



Handbuch „Erste Schritte“

Amazon Redshift



Amazon Redshift: Handbuch „Erste Schritte“

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

Die Handelsmarken und Handelsaufmachung von Amazon dürfen nicht in einer Weise in Verbindung mit nicht von Amazon stammenden Produkten oder Services verwendet werden, auf eine Art und Weise, dass Kunden irreführt werden könnten oder Amazon schlecht gemacht oder diskreditiert werden könnte. Alle anderen Handelsmarken, die nicht im Besitz von Amazon sind, gehören den jeweiligen Besitzern, die möglicherweise zu Amazon gehören oder nicht, mit Amazon verbunden sind oder von Amazon gesponsert werden.

Table of Contents

.....	v
Erste Schritte mit Serverless-Data Warehouses	1
Melde dich an für AWS	1
Erstellen eines Data Warehouse mit Amazon Redshift Serverless	2
Laden von Beispieldaten	4
Ausführen von Beispielabfragen	7
Laden von Daten aus Amazon S3	9
Erste Schritte mit bereitgestellten Data Warehouses	17
Melden Sie sich an für AWS	20
Festlegen von Firewall-Regeln	20
Schritt 1: Einen Beispiel-Cluster erstellen	21
Schritt 2: Regeln für eingehenden Datenverkehr für SQL-Clients konfigurieren	24
Schritt 3: Gewähren von Zugriff auf einen SQL-Client und Ausführen von Abfragen	25
Gewähren des Zugriffs auf Abfrage-Editor v2	26
Schritt 4: Daten aus Amazon S3 in Amazon Redshift laden	27
Laden von Daten aus Amazon S3 mit SQL-Befehlen	27
Laden von Daten aus Amazon S3 mit Query Editor v2	29
Erstellen von TICKIT-Daten in Ihrem Cluster	30
Schritt 5: Beispielabfragen mit dem Abfrage-Editor testen	31
Schritt 6: Umgebung zurücksetzen	32
Definieren und Verwenden einer Datenbank in Ihrem Data Warehouse	34
Verbinden mit Amazon Redshift	35
Erstellen einer -Datenbank	36
Erstellen eines Benutzers	37
Erstellen eines Schemas	37
Erstellen einer Tabelle	39
Einfügen von Datenzeilen in eine Tabelle	40
Auswahl von Daten aus einer Tabelle	40
Laden von Daten	41
Abfragen der Systemtabellen und Ansichten	41
Anzeigen einer Liste von Tabellennamen	42
Anzeigen von Benutzern	43
Anzeigen aktueller Abfragen	44
Feststellen der Sitzungs-ID einer laufenden Abfrage	45

Abbrechen einer Abfrage	45
Abbrechen einer Abfrage mit der Superuser-Warteschlange	48
Abfragen von Daten, die sich nicht in Ihrer Amazon-Redshift-Datenbank befinden	49
Abfragen von Data Lakes	49
Abfragen von Remote-Datenquellen	50
Zugriff auf Daten in anderen Datenbanken	50
Training von ML-Modellen mit Redshift-Daten	51
Über Amazon-Redshift-Konzepte	52
Weitere Lernressourcen	56
Dokumentverlauf	58

Amazon Redshift unterstützt UDFs ab Patch 198 nicht mehr die Erstellung von neuem Python. Das bestehende Python UDFs wird bis zum 30. Juni 2026 weiterhin funktionieren. Weitere Informationen finden Sie im [Blog-Posting](#).

Die vorliegende Übersetzung wurde maschinell erstellt. Im Falle eines Konflikts oder eines Widerspruchs zwischen dieser übersetzten Fassung und der englischen Fassung (einschließlich infolge von Verzögerungen bei der Übersetzung) ist die englische Fassung maßgeblich.

Erste Schritte mit Data Warehouses von Amazon Redshift Serverless

Wenn Sie Amazon Redshift Serverless zum ersten Mal verwenden, empfehlen wir Ihnen, die folgenden Abschnitte zu lesen, um Ihnen den Einstieg in die Verwendung von Amazon Redshift Serverless zu erleichtern. Der grundlegende Ablauf von Amazon Redshift Serverless besteht darin, Serverless-Ressourcen zu erstellen, eine Verbindung zu Amazon Redshift Serverless herzustellen, Beispieldaten zu laden und dann Abfragen für die Daten auszuführen. Bei Verwendung dieses Handbuchs haben Sie die Möglichkeit, Beispieldaten aus Amazon Redshift Serverless oder aus einem Amazon-S3-Bucket zu laden. Die Beispieldaten werden in der gesamten Amazon-Redshift-Dokumentation verwendet, um Features zu demonstrieren. Für die ersten Schritte bei der Verwendung von Amazon Redshift bereitgestellter Data Warehouses vgl. [Erste Schritte mit von Amazon Redshift bereitgestellten Data Warehouses](#).

- [the section called “Melde dich an für AWS”](#)
- [the section called “Erstellen eines Data Warehouse mit Amazon Redshift Serverless”](#)
- [the section called “Laden von Daten aus Amazon S3”](#)

Melde dich an für AWS

Wenn Sie noch kein AWS Konto haben, registrieren Sie sich für eines. Wenn Sie bereits ein Konto besitzen, können Sie diesen Schritt überspringen und Ihr vorhandenes Konto verwenden.

1. Öffne <https://portal.aws.amazon.com/billing/die-Registrierung>.
2. Folgen Sie den Online-Anweisungen.

Wenn Sie sich für ein AWS Konto registrieren, wird ein Root-Benutzer für das AWS Konto erstellt. Der Root-Benutzer hat Zugriff auf alle AWS Dienste und Ressourcen im Konto. Als bewährte Methode zur Gewährleistung der Sicherheit sollten Sie den [administrativen Zugriff einem administrativen Benutzer zuweisen](#) und nur den Root-Benutzer verwenden, um [Aufgaben auszuführen, die Root-Benutzerzugriff erfordern](#).

Erstellen eines Data Warehouse mit Amazon Redshift Serverless

Wenn Sie sich zum ersten Mal bei der Amazon-Redshift-Serverless-Konsole anmelden, werden Sie aufgefordert, auf die Informationen zu den ersten Schritten zuzugreifen, die Sie zum Erstellen und Verwalten von Serverless-Ressourcen verwenden können. In diesem Handbuch werden Sie Serverless-Ressourcen unter Verwendung der Standardeinstellungen von Amazon Redshift Serverless erstellen.

Wenn Sie Ihre Einrichtung genauer kontrollieren möchten, wählen Sie **Customize settings** (Einstellungen anpassen) aus.

Note

Redshift Serverless erfordert eine Amazon-VPC mit drei Subnetzen in drei verschiedenen Availability Zones. Redshift Serverless benötigt außerdem mindestens 3 verfügbare IP-Adressen. Stellen Sie sicher, dass die Amazon VPC, die Sie für Redshift Serverless verwenden, drei Subnetze in drei verschiedenen Availability Zones und mindestens 3 verfügbare IP-Adressen hat, bevor Sie fortfahren. Weitere Informationen zum Erstellen von Subnetzen finden Sie unter [Erstellen eines Subnetzes](#) im Benutzerhandbuch von Amazon Virtual Private Cloud. Weitere Informationen zu IP-Adressen in einer Amazon VPC finden Sie unter [IP-Adressierung für Ihre VPCs und Subnetze](#).

So nehmen Sie die Konfiguration mit Standardeinstellungen vor:

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die Amazon Redshift Redshift-Konsole unter <https://console.aws.amazon.com/redshiftv2/>.

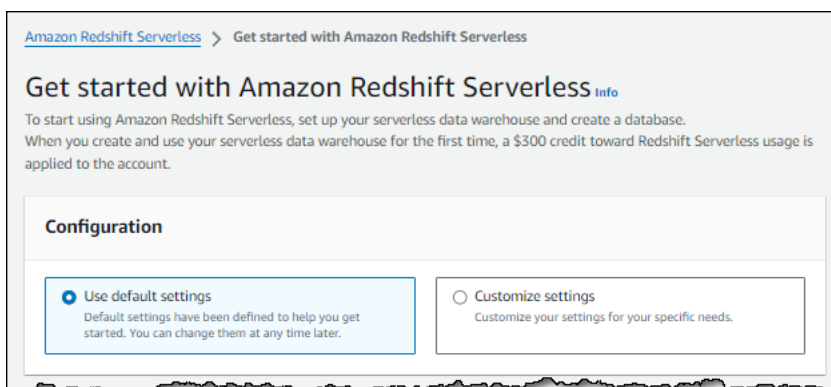
Wählen Sie **Redshift Serverless Free Trial testen**.

2. Wählen Sie unter **Configuration** (Konfiguration) die Option **Use default settings** (Standardeinstellungen verwenden) aus. Amazon Redshift Serverless erstellt einen Standard-Namespace mit einer zugeordneten Standardarbeitsgruppe. Wählen Sie **Save configuration** (Konfiguration speichern) aus.

Note

Ein Namespace ist eine Sammlung von Datenbankobjekten und Benutzern. In Namespaces sind alle Ressourcen zusammengefasst, die Sie in Amazon Redshift Serverless verwenden, wie Schemata, Tabellen, Benutzer, Datashares und Snapshots. Eine Arbeitsgruppe ist eine Sammlung von Rechenressourcen. In Arbeitsgruppen sind Rechenressourcen enthalten, die Amazon Redshift Serverless zur Ausführung von Datenverarbeitungsaufgaben verwendet.

Der folgende Screenshot zeigt die Standardeinstellungen für Amazon Redshift Serverless.



3. Nachdem die Einrichtung abgeschlossen ist, wählen Sie Continue (Weiter), um zu Serverless Dashboard zu wechseln. Wie Sie sehen, sind die Serverless-Arbeitsgruppe und der Serverless-Namespace verfügbar.

Serverless dashboard Info

Namespace overview Info

Namespace data from your account

Total snapshots

0

Datashares in my account

0

Datashares requiring authorization

0

Datashares fr

0

Namespaces / Workgroups Info

Namespace

Status

Workgroup

Status

default

✔ Available

default

✔ Available

Note

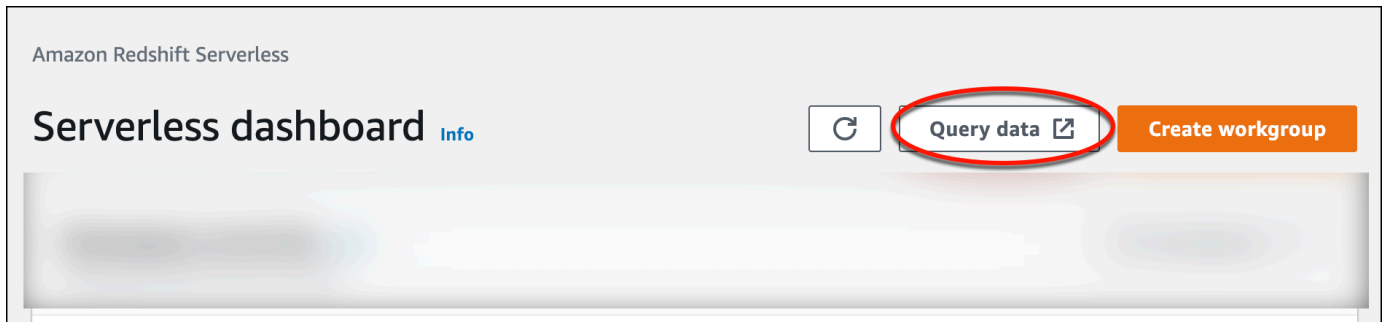
Wenn Redshift Serverless die Arbeitsgruppe nicht erfolgreich erstellt, können Sie wie folgt vorgehen:

- Beheben Sie alle Fehler, die Redshift Serverless meldet, wie z. B. zu wenige Subnetze in Ihrer Amazon VPC.
- Löschen Sie den Namespace, indem Sie default-namespace im Redshift-Serverless-Dashboard und dann Aktionen, Namespace löschen auswählen. Das Löschen eines Namespaces dauert mehrere Minuten.
- Wenn Sie die Redshift-Serverless-Konsole erneut öffnen, wird der Begrüßungsbildschirm angezeigt.

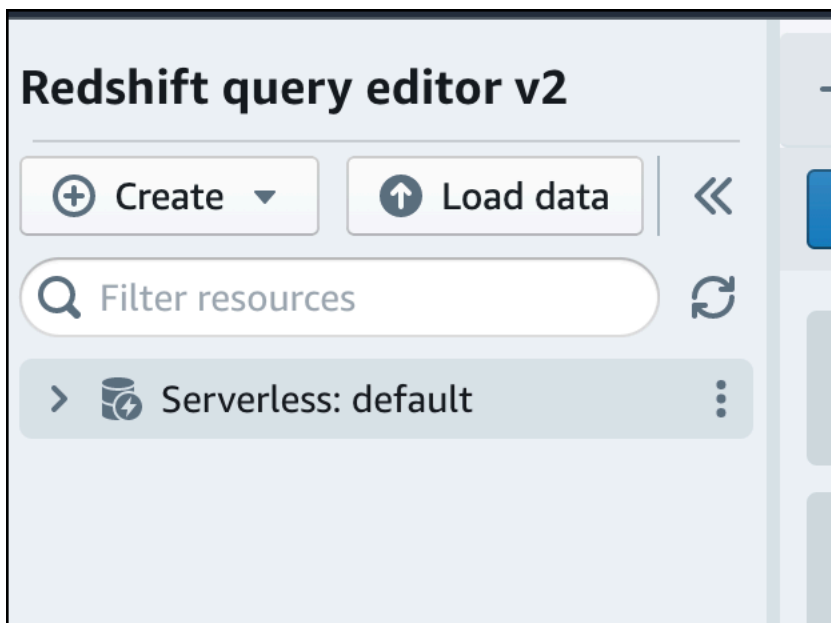
Laden von Beispieldaten

Nachdem Sie Ihr Data Warehouse mit Amazon Redshift Serverless eingerichtet haben, können Sie den Amazon Redshift Query Editor v2 verwenden, um Beispieldaten zu laden.

1. Um Query Editor v2 über die Amazon-Redshift-Serverless-Konsole zu starten, wählen Sie Daten abfragen aus. Wenn Sie den Abfrage-Editor v2 über die Amazon-Redshift-Serverless-Konsole aufrufen, wird er in einer neuen Browser-Registerkarte geöffnet. Der Abfrage-Editor v2 stellt eine Verbindung von Ihrem Clientcomputer mit der Amazon-Redshift-Serverless-Umgebung her.

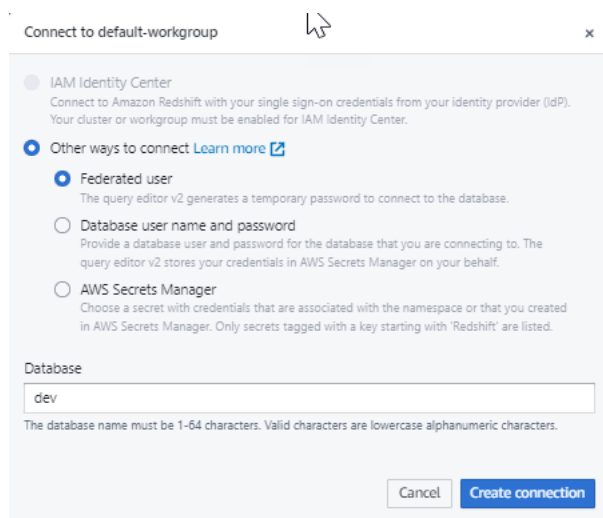


2. Für dieses Handbuch verwenden Sie Ihr AWS Administratorkonto und das AWS KMS key Standardkonto. Informationen zur Konfiguration von Query Editor v2, einschließlich der erforderlichen Berechtigungen, finden Sie unter [Konfigurieren Ihres AWS-Konto](#) im Managementleidfaden zu Amazon Redshift. Informationen zur Konfiguration von Amazon Redshift für die Verwendung eines vom Kunden verwalteten Schlüssels oder zur Änderung des von Amazon Redshift verwendeten KMS-Schlüssels finden Sie unter [Ändern des AWS KMS Schlüssels für einen Namespace](#).
3. Um eine Verbindung zu einer Arbeitsgruppe herzustellen, wählen Sie den Namen der Arbeitsgruppe im Strukturansichtsbereich aus.



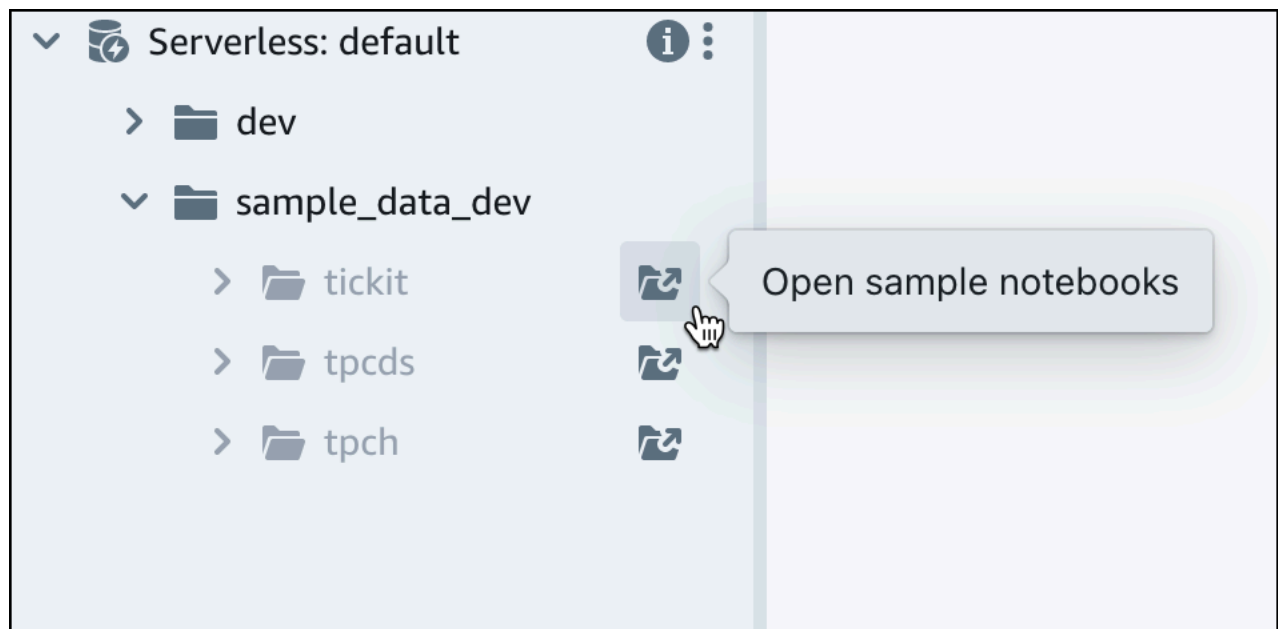
4. Wenn Sie in Query Editor v2 zum ersten Mal eine Verbindung zu einer neuen Arbeitsgruppe herstellen, müssen Sie den Authentifizierungstyp auswählen, der für die Verbindung

zur Arbeitsgruppe verwendet werden soll. Lassen Sie für diese Anleitung die Option **Verbundbenutzer** ausgewählt und wählen Sie **Verbindung erstellen** aus.



Sobald Sie verbunden sind, können Sie Beispieldaten aus Amazon Redshift Serverless oder aus einem Amazon-S3-Bucket laden.

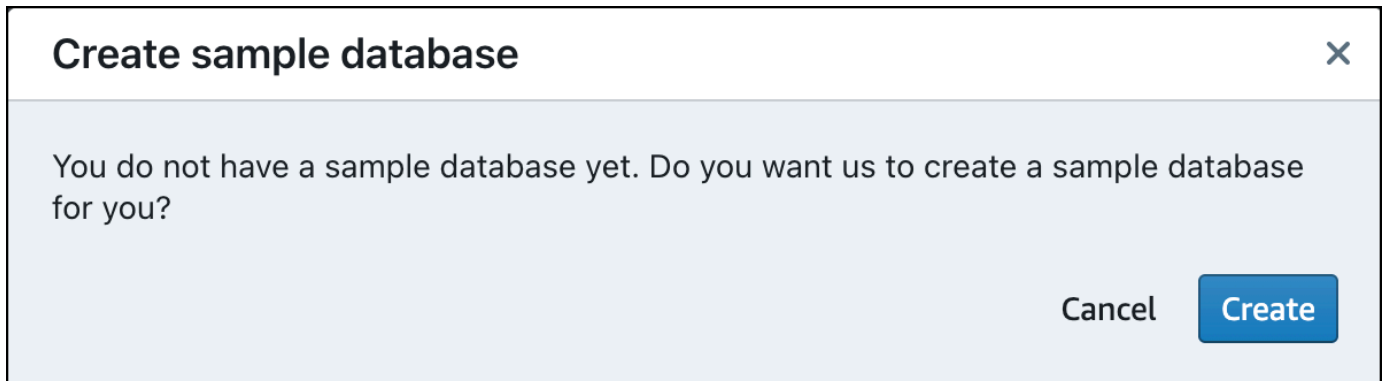
- Erweitern Sie unter der Standardarbeitsgruppe von Amazon Redshift Serverless die Datenbank `sample_data_dev`. Es gibt drei Beispielschemata, die drei Beispieldatensätzen entsprechen, die Sie in die Amazon-Redshift-Serverless-Datenbank laden können. Wählen Sie den Beispieldatensatz, den Sie laden möchten, und dann Beispiel-Notebooks öffnen aus.



Note

Ein SQL-Notizbuch enthält SQL- und Markdown-Zellen. Sie können damit mehrere SQL-Befehle in einem einzigen Dokument organisieren, kommentieren und freigeben.

6. Wenn Sie zum ersten Mal Daten laden, fordert Query Editor v2 Sie auf, eine Beispieldatenbank zu erstellen. Wählen Sie Erstellen aus.



Ausführen von Beispielabfragen

Nachdem Sie Amazon Redshift Serverless eingerichtet haben, können Sie einen Beispieldatensatz in Amazon Redshift Serverless verwenden. Amazon Redshift Serverless lädt den Beispieldatensatz, z. B. den Ticket-Datensatz, automatisch und Sie können die Daten sofort abfragen.

- Sobald Amazon Redshift Serverless mit dem Laden der Beispieldaten fertig ist, werden alle Beispielabfragen in den Editor geladen. Sie können Alle ausführen auswählen, um alle Abfragen aus den Beispiel-Notebooks auszuführen.

The screenshot shows the Amazon Redshift console interface. At the top, the query title is "Sales per event". Below the title, there are controls for "Run" and "Limit 100". The SQL query is displayed in a text area:

```

1 SET search_path to tickit;
2 SELECT eventname, total_price
3 FROM (SELECT eventid, total_price, ntile(1000) over(order by total_price desc) as percentile
4       FROM (SELECT eventid, sum(pricepaid) total_price
5             FROM tickit.sales
6             GROUP BY eventid)) Q, tickit.event E
7 WHERE Q.eventid = E.eventid
8      AND percentile = 1
9 ORDER BY total_price desc;

```

Below the query, the results are shown in a table with two columns: "eventname" and "total_price". The table contains 9 rows of data:

eventname	total_price
Adriana Lecouvreur	51846
Janet Jackson	51049
Phantom of the Opera	50301
The Little Mermaid	49956
Citizen Cope	49823
Sevendust	48020
Electra	47883
Mary Poppins	46780
Live	46661

At the bottom right of the results area, it says "Elapsed time: 401 ms Total rows: 9".

Sie können die Ergebnisse auch als JSON- oder CSV-Datei exportieren oder die Ergebnisse in einem Diagramm anzeigen.

This screenshot shows the same Amazon Redshift console interface as the previous one, but with the "Export" and "Chart" options highlighted with a red circle. The "Export" button is a dropdown menu, and the "Chart" button is a toggle switch that is currently turned on. Below the query, the "Structure" panel is visible on the left, and a "Traces" panel is open, showing a "trace 0" with a "Line" type. The main area shows a chart with a blue line graph. The Y-axis is labeled "Click to enter Y axis title" and has values 2, 4, 6, 8. The X-axis is labeled "Click to enter X axis title".

Sie können Daten auch aus einem Amazon-S3-Bucket laden. Weitere Informationen hierzu finden Sie unter [the section called “Laden von Daten aus Amazon S3”](#).

Laden von Daten aus Amazon S3

Nachdem Sie Ihr Data Warehouse erstellt haben, können Sie Daten aus Amazon S3 laden.

An diesem Punkt verfügen Sie über eine Datenbank namens dev. Als Nächstes legen Sie Tabellen in der Datenbank an, laden Daten in die Tabellen hoch und führen testweise eine Abfrage durch. Die Beispieldaten werden der Einfachheit halber in einem Amazon-S3-Bucket bereitgestellt.

1. Vor dem Laden von Daten aus Amazon S3 müssen Sie zunächst eine IAM-Rolle mit den erforderlichen Berechtigungen erstellen und Ihrem Serverless-Namespace hinzufügen. Kehren Sie dazu zur Redshift Serverless-Konsole zurück und wählen Sie Namespace-Konfiguration. Wählen Sie dazu im Navigationsmenü Ihren Namespace und dann Sicherheit und Verschlüsselung. Wählen Sie IAM-Rollen verwalten.

The screenshot shows the Amazon Redshift console interface for a namespace named 'default'. The 'Security and encryption' tab is selected and highlighted with a red circle. The console displays general information about the namespace, including its ID, ARN, status, and storage usage.

General information	
Namespace	default
Namespace ID	example-namespace-id
Namespace ARN	example-namespace-arn
Status	Available
Date created	March 02, 2023, 12:11 (UTC-08:00)
Storage used	18.9 GB

Navigation tabs: Workgroup, Data backup, **Security and encryption**, Datashares, Tags

Workgroup name
Set up compute resources for your workgroup.

Workgroup	default
Status	Available

2. Erweitern Sie das Menü IAM-Rollen verwalten und wählen Sie IAM-Rolle erstellen aus.

Manage IAM roles

Permissions

i Associate an IAM role so that your serverless endpoint can LOAD and UNLOAD data. You can create an IAM role as the default for this configuration that has the [AmazonRedshiftAllCommandsFullAccess](#) policy attached. This policy includes permissions to run SQL commands to COPY, UNLOAD, and query data with Amazon Redshift Serverless. This policy also grants permissions to run SELECT statements for related services, such as Amazon S3, Amazon CloudWatch logs, Amazon SageMaker, and AWS Glue. You won't be able to run these SQL commands without an IAM role attached to your namespace.

Associated IAM roles (1)

Create, associate, or remove an IAM role. You can associate up to 50 IAM roles. You can also choose an IAM role and set it as the default.

Set default ▼

Manage IAM roles ▲



Search for associa

Associate IAM roles

Create IAM role

Remove IAM roles

or role type

< 1 >



IAM roles [↗](#)



Status



Role
type



3. Wählen Sie die Ebene des S3-Bucket-Zugriffs aus, die Sie dieser Rolle gewähren möchten, und wählen Sie IAM-Rolle als Standard erstellen aus.

Create the default IAM role ✕

i Associate an IAM role so that your serverless endpoint can LOAD and UNLOAD data. You can create an IAM role as the default for this configuration that has the [AmazonRedshiftAllCommandsFullAccess](#) policy attached. This policy includes permissions to run SQL commands to COPY, UNLOAD, and query data with Amazon Redshift Serverless. This policy also grants permissions to run SELECT statements for related services, such as Amazon S3, Amazon CloudWatch logs, Amazon SageMaker, and AWS Glue. You won't be able to run these SQL commands without an IAM role attached to your namespace.

Specify an S3 bucket for the IAM role to access
To create a new bucket, [visit S3](#)

No additional S3 bucket
Create the IAM role without specifying S3 buckets.

Any S3 bucket
Allow users that have access to your Redshift Serverless data to also access any S3 bucket and its contents in your AWS account.

Specific S3 buckets
Specify one or more S3 buckets that the IAM role being created has permission to access.

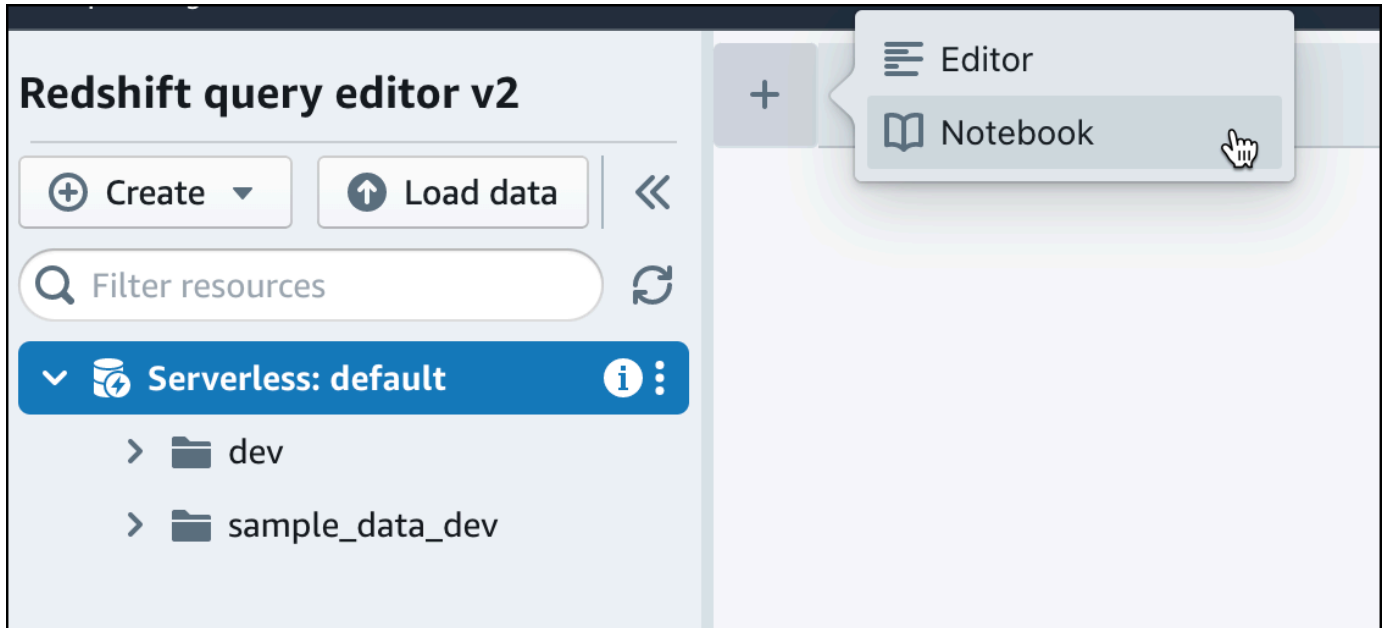
4. Wählen Sie **Änderungen speichern** aus. Sie können jetzt Beispieldaten aus Amazon S3 laden.

In den folgenden Schritten werden Daten in einem öffentlichen Amazon-Redshift-S3-Bucket verwendet, Sie können jedoch dieselben Schritte unter Verwendung Ihres eigenen S3-Buckets und eigener SQL-Befehle wiederholen.

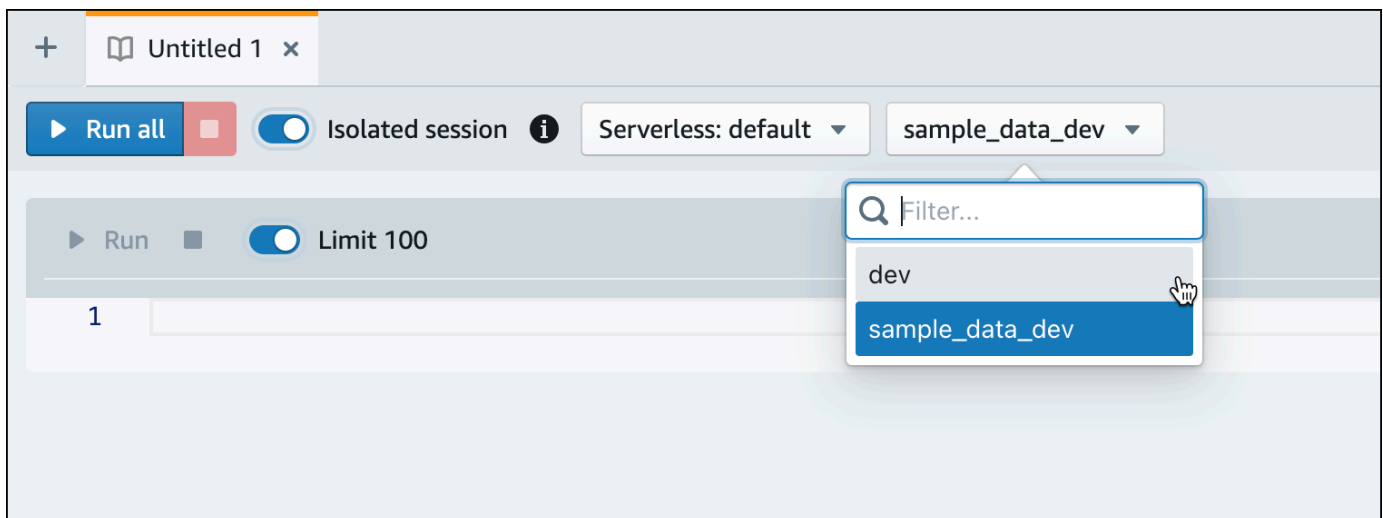
Laden von Beispieldaten aus Amazon S3

1. Wählen Sie in Query Editor v2 „
+

hinzufügen“ und dann Notebook aus, um ein neues SQL-Notebook zu erstellen.



2. Wechseln Sie zur dev-Datenbank.



3. Erstellen Sie Tabellen.

Wenn Sie Query Editor v2 verwenden, kopieren Sie die folgenden Create-Table-Anweisungen und führen Sie sie aus, um Tabellen in der dev-Datenbank zu erstellen. Weitere Informationen

zur Syntax finden Sie unter [CREATE TABLE](#) im Datenbankentwicklerhandbuch zu Amazon Redshift.

```
create table users(  
userid integer not null distkey sortkey,  
username char(8),  
firstname varchar(30),  
lastname varchar(30),  
city varchar(30),  
state char(2),  
email varchar(100),  
phone char(14),  
likesports boolean,  
liketheatre boolean,  
likeconcerts boolean,  
likejazz boolean,  
likeclassical boolean,  
likeopera boolean,  
likerock boolean,  
likevegas boolean,  
likebroadway boolean,  
likemusicals boolean);
```

```
create table event(  
eventid integer not null distkey,  
venueid smallint not null,  
catid smallint not null,  
dateid smallint not null sortkey,  
eventname varchar(200),  
starttime timestamp);
```

```
create table sales(  
salesid integer not null,  
listid integer not null distkey,  
sellerid integer not null,  
buyerid integer not null,  
eventid integer not null,  
dateid smallint not null sortkey,  
qtysold smallint not null,  
pricepaid decimal(8,2),  
commission decimal(8,2),  
saletime timestamp);
```

4. Erstellen Sie in Query Editor v2 eine neue SQL-Zelle in Ihrem Notebook.



5. Verwenden Sie nun den Befehl COPY in Query Editor v2, um große Datensätze aus Amazon S3 oder Amazon DynamoDB in Amazon Redshift zu laden. Weitere Informationen zur COPY-Syntax finden Sie unter [COPY](#) im Datenbankentwicklerhandbuch zu Amazon Redshift.

Sie können den Befehl COPY mit Beispieldaten ausführen, die in einem öffentlichen S3-Bucket verfügbar sind. Führen Sie die folgenden SQL-Befehle in Query Editor v2 aus.

```
COPY users
FROM 's3://redshift-downloads/tickit/allusers_pipe.txt'
DELIMITER '|'
TIMEFORMAT 'YYYY-MM-DD HH:MI:SS'
IGNOREHEADER 1
REGION 'us-east-1'
IAM_ROLE default;

COPY event
FROM 's3://redshift-downloads/tickit/allevents_pipe.txt'
DELIMITER '|'
TIMEFORMAT 'YYYY-MM-DD HH:MI:SS'
IGNOREHEADER 1
REGION 'us-east-1'
IAM_ROLE default;

COPY sales
FROM 's3://redshift-downloads/tickit/sales_tab.txt'
DELIMITER '\t'
TIMEFORMAT 'MM/DD/YYYY HH:MI:SS'
IGNOREHEADER 1
REGION 'us-east-1'
IAM_ROLE default;
```

6. Erstellen Sie nach dem Laden der Daten eine weitere SQL-Zelle in Ihrem Notebook und probieren Sie einige Beispielabfragen aus. Weitere Informationen zur Verwendung des SELECT-Befehls finden Sie unter [SELECT](#) im Amazon-Redshift-Entwicklerhandbuch. Verwenden Sie Query Editor v2, um die Struktur und die Schemata der Beispieldaten zu verstehen.

```
-- Find top 10 buyers by quantity.
SELECT firstname, lastname, total_quantity
FROM (SELECT buyerid, sum(qtysold) total_quantity
      FROM sales
      GROUP BY buyerid
      ORDER BY total_quantity desc limit 10) Q, users
WHERE Q.buyerid = userid
ORDER BY Q.total_quantity desc;

-- Find events in the 99.9 percentile in terms of all time gross sales.
SELECT eventname, total_price
FROM (SELECT eventid, total_price, ntile(1000) over(order by total_price desc) as
      percentile
      FROM (SELECT eventid, sum(pricepaid) total_price
            FROM sales
            GROUP BY eventid)) Q, event E
WHERE Q.eventid = E.eventid
AND percentile = 1
ORDER BY total_price desc;
```

Nachdem Sie nun Daten geladen und einige Beispielabfragen ausgeführt haben, können Sie andere Bereiche von Amazon Redshift Serverless erkunden. In der folgenden Übersicht erfahren Sie mehr über die Verwendungsmöglichkeiten von Amazon Redshift Serverless.

- Sie können Daten aus einem Amazon-S3-Bucket laden. Weitere Informationen finden Sie unter [Laden von Daten aus Amazon S3](#).
- Sie können Query Editor v2 verwenden, um Daten aus einer lokalen zeichengetrennten Datei mit weniger als 5 MB zu laden. Weitere Informationen finden Sie unter [Laden von Daten aus einer lokalen Datei](#).
- Sie können eine Verbindung zu Amazon Redshift Serverless mit SQL-Tools von Drittanbietern mit dem JDBC- und ODBC-Treiber herstellen. Weitere Informationen finden Sie unter [Verbinden mit Amazon Redshift Serverless](#).

- Sie können die Amazon-Redshift-Daten-API auch verwenden, um eine Verbindung mit Amazon Redshift Serverless herzustellen. Weitere Informationen finden Sie unter [Verwenden der Amazon-Redshift-Daten-API](#).
- Sie können Ihre Daten in Amazon Redshift Serverless mit Redshift ML verwenden, um Machine-Learning-Modelle mit dem Befehl CREATE MODEL zu erstellen. Im [Tutorial: Erstellen von Kundenabwanderungsmodellen](#) erfahren Sie, wie Sie ein Redshift-ML-Modell erstellen.
- Sie können Daten aus einem Amazon S3 Data Lake abfragen, ohne Daten in Amazon Redshift Serverless laden zu müssen. Weitere Informationen finden Sie unter [Abfragen eines Data Lake](#).

Erste Schritte mit von Amazon Redshift bereitgestellten Data Warehouses

Wenn Sie Amazon Redshift zum ersten Mal verwenden, empfehlen wir Ihnen, zunächst die folgenden Abschnitte zu lesen, die Ihnen bei den ersten Schritten mit der Verwendung bereitgestellter Cluster helfen. Der grundlegende Ablauf von Amazon Redshift besteht darin, bereitgestellte Ressourcen zu erstellen, eine Verbindung zu Amazon Redshift herzustellen, Beispieldaten zu laden und dann Abfragen für die Daten auszuführen. Bei Verwendung dieses Leitfadens haben Sie die Möglichkeit, Beispieldaten aus Amazon Redshift oder aus einem Amazon-S3-Bucket zu laden. Die Beispieldaten werden in der gesamten Amazon-Redshift-Dokumentation verwendet, um Features zu demonstrieren.

Dieses Tutorial zeigt, wie Sie von Amazon Redshift bereitgestellte Cluster verwenden, bei denen es sich um AWS Data Warehouse-Objekte handelt, für die Sie Systemressourcen verwalten. Sie können Amazon Redshift auch mit Serverless-Arbeitsgruppen verwenden. Dabei handelt es sich um Data Warehouse-Objekte, die je nach Nutzung automatisch skaliert werden. Zu den ersten Schritten mit Redshift Serverless vgl. [Erste Schritte mit Data Warehouses von Amazon Redshift Serverless](#).

Nachdem Sie sich bei der von Amazon Redshift bereitgestellten Konsole angemeldet haben, können Sie alle Amazon-Redshift-Objekte erstellen und verwalten, einschließlich Cluster, Knoten und Datenbanken. Sie können auch Abfragen ausführen und anzeigen und mit einem SQL-Client andere SQL Data Definition Language (DDL)- und Data Manipulation Language (DML)-Operationen durchführen.

Important

Der Cluster, den Sie für diese Übung bereitstellen, wird in einer Live-Umgebung ausgeführt. Solange es läuft, fallen Gebühren für Sie an. AWS-Konto Informationen zu Preisen finden Sie auf der [Amazon-Redshift-Preisseite](#).

Um unnötige Kosten zu vermeiden, sollten Sie den Cluster löschen, wenn Sie damit fertig sind. Der letzte Abschnitt dieses Kapitels erläutert, wie dies geht.

Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die Amazon Redshift Redshift-Konsole unter <https://console.aws.amazon.com/redshiftv2/>.

Wir empfehlen Ihnen, zunächst das Dashboard für bereitgestellte Cluster aufzurufen, um mit der Nutzung der Amazon-Redshift-Konsole zu beginnen.

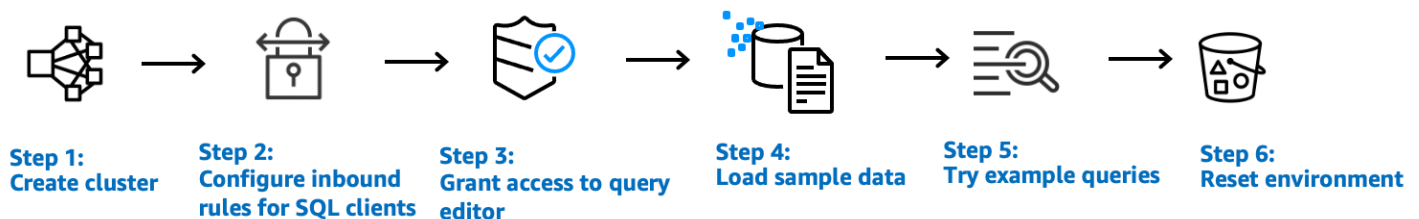
Abhängig von Ihrer Konfiguration werden die folgenden Elemente im Navigationsbereich der von Amazon Redshift bereitgestellten Konsole angezeigt:

- Redshift Serverless – Greifen Sie auf Daten zu und analysieren Sie sie, ohne dass Sie von Amazon Redshift bereitgestellte Cluster einrichten, optimieren und verwalten müssen.
- Dashboard für bereitgestellte Cluster — Sehen Sie sich die Liste der Cluster in Ihrer AWS-Region an. Unter Cluster-Metriken und Abfrageübersicht finden Sie Einblicke in Metrikdaten (wie CPU-Auslastung) und Abfrageinformationen. Damit können Sie feststellen, ob Ihre Leistungsdaten über einen bestimmten Zeitraum anomal sind.
- Cluster — Sehen Sie sich hier Ihre Clusterliste an AWS-Region, wählen Sie einen Cluster aus, um mit der Abfrage zu beginnen, oder führen Sie clusterbezogene Aktionen durch. Sie können von dieser Seite aus auch einen neuen Cluster erstellen.
- Query editor – Führen Sie Abfragen für Datenbanken aus, die von Ihrem Amazon-Redshift-Cluster gehostet werden. Wir empfehlen stattdessen die Verwendung von Query Editor v2.
- Query Editor v2 – Amazon Redshift Query Editor v2 ist eine separate webbasierte SQL-Clientanwendung, mit der Sie Abfragen für Ihr Data Warehouse von Amazon Redshift erstellen und ausführen können. Sie können Ihre Ergebnisse in Diagrammen darstellen und Abfragen mit anderen im Team teilen.
- Queries and loads (Abfragen und Ladevorgänge) – Rufen Sie Informationen als Referenz oder zur Fehlerbehebung auf, z. B. eine Liste der letzten Abfragen und den SQL-Text für jede Abfrage.
- Datashares – Als Administratoren eines Produzentenkontos können Sie entweder Konsumentenkonten zum Zugreifen auf Datashares autorisieren oder keinen Zugriff autorisieren. Um eine autorisierte Datenfreigabe zu verwenden, kann ein Administrator eines Benutzerkontos die Datenfreigabe entweder einem ganzen AWS-Konto oder bestimmten Cluster-Namespaces in einem Konto zuordnen. Ein Administrator kann eine Datenfreigabe auch ablehnen.
- Null-ETL-Integrationen – Verwalten Sie Integrationen, die Transaktionsdaten in Amazon Redshift verfügbar machen, nachdem sie in unterstützte Quellen geschrieben wurden.
- IAM Identity Center-Verbindungen – Konfigurieren Sie die Verbindung zwischen Amazon Redshift und IAM Identity Center.
- Configurations (Konfigurationen) – Stellen Sie über Java-Database-Connectivity(JDBC)- und Open-Database-Connectivity(ODBC)-Verbindungen eine Verbindung von SQL-Client-Tools zu Amazon-Redshift-Clustern her. Sie können auch einen von Amazon Redshift verwalteten Virtual Private

Cloud (VPC)-Endpunkt einrichten. Dadurch wird eine private Verbindung hergestellt zwischen einer VPC, die auf dem Amazon-VPC-Service basiert, der einen Cluster enthält, und einer anderen VPC, in der ein Client-Tool ausgeführt wird.

- AWS Partnerintegration — Erstellen Sie eine Integration mit einem unterstützten AWS Partner.
- Advisor – Erhalten Sie spezifische Empfehlungen zu Änderungen, die Sie an Ihrem Amazon-Redshift-Cluster vornehmen können, um Ihre Optimierungen zu priorisieren.
- AWS Marketplace — Informieren Sie sich über andere Tools oder AWS Services, die mit Amazon Redshift funktionieren.
- Alarms (Alarme) – Erstellen Sie Alarms für Cluster-Metriken, um Leistungsdaten anzuzeigen und Metriken über einen von Ihnen festgelegten Zeitraum nachzuverfolgen.
- Events (Ereignisse) – Verfolgen Sie Ereignisse nach und rufen Sie Berichte mit Informationen wie dem Datum des Ereignisses, einer Beschreibung oder der Ereignisquelle ab.
- What's new (Neuerungen) – Zeigen Sie neue Funktionen und Produktaktualisierungen von Amazon Redshift an.

In diesem Tutorial führen Sie die folgenden Schritte durch:



Themen

- [Melden Sie sich an für AWS](#)
- [Festlegen von Firewall-Regeln](#)
- [Schritt 1: Erstellen eines Amazon-Redshift-Beispielclusters](#)
- [Schritt 2: Regeln für eingehenden Datenverkehr für SQL-Clients konfigurieren](#)
- [Schritt 3: Gewähren von Zugriff auf einen SQL-Client und Ausführen von Abfragen](#)
- [Schritt 4: Daten aus Amazon S3 in Amazon Redshift laden](#)
- [Schritt 5: Beispielabfragen mit dem Abfrage-Editor testen](#)
- [Schritt 6: Umgebung zurücksetzen](#)

Melden Sie sich an für AWS

Wenn Sie noch keine haben AWS-Konto, melden Sie sich für eine an. Wenn Sie bereits ein Konto besitzen, können Sie diesen Schritt überspringen und Ihr vorhandenes Konto verwenden.

1. Öffne <https://portal.aws.amazon.com/billing/die-Anmeldung>.
2. Folgen Sie den Online-Anweisungen.

Während der Anmeldung erhalten Sie einen Telefonanruf oder eine Textnachricht und müssen einen Verifizierungscode über die Telefontasten eingeben.

Wenn Sie sich für eine anmelden AWS-Konto, Root-Benutzer des AWS-Kontos wird eine erstellt. Der Root-Benutzer hat Zugriff auf alle AWS-Services und Ressourcen des Kontos. Als bewährte Sicherheitsmethode weisen Sie einem Benutzer Administratorzugriff zu und verwenden Sie nur den Root-Benutzer, um [Aufgaben auszuführen, die Root-Benutzerzugriff erfordern](#).

Festlegen von Firewall-Regeln

Note

In diesem Tutorial wird davon ausgegangen, dass Ihr Cluster den Standardport 5439 verwendet und Amazon Redshift Query Editor v2 zur Ausführung von SQL-Befehlen verwendet werden kann. Es geht nicht auf Details zu Netzwerkkonfigurationen oder zur Einrichtung eines SQL-Clients ein, die in Ihrer Umgebung erforderlich sein könnten.

In manchen Umgebungen geben Sie einen Port für den Start des Amazon-Redshift-Clusters an. Sie verwenden diesen Port zusammen mit der Endpunkt-URL des Clusters, um auf den Cluster zuzugreifen. Des Weiteren erstellen Sie eine Zugangsregel für eingehenden Datenverkehr in einer Sicherheitsgruppe, die den Zugriff auf den Cluster über den Port regelt.

Wenn der Client-Computer durch eine Firewall geschützt ist, müssen Sie wissen, welcher Port offen ist. Über diesen offenen Port können Sie von einem SQL-Client-Tool eine Verbindung zum Cluster herstellen und Abfragen ausführen. Sollten Sie keinen offenen Port kennen, bitten Sie jemanden, der sich mit den Firewall-Regeln in Ihrem Netzwerk auskennt, einen offenen Port in der Firewall zu ermitteln.

Standardmäßig verwendet Amazon Redshift Port 5439. Eine Verbindung ist aber nur möglich, wenn dieser Port in der Firewall auch geöffnet ist. Sie können die Portnummer für den Amazon-Redshift-Cluster nach der Erstellung nicht mehr ändern. Stellen Sie daher sicher, dass Sie einen offenen Port angeben, der in Ihrer Umgebung beim Startvorgang funktioniert.

Schritt 1: Erstellen eines Amazon-Redshift-Beispielclusters

In diesem Tutorial führen Sie den Prozess zum Erstellen eines Amazon-Redshift-Clusters mit einer Datenbank durch. Anschließend laden Sie einen Datensatz von Amazon S3 in Tabellen in Ihrer Datenbank. Anhand dieses Beispiel-Clusters können Sie den Dienst Amazon Redshift auswerten.

Vergewissern Sie sich vor der Einrichtung eines Amazon-Redshift-Clusters, dass Sie die erforderlichen Voraussetzungen durchgeführt haben, wie etwa [Melden Sie sich an für AWS](#) und [Festlegen von Firewall-Regeln](#).

Für jeden Vorgang, der auf Daten von einer anderen AWS Ressource zugreift, benötigt Ihr Cluster die Erlaubnis, in Ihrem Namen auf die Ressource und die Daten auf der Ressource zuzugreifen. Ein Beispiel ist die Ausführung des SQL-Befehls COPY zum Laden von Daten aus Amazon Simple Storage Service (Amazon S3). Sie stellen diese Berechtigungen mithilfe von AWS Identity and Access Management (IAM) bereit. Sie können dies über eine IAM-Rolle tun, die Sie erstellen und die mit Ihrem Cluster verbunden ist. Weitere Informationen zu Anmeldeinformationen und Zugriffsberechtigungen finden Sie unter [Anmeldeinformationen und Zugriffsberechtigungen](#) im Datenbankentwicklerhandbuch zu Amazon Redshift.

So erstellen Sie einen Amazon-Redshift-Cluster


1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die Amazon Redshift Redshift-Konsole unter <https://console.aws.amazon.com/redshiftv2/>.

Important

Wenn Sie IAM-Benutzeranmeldeinformationen verwenden, stellen Sie sicher, dass Sie über die erforderlichen Berechtigungen verfügen, um die Cluster-Operationen durchzuführen. Weitere Informationen finden Sie unter [Sicherheit in Amazon Redshift](#) im Managementleitfaden zu Amazon-Redshift.

2. Wählen Sie auf der AWS Konsole den AWS-Region Ort aus, an dem Sie den Cluster erstellen möchten.

3. Wählen Sie im Navigationsmenü Clusters (Cluster) und dann Create cluster (Cluster erstellen) aus. Die Seite Create Cluster (Cluster erstellen) wird angezeigt.
4. Geben Sie im Bereich Cluster configuration (Cluster-Konfiguration) Werte für Cluster identifier (Cluster-ID), Node type (Knotentyp) und Nodes (Knoten) an:
 - Cluster Identifier (Cluster-ID): Geben Sie für dieses Tutorial **examplecluster** ein. Diese ID muss eindeutig sein. Die ID muss aus 1–63 Zeichen bestehen und darf nur die Buchstaben a–z (nur Kleinschreibung) und - (Bindestriche) enthalten.
 - Wählen Sie eine der folgenden Methoden zur Bestimmung der Größe Ihres Clusters aus:

 Note

Im folgenden Schritt wird von einem System ausgegangen AWS-Region , das RA3 Knotentypen unterstützt. Eine Liste der AWS-Regionen unterstützten RA3 Knotentypen finden Sie unter [Überblick über RA3 Knotentypen](#) im Amazon Redshift Management Guide. Weitere Informationen über die Knotenspezifikationen für die einzelnen Knotentypen und -größen finden Sie unter [Details zu Knotentypen](#).

- Wenn Sie nicht wissen, wie groß Ihr Cluster sein sollte, wählen Sie Help me choose (Hilfe bei der Auswahl) aus. Dadurch wird ein Größenbestimmungsrechner geöffnet, der Ihnen Fragen zur Größe und zu Abfragemerkmalen der Daten stellt, die Sie in Ihrem Data Warehouse speichern möchten.

Wenn Sie die erforderliche Größe Ihres Clusters kennen (d. h. den Knotentyp und die Anzahl der Knoten), wählen Sie I'll choose (Ich entscheide) aus. Wählen Sie den Node type (Knotentyp) und die Anzahl der Nodes (Knoten) aus, um die Größe Ihres Clusters für den Machbarkeitsnachweis zu bestimmen.

Wählen Sie für dieses Tutorial ra3.4xlarge als Knotentyp und 2 als Anzahl der Knoten aus.

Wenn eine AZ-Konfiguration verfügbar ist, wählen Sie Single-AZ.

- Wählen Sie unter Sample data (Beispieldaten) Load sample data (Beispieldaten laden) aus, um den Beispieldatensatz zu verwenden, den Amazon Redshift bereitstellt. Amazon Redshift lädt den Beispieldatensatz Tickit in die standardmäßige dev-Datenbank und das public-Schema.

5. Geben Sie im Bereich Datenbankkonfiguration einen Wert für Administrator-Benutzername ein. Wählen Sie für Administratorpasswort eine der folgenden Optionen aus:
 - Ein Passwort erstellen – Verwendung eines von Amazon Redshift generierten Passworts.
 - Administratorpasswort manuell hinzufügen – Verwendung Ihres eigenen Passworts.
 - Administratoranmeldedaten verwalten in AWS Secrets Manager — Amazon Redshift verwendet AWS Secrets Manager , um Ihr Administratorkennwort zu generieren und zu verwalten. Für AWS Secrets Manager die Generierung und Verwaltung Ihres Passworts fällt eine Gebühr an. Informationen zu den Preisen für AWS Secrets Manager finden Sie unter [AWS Secrets Manager – Preise](#).

Verwenden Sie für dieses Tutorial folgende Werte:

- Admin user name (Administratorbenutzername): Geben Sie **awsuser** ein.
 - Administratorpasswort: Geben Sie für das Passwort **Changeit1** ein.
6. Erstellen Sie für dieses Tutorial eine IAM-Rolle und legen Sie sie als Standard für Ihren Cluster fest, wie nachfolgend beschrieben. Für einen Cluster kann nur eine Standard-IAM-Rolle festgelegt werden.
 - a. Wählen Sie unter Cluster permissions (Cluster-Berechtigungen) bei Manage IAM roles (IAM-Rollen verwalten) die Option Create IAM role (IAM-Rolle erstellen) aus.
 - b. Geben Sie einen Amazon S3 Bucket an, auf den die IAM-Rolle zugreifen soll, indem Sie eine der folgenden Methoden verwenden:
 - Wählen Sie No additional Amazon S3 bucket (Kein zusätzlicher Amazon S3 Bucket) aus, damit die erstellte IAM-Rolle nur auf die Amazon S3 Buckets zugreifen kann, die als `redshift` benannt sind.
 - Wählen Sie Any Amazon S3 bucket (Beliebiger Amazon S3 Bucket) aus, damit die erstellte IAM-Rolle auf alle Amazon S3 Buckets zugreifen kann.
 - Wählen Sie Specific Amazon S3 buckets (Bestimmte Amazon S3 Buckets) aus, um einen oder mehrere Amazon S3 Buckets anzugeben, auf die die erstellte IAM-Rolle Zugriff hat. Wählen Sie dann einen oder mehrere Amazon S3 Buckets aus der Tabelle aus.
 - c. Wählen Sie Create IAM role as default (IAM-Rolle als Standard erstellen) aus. Amazon Redshift erstellt automatisch die Rolle und legt sie als Standard für Ihren Cluster fest.

Da Sie Ihre IAM-Rolle von der Konsole aus erstellt haben, ist ihr die Richtlinie `AmazonRedshiftAllCommandsFullAccess` angefügt. Dadurch kann Amazon Redshift Daten von Amazon-Ressourcen in Ihrem IAM-Konto kopieren, laden, abfragen und analysieren.

Informationen zum Verwalten der Standard-IAM-Rolle für einen Cluster finden Sie unter [Erstellen einer IAM-Rolle als Standard für Amazon Redshift](#) im Managementleitfaden zu Amazon Redshift.

7. (Optional) Deaktivieren Sie im Bereich **Zusätzliche Konfigurationen** die Option **Standardwerte verwenden**, um die Einstellungen **Netzwerk und Sicherheit**, **Datenbankkonfiguration**, **Wartung**, **Überwachung** und **Backup** anzupassen.

In manchen Fällen können Sie Ihren Cluster mit der Option **Load sample data** (Beispieldaten laden) erstellen. Dabei empfiehlt es sich möglicherweise, erweitertes Amazon-VPC-Routing zu aktivieren. In diesem Fall benötigt der Cluster in Ihrer Virtual Private Cloud (VPC) Zugriff auf den Amazon-S3-Endpunkt, damit Daten geladen werden können.

Um den Cluster öffentlich zugänglich zu machen, haben Sie zwei Möglichkeiten. Sie können eine NAT-Adresse (Network Address Translation) in Ihrer VPC konfigurieren, damit der Cluster auf das Internet zugreifen kann. Oder Sie können einen Amazon-S3-VPC-Endpunkt in Ihrer VPC konfigurieren. Weitere Informationen zu erweitertem Amazon-VPC-Routing finden Sie unter [Aktivieren von erweitertem Amazon-VPC-Routing](#) im Managementleitfaden zu Amazon Redshift.

8. Wählen Sie **Cluster erstellen**. Warten Sie, bis Ihr Cluster mit dem Status `Available` auf der Seite **Cluster erstellt** wurde.

Schritt 2: Regeln für eingehenden Datenverkehr für SQL-Clients konfigurieren

Note

Wir empfehlen Ihnen, diesen Schritt zu überspringen und mit Amazon Redshift Query Editor v2 auf Ihren Cluster zuzugreifen.

Im weiteren Verlauf dieses Tutorials greifen Sie aus einer Virtual Private Cloud (VPC) auf Grundlage des Amazon-VPC-Service heraus auf Ihren Cluster zu. Wenn Sie einen SQL-Client von außerhalb

Ihrer Firewall für den Zugriff auf den Cluster verwenden, müssen Sie jedoch den eingehenden Zugriff gewähren.

So überprüfen Sie Ihre Firewall und gewähren eingehenden Zugriff auf Ihren Cluster:

1. Überprüfen Sie Ihre Firewall-Regeln, wenn auf Ihren Cluster von außerhalb einer Firewall zugegriffen werden muss. Ihr Client kann beispielsweise eine Amazon Elastic Compute Cloud (Amazon EC2)-Instance oder ein externer Computer sein.

Weitere Informationen finden Sie unter [Sicherheitsgruppenregeln](#) im Amazon-EC2-Benutzerhandbuch.

2. Um von einem externen Amazon-EC2-Client aus darauf zuzugreifen, fügen Sie der an Ihren Cluster angefügten Sicherheitsgruppe, die eingehenden Datenverkehr zulässt, eine Eingangsregel hinzu. Sie fügen Amazon-EC2-Sicherheitsgruppenregeln in der Amazon-EC2-Konsole hinzu. Ein Wert CIDR/IP von 192.0.2.0/24 ermöglicht es beispielsweise Clients in diesem IP-Adressbereich, eine Verbindung zu Ihrem Cluster herzustellen. Finden Sie heraus, was für Ihre Umgebung richtig CIDR/IP ist.

Schritt 3: Gewähren von Zugriff auf einen SQL-Client und Ausführen von Abfragen

Für die Abfrage von Ihrem Amazon-Redshift-Cluster gehosteter Datenbanken gibt es verschiedene Möglichkeiten für SQL-Clients. Dazu zählen:

- Verbinden Sie sich mit Ihrem Cluster und führen Sie Abfragen mit Amazon Redshift Query Editor v2 durch.

Wenn Sie Query Editor v2 verwenden, müssen Sie keine SQL-Clientanwendung herunterladen und einrichten. Sie starten Amazon Redshift Query Editor v2 über die Amazon-Redshift-Konsole.

- Herstellen einer Verbindung zu Ihrem Cluster mit RSQL. Weitere Informationen finden Sie unter [Verbindung mit Amazon Redshift RSQL](#) im Managementleitfaden zu Amazon Redshift.
- Stellen Sie über ein SQL-Client-Tool wie SQL Workbench/J eine Verbindung zu Ihrem Cluster her. Weitere Informationen finden Sie unter [Herstellen einer Verbindung mit Ihrem Cluster unter Verwendung von SQL Workbench/J](#) im Managementleitfaden zu Amazon Redshift.

Dieses Tutorial verwendet Amazon Redshift Query Editor v2 als einfachen Weg, Abfragen für Datenbanken auszuführen, die von Ihrem Amazon-Redshift-Cluster gehostet werden. Nachdem Sie Ihren Cluster erstellt haben, können Sie sofort Abfragen ausführen. Details zu Überlegungen bei der Verwendung von Amazon Redshift Query Editor v2 finden Sie unter [Überlegungen zur Arbeit mit Query Editor v2](#) im Managementleitfaden zu Amazon Redshift.

Gewähren des Zugriffs auf Abfrage-Editor v2

Wenn ein Administrator den Query Editor v2 zum ersten Mal für Sie konfiguriert, wählt er den aus AWS-Konto AWS KMS key , der zum Verschlüsseln der Query Editor v2-Ressourcen verwendet wird. Zu den Ressourcen von Amazon Redshift Query Editor v2 gehören gespeicherte Abfragen, Notizbücher und Diagramme. Standardmäßig werden die Ressourcen mit einem AWS -eigenen Schlüssel verschlüsselt. Alternativ dazu kann ein Administrator einen kundenseitig verwalteten Schlüssel verwenden, indem er den Amazon-Ressourcennamen (ARN) als Schlüssel auf der Konfigurationsseite auswählt. Nachdem Sie ein Konto konfiguriert haben, können die AWS KMS Verschlüsselungseinstellungen nicht mehr geändert werden. Weitere Informationen finden Sie unter [Konfigurieren Ihres AWS-Konto](#) im Managementleitfaden zu Amazon Redshift.

Um den Abfrage-Editor v2 aufzurufen, benötigen Sie eine Berechtigung. Ein Administrator kann eine der von AWS verwalteten Richtlinien für Amazon Redshift Query Editor v2 der IAM-Rolle oder dem Benutzer anfügen, um Berechtigungen zu gewähren. Diese AWS verwalteten Richtlinien verfügen über verschiedene Optionen, mit denen gesteuert wird, wie das Taggen von Ressourcen die gemeinsame Nutzung von Abfragen ermöglicht. Sie können die IAM-Konsole (<https://console.aws.amazon.com/iam/>) verwenden, um IAM-Richtlinien anzuhängen. Weitere Informationen zu diesen Richtlinien finden Sie unter [Zugriff auf Query Editor v2](#) im Managementleitfaden zu Amazon Redshift.

Sie können auch Ihre eigene Richtlinie erstellen, basierend auf den zulässigen und verweigerten Berechtigungen in den bereitgestellten verwalteten Richtlinien. Wenn Sie den IAM-Konsolenrichtlinien-Editor verwenden, um Ihre eigene Richtlinie zu erstellen, wählen Sie SQL Workbench als Service aus, für den Sie die Richtlinie im visuellen Editor erstellen. Der Abfrage-Editor v2 verwendet den Dienstnamen AWS SQL Workbench im Visual Editor und im IAM Policy Simulator.

Weitere Informationen finden Sie unter [Arbeiten mit dem Abfrage-Editor v2](#) im Amazon-Redshift-Verwaltungshandbuch.

Schritt 4: Daten aus Amazon S3 in Amazon Redshift laden

Nachdem Sie Ihren Cluster erstellt haben, können Sie Daten aus Amazon S3 in Ihre Datenbanktabellen laden. Es gibt mehrere Möglichkeiten, Daten aus Amazon S3 zu laden.

- Sie können einen SQL-Client verwenden, um den SQL-Befehl `CREATE TABLE` auszuführen, um eine Tabelle in Ihrer Datenbank zu erstellen, und dann den Befehl `SQL COPY`, um Daten aus Amazon S3 zu laden. Amazon Redshift Query Editor v2 ist ein SQL-Client.
- Sie können den Ladeassistenten von Amazon Redshift Query Editor v2 verwenden.

Dieses Tutorial zeigt, wie Sie Amazon Redshift Query Editor v2 verwenden, um SQL-Befehle auszuführen, um Tabellen zu erstellen (`CREATE`) und Daten zu kopieren (`COPY`). Starten Sie Query Editor v2 über den Navigationsbereich der Amazon-Redshift-Konsole. Stellen Sie in Query Editor v2 eine Verbindung zum `examplecluster`-Cluster und zur Datenbank mit dem Namen `dev` mit Ihrem Admin-Benutzer `awsuser` her. Wählen Sie für dieses Tutorial Temporäre Anmeldeinformationen mit einem Datenbankbenutzernamen, wenn Sie die Verbindung herstellen. Details zur Verwendung von Amazon Redshift Query Editor v2 finden Sie unter [Herstellen einer Verbindung mit einer Amazon-Redshift-Datenbank](#) im Managementleitfaden zu Amazon Redshift.

Laden von Daten aus Amazon S3 mit SQL-Befehlen

Vergewissern Sie sich im Fenster von Query Editor v2, dass Sie mit dem `examplecluster`-Cluster und der `dev`-Datenbank verbunden sind. Erstellen Sie dann Tabellen in der Datenbank und laden Daten in die Tabellen. In diesem Tutorial sind die Daten, die Sie laden, in einem Amazon S3 S3-Bucket verfügbar, auf den von vielen aus zugegriffen werden kann AWS-Regionen.

Mit dem folgenden Verfahren werden Tabellen erstellt und Daten aus einem öffentlichen Amazon-S3-Bucket geladen.

Verwenden Sie Amazon Redshift Query Editor v2 verwenden, um die folgende Anweisung zum Erstellen einer Tabelle im `public`-Schema der `dev`-Datenbank zu kopieren und auszuführen. Weitere Informationen zur Syntax finden Sie unter [CREATE TABLE](#) im Datenbankentwicklerhandbuch zu Amazon Redshift.

So erstellen und laden Sie Daten mit einem SQL-Client wie Query Editor v2:

1. Verwenden Sie den folgenden SQL-Befehl, um die `sales` Tabelle zu erstellen (`CREATE`).

```
drop table if exists sales;
create table sales(
salesid integer not null,
listid integer not null distkey,
sellerid integer not null,
buyerid integer not null,
eventid integer not null,
dateid smallint not null sortkey,
qtysold smallint not null,
pricepaid decimal(8,2),
commission decimal(8,2),
saletime timestamp);
```

2. Verwenden Sie den folgenden SQL-Befehl, um die date Tabelle zu erstellen (CREATE).

```
drop table if exists date;
create table date(
dateid smallint not null distkey sortkey,
caldate date not null,
day character(3) not null,
week smallint not null,
month character(5) not null,
qtr character(5) not null,
year smallint not null,
holiday boolean default('N'));
```

3. Laden Sie die sales-Tabelle mit dem COPY-Befehl aus Amazon S3.

Note

Es wird empfohlen, zum Laden großer Datensätze in Amazon Redshift aus Amazon S3 den COPY-Befehl zu verwenden. Weitere Informationen zur COPY-Syntax finden Sie unter [COPY](#) im Datenbankentwicklerhandbuch zu Amazon Redshift.

Stellen Sie Authentifizierung für Ihren Cluster um Zugriff auf Amazon S3 in Ihrem Namen bereit, um die Beispieldaten zu laden. Sie stellen die Authentifizierung bereit, indem Sie auf die IAM-Rolle verweisen, die Sie erstellt und als default für Ihren Cluster mit der Auswahl von IAM-Rolle als Standard erstellen bei der Erstellung des Clusters festgelegt haben.

Laden Sie die sales-Tabelle mit dem folgenden SQL-Befehl. Optional können Sie die [Quelldaten für die sales-Tabelle](#) aus Amazon S3 herunterladen und anzeigen. .

```
COPY sales
FROM 's3://redshift-downloads/ticket/sales_tab.txt'
DELIMITER '\t'
TIMEFORMAT 'MM/DD/YYYY HH:MI:SS'
REGION 'us-east-1'
IAM_ROLE default;
```

4. Laden Sie die date-Tabelle mit dem folgenden SQL-Befehl. Optional können Sie die [Quelldaten für die date-Tabelle](#) aus Amazon S3 herunterladen und anzeigen. .

```
COPY date
FROM 's3://redshift-downloads/ticket/date2008_pipe.txt'
DELIMITER '|'
REGION 'us-east-1'
IAM_ROLE default;
```

Laden von Daten aus Amazon S3 mit Query Editor v2

In diesem Abschnitt wird beschrieben, wie Sie Ihre eigenen Daten in einen Amazon-Redshift-Cluster laden. Query Editor v2 vereinfacht das Laden von Daten mit dem Load data-Assistenten. Der COPY-Befehl, der im Load data-Assistenten von Query Editor v2 generiert und verwendet wird, unterstützt alle Parameter, die für die COPY-Befehlssyntax zum Laden von Daten aus Amazon S3 verfügbar sind. Weitere Informationen zum COPY-Befehl und zu seinen Optionen zum Kopieren und Laden aus Amazon S3 finden Sie unter [COPY aus dem Amazon Simple Storage Service](#) im Datenbankentwicklerhandbuch zu Amazon Redshift.

Um Ihre eigenen Daten aus Amazon S3 in Amazon Redshift zu laden, erfordert Amazon Redshift eine IAM-Rolle, die über die benötigten Berechtigungen zum Laden von Daten aus dem angegebenen Amazon S3 Bucket verfügt.

Zum Laden Ihrer Daten aus Amazon S3 in Amazon Redshift mit Query Editor v2 können Sie den Assistenten zum Laden von Daten von Query Editor v2 verwenden. Weitere Informationen zur Verwendung dieses Assistenten finden Sie unter [Laden von Daten aus Amazon S3](#) im Managementleitfaden zu Amazon Redshift.

Erstellen von TICKIT-Daten in Ihrem Cluster

TICKIT ist eine Beispieldatenbank, die Sie optional in Ihren Amazon-Redshift-Cluster laden können, um zu lernen, wie Sie Daten in Amazon Redshift abfragen. Sie können den vollständigen Satz von TICKIT-Tabellen erstellen und Daten auf folgende Weise in Ihren Cluster laden:

- Wenn Sie einen Cluster in der Amazon-Redshift-Konsole erstellen, haben Sie die Möglichkeit, gleichzeitig Beispieldaten für TICKIT zu laden. Wählen Sie in der Amazon-Redshift-Konsole Cluster, Cluster erstellen aus. Wählen Sie im Abschnitt Beispieldaten die Option Beispieldaten laden aus. Amazon Redshift lädt den Beispieldatensatz während der Clustererstellung automatisch in Ihre dev Amazon-Redshift-Clusterdatenbank.
- So stellen Sie eine Verbindung zu einem vorhandenen Cluster her:
 - Wählen Sie in der Amazon-Redshift-Konsole im Navigationsbereich Cluster aus.
 - Wählen Sie Ihren Cluster im Fenster Cluster aus.
 - Wählen Sie Daten abfragen, Abfrage in Query Editor v2 aus.
 - Erweitern Sie examplecluster in der Ressourcenliste. Wenn Sie zum ersten Mal eine Verbindung zu Ihrem Cluster herstellen, wird Verbindung zu examplecluster angezeigt. Wählen Sie Datenbank-Benutzername und Passwort. Belassen Sie die Datenbank als **dev**. Geben Sie **awsuser** als Benutzernamen und **Changeit1** als Passwort an.
 - Wählen Sie Create Connection (Verbindung erstellen) aus.
- Mit Amazon Redshift Query Editor v2 können Sie TICKIT-Daten in eine Beispieldatenbank mit dem Namen sample_data_dev laden. Wählen Sie die Datenbank sample_data_dev in der Ressourcenliste aus. Wählen Sie neben dem Tickit-Knoten das Symbol Beispielnotizbücher öffnen aus. Bestätigen Sie, dass Sie die Beispieldatenbank erstellen möchten.
- Der Amazon Redshift Query Editor v2 erstellt die Beispieldatenbank zusammen mit einem Beispielnotizbuch mit dem Namen tickit-sample-notebook. Sie können Alle ausführen wählen, um dieses Notizbuch auszuführen und Daten in der Beispieldatenbank abzufragen.

Weitere Informationen zu den TICKIT-Daten finden Sie unter [Beispieldatenbank](#) im Datenbankentwicklerhandbuch zu Amazon Redshift.

Schritt 5: Beispielabfragen mit dem Abfrage-Editor testen

Informationen zur Einrichtung und Verwendung von Amazon Redshift Query Editor v2 zur Abfrage von Datenbanken finden Sie unter [Arbeiten mit Query Editor v2](#) im Managementleitfaden zu Amazon Redshift.

Testen Sie jetzt einige Beispielabfragen wie folgt. Um neue Abfragen in Query Editor v2 zu erstellen, wählen Sie das +-Symbol oben rechts im Abfragebereich und dann SQL. Eine neue Abfrageseite wird angezeigt, auf der Sie die folgenden SQL-Abfragen kopieren und einfügen können.

Note

Stellen Sie sicher, dass Sie zuerst die erste Abfrage im Notizbuch ausführen, wodurch der `search_path`-Serverkonfigurationswert mit dem folgenden SQL-Befehl auf das `ticket`-Schema festgelegt wird:

```
set search_path to ticket;
```

Weitere Informationen zur Verwendung des `SELECT`-Befehls finden Sie unter [SELECT](#) im Datenbankentwicklerhandbuch zu Amazon Redshift.

```
-- Get definition for the sales table.
SELECT *
FROM pg_table_def
WHERE tablename = 'sales';
```

```
-- Find total sales on a given calendar date.
SELECT sum(qtysold)
FROM sales, date
WHERE sales.dateid = date.dateid
AND caldate = '2008-01-05';
```

```
-- Find top 10 buyers by quantity.
SELECT firstname, lastname, total_quantity
FROM (SELECT buyerid, sum(qtysold) total_quantity
      FROM sales
      GROUP BY buyerid
```

```
ORDER BY total_quantity desc limit 10) Q, users
WHERE Q.buyerid = userid
ORDER BY Q.total_quantity desc;
```

```
-- Find events in the 99.9 percentile in terms of all time gross sales.
SELECT eventname, total_price
FROM (SELECT eventid, total_price, ntile(1000) over(order by total_price desc) as
percentile
FROM (SELECT eventid, sum(pricepaid) total_price
FROM sales
GROUP BY eventid)) Q, event E
WHERE Q.eventid = E.eventid
AND percentile = 1
ORDER BY total_price desc;
```

Schritt 6: Umgebung zurücksetzen

In den vorherigen Schritten haben Sie erfolgreich einen Amazon-Redshift-Cluster erstellt, Daten in Tabellen geladen und Daten mit einem SQL-Client wie Amazon Redshift Query Editor v2 abgefragt.

Wenn Sie dieses Tutorial abgeschlossen haben, empfehlen wir, dass Sie Ihre Umgebung auf den vorherigen Zustand zurücksetzen, indem Sie Ihren Beispielcluster löschen. Es fallen so lange Amazon-Redshift-Nutzungsgebühren, bis Sie den Cluster löschen.

Sie können den Beispielcluster jedoch aktiviert lassen, wenn Sie Aufgaben in anderen Amazon-Redshift-Leitfäden oder in [Ausführen von Befehlen, um eine Datenbank in Ihrem Data Warehouse zu definieren und zu verwenden](#) beschriebene Aufgaben ausprobieren möchten.

So löschen Sie einen Cluster

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die Amazon Redshift Redshift-Konsole unter <https://console.aws.amazon.com/redshiftv2/>.
2. Wählen Sie im Navigationsmenü Clusters (Cluster) aus, um Ihre Liste der Cluster anzuzeigen.
3. Wählen Sie den Cluster `examplecluster` aus. Klicken Sie bei Actions auf Delete. Die Seite Beispielcluster löschen? wird angezeigt.
4. Bestätigen Sie, dass der Cluster gelöscht werden soll, deaktivieren Sie die Einstellung Endgültigen Snapshot erstellen und geben Sie dann **delete** ein, um das Löschen zu bestätigen. Wählen Sie Delete cluster (Cluster löschen) aus.

Auf der Seite mit der Clusterliste wird der Clusterstatus aktualisiert, wenn der Cluster gelöscht wird.

Nach Abschluss dieses Tutorials finden Sie weitere Informationen über Amazon Redshift sowie die nächsten Schritte unter [Weitere Ressourcen, um mehr über Amazon Redshift zu erfahren..](#)

Ausführen von Befehlen, um eine Datenbank in Ihrem Data Warehouse zu definieren und zu verwenden

Sowohl Data Warehouses von Redshift Serverless als auch von Amazon Redshift bereitgestellte Data Warehouses enthalten Datenbanken. Nachdem Sie Ihr Data Warehouse gestartet haben, können Sie die meisten Datenbankaktionen mithilfe von SQL-Befehlen verwalten. Mit wenigen Ausnahmen sind Funktionalität und Syntax von SQL für alle Amazon-Redshift-Datenbanken identisch. Einzelheiten zu den in Amazon Redshift verfügbaren SQL-Befehlen finden Sie unter [SQL-Befehle](#) im Datenbankentwicklerhandbuch zu Amazon Redshift.

Wenn Sie Ihr Data Warehouse erstellen, erstellt Amazon Redshift in den meisten Szenarien auch die dev-Standarddatenbank. Nachdem Sie eine Verbindung mit der anfänglichen dev-Datenbank hergestellt haben, können Sie eine weitere Datenbank erstellen.

In den folgenden Abschnitten werden allgemeine Datenbankaufgaben bei der Arbeit mit Amazon-Redshift-Datenbanken beschrieben. Die Aufgaben beginnen mit dem Erstellen einer Datenbank. Wenn Sie mit der letzten Aufgabe fortfahren, können Sie alle von Ihnen erstellten Ressourcen löschen, indem Sie die Datenbank löschen.

Die Beispiele in diesem Abschnitt setzen Folgendes voraus:

- Sie haben ein Data Warehouse von Amazon Redshift erstellt.
- Sie haben von Ihrem SQL-Client-Tool wie dem Amazon Redshift Query Editor v2 eine Verbindung mit dem Data Warehouse hergestellt. Weitere Informationen zur Arbeit mit Query Editor v2 finden Sie unter [Abfragen einer Datenbank mit Amazon Redshift Query Editor v2](#) im Managementleitfaden zu Amazon Redshift.

Themen

- [Herstellen einer Verbindung mit einem Data Warehouse von Amazon Redshift](#)
- [Erstellen einer -Datenbank](#)
- [Erstellen eines Benutzers](#)
- [Erstellen eines Schemas](#)
- [Erstellen einer Tabelle](#)
- [Laden von Daten](#)

- [Abfragen der Systemtabellen und Ansichten](#)
- [Abbrechen einer Abfrage](#)

Herstellen einer Verbindung mit einem Data Warehouse von Amazon Redshift

Um eine Verbindung mit Amazon-Redshift-Clustern herzustellen, erweitern Sie auf der Seite Cluster Verbindung zu Amazon-Redshift-Clustern und gehen dann folgendermaßen vor:

- Verwenden Sie Daten abfragen, um in Query Editor v2 Abfragen für Datenbanken auszuführen, die von Ihrem Amazon-Redshift-Cluster gehostet werden. Nachdem Sie Ihren Cluster erstellt haben, können Sie mit dem Abfrageeditor v2 sofort Abfragen ausführen.

Weitere Informationen finden Sie unter [Abfragen einer Datenbank mit Amazon Redshift Query Editor v2](#) im Managementleitfaden zu Amazon Redshift.

- Wählen Sie unter Arbeiten mit Ihren Client-Tools Ihren Cluster, und stellen Sie mithilfe von JDBC- oder ODBC-Treibern eine Verbindung von Ihren Client-Tools zu Amazon Redshift her, indem Sie die URL des JDBC- oder ODBC-Treibers kopieren. Verwenden Sie diese URL von Ihrem Client-Computer oder Ihrer Instance aus. Schreiben Sie Ihre Anwendungen so, dass sie JDBC- oder ODBC-API-Operationen für den Zugriff auf Daten nutzen, oder verwenden Sie SQL-Client-Tools, die JDBC oder ODBC unterstützen.

Weitere Informationen darüber, wie Sie Ihre Cluster-Verbindungszeichenfolge finden, erhalten Sie unter [Suche der Zeichenfolge für die Verbindung mit dem Cluster](#).

- Wenn Ihr SQL-Client-Tool einen Treiber benötigt, können Sie unter Auswahl Ihres JDBC- oder ODBC-Treibers einen betriebssystemspezifischen Treiber herunterladen, um sich von Ihren Client-Tools aus mit Amazon Redshift zu verbinden.

Weitere Informationen zur Installation des entsprechenden Treibers für Ihren SQL-Client finden Sie unter [Konfiguration einer JDBC-Treiberbindung der Version 2.2](#).

Weitere Informationen zum Konfigurieren einer ODBC-Verbindung finden Sie unter [Konfigurierung einer ODBC-Verbindung](#).

Gehen Sie auf der Seite Serverless-Dashboard der Amazon-Redshift-Konsole wie folgt vor, um eine Verbindung zum Data Warehouse von Redshift Serverless herzustellen:

- Verwenden Sie Amazon Redshift Query Editor V2, um Abfragen für Datenbanken auszuführen, die von Ihrem Data Warehouse von Amazon Redshift Serverless gehostet werden. Nachdem Sie Ihr Data Warehouse erstellt haben, können Sie sofort mit Query Editor v2 Abfragen ausführen.

Weitere Informationen finden Sie unter [Abfragen einer Datenbank mit dem Abfrage-Editor v2 von Amazon Redshift](#).

- Stellen Sie mithilfe von JDBC- oder ODBC-Treibern eine Verbindung von Ihren Client-Tools zu Amazon Redshift her, indem Sie die URL des JDBC- oder ODBC-Treibers kopieren.

Um mit Daten in Ihrem Data Warehouse zu arbeiten, benötigen Sie JDBC- oder ODBC-Treiber, um Verbindungen von Ihrem Client-Computer oder Ihrer Instance aus herzustellen. Schreiben Sie Ihre Anwendungen so, dass sie JDBC- oder ODBC-API-Operationen für den Zugriff auf Daten nutzen, oder verwenden Sie SQL-Client-Tools, die JDBC oder ODBC unterstützen.

Weitere Informationen zum Auffinden Ihrer Verbindungszeichenfolge finden Sie unter [Verbindung zu Redshift Serverless](#) im Managementleitfaden zu Amazon Redshift.

Erstellen einer -Datenbank

Nachdem Sie geprüft haben, ob Ihr Data Warehouse ausgeführt wird, können Sie eine Datenbank erstellen. In dieser Datenbank erstellen Sie Tabellen, laden Daten und führen Abfragen aus. Ein Data Warehouse kann mehrere Datenbanken hosten. Sie können beispielsweise eine Datenbank für Vertriebsdaten mit dem Namen SALESDB und eine Datenbank für Auftragsdaten mit dem Namen ORDERSDB in demselben Data Warehouse haben.

Um eine Datenbank mit dem Namen **SALESDB** zu erstellen, führen Sie den folgenden Befehl in Ihrem SQL-Client-Tool aus:

```
CREATE DATABASE salesdb;
```

Note

Stellen Sie nach der Ausführung des Befehls sicher, dass Sie die Liste der Objekte in Ihrem Data Warehouse in Ihrem SQL-Client-Tool aktualisieren, um die neuen salesdb zu sehen.

Für diese Übung übernehmen wir die Standardeinstellungen. Weitere Informationen über Befehlsoptionen finden Sie unter [CREATE DATABASE](#) im Datenbankentwicklerhandbuch zu Amazon

Redshift. Informationen zum Löschen einer Datenbank und ihres Inhalts finden Sie unter [DROP DATABASE](#) im Datenbankentwicklerhandbuch zu Amazon Redshift.

Nachdem Sie die SALESDB-Datenbank erstellt haben, können Sie von Ihrem SQL-Client aus eine Verbindung zu der neuen Datenbank herstellen. Verwenden Sie dieselben Verbindungsparameter wie für Ihre aktuelle Verbindung, ändern Sie aber den Namen der Datenbank in SALESDB.

Erstellen eines Benutzers

Standardmäßig kann nur der Administratorbenutzer, den Sie beim Start des Data Warehouse eingerichtet haben, auf die Standarddatenbank in diesem Data Warehouse zugreifen. Um anderen Benutzern den Zugriff zu gewähren, erstellen Sie ein oder mehrere Benutzerkonten. Datenbankbenutzerkonten gelten global für alle Datenbanken in einem Data Warehouse und gehören nicht zu einzelnen Datenbanken.

Verwenden Sie den Befehl `CREATE USER`, um einen neuen Benutzer zu erstellen. Wenn Sie einen neuen Benutzer erstellen, geben Sie dessen Namen und ein Passwort an. Es wird empfohlen, dass Sie ein Passwort für den Benutzer angeben. Dies muss aus 8–64 Zeichen bestehen und mindestens einen Großbuchstaben, einen Kleinbuchstaben und eine Ziffer enthalten.

Führen Sie beispielsweise zur Erstellung eines Benutzers mit dem Namen **GUEST** und dem Passwort **ABCd4321** den folgenden Befehl aus:

```
CREATE USER GUEST PASSWORD 'ABCd4321';
```

Um eine Verbindung mit der SALESDB-Datenbank als der GUEST-Benutzer herzustellen, verwenden Sie das Passwort, das Sie bei der Benutzererstellung gewählt haben, zum Beispiel ABCd4321.

Informationen über weitere Befehlsoptionen finden Sie unter [CREATE USER](#) im Datenbankentwicklerhandbuch zu Amazon Redshift.

Erstellen eines Schemas

Nach dem Erstellen einer neuen Datenbank können Sie ein neues Schema in der aktuellen Datenbank erstellen. Ein Schema ist ein Namespace, der benannte Datenbankobjekte wie Tabellen, Ansichten und benutzerdefinierte Funktionen () enthält. UDFs Eine Datenbank kann ein oder mehrere Schemata enthalten, und jedes Schema gehört nur zu einer Datenbank. Zwei Schemata können verschiedene Objekte mit demselben Namen haben.

Sie können mehrere Schemata in derselben Datenbank erstellen, um Daten so zu organisieren, wie Sie möchten, oder um Ihre Daten funktional zu gruppieren. Sie können beispielsweise ein Schema erstellen, um alle Staging-Daten zu speichern, und ein anderes Schema zum Speichern aller Berichtstabellen. Sie können auch verschiedene Schemata erstellen, um Daten zu speichern, die für verschiedene Unternehmensgruppen relevant sind, die sich in derselben Datenbank befinden. Jedes Schema kann unterschiedliche Datenbankobjekte wie Tabellen, Ansichten und benutzerdefinierte Funktionen () speichern. UDFs Darüber hinaus können Sie Schemata mit der AUTHORIZATION-Klausel erstellen. Diese Klausel gewährt einem bestimmten angegebenen Benutzer Besitz oder legt ein Kontingent für den maximalen Speicherplatz fest, den das angegebene Schema verwenden kann.

Amazon Redshift erstellt automatisch ein Schema namens `public` für jede neue Datenbank. Wenn Sie den Schemanamen beim Erstellen von Datenbankobjekten nicht angeben, gehen die Objekte in das `public`-Schema über.

Um auf ein Objekt in einem Schema zuzugreifen, qualifizieren Sie das Objekt mithilfe der `schema_name.table_name`-Notation. Der qualifizierte Name des Schemas besteht aus dem Schemanamen und dem Tabellennamen, die durch einen Punkt getrennt sind. Zum Beispiel kann ein `sales`-Schema eine `price`-Tabelle und ein `inventory`-Schema eine `price`-Tabelle haben. Wenn Sie die `price`-Tabelle referenzieren, müssen Sie sie als `sales.price` oder `inventory.price` qualifizieren.

Im folgenden Beispiel wird ein Schema mit dem Namen **SALES** für den Benutzer `GUEST` erstellt.

```
CREATE SCHEMA SALES AUTHORIZATION GUEST;
```

Weitere Informationen über Befehloptionen finden Sie unter [CREATE SCHEMA](#) im Datenbankentwicklerhandbuch zu Amazon Redshift.

Führen Sie den folgenden Befehl aus, um die Liste der Schemata in Ihrer Datenbank anzuzeigen.

```
select * from pg_namespace;
```

Die Ausgabe sollte in etwa folgendermaßen aussehen:

nspname	nspowner	nspacl
sales	100	
pg_toast	1	
pg_internal	1	

```
catalog_history | 1 |
pg_temp_1      | 1 |
pg_catalog     | 1 | {rdsdb=UC/rdsdb,=U/rdsdb}
public         | 1 | {rdsdb=UC/rdsdb,=U/rdsdb}
information_schema | 1 | {rdsdb=UC/rdsdb,=U/rdsdb}
```

Weitere Informationen zum Abfragen von Katalogtabellen finden Sie unter [Abfragen der Katalogtabellen](#) im Datenbankentwicklerhandbuch zu Amazon Redshift.

Verwenden Sie die GRANT-Anweisung, um Benutzern Berechtigungen für die Schemata zu erteilen.

Im folgenden Beispiel werden dem Benutzer GUEST Berechtigungen gewährt, um Daten aus allen Tabellen oder Ansichten im Schema SALES mit einer SELECT-Anweisung auszuwählen.

```
GRANT SELECT ON ALL TABLES IN SCHEMA SALES TO GUEST;
```

Im folgenden Beispiel werden dem Benutzer GUEST alle verfügbaren Rechte auf einmal gewährt.

```
GRANT ALL ON SCHEMA SALES TO GUEST;
```

Erstellen einer Tabelle

Nach dem Erstellen Ihrer neuen Datenbank erstellen Sie Tabellen für Ihre Daten. Geben Sie alle Spalteninformationen an, wenn Sie die Tabelle erstellen.

Zum Beispiel können Sie mit dem folgenden Befehl eine Tabelle namens **DEMO** erstellen.

```
CREATE TABLE Demo (
  PersonID int,
  City varchar (255)
);
```

Standardmäßig werden neue Datenbankobjekte wie Tabellen im Standardschema mit dem Namen `public` während der Erstellung des Data Warehouses erstellt. Sie können ein anderes Schema verwenden, um Datenbankobjekte zu erstellen. Weitere Informationen über Schemata finden Sie unter [Verwalten der Datenbanksicherheit](#) im Datenbankentwicklerhandbuch zu Amazon Redshift.

Darüber hinaus können Sie mit der `schema_name.object_name`-Notation auch eine Tabelle im SALES-Schema erstellen.

```
CREATE TABLE SALES.DEMO (  
  PersonID int,  
  City varchar (255)  
);
```

Um Schemata und ihre Tabellen anzuzeigen und zu überprüfen, können Sie Amazon Redshift Query Editor v2 verwenden. Oder Sie können die Liste der Tabellen in Schemata mithilfe von Systemansichten ansehen. Weitere Informationen finden Sie unter [Abfragen der Systemtabellen und Ansichten](#).

Die Spalten `encoding`, `distkey` und `sortkey` werden von Amazon Redshift für die parallele Verarbeitung verwendet. Für weitere Informationen zum Entwurf von Tabellen mit diesen Elementen siehe [Bewährte Methoden für die Gestaltung von Tabellen mit Amazon Redshift](#).

Einfügen von Datenzeilen in eine Tabelle

Nach der Erstellung der Tabelle fügen Sie Datenzeilen darin ein.

Note

Zeilen werden mit dem Befehl [INSERT](#) in Tabellen eingefügt. Verwenden Sie für Standard-Masseneinfügungen den Befehl [COPY](#). Weitere Informationen finden Sie unter [Verwenden eines COPY-Befehls zum Laden von Daten](#).

Um zum Beispiel Werte in die Tabelle DEMO einzufügen, führen Sie folgenden Befehl aus.

```
INSERT INTO DEMO VALUES (781, 'San Jose'), (990, 'Palo Alto');
```

Führen Sie den folgenden Befehl aus, um Daten in eine Tabelle einzufügen, die sich in einem bestimmten Schema befindet.

```
INSERT INTO SALES.DEMO VALUES (781, 'San Jose'), (990, 'Palo Alto');
```

Auswahl von Daten aus einer Tabelle

Nachdem Sie eine Tabelle erstellt und mit Daten gefüllt haben, verwenden Sie eine `SELECT`-Anweisung, um die in der Tabelle enthaltenen Daten anzuzeigen. Die Anweisung `SELECT *` gibt

alle Spaltennamen und Zeilenwerte für alle Daten in einer Tabelle zurück. Die Verwendung von SELECT ist eine gute Möglichkeit, um zu prüfen, ob kürzlich hinzugefügte Daten korrekt in die Tabelle eingefügt wurden.

Um die Daten anzuzeigen, die Sie in die Tabelle **DEMO** eingegeben haben, führen Sie den folgenden Befehl aus:

```
SELECT * from DEMO;
```

Das Ergebnis sollte wie das folgende aussehen.

```
personid | city
-----+-----
       781 | San Jose
       990 | Palo Alto
(2 rows)
```

Weitere Informationen zur Verwendung der SELECT-Anweisung zur Abfrage von Tabellen finden Sie unter [SELECT](#).

Laden von Daten

Viele der Beispiele in dieser Anleitung verwenden den Beispieldatensatz TICKIT. Sie können die Datei [ticketdb.zip](#) herunterladen. Diese enthält einzelne Beispieldatendateien. Sie können dann die Beispieldaten in Ihren eigenen Amazon-S3-Bucket laden.

Um die Beispieldaten für Ihre Datenbank zu laden, erstellen Sie zuerst die Tabellen. Verwenden Sie dann den COPY-Befehl, um die Tabellen mit Beispieldaten zu laden, die in einem Amazon S3 Bucket gespeichert sind. Weitere Informationen zu den Schritten für die Erstellung von Tabellen und das Laden von Beispieldaten finden Sie unter [Schritt 4: Daten aus Amazon S3 in Amazon Redshift laden](#).

Abfragen der Systemtabellen und Ansichten

Zusätzlich zu den von Ihnen erstellten Tabellen enthält Ihr Data Warehouse eine Reihe von Systemtabellen und Ansichten. Diese Tabellen und Ansichten enthalten Informationen zu Ihrer Installation sowie zu den verschiedenen Abfragen und Prozessen, die auf dem System ausgeführt werden. Sie können diese Systemtabellen und Ansichten abfragen, um Informationen zu Ihrer

Datenbank zu erhalten. Weitere Informationen finden Sie in der [Referenz zu Systemtabellen und Ansichten](#) im Datenbankentwicklerhandbuch zu Amazon Redshift. Die Beschreibung für jede Tabelle oder Ansicht gibt an, ob die Tabelle für alle Benutzer oder nur für Superuser sichtbar ist. Um nur für Superuser sichtbare Tabellen anzuzeigen, melden Sie sich als Superuser an.

Anzeigen einer Liste von Tabellennamen

Um eine Liste aller Tabellen in einem Schema anzuzeigen, können Sie die Systemkatalogtabelle PG_TABLE_DEF abfragen. Sie können zunächst die Einstellung für prüfen search_path.

```
SHOW search_path;
```

Das Ergebnis sollte in etwa wie folgt aussehen.

```
search_path
-----
$user, public
```

Im folgenden Beispiel wird das SALES-Schema dem Suchpfad hinzugefügt und es werden alle Tabellen im SALES-Schema angezeigt.

```
set search_path to '$user', 'public', 'sales';
```

```
SHOW search_path;
```

```
search_path
-----
"$user", public, sales
```

```
select * from pg_table_def where schemaname = 'sales';
```

schemaname	tablename	column	type	encoding	distkey	sortkey	notnull
sales	demo	personid	integer	az64	f		
		0				f	
sales	demo	city	character varying(255)	lzo	f		
		0				f	

Im folgenden Beispiel wird eine Liste aller Tabellen mit dem Namen DEMO in allen Schemata der aktuellen Datenbank angezeigt.

```
set search_path to '$user', 'public', 'sales';
select * from pg_table_def where tablename = 'demo';
```

schemaname	tablename	column	type	encoding	distkey	sortkey	notnull
public	demo	personid	integer	az64	f		f
public	demo	city	character varying(255)	lzo	f		f
sales	demo	personid	integer	az64	f		f
sales	demo	city	character varying(255)	lzo	f		f

Weitere Informationen finden Sie in der Tabelle [PG_TABLE_DEF](#).

Sie können auch Amazon Redshift Query Editor v2 verwenden, um alle Tabellen in einem angegebenen Schema anzuzeigen, indem Sie zuerst eine Datenbank auswählen, mit der Sie eine Verbindung herstellen möchten.

Anzeigen von Benutzern

Sie können den Katalog PG_USER abfragen, um eine Liste aller Benutzer zusammen mit Benutzer-ID (USESYSID) und Benutzerberechtigungen anzuzeigen.

```
SELECT * FROM pg_user;
```

username	usesysid	usecreatedb	usesuper	usecatupd	passwd	valuntil	useconfig
rdsdb	1	true	true	true	*****	infinity	
awsuser	100	true	true	false	*****		
guest	104	true	false	false	*****		

Der Benutzername rdsdb wird intern von Amazon Redshift für Routine-Verwaltungs- und Wartungsaufgaben verwendet. Sie können Ihre Abfrage so filtern, dass nur benutzerdefinierte

Benutzernamen angezeigt werden, indem Sie Ihrer SELECT-Anweisung `where usesysid > 1` hinzufügen.

```
SELECT * FROM pg_user WHERE usesysid > 1;
```

username	usesysid	usecreatedb	usesuper	usecatupd	passwd	valuntil	useconfig
awsuser	100	true	true	false	*****		
guest	104	true	false	false	*****		

Anzeigen aktueller Abfragen

Im vorherigen Beispiel ist die Benutzer-ID (`user_id`) für `adminuser` 100. Um die vier letzten von `adminuser` ausgeführten Abfragen aufzulisten, können Sie die Ansicht `SYS_QUERY_HISTORY` abfragen.

Anhand dieser Ansicht können Sie die Abfrage-ID (`query_id`) oder die Prozess-ID (`session_id`) für eine kürzlich ausgeführte Abfrage finden. Sie können diese Ansicht auch verwenden, um zu überprüfen, wie lange eine Abfrage in Anspruch nahm. `SYS_QUERY_HISTORY` enthält die ersten 4.000 Zeichen der Abfragezeichenfolge (`query_text`), damit Sie eine bestimmte Abfrage besser finden können. Verwenden Sie die `LIMIT`-Klausel mit Ihrer `SELECT`-Anweisung, um die Ergebnisse zu begrenzen.

```
SELECT query_id, session_id, elapsed_time, query_text
FROM sys_query_history
WHERE user_id = 100
ORDER BY start_time desc
LIMIT 4;
```

Das Ergebnis sieht in etwa wie folgt aus.

query_id	session_id	elapsed_time	query_text
892	21046	55868	SELECT query, pid, elapsed, substring from ...
620	17635	1296265	SELECT query, pid, elapsed, substring from ...

```
610      |    17607      |    82555      | SELECT * from DEMO;  
596      |    16762      |   226372      | INSERT INTO DEMO VALUES (100);
```

Feststellen der Sitzungs-ID einer laufenden Abfrage

Sie müssen möglicherweise die zu einer Abfrage gehörende Sitzungs-ID (Prozess-ID) angeben, um Systemtabelleninformationen zu einer Abfrage abzurufen. Möglicherweise müssen Sie aber auch die Sitzungs-ID für eine Abfrage ermitteln, die noch ausgeführt wird. So brauchen Sie etwa die Sitzungs-ID, wenn Sie eine Abfrage abbrechen müssen, die auf einem bereitgestellten Cluster zu viel Zeit in Anspruch nimmt. Sie können die STV_RECENTS-Systemtabelle abfragen, um eine Liste der Sitzungen IDs für laufende Abfragen zusammen mit der entsprechenden Abfragezeichenfolge zu erhalten. Wenn Ihre Abfrage mehrere Sitzungs-IDs ausgibt, können Sie anhand des Abfragetextes bestimmen, welche Sitzungs-ID Sie benötigen.

Um die Sitzungs-ID einer laufenden Abfrage zu bestimmen, führen Sie die folgende SELECT-Anweisung aus.

```
SELECT session_id, user_id, start_time, query_text  
FROM sys_query_history  
WHERE status='running';
```

Abbrechen einer Abfrage

Wenn Sie eine Abfrage ausführen, die zu lange dauert oder zu viele Ressourcen verbraucht, brechen Sie sie ab. Zum Beispiel: Erstellen Sie eine Liste von Ticketverkäufern, die die Namen der Verkäufer und die Anzahl der verkauften Tickets enthält. Die folgende Abfrage wählt Daten aus der SALES-Tabelle und der USERS-Tabelle aus und verbindet beide Tabellen durch den Abgleich von SELLERID und USERID in der WHERE-Klausel.

```
SELECT sellerid, firstname, lastname, sum(qtysold)  
FROM sales, users  
WHERE sales.sellerid = users.userid  
GROUP BY sellerid, firstname, lastname  
ORDER BY 4 desc;
```

Das Ergebnis sieht in etwa wie folgt aus.

```
sellerid | firstname | lastname | sum
```

```
-----+-----+-----+-----  
48950 |   Nayda   |   Hood   | 184  
19123 |   Scott   | Simmons  | 164  
20029 |   Drew    | Mcguire  | 164  
36791 | Emerson  | Delacruz | 160  
13567 |   Imani   | Adams    | 156  
9697  | Dorian    | Ray      | 156  
41579 | Harrison  | Durham   | 156  
15591 | Phyllis   | Clay     | 152  
3008  | Lucas    | Stanley  | 148  
44956 | Rachel    | Villarreal | 148
```

Note

Dies ist eine komplexe Abfrage. Für dieses Tutorial müssen Sie sich über den Aufbau dieser Abfrage keine Gedanken machen.

Die vorherige Abfrage dauert wenige Sekunden und gibt 2 102 Zeilen aus.

Angenommen, Sie hätten die WHERE-Klausel vergessen.

```
SELECT sellerid, firstname, lastname, sum(qtysold)  
FROM sales, users  
GROUP BY sellerid, firstname, lastname  
ORDER BY 4 desc;
```

Der Ergebnissatz enthält dann alle Zeilen in der SALES-Tabelle, multipliziert mit allen Zeilen in der USERS-Tabelle (49989*3766). Dies ist eine so genannte Cartesische Verbindung, die nicht zu empfehlen ist. Das Ergebnis sind mehr als 188 Millionen Zeilen, und die Verarbeitungszeit ist extrem lang.

Um eine laufende Abfrage abubrechen, verwenden Sie den CANCEL-Befehl mit der Sitzungs-ID der Abfrage. Mit Amazon Redshift Query Editor v2 können Sie eine Abfrage abbrechen, indem Sie auf die Schaltfläche „Abbrechen“ klicken, während die Abfrage ausgeführt wird.

Um die Sitzungs-ID zu finden, starten Sie eine neue Sitzung und fragen Sie die Tabelle STV_RECENTS ab, wie im vorherigen Schritt gezeigt. Das folgende Beispiel zeigt, wie Sie die Ergebnisse lesbarer machen können. Verwenden Sie dazu die TRIM-Funktion, um nachfolgende Leerzeichen abzuschneiden, und zeigen Sie nur die ersten 20 Zeichen der Abfragezeichenfolge an.

Um die Sitzungs-ID einer laufenden Abfrage zu bestimmen, führen Sie die folgende SELECT-Anweisung aus.

```
SELECT user_id, session_id, start_time, query_text
FROM sys_query_history
WHERE status='running';
```

Das Ergebnis sieht in etwa wie folgt aus.

```
user_id | session_id | start_time | query_text
-----+-----+-----+-----
+-----+-----+-----+-----
100     | 1073791534 | 2024-03-19 22:26:21.205739 | SELECT user_id, session_id,
start_time, query_text FROM ...
```

Führen Sie zum Abbrechen der Abfrage mit der Sitzungs-ID 1073791534 den folgenden Befehl aus.

```
CANCEL 1073791534;
```

Note

Der Befehl CANCEL stoppt eine Transaktion nicht. Um eine Transaktion zu stoppen oder rückgängig zu machen, müssen Sie den Befehl ABORT oder ROLLBACK verwenden. Um eine mit einer Transaktion verbundene Abfrage abzuberechnen, brechen Sie zuerst die Abfrage ab und stoppen Sie dann die Transaktion.

Wenn die abgebrochene Abfrage mit einer Transaktion verbunden ist, verwenden Sie den Befehl ABORT oder ROLLBACK, um die Transaktion abzuberechnen und alle an den Daten vorgenommen Änderungen zu verwerfen:

```
ABORT;
```

Sie können nur Ihre eigenen Abfragen abbrechen, sofern Sie nicht als Superuser angemeldet sind. Superuser können alle Abfragen abbrechen.

Wenn Ihr Abfragetool nicht die gleichzeitige Ausführung von Abfragen unterstützt, starten Sie zum Abbruch der Abfrage eine weitere Sitzung.

Weitere Informationen zum Abbrechen einer Abfrage finden Sie unter [CANCEL](#) im Datenbankentwicklerhandbuch zu Amazon Redshift.

Abbrechen einer Abfrage mit der Superuser-Warteschlange

Wenn in Ihrer aktuellen Sitzung zu viele Abfragen gleichzeitig ausgeführt werden, können Sie möglicherweise erst dann den CANCEL-Befehl ausführen, wenn eine andere Abfrage abgeschlossen ist. Führen Sie in diesem Fall den CANCEL-Befehl mit einer anderen Workload-Verwaltungs-Abfragewarteschlange aus.

Workload-Verwaltung ermöglicht Ihnen die Ausführung von Abfragen in verschiedenen Abfragewarteschlangen, so dass Sie nicht warten müssen, bis eine andere Abfrage abgeschlossen ist. Der Workload Manager erstellt eine separate Warteschlange mit der Bezeichnung „Superuser-Warteschlange“, die Sie für Fehlerbehebungsziecke verwenden können. Um die Superuser-Warteschlange verwenden zu können, melden Sie sich als Superuser an und setzen Sie die Abfragegruppe mit dem SET-Befehl auf „Superuser“. Setzen Sie nach der Ausführung Ihrer Befehle die Abfragegruppe mit dem RESET-Befehl wieder zurück.

Führen Sie diese Befehle aus, um eine Abfrage mithilfe der Superuser-Warteschlange abzuberechnen:

```
SET query_group TO 'superuser';  
CANCEL 1073791534;  
RESET query_group;
```

Abfragen von Daten, die sich nicht in Ihrer Amazon-Redshift-Datenbank befinden

Nachfolgend erfahren Sie mehr über die ersten Schritte beim Abfragen von Daten in Remote-Quellen, einschließlich Amazon-S3-Daten, externen Datenbank-Managern, externen Amazon-Redshift-Datenbanken sowie zum Training von Machine Learning (ML)-Modellen mit Amazon Redshift.

Themen

- [Abfragen Ihres Data Lake](#)
- [Abfragen von Daten in externen Datenbank-Managern](#)
- [Zugriff auf Daten in anderen Amazon-Redshift-Datenbanken](#)
- [Training von Machine-Learning-Modellen mit Amazon-Redshift-Daten](#)

Abfragen Ihres Data Lake

Mit Amazon Redshift Spectrum können Sie Daten in Amazon-S3-Dateien abfragen, ohne die Daten in Amazon-Redshift-Tabellen laden zu müssen. Amazon Redshift bietet SQL-Funktionen für die schnelle Online-Analyseverarbeitung (OLAP) von sehr großen Datensätzen, die sowohl in Amazon-Redshift-Clustern als auch Amazon-S3-Data-Lakes gespeichert sind. Sie können Daten in vielen Formaten abfragen, darunter Parquet, ORC, RCFile, TextFile, SequenceFile, RegexSerde, OpenCSV und AVRO. Um die Struktur der Dateien in Amazon S3 zu definieren, erstellen Sie externe Schemata und Tabellen. Anschließend verwenden Sie einen externen Datenkatalog wie AWS Glue oder Ihren eigenen Apache-Hive-Metastore. Änderungen an einem der Datenkatalogtypen sind sofort für jeden Ihrer Amazon-Redshift-Cluster verfügbar.

Nachdem Ihre Daten bei einem AWS Glue-Datenkatalog registriert und mit AWS Lake Formation aktiviert wurden, können Sie sie mit Redshift Spectrum abfragen.

Redshift Spectrum befindet sich auf dedizierten Amazon-Redshift-Servern, die von Ihrem Cluster unabhängig sind. Redshift Spectrum verschiebt viele datenverarbeitungsintensive Aufgaben, wie etwa die Prädikatfilterung und -aggregation, auf die Redshift-Spectrum-Ebene. Redshift Spectrum lässt sich auch intelligent skalieren, um die Vorteile der massiv parallelen Verarbeitung zu nutzen.

Sie können die externen Tabellen in einer oder mehreren Spalten partitionieren, um die Abfrageleistung durch Partitionseliminierung zu optimieren. Sie können die externen Tabellen mit

Amazon-Redshift-Tabellen abfragen und verknüpfen. Sie können externe Tabellen von mehreren Amazon-Redshift-Clustern abrufen und die Amazon-S3-Daten aus jedem Cluster in derselben AWS-Region abfragen. Wenn Sie Amazon-S3-Datendateien aktualisieren, stehen diese Daten sofort zur Abfrage von allen Ihren Amazon-Redshift-Clustern aus zur Verfügung.

Weitere Informationen zu Redshift Spectrum, einschließlich zur Arbeit mit Redshift Spectrum und Data Lakes, finden Sie unter [Erste Schritte mit Amazon Redshift Spectrum](#) im Datenbankentwicklerhandbuch zu Amazon Redshift.

Abfragen von Daten in externen Datenbank-Managern

Sie können Daten mithilfe einer Verbundabfrage aus einer Amazon-RDS-Datenbank oder einer Amazon-Aurora-Datenbank mit Daten in Ihrer Amazon-Redshift-Datenbank verknüpfen. Mit Amazon Redshift können Sie Betriebsdaten direkt abfragen (ohne sie zu verschieben), Transformationen anwenden und Daten in Ihre Redshift-Tabellen einfügen. Ein Teil der Berechnung für Verbundabfragen wird an die Remote-Datenquellen verteilt.

Um Verbundabfragen auszuführen, stellt Amazon Redshift zunächst eine Verbindung zur Remote-Datenquelle her. Amazon Redshift ruft dann Metadaten zu den Tabellen in der Remote-Datenquelle ab, gibt Abfragen aus und ruft dann die Ergebniszeilen ab. Amazon Redshift verteilt die Ergebniszeilen dann zur weiteren Verarbeitung an Amazon-Redshift-Rechenknoten.

Weitere Informationen zum Einrichten der Umgebung für Verbundabfragen finden Sie in den folgenden Themen im Datenbankentwicklerhandbuch zu Amazon Redshift:

- [Erste Schritte mit der Verwendung von Verbundabfragen an PostgreSQL](#)
- [Erste Schritte beim Verwenden von Verbundabfragen für MySQL](#)

Zugriff auf Daten in anderen Amazon-Redshift-Datenbanken

Erfahren Sie, wie Sie mithilfe von Amazon-Redshift-Datenfreigabe Live-Daten sicher und einfach über Amazon-Redshift-Cluster oder AWS-Konten hinweg für Lesezwecke freigeben. Sie profitieren von sofortigem, granulearem und leistungsstarkem Zugriff auf Daten in Amazon-Redshift-Clustern, ohne diese manuell zu kopieren oder zu verschieben. Ihren Benutzer werden aktuelle und konsistente Informationen angezeigt, wenn diese in Amazon-Redshift-Clustern aktualisiert werden. Sie können Daten auf verschiedenen Ebenen freigeben, z. B. Datenbanken, Schemata, Tabellen, Ansichten (einschließlich regulärer Ansichten, Ansichten mit späterer Bindung und materialisierter Ansichten) sowie benutzerdefinierten SQL-Funktionen (UDFs).

Die Amazon-Redshift-Datenfreigabe ist besonders für folgende Anwendungsfälle nützlich:

- Zentralisierung geschäftskritischer Workloads – Verwenden Sie einen zentralen Extract, Transform, Load (ETL)-Cluster, der Daten mit mehreren Business Intelligence (BI)- oder Analyse-Clustern gemeinsam verwendet. Dieser Ansatz bietet Lese-Workload-Isolation und Rückbelastung für einzelne Workloads.
- Freigabe von Daten zwischen Umgebungen – Teilen Sie Daten in Entwicklungs-, Test- und Produktionsumgebungen. Sie können die Teamagilität verbessern, indem Sie Daten auf verschiedenen Granularitätsstufen teilen.

Weitere Informationen zur Datenfreigabe finden Sie unter [Umgang mit Datenfreigabeaufgaben](#) im Datenbankentwicklerhandbuch zu Amazon Redshift.

Training von Machine-Learning-Modellen mit Amazon-Redshift-Daten

Mit Amazon Redshift Machine Learning (Amazon Redshift ML) können Sie ein Modell trainieren, indem Sie die Daten an Amazon Redshift bereitstellen. Dann erstellt Amazon Redshift ML Modelle, die Muster in den Eingabedaten erfassen. Sie können diese Modelle dann verwenden, um Prognosen für neue Eingabedaten zu generieren, ohne dass zusätzliche Kosten entstehen. Mithilfe von Amazon Redshift ML können Sie Machine-Learning-Modelle mithilfe von SQL-Anweisungen trainieren und sie in SQL-Abfragen für Prognosen aufrufen. Sie können die Genauigkeit der Prognosen weiter verbessern, indem Sie die Parameter iterativ ändern und Ihre Trainingsdaten verbessern.

Amazon Redshift ML erleichtert SQL-Benutzern das Erstellen, Trainieren und Bereitstellen von Machine-Learning-Modellen mit vertrauten SQL-Befehlen. Mit Amazon Redshift ML können Sie Ihre Daten in Amazon-Redshift-Clustern verwenden, um Modelle mit Amazon SageMaker AI Autopilot zu trainieren und automatisch das beste Modell zu erhalten. Sie können dann die Modelle lokalisieren und Prognosen innerhalb einer Amazon-Redshift-Datenbank erstellen.

Weitere Informationen zu Amazon Redshift ML finden Sie unter [Erste Schritte mit Amazon Redshift ML](#) im Datenbankentwicklerhandbuch zu Amazon Redshift.

Über Amazon-Redshift-Konzepte

Mit Amazon Redshift Serverless können Sie auf Daten zugreifen und diese analysieren, ohne alle Konfigurationen wie bei einem bereitgestellten Data Warehouse vornehmen zu müssen. Ressourcen werden automatisch bereitgestellt und die Data-Warehouse-Kapazität wird intelligent skaliert, um eine schnelle Leistung selbst für anspruchsvollste und unvorhersehbare Workloads zu erzielen. Es fallen keine Kosten an, wenn das Data Warehouse inaktiv ist, Sie zahlen also nur für das, was Sie tatsächlich nutzen. Sie können Daten laden und sofort mit der Abfrage beginnen. Hierfür können Sie Amazon Redshift Query Editor v2 oder Ihr bevorzugtes Business Intelligence (BI)-Tool nutzen. Profitieren Sie von einem optimalen Preis-Leistungs-Verhältnis und nutzen Sie vertraute SQL-Funktionen in einer benutzerfreundlichen Umgebung ohne Verwaltungsaufwand.

Wenn Sie Amazon Redshift zum ersten Mal verwenden, empfehlen wir Ihnen, zunächst die folgenden Abschnitte zu lesen:

- [Übersicht über die Funktionen von Amazon Redshift Serverless](#) – Unter diesem Thema finden Sie eine Übersicht über Amazon Redshift Serverless und seine wichtigsten Funktionen.
- [Service-Merkmale und Preise](#) – Auf dieser Produktdetailseite erfahren Sie mehr zu den Merkmalen und Preisen von Amazon Redshift Serverless.
- [Erste Schritte mit Data Warehouses von Amazon Redshift Serverless](#). – In diesem Thema erfahren Sie, wie Sie ein Data Warehouse von Amazon Redshift Serverless erstellen und mit dem Abfragen von Daten mithilfe von Query Editor v2 beginnen können.

Wenn Sie Ihre Amazon-Redshift-Ressourcen lieber manuell verwalten möchten, können Sie bereitgestellte Cluster für Ihre Datenabfrageanforderungen erstellen. Weitere Informationen finden Sie unter [Amazon-Redshift-Cluster](#).

Wenn Ihre Organisation berechtigt ist und Ihr Cluster in einer AWS-Region erstellt wird, in der Amazon Redshift Serverless nicht verfügbar ist, können Sie möglicherweise im Rahmen des kostenlosen Testprogramms von Amazon Redshift einen Cluster erstellen. Wählen Sie entweder Produktion oder Kostenlose Testversion als Antwort auf die Frage: Wofür möchten Sie diesen Cluster verwenden? Wenn Sie Kostenlose Testversion auswählen, erstellen Sie eine Konfiguration mit dem Knotentyp dc2.large. Weitere Informationen zur Auswahl einer kostenlosen Testversion finden Sie unter [Kostenloses Testprogramm für Amazon Redshift](#). Eine Liste der AWS-Regionen, in denen Amazon Redshift Serverless verfügbar ist, finden Sie in der Übersicht über die Amazon-Redshift-Endpunkte für die [Redshift-Serverless-API](#) in der Allgemeine Amazon Web Services-Referenz.

Im Folgenden sind einige wichtige Konzepte von Amazon Redshift Serverless aufgeführt.

- **Namespace** – Eine Sammlung von Datenbankobjekten und Benutzern. In Namespaces sind alle Ressourcen zusammengefasst, die Sie in Amazon Redshift Serverless verwenden, wie Schemas, Tabellen, Benutzer, Datashares und Snapshots.
- **Arbeitsgruppe** – Eine Sammlung von Rechenressourcen. In Arbeitsgruppen sind Rechenressourcen enthalten, die Amazon Redshift Serverless zur Ausführung von Datenverarbeitungsaufgaben verwendet. Einige Beispiele für solche Ressourcen sind Redshift Processing Units (RPU), Sicherheitsgruppen und Nutzungslimits. Arbeitsgruppen verfügen über Netzwerk- und Sicherheitseinstellungen, die Sie über die Amazon-Redshift-Serverless-Konsole, die AWS Command Line Interface oder die Amazon-Redshift-Serverless-APIs konfigurieren können.

Weitere Informationen zum Konfigurieren von Namespace- und Arbeitsgruppenressourcen finden Sie unter [Arbeiten mit Namespaces](#) und [Arbeiten mit Arbeitsgruppen](#).

Im Folgenden sind einige wichtige Konzepte im Zusammenhang mit von Amazon Redshift bereitgestellten Clustern aufgeführt:

- **Cluster** – Die zentrale Infrastrukturkomponente eines Amazon-Redshift-Data-Warehouse ist ein Cluster.

Ein Cluster besteht aus einem oder mehreren Datenverarbeitungsknoten. Die Datenverarbeitungsknoten führen den kompilierten Code aus.

Wird ein Cluster mit zwei oder mehr Datenverarbeitungsknoten bereitgestellt, koordiniert ein zusätzlicher Führungsknoten die Datenverarbeitungsknoten. Der Führungsknoten übernimmt die externe Kommunikation mit Anwendungen, wie Business-Intelligence-Tools und Abfrage-Editoren. Ihre Client-Anwendung interagiert nur mit dem Führungsknoten direkt. Die Datenverarbeitungsknoten sind für externe Anwendungen transparent.

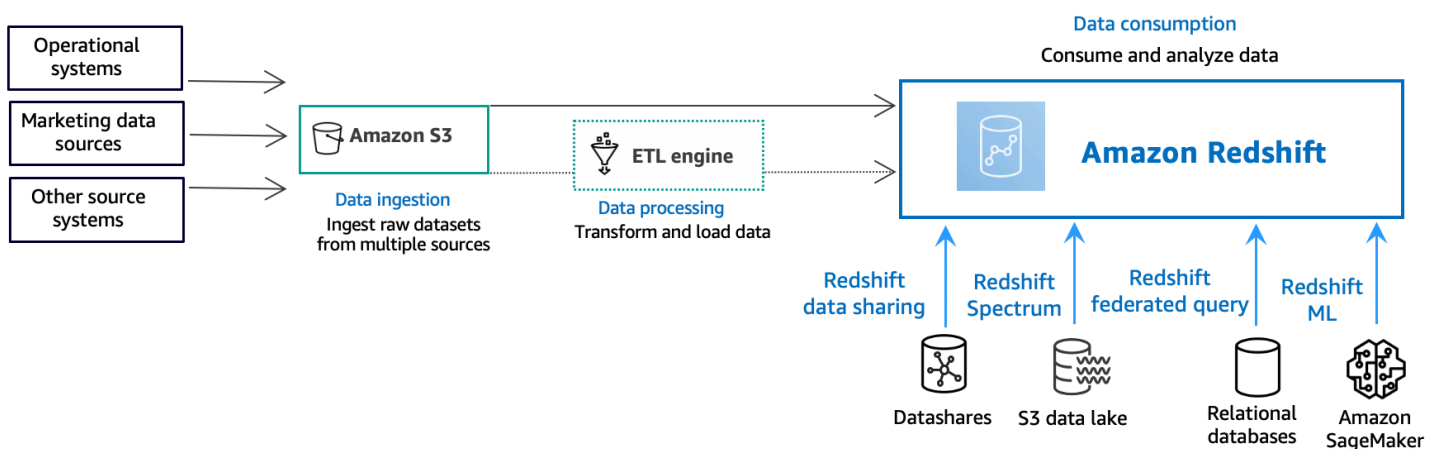
- **Datenbank** – Ein Cluster enthält eine oder mehrere Datenbanken.

Benutzerdaten werden in einer oder mehreren Datenbanken auf den Datenverarbeitungsknoten gespeichert. Ihr SQL-Client kommuniziert mit dem Führungsknoten, der wiederum die Abfrageausführung mit den Datenverarbeitungsknoten koordiniert. Weitere Informationen zu Datenverarbeitungs- und Führungsknoten finden Sie unter [Data-Warehouse-Systemarchitektur](#). Innerhalb einer Datenbank sind Benutzerdaten in einem Schema oder mehreren Schemata organisiert.

Amazon Redshift ist ein relationales Datenbankmanagementsystem (RDBMS) und ist mit anderen RDBMS-Anwendungen kompatibel. Amazon Redshift stellt dieselben Funktionen wie ein typisches RDBMS bereit, einschließlich Funktionen zur Online-Transaktionsverarbeitung (Online Transaction Processing, OLTP), wie das Einfügen und Löschen von Daten. Amazon Redshift ist auch für leistungsfähige Batchanalysen und Berichterstattung von Datensätzen optimiert.

Im Folgenden finden Sie eine Beschreibung des typischen Datenverarbeitungsablaufs in Amazon Redshift sowie Beschreibungen verschiedener Teile im Ablauf. Weitere Informationen zur Amazon-Redshift-Systemarchitektur finden Sie unter [Architektur des Data-Warehouse-Systems](#).

Das folgende Diagramm zeigt einen typischen Datenverarbeitungsablauf in Amazon Redshift.



Ein Amazon-Redshift-Data-Warehouse ist eine Abfrage- und Verwaltungssystem der Enterprise-Klasse für relationale Datenbanken. Amazon Redshift unterstützt Client-Verbindungen mit vielen Arten von Anwendungen, einschließlich Business Intelligence (BI), Berichterstellung, Daten und Analysetools. Bei Analyseabfragen werden große Datenmengen in mehrphasigen Operationen abgerufen, verglichen und bewertet, um ein Endergebnis zurückzugeben.

Auf der Ebene der Datenerfassung laden verschiedene Arten von Datenquellen kontinuierlich strukturierte, halbstrukturierte oder unstrukturierte Daten in die Datenspeicher-Ebene hoch. Dieser Datenspeicherbereich dient als Staging-Bereich, der Daten in verschiedenen Zuständen der Nutzungsbereitschaft speichert. Ein Beispiel für einen solchen Speicher ist ein Amazon Simple Storage Service (Amazon S3)-Bucket.

In der optionalen Ebene Datenverarbeitung durchlaufen die Quelldaten die Vorverarbeitung, Validierung und Transformation über Extract, Transform, Load (ETL)-oder Extract, Load, Transform

(ELT)-Pipelines. Diese Rohdatensätze werden dann mithilfe von ETL-Operationen verfeinert. Ein Beispiel für eine ETL-Engine ist AWS Glue.

In der Ebene Datennutzung werden Daten in Ihren Amazon-Redshift-Cluster geladen, wo Sie Analyse-Workloads ausführen können.

Beispiele für Analyse-Workloads finden Sie unter [Abfragen von externen Datenquellen](#).

Weitere Ressourcen, um mehr über Amazon Redshift zu erfahren.

Wenn Sie mehr über Amazon Redshift Serverless erfahren möchten, empfehlen wir Ihnen, Ihr Wissen über die in diesem Handbuch vorgestellten Konzepte unter Verwendung der folgenden Ressourcen zu Amazon Redshift zu vertiefen:

- **Feature-Videos:** Diese Videos helfen Ihnen, mehr über die Funktionen von Amazon Redshift zu erfahren.
 - Um eine allgemeine Vorstellung von Amazon Redshift Serverless zu erhalten, sehen Sie sich das folgende Video an. [Amazon Redshift Serverless Explained in 90 Seconds](#) (Amazon Redshift Serverless in 90 Sekunden erklärt).
 - Wenn Sie wissen möchten, wie Sie ein Serverless Data Warehouse einrichten und mit dem Abfragen von Daten beginnen können, sehen Sie sich das folgende Video an. [Getting Started with Amazon Redshift Serverless](#) (Erste Schritte mit Amazon Redshift Serverless)
- [Amazon-Redshift-Verwaltungshandbuch](#): Dieser Leitfaden baut auf dem Handbuch Erste Schritte mit Amazon Redshift auf. Sie erhalten darin detaillierte Informationen über die Konzepte und Aufgaben für die Erstellung, Verwaltung und Überwachung von Clustern, die von Amazon Redshift Serverless und Amazon Redshift bereitgestellt werden.
- [Datenbankentwicklerhandbuch zu Amazon Redshift](#): Dieses Handbuch baut ebenfalls auf dem Handbuch Erste Schritte mit Amazon Redshift auf. Es richtet sich an Datenbankentwickler und vermittelt fundierte Kenntnisse auf den Gebieten Entwurf, Entwicklung, Abfrage und Verwaltung von Datenbanken in einem Data Warehouse.
 - [SQL-Referenz](#): In diesem Thema werden SQL-Befehle und Funktionsreferenzen für Amazon Redshift beschrieben.
 - [Referenz zu Systemtabellen und Ansichten](#): In diesem Thema werden Systemtabellen und Ansichten für Amazon Redshift beschrieben.
- **Tutorials für Amazon Redshift:** In diesem Thema werden Tutorials zu den Funktionen von Amazon Redshift angezeigt.
 - [So laden Sie Daten aus Amazon S3](#): In diesem Tutorial wird beschrieben, wie Sie Daten aus Datendateien in einem Amazon-S3-Bucket in Ihre Amazon-Redshift-Datenbanktabellen laden.
 - [Erste Schritte mit der Datenfreigabe](#): In diesem Abschnitt wird beschrieben, wie Sie Daten freigeben und auf Daten in anderen Amazon-Redshift-Clustern zugreifen können.

- [Verwendung von Geo-SQL-Funktionen mit Amazon Redshift](#): In diesem Tutorial wird gezeigt, wie Sie einige der Geo-SQL-Funktionen mit Amazon Redshift verwenden.
- [Abfragen verschachtelter Daten mit Amazon Redshift Spectrum](#): In diesem Tutorial wird beschrieben, wie Sie Redshift Spectrum verwenden, um verschachtelte Daten in den Dateiformaten Parquet, ORC, JSON und Ion mit externen Tabellen abzufragen.
- [Konfigurieren von manuellen Workload-Management \(WLM\)-Warteschlangen](#): In diesem Tutorial wird beschrieben, wie Sie das manuelle Workload-Management (WLM) in Amazon Redshift konfigurieren.
- [Erste Schritte mit Amazon Redshift ML](#): In diesem Abschnitt wird beschrieben, wie Benutzer unter Verwendung von vertrauten SQL-Befehlen Machine-Learning-Modelle erstellen, trainieren und bereitstellen können.
- [Neuerungen](#): Diese Webseite listet neue Funktionen von Amazon Redshift und Produktaktualisierungen auf.

Dokumentverlauf

Note

Eine Beschreibung der neuen Features in Amazon Redshift finden Sie unter [Was ist neu](#).

In der folgenden Tabelle werden wichtige Änderungen seit der letzten Version des Handbuchs Erste Schritte mit Amazon Redshift beschrieben.

Änderungen	Beschreibung	Datum der Veröffentlichung
Aktualisierung der Dokumentation	Der Leitfaden wurde aktualisiert, um die von Query Editor v2 verwalteten Richtlinienänderungen und die verbesserten Zugriffsberechtigungen für Serverless-Namespaces und Arbeitsgruppen widerzuspiegeln.	21. Februar 2024
Aktualisierung der Dokumentation	Die Screenshots und Verfahren wurden aktualisiert, um den neuesten Verbesserungen der Konsoloberfläche und von Query Editor v2 Rechnung zu tragen.	11. März 2023
Neues Feature	Der Leitfaden wurde aktualisiert und enthält jetzt auch Verfahren und Workflows für die ersten Schritte mit Amazon Redshift Serverless. Ein umfassender Abschnitt zur Erstellung und Verwaltung von Serverless-Data Warehouses wurde hinzugefügt.	12. Juli 2022
Aktualisierung der Dokumentation	Der Leitfaden wurde aktualisiert, so dass Query Editor v2 als primäre Abfrageschnittstelle verwendet wird und die Verweise auf den alten Abfrage-Editor ersetzt wurden.	Februar 2022
Aktualisierung der Dokumentation	Das Handbuch wurde aktualisiert, um neue Abschnitte über die ersten Schritte mit allgemeinen Datenbankaufgaben, das Abfragen Ihres Data Lake, das	30. Juni 2021

Änderungen	Beschreibung	Datum der Veröffentlichung
	Abfragen von Daten auf entfernten Quellen, das Freigeben von Daten und das Training von Machine-Learning-Modellen mit Amazon-Redshift-Daten zu enthalten.	
Neues Feature	Aktualisierung des Handbuchs, um das Verfahren des neuen Beispielladevorgangs zu beschreiben.	4. Juni 2021
Aktualisierung der Dokumentation	Die ursprüngliche Amazon-Redshift-Konsole wurde aus dem Handbuch entfernt und der Ablauf der Schritte wurde verbessert.	14. August 2020
New console	Das Handbuch enthält nun eine Beschreibung der neuen Amazon-Redshift-Konsole.	11. November 2019
Neues Feature	Aktualisierung des Handbuchs, um das Verfahren für schnelle Clusterstarts zu beschreiben.	10. August 2018
Neues Feature	Das Handbuch enthält nun Informationen zum Starten von Clustern über das Amazon-Redshift-Dashboard.	28. Juli 2015
Neues Feature	Das Handbuch enthält nun Informationen zur Verwendung von neuen Knotentypnamen.	9. Juni 2015
Aktualisierung der Dokumentation	Die Dokumentation enthält aktualisierte Screenshots und Verfahren zur Konfiguration der VPC-Sicherheitsgruppen.	30. April 2015
Aktualisierung der Dokumentation	Die Dokumentation enthält aktualisierte Screenshots und beschreibt, wie sich die aktuelle Konsole und die Screenshots aufeinander abstimmen lassen.	12. November 2014

Änderungen	Beschreibung	Datum der Veröffentlichung
Aktualisierung der Dokumentation	Zur besseren Auffindbarkeit wurden die Informationen zum Laden von Daten von Amazon S3 in einen separaten Abschnitt verschoben und der Abschnitt mit den nächsten Schritten wurde in den letzten Schritt integriert.	13. Mai 2014
Aktualisierung der Dokumentation	Die Willkommenseite wurde entfernt; ihr Inhalt befindet sich nun auf der Hauptseite der Seite "Erste Schritte".	14. März 2014
Aktualisierung der Dokumentation	Bei dieser Dokumentation handelt es sich um eine neue Version des Handbuchs Erste Schritte mit Amazon Redshift, die Kundenfeedback und Service-Updates berücksichtigt.	14. März 2014
Neues Handbuch	Dies ist die erste Version des Handbuchs Erste Schritte mit Amazon Redshift.	14. Februar 2013