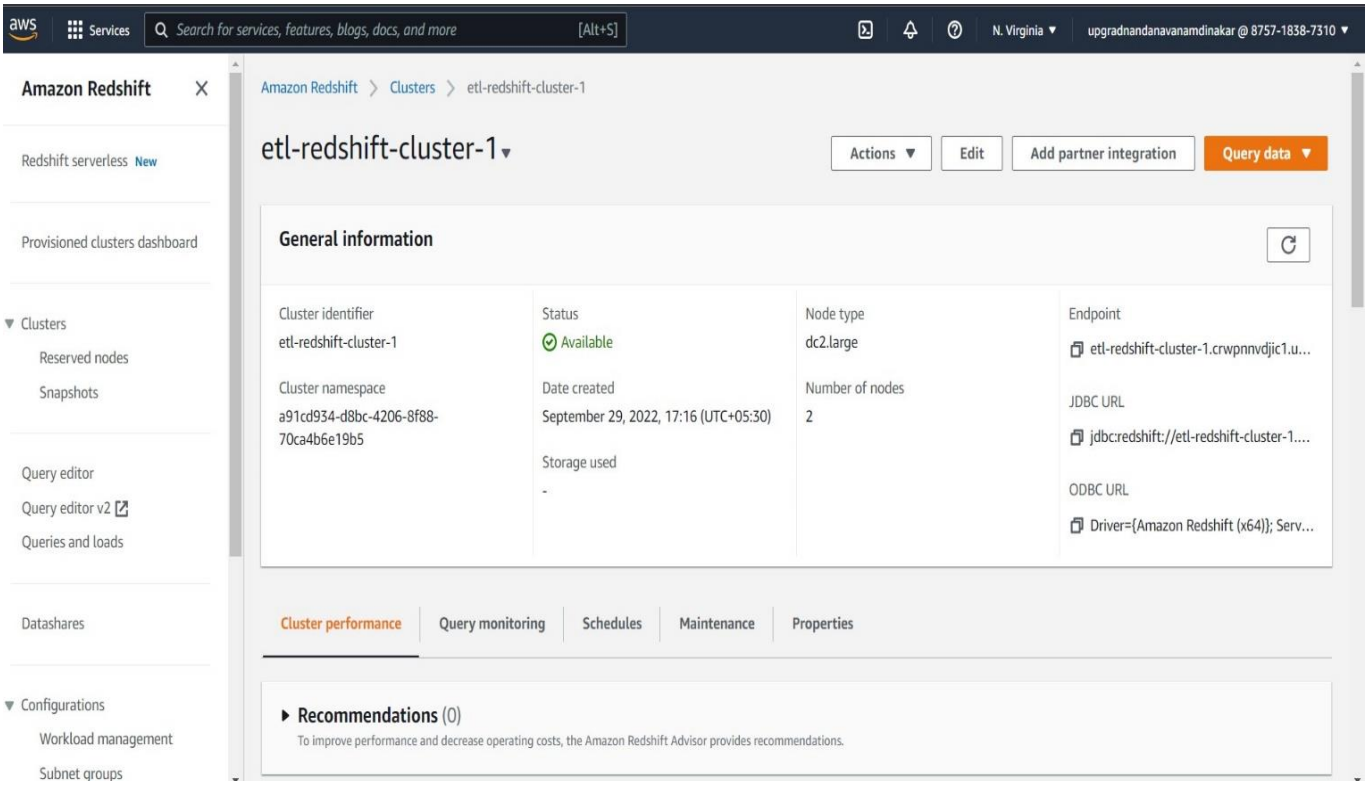
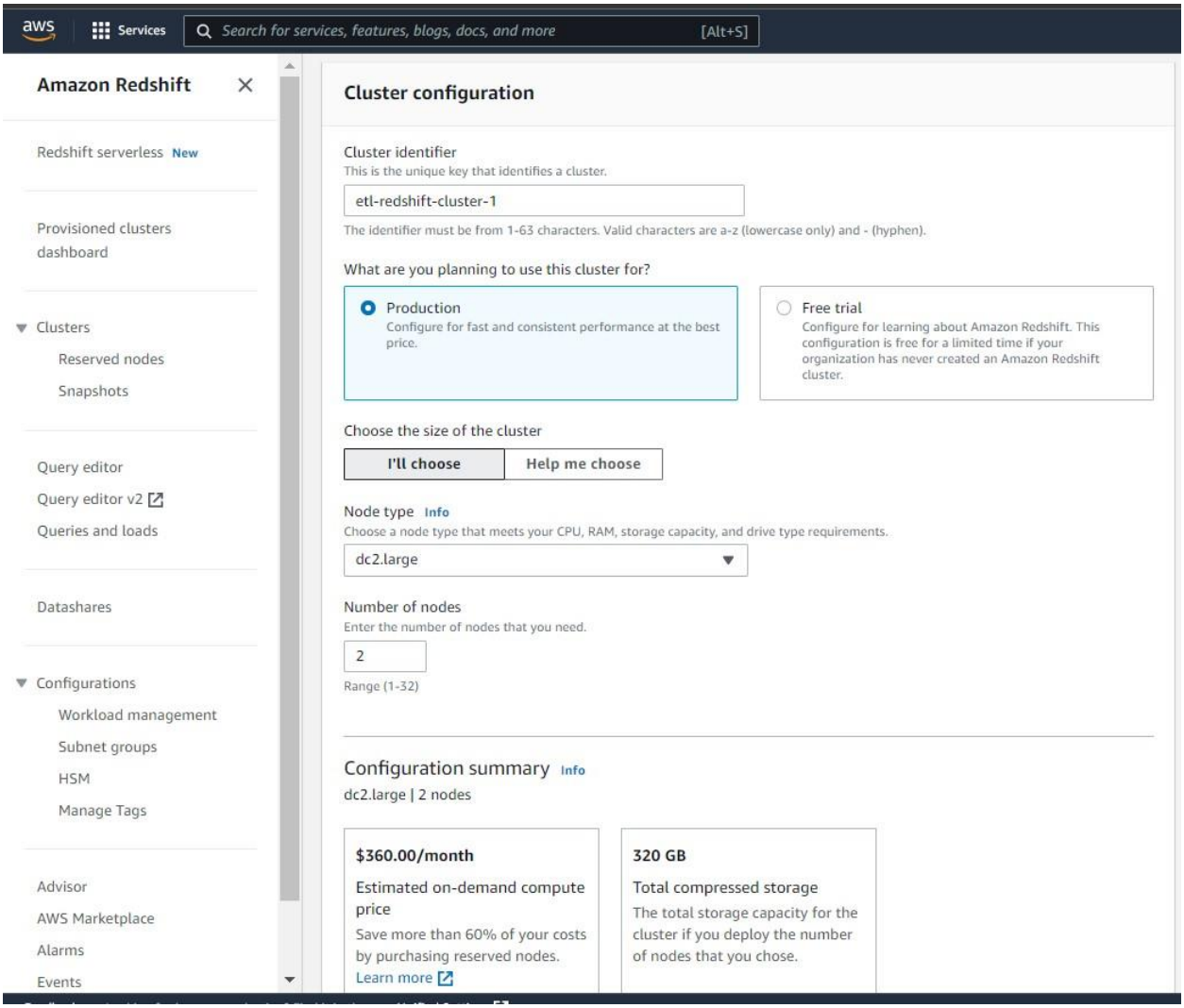
Creation of a Redshift Cluster

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

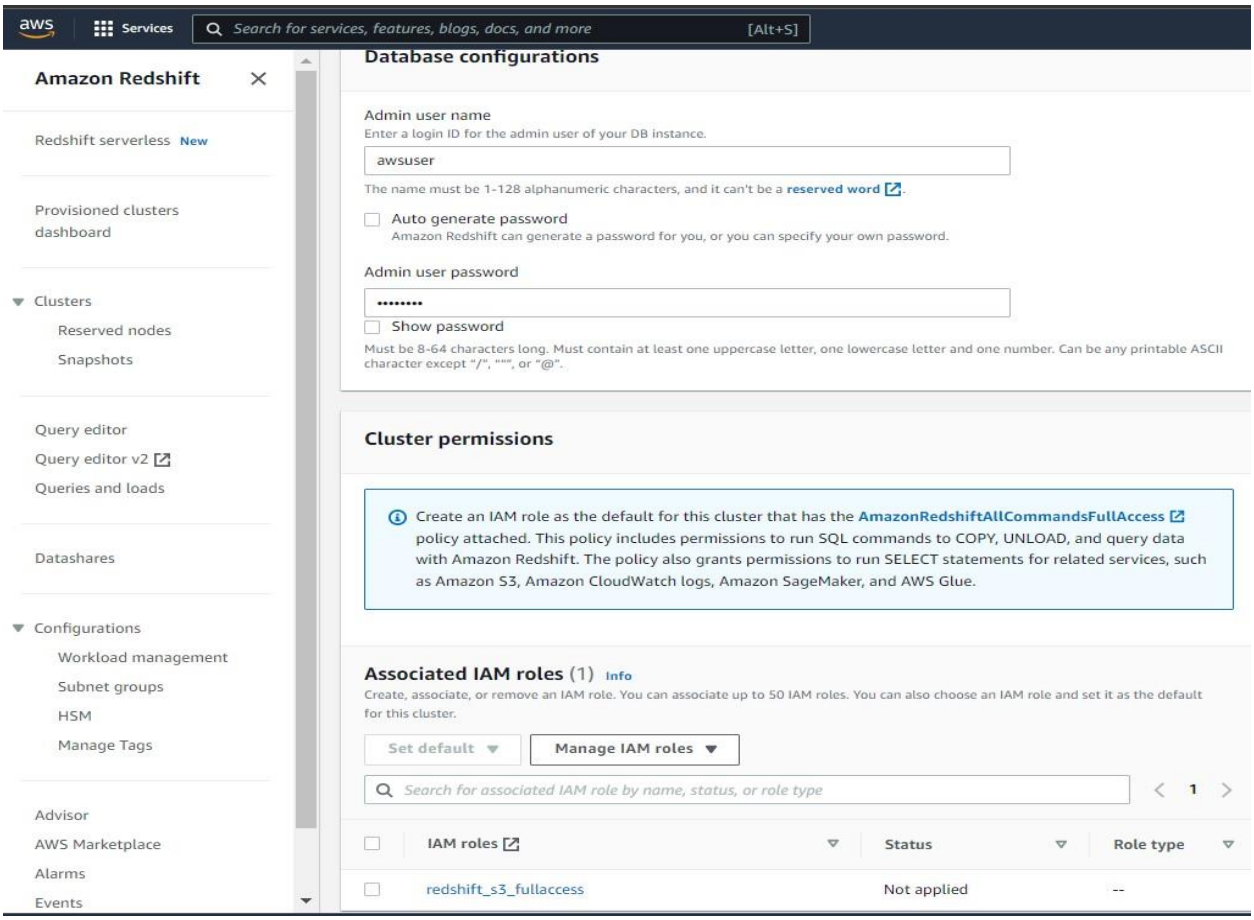
Below Screenshot shows the general information like type of node used and number of nodes used for setting up Redshift cluster.



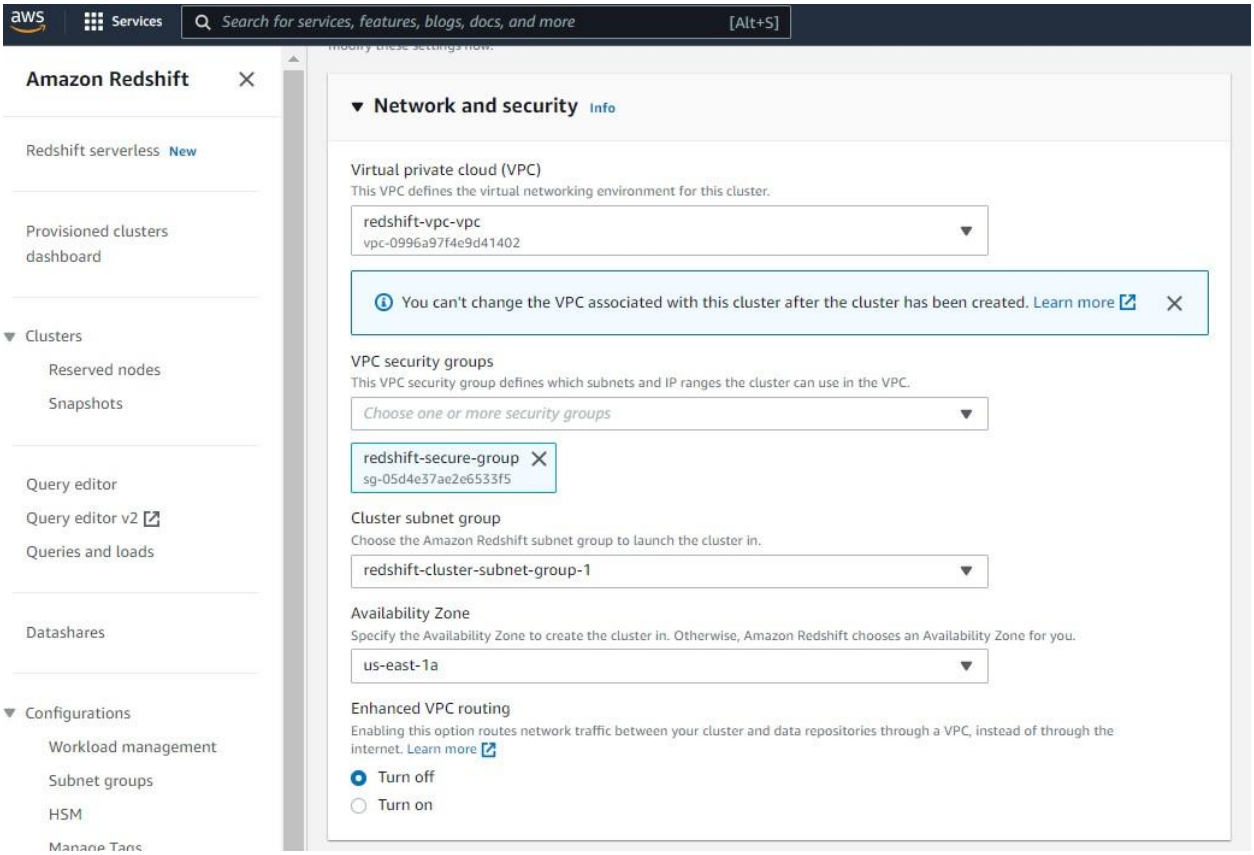
Screenshot of configurations made on cluster creation:

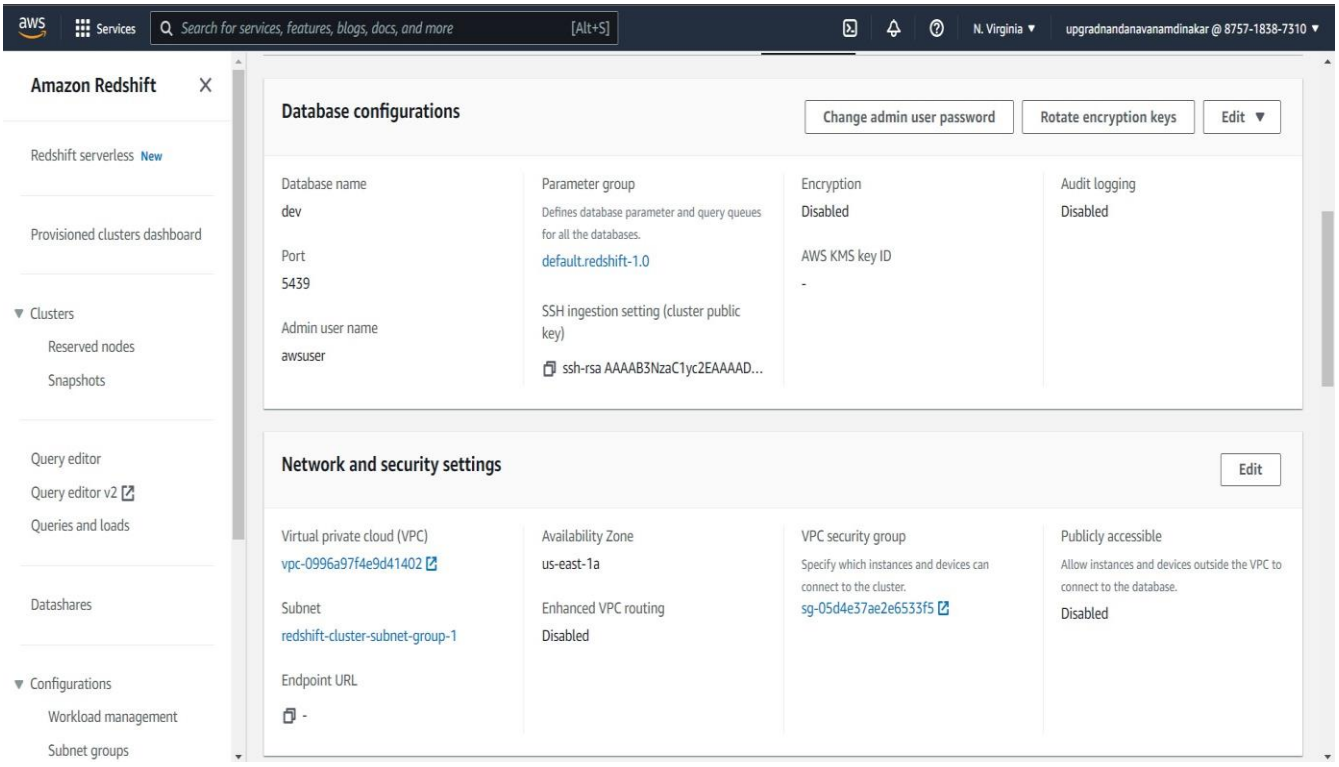


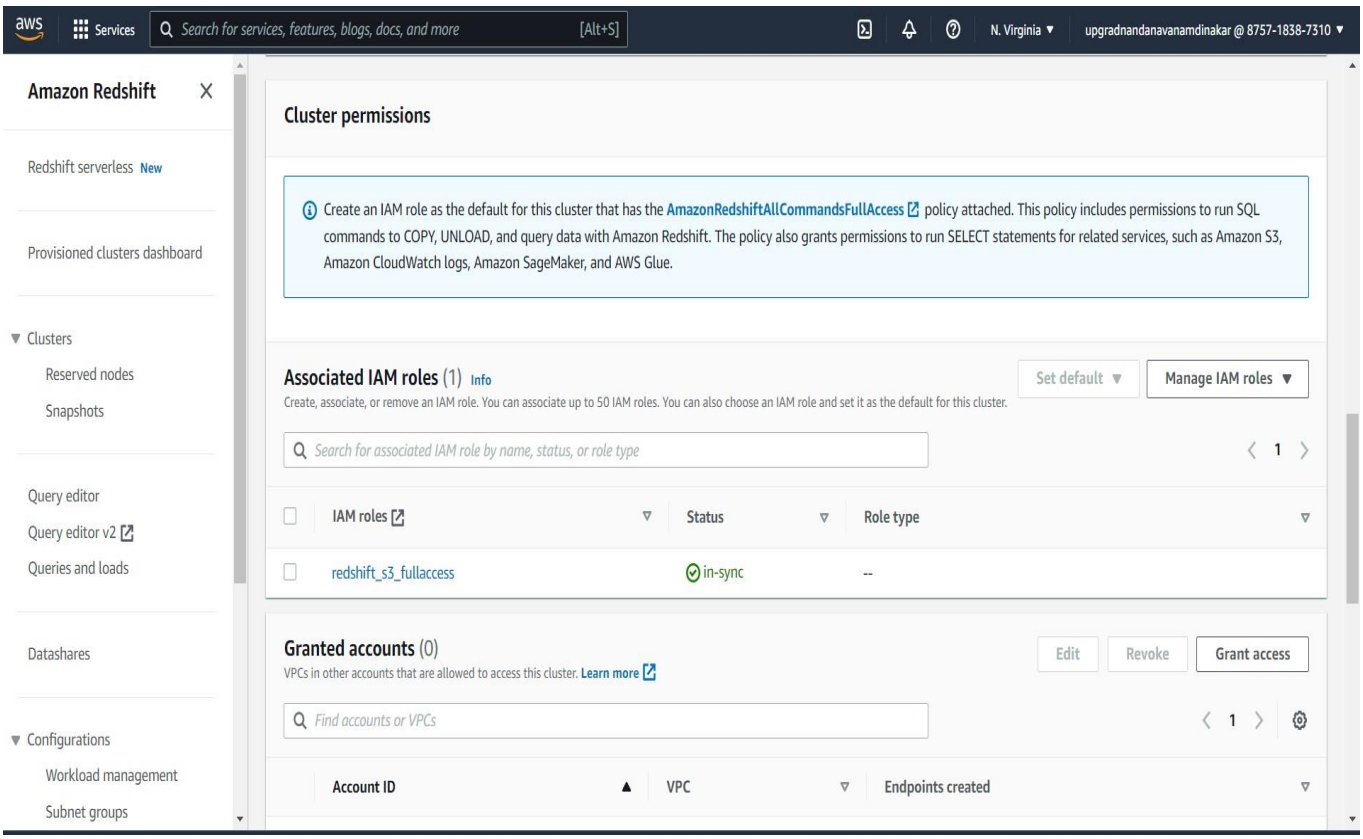
Screenshot showing Database configurations and IAM roles associated with the cluster:



The following screeshots shows the other configurations made on cluster:

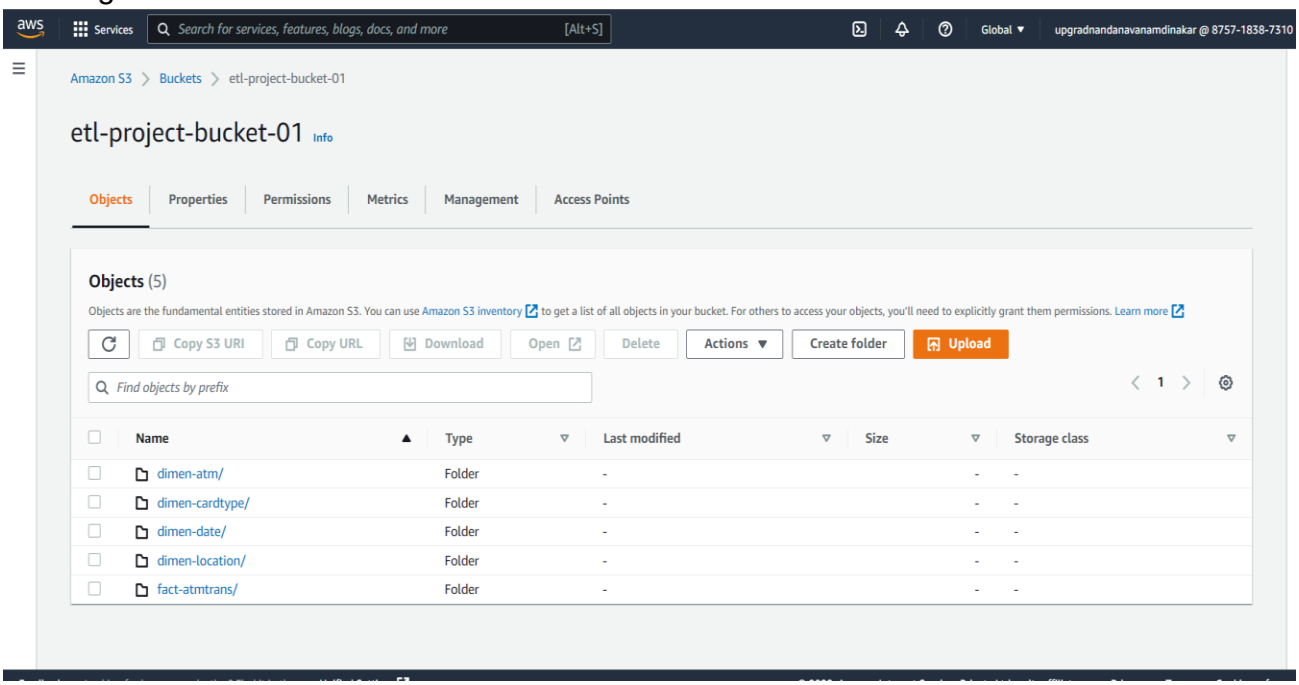






Setting up a database in the Redshift cluster and running queries to create the dimension and fact tables

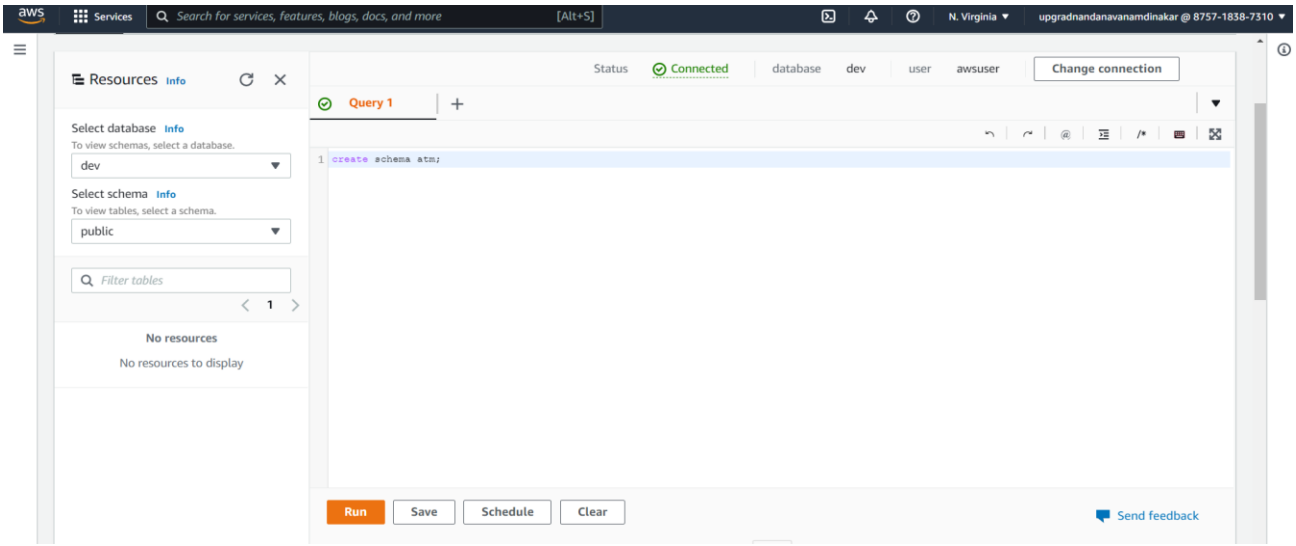
Viewing all the data in Amazon S3 Bucket:



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

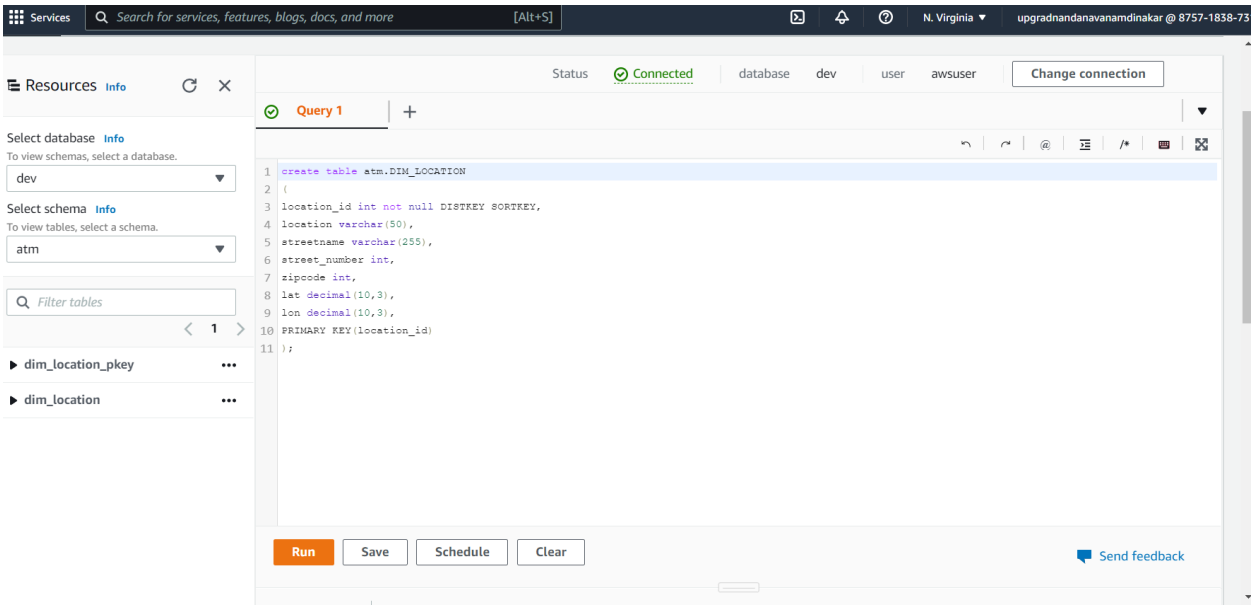
Query to create a schema for the dimension and fact tables:

create schema atm;



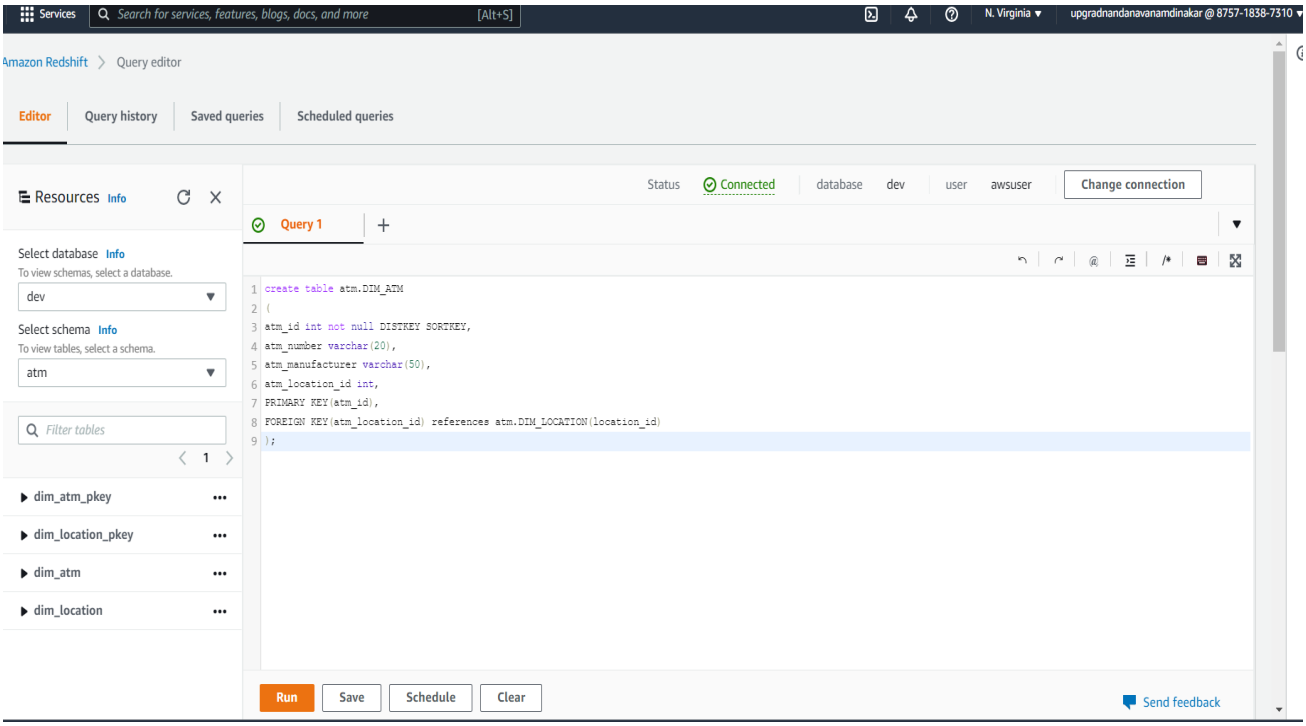
Creating location dimension table:

```
create table atm.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)
);
```



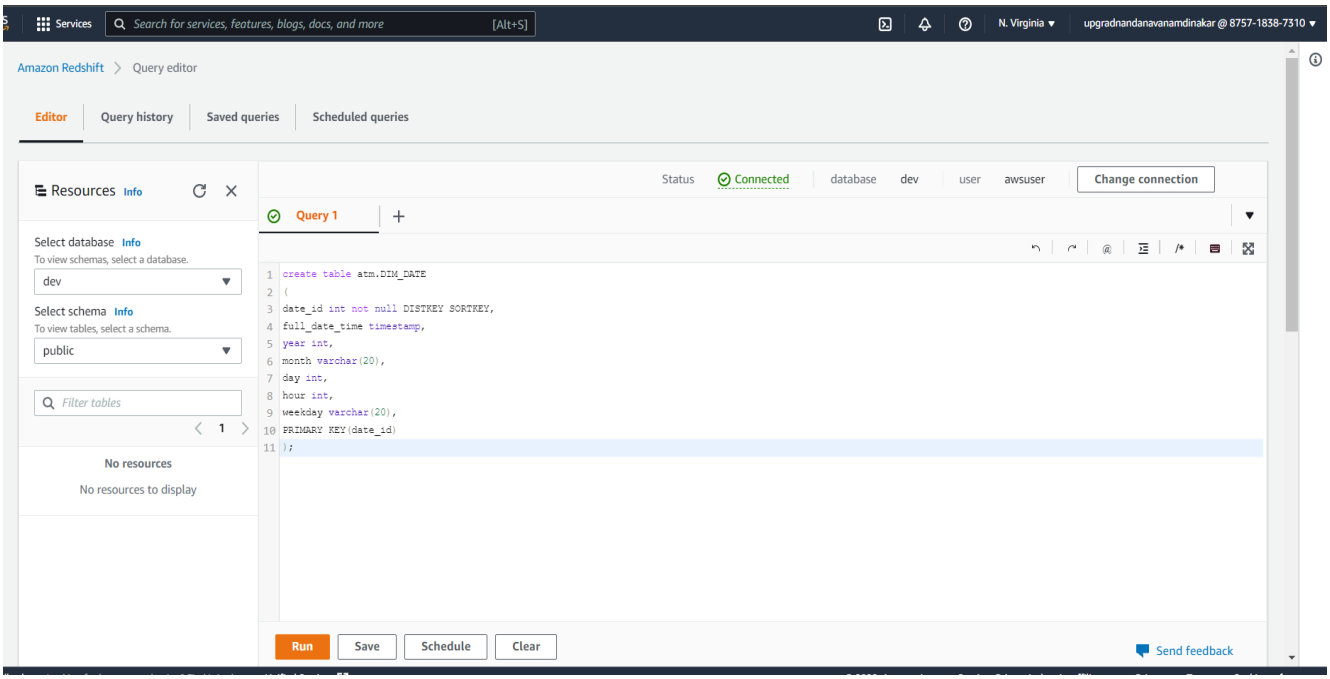
Creating ATM dimension table:

```
create table atm.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.DIM_LOCATION(location_id)
);
```



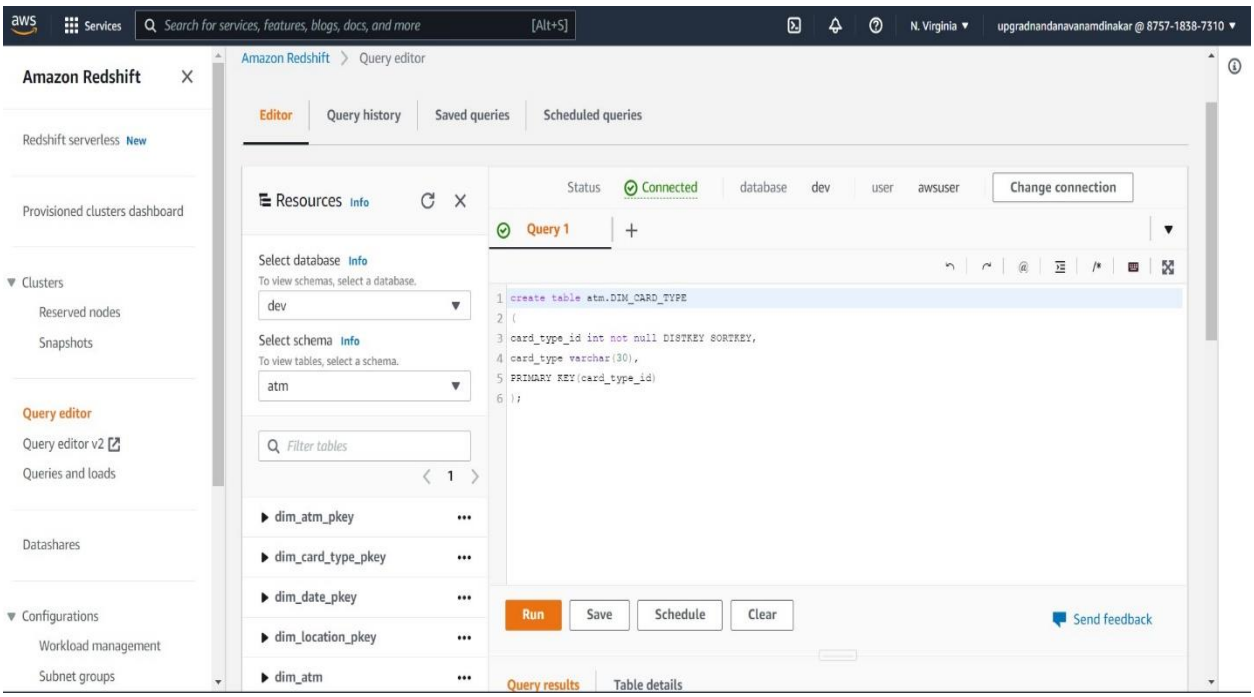
Creating date dimension table:

```
create table atm.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)
);
```



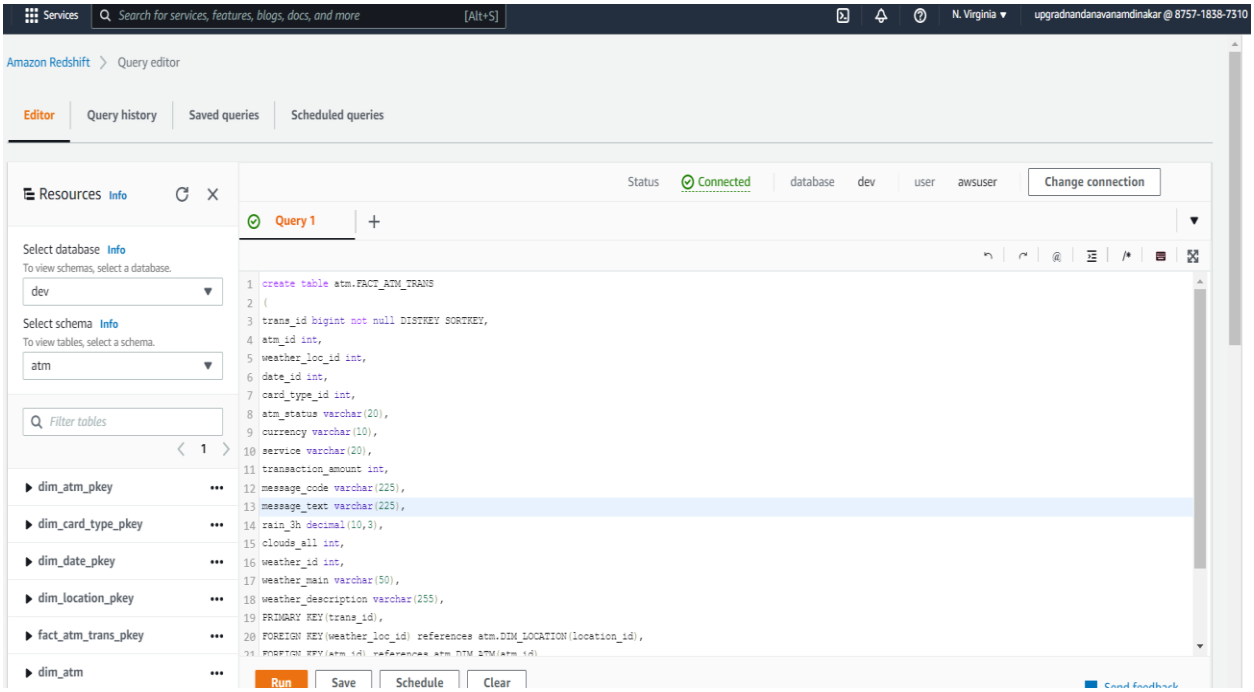
Creating Card type dimension table:

```
create table atm.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.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.DIM_LOCATION(location_id),
FOREIGN KEY(atm_id) references atm.DIM_ATM(atm_id),
FOREIGN KEY(date_id) references atm.DIM_DATE(date_id),
FOREIGN KEY(card_type_id) references atm.DIM_CARD_TYPE(card_type_id)
);
```

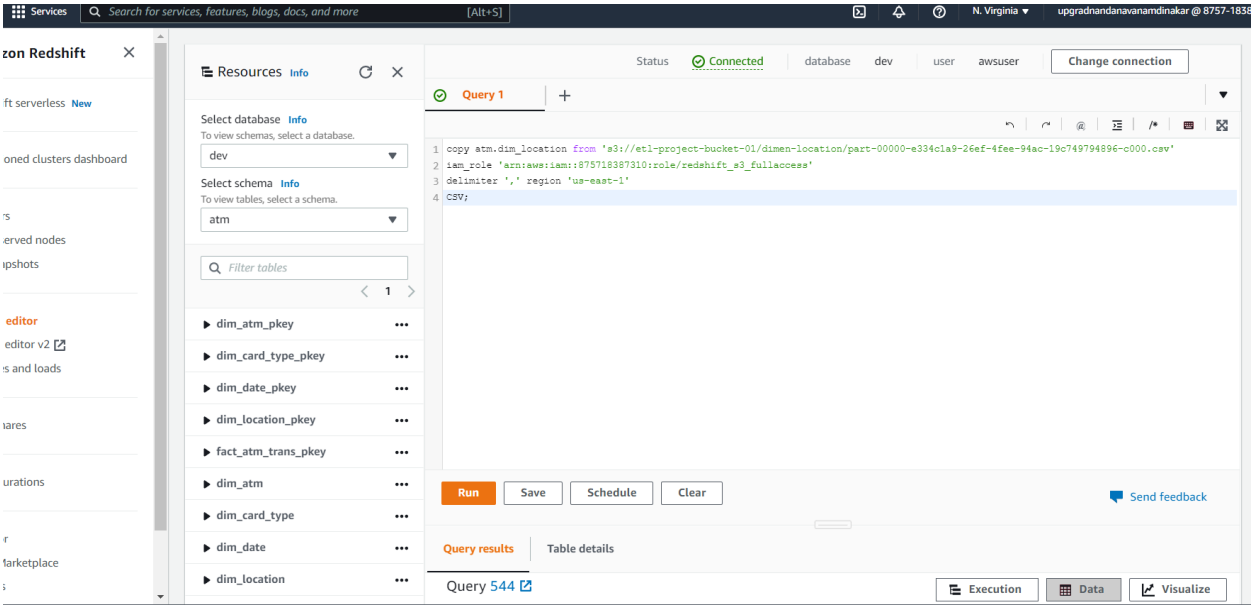


Loading data into a Redshift cluster from Amazon S3 bucket

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

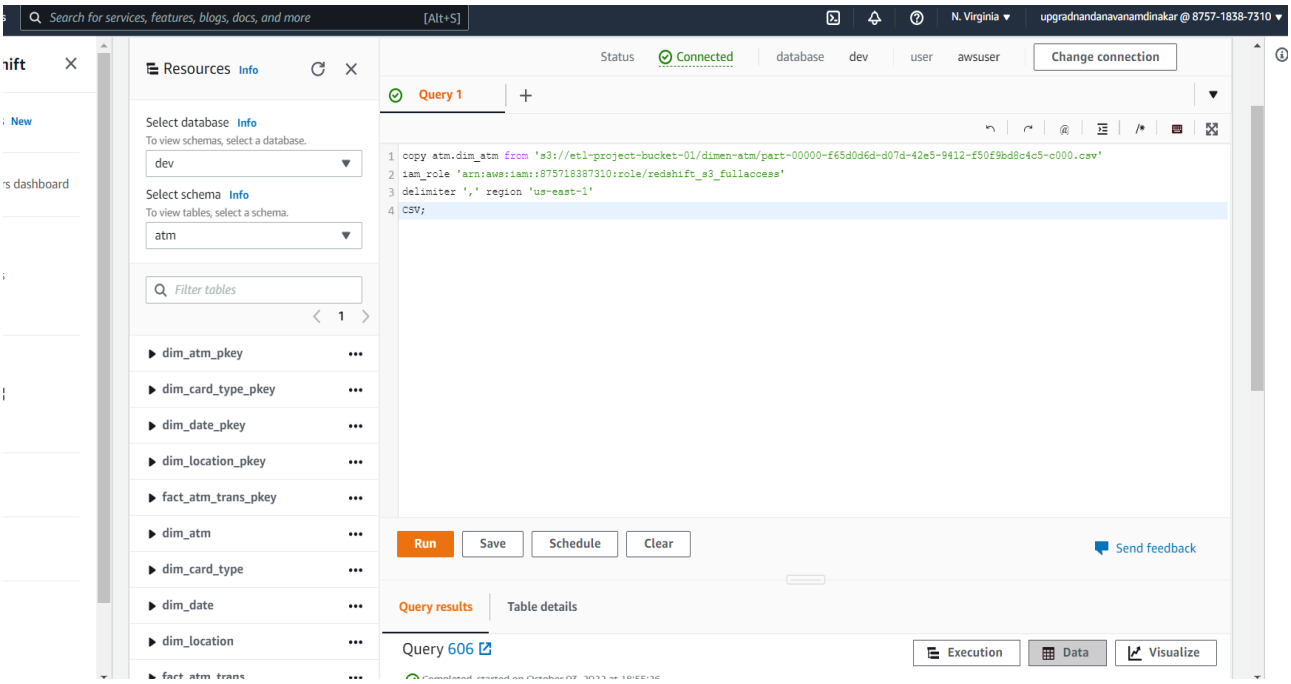
Copying the data from dimen-location folder on S3 bucket to dim_location table:

```
copy atm.dim_location from 's3://etl-project-bucket-01/dimen-location/part-00000-
e334c1a9-26ef-4fee-94ac-19c749794896-c000.csv'
iam_role 'arn:aws:iam::875718387310:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
CSV;
```



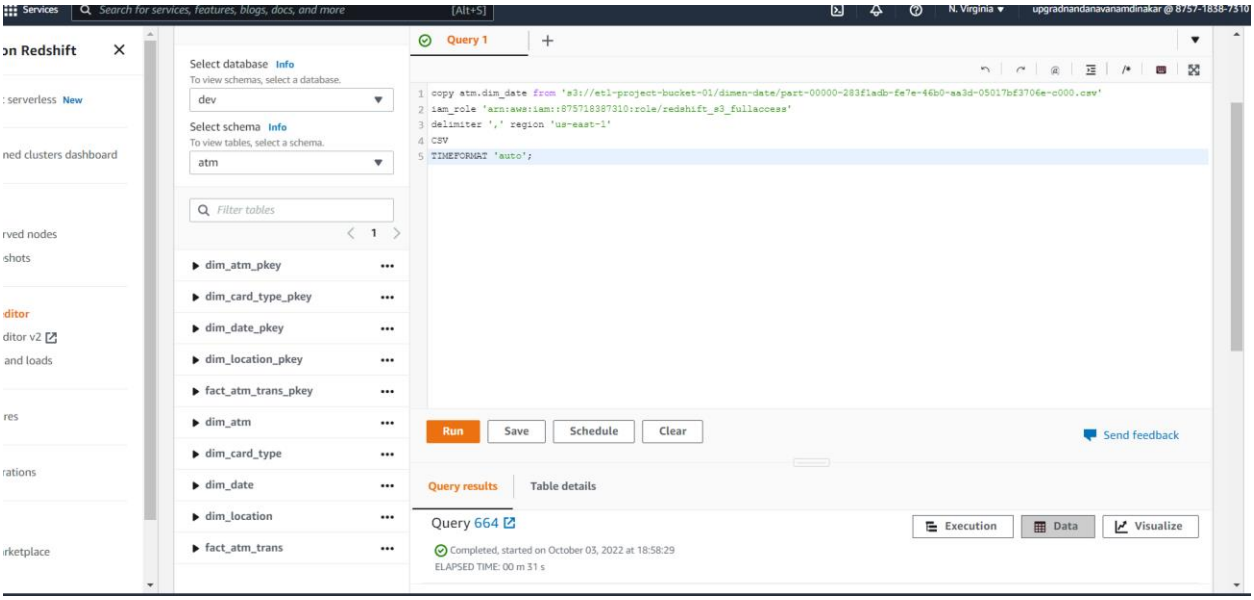
Copying the data from dimen-atm folder on S3 bucket to dim_atm table:

```
copy atm.dim_atm from 's3://etl-project-bucket-01/dimen-atm/part-00000-f65d0d6d-d07d-42e5-9412-f50f9bd8c4c5-c000.csv'
iam_role 'arn:aws:iam::875718387310:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
CSV;
```



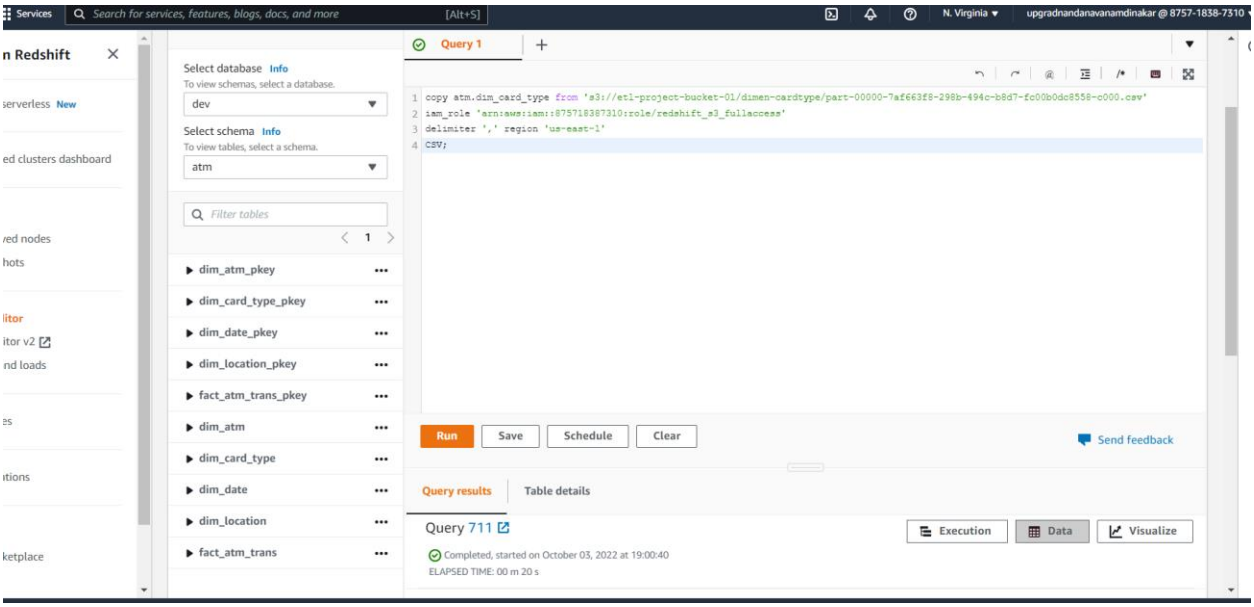
Copying data from dimen-data folder from S3 bucket to dim_date table:

```
copy atm.dim_date from 's3://etl-project-bucket-01/dimen-date/part-00000-283f1adb-fe7e-46b0-aa3d-05017bf3706e-c000.csv'
iam_role 'arn:aws:iam::875718387310:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
CSV
TIMEFORMAT 'auto';
```



Copying data from dimen-cardtype folder from S3 to dim_card_type table:

copy atm.dim_card_type from 's3://etl-project-bucket-01/dimen-cardtype/part-00000-7af663f8-298b-494c-b8d7-fc00b0dc8558-c000.csv'
iam_role 'arn:aws:iam::875718387310:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
CSV;



Copying data from fact-atmtrans folder in S3 to fact_atm_trans table:

```
copy atm.fact_atm_trans from 's3://etl-project-bucket-01/fact-atmtrans/part-00000-
0b34fdda-30f2-4ce7-b707-0467fa1c8255-c000.csv'
iam_role 'arn:aws:iam::875718387310:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
CSV;
```