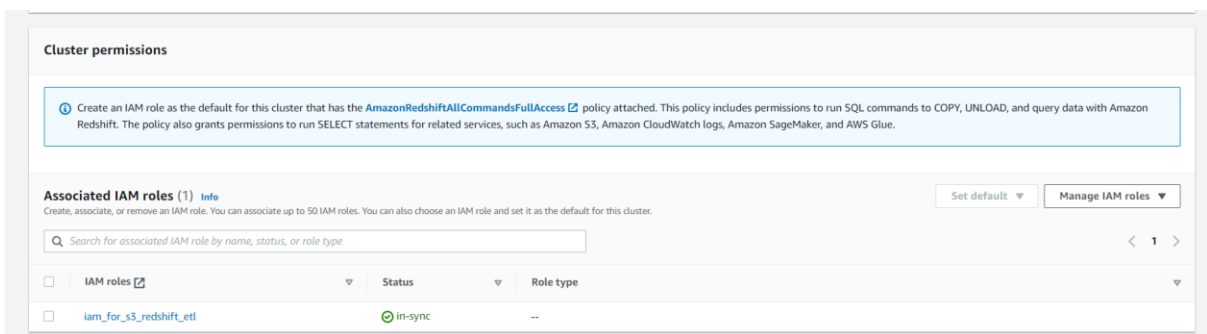
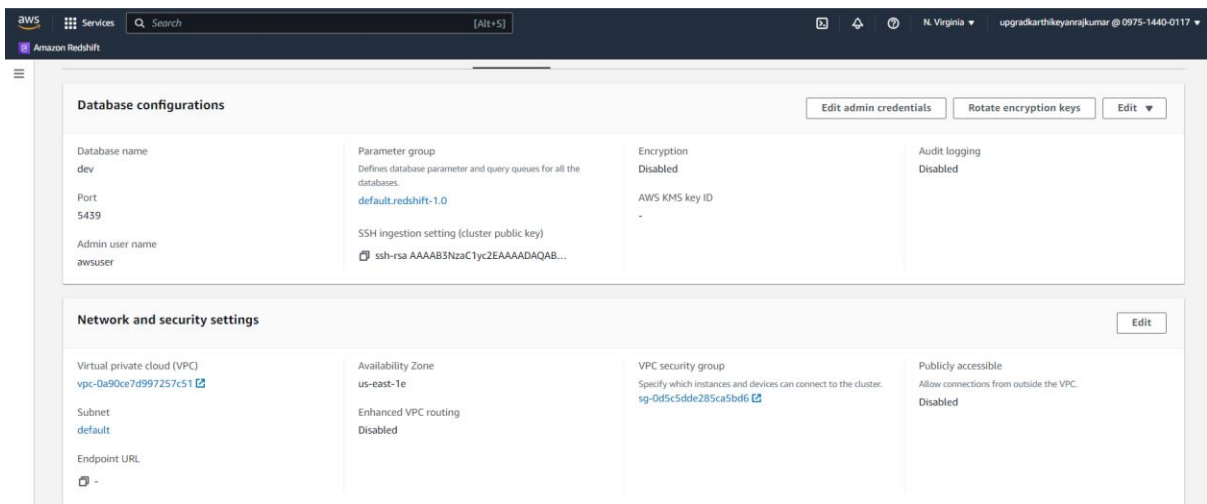
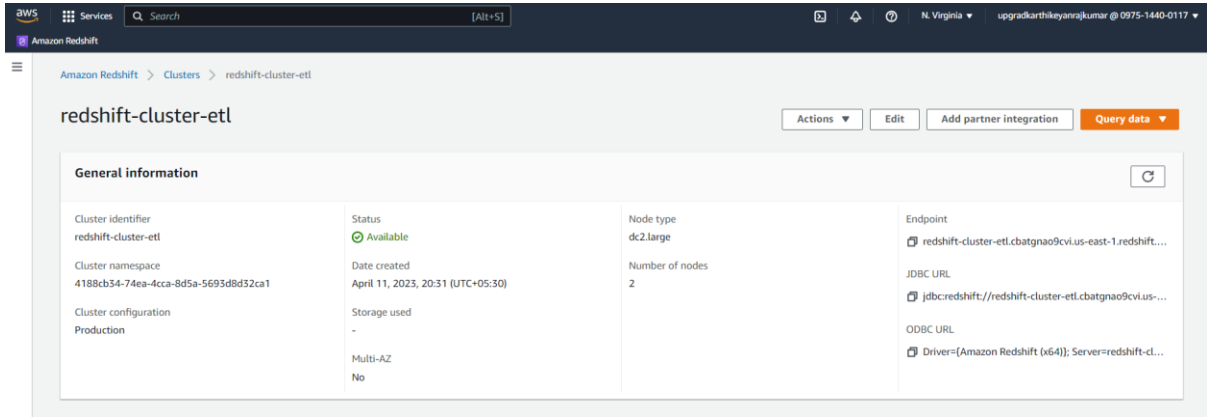


Creation of a Redshift Cluster

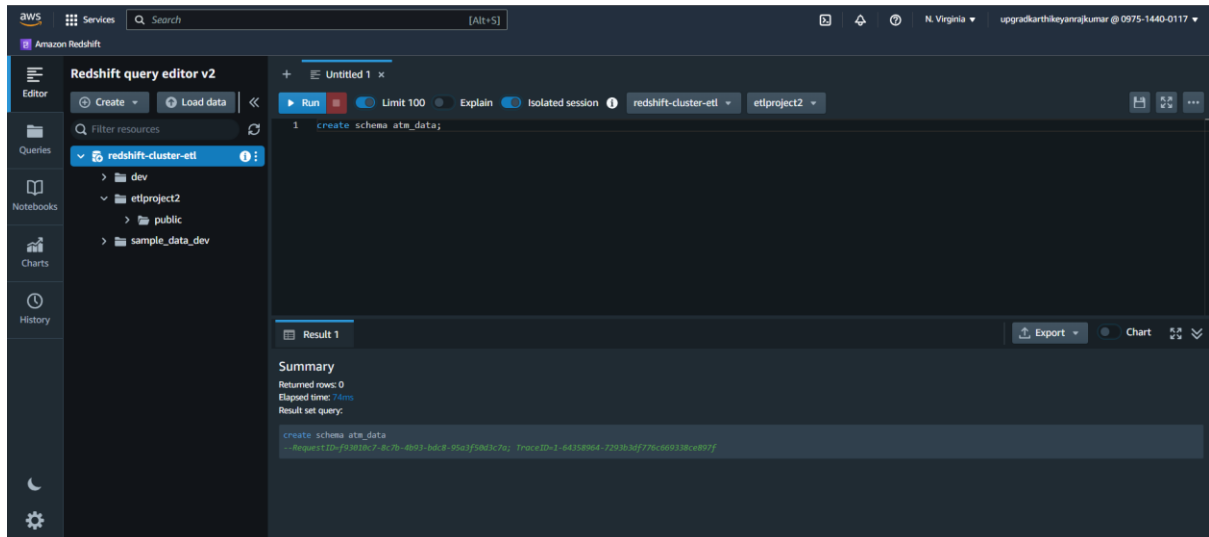
Screenshots of the configuration of the Redshift cluster that I have created:



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

Query for creating schema:

```
create schema atm_data;
```



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,
```

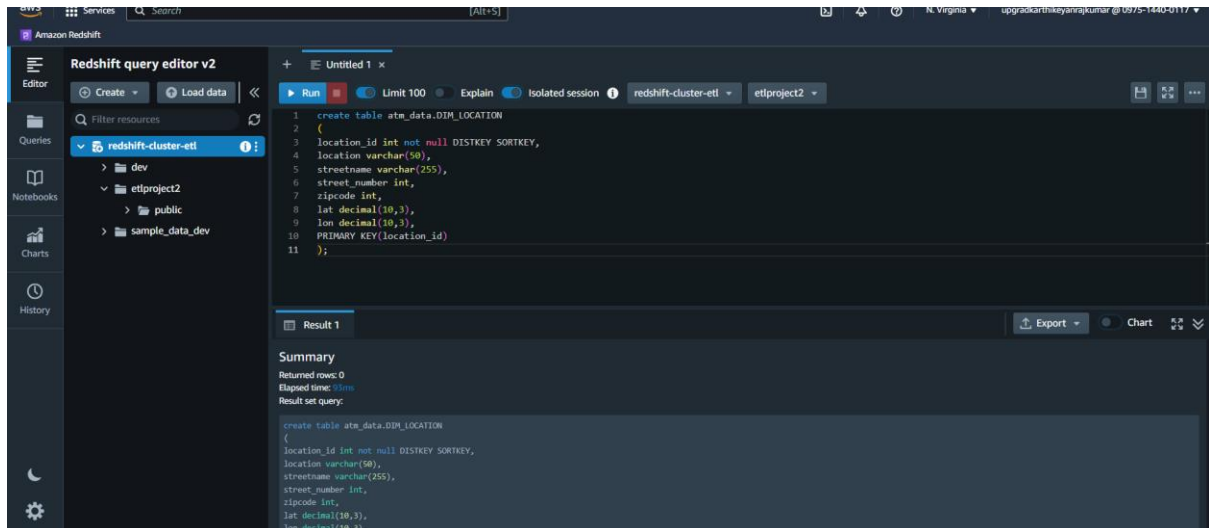
```
zipcode int,
```

```
lat decimal(10,3),
```

```
lon decimal(10,3),
```

PRIMARY KEY(location_id)

);



- **Creating atm dimension table:**

create table atm_data.DIM_ATM

(

atm_id int not null DISTKEY SORTKEY,

atm_number varchar(20),

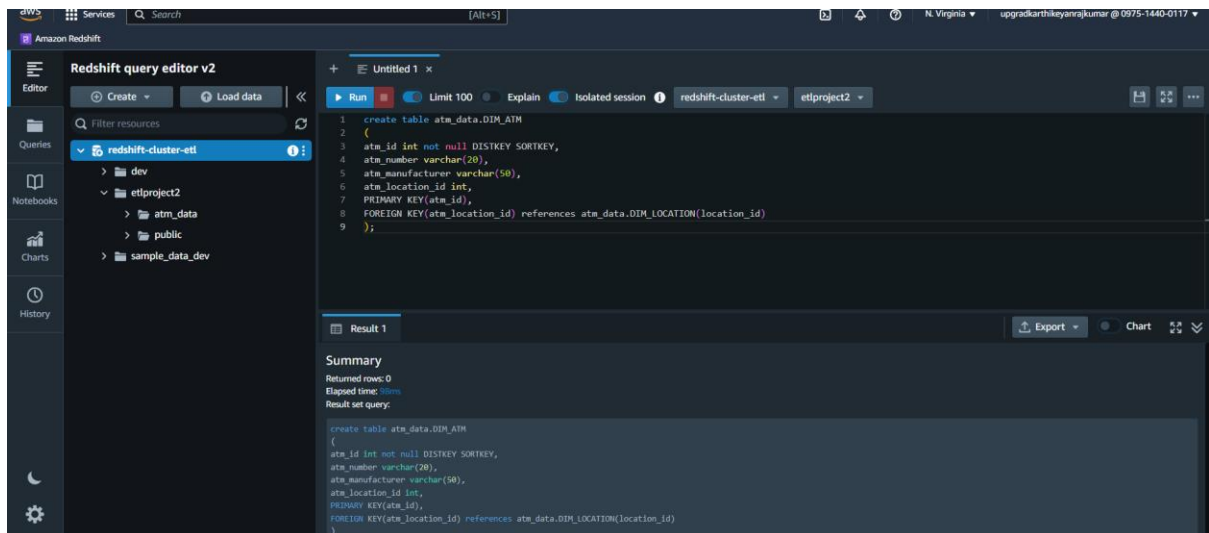
atm_manufacturer varchar(50),

atm_location_id int,

PRIMARY KEY(atm_id),

FOREIGN KEY(atm_location_id) references atm_data.DIM_LOCATION(location_id)

);



- **Creating 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),

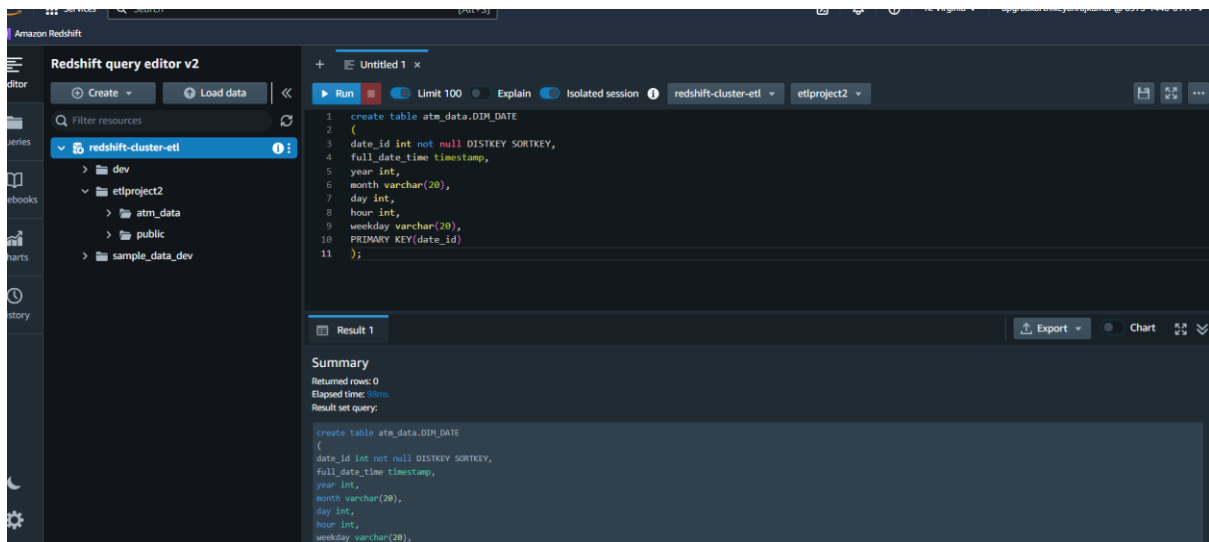
day int,

hour int,

weekday varchar(20),

PRIMARY KEY(date_id)

);



- **Creating card type dimension table:**

create table atm_data.DIM_CARD_TYPE

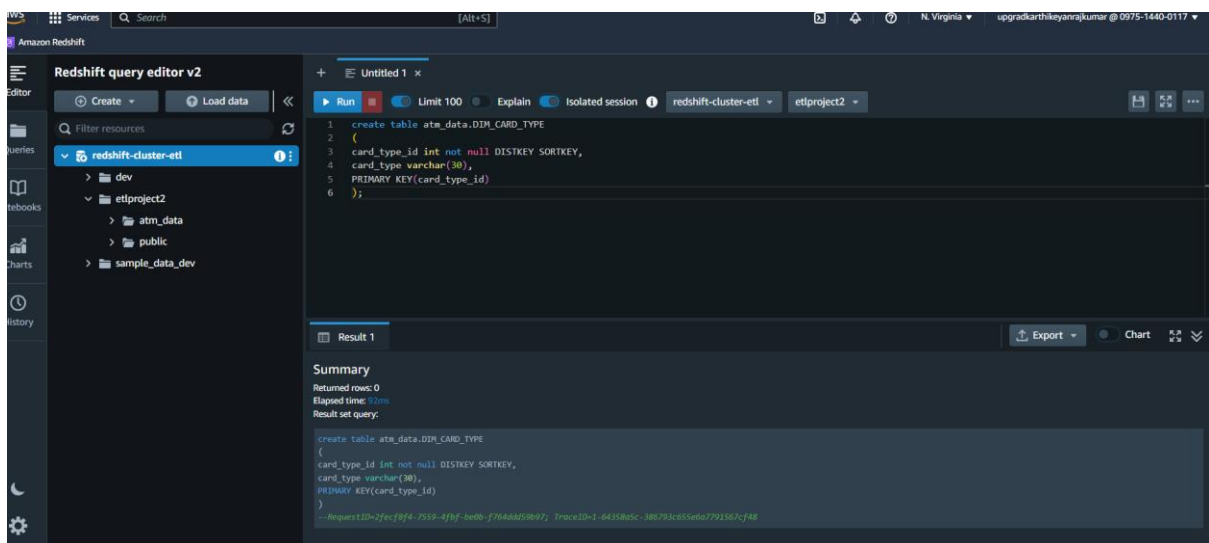
(

card_type_id int not null DISTKEY SORTKEY,

card_type varchar(30),

PRIMARY KEY(card_type_id)

);



- **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_ATM(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)
);
```

Amazon Redshift

Redshift query editor v2

Create Load data

Filter resources

redshift-cluster-etl

dev

etlproject2

atm_data

public

sample_data_dev

Run Limit 100 Explain Isolated session redshift-cluster-etl etlproject2

```
20 service varchar(40),
21 transaction_amount int,
22 message_code varchar(225),
23 message_text varchar(225),
24 rain_3h decimal(10,3),
25 clouds_all int,
26 weather_id int,
27 weather_main varchar(50),
28 weather_description varchar(255),
29 PRIMARY KEY(trans_id),
30 FOREIGN KEY(weather_loc_id) references atm_data.DIM_LOCATION(location_id),
31 FOREIGN KEY(atm_id) references atm_data.DIM_ATM(atm_id),
32 FOREIGN KEY(date_id) references atm_data.DIM_DATE(date_id),
33 FOREIGN KEY(card_type_id) references atm_data.DIM_CARD_TYPE(card_type_id)
34 );
```

Result 1

Export Chart

Summary

Returned rows: 0

Elapsed time: 0ms

Result set query:

```
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);
```

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

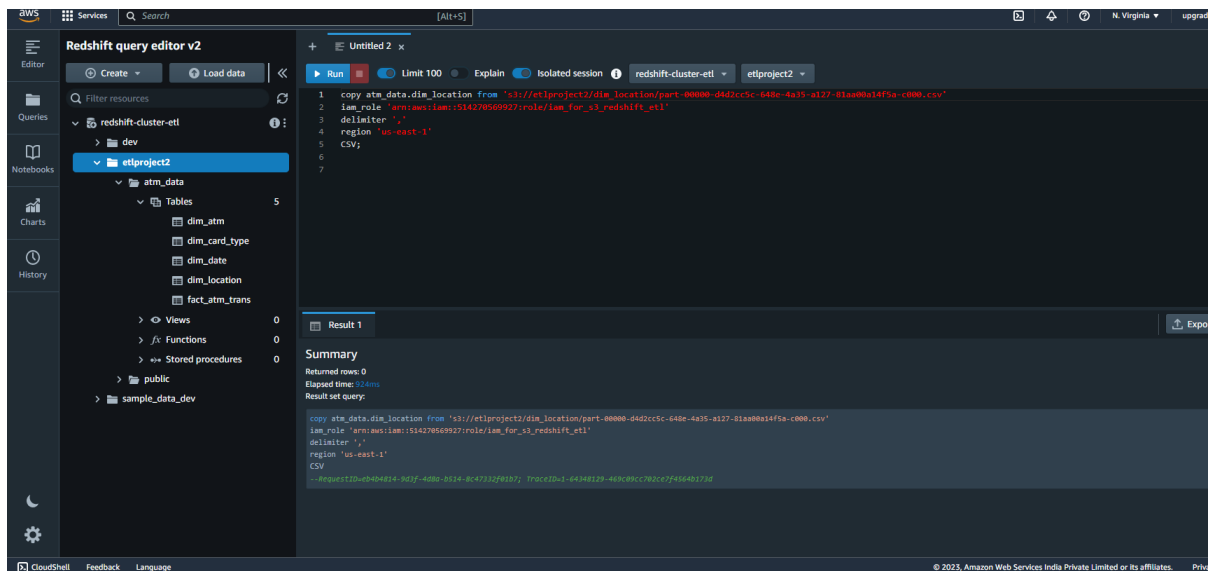
copy atm_data.dim_location from 's3://etlproject2/dim_location/part-00000-d4d2cc5c-648e-4a35-a127-81aa00a14f5a-c000.csv'

iam_role 'arn:aws:iam::514270569927:role/iam_for_s3_redshift_etl'

delimiter ','

region 'us-east-1'

CSV;



Copying the data to dim_atm table

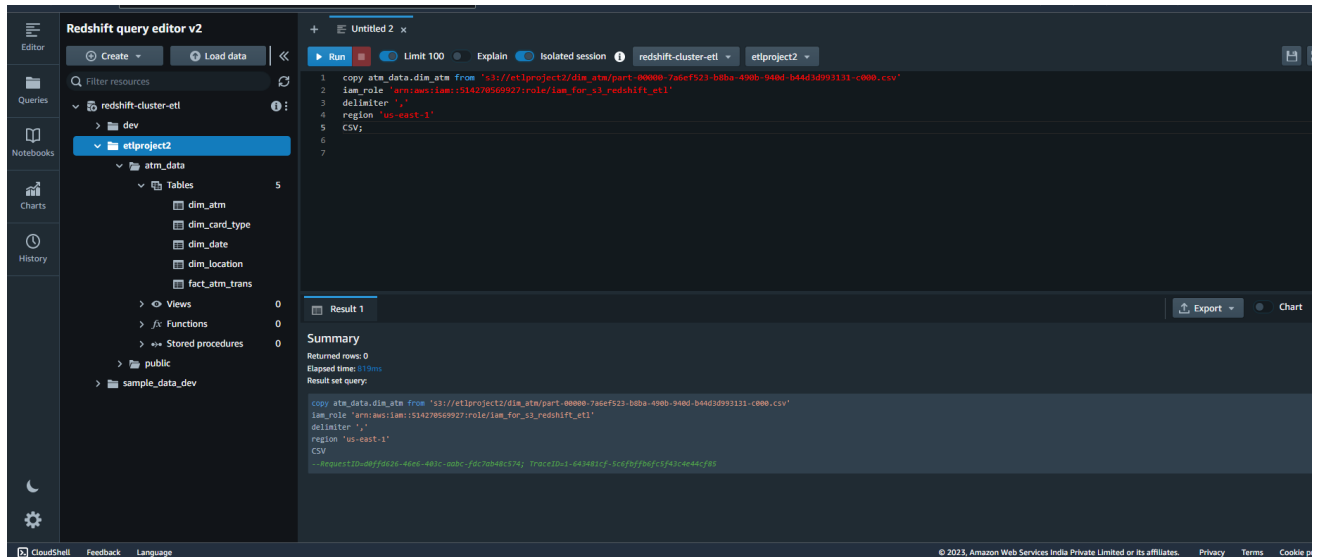
copy atm_data.dim_atm from 's3://etlproject2/dim_atm/part-00000-7a6ef523-b8ba-490b-940d-b44d3d993131-c000.csv'

iam_role 'arn:aws:iam::514270569927:role/iam_for_s3_redshift_etl'

delimiter ','

region 'us-east-1'

CSV;



Copying the data to dim_date table

copy atm_data.dim_date from 's3://etlproject2/dim_date/part-00000-5220ab9d-6ffe-4c3d-adcf-bcd2df98a68b-c000.csv'

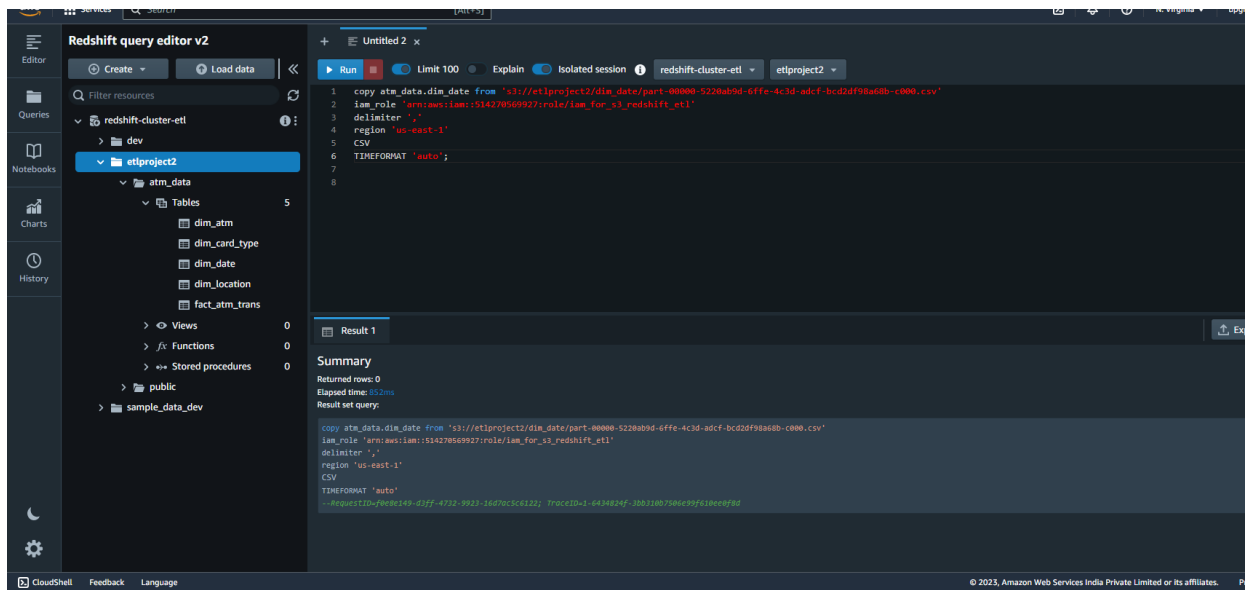
iam_role 'arn:aws:iam::514270569927:role/iam_for_s3_redshift_etl'

delimiter ','

region 'us-east-1'

CSV

TIMEFORMAT 'auto';



Copying the data to dim_card_type table

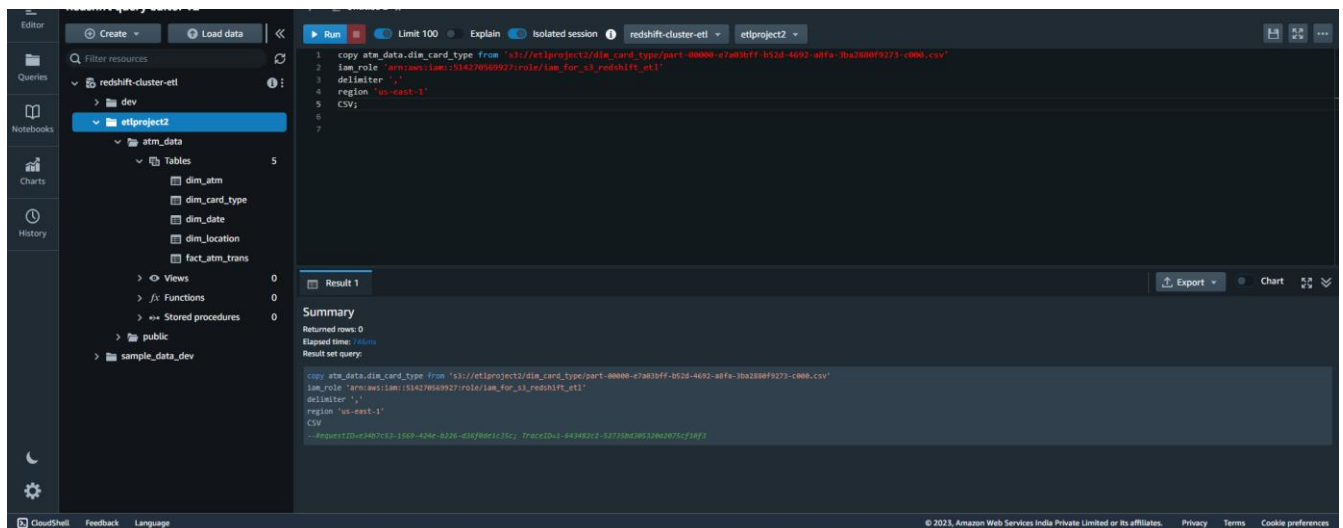
copy atm_data.dim_card_type from 's3://etlproject2/dim_card_type/part-00000-e7a03bff-b52d-4692-a8fa-3ba2880f9273-c000.csv'

iam_role 'arn:aws:iam::514270569927:role/iam_for_s3_redshift_etl'

delimiter ','

region 'us-east-1'

CSV;



Copying the data to fact_atm_trans table

copy atm_data.fact_atm_trans from 's3://etlproject2/fact_atm_trans/part-00000-bc87adae-848d-4915-9cec-0ed34794034f-c000.csv'

iam_role 'arn:aws:iam::514270569927:role/iam_for_s3_redshift_etl'

delimiter ','

region 'us-east-1'

CSV;

The screenshot displays the AWS Redshift query editor v2 interface. On the left, a sidebar shows the 'Filter resources' section with a tree view of the database schema, including 'atm_data' and 'fact_atm_trans'. The main editor area contains a SQL query:

```
1 copy atm_data.fact_atm_trans from 's3://etlproject2/fact_atm_trans/part-00000-bc87adae-848d-4915-9cec-0ed34794034f-c000.csv'
2 iam_role 'arn:aws:iam::514270569927:role/iam_for_s3_redshift_etl'
3 delimiter ','
4 region 'us-east-1'
5 CSV;
6
7
```

 Below the query, the 'Result 1' section shows a 'Summary' of the query execution, including 'Returned rows: 0', 'Elapsed time: 16.5s', and the 'Result set query'. The footer of the interface includes 'CloudShell', 'Feedback', 'Language', and a copyright notice for Amazon Web Services India Private Limited.