

Creation of a Redshift Cluster

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

Amazon Redshift query editor v2 is now available
Query editor v2 provides new features such as multistatement query execution, query parameterization, query versioning, visualizations, and query sharing. [Learn more about query editor v2](#)

Go to query editor v2

Amazon Redshift > Clusters > redshift-cluster-atm

redshift-cluster-atm

Actions Edit Add partner integration Query data

General information

Cluster identifier redshift-cluster-atm	Status Available	Node type dc2.large	Endpoint redshift-cluster-atm.c34p2piwra4q.us-east-1.redshift.amazonaws.com:4321/upgrad
Custom domain name -	Date created May 06, 2024, 22:13 (UTC+05:30)	Number of nodes 2	JDBC URL jdbc:redshift://redshift-cluster-atm.c34p2piwra4q.us-east-1.redshift.amazonaws.com:4321/upgrad
Cluster namespace ARN arn:aws:redshift:us-east-1:533267238611:namespace:7b87f9b0-c7a6-451e-b9e0-052362e77275	Storage used -		ODBC URL Driver=(Amazon Redshift (v64)); Server=redshift-cluster-atm.c34p2piwra4q.us-east-1.redshift.amazonaws.com; Database=upgrad
Cluster configuration Production	Multi-AZ No		

Cluster performance Query monitoring Zero-ETL integrations Resource Policy Schedules Maintenance Properties

Recommendations (0)
To improve performance and decrease operating costs, the Amazon Redshift Advisor provides recommendations.

Cluster performance Query monitoring Zero-ETL integrations Resource Policy Schedules Maintenance Properties

Database configurations

Database name upgrad	Parameter group default.redshift-1.0	Encryption Disabled	Audit logging Disabled
Port 4321	SSH ingestion setting (cluster public key) ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAQBAQC1KD LwvPhzB5SHUWZNOptYukP5VHOMc43UY7Jpc 8u3OPBynfGVRItdCHeq5ITy2GGOORv0Jg7k w3VinevjpZJHRM69AikPvkuBUOWw8X8Vde+L uHUnb4c6eerUa16/rSkyOMENPopG6vTILZWd AHQUMHgd5oEuzPzEFC0w/7v/Rw21QIL5k29N 6/s6J/OYTfioBkczq2MK/4vRd91tHZSnp1brU tsNTqomDIWxrl+NlnJxodWbC+C6cE/dVF+Nhl CsjaE+UXBOEuqPYX5jht9TU1vfwMKmf4q8 Othf/TIL4M47LMixAL1mAbHJGhA7EwtY9a0K	AWS KMS key ID -	
Admin user name awsuser			

Edit admin credentials Rotate encryption keys Edit

Network and security settings

info

Edit

Virtual private cloud (VPC)
vpc-0abd57d9f2729673e

Subnet group
cluster-subnet-group-1

Endpoint URL
-

Availability Zone
us-east-1e

Enhanced VPC routing
Disabled

VPC security group
Specify which instances and devices can connect to the cluster.
sg-02d82e68046e55676

IP address type
-

Publicly accessible
Allow connections from outside the VPC.
Disabled

Cluster permissions

Create an IAM role as the default for this cluster that has the [AmazonRedshiftAllCommandsFullAccess](#) policy attached. This policy includes permissions to run SQL commands to COPY, UNLOAD, and query data with Amazon Redshift. The policy also grants permissions to run SELECT statements for related services, such as Amazon S3, Amazon CloudWatch logs, Amazon SageMaker, and AWS Glue.

Associated IAM roles (1)

Info

Set default

Manage IAM roles

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.

Q

Search for associated IAM role by name, status, or role type

< 1 >

☐

IAM roles

▼

☐

myRedshiftRole

in-sync

--

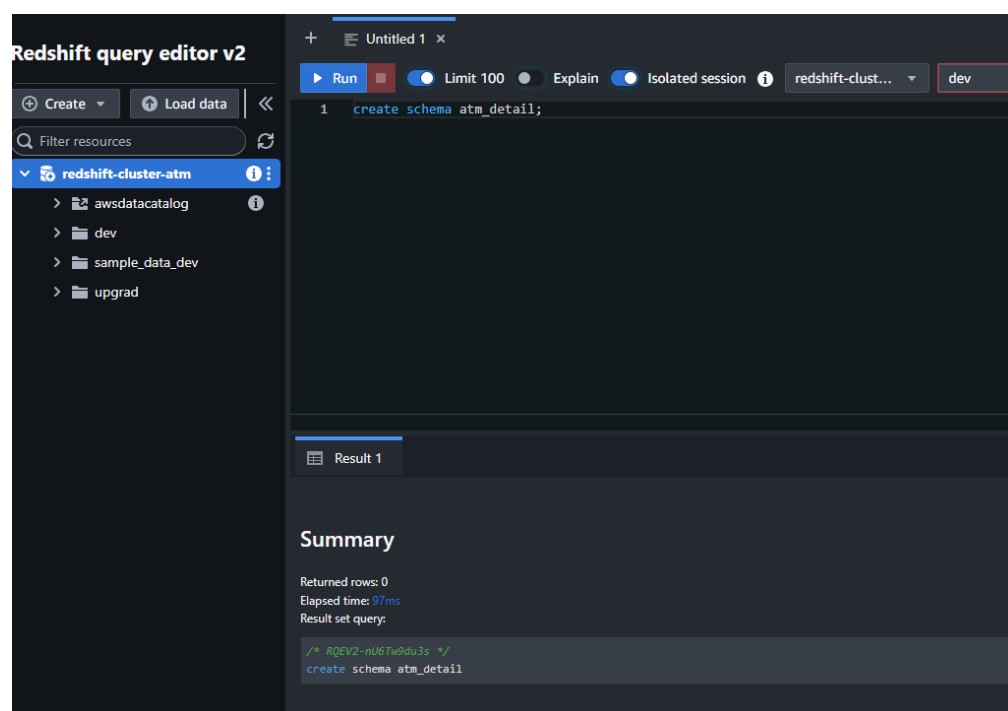
© Copyright. upGrad Education Pvt. Ltd. All rights reserved

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:

- **Create schema:**

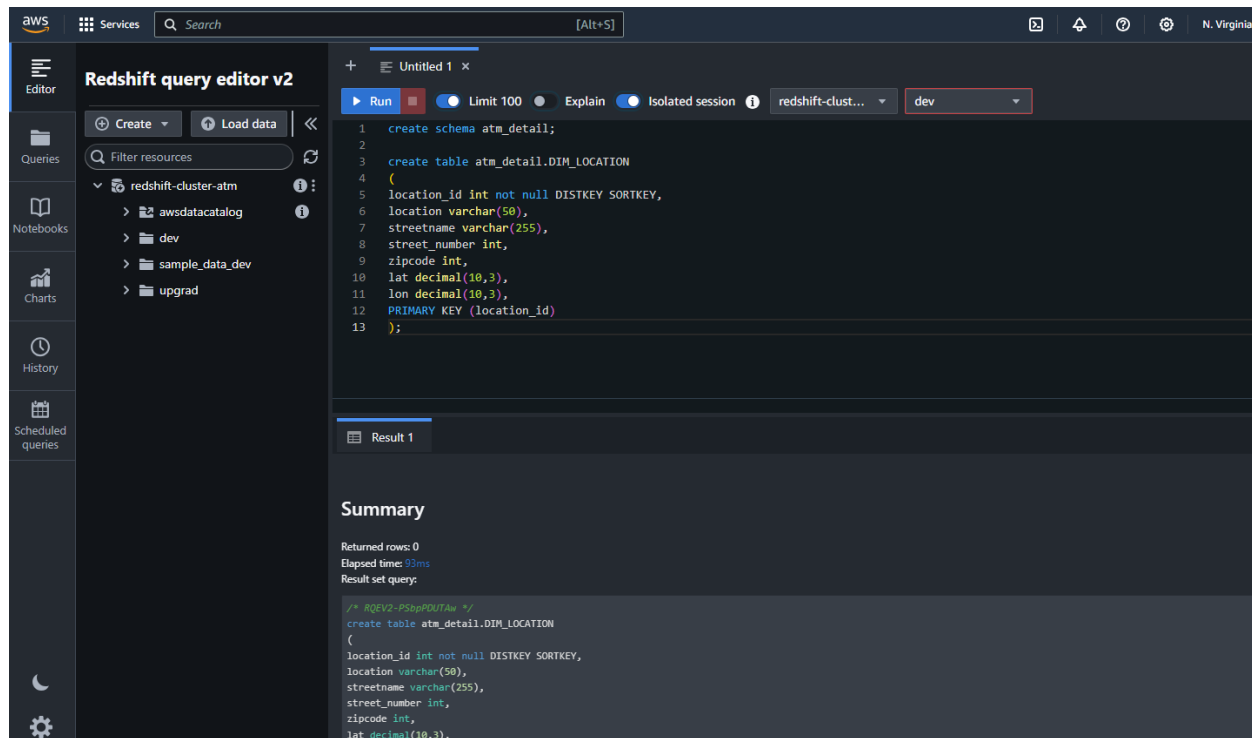
Query: create schema atm_detail;



- **Create DIM_LOCATION table on Redshift:**

Query:

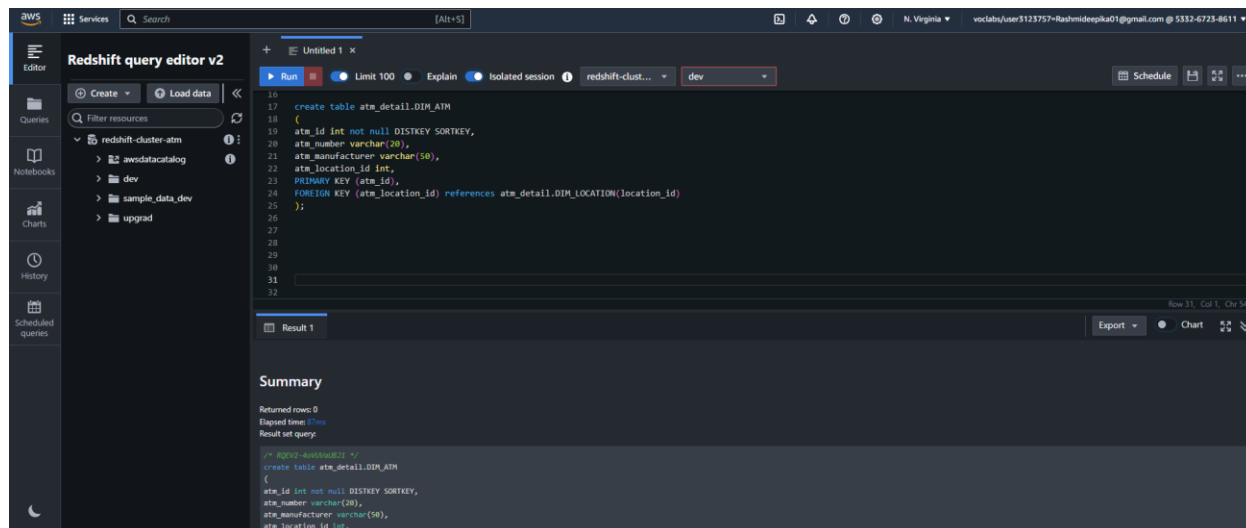
```
create table atm_detail.DIM_LOCATION
(
location_id int not null DISTKEY SORTKEY,
location varchar(50),
streetname varchar(255),
street_number int,
zipcode int,
lat decimal(10,3),
lon decimal(10,3),
PRIMARY KEY (location_id)
);
```



- **Create DIM_ATM table on Redshift:**

Query:

```
create table atm_detail.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_detail.DIM_LOCATION(location_id)
);
```



The screenshot shows the AWS Redshift Query Editor v2 interface. The query editor is open with a new query titled 'Untitled 1'. The query is as follows:

```

16
17 create table atm_detail.DIM_ATM
18 (
19   atm_id int not null DISTKEY SORTKEY,
20   atm_number varchar(20),
21   atm_manufacturer varchar(50),
22   atm_location_id int,
23   PRIMARY KEY (atm_id),
24   FOREIGN KEY (atm_location_id) references atm_detail.DIM_LOCATION(location_id)
25 );
26
27
28
29
30
31
32

```

The 'Summary' section shows that the query returned 0 rows, elapsed time was 0.0ms, and the result set query is the same as the query entered in the editor.

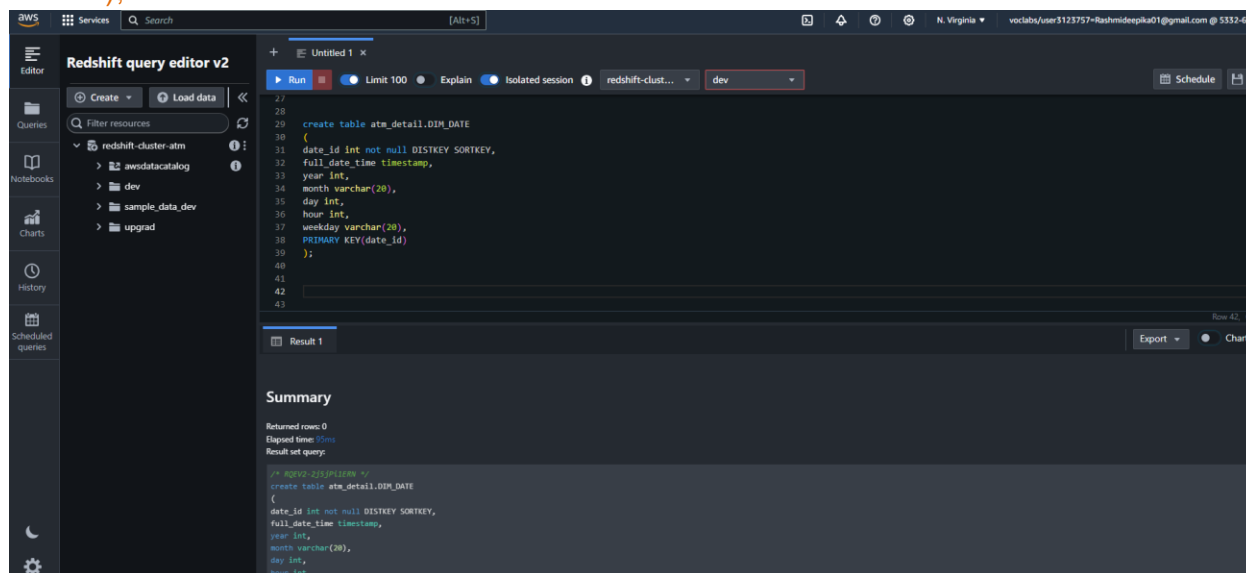
- **Create DIM_DATE on Redshift:**

Query:

```

create table atm_detail.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 the AWS Redshift Query Editor v2 interface. The query editor is open with a new query titled 'Untitled 1'. The query is as follows:

```

27
28
29 create table atm_detail.DIM_DATE
30 (
31   date_id int not null DISTKEY SORTKEY,
32   full_date_time timestamp,
33   year int,
34   month varchar(20),
35   day int,
36   hour int,
37   weekday varchar(20),
38   PRIMARY KEY(date_id)
39 );
40
41
42
43

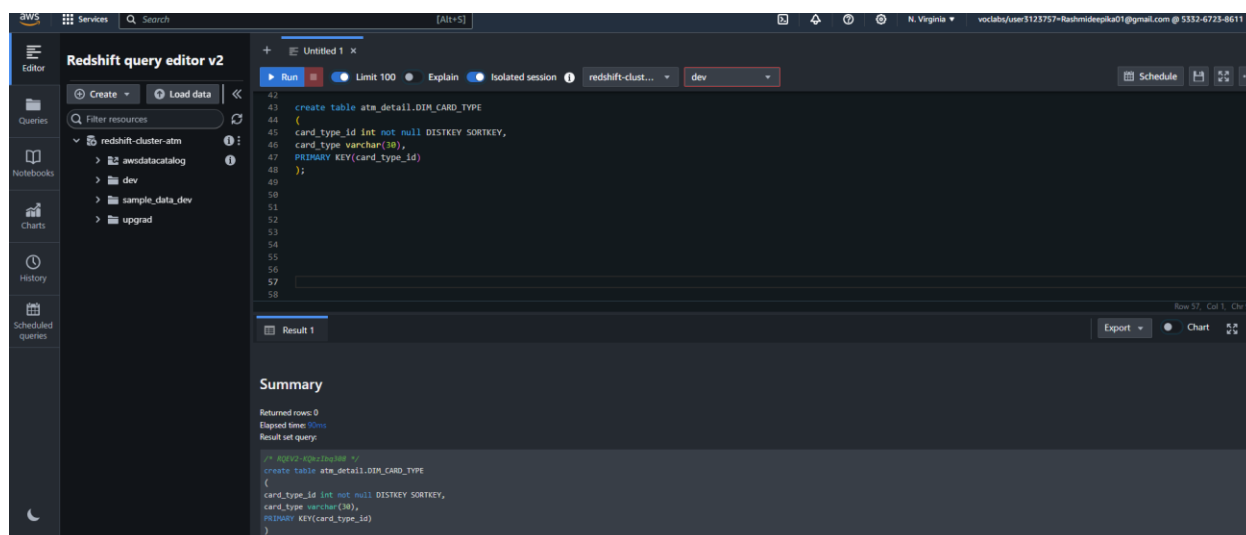
```

The 'Summary' section shows that the query returned 0 rows, elapsed time was 0.0ms, and the result set query is the same as the query entered in the editor.

- **Create DIM_CARD_TYPE on Redshift:**

Query:

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

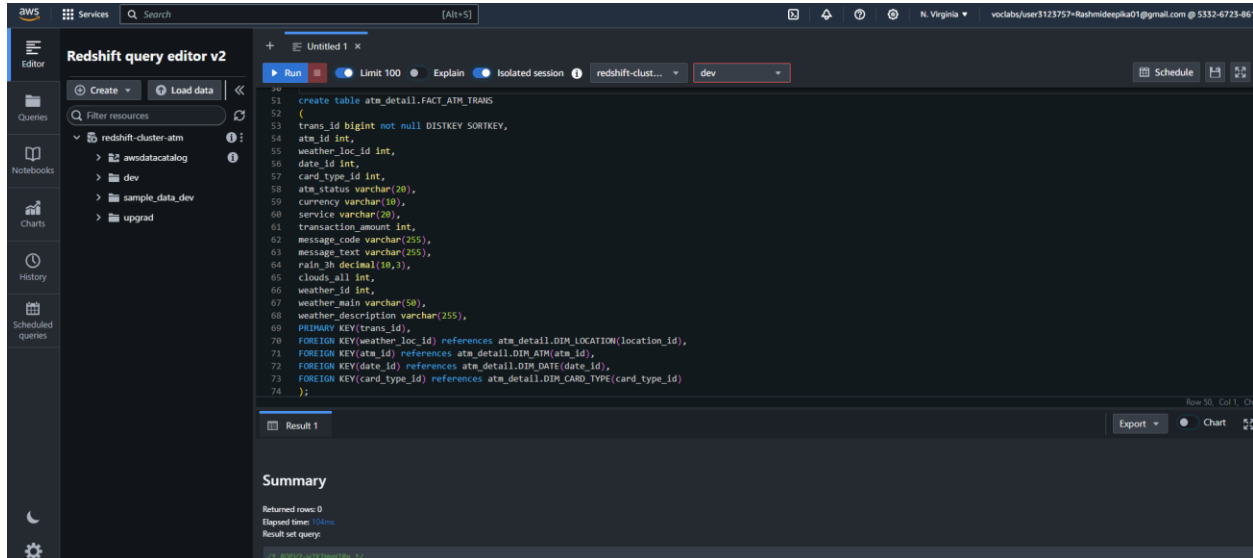


- **Create FACT_ATM_TRANS on Redshift:**

Query:

```
create table atm_detail.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(255),
message_text varchar(255),
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_detail.DIM_LOCATION(location_id),
FOREIGN KEY(atm_id) references atm_detail.DIM_ATM(atm_id),
FOREIGN KEY(date_id) references atm_detail.DIM_DATE(date_id),
FOREIGN KEY(card_type_id) references atm_detail.DIM_CARD_TYPE(card_type_id)
);
```



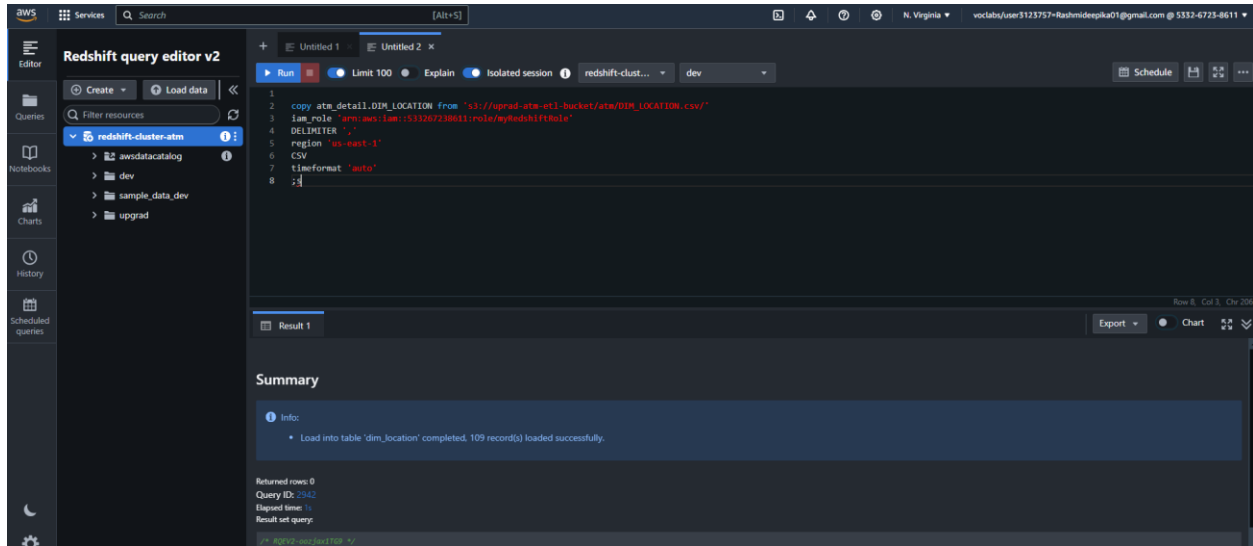
Loading data into a Redshift cluster from Amazon S3 bucket

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

- Load Data to DIM_LOCATION:

Query:

```
copy atm_detail.DIM_LOCATION from 's3://uprad-atm-etl-
bucket/atm/DIM_LOCATION.csv'
iam_role 'arn:aws:iam::533267238611:role/myRedshiftRole'
DELIMITER ','
region 'us-east-1'
CSV
timeformat 'auto'
;
```



```

1 copy atm_detail.DIM_LOCATION from 's3://uprad-atm-etl-bucket/atm/DIM_LOCATION.csv'
2 iam_role 'arn:aws:iam::533267238611:role/myRedshiftRole'
3 DELIMITER ','
4 region 'us-east-1'
5 CSV
6 timeformat 'auto'
7 ;

```

Summary

Info:

- Load into table 'dim_location' completed. 109 record(s) loaded successfully.

Returned rows: 0
Query ID: 2942
Elapsed time:
Result set query:

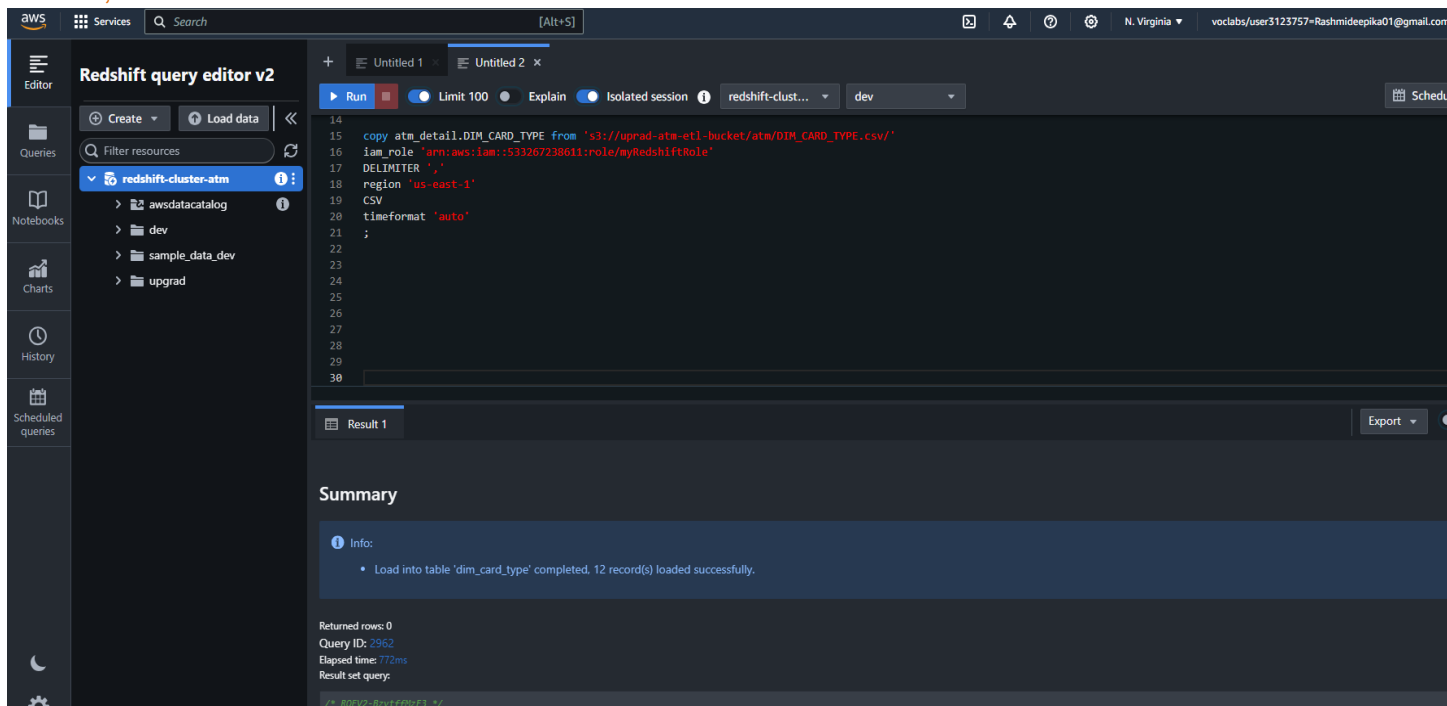
- **Load Data to DIM_CARD_TYPE:**

Query:

```

copy atm_detail.DIM_CARD_TYPE from 's3://uprad-atm-etl-
bucket/atm/DIM_CARD_TYPE.csv/'
iam_role 'arn:aws:iam::533267238611:role/myRedshiftRole'
DELIMITER ','
region 'us-east-1'
CSV
timeformat 'auto'
;

```



```

14
15 copy atm_detail.DIM_CARD_TYPE from 's3://uprad-atm-etl-bucket/atm/DIM_CARD_TYPE.csv/'
16 iam_role 'arn:aws:iam::533267238611:role/myRedshiftRole'
17 DELIMITER ','
18 region 'us-east-1'
19 CSV
20 timeformat 'auto'
21 ;
22
23
24
25
26
27
28
29
30

```

Summary

Info:

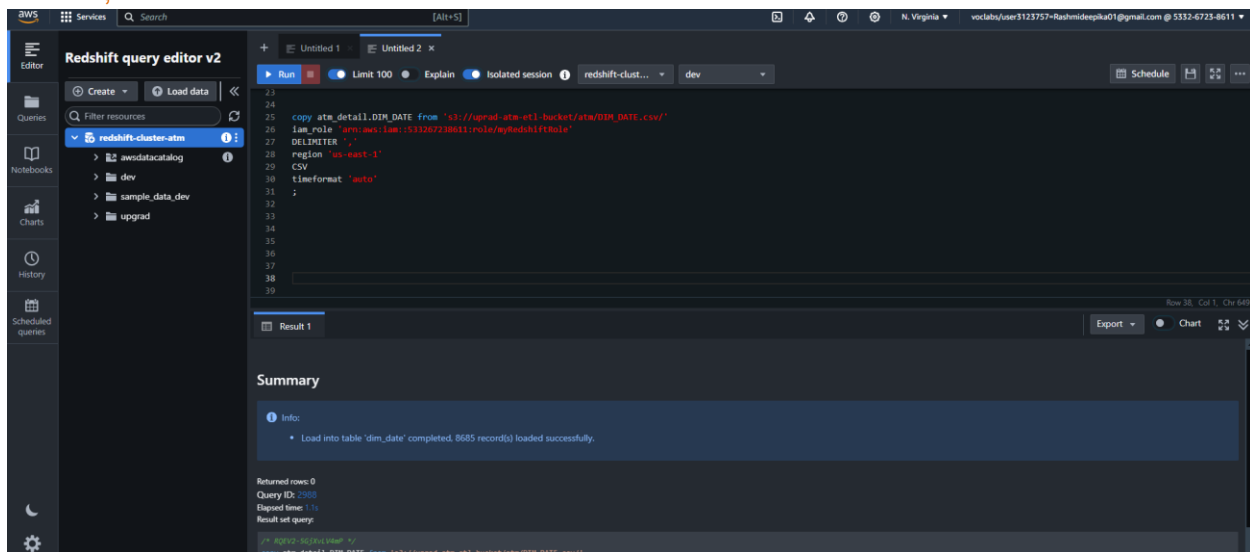
- Load into table 'dim_card_type' completed. 12 record(s) loaded successfully.

Returned rows: 0
Query ID: 2962
Elapsed time: 772ms
Result set query:

- **Load Data to DIM_DATE:**

Query:

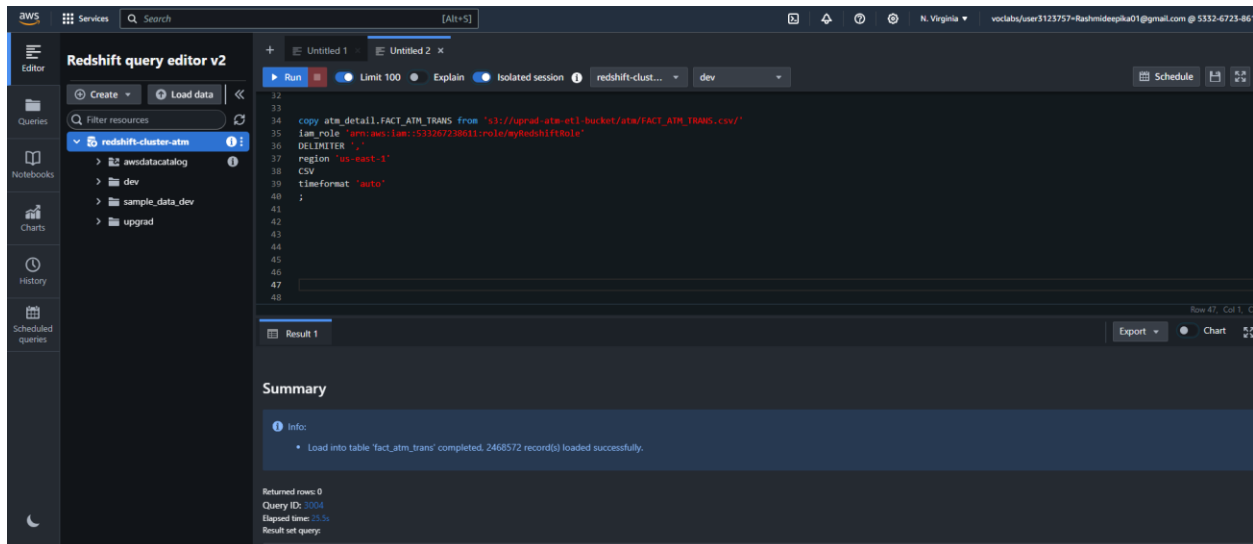
```
copy atm_detail.DIM_DATE from 's3://uprad-atm-etl-bucket/atm/DIM_DATE.csv'
iam_role 'arn:aws:iam::533267238611:role/myRedshiftRole'
DELIMITER ','
region 'us-east-1'
CSV
timeformat 'auto'
;
```



- **Load Data to FACT_ATM_TRANS:**

Query:

```
copy atm_detail.FACT_ATM_TRANS from 's3://uprad-atm-etl-
bucket/atm/FACT_ATM_TRANS.csv'
iam_role 'arn:aws:iam::533267238611:role/myRedshiftRole'
DELIMITER ','
region 'us-east-1'
CSV
timeformat 'auto'
;
```



```

32
33
34 copy_atm_detail.FACT_ATM_TRANS from 's3://upgrad-atm-etl-bucket/atm/FACT_ATM_TRANS.csv/'
35 iam_role 'arn:aws:iam::533267238611:role/myRedshiftRole'
36 DELIMITER ','
37 region 'us-east-1'
38 CSV
39 timeformat 'auto'
40 ;
41
42
43
44
45
46
47
48

```

Summary

Info:

- Load into table 'fact_atm_trans' completed. 2468572 record(s) loaded successfully.

Returned rows: 0
Query ID: 3104
Elapsed time: 25.5s
Result set query:

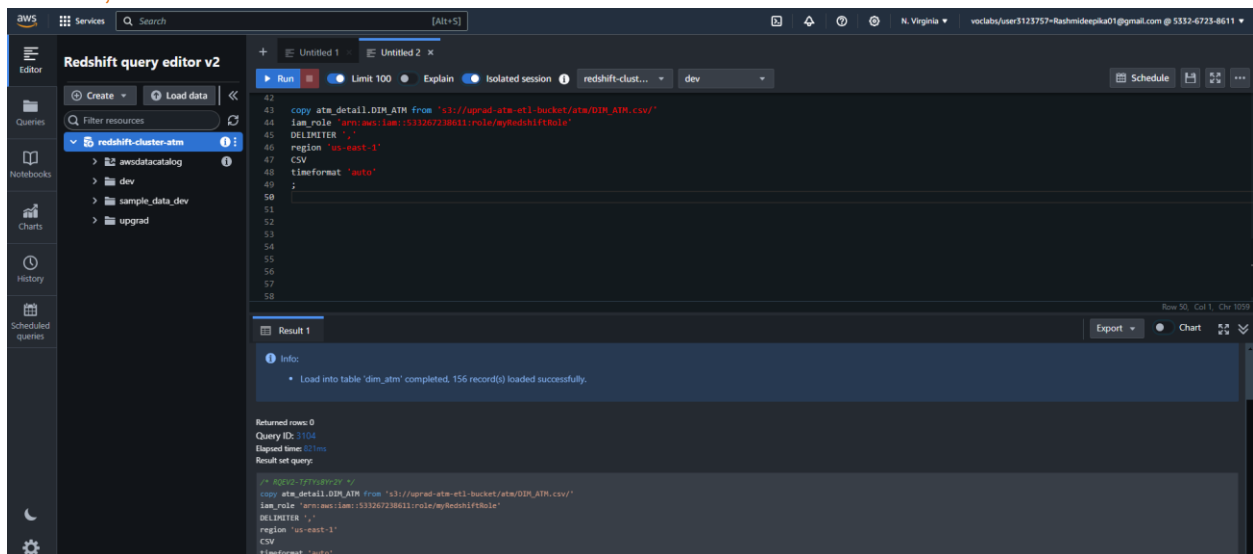
• Load Data to DIM_ATM

Query:

```

copy atm_detail.DIM_ATM from 's3://upgrad-atm-etl-bucket/atm/DIM_ATM.csv/'
iam_role 'arn:aws:iam::533267238611:role/myRedshiftRole'
DELIMITER ','
region 'us-east-1'
CSV
timeformat 'auto'
;

```



```

42
43 copy_atm_detail.DIM_ATM from 's3://upgrad-atm-etl-bucket/atm/DIM_ATM.csv/'
44 iam_role 'arn:aws:iam::533267238611:role/myRedshiftRole'
45 DELIMITER ','
46 region 'us-east-1'
47 CSV
48 timeformat 'auto'
49 ;
50
51
52
53
54
55
56
57
58

```

Summary

Info:

- Load into table 'dim_atm' completed. 156 record(s) loaded successfully.

Returned rows: 0
Query ID: 3104
Elapsed time: 22.1ms
Result set query:

```

-- upgrad-atm-etl-bucket --
copy_atm_detail.DIM_ATM from 's3://upgrad-atm-etl-bucket/atm/DIM_ATM.csv/'
iam_role 'arn:aws:iam::533267238611:role/myRedshiftRole'
DELIMITER ','
region 'us-east-1'
CSV
timeformat 'auto'

```