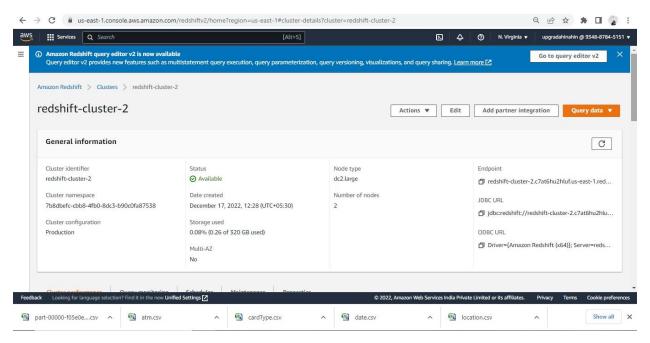
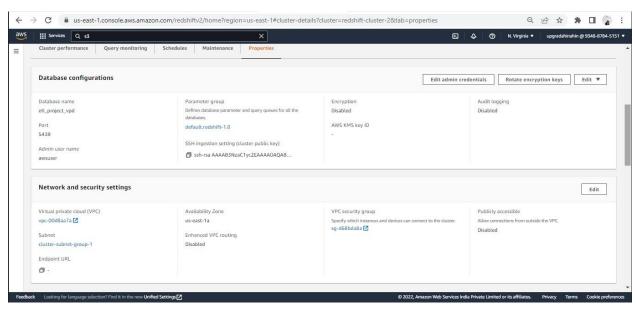
Creation of a Redshift Cluster

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

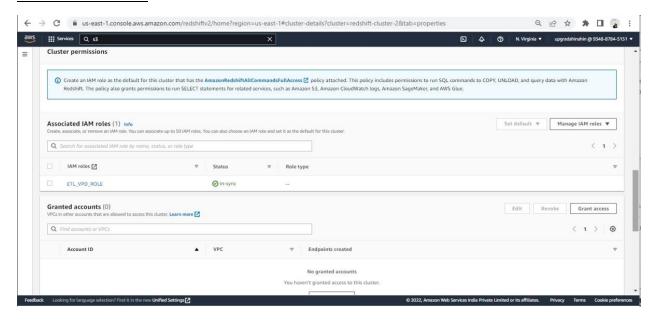
Cluster Configuration-



Database Configuration-

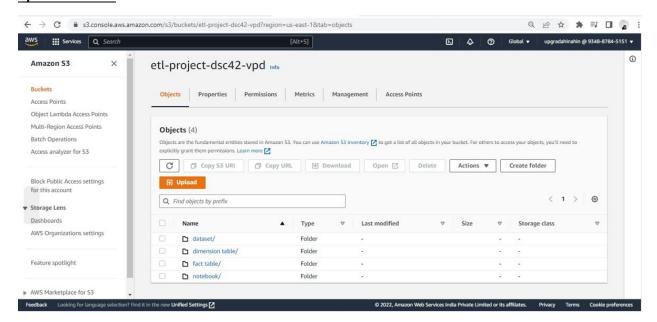


IAM roles associated-

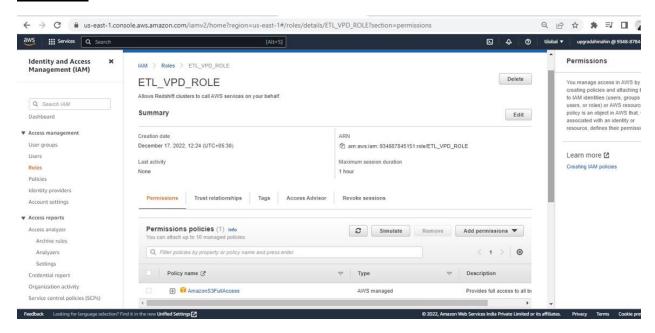


Setting up a database in the Redshift cluster and running queries to create the dimension and fact tables:

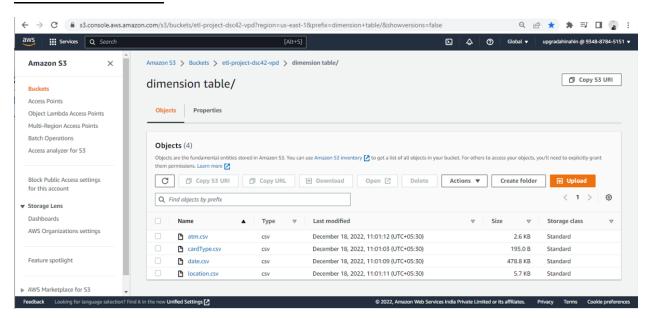
1) S3 Provision



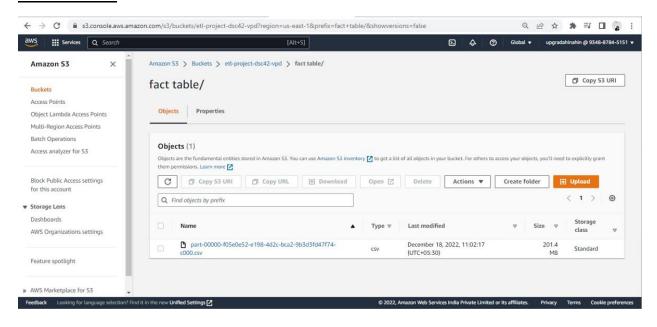
IAM Role



S3 Dimension table

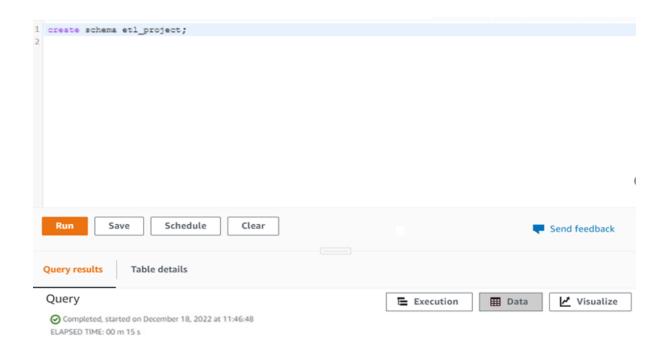


S3 Fact table



2) Creation of schema

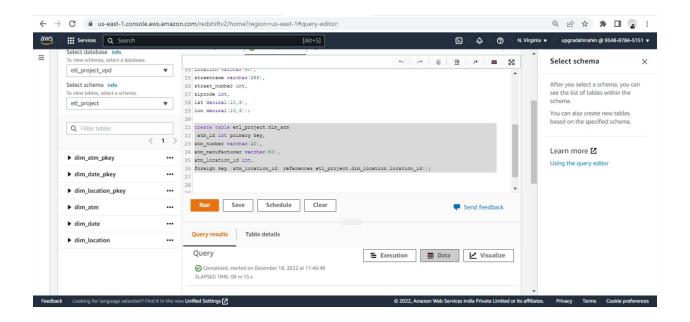
Create schema etl_project;



3) Queries to create the various dimension and fact tables

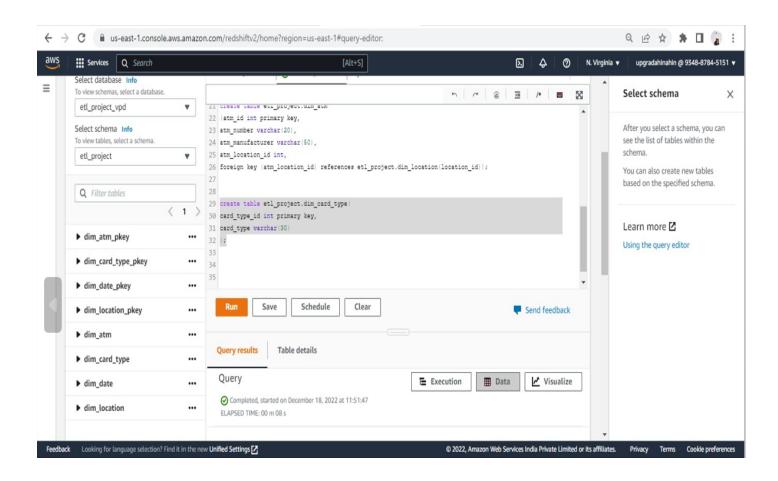
DIM_ATM

```
create table etl_project.dim_atm(
    atm_id int PRIMARY KEY,
    atm_number varchar(20),
    atm_manufacturer varchar(50),
    atm_location_id int,
    FOREIGN KEY(atm_location_id) references etl_project.DIM_LOCATION(location_id)
);
```



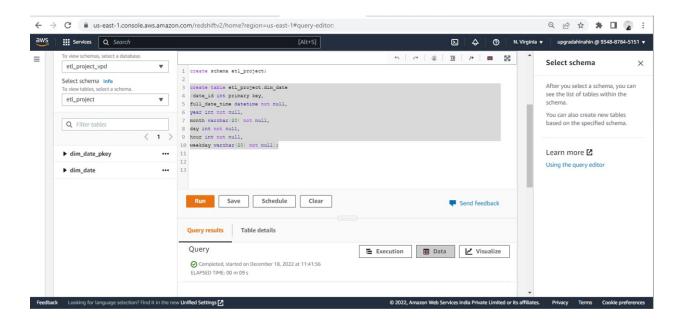
DIM_CARD_TYPE

```
create table etl_project.dim_card_type
(
         card_type_id int PRIMARY KEY,
         card_type varchar(30)
);
```



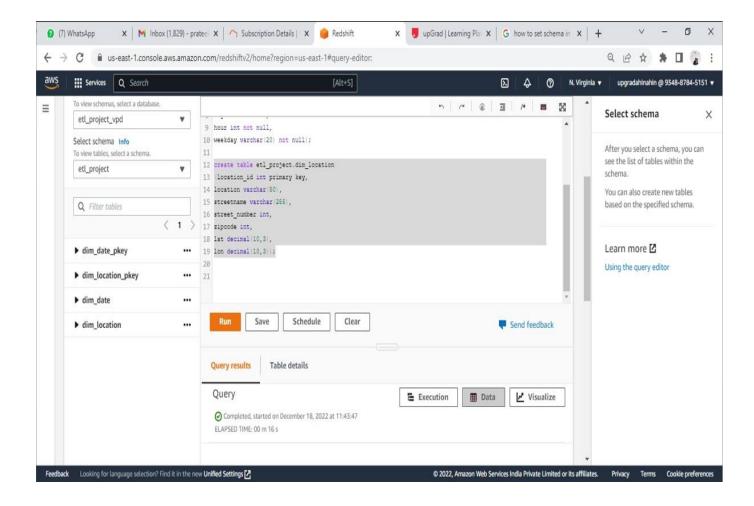
DIM_DATE

```
create table etl_project.dim_date
(
    date_id int PRIMARY KEY,
    full_date_time datetime not null,
    year int not null,
    month varchar(20) not null,
    day int not null,
    hour int not null,
    weekday varchar(20) not null
);
```



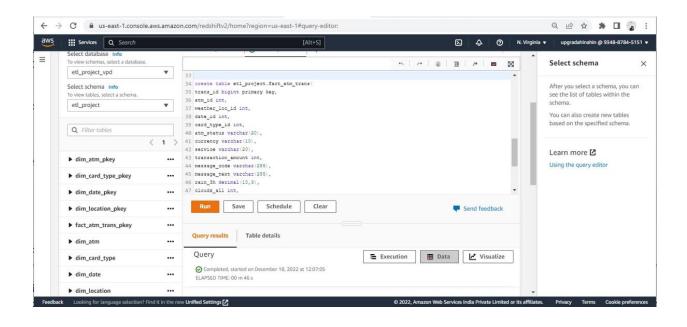
DIM_LOCATION

```
create table etl_project.dim_location
(
    location_id int PRIMARY KEY,
    location varchar(50),
    streetname varchar(255),
    street_number int,
    zipcode int,
    lat decimal(10,3),
    lon decimal(10,3),
);
```



FACT_ATM_TRANS

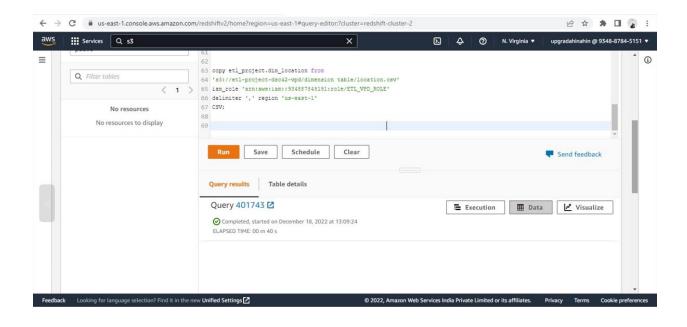
```
create table etl project.fact atm trans
(
       trans id bigint PRIMARY KEY,
       atm id int,
       weather loc id int,
       date id int,
       card type id int,
       atm_status varchar(20),
       currency varchar(10),
       service varchar(20),
       transaction amount int,
      message 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 etl project.dim location(location id),
       FOREIGN KEY(atm id) REFERENCES etl project.dim atm (atm id),
       FOREIGN KEY(date id) REFERENCES etl project.dim date (date id)
);
```



4) Loading data into a Redshift cluster from Amazon S3 bucket

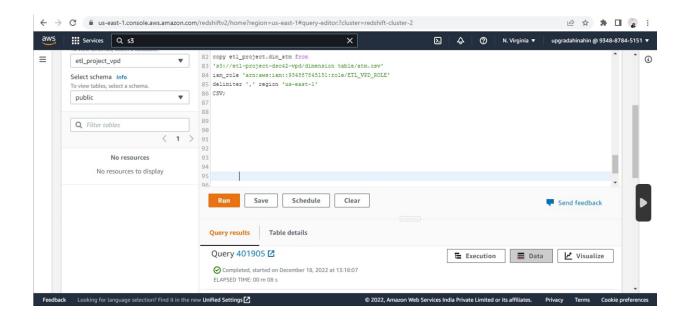
Copying the data to dim_location table

```
copy etl_project.dim_location from
's3://etl-project-dsc42-vpd/dimension table/location.csv'
iam_role 'arn:aws:iam::934887845151:role/ETL_VPD_ROLE'
delimiter ',' region 'us-east-1'
CSV;
```



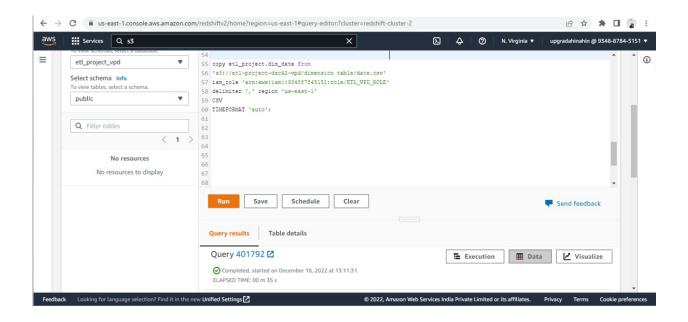
Copying the data to dim_atm table

```
copy etl_project.dim_atm from
's3://etl-project-dsc42-vpd/dimension table/atm.csv'
iam_role 'arn:aws:iam::934887845151:role/ETL_VPD_ROLE'
delimiter ',' region 'us-east-1'
CSV;
```



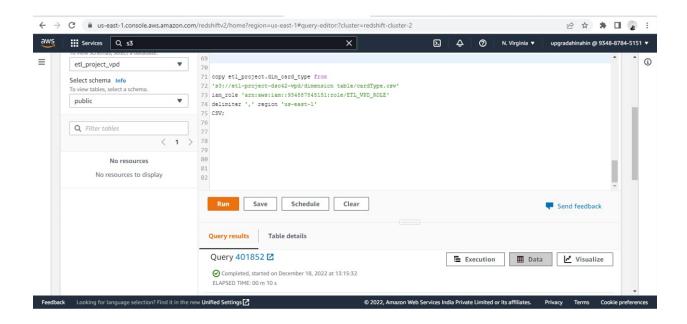
Copying the data to dim_date table

```
copy etl_project.dim_date from
's3://etl-project-dsc42-vpd/dimension table/date.csv'
iam_role 'arn:aws:iam::934887845151:role/ETL_VPD_ROLE'
delimiter ',' region 'us-east-1'
CSV
TIMEFORMAT 'auto';
```



Copying the data to dim_card_type table

```
copy etl_project.dim_card_type from
's3://etl-project-dsc42-vpd/dimension table/cardType.csv'
iam_role 'arn:aws:iam::934887845151:role/ETL_VPD_ROLE'
delimiter ',' region 'us-east-1'
CSV;
```



Copying the data to fact_atm_trans table

copy etl_project.fact_atm_trans from
's3://etl-project-dsc42-vpd/fact table/part-00000-f05e0e52-e198-4d2c-bca2-9b3d3fd47f74-c000.csv'
iam_role 'arn:aws:iam::934887845151:role/ETL_VPD_ROLE'
delimiter ',' region 'us-east-1'
CSV;

