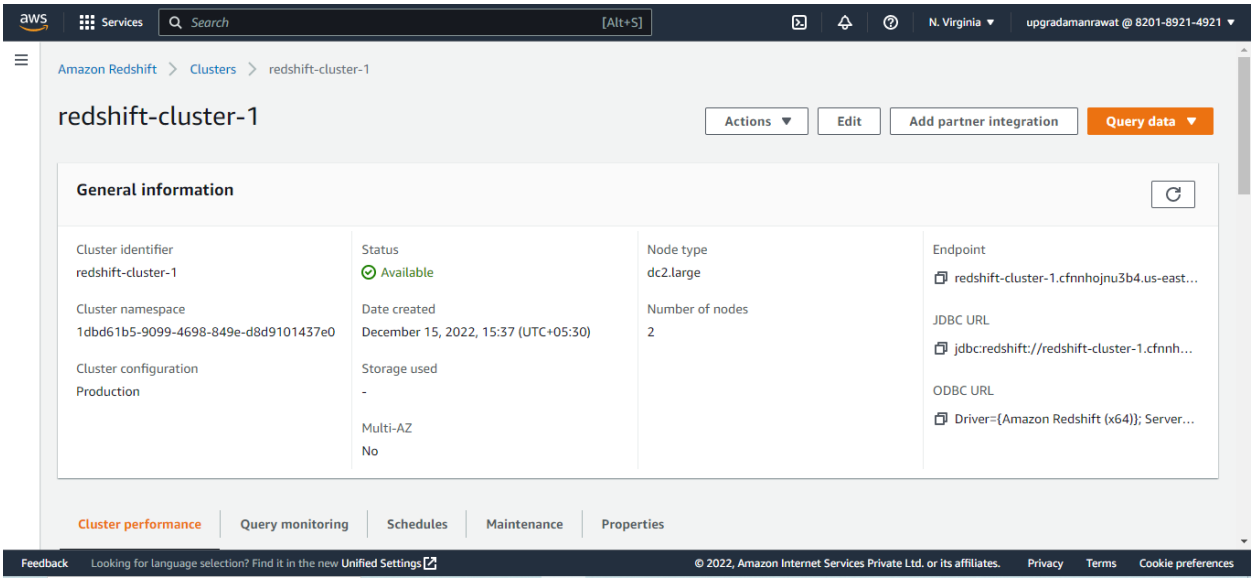
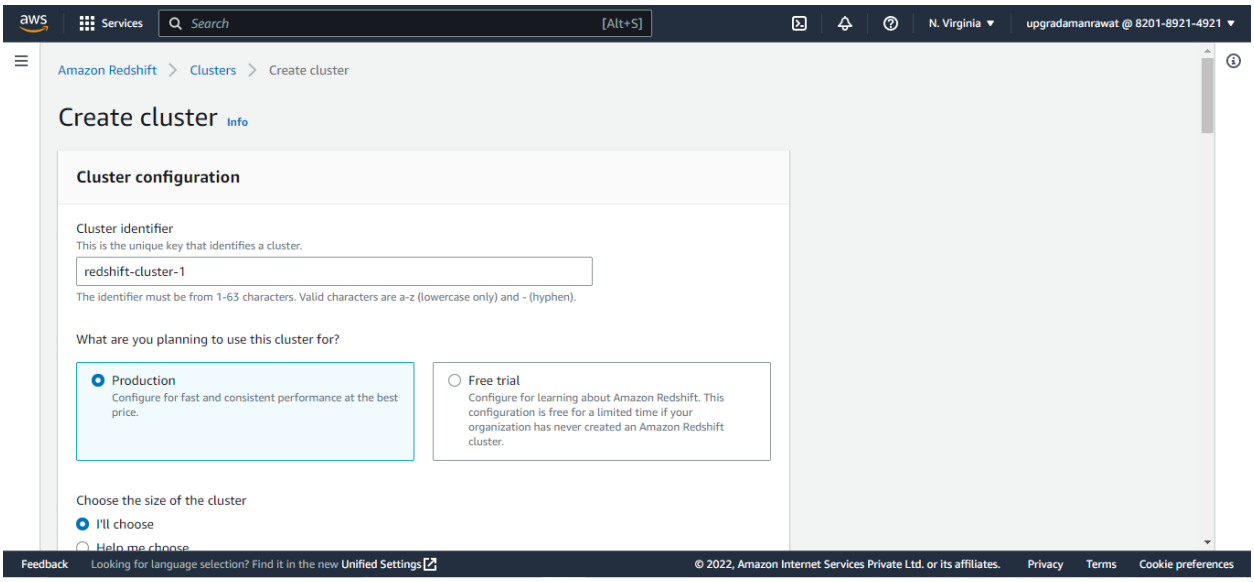
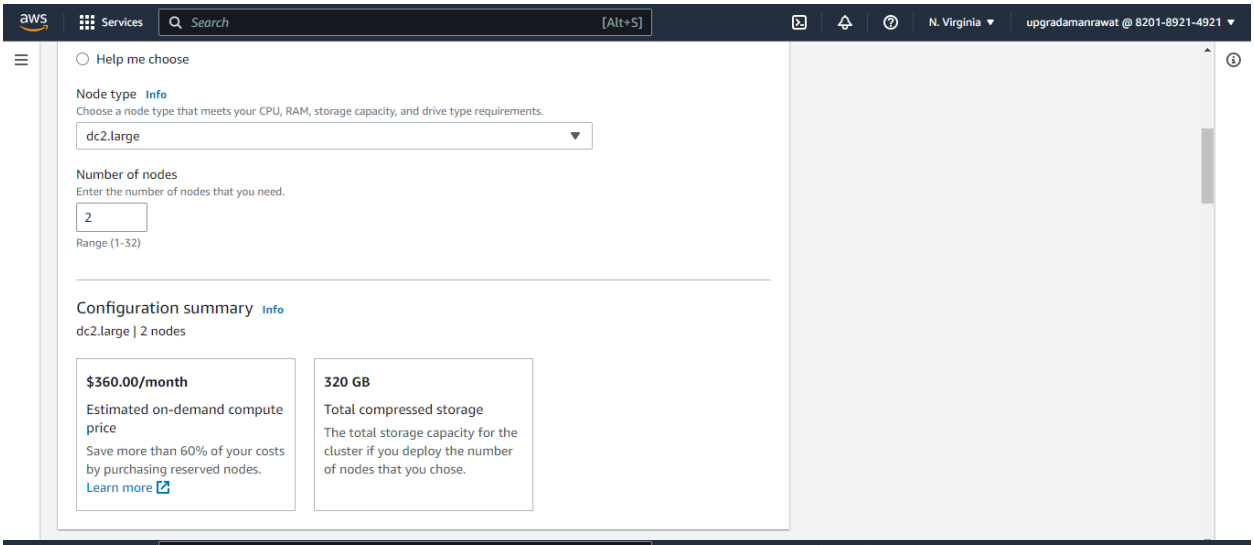
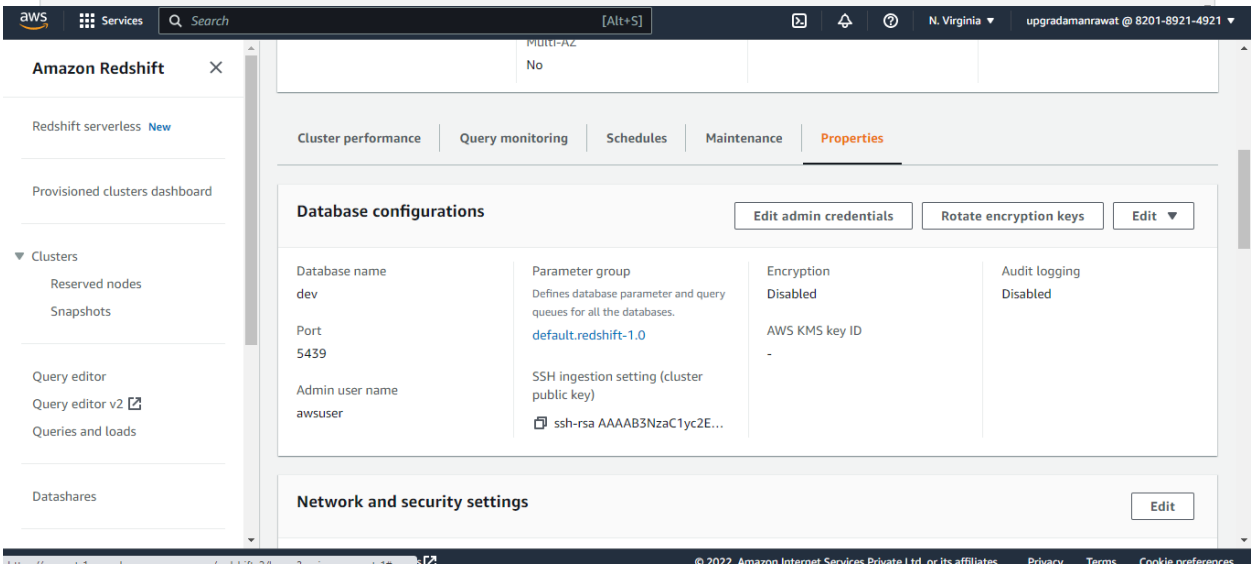
Creation of a Redshift Cluster

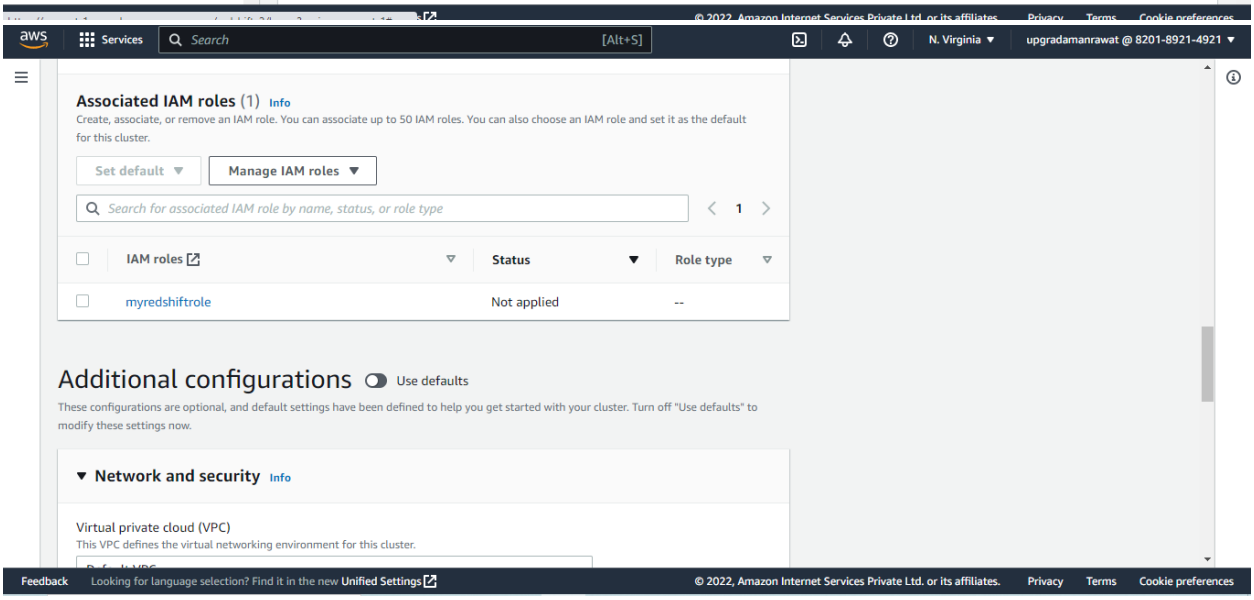
Screenshots of the configuration of the Redshift cluster that you have created:

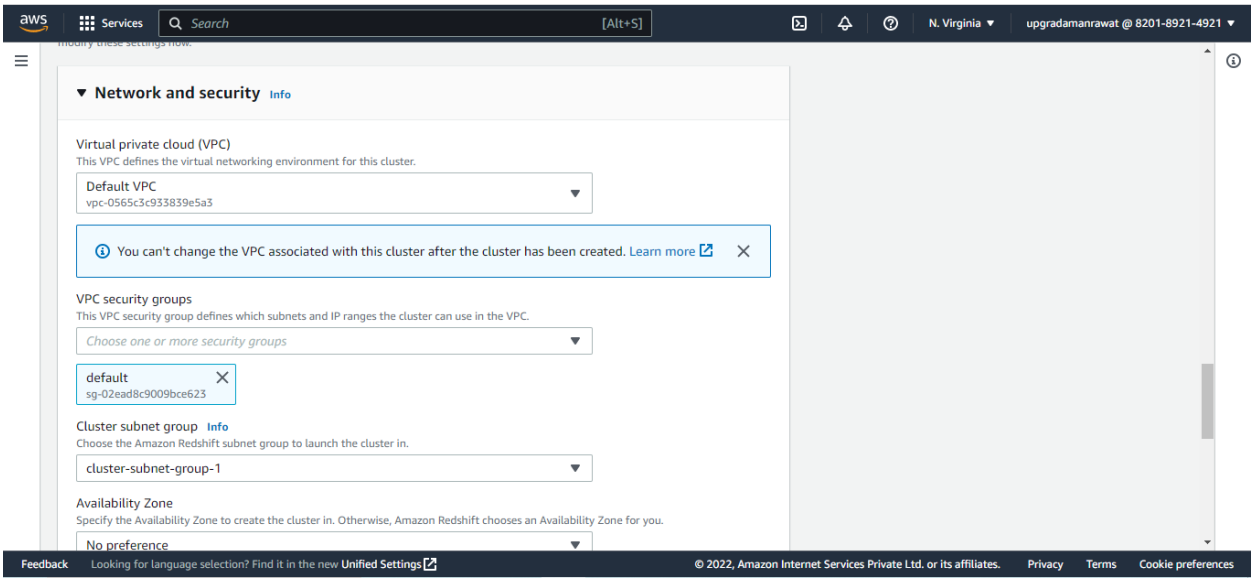


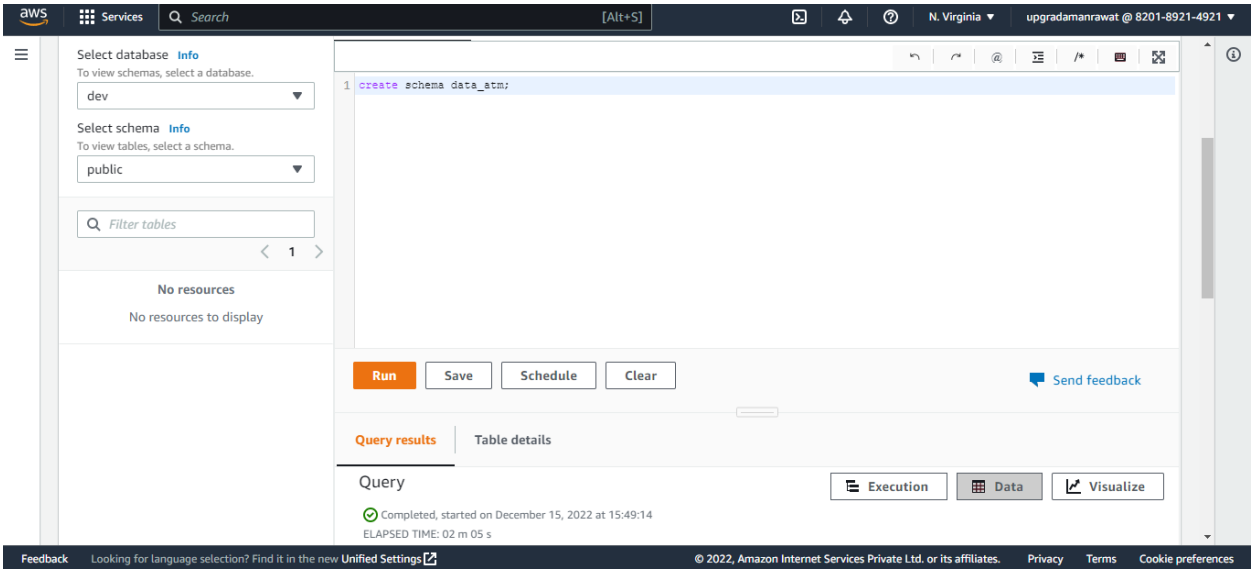


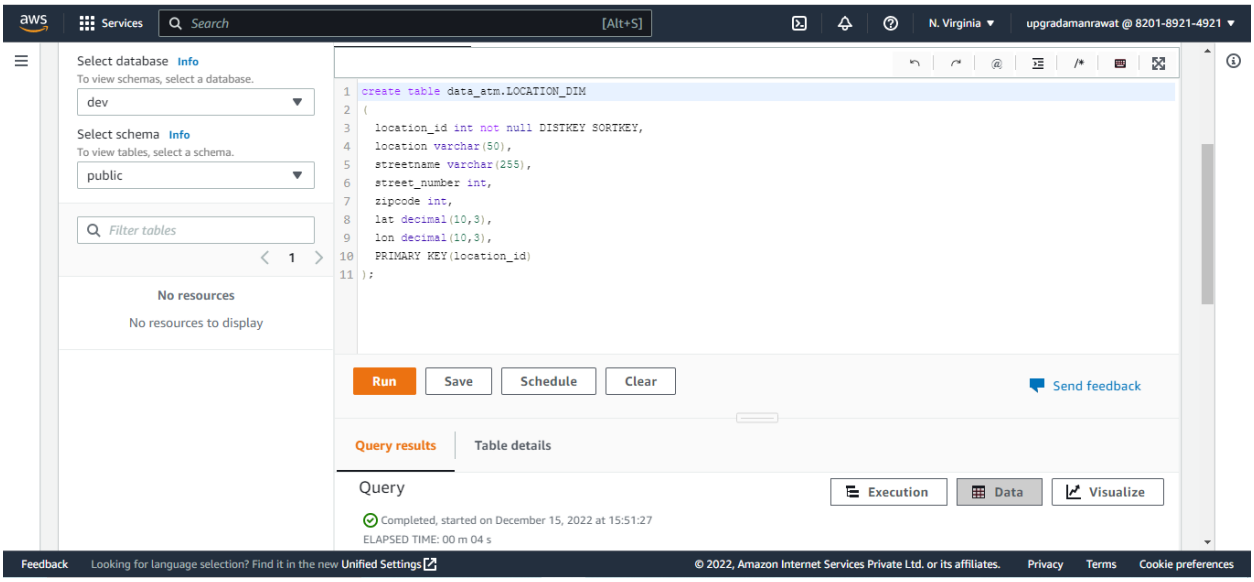








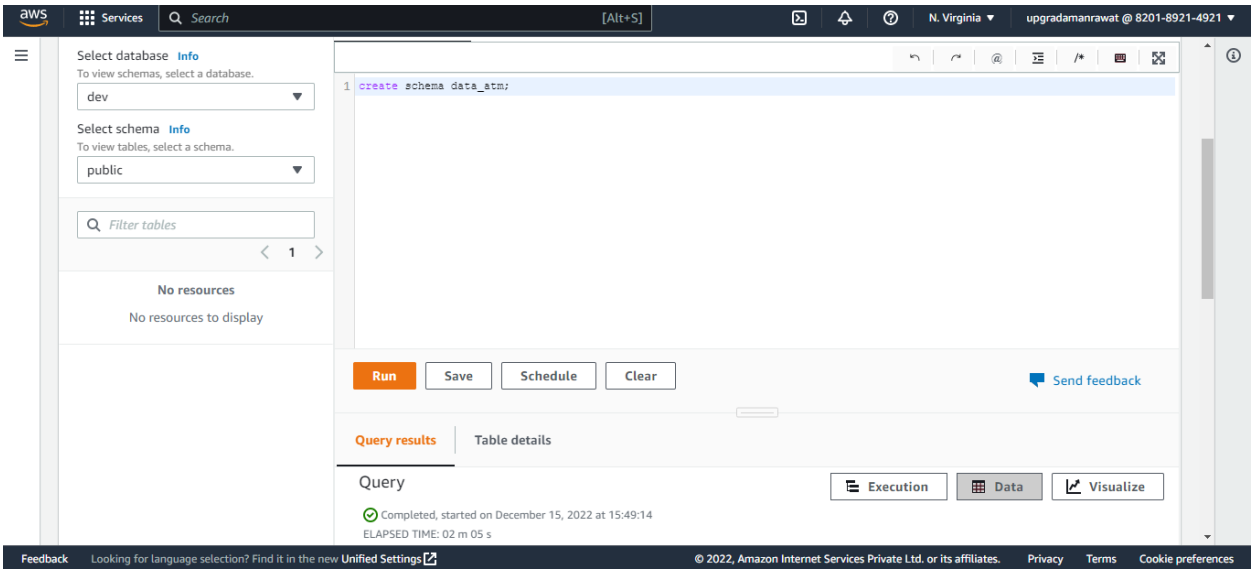




Setting up a database in the Redshift cluster and running queries to create the dimension and fact tables

Query for creating schema:

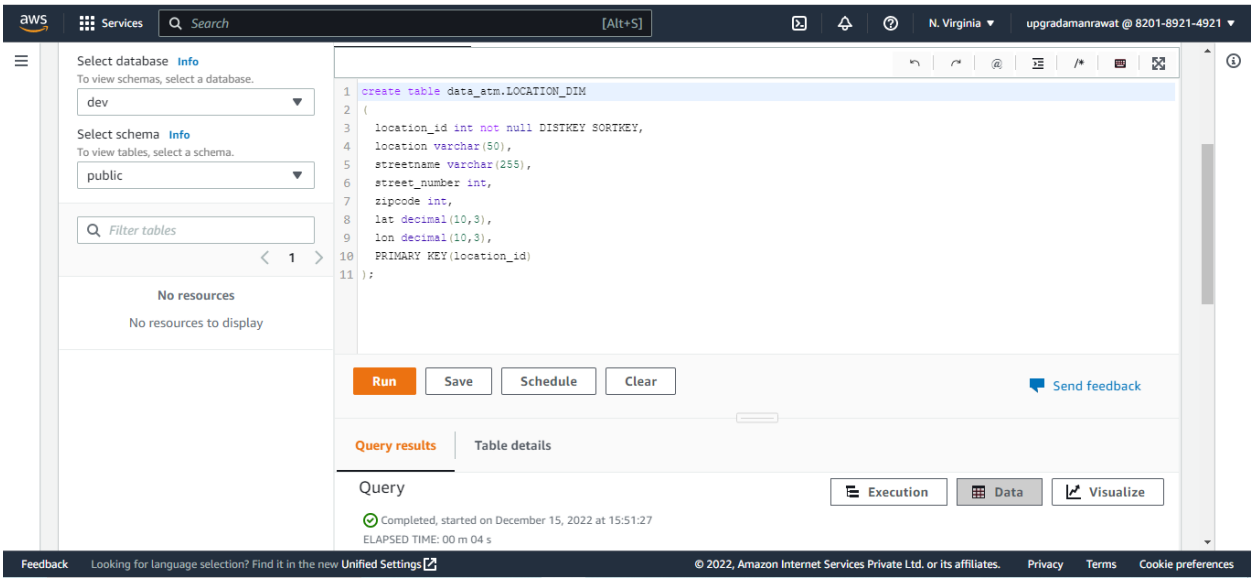
create schema data_atm;



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

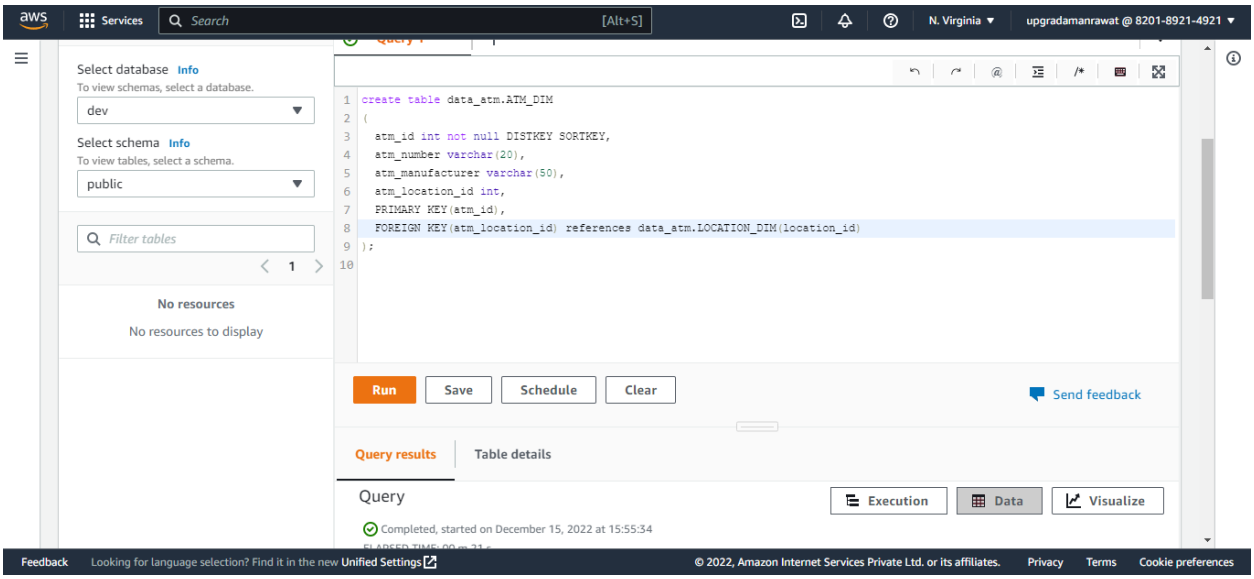
• Creating location dimension table:

```
create table data_atm.LOCATION_DIM
(
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)
);
```



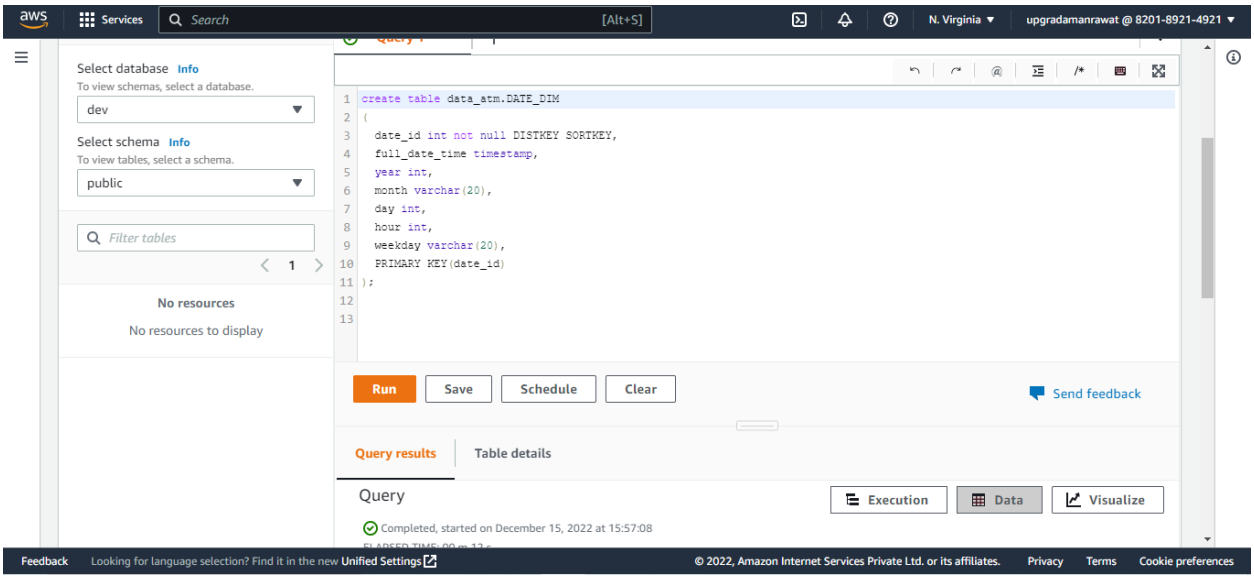
• Creating atm dimension table

```
create table data_atm.ATM_DIM
(
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 data_atm.LOCATION_DIM (location_id)
);
```



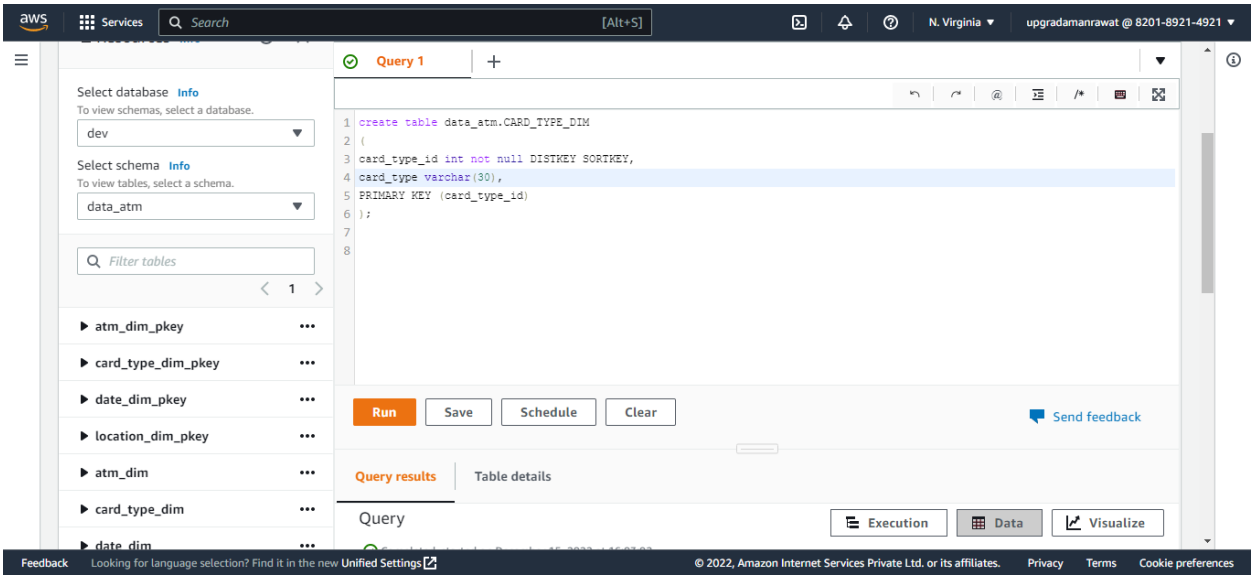
• **Creating date dimension table**

```
create table data_atm.DATE_DIM
(
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)
);
```



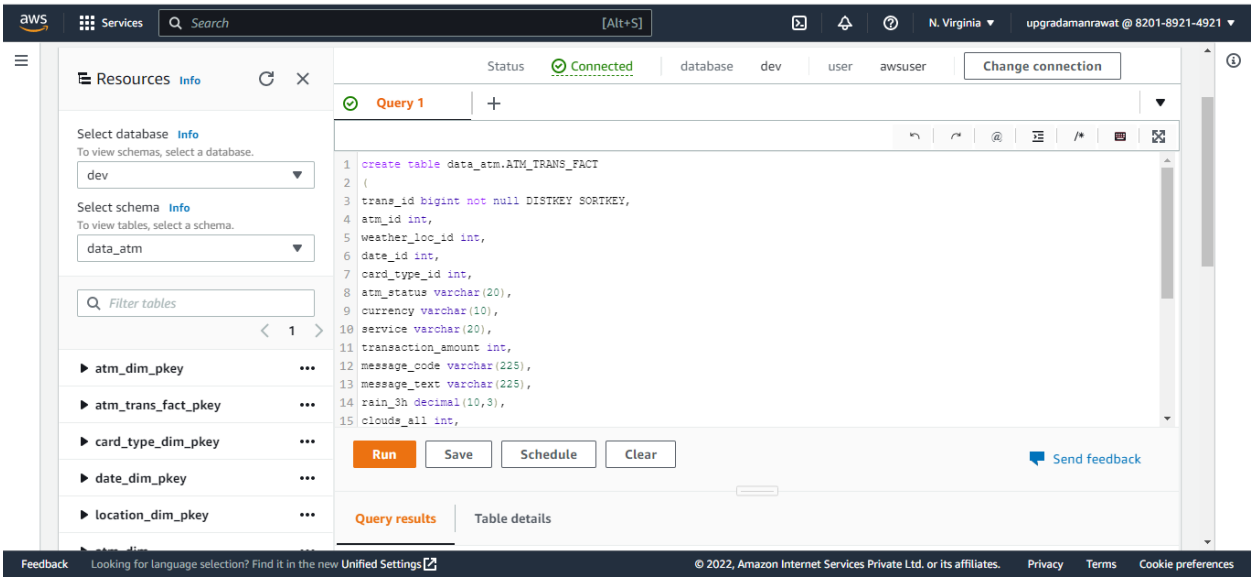
• **Creating card-type dimension table**

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



• Creating atm transactions fact table

```
create table data_atm.ATM_TRANS_FACT
(
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 data_atm.LOCATION_DIM (location_id),
FOREIGN KEY(atm_id) references data_atm.ATM_DIM(atm_id),
FOREIGN KEY(date_id) references data_atm.DATE_DIM(date_id),
FOREIGN KEY(card_type_id) references data_atm.CARD_TYPE_DIM(card_type_id)
);
```

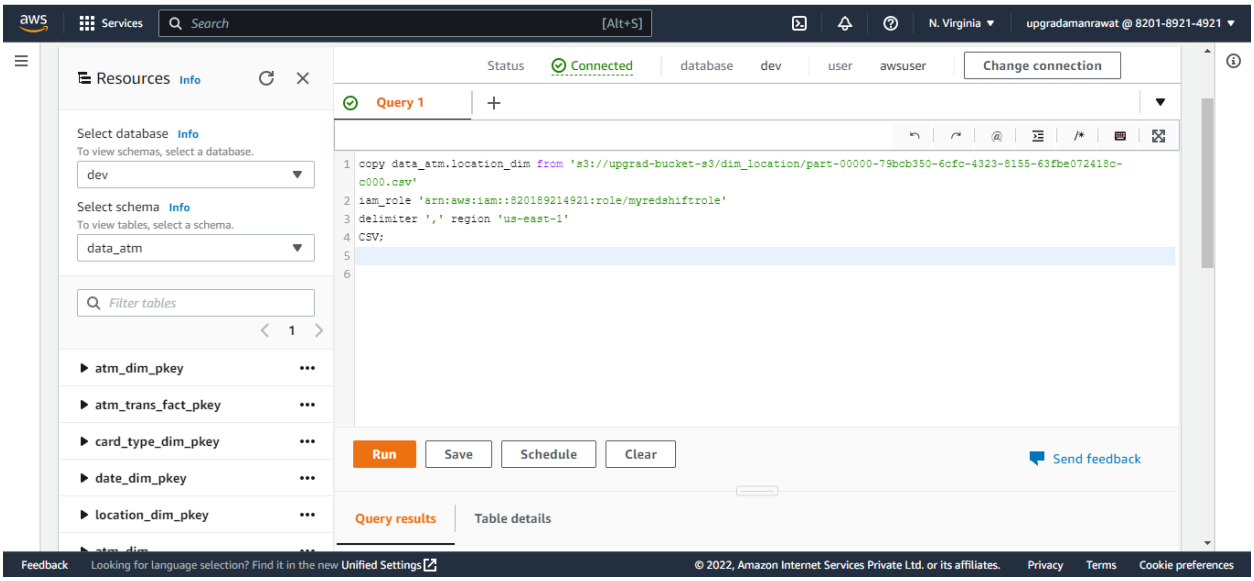


Loading data into a Redshift cluster from Amazon S3 bucket

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

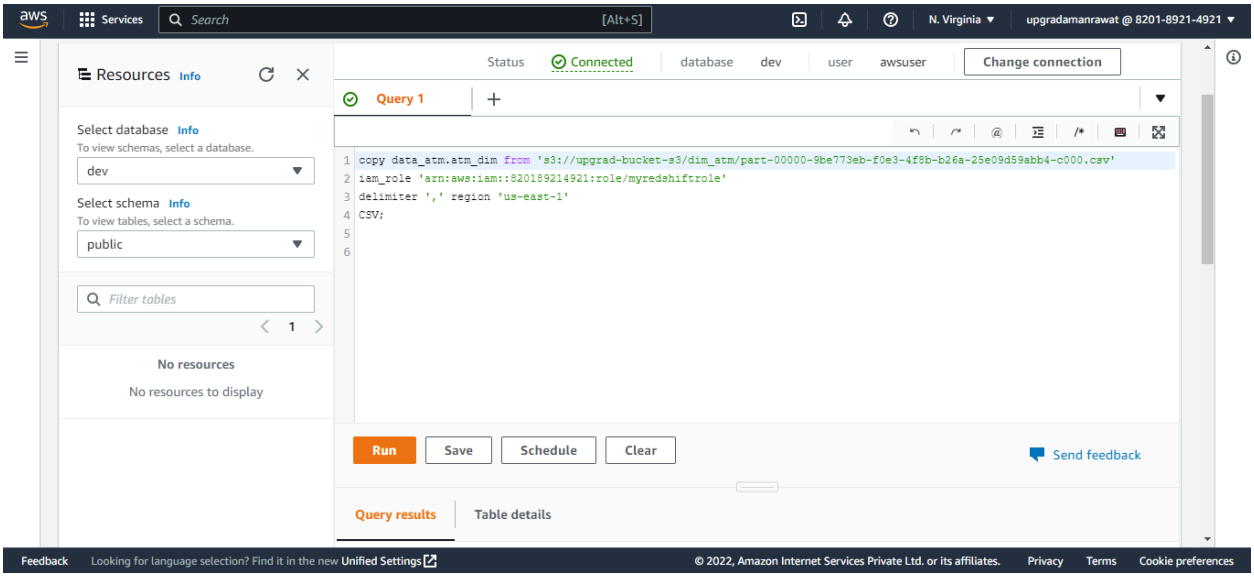
• Copying the data to location_dim table

```
copy data_atm.location_dim from 's3://upgrad-bucket-s3/dim_location/part-00000-
79bcb350-6cfc-4323-8155-63fbe072418c-c000.csv'
iam_role 'arn:aws:iam::820189214921:role/myredshiftrole'
delimiter ',' region 'us-east-1'
CSV;
```



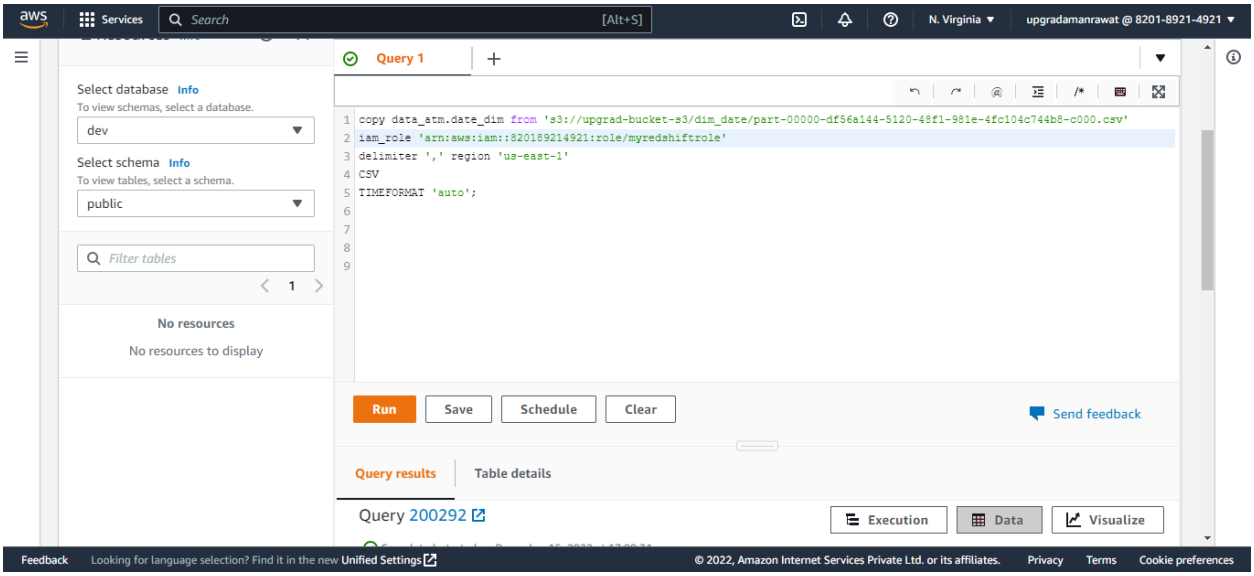
• **Copying the data to atm_dim table**

```
copy data_atm.atm_dim from 's3://upgrad-bucket-s3/dim_atm/part-00000-9be773eb-f0e3-4f8b-b26a-25e09d59abb4-c000.csv'
iam_role 'arn:aws:iam::820189214921:role/myredshiftrole'
delimiter ',' region 'us-east-1'
CSV;
```



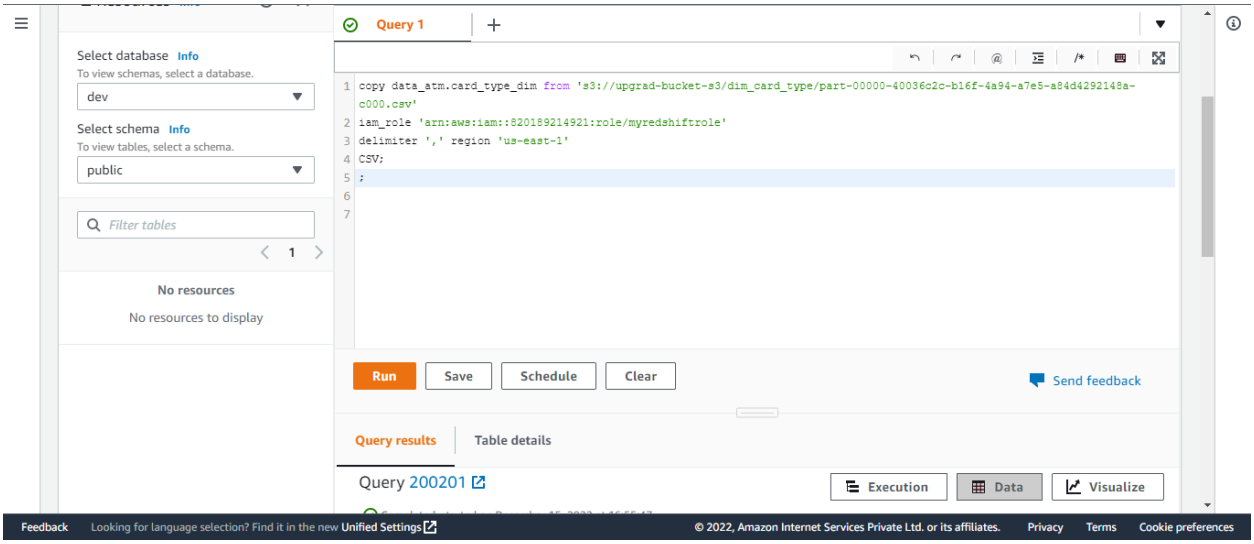
• **Copying the data to date_dim table**

```
copy data_atm.atm_dim from 's3://upgrad-bucket-s3/dim_date/part-00000-df56a144-5120-48f1-981e-4fc104c744b8-c000.csv'
iam_role 'arn:aws:iam::820189214921:role/myredshiftrole'
delimiter ',' region 'us-east-1'
CSV TIMEFORMAT 'auto';
;
```



• Copying the data to card_type_dim table

```
copy data_atm.card_type_dim from 's3://upgrad-bucket-s3/dim_card_type/part-00000-40036c2c-b16f-4a94-a7e5-a84d4292148a-c000.csv'
iam_role 'arn:aws:iam::820189214921:role/myredshiftrole'
delimiter ',' region 'us-east-1'
CSV;
;
```



• **Copying the data to atm_trans_fact table**

```
copy data_atm.atm_trans_fact from 's3://upgrad-bucket-s3/dim_card_type/part-00000-
a1a0ac28-e1de-4c74-af29-afd655f9bf9b-c000.csv'
iam_role 'arn:aws:iam::820189214921:role/myredshiftrole'
delimiter ',' region 'us-east-1'
CSV;
;
```