

# Creation of a Redshift Cluster

## Screenshots of the configuration of the Redshift cluster

[Amazon Redshift](#) > [Clusters](#) > redshift-cluster-1

redshift-cluster-1

Actions Edit Add partner integration Query data

General information Info

Cluster identifier redshift-cluster-1	Status Available	Node type dc2.large	Endpoint redshift-cluster-1.qtzylhhlhb.us-east-1.redshift.amazonaws.com:5439/dev
Custom domain name - new -	Date created October 12, 2023, 12:44 (UTC-04:00)	Number of nodes 2	JDBC URL jdbc:redshift://redshift-cluster-1.qtzylhhlhb.us-east-1.redshift.amazonaws.com:5439/dev
Cluster ARN arn:aws:redshift:us-east-1:163908874715:namespace:35570cff-e282-4f21-907a-107a1b41437c	Storage used -		ODBC URL Driver=(Amazon Redshift (x64)); Server=redshift-cluster-1.qtzylhhlhb.us-east-1.redshift.amazonaws.com; Database=dev
Cluster configuration Production	Multi-AZ No		

Cluster performance Query monitoring Schedules Maintenance Properties

Cluster performance Query monitoring Schedules Maintenance Properties

Database configurations Info

Edit admin credentials Rotate encryption keys Edit

Database name dev	Parameter group Defines database parameter and query queues for all the databases. default.redshift-1.0	Encryption Disabled	Audit logging Disabled
Port 5439	SSH ingestion setting (cluster public key) ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAQOkdyv l3LaVDx1Ao6MZ/Y7D+1QNY8uOP9I/s+wVUZ0 gUO6sp1CmSt3BSlCbj6IEQ9UHV+Fw8Md+GSH laIEJAZBYgWUGx41X+RIWWQV343KPHSc4faR LkeldV4T7o9eHeOK+lbOKNBKZUGHBSJ/RA AE4xyakVMWdV/7++VnAxPvBoB28c8uM0MU8 xejKCvuMMmEbmeVUzlhx1xscvCrYd57pDARX qqS9+fIoYqTh145s2gE/PwHmzQt47VRQeHME xHrGh3tnlwUDXMLJWGYW9tMxmIU/cbsSoC54P kSd/PCZlaRsVi+OC5AhgawwOnLT2g6P8oTpYC1 qmqzSXNAP Amazon-Redshift	AWS KMS key ID -	
Admin user name admin			

Setting up a database in the Redshift cluster and running queries to create the dimension and fact tables

**Queries to create the various dimension and fact tables with appropriate primary and foreign keys:**

## 1. Create Schema

### Query

```
create schema sparnodebankdata;
```

### Result

The screenshot displays the Amazon Redshift console interface. On the left, the 'Resources' sidebar shows the 'dev' database and 'public' schema selected. The main panel, titled 'Query 1', contains the SQL command: `create schema sparnodebankdata;`. Below the query editor, the 'Run' button is highlighted in orange, alongside 'Save', 'Schedule', and 'Clear' buttons. The 'Query results' tab is active, showing a status of 'Completed, started on October 12, 2023 at 12:53:18' with an 'ELAPSED TIME: 00 m 02 s'. At the bottom right, there are buttons for 'Execution', 'Data', and 'Visualize'.

## 2. Create DIM\_LOCATION table

### Query

```
create table sparnodebankdata.DIM_LOCATION  
( location_id int not null DISTKEY SORTKEY,  
  location varchar(50),  
  streetname varchar(255),  
  street_number int,  
  zipcode int,  
  lat decimal(10,3),  
  lon decimal(10,3),  
  PRIMARY KEY(location_id) );
```

### Result



## 4. Create DIM\_DATE table

### Query

```
create table sparnodebankdata.DIM_DATE
(
date_id int not null DISTKEY SORTKEY,
full_date_time timestamp,
year int,
month varchar(20),
day int,
hour int,
weekday varchar(20),
PRIMARY KEY(date_id)
);
```

### Result

The screenshot displays a database management tool interface. On the left, a 'Resources' panel shows the selected database 'dev' and schema 'public'. The main editor area contains a SQL query to create the 'DIM\_DATE' table. Below the query, buttons for 'Run', 'Save', 'Schedule', and 'Clear' are visible. The 'Query results' tab is active, showing a status message: 'Query Completed, started on October 12, 2025 at 15:08:13 ELAPSED TIME: 00 m 03 s'. The 'Table details' tab is also present but empty.

```
1 create table sparnodebankdata.DIM_DATE
2 (
3   date_id int not null DISTKEY SORTKEY,
4   full_date_time timestamp,
5   year int,
6   month varchar(20),
7   day int,
8   hour int,
9   weekday varchar(20),
10  PRIMARY KEY(date_id)
11 );
12
13
```

Query results | Table details

Query

Completed, started on October 12, 2025 at 15:08:13  
ELAPSED TIME: 00 m 03 s

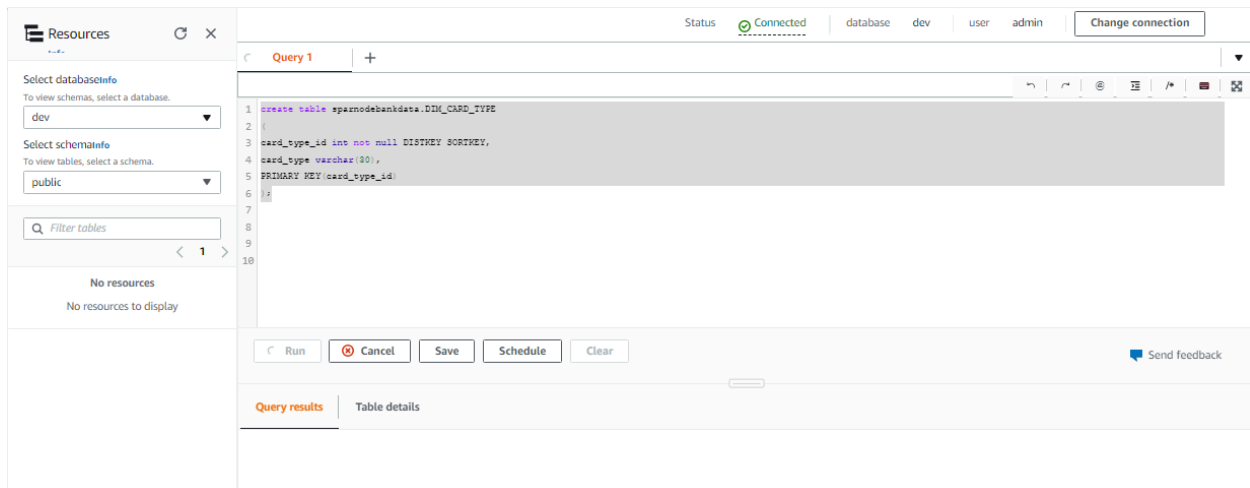
Execution | Data | Visualize

## 5. Create DIM\_CARD\_TYPE table

### Query

```
create table sparnodebankdata.DIM_CARD_TYPE
(
card_type_id int not null DISTKEY SORTKEY,
card_type varchar(30),
PRIMARY KEY(card_type_id)
);
```

### Result

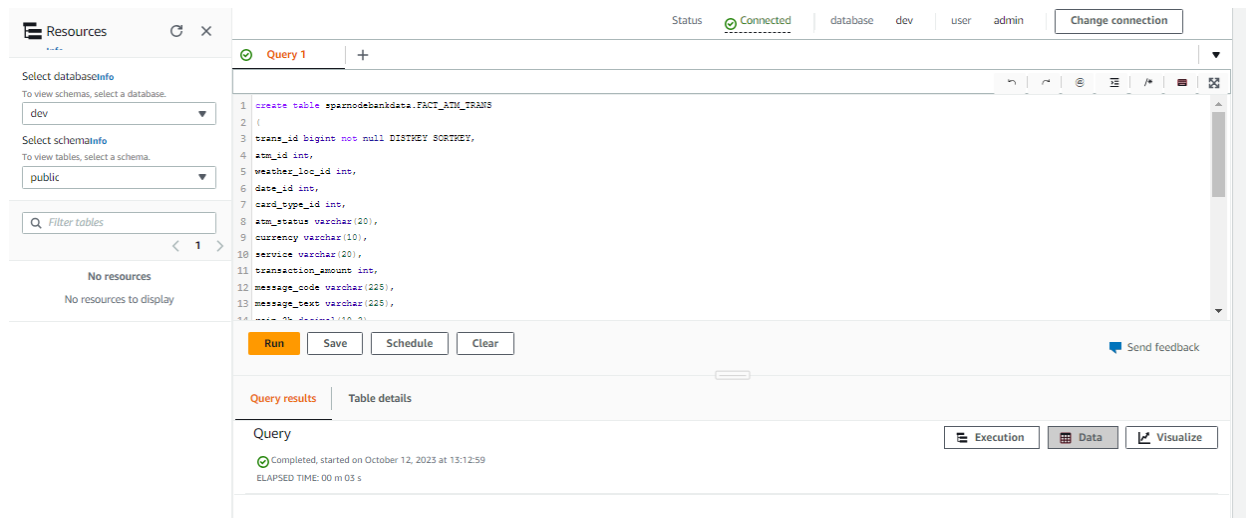


## 6. Create FACT\_ATM\_TRANS table

### Query

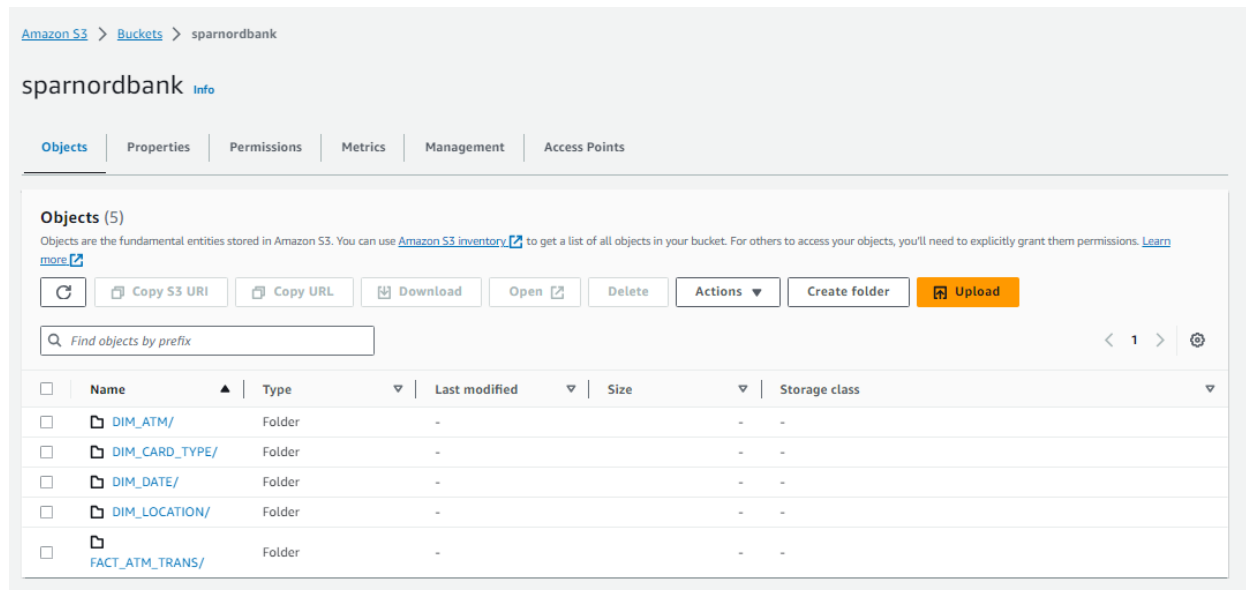
```
create table sparnodebankdata.FACT_ATM_TRANS
(
trans_id bigint not null DISTKEY SORTKEY,
atm_id int,
weather_loc_id int,
date_id int,
card_type_id int,
atm_status varchar(20),
currency varchar(10),
service varchar(20),
transaction_amount int,
message_code varchar(225),
message_text varchar(225),
rain_3h decimal(10,3),
clouds_all int,
weather_id int,
weather_main varchar(50),
weather_description varchar(255),
PRIMARY KEY(trans_id),
FOREIGN KEY(weather_loc_id) references sparnodebankdata.DIM_LOCATION(location_id),
FOREIGN KEY(atm_id) references sparnodebankdata.DIM_ATM(atm_id),
FOREIGN KEY(date_id) references sparnodebankdata.DIM_DATE(date_id),
FOREIGN KEY(card_type_id) references sparnodebankdata.DIM_CARD_TYPE(card_type_id)
);
```

### Result



## Loading data into a Redshift cluster from Amazon S3 bucket

### S3 with files



## Queries to copy the data from S3 buckets to the Redshift cluster in the appropriate tables

### DIM\_LOCATION

copy sparnodebankdata.dim\_location from  
 's3://sparnodebank/DIM\_LOCATION/part-00000-fd2c27e3-e9f5-4ff2-bee7-4366a4ff95dd-c000.csv'  
 iam\_role 'arn:aws:iam::163908874715:role/myRedshiftRole'

```
delimiter ',' region 'us-east-1'  
IGNOREHEADER 1  
CSV;
```

## Result

The screenshot displays the Amazon Redshift Query Editor interface. At the top, there's a navigation bar with 'Amazon Redshift' and 'Query editor'. Below this, a tabbed interface shows 'Editor', 'Query history', 'Saved queries', and 'Scheduled queries'. The 'Editor' tab is active, showing a SQL query in a text area. The query is as follows:

```
1 copy sparnodebankdata.dim_location from  
2 's3://sparnodebank/DIM_LOCATION/part-00000-fd2c27e3-e8f5-4fff-bae7-4366a4ff95dd-c000.csv'  
3 iam_role 'arn:aws:iam::163908874715:role/myRedshiftRole'  
4 delimiter ',' region 'us-east-1'  
5 IGNOREHEADER 1  
6 CSV;  
7  
8
```

Below the query editor, there are buttons for 'Run', 'Save', 'Schedule', and 'Clear'. To the right of these buttons is a 'Send feedback' link. Below the buttons, there's a section for 'Query results' and 'Table details'. The 'Query results' section shows 'Query 2272' with a status of 'Completed, started on October 12, 2023 at 13:36:51' and 'ELAPSED TIME: 00 m 07 s'. To the right of this section are buttons for 'Execution', 'Data', and 'Visualize'.

## DIM\_ATM

```
copy sparnodebankdata.DIM_ATM from  
's3://sparnordbank/DIM_ATM/part-00000-35020052-f23d-4962-a768-685e1a5248da-c000.csv'  
iam_role 'arn:aws:iam::163908874715:role/myRedshiftRole'  
delimiter ',' region 'us-east-1'  
IGNOREHEADER 1  
CSV;
```

## Result

Amazon Redshift > Query editor

Editor | Query history | Saved queries | Scheduled queries

Status Connected | database | dev | user | admin | [Change connection](#)

Query 1

```
1 copy sparnodebankdata.DIM_ATM from
2 's3://sparnodebank/DIM_ATM/part-00000-35020032-f23d-4962-a768-685a1a5248da-c000.csv'
3 iam_role 'arn:aws:iam::163908874715:role/myRedshiftRole'
4 delimiter ',' region 'us-east-1'
5 IGNOREHEADER 1
6 CSV;
```

Run Save Schedule Clear

Send feedback

Query results | Table details

Query 2655 [View](#)

Execution Data Visualize

Completed, started on October 12, 2023 at 14:14:25  
ELAPSED TIME: 00 m 09 s

## DIM\_DATE

```
copy sparnodebankdata.DIM_DATE from
's3://sparnodebank/DIM_DATE/part-00000-d47598eb-7286-4ebd-a4c3-069e7c12ca62-c000.csv'
iam_role 'arn:aws:iam::163908874715:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
IGNOREHEADER 1
CSV
TIMEFORMAT 'auto';
```

## Result

Amazon Redshift > Query editor

Editor | Query history | Saved queries | Scheduled queries

Resources

Select database info  
To view schemas, select a database.  
dev

Select schema info  
To view tables, select a schema.  
public

Filter tables  
No resources  
No resources to display

Status Connected | database | dev | user | admin | [Change connection](#)

Query 1

```
1 copy sparnodebankdata.DIM_DATE from
2 's3://sparnodebank/DIM_DATE/part-00000-d47598eb-7286-4ebd-a4c3-069e7c12ca62-c000.csv'
3 iam_role 'arn:aws:iam::163908874715:role/myRedshiftRole'
4 delimiter ',' region 'us-east-1'
5 IGNOREHEADER 1
6 CSV
7 TIMEFORMAT 'auto';
```

Run Save Schedule Clear

Send feedback

Query results | Table details

Query 3068 [View](#)

Execution Data Visualize

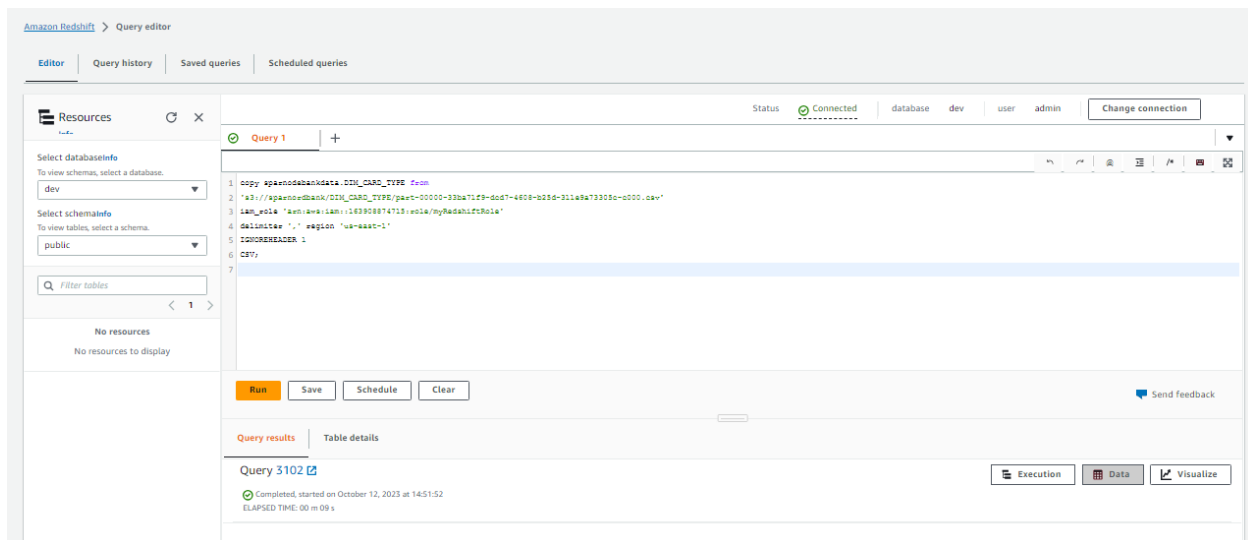
Completed, started on October 12, 2023 at 14:50:27  
ELAPSED TIME: 00 m 04 s



## DIM\_CARD\_TYPE

```
copy sparnodebankdata.DIM_CARD_TYPE from
's3://sparnordbank/DIM_CARD_TYPE/part-00000-33ba71f9-dcd7-4608-b25d-311e9a73305c-
c000.csv'
iam_role 'arn:aws:iam::163908874715:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
IGNOREHEADER 1
CSV;
```

## Result



## FACT\_ATM\_TRANS

```
copy sparnodebankdata.FACT_ATM_TRANS from
's3://sparnordbank/FACT_ATM_TRANS/part-00000-b8d4a013-e60f-45de-86d8-976cd45668ee-
c000.csv'
iam_role 'arn:aws:iam::163908874715:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
IGNOREHEADER 1
CSV;
```

## Result

Resources

Select database info

To view schemas, select a database.

dev

Select schema info

To view tables, select a schema.

public

Filter tables

< 1 >

No resources

No resources to display

Status Connected database dev user admin Change connection

Query 1

```
1 copy s3://aparnodbankdata.FACT_ATM_TRANS from
2 's3://aparnodbank/FACT_ATM_TRANS/part-0000-b8da013-e02f-43da-86d8-976cd43688ea-0000.csv'
3 iam_role 'arn:aws:iam::163908874713:role/MyRedshiftRole'
4 delimiter ',' region 'us-east-1'
5 OVERSAMPLER 1
6 CSV;
```

Run

Save

Schedule

Clear

Send feedback

Query results

Table details

Query 3126

Completed, started on October 12, 2023 at 14:52:48  
ELAPSED TIME: 00 m 13 s

Execution

Data

Visualize