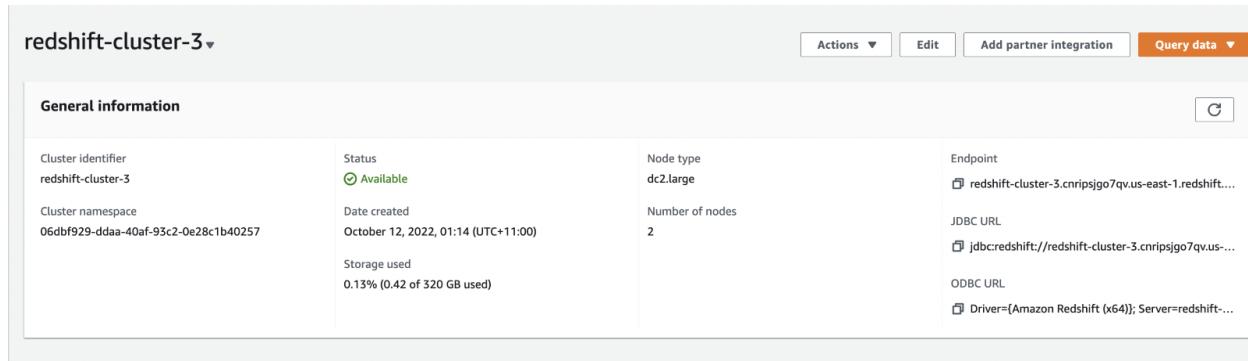


Creation of a Redshift Cluster

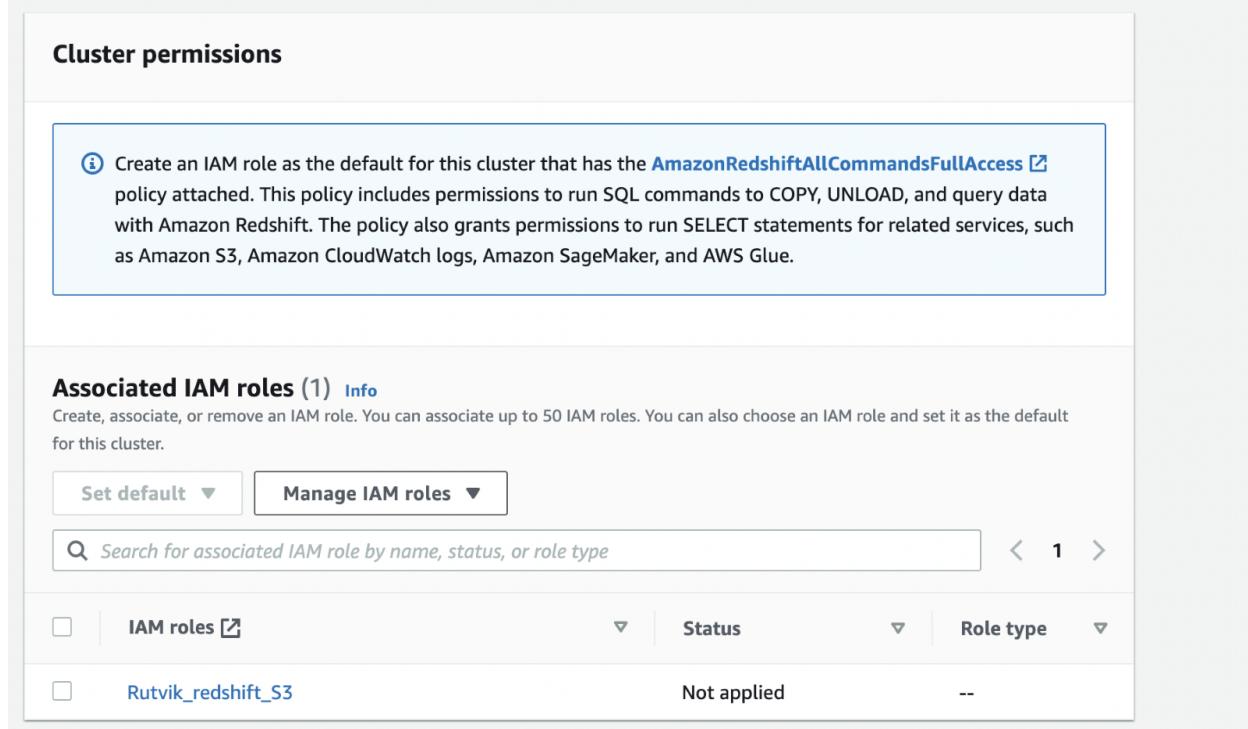
Screenshots of the configuration of the Redshift cluster that you have created:

Number of Nodes = 2 and Machine Selected is dc2



General information			
Cluster identifier redshift-cluster-3	Status Available	Node type dc2.large	Endpoint redshift-cluster-3.cnriipsjgo7qv.us-east-1.redshift.amazonaws.com
Cluster namespace 06dbf929-ddaa-40af-95c2-0e28c1b40257	Date created October 12, 2022, 01:14 (UTC+11:00)	Number of nodes 2	JDBC URL jdbc:redshift://redshift-cluster-3.cnriipsjgo7qv.us-east-1.redshift.amazonaws.com:5439/redshift-cluster-3
	Storage used 0.13% (0.42 of 320 GB used)		ODBC URL Driver={Amazon Redshift (x64)}; Server=redshift-cluster-3.cnriipsjgo7qv.us-east-1.redshift.amazonaws.com; Port=5439; Database=public

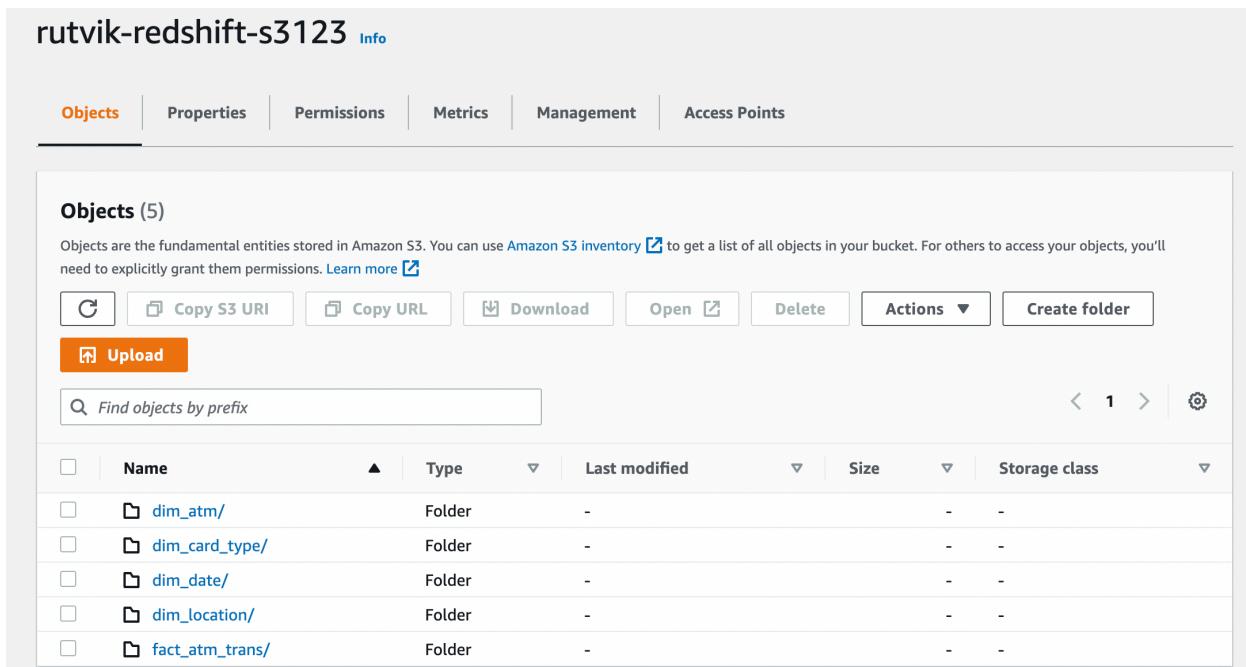
Associated the IAM Role



Associated IAM roles (1) Info			
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 for this cluster.			
<input type="button" value="Set default ▾"/>	<input type="button" value="Manage IAM roles ▾"/>	<input type="text" value="Search for associated IAM role by name, status, or role type"/> < 1 >	
<input type="checkbox"/>	IAM roles 	<input type="button" value="Status ▾"/>	<input type="button" value="Role type ▾"/>
<input type="checkbox"/>	Rutvik_redshift_S3	Not applied	--

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

Viewing all the data in Amazon S3 bucket:



The screenshot shows the AWS S3 console interface for the bucket 'rutvik-redshift-s3123'. The 'Objects' tab is selected. There are 5 objects listed:

Name	Type	Last modified	Size	Storage class
dim_atm/	Folder	-	-	-
dim_card_type/	Folder	-	-	-
dim_date/	Folder	-	-	-
dim_location/	Folder	-	-	-
fact_atm_trans/	Folder	-	-	-

Query to create a schema for the dimension and fact tables:

```
create schema atm_data;
```

```

1 create schema atm_data;
2
3 --Creating location dimension table
4 create table atm_data.DIM_LOCATION (
5 location_id int not null DISTKEY SORTKEY, location varchar(50),
6 streetname varchar(255),
7 street_number int,
8 zipcode int,
9 lat decimal(10,3),
10 lon decimal(10,3), PRIMARY KEY(location_id)
11 );
12
13 --Creating atm dimension table
14 create table atm_data.DIM_ATM (
15 atm_id int not null DISTKEY SORTKEY, atm_number varchar(20), atm_manufacturer varchar(50), atm_location_id int,

```

Query results Table details

Query

🕒 Completed, started on October 12, 2022 at 01:17:54
 ELAPSED TIME: 00 m 02 s

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

Creating location dimension table

`create table atm_data.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)`

`);`

```

3 --Creating location dimension table
4 create table atm_data.DIM_LOCATION (
5 location_id int not null DISTKEY SORTKEY, location varchar(50),
6 streetname varchar(255),
7 street_number int,
8 zipcode int,
9 lat decimal(10,3),
10 lon decimal(10,3), PRIMARY KEY(location_id)
11 );
12
13 --Creating atm dimension table
14 create table atm_data.DIM_ATM (
15 atm_id int not null DISTKEY SORTKEY, atm_number varchar(20), atm_manufacturer varchar(50), atm_location_id int,
16 );

```

[Run](#)
[Save](#)
[Schedule](#)
[Clear](#)
[Send feedback](#)
[Query results](#)
[Table details](#)
[Query](#)
[Execution](#)
[Data](#)
[Visualize](#)

⌚ Completed, started on October 12, 2022 at 01:19:10
 ELAPSED TIME: 00 m 03 s

• Creating atm dimension table

```

create table atm_data.DIM_ATM (
atm_id int not null DISTKEY SORTKEY, atm_number varchar(20),
atm_manufacturer varchar(50), atm_location_id int,
PRIMARY KEY(atm_id),
FOREIGN KEY(atm_location_id) references
atm_data.DIM_LOCATION(location_id ) );

```

```

--Creating atm dimension table
create table atm_data.DIM_ATM (
atm_id int not null DISTKEY SORTKEY, atm_number varchar(20), atm_manufacturer varchar(50), atm_location_id int,
PRIMARY KEY(atm_id),
FOREIGN KEY(atm_location_id) references atm_data.DIM_LOCATION(location_id ) );

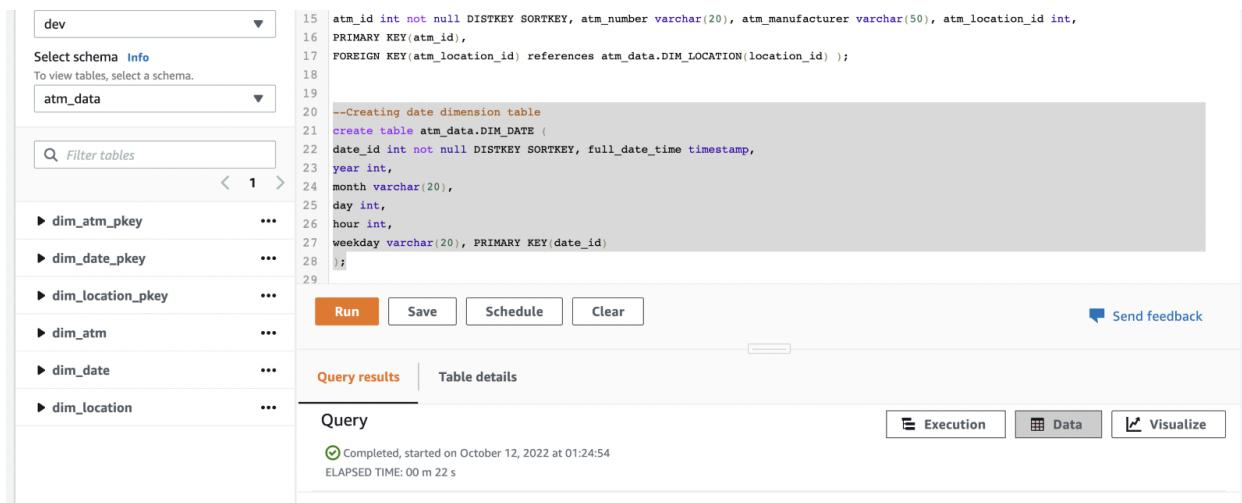
```

[Run](#)
[Save](#)
[Schedule](#)
[Clear](#)
[Send feedback](#)
[Query results](#)
[Table details](#)
[Query](#)
[Execution](#)
[Data](#)
[Visualize](#)

⌚ Completed, started on October 12, 2022 at 01:22:21

Creating date dimension table

```
create table atm_data.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)
);
```



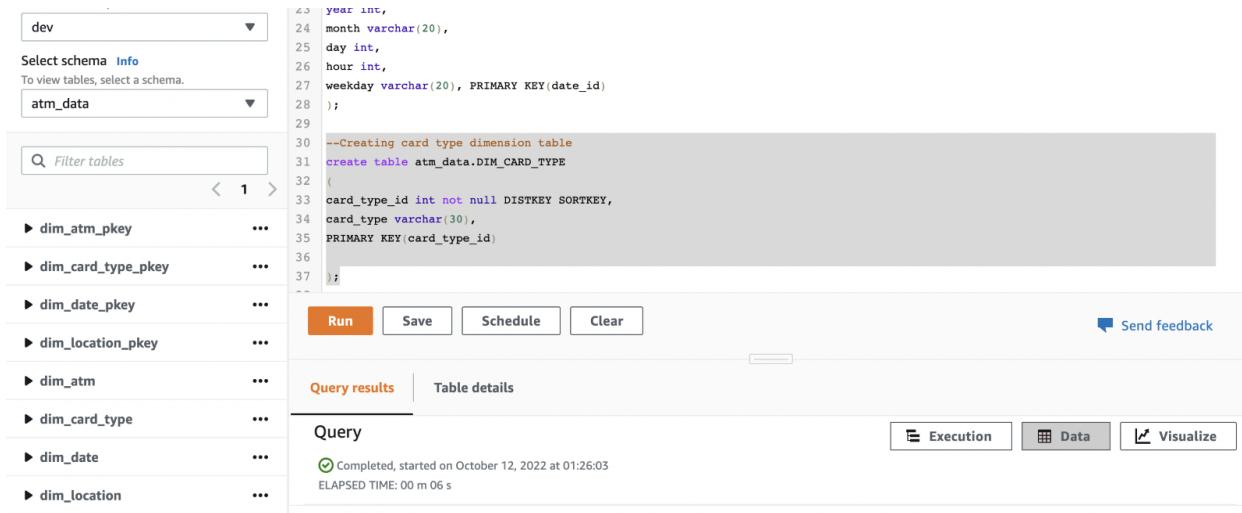
The screenshot shows a database interface with the following details:

- Schemas:** dev (selected), atm_data (selected).
- Tables:** dim_atm_pkey, dim_date_pkey, dim_location_pkey, dim_atm, dim_date, dim_location.
- Query Editor:**
 - Code area (lines 15-29):


```
15 | atm_id int not null DISTKEY SORTKEY, atm_number varchar(20), atm_manufacturer varchar(50), atm_location_id int,
16 | PRIMARY KEY(atm_id),
17 | FOREIGN KEY(atm_location_id) references atm_data.DIM_LOCATION(location_id) );
18 |
19 |
20 |--Creating date dimension table
21 |create table atm_data.DIM_DATE (
22 |date_id int not null DISTKEY SORTKEY, full_date_time timestamp,
23 |year int,
24 |month varchar(20),
25 |day int,
26 |hour int,
27 |weekday varchar(20), PRIMARY KEY(date_id)
28 |);
29 |;
```
 - Buttons: Run, Save, Schedule, Clear, Send feedback.
 - Status: Completed, started on October 12, 2022 at 01:24:54, ELAPSED TIME: 00 m 22 s.
 - Execution, Data, Visualize tabs.

• Creating card type dimension table

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



The screenshot shows a database interface with the schema selected as 'dev'. A query is being run to create a dimension table named 'DIM_CARD_TYPE' with columns for year, month, day, hour, and weekday. The table has a primary key on 'date_id'. Below the code, a list of tables is shown, including 'dim_atm_pkey', 'dim_card_type_pkey', 'dim_date_pkey', 'dim_location_pkey', 'dim_atm', 'dim_card_type', 'dim_date', and 'dim_location'. At the bottom, there are tabs for 'Query results' and 'Table details', and buttons for 'Run', 'Save', 'Schedule', and 'Clear'. A 'Send feedback' button is also present.

```

40     year int,
41     month varchar(20),
42     day int,
43     hour int,
44     weekday varchar(20), PRIMARY KEY(date_id)
45   );
46
47 --Creating card type dimension table
48 create table atm_data.DIM_CARD_TYPE
49 (
50   card_type_id int not null DISTKEY SORTKEY,
51   card_type varchar(30),
52   PRIMARY KEY(card_type_id)
53 );
54
55
    
```

• Creating atm transactions fact table

```
create table atm_data.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
atm_data.DIM_LOCATION(location_id), FOREIGN KEY(atm_id)
references atm_data.DIM_DATA(atm_id),
FOREIGN KEY(date_id) references atm_data.DIM_DATE(date_id),
FOREIGN KEY(card_type_id) references
atm_data.DIM_CARD_TYPE(card_type_id)
    
```

);

```
47     service varchar(2),
48     transaction_amount int,
49     message_code varchar(225),
50     message_text varchar(225),
51     rain_3h decimal(10,3),
52     clouds_all int,
53     weather_id int,
54     weather_main varchar(50),
55     weather_description varchar(255),
56 PRIMARY KEY(trans_id),
57 FOREIGN KEY(weather_loc_id) references atm_data.DIM_LOCATION(location_id),
58 FOREIGN KEY(atm_id) references atm_data.DIM_ATM(atm_id),
59 FOREIGN KEY(date_id) references atm_data.DIM_DATE(date_id),
60 FOREIGN KEY(card_type_id) references atm_data.DIM_CARD_TYPE(card_type_id)
61 );
```

Run

Save

Schedule

Clear

 Send feedback

Query results

Table details

Query

Completed, started on October 12, 2022 at 01:29:54
ELAPSED TIME: 01 m 00 s

E Execution

Data

| Visualize

Loading data into a Redshift cluster from Amazon S3 bucket

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

Copying the data to dim_location table

```
copy atm_data.dim_location from 's3://rutvik-redshift-s3123/dim_location/part-00000-69166982-23bf-4036-bc75-3dda3d031461-c000.csv'
```

iam_role 'arn:aws:iam::814423138586:role/Rutvik_redshift_S3'

delimiter ',' region 'us-east-1'

CSV;

```

64 --Loading data into a RedShift cluster from Amazon S3 bucket
65 --Copying the data to dim_location table
66
67
68 copy atm_data.dim_location from 's3://rutvik-redshift-s3123/dim_location/part-00000-69166982-23bf-4036-bc75-3dda3d031461-c000.csv'
69 iam_role 'arn:aws:iam::814423138586:role/Rutvik_redshift_S3'
70 delimiter ',' region 'us-east-1'
71 CSV;

```

Query results Table details

Query 336 

✔ Completed, started on October 12, 2022 at 01:31:06
 ELAPSED TIME: 00 m 03 s

• Copying the data to dim_atm table

--Copying the data to dim_atm table

```
copy atm_data.dim_atm from 's3://rutvik-redshift-s3123/dim_atm/part-00000-
ba7f7d65-2e20-425c-a5df-8092530ff97d-c000.csv'
```

```
iam_role 'arn:aws:iam::814423138586:role/Rutvik_redshift_S3'
```

```
delimiter ',' region 'us-east-1'
```

CSV;

```

73 --Copying the data to dim_atm table
74 copy atm_data.dim_atm from 's3://rutvik-redshift-s3123/dim_atm/part-00000-ba7f7d65-2e20-425c-a5df-8092530ff97d-c000.csv'
75 iam_role 'arn:aws:iam::814423138586:role/Rutvik_redshift_S3'
76 delimiter ',' region 'us-east-1'
77 CSV;
78

```

Query results Table details

Query 386 

✔ Completed, started on October 12, 2022 at 01:33:32
 ELAPSED TIME: 00 m 08 s

Copying the data to dim_date table

--Copying the data to dim_date table

```
copy atm_data.dim_date from 's3://rutvik-redshift-s3123/dim_date/
part-00000-8fe4c4e5-974f-4927-a830-5c52b49174bd-c000.csv'
```

```
iam_role 'arn:aws:iam::814423138586:role/Rutvik_redshift_S3'
```

```
delimiter ',' region 'us-east-1'
```

CSV;

```

79 --Copying the data to dim_date table
80 copy atm_data.dim_date from 's3://rutvik-redshift-s3123/dim_date/part-00000-8fe4c4e5-974f-4927-a830-5c52b49174bd-c000.csv'
81 iam_role 'arn:aws:iam::814423138586:role/Rutvik_redshift_S3'
82 delimiter ',' region 'us-east-1'
83 CSV;

```

Send feedback

Query results Table details

Query 1601

 Execution Data

- **Copying the data to dim_card_type table**

```
copy atm_data.dim_card_type from 's3://rutvik-redshift-s3123/
dim_card_type/part-00000-731499a7-a568-4d65-8ead-375cb3a84936-
c000.csv'
```

```
iam_role 'arn:aws:iam::814423138586:role/Rutvik_redshift_S3'
```

```
delimiter ',' region 'us-east-1'
```

CSV;

```

85 --Copying the data to dim_card_type table
86 copy atm_data.dim_card_type from 's3://rutvik-redshift-s3123/dim_card_type/part-00000-731499a7-a568-4d65-8ead-375cb3a84936-c000.csv'
87 iam_role 'arn:aws:iam::814423138586:role/Rutvik_redshift_S3'
88 delimiter ',' region 'us-east-1'
89 CSV;

```

Send feedback

Query results Table details

Query 491

 Execution Data Visualize

Completed, started on October 12, 2022 at 01:38:26
ELAPSED TIME: 00 m 03 s

- **Copying the data to fact_atm_trans table**

```
copy atm_data.fact_atm_trans from 's3://rutvik-redshift-s3123/fact_atm_trans/
part-00000-5cd3a475-eb4b-4a3b-a512-b16dbede416c-c000.csv'
iam_role 'arn:aws:iam::814423138586:role/Rutvik_redshift_S3'
delimiter ',' region 'us-east-1'
CSV;
```

```
91 --Copying the data to fact_atm_trans table
92 copy atm_data.fact_atm_trans from 's3://rutvik-redshift-s3123/fact_atm_trans/part-00000-195f78ad-fed2-4634-bf06-427b917alb44-c000.csv'
93 iam_role 'arn:aws:iam::814423138586:role/Rutvik_redshift_S3'
94 delimiter ',' region 'us-east-1'
95 csv;
```

Run **Save** **Schedule** **Clear** **Send feedback**

Query results **Table details**

Query 541 

Completed, started on October 12, 2022 at 01:40:46
ELAPSED TIME: 00 m 14 s

Execution **Data** **Visualize**