan dalam JOIN klausa (dan GROUP BY klausa dalam kondisi tertentu).
- Bahwa Anda hanya JOINing fingerprint kolom dengan nama yang sama jika pengaturan kolaborasi memerlukannya.

Untuk informasi selengkapnya, lihat [the section called “Komputasi kriptografi”](#) dan [the section called “Jenis kolom”](#).

AWS Clean Rooms Spark SQL: Data partisi tidak ada

Semua partisi di dalam juga AWS Glue Data Catalog harus memiliki data di S3. Mesin menggunakan pengaturan Spark `spark.sql.files.ignoreMissingFiles=False`

Untuk informasi lebih lanjut, lihat <https://spark.apache.org/docs/latest/sql-data-sources-generic-options.html#ignore-missing-files>.

Jika Anda mengalami kesalahan ini, Anda akan mendapatkan pesan kesalahan berikut: "Missing partition data: One of the configured tables is partitioned and one or more of the partitions does not have data".

Bandingkan data Anda yang ada di Amazon S3 dengan partisi yang tercantum dalam tabel AWS Glue Data Catalog . Hapus partisi yang tidak memiliki data yang sesuai di S3.

Keamanan di AWS Clean Rooms

Keamanan cloud di AWS adalah prioritas tertinggi. Sebagai AWS pelanggan, Anda mendapat manfaat dari pusat data dan arsitektur jaringan yang dibangun untuk memenuhi persyaratan organisasi yang paling sensitif terhadap keamanan.

Keamanan adalah tanggung jawab bersama antara Anda AWS dan Anda. [Model tanggung jawab bersama](#) menjelaskan hal ini sebagai keamanan cloud dan keamanan dalam cloud:

- Keamanan cloud — AWS bertanggung jawab untuk melindungi infrastruktur yang menjalankan AWS layanan di AWS Cloud. AWS juga memberi Anda layanan yang dapat Anda gunakan dengan aman. Auditor pihak ketiga secara teratur menguji dan memverifikasi efektivitas keamanan kami sebagai bagian dari [Program AWS Kepatuhan](#) [Program AWS Kepatuhan](#) . Untuk mempelajari tentang program kepatuhan yang berlaku AWS Clean Rooms, lihat [AWS Services in Scope by Compliance Program](#) .
- Keamanan di cloud — Tanggung jawab Anda ditentukan oleh AWS layanan yang Anda gunakan. Anda juga bertanggung jawab atas faktor lain, yang mencakup kepekaan data Anda, persyaratan perusahaan, serta peraturan perundangan yang berlaku

Dokumentasi ini membantu Anda memahami cara menerapkan model tanggung jawab bersama saat menggunakan AWS Clean Rooms. Ini menunjukkan kepada Anda cara mengonfigurasi AWS Clean Rooms untuk memenuhi tujuan keamanan dan kepatuhan Anda. Anda juga belajar cara menggunakan AWS layanan lain yang membantu Anda memantau dan mengamankan AWS Clean Rooms sumber daya Anda.

Daftar Isi

- [Perlindungan data di AWS Clean Rooms](#)
- [Menggunakan peran terkait layanan untuk AWS Clean Rooms](#)
- [Retensi data di AWS Clean Rooms](#)
- [Praktik terbaik untuk kolaborasi data di AWS Clean Rooms](#)
- [Identity and Access Management untuk AWS Clean Rooms](#)
- [Validasi kepatuhan untuk AWS Clean Rooms](#)
- [Ketahanan di AWS Clean Rooms](#)
- [Keamanan infrastruktur di AWS Clean Rooms](#)

- [Access AWS Clean Rooms atau AWS Clean Rooms ML menggunakan endpoint antarmuka \(\)AWS PrivateLink](#)

Perlindungan data di AWS Clean Rooms

[Model tanggung jawab AWS bersama model](#) berlaku untuk perlindungan data di AWS Clean Rooms. Seperti yang dijelaskan dalam model AWS ini, bertanggung jawab untuk melindungi infrastruktur global yang menjalankan semua AWS Cloud. Anda bertanggung jawab untuk mempertahankan kendali atas konten yang di-host pada infrastruktur ini. Anda juga bertanggung jawab atas tugas-tugas konfigurasi dan manajemen keamanan untuk Layanan AWS yang Anda gunakan. Lihat informasi yang lebih lengkap tentang privasi data dalam [Pertanyaan Umum Privasi Data](#). Lihat informasi tentang perlindungan data di Eropa di pos blog [Model Tanggung Jawab Bersama dan GDPR AWS](#) di Blog Keamanan AWS .

Untuk tujuan perlindungan data, kami menyarankan Anda melindungi Akun AWS kredensial dan mengatur pengguna individu dengan AWS IAM Identity Center atau AWS Identity and Access Management (IAM). Dengan cara itu, setiap pengguna hanya diberi izin yang diperlukan untuk memenuhi tanggung jawab tugasnya. Kami juga menyarankan supaya Anda mengamankan data dengan cara-cara berikut:

- Gunakan autentikasi multi-faktor (MFA) pada setiap akun.
- Gunakan SSL/TLS untuk berkomunikasi dengan AWS sumber daya. Kami mensyaratkan TLS 1.2 dan menganjurkan TLS 1.3.
- Siapkan API dan logging aktivitas pengguna dengan AWS CloudTrail. Untuk informasi tentang penggunaan CloudTrail jejak untuk menangkap AWS aktivitas, lihat [Bekerja dengan CloudTrail jejak](#) di AWS CloudTrail Panduan Pengguna.
- Gunakan solusi AWS enkripsi, bersama dengan semua kontrol keamanan default di dalamnya Layanan AWS.
- Gunakan layanan keamanan terkelola tingkat lanjut seperti Amazon Macie, yang membantu menemukan dan mengamankan data sensitif yang disimpan di Amazon S3.
- Jika Anda memerlukan modul kriptografi tervalidasi FIPS 140-3 saat mengakses AWS melalui antarmuka baris perintah atau API, gunakan titik akhir FIPS. Lihat informasi selengkapnya tentang titik akhir FIPS yang tersedia di [Standar Pemrosesan Informasi Federal \(FIPS\) 140-3](#).

Kami sangat merekomendasikan agar Anda tidak pernah memasukkan informasi identifikasi yang sensitif, seperti nomor rekening pelanggan Anda, ke dalam tanda atau bidang isian bebas seperti

bidang Nama. Ini termasuk saat Anda bekerja dengan AWS Clean Rooms atau lainnya Layanan AWS menggunakan konsol, API AWS CLI, atau AWS SDKs. Data apa pun yang Anda masukkan ke dalam tanda atau bidang isian bebas yang digunakan untuk nama dapat digunakan untuk log penagihan atau log diagnostik. Saat Anda memberikan URL ke server eksternal, kami sangat menganjurkan supaya Anda tidak menyertakan informasi kredensial di dalam URL untuk memvalidasi permintaan Anda ke server itu.

Enkripsi saat diam

AWS Clean Rooms selalu mengenkripsi semua metadata layanan saat istirahat tanpa memerlukan konfigurasi tambahan apa pun. Enkripsi ini otomatis saat Anda menggunakannya AWS Clean Rooms.

Clean Rooms ML mengenkripsi semua data yang disimpan dalam layanan saat istirahat. AWS KMS Jika Anda memilih untuk memberikan kunci KMS Anda sendiri, konten model mirip Anda dan pekerjaan pembuatan segmen yang mirip dienkripsi saat istirahat dengan kunci KMS Anda.

Saat menggunakan model HTML AWS Clean Rooms kustom, layanan mengenkripsi semua data yang disimpan saat istirahat. AWS KMS AWS Clean Rooms mendukung penggunaan kunci terkelola pelanggan simetris yang Anda buat, miliki, dan kelola untuk mengenkripsi data saat istirahat. Jika kunci yang dikelola pelanggan tidak ditentukan, Kunci milik AWS digunakan secara default.

AWS Clean Rooms menggunakan hibah dan kebijakan utama untuk mengakses kunci yang dikelola pelanggan. Anda dapat mencabut akses ke hibah, atau menghapus akses layanan ke kunci yang dikelola pelanggan kapan saja. Jika Anda melakukannya, AWS Clean Rooms tidak akan dapat mengakses data apa pun yang dienkripsi oleh kunci yang dikelola pelanggan, yang memengaruhi operasi yang bergantung pada data tersebut. Misalnya, jika Anda mencoba membuat model terlatih dari saluran input ML terenkripsi yang tidak AWS Clean Rooms dapat diakses, maka operasi akan mengembalikan kesalahan. `ValidationException`

Note

Anda dapat menggunakan opsi enkripsi di Amazon S3 untuk melindungi data Anda saat istirahat.

Untuk informasi selengkapnya, lihat [Menentukan enkripsi Amazon S3](#) di Panduan Pengguna Amazon S3.

Saat menggunakan tabel pemetaan ID di dalamnya AWS Clean Rooms, layanan mengenkripsi semua data yang disimpan saat istirahat. AWS KMS Jika Anda memilih untuk memberikan kunci KMS Anda sendiri, isi tabel pemetaan ID Anda dienkripsi saat istirahat dengan kunci KMS Anda melalui. Resolusi Entitas AWS Untuk detail selengkapnya tentang izin yang diperlukan untuk bekerja dengan enkripsi dengan alur kerja pemetaan ID, lihat [Membuat peran pekerjaan alur kerja](#) di Panduan Pengguna. Resolusi Entitas AWS Resolusi Entitas AWS

Enkripsi saat bergerak

AWS Clean Rooms menggunakan Transport Layer Security (TLS) untuk enkripsi dalam perjalanan. Komunikasi dengan selalu AWS Clean Rooms dilakukan melalui HTTPS sehingga data Anda selalu dienkripsi saat transit, terlepas dari apakah itu disimpan di Amazon S3, Amazon Athena, atau Snowflake. Ini termasuk semua data dalam perjalanan saat menggunakan Clean Rooms ML.

Mengenkripsi data yang mendasarinya

Untuk informasi selengkapnya tentang cara mengenkripsi data dasar Anda, lihat [Komputasi Kriptografi untuk Clean Rooms](#).

Kebijakan kunci

Kebijakan utama mengontrol akses ke kunci yang dikelola pelanggan Anda. Setiap kunci yang dikelola pelanggan harus memiliki persis satu kebijakan utama, yang berisi pernyataan yang menentukan siapa yang dapat menggunakan kunci dan bagaimana mereka dapat menggunakannya. Saat membuat kunci terkelola pelanggan, Anda dapat menentukan kebijakan kunci. Untuk informasi selengkapnya, lihat Mengelola akses ke kunci terkelola pelanggan di Panduan AWS Key Management Service Pengembang.

Untuk menggunakan kunci terkelola pelanggan Anda dengan model AWS Clean Rooms Custom MS Anda, operasi API berikut harus diizinkan dalam kebijakan kunci:

- `kms:DescribeKey`— Memberikan detail kunci yang dikelola pelanggan untuk memungkinkan AWS Clean Rooms memvalidasi kunci.
- `kms:Decrypt`— Menyediakan akses AWS Clean Rooms untuk mendekripsi data terenkripsi dan menggunakannya dalam pekerjaan terkait.
- `kms:CreateGrant`- Clean Rooms ML mengenkripsi gambar pelatihan dan inferensi saat istirahat di Amazon ECR dengan membuat hibah untuk Amazon ECR. Untuk mempelajari lebih lanjut, lihat [Enkripsi saat Istirahat di Amazon ECR](#). Clean Rooms ML juga menggunakan Amazon SageMaker

AI untuk menjalankan pekerjaan pelatihan dan inferensi, dan menciptakan hibah bagi SageMaker AI untuk mengenkripsi volume Amazon EBS yang dilampirkan ke instans serta data keluaran di Amazon S3. Untuk mempelajari lebih lanjut, lihat [Melindungi Data saat Istirahat Menggunakan Enkripsi di Amazon SageMaker AI](#).

- kms:GenerateDataKey- Clean Rooms MS mengenkripsi data saat istirahat yang disimpan di Amazon S3 menggunakan enkripsi sisi server dengan. AWS KMS keys Untuk mempelajari lebih lanjut, lihat [Menggunakan enkripsi sisi server dengan AWS KMS keys \(SSE-KMS\) di Amazon S3](#).

Berikut ini adalah contoh pernyataan kebijakan yang dapat Anda tambahkan AWS Clean Rooms untuk sumber daya berikut:

Saluran input ML dengan data sintetis

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "Allow access to principals authorized to use AWS Clean Rooms ML",
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::666666666666:role/ExampleRole"
      },
      "Action": [
        "kms:GenerateDataKey",
        "kms:CreateGrant",
        "kms:Decrypt"
      ],
      "Resource": "*",
      "Condition": {
        "StringEquals": {
          "kms:ViaService": "cleanrooms-ml.us-east-1.amazonaws.com"
        },
        "ForAllValues:StringEquals": {
          "kms:GrantOperations": [
            "Decrypt",
            "Encrypt",
            "GenerateDataKeyWithoutPlaintext",
            "ReEncryptFrom",
            "ReEncryptTo",
            "CreateGrant",
            "DescribeKey",
```

```

        "RetireGrant",
        "GenerateDataKey"
    ]
},
    "BoolIfExists": {
        "kms:GrantIsForAWSResource": true
    }
},
{
    "Sid": "Allow describe key for principals authorized to use AWS Clean Rooms
ML",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::444455556666:role/ExampleRole"
    },
    "Action": [
        "kms:DescribeKey"
    ],
    "Resource": "*",
    "Condition": {
        "StringEquals": {
            "kms:ViaService": "cleanrooms-ml.us-east-1.amazonaws.com"
        }
    }
},
{
    "Sid": "Allow grant operations for AWS Clean Rooms ML service principal",
    "Effect": "Allow",
    "Principal": {
        "Service": "cleanrooms-ml.amazonaws.com"
    },
    "Action": [
        "kms:GenerateDataKey",
        "kms:CreateGrant",
        "kms:Decrypt"
    ],
    "Resource": "*",
    "Condition": {
        "ForAllValues:StringEquals": {
            "kms:GrantOperations": [
                "Decrypt",
                "Encrypt",
                "GenerateDataKeyWithoutPlaintext",

```



```

        "ReEncryptFrom",
        "ReEncryptTo",
        "CreateGrant",
        "DescribeKey",
        "RetireGrant",
        "GenerateDataKey"
    ]
}
},
{
    "Sid": "Allow describe key for AWS Clean Rooms ML service principal",
    "Effect": "Allow",
    "Principal": {
        "Service": "cleanrooms-ml.amazonaws.com"
    },
    "Action": [
        "kms:DescribeKey"
    ],
    "Resource": "*"
}
]
}

```

Saluran input ML tanpa data sintetis

JSON

```

{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "Allow access to principals authorized to use AWS Clean Rooms ML",
            "Effect": "Allow",
            "Principal": {
                "AWS": "arn:aws:iam::666666666666:role/ExampleRole"
            },
            "Action": [
                "kms:DescribeKey",
                "kms:GenerateDataKey",
                "kms:Decrypt"
            ],
            "Resource": "*"
        }
    ]
}

```

```

    "Condition": {
      "StringEquals": {
        "kms:ViaService": "cleanrooms-ml.us-east-1.amazonaws.com"
      }
    },
    {
      "Sid": "Allow access to AWS Clean Rooms ML service principal",
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms-ml.amazonaws.com"
      },
      "Action": [
        "kms:DescribeKey",
        "kms:GenerateDataKey",
        "kms:Decrypt"
      ],
      "Resource": "*"
    }
  ]
}

```

Pekerjaan model terlatih atau pekerjaan inferensi model terlatih

JSON

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "Allow grant operations for principals authorized to use AWS  
Clean Rooms ML",
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::666666666666:role/ExampleRole"
      },
      "Action": [
        "kms:GenerateDataKey",
        "kms:CreateGrant",
        "kms:Decrypt"
      ],
      "Resource": "*"
    }
  ]
}

```

```

        "Condition": {
            "StringEquals": {
                "kms:ViaService": "cleanrooms-ml.us-east-1.amazonaws.com"
            },
            "ForAllValues:StringEquals": {
                "kms:GrantOperations": [
                    "Decrypt",
                    "Encrypt",
                    "GenerateDataKeyWithoutPlaintext",
                    "ReEncryptFrom",
                    "ReEncryptTo",
                    "CreateGrant",
                    "DescribeKey",
                    "RetireGrant",
                    "GenerateDataKey"
                ]
            },
            "BoolIfExists": {
                "kms:GrantIsForAWSResource": true
            }
        },
        {
            "Sid": "Allow describe key for principals authorized to use AWS Clean
Rooms ML",
            "Effect": "Allow",
            "Principal": {
                "AWS": "arn:aws:iam::444455556666:role/ExampleRole"
            },
            "Action": [
                "kms:DescribeKey"
            ],
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "kms:ViaService": "cleanrooms-ml.us-east-1.amazonaws.com"
                }
            }
        },
        {
            "Sid": "Allow grant operations for AWS Clean Rooms ML service
principal",
            "Effect": "Allow",
            "Principal": {

```

```

        "Service": "cleanrooms-ml.amazonaws.com"
    },
    "Action": [
        "kms:GenerateDataKey",
        "kms:CreateGrant",
        "kms:Decrypt"
    ],
    "Resource": "*",
    "Condition": {
        "ForAllValues:StringEquals": {
            "kms:GrantOperations": [
                "Decrypt",
                "Encrypt",
                "GenerateDataKeyWithoutPlaintext",
                "ReEncryptFrom",
                "ReEncryptTo",
                "CreateGrant",
                "DescribeKey",
                "RetireGrant",
                "GenerateDataKey"
            ]
        }
    }
},
{
    "Sid": "Allow describe key for AWS Clean Rooms ML service principal",
    "Effect": "Allow",
    "Principal": {
        "Service": "cleanrooms-ml.amazonaws.com"
    },
    "Action": [
        "kms:DescribeKey"
    ],
    "Resource": "*"
}
]
}

```

Clean Rooms ML tidak mendukung penetapan konteks enkripsi layanan atau konteks sumber dalam kebijakan kunci yang dikelola pelanggan. Konteks enkripsi yang digunakan oleh layanan secara internal dapat dilihat oleh pelanggan di CloudTrail.

Menggunakan peran terkait layanan untuk AWS Clean Rooms

AWS Clean Rooms menggunakan AWS Identity and Access Management peran [terkait layanan](#) (IAM). Peran terkait layanan adalah jenis unik peran IAM yang ditautkan langsung ke. AWS Clean Rooms Peran terkait layanan telah ditentukan sebelumnya oleh AWS Clean Rooms dan mencakup semua izin yang diperlukan layanan untuk memanggil AWS layanan lain atas nama Anda.

Peran terkait layanan membuat pengaturan AWS Clean Rooms lebih mudah karena Anda tidak perlu menambahkan izin yang diperlukan secara manual. AWS Clean Rooms mendefinisikan izin peran terkait layanan, dan kecuali ditentukan lain, hanya AWS Clean Rooms dapat mengambil perannya. Izin yang ditentukan mencakup kebijakan kepercayaan dan kebijakan izin, dan kebijakan izin tersebut tidak dapat dilampirkan ke entitas IAM lainnya.

Anda dapat menghapus peran tertaut layanan hanya setelah menghapus sumber daya terkait terlebih dahulu. Ini melindungi AWS Clean Rooms sumber daya Anda karena Anda tidak dapat secara tidak sengaja menghapus izin untuk mengakses sumber daya.

Untuk informasi tentang layanan lain yang mendukung peran terkait layanan, silakan lihat [layanan AWS yang bisa digunakan dengan IAM](#) dan carilah layanan yang memiliki opsi Ya di kolom Peran terkait layanan. Pilih Ya dengan sebuah tautan untuk melihat dokumentasi peran terkait layanan untuk layanan tersebut.

Izin peran terkait layanan untuk AWS Clean Rooms

AWS Clean Rooms menggunakan peran terkait layanan bernama `AWSServiceRoleForAWSCleanRooms` — untuk mempublikasikan CloudWatch metrik terkait Kamar Bersih ke akun Anda. AWS

Peran terkait layanan `AWSService RoleFor AWSClean Rooms` mempercayai layanan berikut untuk mengambil peran:

- `cleanrooms.amazonaws.com`

Kebijakan izin peran bernama `AWSClean RoomsServiceRolePolicy` memungkinkan AWS Clean Rooms untuk menyelesaikan tindakan berikut pada sumber daya yang ditentukan:

- Tindakan: `cloudwatch:PutMetricData` pada `all AWS resources`, `restricted to the AWS Clean Rooms namespace`

Anda harus mengonfigurasi izin agar pengguna, grup, atau peran Anda membuat, mengedit, atau menghapus peran terkait layanan. Untuk informasi selengkapnya, lihat [Izin peran terkait layanan](#) dalam Panduan Pengguna IAM.

Membuat peran terkait layanan untuk AWS Clean Rooms

Anda dapat menggunakan konsol IAM untuk membuat peran terkait layanan dengan kasus penggunaan `AWSServiceRoleForAWSCleanKamar`. Di AWS CLI atau AWS API, buat peran terkait layanan dengan nama `cleanrooms.amazonaws.com` layanan. Untuk informasi selengkapnya, lihat [Membuat peran terkait layanan](#) dalam Panduan Pengguna IAM. Jika Anda menghapus peran terkait layanan ini, Anda dapat mengulang proses yang sama untuk membuat peran tersebut lagi.

Mengedit peran terkait layanan untuk AWS Clean Rooms

AWS Clean Rooms tidak mengizinkan Anda mengedit peran terkait layanan `AWSServiceRoleForAWSCleanKamar`. Setelah membuat peran terkait layanan, Anda tidak dapat mengubah nama peran karena berbagai entitas mungkin merujuk peran tersebut. Namun, Anda dapat mengedit penjelasan peran menggunakan IAM. Untuk informasi selengkapnya, lihat [Mengedit peran terkait layanan](#) dalam Panduan Pengguna IAM.

Menghapus peran terkait layanan untuk AWS Clean Rooms

Jika Anda tidak perlu lagi menggunakan fitur atau layanan yang memerlukan peran terkait layanan, sebaiknya hapus peran tersebut. Dengan begitu, Anda tidak memiliki entitas yang tidak digunakan yang tidak dipantau atau dipelihara secara aktif. Tetapi, Anda harus membersihkan sumber daya peran yang terhubung dengan layanan sebelum menghapusnya secara manual.

Untuk menghapus `AWSServiceRoleForAWSCleanKamar`, Anda harus terlebih dahulu menghapus semua [kolaborasi](#) dan [keanggotaan](#) di Anda. Akun AWS

Note

Jika AWS Clean Rooms layanan menggunakan peran saat Anda mencoba menghapus sumber daya, maka penghapusan mungkin gagal. Jika hal itu terjadi, tunggu beberapa menit dan coba mengoperasikannya lagi.

Untuk menghapus peran terkait layanan secara manual menggunakan IAM

Gunakan konsol IAM, the AWS CLI, atau AWS API untuk menghapus peran terkait layanan AWSService RoleFor AWSClean Rooms. Untuk informasi selengkapnya, lihat [Menghapus peran terkait layanan](#) dalam Panduan Pengguna IAM.

Wilayah yang Didukung untuk AWS Clean Rooms peran terkait layanan

AWS Clean Rooms mendukung penggunaan peran terkait layanan di semua Wilayah tempat layanan tersedia. Untuk informasi selengkapnya, lihat [AWS Wilayah dan titik akhir](#).

Retensi data di AWS Clean Rooms

Setiap data yang sementara dibaca ke dalam AWS Clean Rooms kolaborasi akan dihapus setelah kueri selesai.

Saat Anda membuat model yang mirip, Clean Rooms ML akan membaca data pelatihan Anda, mengubahnya menjadi format yang sesuai untuk model ML kami, dan menyimpan parameter model terlatih di dalam Clean Rooms. Clean Rooms ML tidak menyimpan salinan data pelatihan Anda. AWS Clean Rooms Kueri SQL tidak menyimpan data Anda setelah kueri berjalan. Clean Rooms MS kemudian menggunakan model terlatih untuk meringkas perilaku semua pengguna Anda. Clean Rooms ML menyimpan kumpulan data tingkat pengguna untuk setiap pengguna dalam data Anda selama model mirip Anda aktif.

Saat Anda memulai pekerjaan pembuatan segmen yang mirip, Clean Rooms ML akan membaca data seed, membaca ringkasan perilaku dari model mirip terkait, dan membuat segmen mirip yang disimpan dalam layanan. AWS Clean Rooms Clean Rooms ML tidak menyimpan salinan data benih Anda. Clean Rooms ML menyimpan output tingkat pengguna dari pekerjaan selama pekerjaan itu aktif.

Jika data benih Anda berasal dari kueri SQL, output kueri tersebut hanya disimpan dalam layanan selama durasi pekerjaan. Hasil kueri dienkripsi saat istirahat dan dalam perjalanan.

Jika Anda ingin menghapus model yang mirip atau data pekerjaan pembuatan segmen yang mirip, gunakan API untuk menghapusnya. Clean Rooms MS secara asinkron menghapus semua data yang terkait dengan model atau pekerjaan. Setelah proses ini selesai, Clean Rooms ML menghapus metadata untuk model atau pekerjaan dan tidak lagi terlihat di API. Clean Rooms MS menyimpan data yang dihapus selama 3 hari untuk pencegahan pemulihan bencana. Setelah pekerjaan atau model tidak lagi terlihat di API dan 3 hari berlalu, semua data yang terkait dengan model atau pekerjaan telah dihapus secara permanen.

Praktik terbaik untuk kolaborasi data di AWS Clean Rooms

Topik ini menjelaskan praktik terbaik untuk melakukan kolaborasi data di AWS Clean Rooms.

AWS Clean Rooms mengikuti [Model Tanggung Jawab AWS Bersama](#). AWS Clean Rooms menawarkan [aturan analisis](#) yang dapat Anda konfigurasi untuk memperkuat kemampuan Anda untuk melindungi data sensitif dalam kolaborasi. Aturan analisis yang Anda konfigurasikan AWS Clean Rooms akan memberlakukan pembatasan (kontrol kueri dan kontrol keluaran kueri) yang telah Anda konfigurasikan. Anda bertanggung jawab untuk menentukan batasan dan mengonfigurasi aturan analisis yang sesuai.

Kolaborasi data mungkin melibatkan lebih dari sekedar penggunaan AWS Clean Rooms Anda. Untuk membantu Anda memaksimalkan manfaat kolaborasi data, kami menyarankan Anda melakukan praktik terbaik berikut dengan penggunaan Anda AWS Clean Rooms dan secara khusus dengan aturan analisis.

Topik

- [Praktik terbaik dengan AWS Clean Rooms](#)
- [Praktik terbaik untuk menggunakan aturan analisis di AWS Clean Rooms](#)

Praktik terbaik dengan AWS Clean Rooms

Anda bertanggung jawab untuk menilai risiko setiap kolaborasi data dan membandingkannya dengan persyaratan privasi Anda seperti program dan kebijakan kepatuhan eksternal dan internal. Kami menyarankan Anda mengambil tindakan tambahan dengan penggunaan Anda AWS Clean Rooms. Tindakan ini dapat membantu mengelola risiko lebih lanjut dan membantu mencegah upaya pihak ketiga untuk mengidentifikasi kembali data Anda (misalnya, serangan yang berbeda atau serangan saluran samping).

Misalnya, pertimbangkan untuk melakukan uji tuntas pada kolaborator Anda yang lain dan buat perjanjian hukum dengan mereka sebelum terlibat dalam kolaborasi. Untuk memantau penggunaan data Anda, pertimbangkan juga untuk mengadopsi mekanisme audit lain dengan penggunaan AWS Clean Rooms Anda.

Praktik terbaik untuk menggunakan aturan analisis di AWS Clean Rooms

Aturan analisis AWS Clean Rooms memungkinkan Anda membatasi kueri yang dapat dijalankan dengan menyetel kontrol kueri pada tabel yang dikonfigurasi. Misalnya, Anda dapat mengatur

kontrol kueri untuk bagaimana tabel yang dikonfigurasi dapat digabungkan dan kolom mana yang dapat dipilih. Anda juga dapat membatasi output kueri melalui pengaturan kontrol hasil kueri seperti ambang agregasi pada baris keluaran. Layanan menolak kueri apa pun dan menghapus baris yang tidak sesuai dengan aturan analisis yang ditetapkan oleh anggota pada tabel yang dikonfigurasi dalam kueri.

Kami merekomendasikan 10 praktik terbaik berikut untuk menggunakan aturan analisis pada tabel yang dikonfigurasi:

- Buat tabel terkonfigurasi terpisah untuk kasus penggunaan kueri terpisah (misalnya, perencanaan audiens atau atribusi). Anda dapat membuat beberapa tabel yang dikonfigurasi dengan AWS Glue tabel dasar yang sama.
- Tentukan kolom dalam aturan analisis (misalnya, kolom dimensi, kolom daftar, kolom gabungan) yang diperlukan untuk kueri dalam kolaborasi. Ini dapat membantu mengurangi risiko serangan yang berbeda atau memungkinkan anggota lain untuk merekayasa balik data Anda. Gunakan fitur kolom allowlist untuk mencatat kolom lain yang mungkin ingin Anda jadikan queryable di masa mendatang. Untuk menyesuaikan kolom yang dapat digunakan untuk kolaborasi tertentu, buat tabel tambahan yang dikonfigurasi dengan AWS Glue tabel dasar yang sama.
- Tentukan fungsi dalam aturan analisis yang diperlukan untuk analisis dalam kolaborasi. Ini dapat membantu mengurangi risiko dari kesalahan fungsi langka yang dapat menyajikan informasi pada titik data individu. Untuk menyesuaikan fungsi yang dapat digunakan untuk kolaborasi tertentu, buat tabel tambahan yang dikonfigurasi dengan AWS Glue tabel dasar yang sama.
- Tambahkan batasan agregasi pada kolom mana pun yang nilainya pada tingkat baris sensitif. Ini termasuk kolom dalam tabel yang dikonfigurasi yang juga ada di tabel anggota kolaborasi lainnya dan aturan analisis sebagai kendala agregasi. Ini juga mencakup kolom dalam tabel yang dikonfigurasi yang tidak dapat dikueri, yaitu kolom yang ada di tabel yang dikonfigurasi tetapi tidak ada dalam aturan analisis. Kendala agregasi dapat membantu mengurangi risiko dari mengkorelasikan hasil kueri dengan data di luar kolaborasi.
- Buat kolaborasi pengujian dan aturan analisis untuk menguji batasan yang dibuat dengan aturan analisis yang ditentukan.
- Tinjau tabel yang dikonfigurasi kolaborator dan aturan analisis anggota pada tabel yang dikonfigurasi untuk memeriksa apakah mereka cocok dengan apa yang disepakati untuk kolaborasi. Ini dapat membantu mengurangi risiko dari anggota lain yang merekayasa data mereka sendiri untuk menjalankan kueri yang tidak disepakati.
- Tinjau contoh kueri yang disediakan (khusus konsol) yang diaktifkan pada tabel yang dikonfigurasi setelah Anda mengatur aturan analisis.

 Note

Selain kueri contoh yang disediakan, kueri lain dimungkinkan berdasarkan aturan analisis dan tabel anggota kolaborasi lainnya serta aturan analisis.

- Anda dapat menambahkan atau memperbarui aturan analisis untuk tabel yang dikonfigurasi dalam kolaborasi. Ketika Anda melakukannya, tinjau semua kolaborasi di mana tabel yang dikonfigurasi dikaitkan dan dampaknya. Ini membantu memastikan bahwa tidak ada kolaborasi yang menggunakan aturan analisis usang.
- Tinjau kueri yang dijalankan dalam kolaborasi untuk memeriksa apakah kueri cocok dengan kasus penggunaan atau kueri yang disepakati untuk kolaborasi. (Kueri tersedia di log kueri saat fitur Pencatatan kueri diaktifkan.) Ini dapat membantu mengurangi risiko dari anggota yang menjalankan analisis yang tidak disepakati dan potensi serangan seperti serangan saluran samping.
- Tinjau kolom tabel yang dikonfigurasi yang digunakan dalam aturan analisis anggota kolaborasi dan dalam kueri untuk memeriksa apakah mereka cocok dengan apa yang disepakati dalam kolaborasi. (Kueri tersedia di log kueri saat fitur itu diaktifkan.) Ini dapat membantu mengurangi risiko dari anggota lain yang merekayasa data mereka sendiri untuk melakukan pertanyaan yang tidak disepakati.

Identity and Access Management untuk AWS Clean Rooms

AWS Identity and Access Management (IAM) adalah Layanan AWS yang membantu administrator mengontrol akses ke AWS sumber daya dengan aman. Administrator IAM mengontrol siapa yang dapat diautentikasi (masuk) dan diberi wewenang (memiliki izin) untuk menggunakan sumber daya. AWS Clean Rooms IAM adalah Layanan AWS yang dapat Anda gunakan tanpa biaya tambahan.

Topik

- [Audiens](#)
- [Mengautentikasi dengan identitas](#)
- [Mengelola akses menggunakan kebijakan](#)
- [Bagaimana AWS Clean Rooms bekerja dengan IAM](#)
- [Contoh kebijakan berbasis identitas untuk AWS Clean Rooms](#)
- [AWS kebijakan terkelola untuk AWS Clean Rooms](#)

- [Memecahkan masalah AWS Clean Rooms identitas dan akses](#)
- [Pencegahan "confused deputy" lintas layanan](#)
- [Perilaku IAM untuk AWS Clean Rooms MS](#)
- [Perilaku IAM untuk Kamar Bersih Model Kustom](#)

Audiens

Cara Anda menggunakan AWS Identity and Access Management (IAM) berbeda berdasarkan peran Anda:

- Pengguna layanan - minta izin dari administrator Anda jika Anda tidak dapat mengakses fitur (lihat [Memecahkan masalah AWS Clean Rooms identitas dan akses](#))
- Administrator layanan - tentukan akses pengguna dan mengirimkan permintaan izin (lihat [Bagaimana AWS Clean Rooms bekerja dengan IAM](#))
- Administrator IAM - tulis kebijakan untuk mengelola akses (lihat [Contoh kebijakan berbasis identitas untuk AWS Clean Rooms](#))

Mengautentikasi dengan identitas

Otentikasi adalah cara Anda masuk AWS menggunakan kredensial identitas Anda. Anda harus diautentikasi (masuk ke AWS) sebagai Pengguna root akun AWS, sebagai pengguna IAM, atau dengan mengasumsikan peran IAM.

Anda dapat masuk AWS sebagai identitas federasi dengan menggunakan kredensial yang disediakan melalui sumber identitas. AWS IAM Identity Center Pengguna (IAM Identity Center) atau autentikasi masuk tunggal perusahaan Anda adalah contoh identitas federasi. Saat Anda masuk sebagai identitas terfederasi, administrator Anda sebelumnya menyiapkan federasi identitas menggunakan peran IAM. Ketika Anda mengakses AWS dengan menggunakan federasi, Anda secara tidak langsung mengambil peran.

Bergantung pada jenis pengguna Anda, Anda dapat masuk ke Konsol Manajemen AWS atau portal AWS akses. Untuk informasi selengkapnya tentang masuk AWS, lihat [Cara masuk ke Panduan AWS Sign-In Pengguna Anda Akun AWS](#).

Jika Anda mengakses AWS secara terprogram, AWS sediakan kit pengembangan perangkat lunak (SDK) dan antarmuka baris perintah (CLI) untuk menandatangani permintaan Anda secara

kriptografis menggunakan kredensial Anda. Jika Anda tidak menggunakan AWS alat, Anda harus menandatangani permintaan sendiri. Untuk informasi selengkapnya tentang menggunakan metode yang disarankan untuk menandatangani permintaan sendiri, lihat [proses penandatanganan Versi Tanda Tangan 4](#) di Referensi Umum AWS.

Apa pun metode autentikasi yang digunakan, Anda mungkin diminta untuk menyediakan informasi keamanan tambahan. Misalnya, AWS merekomendasikan agar Anda menggunakan otentikasi multi-faktor (MFA) untuk meningkatkan keamanan akun Anda. Untuk mempelajari selengkapnya, lihat [Autentikasi multi-faktor](#) dalam Panduan Pengguna AWS IAM Identity Center dan [Menggunakan autentikasi multi-faktor \(MFA\) dalam AWS](#) dalam Panduan Pengguna IAM.

Akun AWS pengguna root

Saat Anda membuat Akun AWS, Anda mulai dengan satu identitas masuk yang memiliki akses lengkap ke semua Layanan AWS dan sumber daya di akun. Identitas ini disebut pengguna root Akun AWS dan diakses dengan cara masuk menggunakan alamat email dan kata sandi yang Anda gunakan untuk membuat akun. Kami sangat menyarankan agar Anda tidak menggunakan pengguna root untuk tugas sehari-hari Anda. Lindungi kredensial pengguna root Anda dan gunakan kredensial tersebut untuk melakukan tugas yang hanya dapat dilakukan pengguna root. Untuk daftar lengkap tugas yang mengharuskan Anda masuk sebagai pengguna root, lihat [Pengguna root akun AWS kredensi dan identitas IAM](#) di Referensi Umum AWS

Identitas terfederasi

Sebagai praktik terbaik, mewajibkan pengguna manusia untuk menggunakan federasi dengan penyedia identitas untuk mengakses Layanan AWS menggunakan kredensial sementara.

Identitas federasi adalah pengguna dari direktori perusahaan Anda, penyedia identitas web, atau Directory Service yang mengakses Layanan AWS menggunakan kredensial dari sumber identitas. Identitas terfederasi mengambil peran yang memberikan kredensial sementara.

Untuk manajemen akses terpusat, kami menyarankan AWS IAM Identity Center. Untuk informasi selengkapnya, lihat [Apa itu Pusat Identitas IAM?](#) dalam Panduan Pengguna AWS IAM Identity Center

Pengguna dan grup IAM

[Pengguna IAM](#) adalah identitas dengan izin khusus untuk satu orang atau aplikasi. Sebaiknya gunakan kredensial sementara alih-alih pengguna IAM dengan kredensial jangka panjang. Untuk

informasi selengkapnya, lihat [Mewajibkan pengguna manusia untuk menggunakan federasi dengan penyedia identitas untuk mengakses AWS menggunakan kredensi sementara](#) di Panduan Pengguna IAM.

[Grup IAM](#) menentukan kumpulan pengguna IAM dan mempermudah pengelolaan izin untuk pengguna dalam jumlah besar. Untuk mempelajari selengkapnya, lihat [Kasus penggunaan untuk pengguna IAM](#) dalam Panduan Pengguna IAM.

Peran IAM

[Peran IAM](#) adalah identitas dengan izin khusus yang menyediakan kredensial sementara. Anda dapat mengambil peran dengan [beralih dari pengguna ke peran IAM \(konsol\)](#) atau dengan memanggil operasi AWS CLI atau AWS API. Untuk informasi selengkapnya, lihat [Metode untuk mengambil peran](#) dalam Panduan Pengguna IAM.

Peran IAM berguna untuk akses pengguna terfederasi, izin pengguna IAM sementara, akses lintas akun, akses lintas layanan, dan aplikasi yang berjalan di Amazon EC2. Untuk informasi selengkapnya, lihat [Akses sumber daya lintas akun di IAM](#) dalam Panduan Pengguna IAM.

Mengelola akses menggunakan kebijakan

Anda mengontrol akses AWS dengan membuat kebijakan dan melampirkannya ke AWS identitas atau sumber daya. Kebijakan adalah objek AWS yang, ketika dikaitkan dengan identitas atau sumber daya, menentukan izinnya. AWS mengevaluasi kebijakan ini ketika prinsipal (pengguna, pengguna root, atau sesi peran) membuat permintaan. Izin dalam kebijakan menentukan apakah permintaan diizinkan atau ditolak. Sebagian besar kebijakan disimpan AWS sebagai dokumen JSON. Untuk informasi selengkapnya tentang struktur dan isi dokumen kebijakan JSON, lihat [Gambaran umum kebijakan JSON](#) dalam Panduan Pengguna IAM.

Administrator dapat menggunakan kebijakan AWS JSON untuk menentukan siapa yang memiliki akses ke apa. Yaitu, principal dapat melakukan tindakan pada suatu sumber daya, dan dalam suatu syarat.

Setiap entitas IAM (pengguna atau peran) dimulai tanpa izin. Secara default, pengguna tidak dapat melakukan apa pun, bahkan tidak mengubah kata sandi mereka sendiri. Untuk memberikan izin kepada pengguna untuk melakukan sesuatu, administrator harus melampirkan kebijakan izin kepada pengguna. Atau administrator dapat menambahkan pengguna ke grup yang memiliki izin yang dimaksudkan. Ketika administrator memberikan izin untuk grup, semua pengguna dalam grup tersebut akan diberi izin tersebut.

Kebijakan IAM mendefinisikan izin untuk suatu tindakan terlepas dari metode yang Anda gunakan untuk melakukan operasinya. Misalnya, anggaplah Anda memiliki kebijakan yang mengizinkan tindakan `iam:GetRole`. Pengguna dengan kebijakan tersebut bisa mendapatkan informasi peran dari Konsol Manajemen AWS, API AWS CLI, atau AWS API.

Kebijakan berbasis identitas

Kebijakan berbasis identitas adalah dokumen kebijakan izin JSON yang dapat Anda lampirkan ke sebuah identitas, seperti pengguna IAM, grup pengguna IAM, atau peran IAM. Kebijakan ini mengontrol jenis tindakan yang dapat dilakukan oleh pengguna dan peran, di sumber daya mana, dan berdasarkan kondisi seperti apa. Untuk mempelajari cara membuat kebijakan berbasis identitas, lihat [Tentukan izin IAM kustom dengan kebijakan terkelola pelanggan](#) dalam Panduan Pengguna IAM.

Kebijakan berbasis identitas dapat dikategorikan lebih lanjut sebagai kebijakan inline atau kebijakan yang dikelola. Kebijakan inline disematkan langsung ke satu pengguna, grup, atau peran. Kebijakan yang dikelola adalah kebijakan mandiri yang dapat Anda lampirkan ke beberapa pengguna, grup, dan peran di Akun AWS Anda. Kebijakan AWS terkelola mencakup kebijakan terkelola dan kebijakan yang dikelola pelanggan. Untuk mempelajari cara memilih antara kebijakan yang dikelola atau kebijakan inline, lihat [Memilih antara kebijakan yang dikelola dan kebijakan inline](#) dalam Panduan Pengguna IAM.

Kebijakan berbasis sumber daya

Kebijakan berbasis sumber daya adalah dokumen kebijakan JSON yang Anda lampirkan ke sumber daya. Contoh kebijakan berbasis sumber daya adalah kebijakan kepercayaan peran IAM dan kebijakan bucket Amazon S3. Dalam layanan yang mendukung kebijakan berbasis sumber daya, administrator layanan dapat menggunakannya untuk mengontrol akses ke sumber daya tertentu. Untuk sumber daya tempat kebijakan dilampirkan, kebijakan menentukan tindakan apa yang dapat dilakukan oleh principal tertentu pada sumber daya tersebut dan dalam kondisi apa. Anda harus [menentukan principal](#) dalam kebijakan berbasis sumber daya. Prinsipal dapat mencakup akun, pengguna, peran, pengguna federasi, atau. Layanan AWS

Kebijakan berbasis sumber daya merupakan kebijakan inline yang terletak di layanan tersebut. Anda tidak dapat menggunakan kebijakan AWS terkelola dari IAM dalam kebijakan berbasis sumber daya.

Jenis-jenis kebijakan lain

AWS mendukung jenis kebijakan tambahan yang kurang umum. Jenis-jenis kebijakan ini dapat mengatur izin maksimum yang diberikan kepada Anda oleh jenis kebijakan yang lebih umum.

- **Batasan izin** – Batasan izin adalah fitur lanjutan tempat Anda mengatur izin maksimum yang dapat diberikan oleh kebijakan berbasis identitas ke entitas IAM (pengguna IAM atau peran IAM). Anda dapat menetapkan batas izin untuk suatu entitas. Izin yang dihasilkan adalah persimpangan antara kebijakan berbasis identitas milik entitas dan batas izinnya. Kebijakan berbasis sumber daya yang menentukan pengguna atau peran dalam bidang `Principal` tidak dibatasi oleh batasan izin. Penolakan eksplisit dalam salah satu kebijakan ini akan menggantikan pemberian izin. Untuk informasi selengkapnya tentang batasan izin, lihat [Batasan izin untuk entitas IAM](#) dalam Panduan Pengguna IAM.
- **Kebijakan kontrol layanan (SCPs)** — SCPs adalah kebijakan JSON yang menentukan izin maksimum untuk organisasi atau unit organisasi (OU) di AWS Organizations. AWS Organizations adalah layanan untuk mengelompokkan dan mengelola secara terpusat beberapa Akun AWS yang dimiliki bisnis Anda. Jika Anda mengaktifkan semua fitur dalam suatu organisasi, maka Anda dapat menerapkan kebijakan kontrol layanan (SCPs) ke salah satu atau semua akun Anda. SCP membatasi izin untuk entitas di akun anggota, termasuk masing-masing. Pengguna root akun AWS. Untuk informasi selengkapnya tentang Organizations dan SCPs, lihat [Cara SCPs kerja](#) di Panduan AWS Organizations Pengguna.
- **Kebijakan sesi** – Kebijakan sesi adalah kebijakan lanjutan yang Anda berikan sebagai parameter ketika Anda membuat sesi sementara secara programatis untuk peran atau pengguna terfederasi. Izin sesi yang dihasilkan adalah perpotongan antara kebijakan berbasis identitas pengguna atau peran dan kebijakan sesi. Izin juga bisa datang dari kebijakan berbasis sumber daya. Penolakan eksplisit dalam salah satu kebijakan ini akan menggantikan pemberian izin. Untuk informasi selengkapnya, lihat [Kebijakan sesi](#) dalam Panduan Pengguna IAM.

Berbagai jenis kebijakan

Ketika beberapa jenis kebijakan berlaku pada suatu permintaan, izin yang dihasilkan lebih rumit untuk dipahami. Untuk mempelajari cara AWS menentukan apakah akan mengizinkan permintaan saat beberapa jenis kebijakan terlibat, lihat [Logika evaluasi kebijakan](#) di Panduan Pengguna IAM.

Bagaimana AWS Clean Rooms bekerja dengan IAM

Sebelum Anda menggunakan IAM untuk mengelola akses AWS Clean Rooms, pelajari fitur IAM yang tersedia untuk digunakan. AWS Clean Rooms

Fitur IAM yang dapat Anda gunakan AWS Clean Rooms

Fitur IAM	AWS Clean Rooms dukungan
Kebijakan berbasis identitas	Ya
Kebijakan berbasis sumber daya	Parsial
Tindakan kebijakan	Ya
Sumber daya kebijakan	Ya
kunci-kunci persyaratan kebijakan (spesifik layanan)	Sebagian
ACLs	Tidak
ABAC (tanda dalam kebijakan)	Ya
Kredensial sementara	Ya
Sesi akses teruskan (FAS)	Ya
Peran layanan	Ya
Peran terkait layanan	Tidak

Untuk mendapatkan tampilan tingkat tinggi tentang cara AWS Clean Rooms dan Layanan AWS pekerjaan lainnya dengan sebagian besar fitur IAM, lihat [Layanan AWS yang berfungsi dengan IAM di Panduan Pengguna IAM](#).

Kebijakan berbasis identitas untuk AWS Clean Rooms

Mendukung kebijakan berbasis identitas: Ya

Kebijakan berbasis identitas adalah dokumen kebijakan izin JSON yang dapat Anda lampirkan ke sebuah identitas, seperti pengguna IAM, grup pengguna IAM, atau peran IAM. Kebijakan ini mengontrol jenis tindakan yang dapat dilakukan oleh pengguna dan peran, di sumber daya mana, dan berdasarkan kondisi seperti apa. Untuk mempelajari cara membuat kebijakan berbasis identitas, lihat [Tentukan izin IAM kustom dengan kebijakan terkelola pelanggan](#) dalam Panduan Pengguna IAM.

Dengan kebijakan berbasis identitas IAM, Anda dapat menentukan secara spesifik apakah tindakan dan sumber daya diizinkan atau ditolak, serta kondisi yang menjadi dasar dikabulkan atau ditolaknya tindakan tersebut. Untuk mempelajari semua elemen yang dapat Anda gunakan dalam kebijakan JSON, lihat [Referensi elemen kebijakan JSON IAM](#) dalam Panduan Pengguna IAM.

Contoh kebijakan berbasis identitas untuk AWS Clean Rooms

Untuk melihat contoh kebijakan AWS Clean Rooms berbasis identitas, lihat. [Contoh kebijakan berbasis identitas untuk AWS Clean Rooms](#)

Kebijakan berbasis sumber daya dalam AWS Clean Rooms

Mendukung kebijakan berbasis sumber daya: Sebagian

Kebijakan berbasis sumber daya adalah dokumen kebijakan JSON yang Anda lampirkan ke sumber daya. Contoh kebijakan berbasis sumber daya adalah kebijakan kepercayaan peran IAM dan kebijakan bucket Amazon S3. Dalam layanan yang mendukung kebijakan berbasis sumber daya, administrator layanan dapat menggunakannya untuk mengontrol akses ke sumber daya tertentu. Untuk sumber daya tempat kebijakan dilampirkan, kebijakan menentukan tindakan apa yang dapat dilakukan oleh principal tertentu pada sumber daya tersebut dan dalam kondisi apa. Anda harus [menentukan principal](#) dalam kebijakan berbasis sumber daya. Prinsipal dapat mencakup akun, pengguna, peran, pengguna federasi, atau. Layanan AWS

Untuk mengaktifkan akses lintas akun, Anda dapat menentukan secara spesifik seluruh akun atau entitas IAM di akun lain sebagai principal dalam kebijakan berbasis sumber daya. Untuk informasi selengkapnya, lihat [Akses sumber daya lintas akun di IAM](#) dalam Panduan Pengguna IAM.

AWS Clean Rooms Layanan ini hanya mendukung satu jenis kebijakan berbasis sumber daya yang disebut kebijakan sumber daya terkelola model mirip mirip yang dikonfigurasi, yang dilampirkan ke model mirip yang dikonfigurasi. Kebijakan ini menentukan prinsipal mana yang dapat melakukan tindakan pada model mirip yang dikonfigurasi.

Untuk mempelajari cara melampirkan kebijakan berbasis sumber daya ke model mirip yang dikonfigurasi, lihat. [Perilaku IAM untuk AWS Clean Rooms MS](#)

Tindakan kebijakan untuk AWS Clean Rooms

Mendukung tindakan kebijakan: Ya

Administrator dapat menggunakan kebijakan AWS JSON untuk menentukan siapa yang memiliki akses ke apa. Yaitu, di mana utama dapat melakukan tindakan pada sumber daya, dan dalam kondisi apa.

Elemen `Action` dari kebijakan JSON menjelaskan tindakan yang dapat Anda gunakan untuk mengizinkan atau menolak akses dalam sebuah kebijakan. Sertakan tindakan dalam kebijakan untuk memberikan izin untuk melakukan operasi terkait.

Untuk melihat daftar AWS Clean Rooms tindakan, lihat [Tindakan yang ditentukan oleh AWS Clean Rooms](#) dalam Referensi Otorisasi Layanan.

Tindakan kebijakan AWS Clean Rooms menggunakan awalan berikut sebelum tindakan.

```
cleanrooms
```

Untuk menetapkan secara spesifik beberapa tindakan dalam satu pernyataan, pisahkan tindakan tersebut dengan koma.

```
"Action": [  
    "cleanrooms:action1",  
    "cleanrooms:action2"  
]
```

Untuk melihat contoh kebijakan AWS Clean Rooms berbasis identitas, lihat. [Contoh kebijakan berbasis identitas untuk AWS Clean Rooms](#)

Sumber daya kebijakan untuk AWS Clean Rooms

Mendukung sumber daya kebijakan: Ya

Administrator dapat menggunakan kebijakan AWS JSON untuk menentukan siapa yang memiliki akses ke apa. Yaitu, di mana utama dapat melakukan tindakan pada sumber daya, dan dalam kondisi apa.

Elemen kebijakan JSON `Resource` menentukan objek yang menjadi target penerapan tindakan. Praktik terbaiknya, tentukan sumber daya menggunakan [Amazon Resource Name \(ARN\)](#). Untuk tindakan yang tidak mendukung izin di tingkat sumber daya, gunakan wildcard (*) untuk menunjukkan bahwa pernyataan tersebut berlaku untuk semua sumber daya.

```
"Resource": ""
```

Untuk melihat daftar jenis sumber daya dan jenis AWS Clean Rooms sumber daya ARNs, lihat [Sumber daya yang ditentukan oleh AWS Clean Rooms](#) dalam Referensi Otorisasi Layanan. Untuk mempelajari dengan tindakan mana Anda dapat menentukan ARN setiap sumber daya, lihat [Tindakan yang ditentukan oleh AWS Clean Rooms](#).

Untuk melihat contoh kebijakan AWS Clean Rooms berbasis identitas, lihat. [Contoh kebijakan berbasis identitas untuk AWS Clean Rooms](#)

Kunci kondisi kebijakan untuk AWS Clean Rooms

Mendukung kunci kondisi kebijakan khusus layanan: Sebagian

Administrator dapat menggunakan kebijakan AWS JSON untuk menentukan siapa yang memiliki akses ke apa. Yaitu, principal dapat melakukan tindakan pada suatu sumber daya, dan dalam suatu syarat.

Elemen `Condition` menentukan ketika pernyataan dieksekusi berdasarkan kriteria yang ditetapkan. Anda dapat membuat ekspresi bersyarat yang menggunakan [operator kondisi](#), misalnya sama dengan atau kurang dari, untuk mencocokkan kondisi dalam kebijakan dengan nilai-nilai yang diminta. Untuk melihat semua kunci kondisi AWS global, lihat [kunci konteks kondisi AWS global](#) di Panduan Pengguna IAM.

Untuk mempelajari cara AWS Clean Rooms menggunakan kunci kondisi kebijakan, lihat [Perilaku IAM untuk AWS Clean Rooms MS](#).

ACLs di AWS Clean Rooms

Mendukung ACLs: Tidak

Access control lists (ACLs) mengontrol prinsipal mana (anggota akun, pengguna, atau peran) yang memiliki izin untuk mengakses sumber daya. ACLs mirip dengan kebijakan berbasis sumber daya, meskipun mereka tidak menggunakan format dokumen kebijakan JSON.

ABAC dengan AWS Clean Rooms

Mendukung ABAC (tanda dalam kebijakan): Ya

Kontrol akses berbasis atribut (ABAC) adalah strategi otorisasi yang menentukan izin berdasarkan atribut tanda. Anda dapat melampirkan tag ke entitas dan AWS sumber daya IAM, lalu merancang kebijakan ABAC untuk mengizinkan operasi saat tag prinsipal cocok dengan tag pada sumber daya.

Untuk mengendalikan akses berdasarkan tanda, berikan informasi tentang tanda di [elemen kondisi](#) dari kebijakan menggunakan kunci kondisi `aws:ResourceTag/key-name`, `aws:RequestTag/key-name`, atau `aws:TagKeys`.

Jika sebuah layanan mendukung ketiga kunci kondisi untuk setiap jenis sumber daya, nilainya adalah Ya untuk layanan tersebut. Jika suatu layanan mendukung ketiga kunci kondisi untuk hanya beberapa jenis sumber daya, nilainya adalah Parsial.

Untuk informasi selengkapnya tentang ABAC, lihat [Tentukan izin dengan otorisasi ABAC](#) dalam Panduan Pengguna IAM. Untuk melihat tutorial yang menguraikan langkah-langkah pengaturan ABAC, lihat [Menggunakan kontrol akses berbasis atribut \(ABAC\)](#) dalam Panduan Pengguna IAM.

Menggunakan kredensi sementara dengan AWS Clean Rooms

Mendukung kredensial sementara: Ya

Kredensi sementara menyediakan akses jangka pendek ke AWS sumber daya dan secara otomatis dibuat saat Anda menggunakan federasi atau beralih peran. AWS merekomendasikan agar Anda secara dinamis menghasilkan kredensi sementara alih-alih menggunakan kunci akses jangka panjang. Untuk informasi selengkapnya, lihat [Kredensial keamanan sementara di IAM](#) dan [Layanan AWS yang berfungsi dengan IAM](#) dalam Panduan Pengguna IAM.

Teruskan sesi akses untuk AWS Clean Rooms

Mendukung sesi akses terusan (FAS): Ya

Sesi akses terusan (FAS) menggunakan izin dari pemanggilan utama Layanan AWS, dikombinasikan dengan permintaan Layanan AWS untuk membuat permintaan ke layanan hilir. Untuk detail kebijakan ketika mengajukan permintaan FAS, lihat [Sesi akses terusan](#).

Peran layanan untuk AWS Clean Rooms

Mendukung peran layanan: Ya

Peran layanan adalah [peran IAM](#) yang diambil oleh sebuah layanan untuk melakukan tindakan atas nama Anda. Administrator IAM dapat membuat, mengubah, dan menghapus peran layanan dari

dalam IAM. Untuk informasi selengkapnya, lihat [Buat sebuah peran untuk mendelegasikan izin ke Layanan AWS](#) dalam Panduan pengguna IAM.

 Warning

Mengubah izin untuk peran layanan dapat merusak AWS Clean Rooms fungsionalitas. Edit peran layanan hanya jika AWS Clean Rooms memberikan panduan untuk melakukannya.

Peran terkait layanan untuk AWS Clean Rooms

Mendukung peran terkait layanan: Tidak

Peran terkait layanan adalah jenis peran layanan yang ditautkan ke. Layanan AWS Layanan tersebut dapat menjalankan peran untuk melakukan tindakan atas nama Anda. Peran terkait layanan muncul di Anda Akun AWS dan dimiliki oleh layanan. Administrator IAM dapat melihat, tetapi tidak dapat mengedit izin untuk peran terkait layanan.

Untuk detail tentang pembuatan atau manajemen peran terkait layanan, lihat [Layanan AWS yang berfungsi dengan IAM](#). Cari layanan dalam tabel yang memiliki Yes di kolom Peran terkait layanan. Pilih tautan Ya untuk melihat dokumentasi peran terkait layanan untuk layanan tersebut.

Contoh kebijakan berbasis identitas untuk AWS Clean Rooms

Secara default, pengguna dan peran tidak memiliki izin untuk membuat atau mengubah sumber daya AWS Clean Rooms . Untuk memberikan izin kepada pengguna untuk melakukan tindakan di sumber daya yang mereka perlukan, administrator IAM dapat membuat kebijakan IAM.

Untuk mempelajari cara membuat kebijakan berbasis identitas IAM dengan menggunakan contoh dokumen kebijakan JSON ini, lihat [Membuat kebijakan IAM \(konsol\) di Panduan Pengguna IAM](#).

Untuk detail tentang tindakan dan jenis sumber daya yang ditentukan oleh AWS Clean Rooms, termasuk format ARNs untuk setiap jenis sumber daya, lihat [Kunci tindakan, sumber daya, dan kondisi untuk AWS Clean Rooms](#) dalam Referensi Otorisasi Layanan.

Topik

- [Praktik terbaik kebijakan](#)
- [Menggunakan AWS Clean Rooms konsol](#)
- [Mengizinkan pengguna melihat izin mereka sendiri](#)

Praktik terbaik kebijakan

Kebijakan berbasis identitas menentukan apakah seseorang dapat membuat, mengakses, atau menghapus AWS Clean Rooms sumber daya di akun Anda. Tindakan ini membuat Akun AWS Anda dikenai biaya. Ketika Anda membuat atau mengedit kebijakan berbasis identitas, ikuti panduan dan rekomendasi ini:

- Mulailah dengan kebijakan AWS terkelola dan beralih ke izin hak istimewa paling sedikit — Untuk mulai memberikan izin kepada pengguna dan beban kerja Anda, gunakan kebijakan AWS terkelola yang memberikan izin untuk banyak kasus penggunaan umum. Mereka tersedia di Anda Akun AWS. Kami menyarankan Anda mengurangi izin lebih lanjut dengan menentukan kebijakan yang dikelola AWS pelanggan yang khusus untuk kasus penggunaan Anda. Untuk informasi selengkapnya, lihat [Kebijakan yang dikelola AWS](#) atau [Kebijakan yang dikelola AWS untuk fungsi tugas](#) dalam Panduan Pengguna IAM.
- Menerapkan izin dengan hak akses paling rendah – Ketika Anda menetapkan izin dengan kebijakan IAM, hanya berikan izin yang diperlukan untuk melakukan tugas. Anda melakukannya dengan mendefinisikan tindakan yang dapat diambil pada sumber daya tertentu dalam kondisi tertentu, yang juga dikenal sebagai izin dengan hak akses paling rendah. Untuk informasi selengkapnya tentang cara menggunakan IAM untuk mengajukan izin, lihat [Kebijakan dan izin dalam IAM](#) dalam Panduan Pengguna IAM.
- Gunakan kondisi dalam kebijakan IAM untuk membatasi akses lebih lanjut – Anda dapat menambahkan suatu kondisi ke kebijakan Anda untuk membatasi akses ke tindakan dan sumber daya. Sebagai contoh, Anda dapat menulis kondisi kebijakan untuk menentukan bahwa semua permintaan harus dikirim menggunakan SSL. Anda juga dapat menggunakan ketentuan untuk memberikan akses ke tindakan layanan jika digunakan melalui yang spesifik Layanan AWS, seperti CloudFormation. Untuk informasi selengkapnya, lihat [Elemen kebijakan JSON IAM: Kondisi](#) dalam Panduan Pengguna IAM.
- Gunakan IAM Access Analyzer untuk memvalidasi kebijakan IAM Anda untuk memastikan izin yang aman dan fungsional – IAM Access Analyzer memvalidasi kebijakan baru dan yang sudah ada sehingga kebijakan tersebut mematuhi bahasa kebijakan IAM (JSON) dan praktik terbaik IAM. IAM Access Analyzer menyediakan lebih dari 100 pemeriksaan kebijakan dan rekomendasi yang dapat ditindaklanjuti untuk membantu Anda membuat kebijakan yang aman dan fungsional. Untuk informasi selengkapnya, lihat [Validasi kebijakan dengan IAM Access Analyzer](#) dalam Panduan Pengguna IAM.
- Memerlukan otentikasi multi-faktor (MFA) - Jika Anda memiliki skenario yang mengharuskan pengguna IAM atau pengguna root di Anda, Akun AWS aktifkan MFA untuk keamanan tambahan.

Untuk meminta MFA ketika operasi API dipanggil, tambahkan kondisi MFA pada kebijakan Anda. Untuk informasi selengkapnya, lihat [Amankan akses API dengan MFA](#) dalam Panduan Pengguna IAM.

Untuk informasi selengkapnya tentang praktik terbaik dalam IAM, lihat [Praktik terbaik keamanan di IAM](#) dalam Panduan Pengguna IAM.

Menggunakan AWS Clean Rooms konsol

Untuk mengakses AWS Clean Rooms konsol, Anda harus memiliki set izin minimum. Izin ini harus memungkinkan Anda untuk membuat daftar dan melihat detail tentang AWS Clean Rooms sumber daya di Anda Akun AWS. Jika Anda membuat kebijakan berbasis identitas yang lebih ketat daripada izin minimum yang diperlukan, konsol tidak akan berfungsi sebagaimana mestinya untuk entitas (pengguna atau peran) dengan kebijakan tersebut.

Anda tidak perlu mengizinkan izin konsol minimum untuk pengguna yang melakukan panggilan hanya ke AWS CLI atau AWS API. Sebagai gantinya, izinkan akses hanya ke tindakan yang sesuai dengan operasi API yang coba mereka lakukan.

Untuk memastikan bahwa pengguna dan peran masih dapat menggunakan AWS Clean Rooms konsol, lampirkan juga kebijakan AWS Clean Rooms *FullAccess* atau *ReadOnly* AWS terkelola ke entitas. Untuk informasi selengkapnya, lihat [Menambah izin untuk pengguna](#) dalam Panduan Pengguna IAM.

Mengizinkan pengguna melihat izin mereka sendiri

Contoh ini menunjukkan cara membuat kebijakan yang mengizinkan pengguna IAM melihat kebijakan inline dan terkelola yang dilampirkan ke identitas pengguna mereka. Kebijakan ini mencakup izin untuk menyelesaikan tindakan ini di konsol atau menggunakan API atau secara terprogram. AWS CLI AWS

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "ViewOwnUserInfo",
      "Effect": "Allow",
      "Action": [
        "iam:GetUserPolicy",
        "iam:ListGroupsForUser",
```

```

        "iam:ListAttachedUserPolicies",
        "iam:ListUserPolicies",
        "iam:GetUser"
    ],
    "Resource": ["arn:aws:iam::*:user/${aws:username}"]
},
{
    "Sid": "NavigateInConsole",
    "Effect": "Allow",
    "Action": [
        "iam:GetGroupPolicy",
        "iam:GetPolicyVersion",
        "iam:GetPolicy",
        "iam:ListAttachedGroupPolicies",
        "iam:ListGroupPolicies",
        "iam:ListPolicyVersions",
        "iam:ListPolicies",
        "iam:ListUsers"
    ],
    "Resource": "*"
}
]
}

```

AWS kebijakan terkelola untuk AWS Clean Rooms

Kebijakan AWS terkelola adalah kebijakan mandiri yang dibuat dan dikelola oleh AWS. AWS Kebijakan terkelola dirancang untuk memberikan izin bagi banyak kasus penggunaan umum sehingga Anda dapat mulai menetapkan izin kepada pengguna, grup, dan peran.

Perlu diingat bahwa kebijakan AWS terkelola mungkin tidak memberikan izin hak istimewa paling sedikit untuk kasus penggunaan spesifik Anda karena tersedia untuk digunakan semua pelanggan. AWS Kami menyarankan Anda untuk mengurangi izin lebih lanjut dengan menentukan [kebijakan yang dikelola pelanggan](#) yang khusus untuk kasus penggunaan Anda.

Anda tidak dapat mengubah izin yang ditentukan dalam kebijakan AWS terkelola. Jika AWS memperbarui izin yang ditentukan dalam kebijakan AWS terkelola, pembaruan akan memengaruhi semua identitas utama (pengguna, grup, dan peran) yang dilampirkan kebijakan tersebut. AWS kemungkinan besar akan memperbarui kebijakan AWS terkelola saat baru Layanan AWS diluncurkan atau operasi API baru tersedia untuk layanan yang ada.

Untuk informasi selengkapnya, lihat [Kebijakan terkelola AWS](#) dalam Panduan Pengguna IAM.

AWS kebijakan terkelola: **AWSCleanRoomsReadOnlyAccess**

Anda dapat melampirkan **AWSCleanRoomsReadOnlyAccess** ke kepala IAM Anda.

Kebijakan ini memberikan izin hanya-baca untuk sumber daya dan metadata dalam kolaborasi.

AWSCleanRoomsReadOnlyAccess

Detail izin

Kebijakan ini mencakup izin berikut:

- **CleanRoomsRead**— Memungkinkan kepala sekolah akses hanya-baca ke layanan.
- **ConsoleDisplayTables**— Memungkinkan akses hanya-baca prinsipal ke AWS Glue metadata yang diperlukan untuk menampilkan data tentang tabel yang mendasarinya di konsol. AWS Glue
- **ConsoleLogSummaryQueryLogs**— Memungkinkan kepala sekolah untuk melihat log kueri.
- **ConsoleLogSummaryObtainLogs**— Memungkinkan kepala sekolah untuk mengambil hasil log.

Untuk daftar JSON tentang detail kebijakan, lihat [AWSCleanRoomsReadOnlyAccess](#) di Panduan Referensi Kebijakan AWS Terkelola.

AWS kebijakan terkelola: **AWSCleanRoomsFullAccess**

Anda dapat melampirkan **AWSCleanRoomsFullAccess** ke kepala IAM Anda.

Kebijakan ini memberikan izin administratif yang memungkinkan akses penuh (baca, tulis, dan perbarui) ke sumber daya dan metadata dalam suatu kolaborasi. AWS Clean Rooms Kebijakan ini mencakup akses untuk melakukan kueri.

Detail izin

Kebijakan ini mencakup izin berikut:

- **CleanRoomsAccess**— Memberikan akses penuh ke semua tindakan pada semua sumber daya untuk AWS Clean Rooms.
- **PassServiceRole**— Memberikan akses untuk meneruskan peran layanan hanya ke layanan (**PassedToService** kondisi) yang memiliki "cleanrooms" dalam namanya.
- **ListRolesToPickServiceRole**— Memungkinkan kepala sekolah untuk membuat daftar semua peran mereka untuk memilih peran layanan saat menggunakan. AWS Clean Rooms

- `GetRoleAndListRolePoliciesToInspectServiceRole`— Memungkinkan kepala sekolah untuk melihat peran layanan dan kebijakan terkait di IAM.
- `ListPoliciesToInspectServiceRolePolicy`— Memungkinkan kepala sekolah untuk melihat peran layanan dan kebijakan terkait di IAM.
- `GetPolicyToInspectServiceRolePolicy`— Memungkinkan kepala sekolah untuk melihat peran layanan dan kebijakan terkait di IAM.
- `ConsoleDisplayTables`— Memungkinkan akses hanya-baca prinsipal ke AWS Glue metadata yang diperlukan untuk menampilkan data tentang tabel yang mendasarinya di konsol. AWS Glue
- `ConsolePickQueryResultsBucketListAll`— Memungkinkan kepala sekolah memilih bucket Amazon S3 dari daftar semua bucket S3 yang tersedia di mana hasil kueri mereka ditulis.
- `SetQueryResultsBucket`— Memungkinkan kepala sekolah untuk memilih bucket S3 di mana hasil kueri mereka ditulis.
- `ConsoleDisplayQueryResults`— Memungkinkan kepala sekolah untuk menunjukkan hasil kueri kepada pelanggan, baca dari bucket S3.
- `WriteQueryResults`— Memungkinkan kepala sekolah untuk menulis hasil kueri ke dalam bucket S3 milik pelanggan.
- `EstablishLogDeliveries`— Memungkinkan prinsipal mengirimkan log kueri ke grup CloudWatch log Amazon Logs pelanggan.
- `SetupLogGroupsDescribe`— Memungkinkan kepala sekolah untuk menggunakan proses pembuatan grup CloudWatch log Amazon Logs.
- `SetupLogGroupsCreate`— Memungkinkan kepala sekolah untuk membuat grup CloudWatch log Amazon Logs.
- `SetupLogGroupsResourcePolicy`— Memungkinkan prinsipal untuk menyiapkan kebijakan sumber daya di grup CloudWatch log Amazon Logs.
- `ConsoleLogSummaryQueryLogs`— Memungkinkan kepala sekolah untuk melihat log kueri.
- `ConsoleLogSummaryObtainLogs`— Memungkinkan kepala sekolah untuk mengambil hasil log.

Untuk daftar JSON tentang detail kebijakan, lihat [AWSCleanRoomsFullAccess](#) di Panduan Referensi Kebijakan AWS Terkelola.

AWS kebijakan terkelola: **AWSCleanRoomsFullAccessNoQuerying**

Anda dapat melampirkan `AWSCleanRoomsFullAccessNoQuerying` ke `AndalAM` principals.

Kebijakan ini memberikan izin administratif yang memungkinkan akses penuh (baca, tulis, dan perbarui) ke sumber daya dan metadata dalam suatu kolaborasi. AWS Clean Rooms Kebijakan ini mengecualikan akses untuk melakukan kueri.

Detail izin

Kebijakan ini mencakup izin berikut:

- **CleanRoomsAccess**— Memberikan akses penuh ke semua tindakan pada semua sumber daya untuk AWS Clean Rooms, kecuali untuk kueri dalam kolaborasi.
- **CleanRoomsNoQuerying**— Secara eksplisit menyangkal **StartProtectedQuery** dan **UpdateProtectedQuery** mencegah kueri.
- **PassServiceRole**— Memberikan akses untuk meneruskan peran layanan hanya ke layanan (**PassedToService** kondisi) yang memiliki "cleanrooms" dalam namanya.
- **ListRolesToPickServiceRole**— Memungkinkan kepala sekolah untuk membuat daftar semua peran mereka untuk memilih peran layanan saat menggunakan. AWS Clean Rooms
- **GetRoleAndListRolePoliciesToInspectServiceRole**— Memungkinkan kepala sekolah untuk melihat peran layanan dan kebijakan terkait di IAM.
- **ListPoliciesToInspectServiceRolePolicy**— Memungkinkan kepala sekolah untuk melihat peran layanan dan kebijakan terkait di IAM.
- **GetPolicyToInspectServiceRolePolicy**— Memungkinkan kepala sekolah untuk melihat peran layanan dan kebijakan terkait di IAM.
- **ConsoleDisplayTables**— Memungkinkan akses hanya-baca prinsipal ke AWS Glue metadata yang diperlukan untuk menampilkan data tentang tabel yang mendasarinya di konsol. AWS Glue
- **EstablishLogDeliveries**— Memungkinkan prinsipal mengirimkan log kueri ke grup CloudWatch log Amazon Logs pelanggan.
- **SetupLogGroupsDescribe**— Memungkinkan kepala sekolah untuk menggunakan proses pembuatan grup CloudWatch log Amazon Logs.
- **SetupLogGroupsCreate**— Memungkinkan kepala sekolah untuk membuat grup CloudWatch log Amazon Logs.
- **SetupLogGroupsResourcePolicy**— Memungkinkan prinsipal untuk menyiapkan kebijakan sumber daya di grup CloudWatch log Amazon Logs.
- **ConsoleLogSummaryQueryLogs**— Memungkinkan kepala sekolah untuk melihat log kueri.
- **ConsoleLogSummaryObtainLogs**— Memungkinkan kepala sekolah untuk mengambil hasil log.

- `cleanrooms`— Kelola kolaborasi, templat analisis, tabel yang dikonfigurasi, keanggotaan, dan sumber daya terkait dalam layanan. AWS Clean Rooms Lakukan berbagai operasi seperti membuat, memperbarui, menghapus, mencantumkan, dan mengambil informasi tentang sumber daya ini.
- `iam`— Lulus peran layanan dengan nama yang berisi `cleanrooms` ke AWS Clean Rooms layanan. Buat daftar peran, kebijakan, dan periksa peran dan kebijakan layanan yang terkait dengan AWS Clean Rooms layanan.
- `glue`— Mengambil informasi tentang database, tabel, partisi, dan skema dari. AWS Glue Ini diperlukan agar AWS Clean Rooms layanan dapat menampilkan dan berinteraksi dengan sumber data yang mendasarinya.
- `logs`— Mengelola pengiriman log, grup log, dan kebijakan sumber daya untuk CloudWatch Log. Kueri dan ambil log yang terkait dengan AWS Clean Rooms layanan. Izin ini diperlukan untuk tujuan pemantauan, audit, dan pemecahan masalah dalam layanan.

Kebijakan ini juga secara eksplisit menyangkal tindakan `cleanrooms:StartProtectedQuery` dan `cleanrooms:UpdateProtectedQuery` untuk mencegah pengguna mengeksekusi atau memperbarui kueri yang dilindungi secara langsung, yang harus dilakukan melalui mekanisme yang dikendalikan. AWS Clean Rooms

Untuk daftar JSON tentang detail kebijakan, lihat [AWSCleanRoomsFullAccessNoQuerying](#) di Panduan Referensi Kebijakan AWS Terkelola.

AWS kebijakan terkelola: **AWSCleanRoomsMLReadOnlyAccess**

Anda dapat melampirkan `AWSCleanRoomsMLReadOnlyAccess` ke kepala IAM Anda.

Kebijakan ini memberikan izin hanya-baca untuk sumber daya dan metadata dalam kolaborasi. `AWSCleanRoomsMLReadOnlyAccess`

Kebijakan ini mencakup izin berikut:

- `CleanRoomsConsoleNavigation`— Memberikan akses untuk melihat layar AWS Clean Rooms konsol.
- `CleanRoomsMLRead`— Memungkinkan akses hanya-baca kepala sekolah ke layanan Clean Rooms MS.
- `PassCleanRoomsResources`— Memberikan akses untuk melewati AWS Clean Rooms sumber daya yang ditentukan.

Untuk daftar JSON tentang detail kebijakan, lihat [AWSCleanKamar MLRead OnlyAccess](#) di Panduan Referensi Kebijakan AWS Terkelola.

AWS kebijakan terkelola: **AWSCleanRoomsMLFullAccess**

Anda dapat melampirkan **AWSCleanRoomsMLFullAccess** ke kepala IAM Anda. Kebijakan ini memberikan izin administratif yang memungkinkan akses penuh (baca, tulis, dan perbarui) ke sumber daya dan metadata yang dibutuhkan oleh Clean Rooms.

Detail izin

Kebijakan ini mencakup izin berikut:

- **CleanRoomsMLFullAccess**— Memberikan akses ke semua tindakan Clean Rooms MS.
- **PassServiceRole**— Memberikan akses untuk meneruskan peran layanan hanya ke layanan (**PassedToService** kondisi) yang memiliki "cleanrooms-ml" dalam namanya.
- **CleanRoomsConsoleNavigation**— Memberikan akses untuk melihat layar AWS Clean Rooms konsol.
- **CollaborationMembershipCheck**— Saat Anda memulai pekerjaan pembuatan audiens (segmen mirip) dalam sebuah kolaborasi, layanan Clean Rooms MS memanggil **ListMembers** untuk memeriksa apakah kolaborasi tersebut valid, pemanggil adalah anggota aktif, dan pemilik model audiens yang dikonfigurasi adalah anggota aktif. Izin ini selalu diperlukan; SID navigasi konsol hanya diperlukan untuk pengguna konsol.
- **PassCleanRoomsResources**— Memberikan akses untuk melewati AWS Clean Rooms sumber daya yang ditentukan.
- **AssociateModels**— Memungkinkan kepala sekolah untuk mengaitkan model Clean Rooms MS dengan kolaborasi Anda.
- **TagAssociations**— Memungkinkan prinsipal untuk menambahkan tag ke asosiasi antara model mirip dan kolaborasi.
- **ListRolesToPickServiceRole**— Memungkinkan kepala sekolah untuk membuat daftar semua peran mereka untuk memilih peran layanan saat menggunakan AWS Clean Rooms
- **GetRoleAndListRolePoliciesToInspectServiceRole**— Memungkinkan kepala sekolah untuk melihat peran layanan dan kebijakan terkait di IAM.
- **ListPoliciesToInspectServiceRolePolicy**— Memungkinkan kepala sekolah untuk melihat peran layanan dan kebijakan terkait di IAM.

- `GetPolicyToInspectServiceRolePolicy`— Memungkinkan kepala sekolah untuk melihat peran layanan dan kebijakan terkait di IAM.
- `ConsoleDisplayTables`— Memungkinkan akses hanya-baca prinsipal ke AWS Glue metadata yang diperlukan untuk menampilkan data tentang tabel yang mendasarinya di konsol. AWS Glue
- `ConsolePickOutputBucket`— Memungkinkan kepala sekolah memilih bucket Amazon S3 untuk output model audiens yang dikonfigurasi.
- `ConsolePickS3Location`— Memungkinkan kepala sekolah untuk memilih lokasi dalam ember untuk output model audiens yang dikonfigurasi.
- `ConsoleDescribeECRRepositories`— Memungkinkan kepala sekolah untuk menggambarkan repositori dan gambar Amazon ECR.

Untuk daftar JSON tentang detail kebijakan, lihat [MLFullAkses AWSClean Kamar](#) di Panduan Referensi Kebijakan AWS Terkelola.

AWS Clean Rooms pembaruan kebijakan AWS terkelola

Lihat detail tentang pembaruan kebijakan AWS terkelola AWS Clean Rooms sejak layanan ini mulai melacak perubahan ini. Untuk peringatan otomatis tentang perubahan pada halaman ini, berlangganan umpan RSS di halaman Riwayat AWS Clean Rooms dokumen.

Ubah	Deskripsi	Date
AWSCleanRoomsFullAccessNoQuering — Perbarui ke kebijakan yang ada	Menambahkan kamar bersih: <code>UpdateConfiguredTableAllowedColumns</code> dan kamar bersih: <code>UpdateConfiguredTableReference</code> ke <code>CleanRoomsAccess</code>	Juli 29, 2025
AWSCleanRoomsMLReadOnlyAccess – Pembaruan ke kebijakan yang sudah ada	Menambahkan <code>PassCleanRoomsResources</code> ke <code>AWSCleanRoomsMLReadOnlyAccess</code> . Ditambahkan <code>PassCleanRoomsResources</code> dan <code>ConsoleDescribeECRRepositories</code> ke <code>AWSCleanRoomsMLFullAccess</code> .	Januari 10, 2025

Ubah	Deskripsi	Date
AWSCleanRoomsFullAccessNoQuering – Pembaruan ke kebijakan yang sudah ada	Menambahkan cleanrooms:BatchGetSchemaAnalysisRule ke CleanRoomsAccess.	13 Mei 2024
AWSCleanRoomsFullAccess – Pembaruan ke kebijakan yang sudah ada	Memperbarui ID Pernyataan AWSCleanRoomsFullAccess dari ConsolePickQueryResultsBucket ke SetQueryResultsBucket dalam kebijakan ini untuk merepresentasikan izin dengan lebih baik karena izin diperlukan untuk menyetel bucket hasil kueri baik dengan maupun tanpa konsol.	Maret 21, 2024
AWSCleanRoomsMLReadOnlyAccess – Kebijakan baru AWSCleanRoomsMLFullAccess – Kebijakan baru	Ditambahkan AWSCleanRoomsMLReadOnlyAccess dan AWSCleanRoomsMLFullAccess mendukung AWS Clean Rooms MS.	November 29, 2023
AWSCleanRoomsFullAccessNoQuering – Pembaruan ke kebijakan yang sudah ada	Ditambahkan cleanrooms:CreateAnalysisTemplate,cleanrooms:GetAnalysisTemplate,cleanrooms:UpdateAnalysisTemplate, cleanrooms:DeleteAnalysisTemplate,cleanrooms:ListAnalysisTemplates,cleanrooms:GetCollaborationAnalysisTemplate,cleanrooms:BatchGetCollaborationAnalysisTemplate,, dan cleanrooms:ListCollaborationAnalysisTemplates CleanRoomsAccess untuk mengaktifkan fitur template analisis baru.	31 Juli 2023

Ubah	Deskripsi	Date
AWSCleanRoomsFullAccessNoQuering – Pembaruan ke kebijakan yang sudah ada	Ditambahkancleanrooms:ListTagsForResource,cleanrooms:UntagResource, dan cleanrooms:TagResource CleanRoomsAccess untuk mengaktifkan penandaan sumber daya.	21 Maret 2023
AWS Clean Rooms mulai melacak perubahan	AWS Clean Rooms mulai melacak perubahan untuk kebijakan yang AWS dikelola.	Januari 12, 2023

Memecahkan masalah AWS Clean Rooms identitas dan akses

Gunakan informasi berikut untuk membantu Anda mendiagnosis dan memperbaiki masalah umum yang mungkin Anda temui saat bekerja dengan AWS Clean Rooms dan IAM.

Topik

- [Saya tidak berwenang untuk melakukan tindakan di AWS Clean Rooms](#)
- [Saya tidak berwenang untuk melakukan iam: PassRole](#)
- [Saya ingin mengizinkan orang di luar saya Akun AWS untuk mengakses AWS Clean Rooms sumber daya saya](#)

Saya tidak berwenang untuk melakukan tindakan di AWS Clean Rooms

Jika Anda menerima pesan kesalahan bahwa Anda tidak memiliki otorisasi untuk melakukan tindakan, kebijakan Anda harus diperbarui agar Anda dapat melakukan tindakan tersebut.

Contoh kesalahan berikut terjadi ketika pengguna IAM `mateojackson` mencoba menggunakan konsol untuk melihat detail tentang suatu sumber daya fiktif `my-example-widget`, tetapi tidak memiliki izin fiktif `cleanrooms:GetWidget`.

```
User: arn:aws:iam::123456789012:user/mateojackson is not authorized to perform:
cleanrooms:GetWidget on resource: my-example-widget
```


Dalam hal ini, kebijakan Mateo harus diperbarui untuk memungkinkannya mengakses *my-example-widget* sumber daya menggunakan `cleanrooms:GetWidget` tindakan tersebut.

Jika Anda memerlukan bantuan, hubungi AWS administrator Anda. Administrator Anda adalah orang yang memberi Anda kredensial masuk.

Saya tidak berwenang untuk melakukan `iam:PassRole`

Jika Anda menerima kesalahan yang tidak diizinkan untuk melakukan `iam:PassRole` tindakan, kebijakan Anda harus diperbarui agar Anda dapat meneruskan peran AWS Clean Rooms.

Beberapa Layanan AWS memungkinkan Anda untuk meneruskan peran yang ada ke layanan tersebut alih-alih membuat peran layanan baru atau peran terkait layanan. Untuk melakukannya, Anda harus memiliki izin untuk meneruskan peran ke layanan.

Contoh kesalahan berikut terjadi ketika pengguna IAM bernama `marymajor` mencoba menggunakan konsol tersebut untuk melakukan tindakan di AWS Clean Rooms. Namun, tindakan tersebut memerlukan layanan untuk mendapatkan izin yang diberikan oleh peran layanan. Mary tidak memiliki izin untuk meneruskan peran tersebut pada layanan.

```
User: arn:aws:iam::123456789012:user/marymajor is not authorized to perform:
iam:PassRole
```

Dalam kasus ini, kebijakan Mary harus diperbarui agar dia mendapatkan izin untuk melakukan tindakan `iam:PassRole` tersebut.

Jika Anda memerlukan bantuan, hubungi AWS administrator Anda. Administrator Anda adalah orang yang memberi Anda kredensial masuk.

Saya ingin mengizinkan orang di luar saya Akun AWS untuk mengakses AWS Clean Rooms sumber daya saya

Anda dapat membuat peran yang dapat digunakan pengguna di akun lain atau orang-orang di luar organisasi Anda untuk mengakses sumber daya Anda. Anda dapat menentukan siapa saja yang dipercaya untuk mengambil peran tersebut.

Untuk mempelajari selengkapnya, periksa referensi berikut:

- Untuk mempelajari apakah AWS Clean Rooms mendukung fitur ini, lihat [Bagaimana AWS Clean Rooms bekerja dengan IAM](#).

- Untuk mempelajari cara menyediakan akses ke sumber daya Anda di seluruh sumber daya Akun AWS yang Anda miliki, lihat [Menyediakan akses ke pengguna IAM di pengguna lain Akun AWS yang Anda miliki](#) di Panduan Pengguna IAM.
- Untuk mempelajari cara menyediakan akses ke sumber daya Anda kepada pihak ketiga Akun AWS, lihat [Menyediakan akses yang Akun AWS dimiliki oleh pihak ketiga](#) dalam Panduan Pengguna IAM.
- Untuk mempelajari cara memberikan akses melalui federasi identitas, lihat [Menyediakan akses ke pengguna terautentikasi eksternal \(federasi identitas\)](#) dalam Panduan Pengguna IAM.
- Untuk mempelajari perbedaan antara penggunaan kebijakan peran dan kebijakan berbasis sumber daya untuk akses lintas akun, lihat [Bagaimana peran IAM berbeda dari kebijakan berbasis sumber daya](#) dalam Panduan Pengguna IAM.

Pencegahan "confused deputy" lintas layanan

Masalah "confused deputy" adalah masalah keamanan di mana entitas yang tidak memiliki izin untuk melakukan tindakan dapat memengaruhi entitas yang memiliki hak akses lebih tinggi untuk melakukan tindakan. Pada tahun AWS, peniruan lintas layanan dapat mengakibatkan masalah wakil yang membingungkan. Peniruan identitas lintas layanan dapat terjadi ketika satu layanan (layanan yang dipanggil) memanggil layanan lain (layanan yang dipanggil). Layanan pemanggilan dapat dimanipulasi menggunakan izinnya untuk bertindak pada sumber daya pelanggan lain dengan cara yang seharusnya tidak dilakukannya kecuali bila memiliki izin untuk mengakses. Untuk mencegah hal ini, AWS menyediakan alat yang membantu Anda melindungi data untuk semua layanan dengan principal layanan yang telah diberi akses ke sumber daya di akun Anda.

Sebaiknya gunakan kunci konteks kondisi `aws:SourceArn` global dalam kebijakan sumber daya untuk membatasi izin yang AWS Clean Rooms memberikan layanan lain ke sumber daya. Gunakan `aws:SourceArn` jika Anda ingin hanya satu sumber daya yang akan dikaitkan dengan akses lintas layanan.

Cara paling efektif untuk melindungi dari masalah "confused deputy" adalah dengan menggunakan kunci konteks kondisi global `aws:SourceArn` dengan ARN lengkap sumber daya. Di AWS Clean Rooms, Anda juga harus membandingkan dengan kunci `sts:ExternalId` kondisi.

Nilai `aws:SourceArn` harus diatur ke ARN dari keanggotaan peran yang diasumsikan.

Contoh berikut menunjukkan bagaimana Anda dapat menggunakan kunci konteks kondisi `aws:SourceArn` global AWS Clean Rooms untuk mencegah masalah wakil yang membingungkan.

Note

Kebijakan contoh berlaku untuk kebijakan kepercayaan peran layanan yang AWS Clean Rooms digunakan untuk mengakses data dan metadata untuk tabel yang dikonfigurasi. Nilai untuk `<query-runner-membership-id>` perlu disetel ke ID keanggotaan runner kueri.

Semua anggota kolaborasi dapat melihat metadata tabel yang dikonfigurasi sehingga setiap ARN keanggotaan harus dimasukkan dalam daftar keanggotaan. ARNs

Note

Ketika peran layanan dibuat melalui AWS Clean Rooms konsol, semua anggota kolaborasi saat ini disertakan dalam kondisi wakil yang bingung secara default.

Jika Anda menambahkan anggota baru ke kolaborasi yang telah mengonfigurasi tabel yang terkait dengannya, pastikan Anda memperbarui kondisi wakil peran layanan yang membingungkan dengan ARN keanggotaan anggota baru.

Jika Anda tidak memperbarui kondisi wakil yang membingungkan peran layanan Anda setelah menambahkan anggota baru, anggota baru tersebut tidak akan dapat mengakses informasi AWS Clean Rooms yang diambil menggunakan peran tersebut.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowIfExternalIdMatches",
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "StringLike": {
          "sts:ExternalId": "arn:aws:*:us-east-1:*:dbuser:*/<query-runner-membership-id>*"
        }
      }
    }
  ]
}
```

```

    }
  },
  {
    "Sid": "AllowIfSourceArnMatches",
    "Effect": "Allow",
    "Principal": {
      "Service": "cleanrooms.amazonaws.com"
    },
    "Action": "sts:AssumeRole",
    "Condition": {
      "ForAnyValue:ArnEquals": {
        "aws:SourceArn": [
          "arn:aws:cleanrooms:us-  
east-1:111122223333:membership/<member-1-membership-id>",
          "arn:aws:cleanrooms:us-  
east-1:444455556666:membership/<member-2-membership-id>"
        ]
      }
    }
  }
]
}

```

Perilaku IAM untuk AWS Clean Rooms MS

Lowongan kerja lintas akun

Clean Rooms ML memungkinkan sumber daya tertentu yang dibuat oleh satu orang Akun AWS untuk diakses dengan aman di akun mereka oleh yang lain Akun AWS. Ketika klien di Akun AWS A memanggil `StartAudienceGenerationJob` `ConfiguredAudienceModel` sumber daya yang dimiliki oleh Akun AWS B, Clean Rooms ML menciptakan dua ARNs untuk pekerjaan itu. Satu ARN di Akun AWS A dan satu lagi di B. Akun AWS ARNs Mereka identik kecuali untuk mereka Akun AWS.

Clean Rooms MS menciptakan dua ARNs untuk pekerjaan tersebut untuk memastikan bahwa kedua akun dapat menerapkan kebijakan IAM mereka sendiri untuk pekerjaan tersebut. Misalnya, kedua akun dapat menggunakan kontrol akses berbasis tag dan menerapkan kebijakan dari AWS organisasi mereka. Pekerjaan memproses data dari kedua akun, sehingga kedua akun dapat menghapus pekerjaan dan data terkait. Tidak ada akun yang dapat memblokir akun lain dari menghapus pekerjaan.

Hanya ada satu eksekusi pekerjaan dan kedua akun dapat melihat pekerjaan ketika mereka menelepon `ListAudienceGenerationJobs`. Kedua akun dapat memanggil `Get`, `Delete`, dan `Export` APIs di tempat kerja menggunakan ARN dengan ID mereka sendiri Akun AWS .

Tidak ada yang Akun AWS dapat mengakses pekerjaan saat menggunakan ARN dengan ID lainnya Akun AWS .

Nama pekerjaan harus unik dalam sebuah Akun AWS. Nama dalam Akun AWS B adalah `$accountA-$name`. Nama yang dipilih oleh Akun AWS A diawali dengan Akun AWS A ketika pekerjaan dilihat di Akun AWS B.

Agar lintas akun `StartAudienceGenerationJob` berhasil, Akun AWS B harus mengizinkan tindakan tersebut pada pekerjaan baru di Akun AWS B dan `ConfiguredAudienceModel` di Akun AWS B menggunakan kebijakan sumber daya yang mirip dengan contoh berikut:

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "Clean-Rooms-CAMA-ID",
      "Effect": "Allow",
      "Principal": {
        "AWS": [
          "111122223333"
        ]
      },
      "Action": [
        "cleanrooms-ml:StartAudienceGenerationJob"
      ],
      "Resource": [
        "arn:aws:cleanrooms-ml:us-east-1:444455556666:configured-audience-model/id",
        "arn:aws:cleanrooms-ml:us-east-1:444455556666:audience-generation-job/*"
      ],
      "Condition": {"StringEquals": {"cleanrooms-ml:CollaborationId": "UUID"}}
    }
  ]
}
```

 Note

Kebijakan sumber daya AWS Clean Rooms ML ini mereferensikan dua hal yang berbeda Akun AWS IDs untuk mendukung pembuatan audiens lintas akun:

- 111122223333 - Ini adalah akun yang berisi prinsipal (pengguna, peran, atau layanan) yang berwenang untuk memulai pekerjaan generasi audiens. Akun ini memulai alur kerja pemrosesan ML.
- 444455556666 - Ini adalah akun yang memiliki sumber daya AWS Clean Rooms ML (model audiens yang dikonfigurasi dan pekerjaan generasi audiens). Akun ini menghosting model ML dan mengelola eksekusi pekerjaan.

Catatan Konfigurasi Tambahan:

- Statement ID (Sid): Ganti CAMA-ID dengan pengenalan AWS Clean Rooms Audience Model Application (CAMA) Anda yang sebenarnya untuk membuat pernyataan kebijakan mudah diidentifikasi.
- Sumber Daya IDs: Ganti *id* dengan ID aktual model audiens yang dikonfigurasi, dan *UUID* dengan ID kolaborasi spesifik Anda.
- Kondisi: `cleanrooms-ml:CollaborationId` Kondisi ini memastikan bahwa pekerjaan generasi audiens hanya dapat dimulai dalam konteks AWS Clean Rooms kolaborasi yang ditentukan, memberikan batas keamanan tambahan.

Konfigurasi lintas akun ini memungkinkan skenario di mana satu organisasi mengelola model dan infrastruktur ML sambil memungkinkan mitra resmi untuk memulai proses pembuatan audiens dalam batas-batas perjanjian kolaborasi mereka.

Jika Anda menggunakan [AWS Clean Rooms MLAPI](#) untuk membuat model mirip yang dikonfigurasi dengan `manageResourcePolicies` disetel ke `true`, AWS Clean Rooms buat kebijakan ini untuk Anda.

Selain itu, kebijakan identitas penelepon di Akun AWS A memerlukan `StartAudienceGenerationJob` izin. `arn:aws:cleanrooms-ml:us-west-1:AccountA:audience-generation-job/*` Jadi ada tiga Sumber Daya IAM untuk

Tindakan `StartAudienceGenerationJob`: pekerjaan Akun AWS A, pekerjaan Akun AWS B, dan Akun AWS B `ConfiguredAudienceModel`.

Warning

Akun AWS Yang memulai pekerjaan menerima peristiwa log AWS CloudTrail audit tentang pekerjaan itu. Akun AWS Yang memiliki `ConfiguredAudienceModel` tidak menerima peristiwa log AWS CloudTrail audit.

Lowongan kerja Tagging

Saat Anda menyetel `childResourceTagOnCreatePolicy=FROM_PARENT_RESOURCE` parameter `CreateConfiguredAudienceModel`, semua pekerjaan pembuatan segmen mirip dalam akun Anda yang dibuat dari model mirip mirip yang dikonfigurasi secara default untuk memiliki tag yang sama dengan model mirip yang dikonfigurasi. Model mirip yang dikonfigurasi adalah induknya dan pekerjaan pembuatan segmen yang mirip adalah anak.

Jika Anda membuat pekerjaan di dalam akun Anda sendiri, tag permintaan pekerjaan akan menggantikan tag induk. Pekerjaan yang dibuat oleh akun lain tidak pernah membuat tag di akun Anda. Jika Anda menetapkan `childResourceTagOnCreatePolicy=FROM_PARENT_RESOURCE` dan akun lain membuat pekerjaan, ada dua salinan pekerjaan. Salinan di akun Anda memiliki tag sumber daya induk dan salinan di akun pengirim pekerjaan memiliki tag dari permintaan.

Memvalidasi kolaborator

Saat memberikan izin kepada anggota AWS Clean Rooms kolaborasi lainnya, kebijakan sumber daya harus menyertakan kunci kondisi. `cleanrooms-ml:CollaborationId` Ini memberlakukan bahwa `collaborationId` parameter disertakan dalam [StartAudienceGenerationJob](#) permintaan. Ketika `collaborationId` parameter disertakan dalam permintaan, Clean Rooms MS memvalidasi bahwa kolaborasi ada, pengirim pekerjaan adalah anggota aktif kolaborasi, dan pemilik model mirip yang dikonfigurasi adalah anggota aktif kolaborasi.

Saat AWS Clean Rooms mengelola kebijakan sumber daya model mirip mirip yang dikonfigurasi (`manageResourcePolicies` parameternya `TRUE` dalam [CreateConfiguredAudienceModelAssociation](#) permintaan), kunci kondisi ini akan disetel dalam kebijakan sumber daya. Oleh karena itu, Anda harus menentukan `collaborationId` in [StartAudienceGenerationJob](#).

Akses lintas akun

Hanya `StartAudienceGenerationJob` dapat dipanggil di seluruh akun. Semua Clean Rooms ML lainnya hanya APIs dapat digunakan dengan sumber daya di akun Anda sendiri. Ini memastikan bahwa data pelatihan Anda, konfigurasi model yang mirip, dan informasi lainnya tetap pribadi.

Clean Rooms tidak pernah mengungkapkan Amazon S3 atau AWS Glue lokasi di seluruh akun. Lokasi data pelatihan, lokasi keluaran model mirip yang dikonfigurasi, dan lokasi benih pekerjaan pembuatan segmen yang mirip tidak pernah terlihat di seluruh akun. Kecuali pencatatan kueri diaktifkan dalam kolaborasi, apakah data benih berasal dari kueri SQL dan kueri itu sendiri tidak terlihat di seluruh akun. Jika Anda `Get` memiliki pekerjaan pembuatan audiens yang dikirimkan oleh akun lain, layanan tidak menampilkan lokasi benih.

Perilaku IAM untuk Kamar Bersih Model Kustom

Lowongan kerja lintas akun

Clean Rooms ML memungkinkan sumber daya tertentu yang terkait dengan kolaborasi yang dibuat oleh satu orang Akun AWS untuk diakses dengan aman di akun mereka oleh yang lain Akun AWS. Klien di Akun AWS A dengan kemampuan anggota untuk menjalankan kueri dapat memanggil `CreateTrainedModel`, `CreateMLInputChannel`, atau `StartTrainedModelInferenceJob` pada `ConfiguredModelAlgorithmAssociation` sumber daya yang dimiliki oleh anggota lain dalam kolaborasi, asalkan `ConfiguredModelAlgorithmAssociation` diizinkan oleh aturan analisis kustom yang dibuat dengan `CreateConfiguredTableAnalysisRule`.

Selain itu, setiap anggota aktif kolaborasi dapat menghapus data yang terkait dengan model terlatih atau saluran input ML melalui `DeleteTrainedModelOutput` dan `DeleteMLInputChannelData` APIs.

Akses lintas akun

Clean Rooms ML memungkinkan pengguna untuk mengambil metadata tentang sumber daya yang dibuat oleh akun lain melalui `GetCollaboration` dan `ListCollaboration` APIs Clean Rooms ML tidak mengungkapkan kunci KMS ARNs, tag, variabel lingkungan, atau hiperparameter (untuk `TrainedModel` tindakan) ke akun lain.

Akses keanggotaan dan kolaborasi

Saat mengakses sumber daya keanggotaan dan kolaborasi dalam konteks model kustom Clean Rooms, kebijakan identitas pengguna memerlukan izin untuk tindakan `cleanrooms:PassMembership` atau `cleanrooms:PassCollaboration`, atau keduanya. Semua APIs yang menerima `membershipId` membutuhkan `cleanrooms:PassMembership` izin, dan semua APIs yang menerima `collaborationId` membutuhkan `cleanrooms:PassCollaboration` izin. Kebijakan identitas sampel untuk peran yang dapat dipanggil `createTrainedModel` dalam konteks ID keanggotaan yang dapat memanggil `GetCollaborationTrainedModel` dalam konteks ID kolaborasi disediakan.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowCleanroomsMLActions",
      "Effect": "Allow",
      "Action": [
        "cleanrooms:PassCollaboration",
        "cleanrooms:PassMembership"
      ],
      "Resource": [
        "*"
      ]
    },
    {
      "Sid": "AllowMembershipAccess",
      "Effect": "Allow",
      "Action": [
        "cleanrooms:GetMembership"
      ],
      "Resource": [
        "arn:aws:cleanrooms:us-east-1:111122223333:membership/memberId"
      ]
    },
    {
      "Sid": "AllowCollaborationAccess",
      "Effect": "Allow",
      "Action": [
```

```
        "cleanrooms:GetCollaboration"
      ],
      "Resource": [
        "arn:aws:cleanrooms:us-  
east-1:111122223333:collaboration/collaborationId"
      ]
    }
  ]
}
```

Validasi kepatuhan untuk AWS Clean Rooms

Untuk mempelajari apakah an Layanan AWS berada dalam lingkup program kepatuhan tertentu, lihat [Layanan AWS di Lingkup oleh Program Kepatuhan Layanan AWS](#) dan pilih program kepatuhan yang Anda minati. Untuk informasi umum, lihat [Program AWS Kepatuhan Program AWS](#) .

Anda dapat mengunduh laporan audit pihak ketiga menggunakan AWS Artifact. Untuk informasi selengkapnya, lihat [Mengunduh Laporan di AWS Artifact](#) .

Tanggung jawab kepatuhan Anda saat menggunakan Layanan AWS ditentukan oleh sensitivitas data Anda, tujuan kepatuhan perusahaan Anda, dan hukum dan peraturan yang berlaku. Untuk informasi selengkapnya tentang tanggung jawab kepatuhan Anda saat menggunakan Layanan AWS, lihat [Dokumentasi AWS Keamanan](#).

Ketahanan di AWS Clean Rooms

Infrastruktur AWS global dibangun di sekitar AWS Wilayah dan Zona Ketersediaan. Wilayah memberikan beberapa Zona Ketersediaan yang terpisah dan terisolasi secara fisik, yang terkoneksi melalui jaringan latensi rendah, throughput tinggi, dan sangat redundan. Dengan Zona Ketersediaan, Anda dapat merancang serta mengoperasikan aplikasi dan basis data yang secara otomatis melakukan fail over di antara zona tanpa gangguan. Zona Ketersediaan memiliki ketersediaan dan toleransi kesalahan yang lebih baik, dan dapat diskalakan dibandingkan infrastruktur pusat data tunggal atau multi tradisional.

Untuk informasi selengkapnya tentang AWS Wilayah dan Availability Zone, lihat [Infrastruktur AWS Global](#).

Keamanan infrastruktur di AWS Clean Rooms

Sebagai layanan terkelola, AWS Clean Rooms dilindungi oleh keamanan jaringan AWS global. Untuk informasi tentang layanan AWS keamanan dan cara AWS melindungi infrastruktur, lihat [Keamanan AWS Cloud](#). Untuk mendesain AWS lingkungan Anda menggunakan praktik terbaik untuk keamanan infrastruktur, lihat [Perlindungan Infrastruktur dalam Kerangka Kerja](#) yang AWS Diarsiteksikan dengan Baik Pilar Keamanan.

Anda menggunakan panggilan API yang AWS dipublikasikan untuk mengakses AWS Clean Rooms melalui jaringan. Klien harus mendukung hal-hal berikut:

- Keamanan Lapisan Pengangkutan (TLS). Kami mensyaratkan TLS 1.2 dan menganjurkan TLS 1.3.
- Sandi cocok dengan sistem kerahasiaan maju sempurna (perfect forward secrecy, PFS) seperti DHE (Ephemeral Diffie-Hellman) atau ECDHE (Elliptic Curve Ephemeral Diffie-Hellman). Sebagian besar sistem modern seperti Java 7 dan versi lebih baru mendukung mode-mode ini.

Keamanan jaringan

Saat AWS Clean Rooms membaca dari bucket S3 Anda selama eksekusi kueri, lalu lintas antara AWS Clean Rooms dan Amazon S3 dirutekan dengan aman melalui jaringan pribadi. AWS Lalu lintas dalam penerbangan ditandatangani menggunakan protokol Amazon Signature Version 4 (SIGv4) dan dienkripsi menggunakan HTTPS. Lalu lintas ini diotorisasi berdasarkan peran layanan IAM yang telah Anda siapkan untuk tabel yang dikonfigurasi.

Anda dapat terhubung secara terprogram ke AWS Clean Rooms melalui titik akhir. Untuk daftar titik akhir layanan, lihat [AWS Clean Rooms titik akhir dan kuota](#) di. Referensi Umum AWS

Semua titik akhir layanan hanya HTTP. Anda dapat menggunakan titik akhir Amazon Virtual Private Cloud (VPC) jika Anda ingin terhubung dari VPC AWS Clean Rooms Anda dan tidak ingin memiliki konektivitas internet. Untuk informasi selengkapnya, lihat [Akses AWS layanan melalui AWS PrivateLinkAWS PrivateLink](#) Panduan.

Anda dapat menetapkan kebijakan IAM ke kepala sekolah IAM Anda yang menggunakan [kunci SourceVpce konteks aws](#): untuk membatasi prinsipal IAM Anda agar hanya dapat melakukan panggilan melalui titik akhir VPC dan bukan melalui AWS Clean Rooms internet.

Access AWS Clean Rooms atau AWS Clean Rooms ML menggunakan endpoint antarmuka ()AWS PrivateLink

Anda dapat menggunakan AWS PrivateLink untuk membuat koneksi pribadi antara virtual private cloud (VPC) dan AWS Clean Rooms atau AWS Clean Rooms ML. Anda dapat mengakses AWS Clean Rooms atau AWS Clean Rooms ML seolah-olah berada di VPC Anda, tanpa menggunakan gateway internet, perangkat NAT, koneksi VPN, atau koneksi. Direct Connect Instans di VPC Anda tidak memerlukan alamat IP publik untuk mengakses. AWS Clean Rooms

Anda membuat koneksi pribadi ini dengan membuat titik akhir antarmuka, yang didukung oleh AWS PrivateLink. Kami membuat antarmuka jaringan endpoint di setiap subnet yang Anda aktifkan untuk titik akhir antarmuka. Ini adalah antarmuka jaringan yang dikelola pemohon yang berfungsi sebagai titik masuk untuk lalu lintas yang ditakdirkan. AWS Clean Rooms

Untuk informasi selengkapnya, lihat [Mengakses Layanan AWS melalui AWS PrivateLink](#) di Panduan AWS PrivateLink .

Pertimbangan untuk AWS Clean Rooms

Sebelum Anda menyiapkan titik akhir antarmuka AWS Clean Rooms, tinjau [Pertimbangan](#) dalam Panduan.AWS PrivateLink

AWS Clean Rooms dan dukungan AWS Clean Rooms ML membuat panggilan ke semua tindakan API mereka melalui titik akhir antarmuka.

Kebijakan titik akhir VPC tidak didukung untuk AWS Clean Rooms atau ML. AWS Clean Rooms Secara default, akses penuh ke AWS Clean Rooms dan AWS Clean Rooms ML diizinkan melalui titik akhir antarmuka. Atau, Anda dapat mengaitkan grup keamanan dengan antarmuka jaringan titik akhir untuk mengontrol lalu lintas ke AWS Clean Rooms atau AWS Clean Rooms ML melalui titik akhir antarmuka.

Buat titik akhir antarmuka untuk AWS Clean Rooms

Anda dapat membuat titik akhir antarmuka untuk AWS Clean Rooms atau AWS Clean Rooms ML menggunakan konsol Amazon VPC atau AWS Command Line Interface ().AWS CLI Untuk informasi selengkapnya, lihat [Membuat titik akhir antarmuka](#) di AWS PrivateLink Panduan.

Buat titik akhir antarmuka untuk AWS Clean Rooms menggunakan nama layanan berikut.

```
com.amazonaws.region.cleanrooms
```

Buat endpoint antarmuka untuk AWS Clean Rooms ML menggunakan nama layanan berikut.

```
com.amazonaws.region.cleanrooms-ml
```

Jika Anda mengaktifkan DNS pribadi untuk titik akhir antarmuka, Anda dapat membuat permintaan API untuk AWS Clean Rooms menggunakan nama DNS Regional default. Misalnya, `cleanrooms-ml.us-east-1.amazonaws.com`.

Pemantauan AWS Clean Rooms

Pemantauan adalah bagian penting dari menjaga keandalan, ketersediaan, dan kinerja AWS Clean Rooms dan AWS solusi Anda yang lain. AWS menyediakan alat pemantauan berikut untuk menonton AWS Clean Rooms, melaporkan ketika ada sesuatu yang salah, dan mengambil tindakan otomatis bila perlu:

- Amazon CloudWatch Logs memungkinkan Anda memantau, menyimpan, dan mengakses file log Anda dari instans Amazon EC2, AWS CloudTrail, dan sumber lainnya. Amazon CloudWatch Logs dapat memantau informasi dalam file log dan memberi tahu Anda ketika ambang batas tertentu terpenuhi. Anda juga dapat mengarsipkan data log dalam penyimpanan yang sangat durabel. Untuk informasi selengkapnya, lihat [Panduan Pengguna Amazon CloudWatch Logs](#).

Clean Rooms MS memungkinkan pekerjaan lintas akun untuk tindakan API tertentu. Akun AWS Yang memulai pekerjaan menerima peristiwa log AWS CloudTrail audit untuk pekerjaan itu. Untuk informasi selengkapnya, lihat [Perilaku IAM untuk AWS Clean Rooms MS](#)

- AWS CloudTrail menangkap panggilan API dan peristiwa terkait yang dibuat oleh atau atas nama Anda Akun AWS dan mengirimkan file log ke bucket Amazon S3 yang Anda tentukan. Anda dapat mengidentifikasi pengguna dan akun mana yang dipanggil AWS, alamat IP sumber dari mana panggilan dilakukan, dan kapan panggilan terjadi. Untuk informasi selengkapnya, silakan lihat [Panduan Pengguna AWS CloudTrail](#).
- Amazon EventBridge adalah layanan bus acara tanpa server yang memudahkan untuk menghubungkan aplikasi Anda dengan data dari berbagai sumber. EventBridge mengirimkan aliran data real-time dari aplikasi Anda sendiri, aplikasi Software-as-a-Service (SaaS), AWS dan layanan dan rute data tersebut ke target seperti Lambda. Hal ini memungkinkan Anda memantau kejadian yang terjadi dalam layanan, dan membangun arsitektur yang didorong kejadian. Untuk informasi selengkapnya, lihat [Panduan EventBridge Pengguna Amazon](#) dan [Referensi EventBridge Acara Amazon](#).
- AWS Manajemen Penagihan dan Biaya adalah layanan yang menyediakan fitur untuk membantu Anda membayar tagihan dan mengoptimalkan biaya Anda. Amazon Web Services menagih akun Anda berdasarkan penggunaan, memastikan bahwa Anda hanya membayar untuk apa yang Anda gunakan. AWS juga memungkinkan Anda menerapkan [tag alokasi biaya](#) ke sumber daya Anda untuk melacak dan mengkategorikan biaya AWS Anda. Misalnya, di AWS Clean Rooms, Anda dapat menerapkan tag ke keanggotaan kolaborasi untuk melacak biaya per kolaborasi. Untuk informasi selengkapnya, lihat [Panduan Pengguna AWS Billing](#).

Analisis masuk AWS Clean Rooms

Analisis logging adalah fitur di AWS Clean Rooms. Saat Anda [membuat kolaborasi](#) dan mengaktifkan pencatatan Analisis, anggota dapat menyimpan log yang relevan dari kueri atau log dari pekerjaan di CloudWatch Log Amazon.

Dengan log kueri dan log pekerjaan, anggota dapat menentukan apakah kueri sesuai dengan aturan analisis dan selaras dengan perjanjian kolaborasi. Selain itu, log kueri membantu mendukung audit.

Saat opsi analisis logging diaktifkan di AWS Clean Rooms konsol, log kueri menyertakan yang berikut:

- `analysisRule`— Aturan analisis untuk tabel yang dikonfigurasi.
- `analysisTemplateArn`— Template analisis yang dijalankan (muncul tergantung pada aturan analisis).
- `collaborationId`— Pengidentifikasi unik untuk kolaborasi di mana kueri dijalankan.
- `configuredTableID`— Pengidentifikasi unik untuk tabel dikonfigurasi direferensikan dalam kueri.
- `directQueryAnalysisRulePolicy.custom.allowedAnalysis`— Template analisis diizinkan untuk berjalan pada tabel yang dikonfigurasi (muncul tergantung pada aturan analisis).
- `directQueryAnalysisRulePolicy.v1.custom.allowedAnalysisProviders`— Penyedia kueri diizinkan untuk membuat kueri (muncul tergantung pada aturan analisis).
- `errorCode`— Kode kesalahan ketika kueri gagal dijalankan dengan benar.
- `errorMessage`— Pesan kesalahan ketika kueri gagal dijalankan dengan benar.
- `eventID`— Pengidentifikasi unik untuk menjalankan kueri. Setelah 31 Agustus 2023, pengidentifikasi unik sama dengan `protectedQueryID`.
- `eventTimestamp`— Query run time.
- `parameters.parametervalue`— Nilai parameter (muncul tergantung pada teks kueri).
- `queryText`— Definisi SQL dari query run. Jika ada parameter, mereka diberi label sebagai `:parametervalue`.
- `queryValidationErrors`— Kesalahan kueri pada validasi kueri.
- `schemaName`— Nama asosiasi tabel yang dikonfigurasi direferensikan dalam kueri.
- `status`— Status eksekusi kueri.

Menerima kueri dan log pekerjaan

Anda tidak perlu melakukan tindakan apa pun di luar AWS Clean Rooms untuk menyiapkan log kueri dan log pekerjaan. AWS Clean Rooms membuat grup log untuk kolaborasi setelah setiap anggota kolaborasi [membuat keanggotaan](#).

Anggota yang dapat melakukan kueri, anggota yang dapat menjalankan kueri dan pekerjaan, anggota yang dapat menerima hasil, dan anggota yang tabel konfigurasinya direferensikan dalam kueri akan menerima log kueri atau log pekerjaan.

Anggota yang dapat melakukan kueri dan anggota yang dapat menerima hasil akan menerima log kueri untuk setiap tabel yang dikonfigurasi yang direferensikan dalam kueri. Jika mereka tidak memiliki tabel yang dikonfigurasi, mereka tidak akan dapat melihat ID tabel yang dikonfigurasi (`configuredTableID`).

Anggota yang dapat menjalankan kueri dan pekerjaan serta anggota yang dapat menerima hasil akan menerima log pekerjaan untuk setiap tabel yang dikonfigurasi yang direferensikan dalam pekerjaan. Jika mereka tidak memiliki tabel yang dikonfigurasi, mereka tidak akan dapat melihat ID tabel yang dikonfigurasi (`configuredTableID`).

Jika anggota memiliki beberapa asosiasi tabel dikonfigurasi yang direferensikan dalam kueri, mereka akan menerima log kueri untuk setiap tabel yang dikonfigurasi.

Jika anggota memiliki beberapa asosiasi tabel dikonfigurasi yang direferensikan dalam pekerjaan, mereka akan menerima log pekerjaan untuk setiap tabel yang dikonfigurasi.

Log dibuat untuk kueri yang berisi SQL yang tidak didukung dan didukung di. AWS Clean Rooms Untuk detail selengkapnya, lihat [Referensi AWS Clean Rooms SQL](#).

Log juga dibuat ketika query atau lowongan referensi dikonfigurasi tabel yang tidak terkait dengan kolaborasi.

Log mungkin berisi informasi tentang SQL yang salah.

Kueri dan log pekerjaan menunjukkan status kueri tetapi tidak melaporkan apakah output kueri dikirim. Mereka mengkonfirmasi bahwa permintaan atau pekerjaan telah dikirimkan oleh anggota yang dapat menanyakan. Log kueri juga mengonfirmasi bahwa kueri berisi SQL yang didukung AWS Clean Rooms dan referensi tabel yang dikonfigurasi terkait dengan kolaborasi.

Example

Misalnya, log tidak dihasilkan jika kueri dibatalkan setelah AWS Clean Rooms memvalidasi kepatuhannya dengan aturan analisis dan selama pemrosesan kueri.

Jika Anda menghapus grup log, Anda harus membuat ulang grup log secara manual dengan nama grup log yang sama (ID kolaborasi kolaborasi). Atau, Anda dapat mematikan dan mengaktifkan log dalam keanggotaan Anda.

Untuk informasi selengkapnya tentang cara mengaktifkan pencatatan analisis, lihat [Menciptakan kolaborasi](#).

Untuk informasi selengkapnya tentang CloudWatch Log Amazon, lihat [Panduan Pengguna CloudWatch Log Amazon](#).

Tindakan yang disarankan untuk kueri dan log pekerjaan

Kami menyarankan agar anggota secara berkala mengambil tindakan berikut:

- Untuk memverifikasi bahwa kueri dan pekerjaan cocok dengan kasus penggunaan atau kueri yang disepakati untuk kolaborasi, tinjau kueri dan pekerjaan yang dijalankan dalam kolaborasi.

Untuk informasi selengkapnya tentang cara melihat kueri terbaru, lihat [Melihat kueri terbaru](#).

Untuk informasi selengkapnya tentang cara melihat lowongan terbaru, lihat [Melihat pekerjaan terbaru](#).

- Untuk memverifikasi bahwa kolom tabel yang dikonfigurasi cocok dengan apa yang telah disepakati untuk kolaborasi, tinjau kolom tabel yang dikonfigurasi yang digunakan dalam aturan analisis anggota kolaborasi dan dalam kueri.

Untuk informasi selengkapnya tentang cara melihat kolom yang dikonfigurasi, lihat [Melihat tabel dan aturan analisis](#).

Pemantauan terperinci dengan CloudWatch in AWS Clean Rooms

Pemantauan terperinci adalah fitur keikutsertaan di AWS Clean Rooms. Saat Anda [membuat kolaborasi](#) dan mengaktifkan Pemantauan terperinci, anggota yang dapat menjalankan kueri dan pembayar yang dikonfigurasi dapat memilih untuk menerima metrik pengamatan terperinci di akun mereka. CloudWatch

Dengan pemantauan terperinci, anggota dapat memantau kinerja kueri, melacak pemanfaatan sumber daya, dan membuat dasbor untuk wawasan operasional. Metrik ini membantu perencanaan kapasitas, analisis biaya, dan pemecahan masalah kinerja.

Note

Ketersediaan metrik bervariasi berdasarkan sumber data tabel yang dikonfigurasi:

- Untuk analisis yang mereferensikan tabel yang dikonfigurasi dengan Athena sebagai sumber data, BytesRead metrik QueryFileScanned dan selalu mengembalikan 0.
- Untuk analisis yang referensi dikonfigurasi tabel dengan Snowflake sebagai sumber data,, BytesReadQueryFileScanned, dan RecordsRead metrik selalu mengembalikan 0.
- Untuk analisis yang mereferensikan beberapa tabel yang dikonfigurasi dengan sumber data yang berbeda (misalnya, tabel yang didukung Snowflake dan S3), BytesRead dan QueryFileScanned mengembalikan metrik hanya untuk AWS Glue tabel yang didukung S3. AWS Glue RecordsReadmengembalikan metrik untuk tabel yang didukung S3 AWS Glue dan tabel Athena.

Tipe metrik

AWS Clean Roomsmenerbitkan dua kategori metrik untuk: CloudWatch

Metrik pemantauan terperinci agregat

Metrik ini memberikan visibilitas tingkat agregat di seluruh analisis:

AnalysisRuntime

Menunjukkan total durasi analisis dalam milidetik.

CRPUConsumed

CRPU (Unit Pemrosesan Ruang Bersih) dikonsumsi untuk analisis tertentu. Untuk informasi selengkapnya tentang perhitungan CRPU, lihat [AWS Clean RoomsHarga](#).

ConcurrentVCPUs

VCPU bersamaan digunakan pada titik waktu tertentu (terkait dengan kuota layanan).

AnalysesCount

Jumlah analisis berjalan selama periode waktu untuk dimensi tertentu.

ConcurrentQueries

Jumlah kueri bersamaan pada titik waktu tertentu (terkait dengan kuota layanan).

Metrik pemantauan terperinci tingkat kueri

Metrik ini memberikan wawasan terperinci tentang eksekusi kueri individual:

BytesRead

Total data (dalam byte) dibaca dari tabel sumber untuk analisis.

BytesWritten

Total data (dalam byte) ditulis untuk analisis.

RecordsRead

Jumlah total catatan yang dibaca dari tabel sumber untuk analisis.

RecordsWritten

Jumlah total catatan yang dikembalikan dalam hasil analisis.

QueryFileScanned

Jumlah total file data yang dipindai untuk analisis.

Memory utilization

Sumber daya memori yang digunakan untuk analisis.

vCPU utilization

Sumber daya vCPU digunakan untuk analisis.

Disk utilization

Sumber daya penyimpanan yang digunakan untuk analisis.

Dimensi metrik

Metrik diatur oleh dimensi berikut:

- Status Analisis: Selesai
- Jenis Analisis: SparkSQL
- Keanggotaan: ID keanggotaan
- Kueri: ID kueri individual (untuk metrik tingkat kueri)
- Akun/Wilayah: Anda Akun AWS dan Wilayah

Semua metrik dipublikasikan ke `AWS/CleanRooms` namespace di CloudWatch

Siapa yang dapat mengakses metrik

Metrik pemantauan terperinci tersedia untuk:

- Pelari kueri: Anggota yang dapat menjalankan kueri dalam kolaborasi
- Pembayar yang dikonfigurasi: Anggota yang ditunjuk untuk membayar biaya eksekusi kueri

Anggota yang hanya menyumbangkan data tidak dapat mengakses metrik pemantauan terperinci.

Penetapan harga

AWS Clean Rooms tidak mengenakan biaya untuk menerbitkan metrik ke CloudWatch. Namun, CloudWatch biaya standar berlaku untuk penyimpanan metrik, penggunaan dasbor, dan alarm. Untuk informasi selengkapnya, lihat [CloudWatch Harga Amazon](#).

Pencatatan panggilan AWS Clean Rooms API menggunakan AWS CloudTrail

AWS Clean Rooms terintegrasi dengan AWS CloudTrail, layanan yang menyediakan catatan tindakan yang diambil oleh pengguna, peran, atau Layanan AWS dalam AWS Clean Rooms. CloudTrail menangkap semua panggilan API untuk AWS Clean Rooms sebagai peristiwa. Panggilan yang diambil termasuk panggilan dari AWS Clean Rooms konsol dan panggilan kode ke operasi AWS Clean Rooms API. Jika Anda membuat jejak, Anda dapat mengaktifkan pengiriman CloudTrail acara secara berkelanjutan ke bucket Amazon S3, termasuk acara untuk AWS Clean Rooms. Jika Anda tidak mengonfigurasi jejak, Anda masih dapat melihat peristiwa terbaru di CloudTrail konsol dalam Riwayat acara. Dengan menggunakan informasi yang dikumpulkan oleh CloudTrail, Anda dapat menentukan permintaan yang dibuat AWS Clean Rooms, alamat IP dari mana permintaan dibuat, siapa yang membuat permintaan, kapan dibuat, dan detail tambahan.

Untuk mempelajari selengkapnya CloudTrail, lihat [Panduan AWS CloudTrail Pengguna](#).

AWS Clean Rooms informasi di CloudTrail

CloudTrail diaktifkan pada Akun AWS saat Anda membuat akun. Ketika aktivitas terjadi AWS Clean Rooms, aktivitas tersebut dicatat dalam suatu CloudTrail peristiwa bersama dengan Layanan AWS peristiwa lain dalam sejarah Peristiwa. Anda dapat melihat, mencari, dan mengunduh peristiwa terbaru di Akun AWS Anda. Untuk informasi selengkapnya, lihat [Melihat peristiwa dengan Riwayat CloudTrail acara](#).

Untuk catatan acara yang sedang berlangsung di Anda Akun AWS, termasuk acara untuk AWS Clean Rooms, buat jejak. Jejak memungkinkan CloudTrail untuk mengirimkan file log ke bucket Amazon S3. Secara default, saat Anda membuat jejak di konsol, jejak tersebut berlaku untuk semua Wilayah AWS. Jejak mencatat peristiwa dari semua Wilayah di AWS partisi dan mengirimkan file log ke bucket Amazon S3 yang Anda tentukan. Selain itu, Anda dapat mengonfigurasi lainnya Layanan AWS untuk menganalisis lebih lanjut dan menindaklanjuti data peristiwa yang dikumpulkan dalam CloudTrail log. Untuk informasi selengkapnya, lihat berikut:

- [Gambaran umum untuk membuat jejak](#)
- [CloudTrail layanan dan integrasi yang didukung](#)
- [Mengonfigurasi notifikasi Amazon SNS untuk CloudTrail](#)
- [Menerima file CloudTrail log dari beberapa Wilayah](#)
- [Menerima file CloudTrail log dari beberapa akun](#)

Semua AWS Clean Rooms tindakan dicatat oleh CloudTrail dan didokumentasikan dalam [Referensi AWS Clean Rooms API](#).

Setiap entri peristiwa atau log berisi informasi tentang entitas yang membuat permintaan tersebut. Informasi identitas tersebut membantu Anda menentukan hal berikut:

- Apakah permintaan tersebut dibuat dengan kredensial pengguna root atau pengguna IAM.
- Apakah permintaan tersebut dibuat dengan kredensial keamanan sementara untuk peran atau pengguna gabungan.
- Apakah permintaan tersebut dibuat oleh Layanan AWS lain.

Untuk informasi selengkapnya, lihat [Elemen userIdentity CloudTrail](#).

Memahami entri file AWS Clean Rooms log

Trail adalah konfigurasi yang memungkinkan pengiriman peristiwa sebagai file log ke bucket Amazon S3 yang Anda tentukan. CloudTrail file log berisi satu atau lebih entri log. Peristiwa mewakili permintaan tunggal dari sumber mana pun dan mencakup informasi tentang tindakan yang diminta, tanggal dan waktu tindakan, parameter permintaan, dan sebagainya. CloudTrail file log bukanlah jejak tumpukan yang diurutkan dari panggilan API publik, jadi file tersebut tidak muncul dalam urutan tertentu.

Contoh AWS Clean Rooms CloudTrail peristiwa

Contoh-contoh berikut menunjukkan CloudTrail peristiwa untuk:

Topik

- [StartProtectedQuery \(berhasil\)](#)
- [StartProtectedQuery\(gagal\)](#)

StartProtectedQuery (berhasil)

```
{
  "eventVersion": "1.08",
  "userIdentity": {
    "type": "AssumedRole",
    "principalId": "EXAMPLE_PRINCIPAL_ID",
    "arn": "arn:aws:sts::123456789012:assumed-role/query-runner/jdoe",
    "accountId": "123456789012",
    "accessKeyId": "EXAMPLE_KEY_ID",
    "sessionContext": {
      "sessionIssuer": {
        "type": "Role",
        "principalId": "EXAMPLE_PRINCIPAL_ID",
        "arn": "arn:aws:iam::123456789012:role/query-runner",
        "accountId": "123456789012",
        "userName": "query-runner"
      },
      "webIdFederationData": {},
      "attributes": {
        "creationDate": "2023-04-07T19:34:32Z",
        "mfaAuthenticated": "false"
      }
    }
  }
}
```

```

    }
  },
  "eventTime": "2023-04-07T19:53:32Z",
  "eventSource": "cleanrooms.amazonaws.com",
  "eventName": "StartProtectedQuery",
  "awsRegion": "us-east-2",
  "sourceIPAddress": "203.0.113.1",
  "userAgent": "aws-internal/3",
  "requestParameters": {
    "resultConfiguration": {
      "outputConfiguration": {
        "s3": {
          "resultFormat": "CSV",
          "bucket": "cleanrooms-queryresults-jdoe-test",
          "keyPrefix": "test"
        }
      }
    }
  },
  "sqlParameters": "****",
  "membershipIdentifier": "a1b2c3d4-5678-90ab-cdef-EXAMPLE11111",
  "type": "SQL"
},
"responseElements": {
  "Access-Control-Expose-Headers": "x-amzn-RequestId,x-amzn-ErrorType,x-amzn-ErrorMessage,Date",
  "protectedQuery": {
    "createTime": 1680897212.279,
    "id": "f5988bf1-771a-4141-82a8-26fcc4e41c9f",
    "membershipArn": "arn:aws:cleanrooms:us-east-2:123456789012:membership/a1b2c3d4-5678-90ab-cdef-EXAMPLE11111",
    "membershipId": "a1b2c3d4-5678-90ab-cdef-EXAMPLE11111",
    "resultConfiguration": {
      "outputConfiguration": {
        "s3": {
          "bucket": "cleanrooms-queryresults-jdoe-test",
          "keyPrefix": "test",
          "resultFormat": "CSV"
        }
      }
    }
  },
  "sqlParameters": "****",
  "status": "SUBMITTED"
}
},

```

```

"requestID": "7464211b-2277-4b55-9723-fb4f259aefd2",
"eventID": "f7610f5e-74b9-420f-ae43-206571ebcbf7",
"readOnly": false,
"eventType": "AwsApiCall",
"managementEvent": true,
"recipientAccountId": "123456789012",
"eventCategory": "Management"
}

```

StartProtectedQuery(gagal)

```

{
  "eventVersion": "1.08",
  "userIdentity": {
    "type": "AssumedRole",
    "principalId": "EXAMPLE_PRINCIPAL_ID",
    "arn": "arn:aws:sts::123456789012:assumed-role/query-runner/jdoe",
    "accountId": "123456789012",
    "accessKeyId": "EXAMPLE_KEY_ID",
    "sessionContext": {
      "sessionIssuer": {
        "type": "Role",
        "principalId": "EXAMPLE_PRINCIPAL_ID",
        "arn": "arn:aws:iam::123456789012:role/query-runner",
        "accountId": "123456789012",
        "userName": "query-runner"
      },
      "webIdFederationData": {},
      "attributes": {
        "creationDate": "2023-04-07T19:34:32Z",
        "mfaAuthenticated": "false"
      }
    }
  },
  "eventTime": "2023-04-07T19:47:27Z",
  "eventSource": "cleanrooms.amazonaws.com",
  "eventName": "StartProtectedQuery",
  "awsRegion": "us-east-2",
  "sourceIPAddress": "203.0.113.1",
  "userAgent": "aws-internal/3",
  "errorCode": "ValidationException",
  "requestParameters": {
    "resultConfiguration": {

```



```

        "outputConfiguration": {
            "s3": {
                "resultFormat": "CSV",
                "bucket": "cleanrooms-queryresults-jdoe-test",
                "keyPrefix": "test"
            }
        },
        "sqlParameters": "****",
        "membershipIdentifier": "a1b2c3d4-5678-90ab-cdef-EXAMPLE11111",
        "type": "SQL"
    },
    "responseElements": {
        "Access-Control-Expose-Headers": "x-amzn-RequestId,x-amzn-ErrorType,x-amzn-ErrorMessage,Date",
        "message": "Column(s) [identifier] is not allowed in select"
    },
    "requestID": "e29f9f74-8299-4a83-9d18-5ddce7302f07",
    "eventID": "c8ee3498-8e4e-44b5-87e4-ab9477e56eb5",
    "readOnly": false,
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "recipientAccountId": "123456789012",
    "eventCategory": "Management"
}

```

Mengintegrasikan AWS Clean Rooms ke dalam aplikasi berbasis peristiwa menggunakan Amazon EventBridge

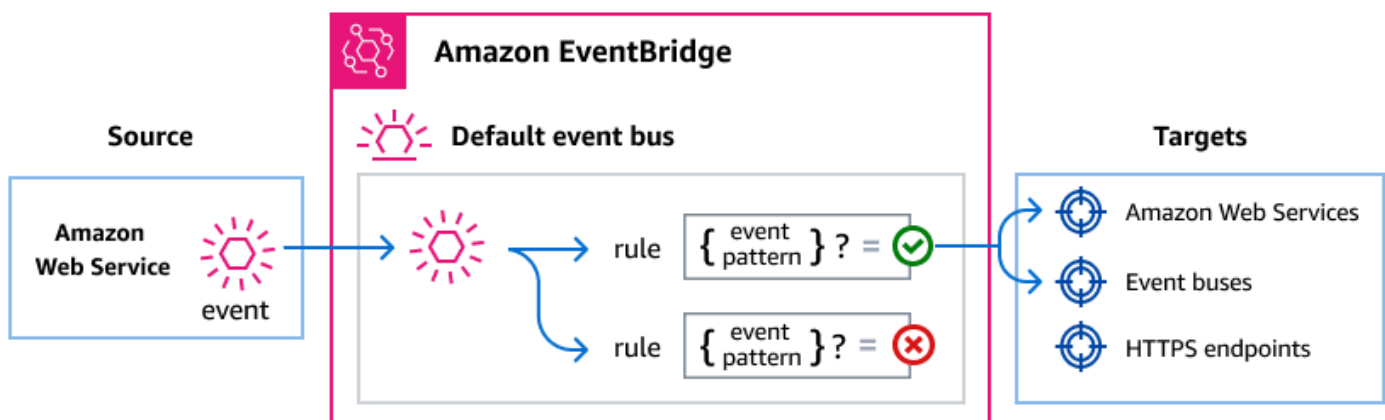
Anda dapat memasukkan AWS Clean Rooms ke dalam aplikasi berbasis peristiwa (EDAs) yang menggunakan peristiwa yang terjadi AWS Clean Rooms untuk berkomunikasi antara komponen aplikasi dan memulai proses hilir. Anda melakukan ini dengan menggunakan Amazon EventBridge untuk merutekan acara dari AWS Clean Rooms ke komponen perangkat lunak lainnya. Amazon EventBridge adalah layanan tanpa server yang menggunakan peristiwa untuk menghubungkan komponen aplikasi bersama-sama, sehingga memudahkan Anda untuk mengintegrasikan AWS layanan seperti AWS Clean Rooms ke dalam arsitektur berbasis peristiwa tanpa kode dan operasi tambahan.

Arsitektur berbasis peristiwa adalah gaya membangun sistem perangkat lunak yang digabungkan secara longgar yang bekerja sama dengan memancarkan dan menanggapi peristiwa. Dalam model ini, peristiwa mewakili perubahan dalam sumber daya atau lingkungan.

Berikut cara EventBridge bekerja dengan AWS Clean Rooms:

Seperti banyak AWS layanan, AWS Clean Rooms menghasilkan dan mengirim acara ke bus acara EventBridge default. Bus acara adalah router yang menerima acara dan merutekan mereka ke tujuan, atau target, yang Anda tentukan. Target dapat mencakup AWS layanan lain, aplikasi khusus, dan aplikasi mitra SaaS.

EventBridge rute acara sesuai dengan aturan yang Anda buat di bus acara. Untuk setiap aturan, Anda menentukan filter, atau pola acara, untuk memilih hanya peristiwa yang Anda inginkan. Setiap kali acara dikirim ke bus acara, EventBridge bandingkan dengan setiap aturan. Jika acara cocok dengan aturan, EventBridge rutekan acara ke target yang ditentukan.



Misalnya, Anda ingin tahu setiap kali AWS Clean Rooms kolaborasi baru dibuat di akun Anda. Anda dapat membuat aturan pada bus acara default. Dalam aturan, Anda akan membuat pola acara yang menetapkan peristiwa dari AWS Clean Rooms bernama **Collaboration Created**. Setiap kali EventBridge menerima acara yang cocok dengan properti tersebut, itu akan merutekan acara ke alur kerja yang ditentukan.

AWS Clean Rooms acara

AWS layanan dapat mengirim acara langsung ke bus acara EventBridge default. Selain itu, AWS CloudTrail mengirimkan acara yang berasal dari berbagai AWS layanan ke EventBridge juga. Peristiwa ini dapat mencakup panggilan API, login dan tindakan konsol, peristiwa layanan, dan

CloudTrail Wawasan. Untuk informasi selengkapnya, lihat [acara AWS layanan yang disampaikan melalui AWS CloudTrail](#) Panduan EventBridge Pengguna.

Untuk daftar lengkap AWS Clean Rooms acara yang dikirim ke EventBridge, lihat AWS Clean Rooms topik di [Referensi EventBridge Acara](#).

Jenis detail acara	Deskripsi
Template Analisis Dibuat	Pemilik Template Analisis dan semua anggota aktif kolaborasi diberi tahu saat Template Analisis Dibuat.
Template Analisis Diperbarui	Pemilik Template Analisis dan semua anggota aktif kolaborasi yang memiliki visibilitas ke pembaruan akan diberi tahu saat Template Analisis Diperbarui.
Template Analisis Dihapus	Pemilik Template Analisis dan semua anggota aktif kolaborasi akan diberi tahu saat Template Analisis Dihapus.
Kolaborasi Dibuat	Pemilik Kolaborasi diberi tahu saat Kolaborasi Dibuat.
Kolaborasi Diperbarui	Pemilik Kolaborasi dan semua anggota aktif kolaborasi yang memiliki visibilitas ke pembaruan akan diberi tahu saat Kolaborasi Diperbarui.
Permintaan Perubahan Kolaborasi Dibuat	Pemilik Kolaborasi dan semua anggota aktif kolaborasi akan diberi tahu saat Permintaan Perubahan Kolaborasi Dibuat.
Permintaan Perubahan Kolaborasi Disetujui	Pemilik Kolaborasi dan semua anggota aktif kolaborasi akan diberi tahu saat Permintaan Perubahan Kolaborasi Disetujui.
Permintaan Perubahan Kolaborasi Dibatalkan	Pemilik Kolaborasi dan semua anggota aktif kolaborasi akan diberi tahu saat Permintaan Perubahan Kolaborasi Dibatalkan.
Permintaan Perubahan Kolaborasi Berkomitmen	Pemilik Kolaborasi dan semua anggota aktif kolaborasi diberi tahu saat Permintaan Perubahan Kolaborasi dilakukan.
Asosiasi Tabel Dikonfigurasi Dibuat	Pemilik Asosiasi Tabel yang Dikonfigurasi dan semua anggota aktif kolaborasi akan diberi tahu saat Asosiasi Tabel Terkonfigurasi Dibuat.

Jenis detail acara	Deskripsi
Asosiasi Tabel Dikonfigurasi Diperbarui	Pemilik Asosiasi Tabel Terkonfigurasi dan semua anggota aktif kolaborasi yang memiliki visibilitas ke dalam pembaruan akan diberi tahu saat Asosiasi Tabel Terkonfigurasi Diperbarui.
Asosiasi Tabel yang Dikonfigurasi Dihapus	Pemilik Asosiasi Tabel yang Dikonfigurasi dan semua anggota aktif kolaborasi akan diberi tahu saat Asosiasi Tabel Terkonfigurasi Dihapus.
Aturan Analisis Asosiasi Tabel yang Dikonfigurasi Dibuat	Pemilik Aturan Analisis Asosiasi Tabel yang Dikonfigurasi dan semua anggota aktif kolaborasi diberi tahu saat Aturan Analisis Asosiasi Tabel yang Dikonfigurasi Dibuat.
Aturan Analisis Asosiasi Tabel yang Dikonfigurasi Diperbarui	Pemilik Aturan Analisis Asosiasi Tabel yang Dikonfigurasi dan semua anggota aktif kolaborasi yang memiliki visibilitas ke dalam pembaruan akan diberi tahu saat Aturan Analisis Asosiasi Tabel yang Dikonfigurasi Diperbarui.
Aturan Analisis Asosiasi Tabel yang Dikonfigurasi Dihapus	Pemilik Aturan Analisis Asosiasi Tabel yang Dikonfigurasi dan semua anggota aktif kolaborasi diberi tahu saat Aturan Analisis Asosiasi Tabel yang Dikonfigurasi Dihapus.
Tabel Pemetaan Id Dibuat	Pemilik Tabel Pemetaan Id dan semua anggota aktif kolaborasi diberi tahu saat Tabel Pemetaan Id Dibuat.
Tabel Pemetaan Id Diperbarui	Pemilik Tabel Pemetaan Id dan semua anggota aktif kolaborasi yang memiliki visibilitas ke dalam pembaruan akan diberi tahu saat Tabel Pemetaan Id Diperbarui.
Tabel Pemetaan Id Dihapus	Pemilik Tabel Pemetaan Id dan semua anggota aktif kolaborasi akan diberi tahu saat Tabel Pemetaan Id Dihapus.
Asosiasi Namespace Id Dibuat	Pemilik Asosiasi Namespace Id dan semua anggota aktif kolaborasi akan diberi tahu saat Asosiasi Namespace Id Dibuat.
Asosiasi Namespace Id Diperbarui	Pemilik Asosiasi Namespace Id dan semua anggota aktif kolaborasi yang memiliki visibilitas ke dalam pembaruan akan diberi tahu saat Asosiasi Namespace Id Diperbarui.

Jenis detail acara	Deskripsi
Asosiasi Namespace Id Dihapus	Pemilik Asosiasi Namespace Id dan semua anggota aktif kolaborasi akan diberi tahu saat Asosiasi Namespace Id Dihapus.
Diundang Untuk Kolaborasi	Anggota yang diundang diberitahu ketika mereka diundang ke Kolaborasi.
Keanggotaan Dibuat	Pemilik Keanggotaan dan semua anggota aktif kolaborasi diberi tahu saat Keanggotaan Dibuat.
Keanggotaan Diperbarui	Pemilik Keanggotaan diberi tahu ketika Keanggotaan Diperbarui, kecuali keanggotaan telah dihapus dari kolaborasi dalam hal ini semua anggota aktif kolaborasi diberitahu.
Keanggotaan Dihapus	Pemilik Keanggotaan dan semua anggota aktif kolaborasi akan diberi tahu saat Keanggotaan Dihapus.
Protected Job Dikirim	Job Runner, Job Payer, dan Results Receiver dari Pekerjaan yang Dilindungi diberi tahu ketika Job yang Dilindungi Diserahkan.
Protected Job Dimulai	Job Runner, Job Payer dan Results Receiver dari Protected Job akan diberi tahu ketika Job yang Dilindungi Dimulai.
Pembatalan Job yang Dilindungi	Job Runner, Job Payer dan Results Receiver dari Protected Job akan diberi tahu ketika Job yang Dilindungi dibatalkan.
Protected Job Dibatalkan	Job Runner, Job Payer dan Results Receiver dari Protected Job akan diberi tahu ketika Job yang Dilindungi Dibatalkan.
Protected Job Berhasil	Job Runner, Job Payer dan Results Receiver dari Protected Job diberi tahu ketika Job yang Dilindungi telah Berhasil.
Pekerjaan yang Dilindungi Gagal	Job Runner, Job Payer dan Results Receiver dari Pekerjaan yang Dilindungi diberi tahu ketika Job yang Dilindungi telah Gagal.

Jenis detail acara	Deskripsi
Kueri yang Dilindungi Dikirim	Query Runner, Query Payer, dan Results Receiver dari Query yang Dilindungi akan diberi tahu saat Kueri yang Dilindungi Dikirim.
Kueri yang Dilindungi Dimulai	Query Runner, Query Payer, dan Results Receiver dari Query yang Dilindungi akan diberi tahu saat Kueri Terproteksi Dimulai.
Pembatalan Kueri Terproteksi	Query Runner, Query Payer, dan Results Receiver dari Query yang Dilindungi akan diberi tahu saat Kueri yang Dilindungi Membatalkan.
Kueri Terproteksi Dibatalkan	Query Runner, Query Payer, dan Results Receiver dari Query yang Dilindungi akan diberi tahu saat Kueri yang Dilindungi Dibatalkan.
Kueri yang Dilindungi Berhasil	Query Runner, Query Payer, dan Results Receiver dari Query yang Dilindungi akan diberi tahu saat Kueri yang Dilindungi telah Berhasil.
Kueri yang Dilindungi Gagal	Query Runner, Query Payer, dan Results Receiver dari Query yang Dilindungi akan diberi tahu saat Kueri yang Dilindungi telah Gagal.
Waktu Kueri yang Dilindungi Habis	Query Runner, Query Payer, dan Results Receiver dari Query yang Dilindungi akan diberi tahu saat Kueri yang Dilindungi telah Habis Waktu.

Perutean AWS Clean Rooms acara menggunakan EventBridge

Untuk memiliki EventBridge rute AWS Clean Rooms acara ke target, Anda harus membuat aturan. Setiap aturan berisi pola acara, yang EventBridge cocok dengan setiap acara yang diterima di bus acara. Jika data peristiwa cocok dengan pola peristiwa yang ditentukan, EventBridge rute acara tersebut ke target aturan.

Untuk petunjuk komprehensif tentang cara membuat aturan bus acara, lihat [Membuat aturan yang bereaksi terhadap peristiwa](#) di Panduan EventBridge Pengguna.

Membuat pola acara yang cocok dengan AWS Clean Rooms acara

Setiap pola acara adalah objek JSON yang berisi:

- (Opsional): `source` Atribut yang mengidentifikasi layanan yang mengirim acara. Untuk AWS Clean Rooms acara, sumbernya adalah `aws.cleanrooms`.
- (Opsional): `detail-type` Atribut yang berisi array nama acara untuk dicocokkan.
- (Opsional): `detail` Atribut yang berisi data acara lain yang cocok.

Misalnya, pola acara berikut cocok dengan semua peristiwa yang diperbarui Asosiasi Namespace Id tempat kolaborasi dihapus: AWS Clean Rooms

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Id Namespace Association Updated"],
  "detail": {
    "status": ["COLLABORATION_DELETED"]
  }
}
```

Untuk informasi selengkapnya tentang penulisan pola acara, lihat [Pola acara](#) di Panduan EventBridge Pengguna.

AWS Clean Rooms referensi detail acara

Semua peristiwa dari AWS layanan memiliki seperangkat bidang umum yang berisi metadata tentang acara tersebut, seperti AWS layanan yang merupakan sumber acara, waktu acara dibuat, akun dan wilayah tempat acara berlangsung, dan lainnya. Untuk definisi bidang umum ini, lihat [Struktur acara](#) di Referensi EventBridge Acara Amazon.

Selain itu, setiap acara memiliki `detail` bidang yang berisi data khusus untuk peristiwa tertentu. Referensi di bawah ini mendefinisikan bidang detail untuk berbagai AWS Clean Rooms acara.

Saat menggunakan EventBridge untuk memilih dan mengelola AWS Clean Rooms acara, penting untuk mengingat hal berikut:

- `sourceBidang` untuk semua acara dari AWS Clean Rooms diatur ke `aws.cleanrooms`.
- `detail-typeBidang` menentukan jenis acara.

Misalnya, Collaboration Created.

- detailBidang berisi data yang spesifik untuk peristiwa tertentu.

Untuk informasi tentang membuat pola peristiwa yang memungkinkan aturan untuk mencocokkan AWS Clean Rooms peristiwa, lihat [Pola peristiwa](#) di Panduan EventBridge Pengguna Amazon.

Template Analisis Acara yang dibuat

Di bawah ini adalah bidang detail untuk Analysis Template Created acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Analysis Template Created"]
}
```

Template Analisis Acara yang diperbarui

Di bawah ini adalah bidang detail untuk Analysis Template Updated acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Analysis Template Updated"]
}
```

Template Analisis Acara yang dihapus

Di bawah ini adalah bidang detail untuk Analysis Template Deleted acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Analysis Template Deleted"]
}
```

Kolaborasi Acara yang dibuat

Di bawah ini adalah bidang detail untuk Collaboration Created acara tersebut.

```
{
```



```
"source": ["aws.cleanrooms"],  
"detail-type": ["Collaboration Created"]  
}
```

Kolaborasi Acara yang diperbarui

Di bawah ini adalah bidang detail untuk Collaboration Updated acara tersebut.

```
{  
  "source": ["aws.cleanrooms"],  
  "detail-type": ["Collaboration Updated"]  
}
```

Permintaan Perubahan Kolaborasi Acara yang dibuat

Di bawah ini adalah bidang detail untuk Collaboration Change Request Created acara tersebut.

```
{  
  "source": ["aws.cleanrooms"],  
  "detail-type": ["Collaboration Change Request Created"]  
}
```

Permintaan Perubahan Kolaborasi Acara yang disetujui

Di bawah ini adalah bidang detail untuk Collaboration Change Request Approved acara tersebut.

```
{  
  "source": ["aws.cleanrooms"],  
  "detail-type": ["Collaboration Change Request Approved"]  
}
```

Permintaan Perubahan Kolaborasi Acara dibatalkan

Di bawah ini adalah bidang detail untuk Collaboration Change Request Cancelled acara tersebut.

```
{  
  "source": ["aws.cleanrooms"],
```

```
"detail-type": ["Collaboration Change Request Cancelled"]
}
```

Permintaan Perubahan Kolaborasi Acara yang berkomitmen

Di bawah ini adalah bidang detail untuk Collaboration Change Request Committed acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Collaboration Change Request Committed"]
}
```

Asosiasi Tabel Dikonfigurasi Acara yang dibuat

Di bawah ini adalah bidang detail untuk Configured Table Association Created acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Configured Table Association Created"]
}
```

Asosiasi Tabel Dikonfigurasi Acara yang diperbarui

Di bawah ini adalah bidang detail untuk Configured Table Association Updated acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Configured Table Association Updated"]
}
```

Asosiasi Tabel yang Dikonfigurasi Acara yang dihapus

Di bawah ini adalah bidang detail untuk Configured Table Association Deleted acara tersebut.

```
{
```

```
"source": ["aws.cleanrooms"],
"detail-type": ["Configured Table Association Deleted"]
}
```

Aturan Analisis Asosiasi Tabel Dikonfigurasi Acara yang dibuat

Di bawah ini adalah bidang detail untuk Configured Table Association Analysis Rule Created acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Configured Table Association Analysis Rule Created"]
}
```

Aturan Analisis Asosiasi Tabel Dikonfigurasi Acara yang diperbarui

Di bawah ini adalah bidang detail untuk Configured Table Association Analysis Rule Updated acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Configured Table Association Analysis Rule Updated"]
}
```

Aturan Analisis Asosiasi Tabel Dikonfigurasi Acara yang dihapus

Di bawah ini adalah bidang detail untuk Configured Table Association Analysis Rule Deleted acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Configured Table Association Analysis Rule Deleted"]
}
```

Tabel Pemetaan Id Acara yang dibuat

Di bawah ini adalah bidang detail untuk Id Mapping Table Created acara tersebut.

```
{
```

```
"source": ["aws.cleanrooms"],  
"detail-type": ["Id Mapping Table Created"]  
}
```

Tabel Pemetaan Id Acara yang diperbarui

Di bawah ini adalah bidang detail untuk Id Mapping Table Updated acara tersebut.

```
{  
  "source": ["aws.cleanrooms"],  
  "detail-type": ["Id Mapping Table Updated"]  
}
```

Tabel Pemetaan Id Acara yang dihapus

Di bawah ini adalah bidang detail untuk Id Mapping Table Deleted acara tersebut.

```
{  
  "source": ["aws.cleanrooms"],  
  "detail-type": ["Id Mapping Table Deleted"]  
}
```

Id Namespace Association Membuat acara

Di bawah ini adalah bidang detail untuk Id Namespace Association Created acara tersebut.

```
{  
  "source": ["aws.cleanrooms"],  
  "detail-type": ["Id Namespace Association Created"]  
}
```

Acara yang diperbarui Asosiasi Namespace Id

Di bawah ini adalah bidang detail untuk Id Namespace Association Updated acara tersebut.

```
{  
  "source": ["aws.cleanrooms"],  
  "detail-type": ["Id Namespace Association Updated"]  
}
```

Note

Saat kolaborasi dihapus, acara akan ditangkap dalam acara Id Namespace Association Updated di detail.status bidang.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Id Namespace Association Updated"],
  "detail": {
    "status": ["COLLABORATION_DELETED"]
  }
}
```

Asosiasi Namespace Id Acara yang dihapus

Di bawah ini adalah bidang detail untuk Id Namespace Association Deleted acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Id Namespace Association Deleted"]
}
```

Diundang Ke Acara Kolaborasi

Di bawah ini adalah bidang detail untuk Invited To Collaboration acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Invited To Collaboration"]
}
```

Keanggotaan Acara yang dibuat

Di bawah ini adalah bidang detail untuk Membership Created acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Membership Created"]
}
```

Keanggotaan Acara yang diperbarui

Di bawah ini adalah bidang detail untuk Membership Updated acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Membership Updated"]
}
```

Keanggotaan Acara yang dihapus

Di bawah ini adalah bidang detail untuk Membership Deleted acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Membership Deleted"]
}
```

Acara Protected Job Diserahkan

Di bawah ini adalah bidang detail untuk Protected Job Submitted acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Job Submitted"]
}
```

Acara Protected Job Started

Di bawah ini adalah bidang detail untuk Protected Job Started acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Job Started"]
}
```

Acara Pembatalan Pekerjaan yang Dilindungi

Di bawah ini adalah bidang detail untuk Protected Job Cancelling acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Job Cancelling"]
}
```

Acara Protected Job Cancelled

Di bawah ini adalah bidang detail untuk Protected Job Cancelled acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Job Cancelled"]
}
```

Acara Protected Job Berhasil

Di bawah ini adalah bidang detail untuk Protected Job Succeeded acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Job Succeeded"]
}
```

Acara Protected Job Failed

Di bawah ini adalah bidang detail untuk Protected Job Failed acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Job Failed"]
}
```

Pertanyaan yang Dilindungi Acara yang dikirimkan

Di bawah ini adalah bidang detail untuk Protected Query Submitted acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Query Submitted"]
}
```

Permintaan Terproteksi Memulai acara

Di bawah ini adalah bidang detail untuk Protected Query Started acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Query Started"]
}
```

Acara Pembatalan Kueri Terproteksi

Di bawah ini adalah bidang detail untuk Protected Query Cancelling acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Query Cancelling"]
}
```

Permintaan Terproteksi Acara dibatalkan

Di bawah ini adalah bidang detail untuk Protected Query Cancelled acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Query Cancelled"]
}
```

Peristiwa Query yang Dilindungi Berhasil

Di bawah ini adalah bidang detail untuk Protected Query Succeeded acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Query Succeeded"]
}
```

Peristiwa Gagal Kueri yang Dilindungi

Di bawah ini adalah bidang detail untuk Protected Query Failed acara tersebut.


```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Query Failed"]
}
```

Acara Timed Out Query yang Dilindungi

Di bawah ini adalah bidang detail untuk Protected Query Timed Out acara tersebut.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Query Timed Out"]
}
```

Penandaan Alokasi Biaya

AWS Clean Rooms mendukung penggunaan Tag Alokasi Biaya untuk melacak AWS biaya Anda. Anda mengaktifkan tag ini di AWS Manajemen Penagihan dan Biaya dasbor. AWS menggunakan tag untuk mengkategorikan biaya Anda dan mengirimkan laporan alokasi biaya bulanan kepada Anda. Tag alokasi biaya yang ditentukan pengguna dapat diterapkan ke AWS Clean Rooms sumber daya untuk membantu melacak dan mengalokasikan biaya di seluruh kolaborasi Anda.

Sumber Daya Taggable untuk Alokasi Biaya

Tabel berikut mencantumkan AWS Clean Rooms sumber daya yang dikenakan biaya dan dapat ditandai dengan tag alokasi biaya yang ditentukan pengguna untuk melacak dan mengatur biaya.

Sumber Daya yang Ditagih	Sumber daya yang ditandai
Kueri SQL	Keanggotaan
PySpark Job	Keanggotaan
Job Pelatihan ML	Keanggotaan
Lowongan Inferensi ML	Keanggotaan
Pembuatan Data Sintetis	Keanggotaan

Sumber Daya yang Ditagih	Sumber daya yang ditandai
Pelatihan Model Lookalike	Model Mirip
Ekspor Segmen Lookalike	Segmen Lookalike

Menciptakan AWS Clean Rooms sumber daya dengan AWS CloudFormation

AWS Clean Rooms terintegrasi dengan AWS CloudFormation, layanan yang membantu Anda memodelkan dan mengatur AWS sumber daya Anda. Sebagai hasil dari integrasi ini, Anda dapat menghabiskan lebih sedikit waktu untuk membuat dan mengelola sumber daya dan infrastruktur Anda. Anda membuat template yang menjelaskan semua AWS sumber daya yang Anda inginkan, dan menyediakan serta CloudFormation mengonfigurasi sumber daya tersebut untuk Anda. Contoh sumber daya termasuk kolaborasi, tabel yang dikonfigurasi, asosiasi tabel yang dikonfigurasi, dan keanggotaan.

Ketika Anda menggunakan CloudFormation, Anda dapat menggunakan kembali template Anda untuk mengatur AWS Clean Rooms sumber daya Anda secara konsisten dan berulang kali. Jelaskan sumber daya Anda sekali, lalu sediakan sumber daya yang sama berulang-ulang dalam beberapa Akun AWS dan Wilayah AWS.

AWS Clean Rooms dan CloudFormation template

Untuk menyediakan dan mengonfigurasi sumber daya untuk AWS Clean Rooms dan layanan terkait, Anda harus memahami [CloudFormation templat](#). Templat adalah file teks dengan format JSON atau YAML. Template ini menjelaskan sumber daya yang ingin Anda sediakan di CloudFormation tumpukan Anda. Jika Anda tidak terbiasa dengan JSON atau YAMAL, Anda dapat menggunakan CloudFormation Designer untuk membantu Anda memulai dengan template. CloudFormation Untuk informasi selengkapnya, lihat [Apa itu CloudFormation Designer?](#) di Panduan Pengguna AWS CloudFormation .

AWS Clean Rooms mendukung pembuatan kolaborasi, tabel yang dikonfigurasi, asosiasi tabel yang dikonfigurasi, dan keanggotaan di. CloudFormation Untuk informasi selengkapnya, termasuk contoh templat JSON dan YAMAL untuk kolaborasi, tabel yang dikonfigurasi, asosiasi tabel yang dikonfigurasi, dan keanggotaan, lihat referensi jenis sumber daya [AWS Clean Rooms](#) [AWS Clean Rooms dan AWS](#) di Panduan Pengguna.AWS CloudFormation

Pelajari lebih lanjut tentang CloudFormation

Untuk mempelajari selengkapnya CloudFormation, lihat sumber daya berikut:

- [AWS CloudFormation](#)
- [AWS CloudFormation Panduan Pengguna](#)
- [CloudFormation Referensi API](#)
- [Panduan Pengguna Antarmuka Baris Perintah AWS CloudFormation](#)

Kuota untuk AWS Clean Rooms

Anda Akun AWS memiliki kuota default, sebelumnya disebut sebagai batas, untuk masing-masing. Layanan AWS Kecuali dinyatakan lain, setiap kuota khusus untuk. AWS Region Anda dapat meminta kenaikan untuk beberapa kuota, dan kuota lainnya tidak dapat ditingkatkan.

Untuk melihat kuota AWS Clean Rooms, buka konsol [Service Quotas](#). Di panel navigasi, pilih layanan AWS dan pilih AWS Clean Rooms.

Untuk meminta peningkatan kuota, lihat [Meminta Peningkatan Kuota](#) dalam Panduan Pengguna Service Quotas. Jika kuota belum tersedia di Service Quotas, gunakan formulir peningkatan [batas Layanan](#).

Topik

- [AWS Clean Rooms kuota](#)
- [Kuota AWS Clean Rooms](#)

AWS Clean Rooms kuota

Anda Akun AWS memiliki kuota berikut yang terkait AWS Clean Rooms dengan.

Nama	Default	Dapat disesu an	Deskripsi
Ukuran aturan analisis	Setiap Wilayah yang didukung: 100 Kilobyte	Tidak	Ukuran maksimum JSON untuk aturan analisis
Templat analisis per keanggotaan	Setiap Wilayah yang didukung: 25	Ya	Jumlah maksimum templat analisis per keanggotaan
Kolaborasi dibuat per akun	Setiap Wilayah yang didukung: 10	Ya	Jumlah maksimum kolaborasi yang dibuat per akun

Nama	Default	Dapat disesuaikan	Deskripsi
Kolom per daftar izinkan tabel yang dikonfigurasi	Setiap Wilayah yang didukung: 100	Ya	Jumlah maksimum kolom yang dapat diizinkan terdaftar per tabel yang dikonfigurasi
Penggunaan vCPU PySpark pekerjaan bersamaan per akun	us-east-1:1.024 Masing-masing Wilayah yang didukung lainnya: 512	Ya	Total penggunaan vCPU maksimum dari semua pekerjaan yang berjalan secara bersamaan per akun PySpark
PySpark Lowongan kerja bersamaan per akun	us-east-1:5 Masing-masing Wilayah yang didukung lainnya: 2	Ya	Jumlah maksimum PySpark pekerjaan yang berjalan secara bersamaan per akun
Kueri SQL bersamaan per akun	us-east-1:5 Masing-masing Wilayah yang didukung lainnya: 2	Ya	Jumlah maksimum kueri SQL yang berjalan secara bersamaan per akun
Penggunaan vCPU kueri SQL bersamaan per akun	Setiap Wilayah yang didukung: 512	Ya	Total penggunaan vCPU maksimum dari semua kueri SQL yang berjalan secara bersamaan per akun

Nama	Default	Dapat disesuaikan	Deskripsi
Pekerjaan berkelanjutan bersamaan per keanggotaan	Setiap Wilayah yang didukung: 1	Tidak	Jumlah maksimum pekerjaan yang sedang berlangsung bersamaan per keanggotaan
Kueri yang sedang berlangsung bersamaan per keanggotaan	Setiap Wilayah yang didukung: 5	Tidak	Jumlah maksimum kueri yang sedang berlangsung bersamaan per keanggotaan
Asosiasi model audiens yang dikonfigurasi per keanggotaan	Setiap Wilayah yang didukung: 5	Ya	Jumlah maksimum asosiasi model audiens yang dikonfigurasi per keanggotaan
Tabel yang dikonfigurasi per akun	Setiap Wilayah yang didukung: 250	Ya	Jumlah maksimum tabel yang dikonfigurasi yang dibuat per akun
Tabel yang dikonfigurasi per kueri yang dilindungi	Setiap Wilayah yang didukung: 15	Ya	Jumlah maksimum tabel yang dikonfigurasi dalam kueri yang dilindungi
Tabel pemetaan ID per keanggotaan	Setiap Wilayah yang didukung: 5	Ya	Jumlah maksimum tabel pemetaan ID per keanggotaan
Asosiasi namespace ID per keanggotaan	Setiap Wilayah yang didukung: 10	Ya	Jumlah maksimum asosiasi namespace ID per keanggotaan
Anggota diundang per kolaborasi	Setiap Wilayah yang didukung: 5	Ya	Jumlah maksimum anggota yang diundang per kolaborasi

Nama	Default	Dapat disesuaikan	Deskripsi
Keanggotaan per akun	Setiap Wilayah yang didukung: 100	Ya	Jumlah maksimum keanggotaan per akun
Templat anggaran privasi untuk anggaran akses per kolaborasi	Setiap Wilayah yang didukung: 25	Ya	Jumlah maksimum templat anggaran privasi untuk anggaran akses per kolaborasi
Panjang teks kueri (menggunakan privasi diferensial)	Setiap Wilayah yang didukung: 8 Kilobyte	Tidak	Panjang teks maksimum untuk pernyataan kueri SQL menggunakan privasi diferensial
Kueri panjang teks pada mesin analitik SQL Clean Rooms	Setiap Wilayah yang didukung: 90 Kilobyte	Tidak	Panjang teks maksimum untuk pernyataan kueri SQL pada mesin analitik SQL Clean Rooms
Kueri panjang teks pada mesin analitik Spark	Setiap Wilayah yang didukung: 500 Kilobyte	Tidak	Panjang teks maksimum untuk pernyataan kueri SQL di mesin analitik Spark
Tingkat BatchGetCollaborationAnalysisTemplate permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan BatchGetCollaborationAnalysisTemplate API per detik
Tingkat BatchGetSchema permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan BatchGetSchema API per detik

Nama	Default	Dapat disesuaikan	Deskripsi
Tingkat BatchGetSchemaAnalysisRule permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan BatchGetSchemaAnalysisRule API per detik
Tingkat CreateAnalysisTemplate permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan CreateAnalysisTemplate API per detik
Tingkat CreateCollaboration permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan CreateCollaboration API per detik
Tingkat CreateCollaborationChangeRequest permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan CreateCollaborationChangeRequest API per detik
Tingkat CreateConfiguredAudienceModelAssociation permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan CreateConfiguredAudienceModelAssociation API per detik
Tingkat CreateConfiguredTable permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan CreateConfiguredTable API per detik
Tingkat CreateConfiguredTableAnalysisRule permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan CreateConfiguredTableAnalysisRule API per detik

Nama	Default	Dapat disesuaikan	Deskripsi
Tingkat CreateConfiguredTableAssociation permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan CreateConfiguredTableAssociation API per detik
Tingkat CreateConfiguredTableAssociationAnalysisRule permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan CreateConfiguredTableAssociationAnalysisRule API per detik
Tingkat CreateIdMappingTable permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan CreateIdMappingTable API per detik
Tingkat CreateIdNamespaceAssociation permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan CreateIdNamespaceAssociation API per detik
Tingkat CreateMembership permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan CreateMembership API per detik
Tingkat CreatePrivacyBudgetTemplate permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan CreatePrivacyBudgetTemplate API per detik
Tingkat DeleteAnalysisTemplate permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan DeleteAnalysisTemplate API per detik

Nama	Default	Dapat disesuaikan	Deskripsi
Tingkat DeleteCollaboration permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan DeleteCollaboration API per detik
Tingkat DeleteConfiguredAudienceModelAssociation permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan DeleteConfiguredAudienceModelAssociation API per detik
Tingkat DeleteConfiguredTable permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan DeleteConfiguredTable API per detik
Tingkat DeleteConfiguredTableAnalysisRule permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan DeleteConfiguredTableAnalysisRule API per detik
Tingkat DeleteConfiguredTableAssociation permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan DeleteConfiguredTableAssociation API per detik
Tingkat DeleteConfiguredTableAssociationAnalysisRule permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan DeleteConfiguredTableAssociationAnalysisRule API per detik
Tingkat DeleteIdMappingTable permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan DeleteIdMappingTable API per detik

Nama	Default	Dapat disesuaikan	Deskripsi
Tingkat DeleteIdNamespaceAssociation permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan DeleteIdNamespaceAssociation API per detik
Tingkat DeleteMember permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan DeleteMember API per detik
Tingkat DeleteMembership permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan DeleteMembership API per detik
Tingkat DeletePrivacyBudgetTemplate permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan DeletePrivacyBudgetTemplate API per detik
Tingkat GetAnalysisTemplate permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan GetAnalysisTemplate API per detik
Tingkat GetCollaboration permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan GetCollaboration API per detik
Tingkat GetCollaborationAnalysisTemplate permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan GetCollaborationAnalysisTemplate API per detik

Nama	Default	Dapat disesuaikan	Deskripsi
Tingkat GetCollaborationChangeRequest permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan GetCollaborationChangeRequest API per detik
Tingkat GetCollaborationConfiguredAudienceModelAssociation permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan GetCollaborationConfiguredAudienceModelAssociation API per detik
Tingkat GetCollaborationIdNamespace Association permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan GetCollaborationIdNamespace Association API per detik
Tingkat GetCollaborationPrivacyBudgetTemplate permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan GetCollaborationPrivacyBudgetTemplate API per detik
Tingkat GetConfiguredAudienceModelAssociation permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan GetConfiguredAudienceModelAssociation API per detik
Tingkat GetConfiguredTable permintaan	Setiap Wilayah yang didukung: 20 per detik	Ya	Jumlah maksimum panggilan GetConfiguredTable API per detik
Tingkat GetConfiguredTableAnalysisRule permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan GetConfiguredTableAnalysisRule API per detik

Nama	Default	Dapat disesuaikan	Deskripsi
Tingkat GetConfiguredTableAssociation permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan GetConfiguredTableAssociation API per detik
Tingkat GetConfiguredTableAssociationAnalysisRule permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan GetConfiguredTableAssociationAnalysisRule API per detik
Tingkat GetIdMappingTable permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan GetIdMappingTable API per detik
Tingkat GetIdNamespaceAssociation permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan GetIdNamespaceAssociation API per detik
Tingkat GetMembership permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan GetMembership API per detik
Tingkat GetPrivacyBudgetTemplate permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan GetPrivacyBudgetTemplate API per detik
Tingkat GetProtectedJob permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan GetProtectedJob API per detik

Nama	Default	Dapat disesuaikan	Deskripsi
Tingkat GetProtectedQuery permintaan	Setiap Wilayah yang didukung: 20 per detik	Ya	Jumlah maksimum panggilan GetProtectedQuery API per detik
Tarif GetSchema permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan GetSchema API per detik
Tarif GetSchemaAnalysisRule permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan GetSchemaAnalysisRule API per detik
Tarif ListAnalysisTemplates permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListAnalysisTemplates API per detik
Tarif ListCollaborationAnalysisTemplates permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListCollaborationAnalysisTemplates API per detik
Tarif ListCollaborationChangeRequests permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListCollaborationChangeRequests API per detik
Tarif ListCollaborationConfiguredAudienceModelAssociations permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListCollaborationConfiguredAudienceModelAssociations API per detik

Nama	Default	Dapat disebutkan	Deskripsi
Tarif ListCollaborationIdNamespacesAssociations permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListCollaborationIdNamespacesAssociations API per detik
Tarif ListCollaborationPrivacyBudgetTemplates permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListCollaborationPrivacyBudgetTemplates API per detik
Tarif ListCollaborationPrivacyBudgets permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListCollaborationPrivacyBudgets API per detik
Tarif ListCollaborations permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListCollaborations API per detik
Tarif ListConfiguredAudienceModelAssociations permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListConfiguredAudienceModelAssociations API per detik
Tarif ListConfiguredTableAssociations permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListConfiguredTableAssociations API per detik

Nama	Default	Dapat disesuaikan	Deskripsi
Tarif ListConfiguredTables permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListConfiguredTables API per detik
Tarif ListIdMappingTables permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListIdMappingTables API per detik
Tarif ListIdNamespaceAssociations permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListIdNamespaceAssociations API per detik
Tarif ListMembers permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListMembers API per detik
Tarif ListMemberships permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListMemberships API per detik
Tarif ListPrivacyBudgetTemplates permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListPrivacyBudgetTemplates API per detik
Tarif ListPrivacyBudgets permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListPrivacyBudgets API per detik
Tarif ListProtectedJobs permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListProtectedJobs API per detik

Nama	Default	Dapat disesuaikan	Deskripsi
Tarif ListProtectedQueries permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListProtectedQueries API per detik
Tarif ListSchemas permintaan	Setiap Wilayah yang didukung: 5	Ya	Jumlah maksimum panggilan ListSchemas API per detik
Tarif ListTagsForResource permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan ListTagsForResource API per detik
Tarif PopulateIdMappingTable permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan PopulateIdMappingTable API per detik
Tarif PreviewPrivacyImpact permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan PreviewPrivacyImpact API per detik
Tarif StartProtectedJob permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan StartProtectedJob API per detik
Tarif StartProtectedQuery permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan StartProtectedQuery API per detik
Tarif TagResource permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan TagResource API per detik

Nama	Default	Dapat disesuaikan	Deskripsi
Tarif UntagResource permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan UntagResource API per detik
Tarif UpdateAnalysisTemplate permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan UpdateAnalysisTemplate API per detik
Tarif UpdateCollaboration permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan UpdateCollaboration API per detik
Tarif UpdateCollaborationChangeRequest permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan UpdateCollaborationChangeRequest API per detik
Tarif UpdateConfiguredAudienceModelAssociation permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan UpdateConfiguredAudienceModelAssociation API per detik
Tarif UpdateConfiguredTable permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan UpdateConfiguredTable API per detik
Tarif UpdateConfiguredTableAnalysisRule permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan UpdateConfiguredTableAnalysisRule API per detik

Nama	Default	Dapat disesuaikan	Deskripsi
Tarif UpdateConfiguredTableAssociation permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan UpdateConfiguredTableAssociation API per detik
Tarif UpdateConfiguredTableAssociationAnalysisRule permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan UpdateConfiguredTableAssociationAnalysisRule API per detik
Tarif UpdateIdMappingTable permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan UpdateIdMappingTable API per detik
Tarif UpdateIdNamespaceAssociation permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan UpdateIdNamespaceAssociation API per detik
Tarif UpdatePrivacyBudgetTemplate permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan UpdatePrivacyBudgetTemplate API per detik
Tarif UpdateProtectedJob permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan UpdateProtectedJob API per detik
Tarif UpdateProtectedQuery permintaan	Setiap Wilayah yang didukung: 5 per detik	Ya	Jumlah maksimum panggilan UpdateProtectedQuery API per detik

Nama	Default	Dapat disesuaikan	Deskripsi
Asosiasi tabel per keanggotaan	Setiap Wilayah yang didukung: 25	Ya	Jumlah maksimum asosiasi tabel per keanggotaan

AWS Clean Rooms batas parameter sumber daya

Sumber daya	Default	Deskripsi
Panjang teks kueri	90 KB	Panjang teks maksimum untuk pernyataan kueri SQL
Panjang teks kueri (menggunakan privasi diferensial)	90 KB	Panjang teks maksimum untuk pernyataan kueri SQL menggunakan privasi diferensial
Waktu berjalan kueri	24 jam	Durasi maksimum kueri dijalankan sebelum batas waktu

Kuota AWS Clean Rooms

Anda Akun AWS memiliki kuota default, yang sebelumnya disebut sebagai limit, untuk AWS Clean Rooms ML.

Untuk melihat kuota layanan AWS Clean Rooms, lakukan salah satu hal berikut:

- Ikuti langkah-langkah di [Melihat kuota layanan](#) di Panduan Pengguna Service Quotas dan AWS Clean Rooms pilih ML sebagai layanan.
- Lihat [kuota layanan AWS Clean Rooms AWS](#) di. Referensi Umum Amazon Web Services

Untuk menjaga kinerja layanan dan memastikan penggunaan AWS Clean Rooms yang tepat, kuota default yang ditetapkan ke akun dapat diperbarui tergantung pada faktor Regional, riwayat pembayaran, penggunaan penipuan, and/or persetujuan permintaan peningkatan kuota.

Nama	Default	Dapat disesu an	Deskripsi
Pekerjaan ekspor audiens aktif per pekerjaan pembuatan audiens	Setiap Wilayah yang didukung: 25	Tidak	Jumlah maksimum pekerjaan ekspor audiens aktif untuk pekerjaan generasi audiens
Asosiasi algoritma model yang dikonfigu rasi aktif per keanggotaan	Setiap Wilayah yang didukung: 1.000	Ya	Jumlah maksimum asosiasi algoritma model yang dikonfigurasi aktif per keanggotaan
Algoritma model yang dikonfigurasi aktif per keanggotaan	Setiap Wilayah yang didukung: 1.000	Ya	Jumlah maksimum algoritma model yang dikonfigurasi aktif per keanggotaan
Saluran input model kustom aktif per keanggotaan	Setiap Wilayah yang didukung: 100	Ya	Jumlah maksimum saluran input model kustom aktif per keanggotaan
Jumlah maksimum instans pelatihan aktif per akun	Setiap Wilayah yang didukung: 10	Ya	Jumlah maksimum instans pelatihan aktif yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum versi model terlatih active/pending/in kemajuan per model terlatih	Setiap Wilayah yang didukung: 100	Ya	Jumlah maksimum versi model terlatih active/pending/in kemajuan yang

Nama	Default	Dapat disesuaikan	Deskripsi
			dapat dibuat per model terlatih.
Jumlah maksimum kolom input untuk menghasilkan data sintetis	Setiap Wilayah yang didukung: 1.000	Tidak	Jumlah maksimum kolom input untuk menghasilkan data sintetis
Jumlah maksimum baris input untuk menghasilkan data sintetis	Setiap Wilayah yang didukung: 2.500.000	Tidak	Jumlah maksimum baris input untuk menghasilkan data sintetis
Jumlah maksimum instans pelatihan ml.c4.2xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c4.2xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c4.4xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c4.4xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c4.8xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c4.8xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c4.xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c4.xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.c5.18xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c5.18xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c5.2xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c5.2xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c5.4xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c5.4xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c5.9xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c5.9xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c5.xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c5.xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.c5n.18xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.c5n.18xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c5n.2xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.c5n.2xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c5n.4xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c5n.4xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c5n.9xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.c5n.9xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c5n.xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.c5n.xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.c6i.12xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c6i.12xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c6i.16xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c6i.16xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c6i.24xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c6i.24xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c6i.2xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.c6i.2xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c6i.32xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c6i.32xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.c6i.4xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c6i.4xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c6i.8xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c6i.8xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c6i.xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.c6i.xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c7i.12xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c7i.12xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c7i.16xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c7i.16xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.c7i.24xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c7i.24xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c7i.2xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.c7i.2xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c7i.48xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c7i.48xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c7i.4xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c7i.4xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c7i.8xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.c7i.8xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.c7i.large per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.c7i.large yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.c7i.xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.c7i.xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g4dn.12xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g4dn.12xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g4dn.16xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.g4dn.16xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g4dn.2xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.g4dn.2xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.g4dn.4xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g4dn.4xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g4dn.8xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g4dn.8xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g4dn.xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.g4dn.xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g5.12xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g5.12xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.g5.16xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g5.16xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g5.24xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g5.24xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g5.2xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g5.2xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g5.48xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g5.48xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g5.4xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g5.4xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.g5.8xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g5.8xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g5.xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g5.xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g6.12xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g6.12xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g6.16xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g6.16xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g6.24xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g6.24xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.g6.2xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g6.2xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g6.48xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g6.48xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g6.4xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g6.4xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g6.8xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g6.8xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g6.xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g6.xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.g6e.12xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g6e.12xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g6e.16xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g6e.16xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g6e.24xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.g6e.24xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g6e.2xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.g6e.2xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.g6e.48xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g6e.48xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g6e.4xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g6e.4xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g6e.8xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g6e.8xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.g6e.xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.g6e.xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m4.10xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m4.10xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.m4.16xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m4.16xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m4.2xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m4.2xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m4.4xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m4.4xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m4.xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m4.xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m5.12xlarge per akun	Setiap Wilayah yang didukung: 3	Ya	Jumlah maksimum instans pelatihan ml.m5.12xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.m5.24xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m5.24xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m5.2xlarge per akun	Setiap Wilayah yang didukung: 3	Ya	Jumlah maksimum instans pelatihan ml.m5.2xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m5.4xlarge per akun	Setiap Wilayah yang didukung: 3	Ya	Jumlah maksimum instans pelatihan ml.m5.4xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m5.large per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m5.large yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m5.xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m5.xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.m6i.12xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m6i.12xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m6i.16xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m6i.16xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m6i.24xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m6i.24xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m6i.2xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.m6i.2xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.m6i.32xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m6i.32xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m6i.4xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m6i.4xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m6i.8xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m6i.8xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m6i.large per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.m6i.large yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m6i.xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.m6i.xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.m7i.12xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m7i.12xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m7i.16xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m7i.16xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m7i.24xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m7i.24xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m7i.2xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m7i.2xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.m7i.48xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m7i.48xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m7i.4xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m7i.4xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m7i.8xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.m7i.8xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m7i.large per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.m7i.large yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.m7i.xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.m7i.xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.p2.16xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.p2.16xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.p2.8xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.p2.8xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.p2.xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.p2.xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.p4d.24xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.p4d.24xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.p4de.24xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.p4de.24xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.p5.48xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.p5.48xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.p5en.48xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.p5en.48xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.r5.12xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.r5.12xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.r5.16xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.r5.16xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.r5.24xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.r5.24xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.r5.2xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.r5.2xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.r5.4xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.r5.4xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.r5.8xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.r5.8xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum ml.r5.large instance pelatihan per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.r5.large yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.r5.xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.r5.xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.r5d.12xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.r5d.12xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.r5d.16xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.r5d.16xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.r5d.24xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.r5d.24xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.r5d.2xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.r5d.2xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.r5d.4xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.r5d.4xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.r5d.8xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.r5d.8xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.r5d.large per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.r5d.large yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.r5d.xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.r5d.xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.r7i.12xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.r7i.12xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.r7i.16xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.r7i.16xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.r7i.24xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.r7i.24xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.r7i.2xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.r7i.2xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.r7i.48xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.r7i.48xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.r7i.4xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.r7i.4xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.r7i.8xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.r7i.8xlarge yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah maksimum instans pelatihan ml.r7i.large per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.r7i.large yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.r7i.xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.r7i.xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.t3.2xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.t3.2xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.t3.large per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.t3.large yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.t3.medium per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.t3.medium yang dapat digunakan per akun untuk membuat model terlatih.

Nama	Default	Dapat disesu an	Deskripsi
Jumlah maksimum instans pelatihan ml.t3.xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.t3.xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.trn1.2xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.trn1.2xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.trn1.32xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instans pelatihan ml.trn1.32xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Jumlah maksimum instans pelatihan ml.trn1n.32xlarge per akun	Setiap Wilayah yang didukung: 0	Ya	Jumlah maksimum instance pelatihan ml.trn1n.32xlarge yang dapat digunakan per akun untuk membuat model terlatih.
Nilai unik maksimum di kolom prediksi kategoris	Setiap Wilayah yang didukung: 100	Tidak	Nilai unik maksimum dalam kolom prediksi kategoris

Nama	Default	Dapat disesuaikan	Deskripsi
Jumlah minimum kolom input yang diperlukan untuk menghasilkan data sintetis	Setiap Wilayah yang didukung: 5	Tidak	Jumlah minimum kolom input yang diperlukan untuk menghasilkan data sintetis
Jumlah minimum baris input yang diperlukan untuk menghasilkan data sintetis	Setiap Wilayah yang didukung: 1.500	Tidak	Jumlah minimum baris input yang diperlukan untuk menghasilkan data sintetis
Pekerjaan ekspor audiens yang tertunda/sedang berlangsung per pelanggan	Setiap Wilayah yang didukung: 20	Tidak	Jumlah maksimum pekerjaan ekspor audiens yang sedang berlangsung per pelanggan
Pekerjaan pembuatan audiens yang tertunda/dalam proses per pelanggan	Setiap Wilayah yang didukung: 10	Ya	Jumlah maksimum pekerjaan pembuatan audiens yang tertunda/dalam proses per pelanggan
Model audiens yang tertunda/dalam proses per pelanggan	Setiap Wilayah yang didukung: 2	Ya	Jumlah maksimum pekerjaan pelatihan model audiens yang sedang berlangsung per pelanggan
Pekerjaan inferensi model kustom yang tertunda/sedang berlangsung per akun	Setiap Wilayah yang didukung: 10	Ya	Jumlah maksimum pekerjaan inferensi model kustom yang tertunda dalam proses per akun

Nama	Default	Dapat disesuaikan	Deskripsi
Pekerjaan inferensi model kustom yang tertunda/sedang berlangsung per keanggotaan	Setiap Wilayah yang didukung: 5	Ya	Jumlah maksimum pekerjaan inferensi model kustom yang tertunda dalam proses per keanggotaan
Pekerjaan pelatihan model kustom yang tertunda/sedang berlangsung per akun	Setiap Wilayah yang didukung: 10	Ya	Jumlah maksimum pekerjaan pelatihan model kustom yang sedang berlangsung per akun
Pekerjaan pelatihan model kustom yang tertunda/sedang berlangsung per keanggotaan	Setiap Wilayah yang didukung: 5	Ya	Jumlah maksimum pekerjaan pelatihan model kustom yang sedang berlangsung per keanggotaan
Pekerjaan pembuatan MLIC sintetis yang sedang berlangsung per akun	Setiap Wilayah yang didukung: 2	Ya	Jumlah maksimum pekerjaan pembuatan MLIC sintetis yang sedang berlangsung per akun

AWS Clean Rooms MS memiliki kuota tambahan yang tercantum di bawah ini

Kuota Kamar Bersih

Sumber daya	Default	Deskripsi
Durasi maksimum pekerjaan inferensi pemodelan khusus	25 jam	

Sumber daya	Default	Deskripsi
Jumlah maksimum interaksi	20 miliar	Jumlah maksimum interaksi yang diizinkan dalam data pelatihan. Input yang lebih besar diambil sampelnya.
Minimal jumlah interaksi	1 juta.	
Jumlah maksimum pengguna berbeda untuk pelatihan model mirip	100 juta	Jika lebih banyak dimasukkan, hanya 100 juta teratas yang digunakan, diberi peringkat berdasarkan jumlah interaksi.
Jumlah minimum pengguna yang berbeda untuk pelatihan model mirip	100.000	
Jumlah minimum pengguna untuk pekerjaan segmen mirip ekspor (audiens)	10.000	
Jumlah maksimum item berbeda yang digunakan untuk pelatihan model.	1 juta.	Anda dapat memasukkan hingga 50 juta item, tetapi hanya 1 juta yang paling populer yang digunakan.
Jumlah maksimum kolom fitur dalam kumpulan data pelatihan.	10	
Jumlah minimum item berbeda per pengguna	2	AWS Clean Rooms ML mengharuskan setiap baris atau pengguna memiliki dua item atau lebih, termasuk item berulang.
Ukuran maksimum audiens benih	500.000	

Sumber daya	Default	Deskripsi
Ukuran minimum audiens benih	500	Penyedia data pelatihan dapat menetapkan nilai ini serendah 25.
APIs	per pelanggan	
Jumlah total kumpulan data pelatihan aktif	500	
Jumlah total model mirip aktif (model audiens)	500	
Jumlah total model mirip yang dikonfigurasi aktif (model audiens)	10.000	
Jumlah total pekerjaan pembuatan segmen mirip (audiens) yang diselesaikan	Tidak ada batas	
Jumlah total pekerjaan segmen mirip ekspor (audiens) yang diselesaikan	Tidak ada batas	
Durasi maksimum pekerjaan pembuatan model mirip (model audiens)	1 hari (24 jam)	

Sumber daya	Default	Deskripsi
Durasi maksimum pekerjaan pembuatan segmen mirip (audiens)	10 jam	Setelah Anda memberikan benih, Clean Rooms ML membutuhkan waktu maksimal 10 jam untuk menghasilkan segmen yang mirip. Jika Anda menggunakan kueri SQL sebagai data benih, dibutuhkan waktu hingga 12 jam untuk menjalankan kueri selain 10 jam untuk menghasilkan segmen mirip.
Persentase minimum untuk bin ukuran segmen (audiens)	1%	
Persentase maksimum untuk tempat sampah ukuran segmen (audiens)	20%	
Ukuran absolut minimum untuk bin ukuran segmen (audiens)	1% dari jumlah pengguna yang berbeda	
Ukuran absolut maksimum untuk tempat sampah ukuran segmen (audiens)	20% dari jumlah pengguna yang berbeda	

Kuota pembatasan API Clean Rooms

Anda Akun AWS memiliki transaksi API per detik (TPS) per akun per kuota titik akhir berikut.

Sumber daya	Batas tarif	Deskripsi
Permintaan tarif CreateAudienceModel	1 tingkat TPS, 3 TPS meledak	Jumlah maksimum panggilan CreateAudienceModel API per detik
Permintaan tarif CreateConfiguredAudienceModel	10 TPS	Jumlah maksimum panggilan CreateConfiguredAudienceModel API per detik
Permintaan tarif CreateTrainingDataset	10 TPS	Jumlah maksimum panggilan CreateTrainingDataset API per detik
Permintaan tarif DeleteAudienceGenerationJob	2 tingkat TPS, 10 TPS meledak	Jumlah maksimum panggilan DeleteAudienceGenerationJob API per detik
Permintaan tarif DeleteAudienceModel	2 tingkat TPS, 10 TPS meledak	Jumlah maksimum panggilan DeleteAudienceModel API per detik
Permintaan tarif DeleteConfiguredAudienceModel	10 TPS	Jumlah maksimum panggilan DeleteConfiguredAudienceModel API per detik
Permintaan tarif DeleteConfiguredAudienceModelPolicy	25 TPS	Jumlah maksimum panggilan DeleteConfiguredAudienceModelPolicy API per detik
Permintaan tarif DeleteTrainingDataset	10 TPS	Jumlah maksimum panggilan DeleteTrainingDataset API per detik

Sumber daya	Batas tarif	Deskripsi
Permintaan tarif <code>GetAudienceGenerationJob</code>	50 TPS	Jumlah maksimum panggilan <code>GetAudienceGenerationJob</code> API per detik
Permintaan tarif <code>GetAudienceModel</code>	50 TPS	Jumlah maksimum panggilan <code>GetAudienceModel</code> API per detik
Permintaan tarif <code>GetConfiguredAudienceModel</code>	50 TPS	Jumlah maksimum panggilan <code>GetConfiguredAudienceModel</code> API per detik
Permintaan tarif <code>GetConfiguredAudienceModelPolicy</code>	50 TPS	Jumlah maksimum panggilan <code>GetConfiguredAudienceModelPolicy</code> API per detik
Permintaan tarif <code>GetTrainingDataset</code>	50 TPS	Jumlah maksimum panggilan <code>GetTrainingDataset</code> API per detik
Permintaan tarif <code>ListAudienceExportJobs</code>	50 TPS	Jumlah maksimum panggilan <code>ListAudienceExportJobs</code> API per detik
Permintaan tarif <code>ListAudienceGenerationJobs</code>	50 TPS	Jumlah maksimum panggilan <code>ListAudienceGenerationJobs</code> API per detik
Permintaan tarif <code>ListAudienceModels</code>	50 TPS	Jumlah maksimum panggilan <code>ListAudienceModels</code> API per detik
Permintaan tarif <code>ListConfiguredAudienceModels</code>	50 TPS	Jumlah maksimum panggilan <code>ListConfiguredAudienceModels</code> API per detik

Sumber daya	Batas tarif	Deskripsi
Permintaan tarif ListTagsForResource	50 TPS	Jumlah maksimum panggilan ListTagsForResource API per detik
Permintaan tarif ListTrainingDatasets	50 TPS	Jumlah maksimum panggilan ListTrainingDatasets API per detik
Permintaan tarif PutConfiguredAudienceModelPolicy	25 TPS	Jumlah maksimum panggilan PutConfiguredAudienceModelPolicy API per detik
Permintaan tarif StartAudienceExportJob	1 tingkat TPS, 3 TPS meledak	Jumlah maksimum panggilan StartAudienceExportJob API per detik
Permintaan tarif StartAudienceGenerationJob	1 tingkat TPS, 5 TPS meledak	Jumlah maksimum panggilan StartAudienceGenerationJob API per detik
Permintaan tarif TagResource	10 TPS	Jumlah maksimum panggilan TagResource API per detik
Permintaan tarif UntagResource	50 TPS	Jumlah maksimum panggilan UntagResource API per detik
Permintaan tarif UpdateConfiguredAudienceModel	10 TPS	Jumlah maksimum panggilan UpdateConfiguredAudienceModel API per detik

Sumber daya	Batas tarif	Deskripsi
Permintaan tarif CreateConfiguredModelAlgorithm	10 TPS	Jumlah maksimum panggilan CreateConfiguredModelAlgorithm API per detik.
Permintaan tarif CreateConfiguredModelAlgorithmAssociation	10 TPS	Jumlah maksimum panggilan CreateConfiguredModelAlgorithmAssociation API per detik.
Permintaan tarif PutMLConfiguration	10 TPS	Jumlah maksimum panggilan PutMLConfiguration API per detik.
Permintaan tarif CreateTrainedModel	1 tingkat TPS, 3 TPS meledak	Jumlah maksimum panggilan CreateTrainedModel API per detik.
Permintaan tarif StartTrainedModelExportJob	10 TPS	Jumlah maksimum panggilan StartTrainedModelExportJob API per detik.
Permintaan tarif StartTrainedModelInferenceJob	1 tingkat TPS, tingkat 3 TPS	Jumlah maksimum panggilan StartTrainedModelInferenceJob API per detik.
Tingkat GetConfiguredModelAlgorithm permintaan	50 TPS	Jumlah maksimum panggilan GetConfiguredModelAlgorithm API per detik.
Tingkat GetConfiguredModelAlgorithmAssociation permintaan	50 TPS	Jumlah maksimum panggilan GetConfiguredModelAlgorithmAssociation API per detik.

Sumber daya	Batas tarif	Deskripsi
Permintaan tarif <code>GetTrainedModel</code>	50 TPS	Jumlah maksimum panggilan <code>GetTrainedModel</code> API per detik.
Permintaan tarif <code>GetMLConfiguration</code>	50 TPS	Jumlah maksimum panggilan <code>GetMLConfiguration</code> API per detik.
Permintaan tarif <code>GetTrainedModelInferenceJob</code>	50 TPS	Jumlah maksimum panggilan <code>GetTrainedModelInferenceJob</code> API per detik.
Permintaan tarif <code>ListConfiguredModelAlgorithm</code>	50 TPS	Jumlah maksimum panggilan <code>ListConfiguredModelAlgorithm</code> API per detik.
Permintaan tarif <code>ListConfiguredModelAlgorithmAssociations</code>	50 TPS	Jumlah maksimum panggilan <code>ListConfiguredModelAlgorithmAssociations</code> API per detik.
Permintaan tarif <code>ListTrainedModels</code>	50 TPS	Jumlah maksimum panggilan <code>ListTrainedModels</code> API per detik.
Permintaan tarif <code>ListCollaborationTrainedModelExportJobs</code>	50 TPS	Jumlah maksimum panggilan <code>ListCollaborationTrainedModelExportJobs</code> API per detik.
Permintaan tarif <code>ListCollaborationTrainedModelInferenceJobs</code>	50 TPS	Jumlah maksimum panggilan <code>ListCollaborationTrainedModelInferenceJobs</code> API per detik.

Sumber daya	Batas tarif	Deskripsi
Permintaan tarif DeleteConfiguredModelAlgorithm	2 tingkat TPS, 10 TPS meledak	Jumlah maksimum panggilan DeleteConfiguredModelAlgorithm API per detik.
Permintaan tarif DeleteConfiguredModelAlgorithmAssociation	2 tingkat TPS, 10 TPS meledak	Jumlah maksimum permintaan DeleteConfiguredModelAlgorithmAssociation API per detik.
Permintaan tarif DeleteMLConfiguration	2 tingkat TPS, 10 TPS meledak	Jumlah maksimum permintaan DeleteMLConfiguration API per detik.
Permintaan tarif DeleteTrainedModelOutput	2 tingkat TPS, 10 TPS meledak	Jumlah maksimum permintaan DeleteTrainedModelOutput API per detik.

Riwayat dokumen untuk Panduan AWS Clean Rooms Pengguna

Tabel berikut menjelaskan rilis dokumentasi untuk AWS Clean Rooms.

Untuk notifikasi tentang pembaruan-pembaruan dokumentasi ini, Anda dapat berlangganan ke sebuah umpan RSS. Untuk berlangganan pembaruan RSS, Anda harus mengaktifkan plug-in RSS untuk browser yang Anda gunakan.

Perubahan	Deskripsi	Tanggal
Dokumentasi yang Diperbarui untuk Bagian Penandaan Alokasi Biaya	Menambahkan dokumentasi untuk AWS Manajemen Penagihan dan Biaya dan membuat bagian untuk alokasi biaya terkait sumber daya taggable.	Februari 20, 2026
Support untuk konfigurasi komputasi spark cleanrooms-mL	Pelanggan sekarang dapat menyesuaikan perilaku runtime Spark dan mengoptimalkan kinerja SQL dengan mengonfigurasi properti Spark yang didukung untuk kueri SQL. Kemampuan ini sekarang tersedia untuk pembuatan lapangan kerja generasi pemirsa dan pembuatan MLInput Saluran .	Januari 5, 2026
Support untuk pemantauan terperinci dengan CloudWatch	Pelanggan sekarang dapat ikut serta dalam pemantauan terperinci dengan CloudWatch metrik untuk pemantauan operasional kolaborasi,	Januari 2, 2026

termasuk kinerja kueri dan pemanfaatan sumber daya.

[Support untuk acara EventBridge notifikasi Permintaan Perubahan Kolaborasi](#)

Menambahkan EventBridge acara baru untuk acara siklus hidup Permintaan Perubahan Kolaborasi.

Desember 19, 2025

[Dukungan untuk permintaan perubahan untuk mengelola keanggotaan dan pengaturan kolaborasi](#)

Pembuat kolaborasi sekarang dapat mengirimkan permintaan perubahan untuk menambahkan anggota baru, memperbaiki kemampuan anggota yang ada, dan mengubah pengaturan persetujuan otomatis. Semua anggota kolaborasi harus menyetujui permintaan perubahan sebelum perubahan diterapkan.

Desember 18, 2025

[Akhir garis waktu dukungan untuk mesin analitik AWS Clean Rooms SQL](#)

Support untuk mesin analitik AWS Clean Rooms SQL berakhir pada 17 Desember 2025. Pelanggan sekarang dapat menggunakan mesin analitik Spark SQL saja.

Desember 17, 2025

Support untuk acara EventBridge notifikasi tambahan	Menambahkan EventBridge peristiwa baru untuk Template Analisis, Asosiasi Tabel Terkonfigurasi, Aturan Analisis Asosiasi Tabel yang Dikonfigurasi, Tabel Pemetaan Id, Asosiasi Namespace Id, Pekerjaan yang Dilindungi, dan Kueri Terlindungi.	Desember 15, 2025
Support untuk pembuatan dataset sintetis yang disempurnakan privasi untuk melatih model pembelajaran mesin kustom	Pelanggan sekarang dapat menghasilkan kumpulan data sintetis yang ditingkatkan privasi untuk melatih model pembelajaran mesin khusus sambil melindungi data sensitif.	November 30, 2025
Support untuk properti Spark yang dapat dikonfigurasi dalam kueri SQL	Pelanggan sekarang dapat menyesuaikan perilaku runtime Spark dengan mengonfigurasi properti Spark yang didukung untuk kueri SQL.	Oktober 30, 2025
Support untuk kolaborasi lintas wilayah	AWS Clean Rooms sekarang mendukung kolaborasi lintas wilayah untuk AWS Glue dan tabel data Athena, memungkinkan pelanggan untuk memproses data yang disimpan dalam AWS Region satu dan memberikan hasil ke Wilayah yang disetujui oleh semua anggota kolaborasi.	Oktober 3, 2025

[Support untuk anggaran akses data](#)

Pelanggan sekarang dapat mengatur dan mengelola anggaran akses data untuk mengontrol penggunaan tabel.

Oktober 1, 2025

[Isi tabel pemetaan ID - perbarui](#)

Pelanggan sekarang dapat menggunakan pemrosesan inkremental dalam alur kerja pemetaan ID berbasis aturan untuk memproses kumpulan data besar secara lebih efisien. Pelanggan juga dapat menghapus catatan dari alur kerja pemetaan ID untuk membantu mematuhi peraturan manajemen data.

September 23, 2025

[Support untuk ukuran beban kerja variabel untuk PySpark pekerjaan](#)

Pelanggan sekarang dapat memilih dan menyesuaikan ukuran komputasi mereka untuk menjalankan PySpark pekerjaan dengan lebih efektif.

September 4, 2025

[Support untuk menambahkan anggota baru ke kolaborasi yang sudah ada](#)

Pelanggan sekarang dapat menambahkan anggota baru ke kolaborasi yang ada. Fitur ini mencakup opsi untuk mengonfigurasi pengaturan persetujuan otomatis dan memblokir permintaan keanggotaan.

September 3, 2025

Menambahkan dukungan untuk mengonfigurasi ringkasan kesalahan dan redaksi untuk log model khusus	Pelanggan sekarang dapat mengonfigurasi ringkasan kesalahan dan redaksi untuk log model khusus dalam asosiasi algoritma model yang dikonfigurasi.	September 3, 2025
Dukungan untuk opsi otentikasi Snowflake	Pelanggan sekarang dapat menggunakan otentikasi i pasangan kunci dengan kunci pribadi Snowflake PEM saat menyimpan kredensial Snowflake mereka untuk koneksi. AWS Secrets Manager AWS Clean Rooms	Agustus 29, 2025
Menambahkan konfigurasi kesalahan untuk template PySpark analisis	Pelanggan sekarang dapat mengonfigurasi tingkat pesan kesalahan dasar atau terperinci dalam templat PySpark analisis untuk mengelola visibilitas kesalahan dan kemampuan pemecahan masalah.	Agustus 19, 2025
Support untuk EventBridge	Pelanggan sekarang dapat menggunakan Amazon EventBridge untuk merutekan acara dari AWS Clean Rooms ke komponen perangkat lunak lainnya.	Juli 31, 2025

[Perbarui ke kebijakan yang ada](#)

Izin baru berikut telah ditambahkan ke kebijakan `AWSCleanRoomsFullAccessNoQuerying` terkelola: `cleanrooms:UpdateConfiguredTableAllowedColumns` dan `cleanrooms:UpdateConfiguredTableReference` .

Juli 29, 2025

[Peningkatan kemampuan pengeditan untuk tabel yang dikonfigurasi](#)

Pelanggan sekarang dapat memodifikasi kolom yang diizinkan dan bidang referensi tabel untuk tabel yang dikonfigurasi.

Juli 29, 2025

[Saluran masukan ML mendukung jenis file Parquet](#)

Pelanggan sekarang dapat memilih jenis file Parquet sebagai format data yang dapat digunakan saluran input ML.

Juli 17, 2025

[Support untuk privasi diferensial pada mesin analisis Spark SQL](#)

Pelanggan sekarang dapat menggunakan privasi diferensial dengan mesin analitik Spark SQL. Selain itu, panjang teks maksimum untuk pernyataan kueri SQL menggunakan privasi diferensial sekarang 90 KB.

Juli 16, 2025

[Akhir garis waktu dukungan untuk mesin analitik AWS Clean Rooms SQL](#)

Support untuk kolaborasi baru menggunakan mesin analitik SQL Clean Rooms berakhir pada 16 Juli 2025. Mulai 17 Juli 2025, Anda tidak dapat membuat kolaborasi baru dengan mesin analitik SQL Clean Rooms lama. AWS Clean Rooms mengakhiri dukungan untuk mesin analitik SQL Clean Rooms pada 17 Desember 2025.

Juli 16, 2025

[Support untuk pelatihan inkremental dan terdistribusi](#)

Pelanggan sekarang dapat menggunakan pelatihan tambahan dan pelatihan terdistribusi untuk model khusus mereka.

Juli 1, 2025

[Akhir dari garis waktu dukungan untuk mesin analitik SQL Clean Rooms](#)

Mulai sekarang hingga 16 Juli 2025, kolaborasi baru menggunakan mesin analitik SQL Clean Rooms lama memerlukan peningkatan batas. Mulai 17 Juli 2025, Anda tidak dapat membuat kolaborasi baru dengan mesin ini. AWS Clean Rooms mengakhiri dukungan untuk mesin analitik SQL Clean Rooms pada 17 Desember 2025.

22 Mei 2025

[Perbarui ke templat PySpark analisis](#)

Pelanggan sekarang dapat memberi nama file skrip pengguna dengan nama file Python yang valid.

15 Mei 2025

[Support untuk beberapa penerima hasil](#)

Pelanggan sekarang dapat memilih beberapa anggota untuk menerima hasil kueri dari satu kueri.

April 30, 2025

[Support untuk migrasi kolaborasi ke Spark SQL](#)

AWS Clean Rooms SQL sekarang mendukung agregasi dan aturan analisis daftar, selain aturan analisis kustom. Selain itu, pelanggan dapat memperbarui kolaborasi yang ada untuk menggunakan mesin analitik Spark yang mendukung Spark SQL.

April 2, 2025

[Support untuk PySpark pekerjaan](#)

Pelanggan sekarang dapat menganalisis data dengan menjalankan pekerjaan menggunakan templat PySpark analisis yang disetujui.

Maret 18, 2025

[Perbarui ke kebijakan yang ada](#)

Izin baru berikut telah ditambahkan ke kebijakan `AWSCleanRoomsMLReadOnlyAccess` terkelola: `PassCleanRoomsResources`. Izin baru berikut telah ditambahkan ke kebijakan `AWSCleanRoomsMLFullAccess` terkelola: `PassCleanRoomsResources` dan `ConsoleDescribeECRRepositories`.

Januari 10, 2025

<u>Support untuk beberapa pekerja komputasi</u>	Pelanggan sekarang dapat menentukan jenis pekerja komputasi dan berapa banyak yang harus disediakan saat membuat segmen yang mirip.	Desember 17, 2024
<u>Support untuk berbagai sumber data dan cloud</u>	Pelanggan sekarang dapat menggunakan beberapa sumber data dan cloud, seperti Amazon Athena dan Snowflake, untuk berkolaborasi dengan kumpulan data mitra mereka.	Desember 1, 2024
<u>Clean Rooms MS Custom Modeling sekarang tersedia</u>	Pelanggan sekarang dapat menggunakan model ML kustom mereka sendiri dalam sebuah kolaborasi.	November 7, 2024
<u>Mesin analitik baru</u>	Pelanggan yang memiliki kumpulan data besar sekarang dapat menjalankan kueri kompleks menggunakan fungsi SQL yang didukung oleh mesin analitik Spark SQL.	Oktober 29, 2024
<u>Perlindungan privasi yang ditingkatkan, bangun pemirsa yang mirip, pilih beberapa penerima hasil</u>	Anda dapat melindungi data Anda sambil juga mengizinkan kueri aktivasi kompleks menggunakan Analisis tambahan dan aturan analisis kolaborasi. Anda dapat membuat model audiens yang mirip dari kueri SQL atau templat analisis. Anda dapat memilih beberapa anggota untuk menerima hasil.	Juli 24, 2024

[Resolusi entitas di AWS Clean Rooms](#)

Dengan Resolusi Entitas AWS in AWS Clean Rooms, Anda dapat membuat tabel pemetaan ID antara dua ruang nama ID untuk menanyakan data peristiwa di seluruh ruang identitas yang berbeda.

Juli 23, 2024

[Perbarui ke kebijakan yang ada](#)

Izin baru berikut telah ditambahkan ke kebijakan `AWSCleanRoomsFullAccessNoQuerying` terkelola: `cleanrooms:BatchGetSchemaAnalysisRule`.

13 Mei 2024

[AWS Clean Rooms ML sekarang tersedia sepenuhnya](#)

AWS Clean Rooms ML menyediakan metode peningkatan privasi bagi dua pihak untuk mengidentifikasi pengguna serupa dalam data mereka tanpa perlu berbagi data satu sama lain.

April 3, 2024

[Perbarui ke kebijakan yang ada](#)

ID Pernyataan dalam kebijakan `AWSCleanRoomsFullAccess` terkelola telah diperbarui dari `ConsolePickQueryResultsBucket` ke `SetQueryResultsBucket` untuk mewakili izin dengan lebih baik sejak izin.

Maret 21, 2024

Kebijakan terkelola baru untuk AWS Clean Rooms	Dua kebijakan terkelola baru telah ditambahkan: <code>AWSCleanRoomsMLReadOnlyAccess</code> dan <code>AWSCleanRoomsMLFullAccess</code> .	November 29, 2023
AWS Clean Rooms ML (pratinjau)	AWS Clean Rooms ML menyediakan metode peningkatan privasi bagi dua pihak untuk mengidentifikasi pengguna serupa dalam data mereka tanpa perlu berbagi data satu sama lain.	November 29, 2023
AWS Clean Rooms Privasi Diferensial (pratinjau)	Pelanggan sekarang dapat menggunakan Privasi AWS Clean Rooms Diferensial untuk membantu melindungi privasi pengguna mereka.	November 29, 2023
Konfigurasi pembayaran	Pembuat kolaborasi sekarang dapat mengonfigurasi anggota yang dapat menjalankan kueri atau anggota lain dalam kolaborasi yang akan ditagih untuk biaya komputasi kueri.	14 November 2023
Waktu berjalan kueri - diperbarui	Durasi maksimum kueri dijalankan sebelum batas waktu diperbarui dari 4 jam menjadi 12 jam.	Oktober 6, 2023

[CloudFormation sumber daya
- perbarui](#)

AWS Clean Rooms telah menambahkan sumber daya baru berikut: `AWS::CleanRooms::MembershipProtectedQueryOutputConfiguration`, `AWS::CleanRooms::MembershipProtectedQueryResultConfiguration`, dan `AWS::CleanRooms::MembershipProtectedQueryS3OutputConfiguration`.

7 September 2023

[CloudFormation sumber daya
- perbarui](#)

AWS Clean Rooms telah menambahkan sumber daya baru berikut: `AWS::CleanRooms::AnalysisTemplate` dan `AWS::CleanRooms::ConfiguredTableAnalysisRuleCustom`.

31 Agustus 2023

[Kemampuan anggota terpisah](#)

Pembuat kolaborasi sekarang dapat menunjuk satu anggota sebagai anggota yang dapat meminta dan anggota lain sebagai anggota yang dapat menerima hasil. Ini memberi pembuat kolaborasi kemampuan untuk memastikan bahwa anggota yang dapat melakukan kueri tidak memiliki akses ke hasil kueri.

Agustus 30, 2023

AWS Clean Rooms Glosarium	Pembaruan khusus dokumentasi untuk menambahkan glosarium istilah. AWS Clean Rooms	Agustus 30, 2023
Support untuk Apache Iceberg tabel (pratinjau)	AWS Clean Rooms sekarang mendukung Apache Iceberg tabel (pratinjau).	Agustus 25, 2023
Pembaruan kuota	Bagian Kuota telah diperbarui untuk mencerminkan kuota default baru untuk keanggotaan per akun.	9 Agustus 2023
Perbarui ke kebijakan yang ada	Izin baru berikut telah ditambahkan ke kebijakan <code>AWSCleanRoomsFullAccessNoQuerying</code> terkelola: <code>cleanrooms:CreateAnalysisTemplate</code> , <code>cleanrooms:GetAnalysisTemplate</code> , <code>cleanrooms:UpdateAnalysisTemplate</code> , <code>cleanrooms>DeleteAnalysisTemplate</code> , <code>cleanrooms>ListAnalysisTemplates</code> , <code>cleanrooms:GetCollaborationAnalysisTemplate</code> , <code>cleanrooms:BatchGetCollaborationAnalysisTemplate</code> , dan <code>cleanrooms>ListCollaborationAnalysisTemplates</code> .	31 Juli 2023

<u>Template analisis dan aturan analisis kustom</u>	AWS Clean Rooms sekarang mendukung template analisis dan aturan analisis Kustom. Template analisis memungkinkan kolaborator untuk membangun atau mengimpor kueri SQL kustom mereka sendiri untuk digunakan dalam kolaborasi. Dengan aturan analisis Kustom, pemilik tabel dapat menyetujui kueri SQL kustom pada tabel yang dikonfigurasi.	31 Juli 2023
<u>Aturan analisis mendukung kondisi OR logis</u>	AWS Clean Rooms aturan analisis sekarang mendukung kondisi OR logis dalam JOIN klausa.	29 Juni 2023
<u>CloudFormation integrasi</u>	AWS Clean Rooms sekarang terintegrasi dengan CloudFormation.	15 Juni 2023
<u>Pembuat analisis</u>	Anggota yang dapat menanyakan dan menerima hasil sekarang memiliki kemampuan untuk menjalankan kueri pada beberapa tabel tanpa menulis kode SQL dengan menggunakan UI pembuat Analisis.	15 Juni 2023
<u>Fungsi SQL</u>	Pembaruan khusus dokumentasi untuk memperjelas fungsi SQL yang didukung.	5 Mei 2023

Pemecahan Masalah	Pembaruan khusus dokumentasi untuk menambahkan bagian Pemecahan Masalah untuk masalah umum.	27 April 2023
Tipe data yang didukung untuk AWS Clean Rooms	Pembaruan khusus dokumentasi untuk menambahkan bagian baru yang mencantumkan tipe data yang didukung AWS Glue Data Catalog .	April 26, 2023
Contoh AWS CloudTrail acara	Pembaruan khusus dokumentasi untuk menambahkan contoh CloudTrail peristiwa untuk StartProtectedQuery (berhasil) dan StartProtectedQuery (gagal).	20 April 2023
Perbarui ke kebijakan yang ada	Izin baru berikut telah ditambahkan ke kebijakan <code>AWSCleanRoomsFullAccessNoQuerying</code> <code>terkelola:cleanrooms:ListTagsForResource, cleanrooms:UntagResource , dan cleanrooms:TagResource</code> . Untuk informasi selengkapnya, lihat Kebijakan terkelola AWS .	21 Maret 2023
Ketersediaan umum	AWS Clean Rooms sekarang tersedia secara umum.	21 Maret 2023

[Rilis pratinjau](#)

Pratinjau rilis Panduan AWS
Clean Rooms Pengguna

Januari 12, 2023

AWS Clean Rooms Glosarium

Konsultasikan glosarium ini untuk menjadi akrab dengan terminologi yang digunakan untuk AWS Clean Rooms.

Aturan analisis agregasi

Pembatasan kueri yang memungkinkan kueri yang menggunakan analisis agregat COUNT, SUM, atau AVG berfungsi sepanjang dimensi opsional. Kueri ini tidak akan mengungkapkan informasi tingkat baris.

Mendukung kasus penggunaan seperti perencanaan kampanye, jangkauan media, frekuensi, dan pengukuran konversi.

Jenis aturan analisis lainnya adalah [kustom](#) dan [daftar](#).

Aturan analisis

Pembatasan kueri yang mengotorisasi jenis kueri tertentu.

Jenis aturan analisis menentukan jenis analisis apa yang dapat dijalankan pada tabel yang dikonfigurasi. Setiap jenis memiliki struktur kueri yang telah ditentukan sebelumnya. Anda mengontrol bagaimana kolom tabel Anda dapat digunakan dalam struktur melalui kontrol kueri.

Jenis aturan analisis adalah [agregasi](#), [daftar](#), dan [kustom](#).

Template analisis

Kueri khusus kolaborasi dan disetujui sebelumnya yang dapat digunakan kembali.

Format yang didukung: kode SQL atau kode Python untuk Spark.

Jika menggunakan SQL, template analisis dapat berisi parameter di mana pun nilai literal biasanya dapat muncul dalam kueri SQL. Untuk informasi selengkapnya tentang tipe parameter yang didukung, lihat [Tipe data](#) di Referensi AWS Clean Rooms SQL.

Template analisis hanya berfungsi dengan [aturan analisis khusus](#).

Klien enkripsi C3R

Klien enkripsi Cryptographic Computing for Clean Rooms (C3R).

Digunakan untuk mengenkripsi dan mendekripsi data, C3R adalah SDK enkripsi sisi klien dengan antarmuka baris perintah.

Kolom Cleartext

Kolom yang tidak dilindungi secara kriptografis untuk konstruksi JOIN atau SELECT SQL.

Kolom Cleartext dapat digunakan di bagian manapun dari query SQL.

Kolaborasi

Batas logis yang aman AWS Clean Rooms di mana anggota dapat melakukan kueri SQL pada tabel yang dikonfigurasi.

Kolaborasi dibuat oleh [pencipta kolaborasi](#).

Hanya anggota yang telah diundang ke kolaborasi yang dapat bergabung dalam kolaborasi.

Kolaborasi hanya dapat memiliki satu [anggota yang dapat melakukan kueri](#) data atau satu [anggota yang dapat menjalankan kueri dan pekerjaan](#).

Kolaborasi hanya dapat memiliki satu [anggota yang dapat menerima hasil](#).

Kolaborasi hanya dapat memiliki satu [anggota yang membayar biaya komputasi kueri atau satu anggota yang membayar biaya kueri dan komputasi pekerjaan](#).

Semua anggota dapat melihat daftar peserta yang diundang dalam kolaborasi sebelum mereka bergabung dalam kolaborasi.

Pencipta kolaborasi

Anggota yang menciptakan kolaborasi.

Hanya ada satu pembuat kolaborasi per kolaborasi.

Hanya pembuat kolaborasi yang dapat menghapus anggota dari kolaborasi atau menghapus kolaborasi.

Tabel yang dikonfigurasi

Setiap tabel dikonfigurasi mewakili referensi ke tabel yang ada di AWS Glue Data Catalog yang telah dikonfigurasi untuk digunakan dalam AWS Clean Rooms. Tabel yang dikonfigurasi berisi aturan analisis yang menentukan bagaimana data dapat digunakan.

Saat ini, AWS Clean Rooms mendukung data asosiasi yang disimpan di Amazon Simple Storage Service (Amazon S3) yang dikatalogkan melalui katalog AWS Glue.

Untuk informasi selengkapnya AWS Glue, lihat [Panduan AWS Glue Pengembang](#).

Tabel yang dikonfigurasi dapat dikaitkan dengan satu atau lebih kolaborasi.

Note

AWS Clean Rooms saat ini tidak mendukung lokasi bucket Amazon S3 yang terdaftar. AWS Lake Formation

Aturan analisis kustom

Pembatasan kueri yang memungkinkan serangkaian kueri tertentu yang telah disetujui sebelumnya ([templat analisis](#)) atau memungkinkan serangkaian akun tertentu yang dapat memberikan kueri atau pekerjaan yang menggunakan data Anda.

Mendukung kasus penggunaan seperti atribusi sentuhan pertama, analisis inkremental, dan analisis penemuan audiens.

Mendukung privasi diferensial.

Jenis aturan analisis lainnya adalah [agregasi](#) dan [daftar](#).

Dekripsi

Proses mengubah data terenkripsi kembali ke bentuk aslinya. Dekripsi hanya dapat dilakukan jika Anda memiliki akses ke kunci rahasia.

Privasi diferensial

Teknik matematikal-ketat yang melindungi data kolaborasi dari anggota yang dapat menerima hasil belajar tentang individu tertentu.

Enkripsi

Proses pengkodean data ke dalam bentuk yang muncul acak menggunakan nilai rahasia yang disebut kunci. Tidak mungkin untuk menentukan plaintext asli tanpa akses ke kunci.

Kolom sidik jari

Kolom yang dilindungi secara kriptografi untuk konstruksi JOIN SQL.

Metode alur kerja pemetaan ID

Bagaimana Anda ingin pemetaan ID dilakukan.

Ada dua metode alur kerja pemetaan ID:

- Pemetaan ID berbasis aturan — Metode yang digunakan untuk menggunakan aturan pencocokan untuk menerjemahkan data pihak pertama dari sumber ke target dalam alur kerja pemetaan ID.
- Pemetaan ID layanan penyedia — Metode yang digunakan untuk menggunakan layanan penyedia untuk menerjemahkan data yang disandikan pihak ketiga dari sumber ke target dalam alur kerja pemetaan ID.

AWS Clean Rooms saat ini mendukung LiveRamp sebagai metode alur kerja pemetaan ID berbasis layanan penyedia. Anda harus berlangganan AWS Data Exchange untuk LiveRamp menggunakan metode ini. Untuk informasi selengkapnya, lihat [Berlangganan layanan penyedia AWS Data Exchange di](#) Panduan Resolusi Entitas AWS Pengguna.

Tabel pemetaan ID

Sumber daya AWS Clean Rooms yang memungkinkan aturan pencocokan pihak pertama atau transcoding identitas multi-pihak dalam kolaborasi.

Tabel pemetaan ID adalah referensi ke tabel yang ada di AWS Glue Data Catalog. Ini berisi [aturan analisis tabel pemetaan ID](#) yang menentukan bagaimana data dapat ditanyakan. AWS Clean Rooms Tabel pemetaan ID dapat dikaitkan dengan satu atau beberapa kolaborasi.

Aturan analisis tabel pemetaan ID

Jenis aturan analisis yang dikelola oleh AWS Clean Rooms dan digunakan untuk menggabungkan data identitas yang berbeda untuk memfasilitasi kueri. Ini secara otomatis ditambahkan ke [tabel pemetaan ID](#) dan tidak dapat diedit. Ini mewarisi perilaku aturan analisis lain dalam kolaborasi — selama aturan analisis tersebut homogen.

Alur kerja pemetaan ID

Pekerjaan pemrosesan data yang memetakan data dari sumber ke target berdasarkan metode [alur kerja pemetaan ID](#) yang ditentukan. Ini menghasilkan [tabel pemetaan ID](#).

Ruang nama ID

[Sumber daya AWS Clean Rooms yang berisi metadata yang menjelaskan kumpulan data di beberapa Akun AWS dan cara menggunakan kumpulan data ini dalam alur kerja pemetaan ID.](#)

Asosiasi namespace ID

Asosiasi sumber daya namespace ID yang membantu Anda menemukan input ke dalam alur kerja [pemetaan ID](#) mereka.

Pekerjaan

Metode untuk mengakses dan menganalisis tabel yang dikonfigurasi dalam kolaborasi menggunakan serangkaian fungsi, kelas, dan variabel yang didukung.

AWS Clean Rooms saat ini mendukung jenis PySpark pekerjaan.

AWS Clean Rooms saat ini mendukung menjalankan pekerjaan menggunakan template PySpark analisis.

Aturan analisis daftar

Pembatasan kueri yang memungkinkan kueri yang menampilkan analisis atribut tingkat baris dari tumpang tindih antara tabel ini dan tabel anggota yang dapat melakukan kueri.

Mendukung kasus penggunaan seperti pengayaan dan pembangunan audiens atau penindasan.

Jenis aturan analisis lainnya adalah [agregasi](#) dan [kebiasaan](#).

Model mirip

Model data penyedia data pelatihan yang memungkinkan penyedia data benih untuk membuat [segmen serupa](#) dari data penyedia data pelatihan yang paling mirip dengan data [benih](#) mereka.

Segmen mirip

Subset dari data pelatihan yang paling mirip dengan [data benih](#).

Anggota

AWS Pelanggan yang merupakan peserta dalam [kolaborasi](#).

Seorang anggota diidentifikasi menggunakan mereka Akun AWS.

Semua anggota dapat menyumbangkan data.

Anggota yang dapat menanyakan

Anggota yang dapat meminta data dalam [kolaborasi](#).

Hanya ada satu anggota yang dapat menanyakan per kolaborasi, dan anggota itu tidak dapat diubah.

Pengguna administratif dapat menggunakan izin AWS Identity and Access Management (IAM) untuk mengontrol prinsip IAM mereka (seperti pengguna atau peran) yang dapat menanyakan data dalam kolaborasi. Untuk informasi selengkapnya, lihat [Membuat peran layanan untuk membaca data dari Amazon S3](#).

Anggota yang dapat menjalankan kueri dan pekerjaan

Anggota yang dapat menjalankan kueri dan pekerjaan pada data dalam [kolaborasi](#).

Hanya ada satu anggota yang dapat menjalankan kueri dan pekerjaan per kolaborasi, dan anggota itu tidak dapat diubah.

Pengguna administratif dapat menggunakan izin AWS Identity and Access Management (IAM) untuk mengontrol prinsip IAM mereka (seperti pengguna atau peran) yang dapat menjalankan kueri dan pekerjaan dalam kolaborasi. Untuk informasi selengkapnya, lihat [Membuat peran layanan untuk membaca data dari Amazon S3](#).

Anggota yang dapat menerima hasil

Anggota yang dapat menerima hasil kueri. Anggota yang dapat menerima hasil menentukan setelan hasil kueri untuk tujuan Amazon S3 dan format hasil kueri (CSV atau Parquet).

Untuk analisis dengan mesin analitik Spark, anggota yang dapat menerima hasil juga menentukan apakah file harus dikeluarkan sebagai file tunggal atau sebagai beberapa file.

Ada lebih dari satu anggota yang dapat menerima hasil dalam kolaborasi.

Anggota membayar biaya komputasi kueri

Anggota yang bertanggung jawab untuk membayar biaya komputasi kueri.

Hanya ada satu anggota yang bertanggung jawab untuk membayar biaya komputasi kueri per kolaborasi, dan anggota itu tidak dapat diubah.

Jika pembuat kolaborasi belum menetapkan siapa pun sebagai anggota yang membayar biaya komputasi kueri, maka [anggota yang dapat melakukan kueri](#) adalah pembayar default.

Anggota yang membayar biaya komputasi kueri menerima tagihan untuk kueri yang telah dijalankan dalam kolaborasi.

Anggota membayar biaya kueri dan komputasi pekerjaan

Anggota yang bertanggung jawab untuk membayar permintaan dan biaya komputasi pekerjaan.

Hanya ada satu anggota yang bertanggung jawab untuk membayar kueri dan biaya komputasi pekerjaan per kolaborasi, dan anggota itu tidak dapat diubah.

Jika pembuat kolaborasi belum menentukan siapa pun sebagai anggota yang membayar biaya kueri dan komputasi pekerjaan, maka [anggota yang dapat melakukan kueri](#) adalah pembayar default.

Anggota yang membayar biaya kueri dan komputasi pekerjaan menerima tagihan untuk kueri yang telah dijalankan dalam kolaborasi.

Keanggotaan

Sumber daya yang dibuat saat [anggota](#) bergabung dengan [kolaborasi](#).

Semua sumber daya yang diasosiasikan anggota untuk kolaborasi adalah bagian dari keanggotaan atau terkait dengan keanggotaan.

Hanya anggota yang memiliki keanggotaan yang dapat menambah, menghapus, atau mengedit sumber daya dalam keanggotaan tersebut.

Kolom tertutup

Kolom yang dilindungi secara kriptografi untuk konstruksi SELECT SQL.

Data benih

Data penyedia data benih, yang digunakan untuk membuat segmen yang [mirip](#). Data benih dapat diberikan secara langsung atau dapat berasal dari hasil AWS Clean Rooms kueri. Output segmen mirip adalah sekumpulan pengguna dari data pelatihan yang paling mirip dengan pengguna benih.

Mesin analitik percikan

Opsi analitik AWS Clean Rooms yang memungkinkan pelanggan menjalankan kueri kompleks pada kumpulan data besar yang disimpan di Amazon S3, Amazon Athena, atau Snowflake menggunakan fungsi Apache Spark SQL. Ini juga mendukung PySpark analisis di AWS Clean Rooms.

Saat Anda membuat kolaborasi menggunakan [CreateCollaborationAPI](#), nilai mesin analitik Spark adalah SPARK.

Kueri

Metode untuk mengakses dan menganalisis tabel yang dikonfigurasi dalam kolaborasi, menggunakan serangkaian fungsi, kelas, dan variabel yang didukung.

AWS Clean Rooms saat ini mendukung bahasa query SQL.

AWS Clean Rooms saat ini mendukung menjalankan kueri SQL langsung atau menjalankan kueri menggunakan template analisis SQL.

Terjemahan disediakan oleh mesin penerjemah. Jika konten terjemahan yang diberikan bertentangan dengan versi bahasa Inggris aslinya, utamakan versi bahasa Inggris.