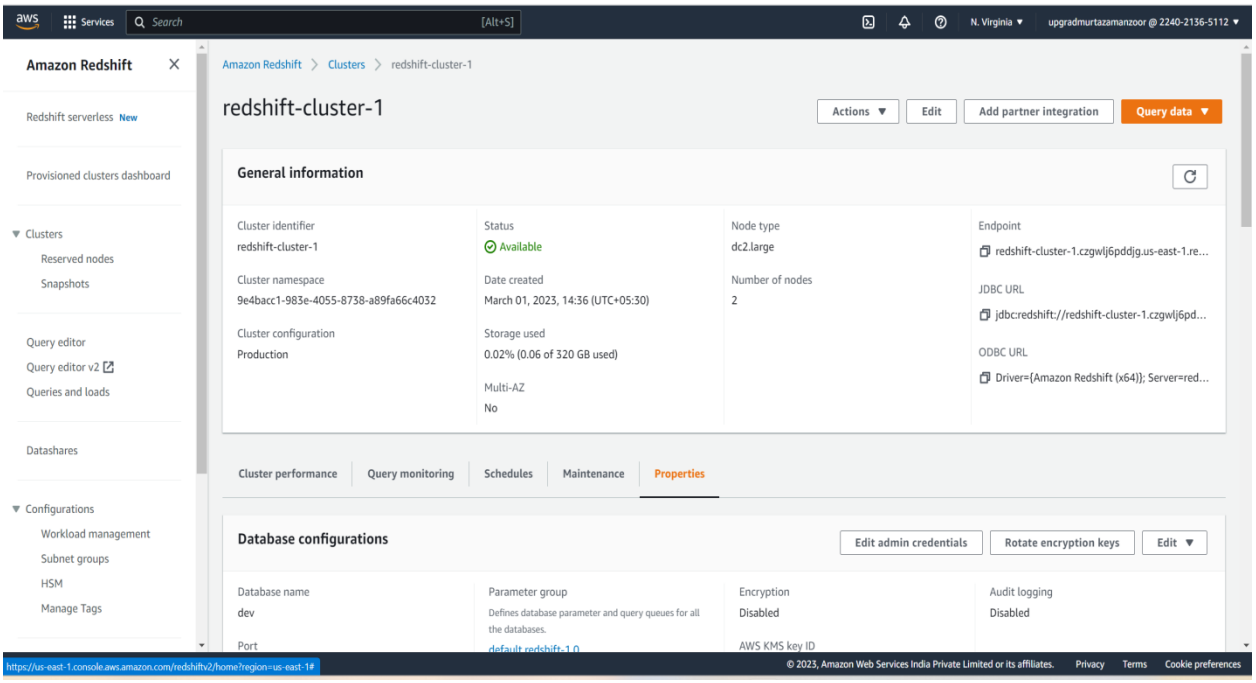
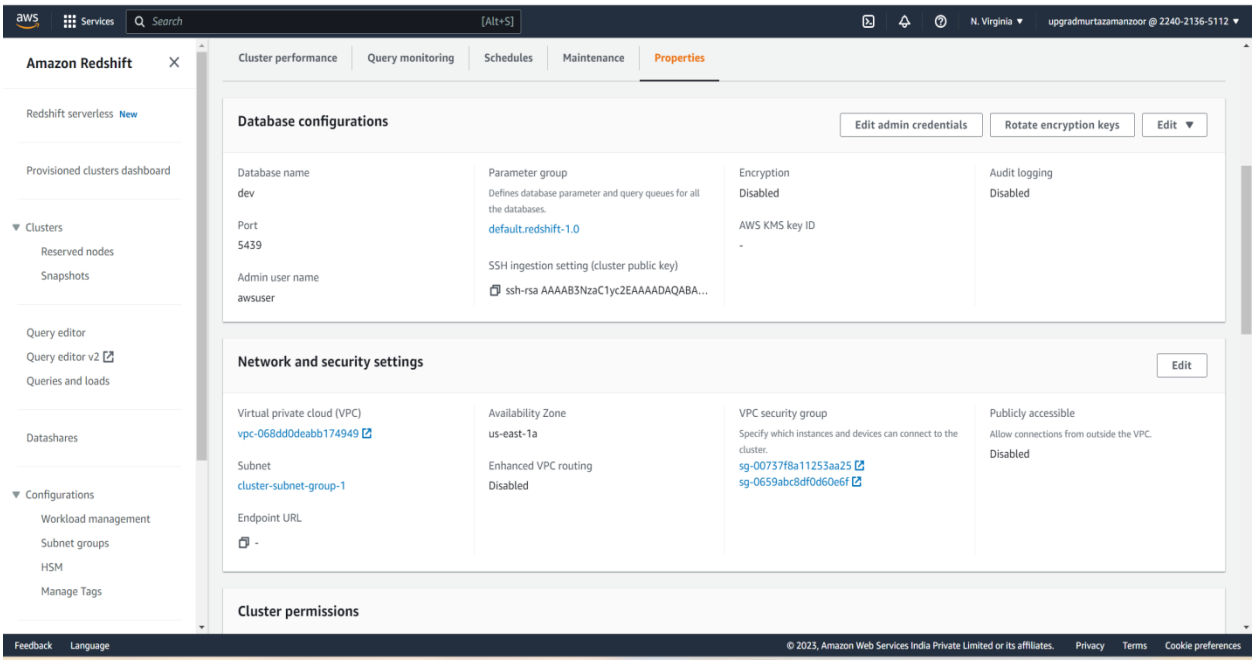
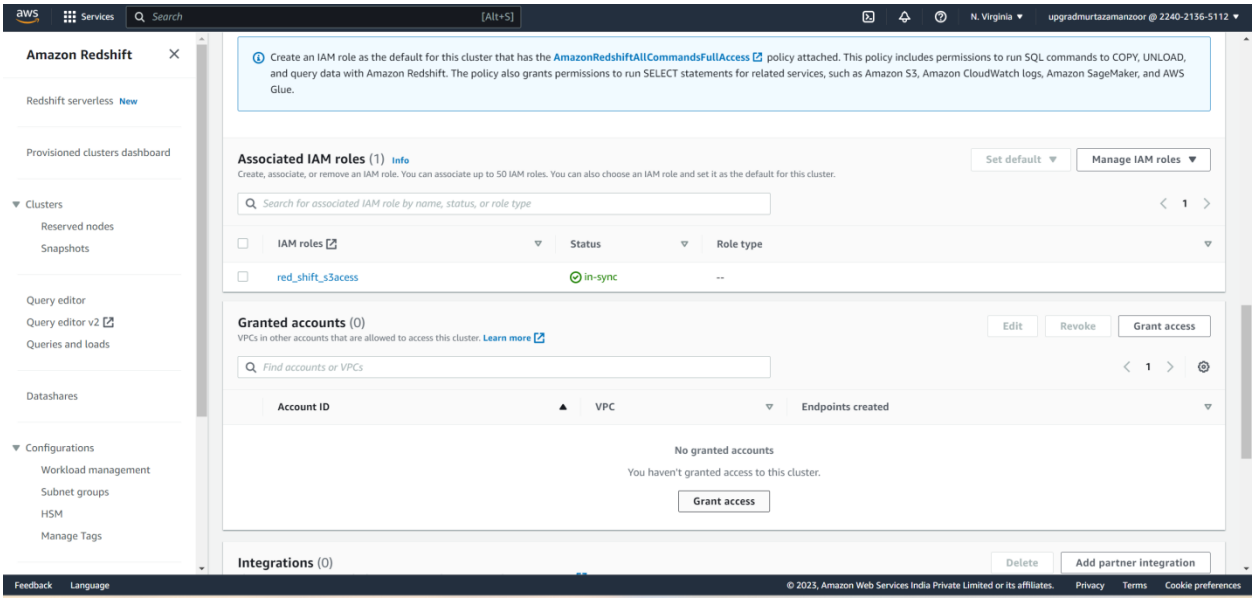
Creation of a Redshift Cluster

Screenshots of the configuration of the Redshift cluster that has been created:

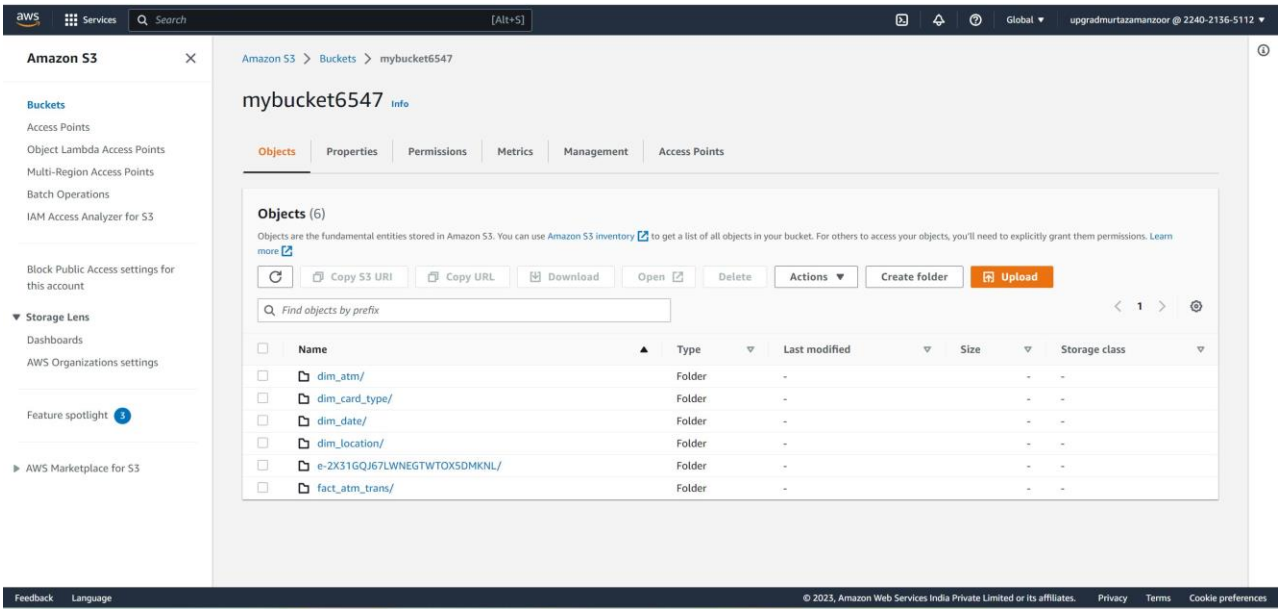






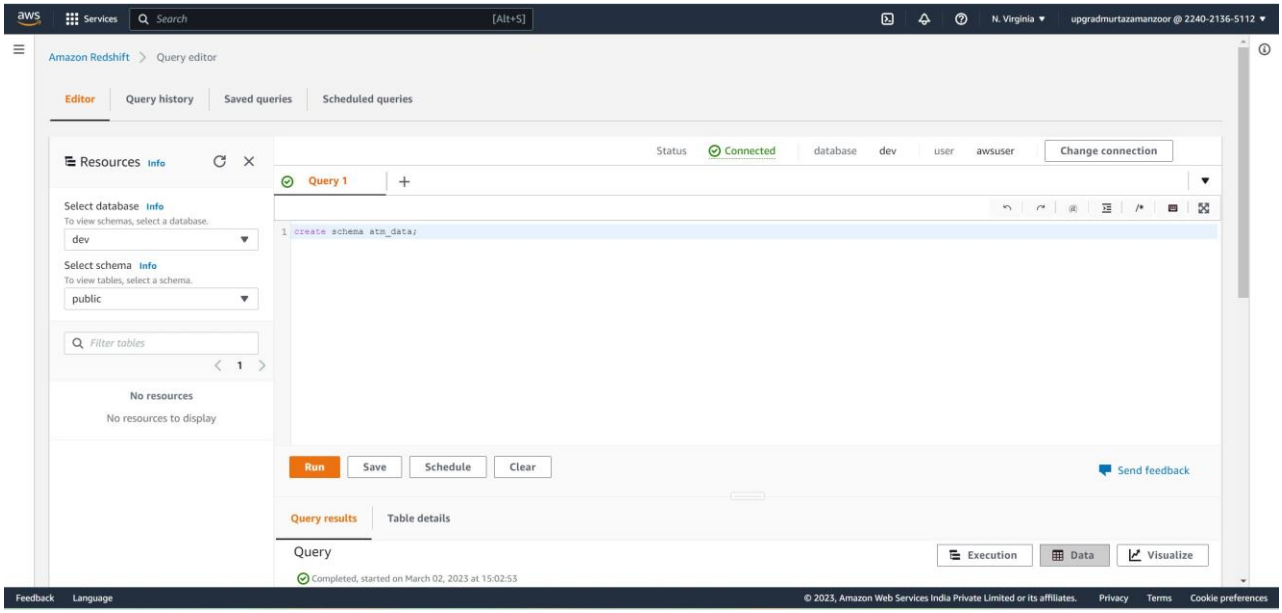
Setting up a database in the Redshift cluster and running queries to createthe dimension and fact tables.

S3 bucket containing Files:



Query for creating schema:

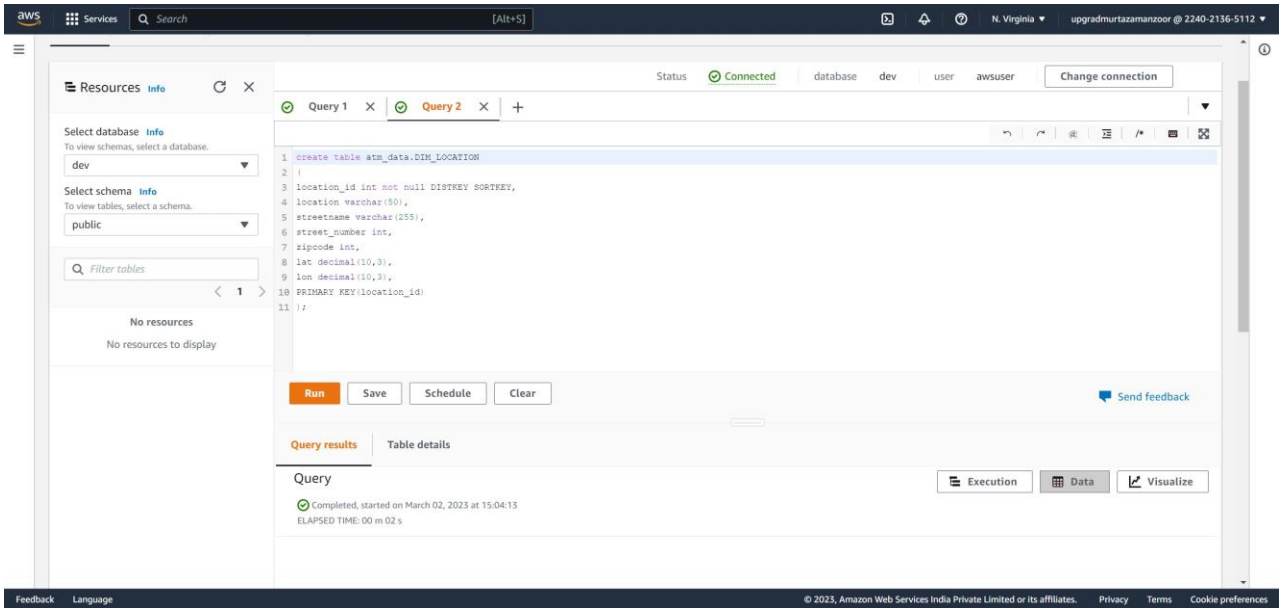
```
create schema atm_data;
```



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

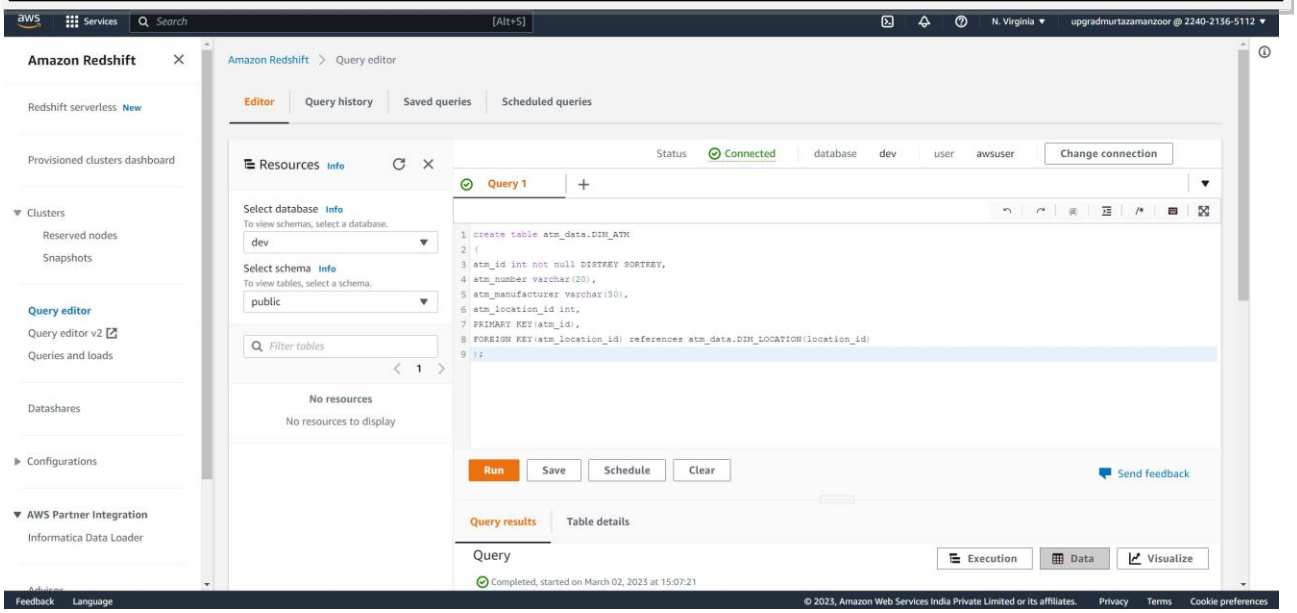
• Creating location dimension table

```
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)
);
```



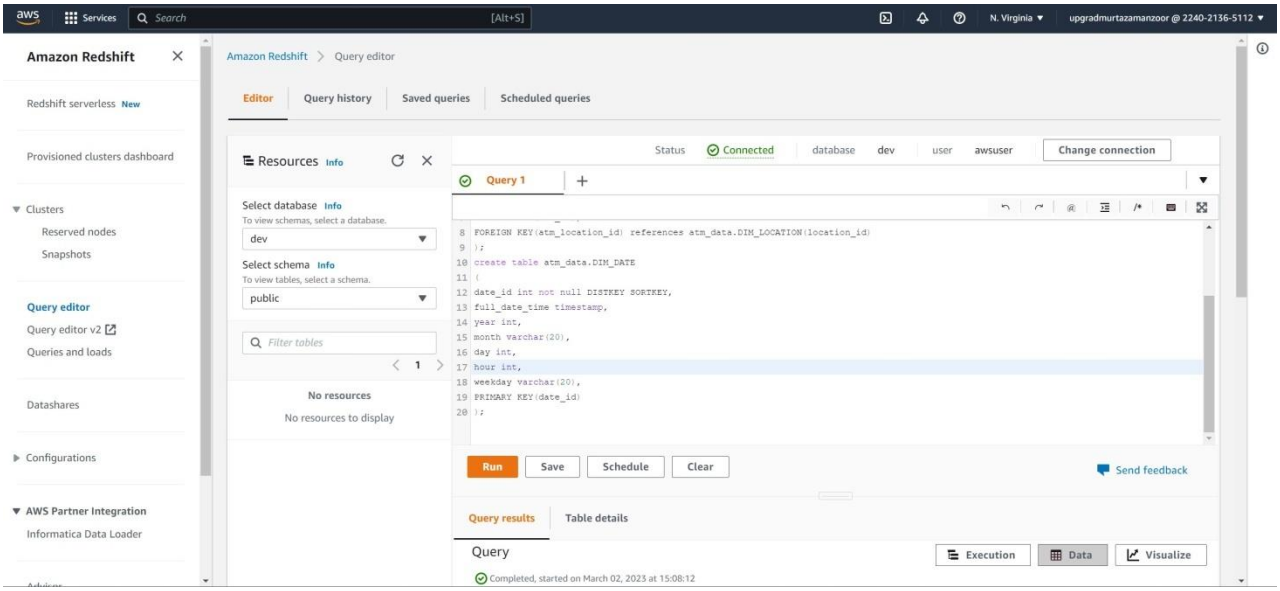
- **Creating atm dimension table**

```
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)
);
```



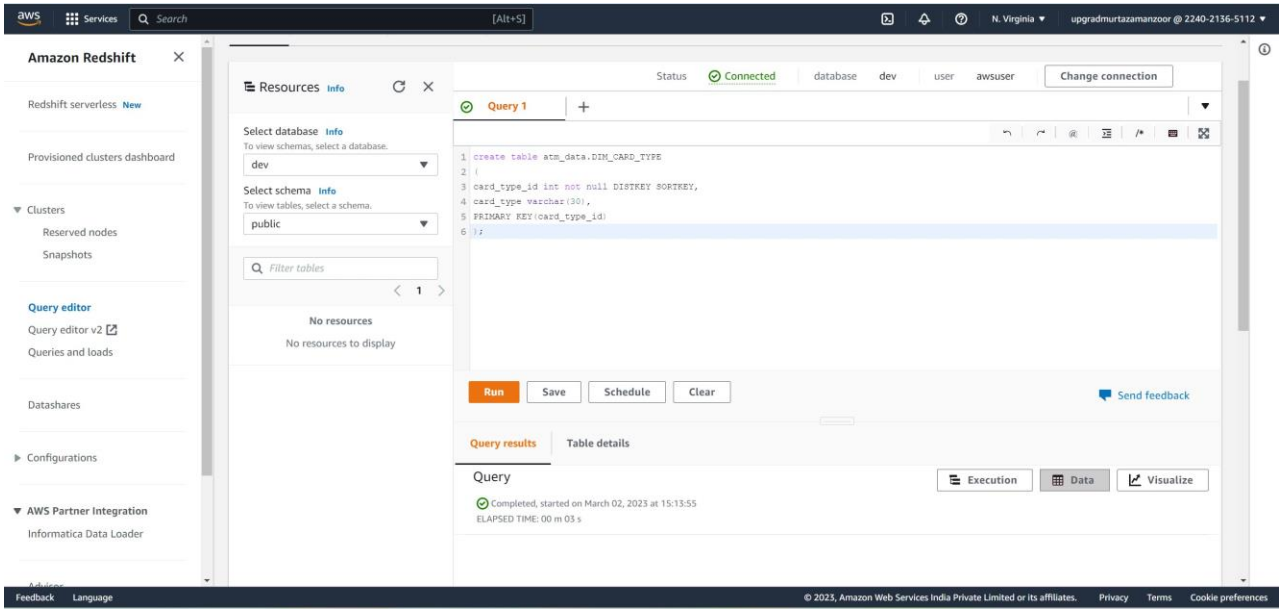
- **Creating date dimension table**

```
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)
);
```



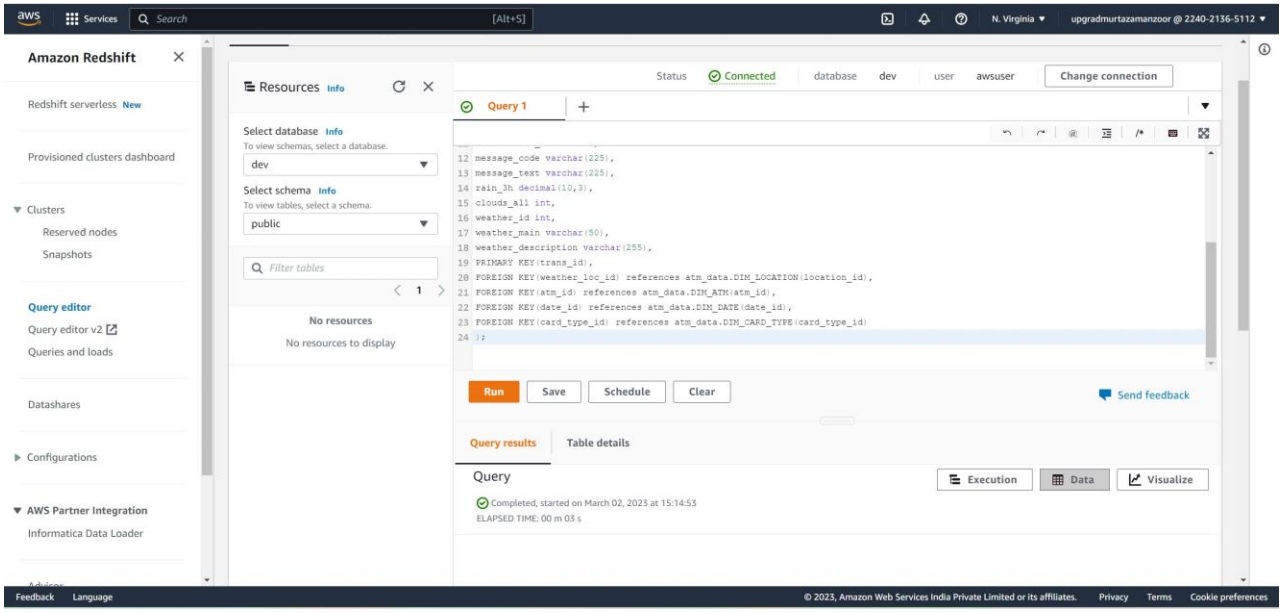
- **Creating card type dimension table**

```
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)
);
```

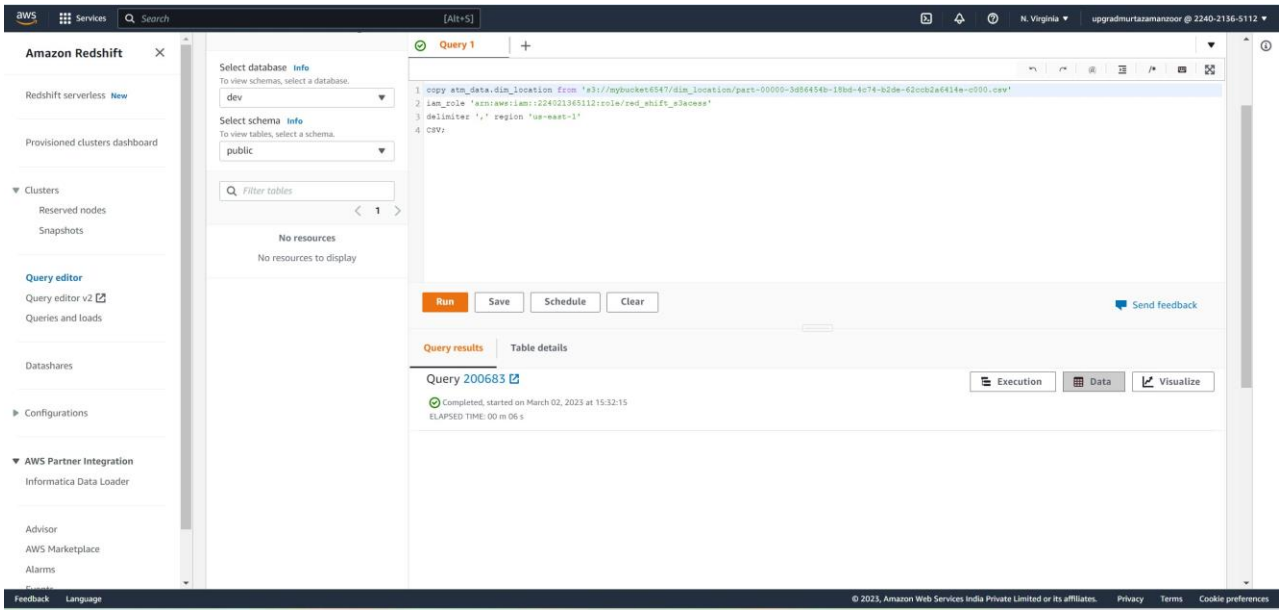


Loading data into a Redshift cluster from Amazon S3 bucket

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

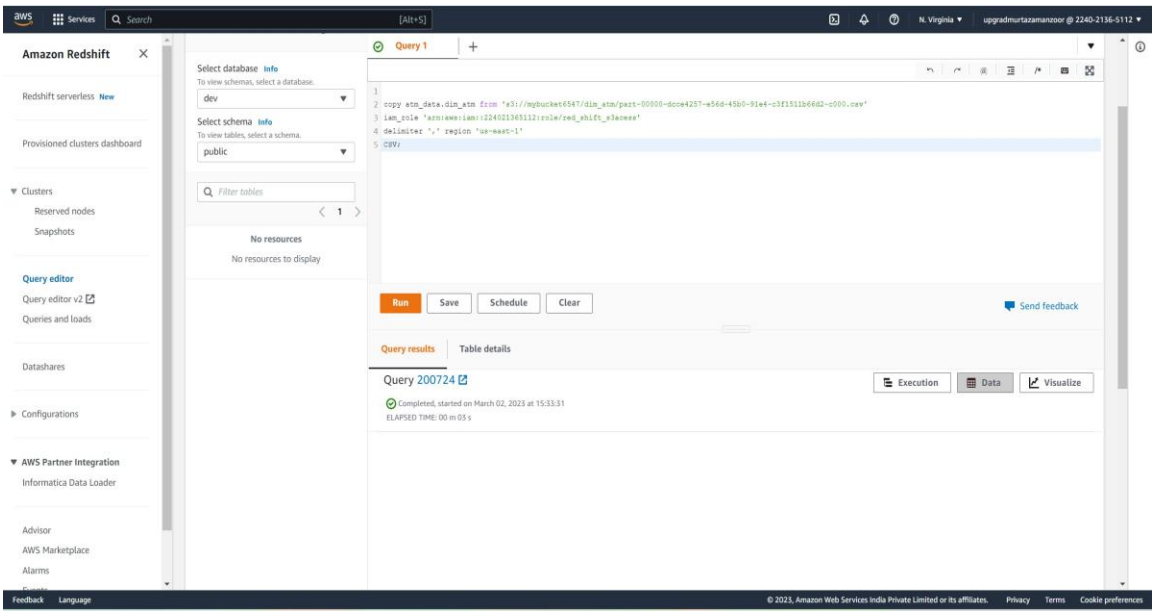
• Copying the data to dim_location table

```
copy atm_data.dim_location from 's3://mybucket6547/dim_location/part-00000-
3d86454b-18bd-4c74-b2de-62ccb2a6414e-c000.csv'
iam_role 'arn:aws:iam::224021365112:role/red_shift_s3acess'
delimiter ',' region 'us-east-1'
CSV;
```



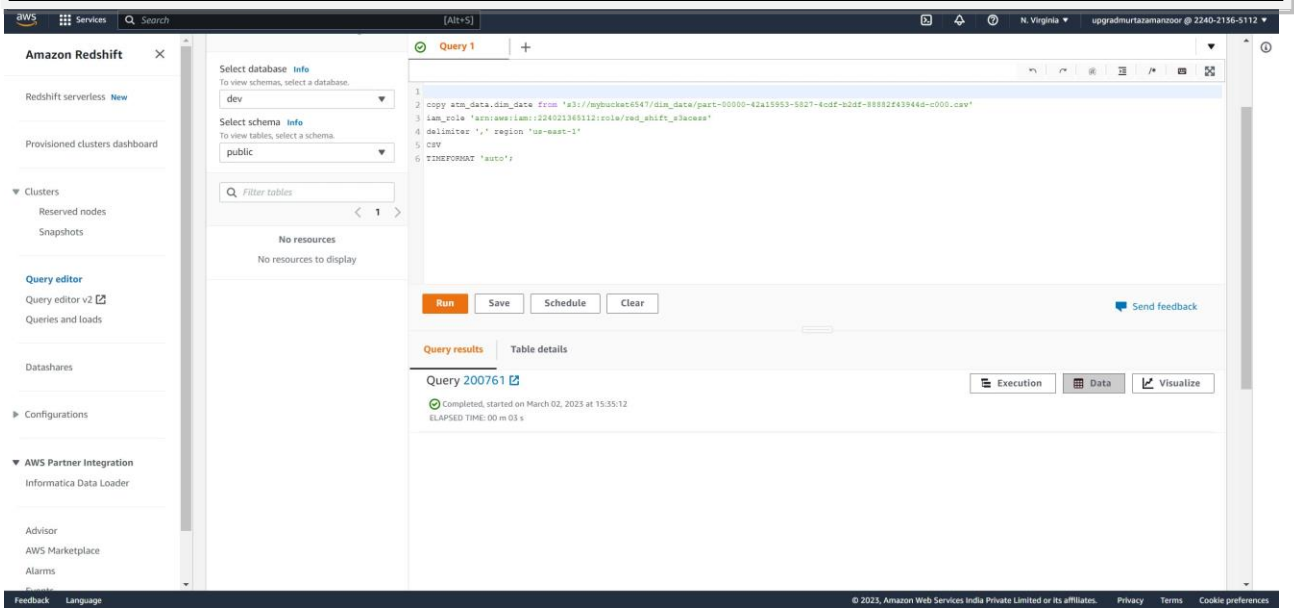
- **Copying the data to dim_atm table**

```
copy atm_data.dim_atm from 's3://mybucket6547/dim_atm/part-00000-dcce4257-e56d-45b0-91e4-c3f1511b66d2-c000.csv'
iam_role 'arn:aws:iam::224021365112:role/red_shift_s3acess'
delimiter ',' region 'us-east-1'
CSV;
```



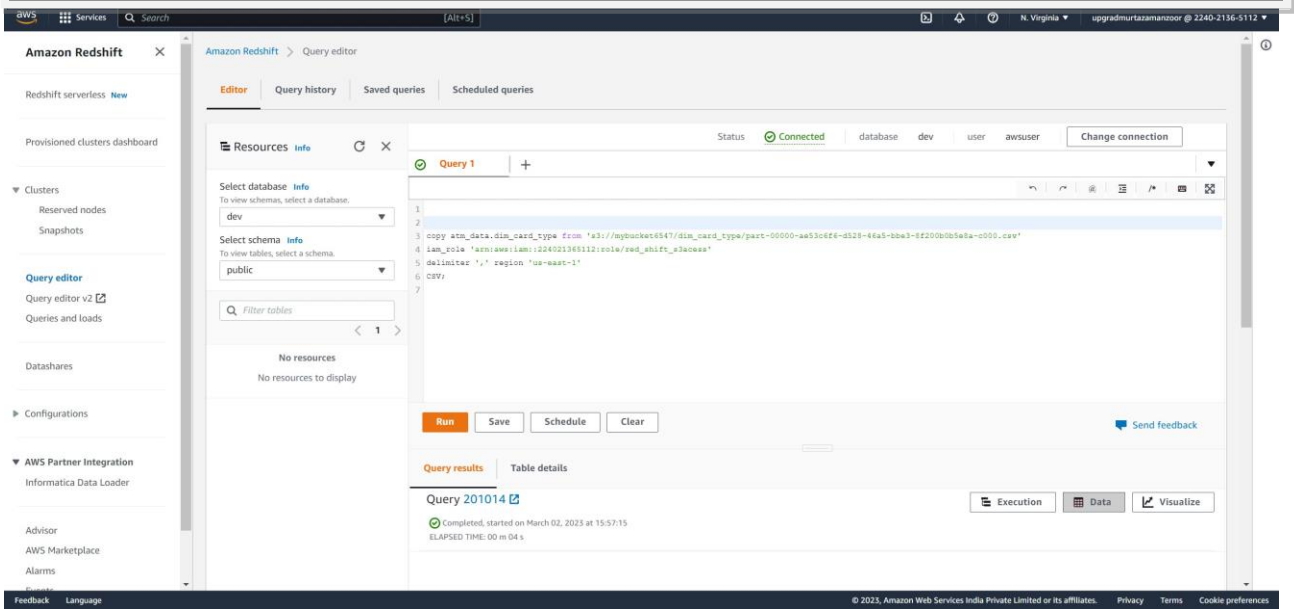
- **Copying the data to dim_date table**

```
copy atm_data.dim_date from 's3://mybucket6547/dim_date/part-00000-42a15953-5827-4cdf-b2df-88882f43944d-c000.csv'
iam_role 'arn:aws:iam::224021365112:role/red_shift_s3acess'
delimiter ',' region 'us-east-1'
CSV
TIMEFORMAT  'auto';
```



- **Copying the data to dim_card_type table**

```
copy atm_data.dim_card_type from 's3://mybucket6547/dim_card_type/part-00000-
ae53c6f6-d528-46a5-bbe3-8f200b0b5e8a-c000.csv'
iam_role 'arn:aws:iam::224021365112:role/red_shift_s3acess'
delimiter ',' region 'us-east-1'
CSV;
```



- **Copying the data to fact_atm_trans table**

```
copy atm_data.fact_atm_trans from 's3://mybucket6547/fact_atm_trans/part-00000-
b229e2cd-f4d4-4158-90a5-22ce70a893eb-c000.csv'
iam_role 'arn:aws:iam::224021365112:role/red_shift_s3acess'
delimiter ',' region 'us-east-1'
CSV;
```