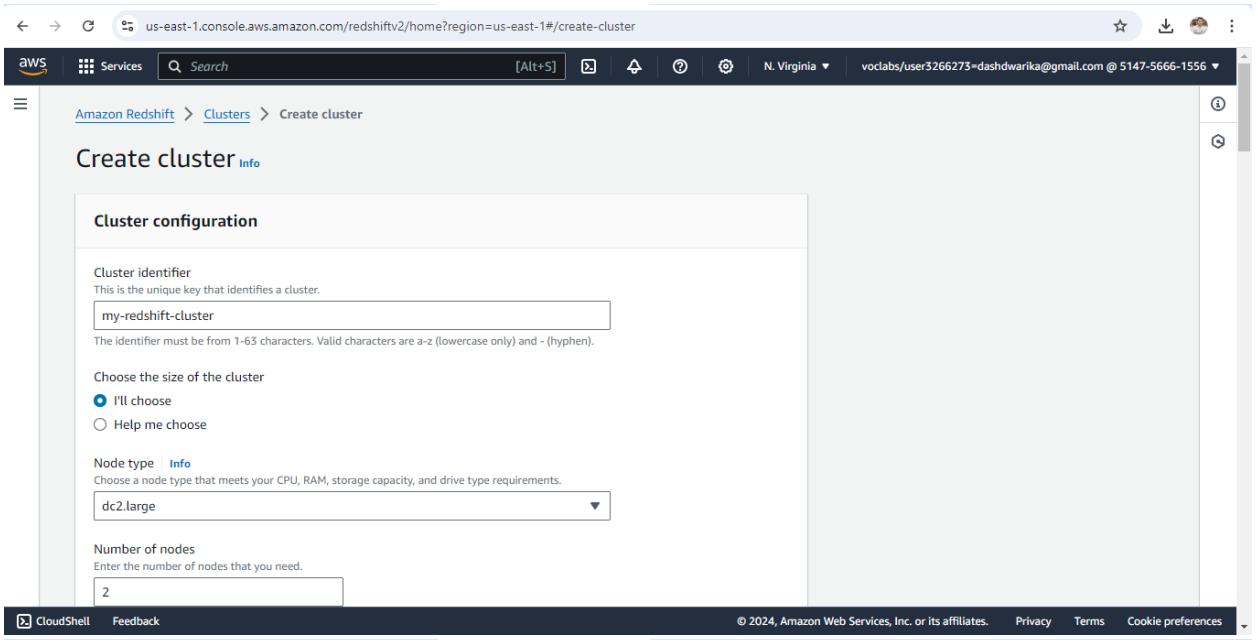
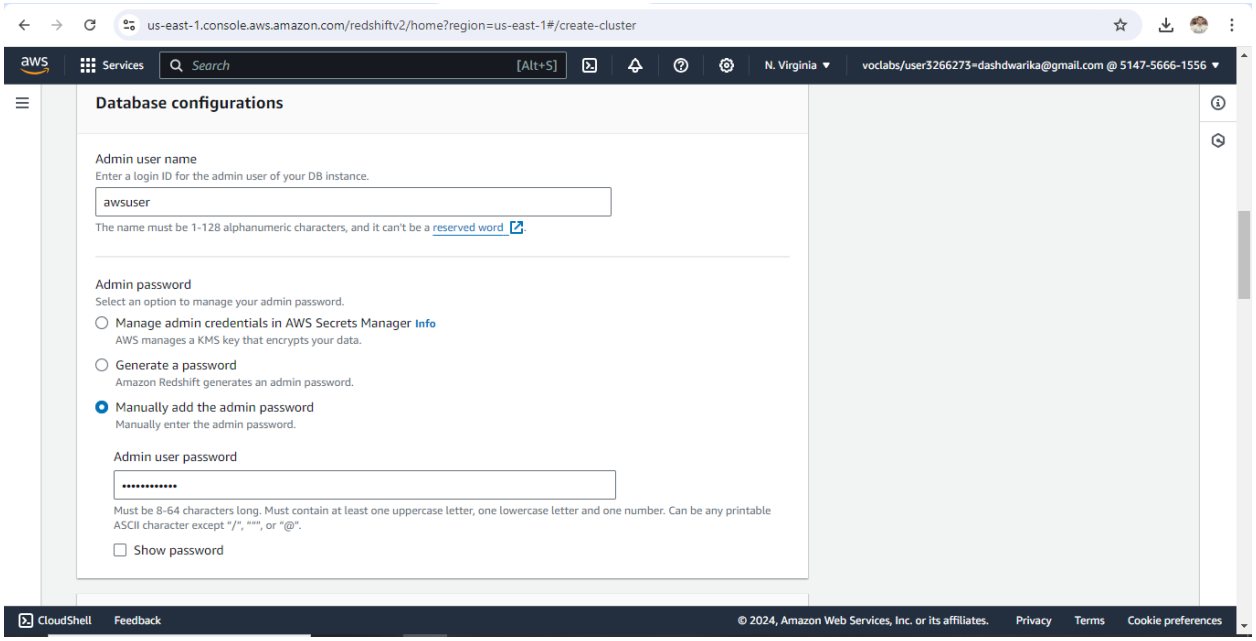
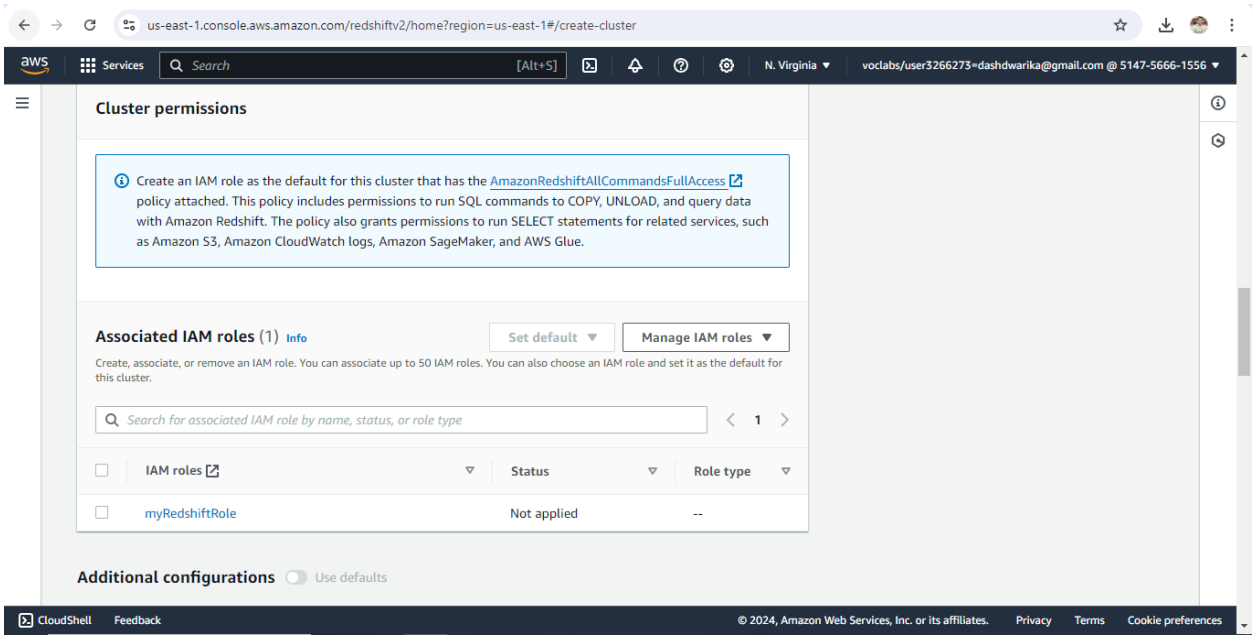
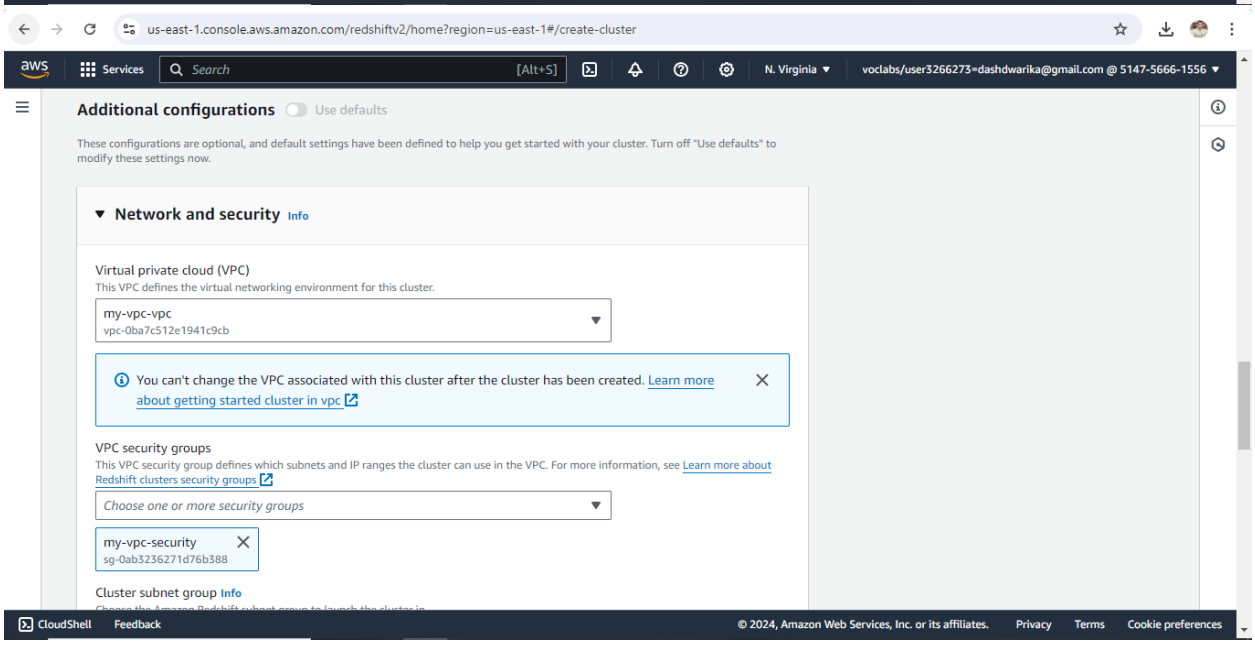
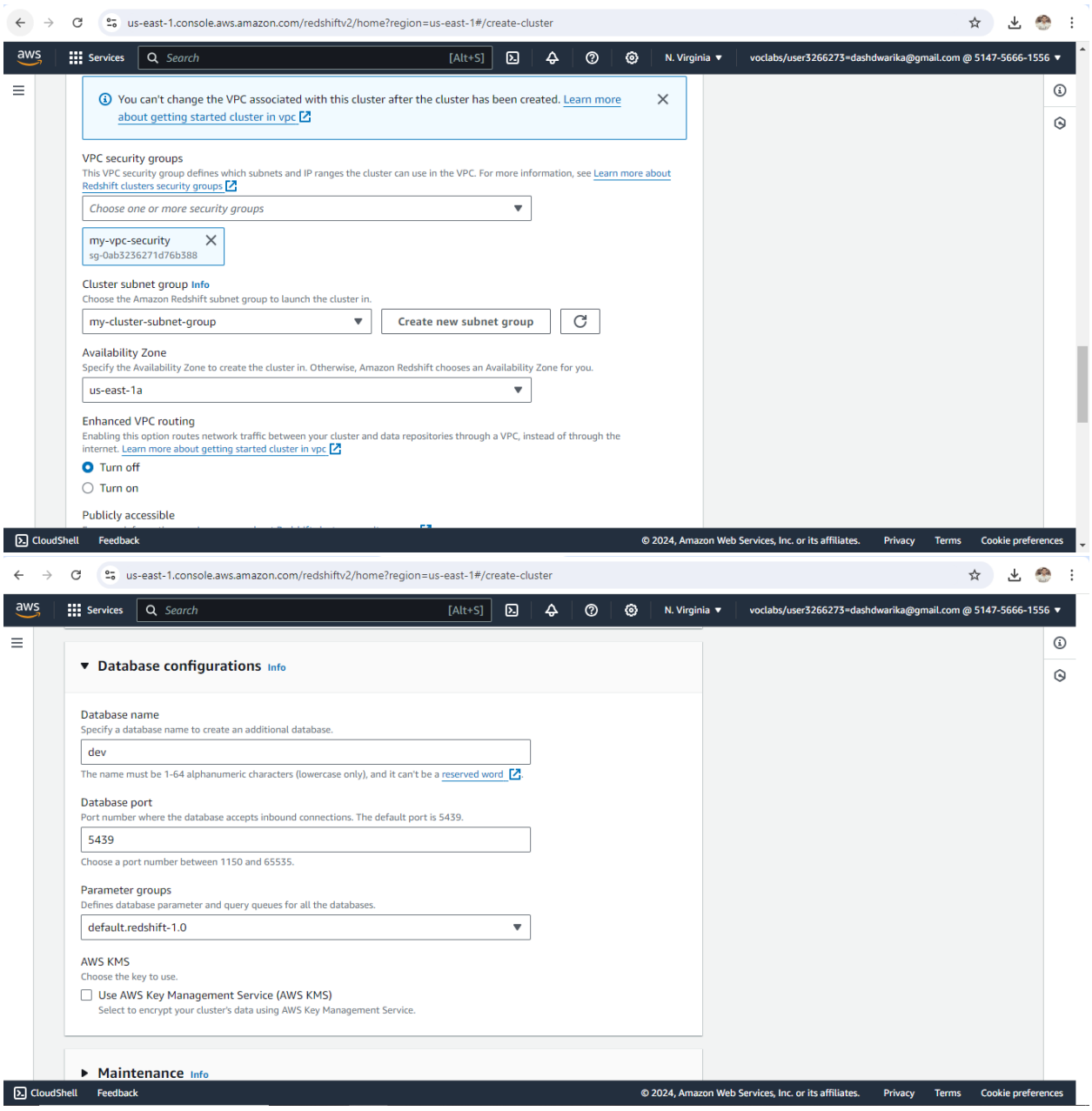
Creation of a Redshift Cluster

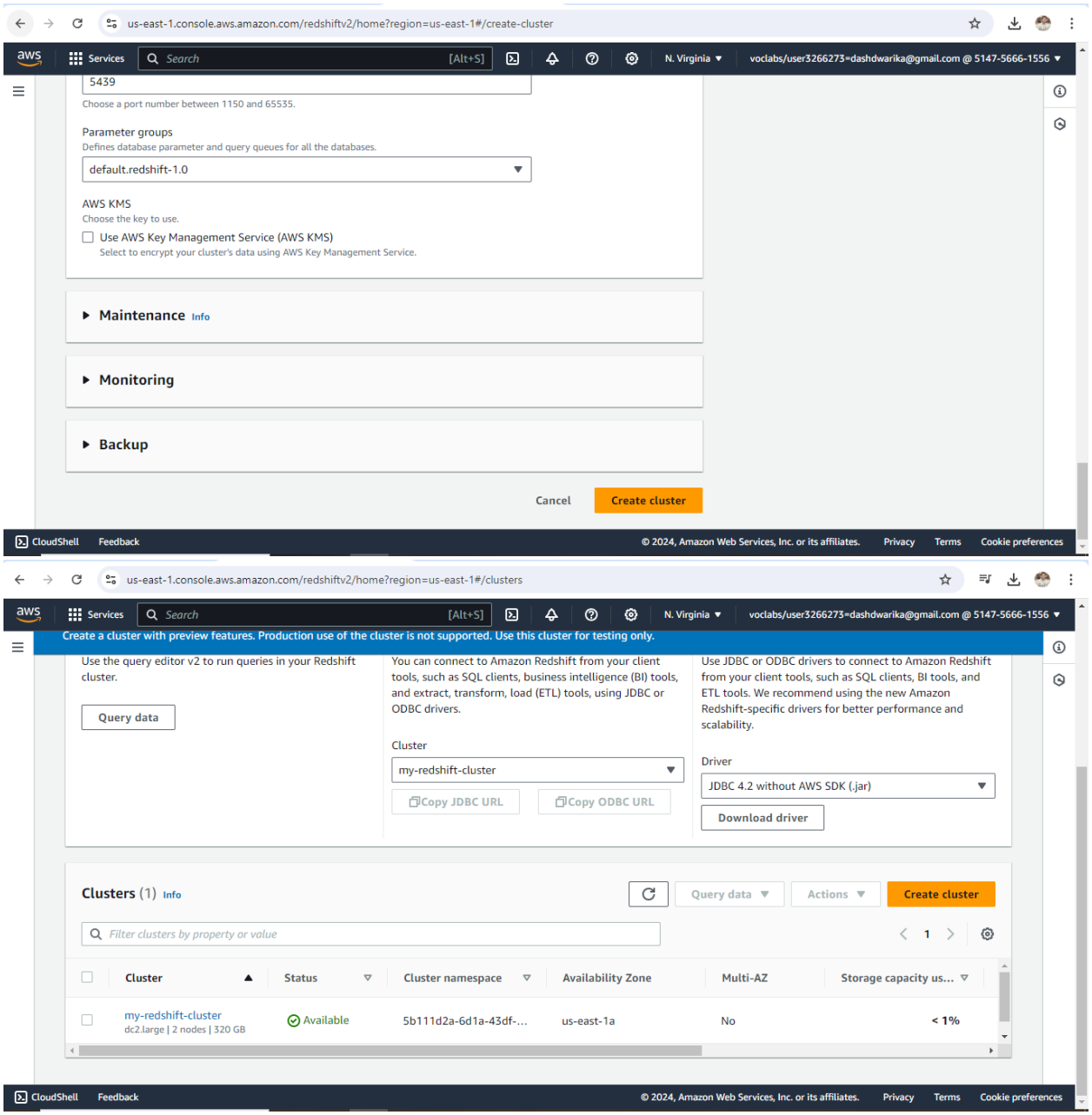




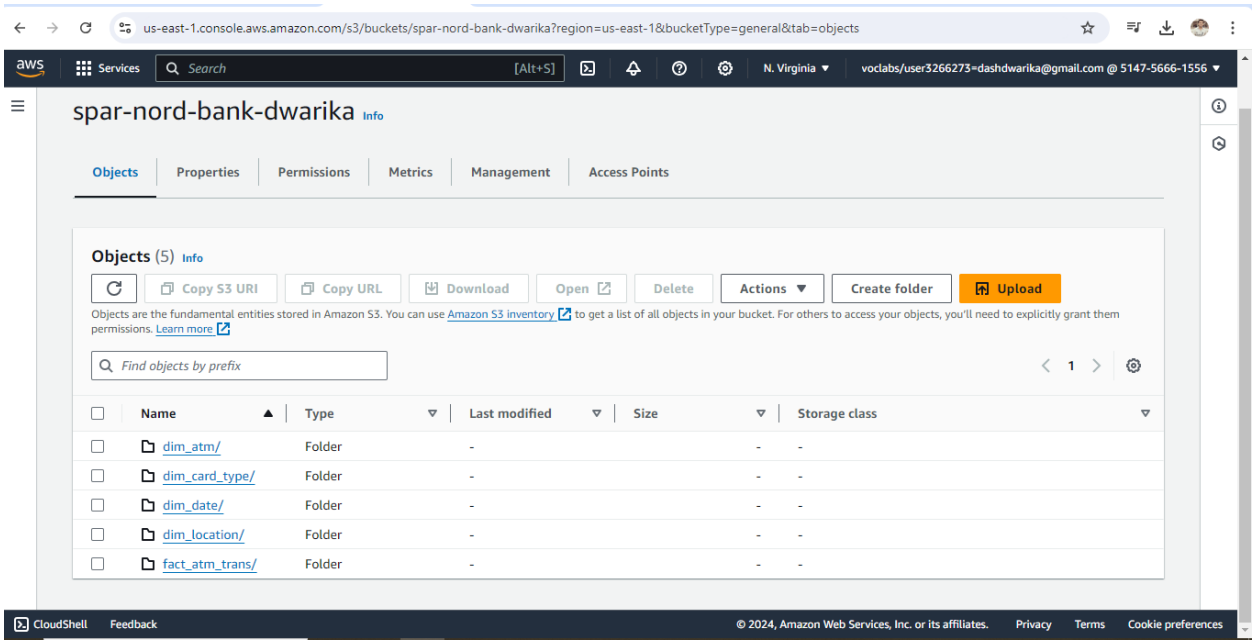








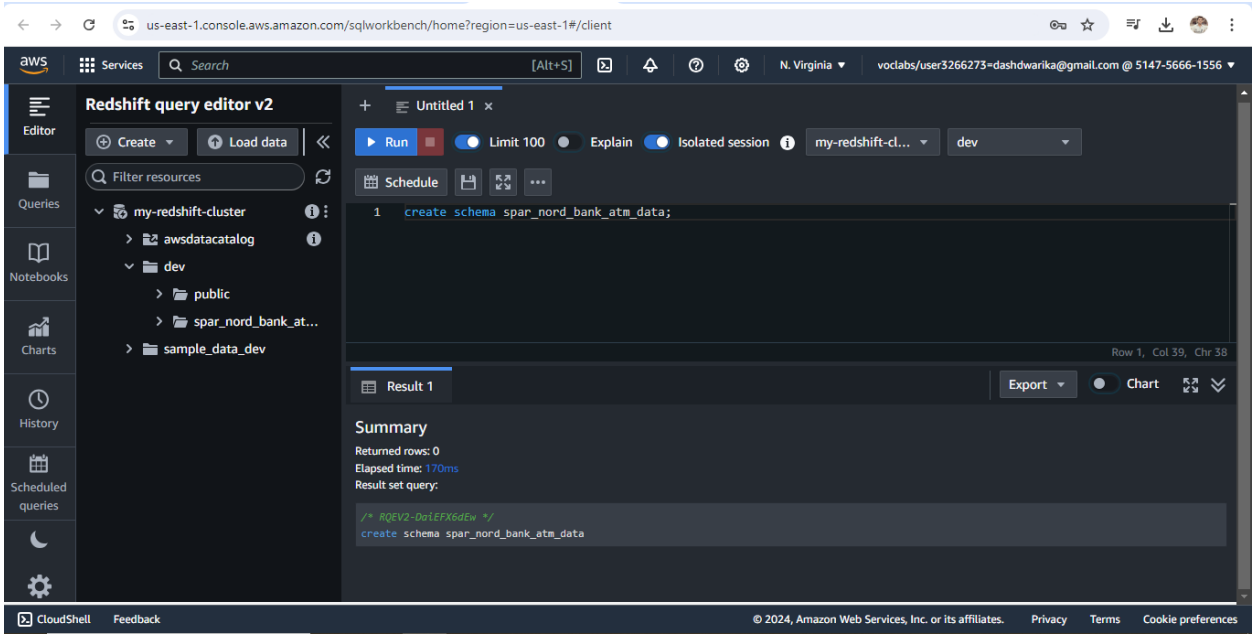
S3 bucket containing files:



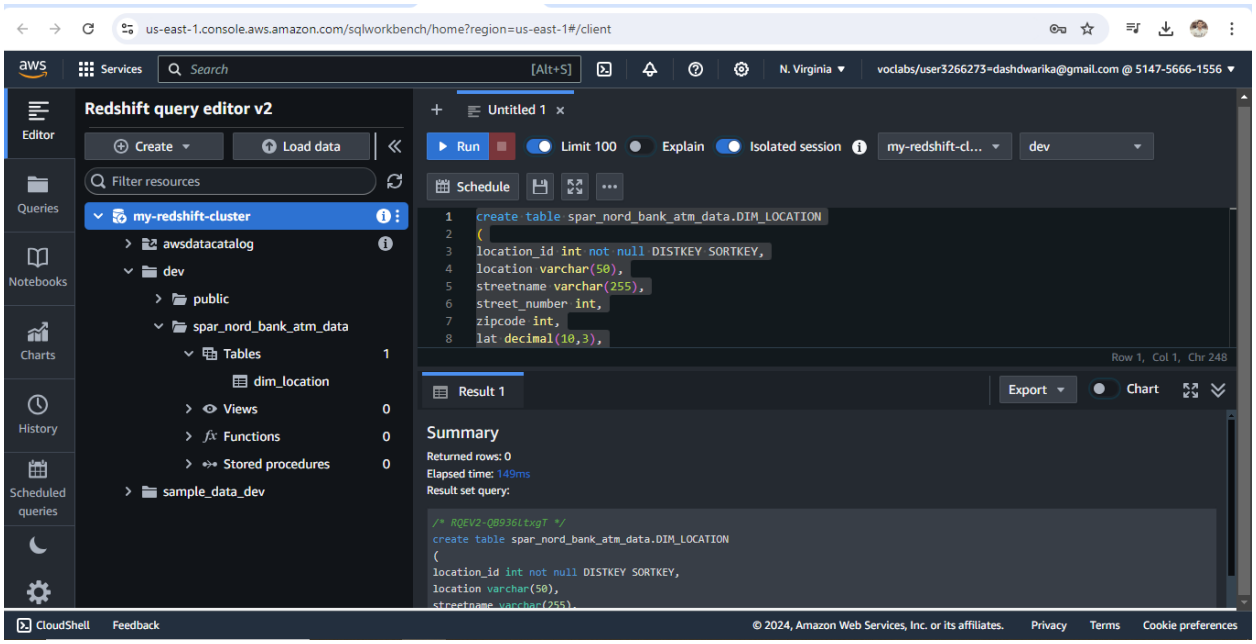
Setting up a database in the Redshift cluster and running queries to create the dimension and fact tables

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

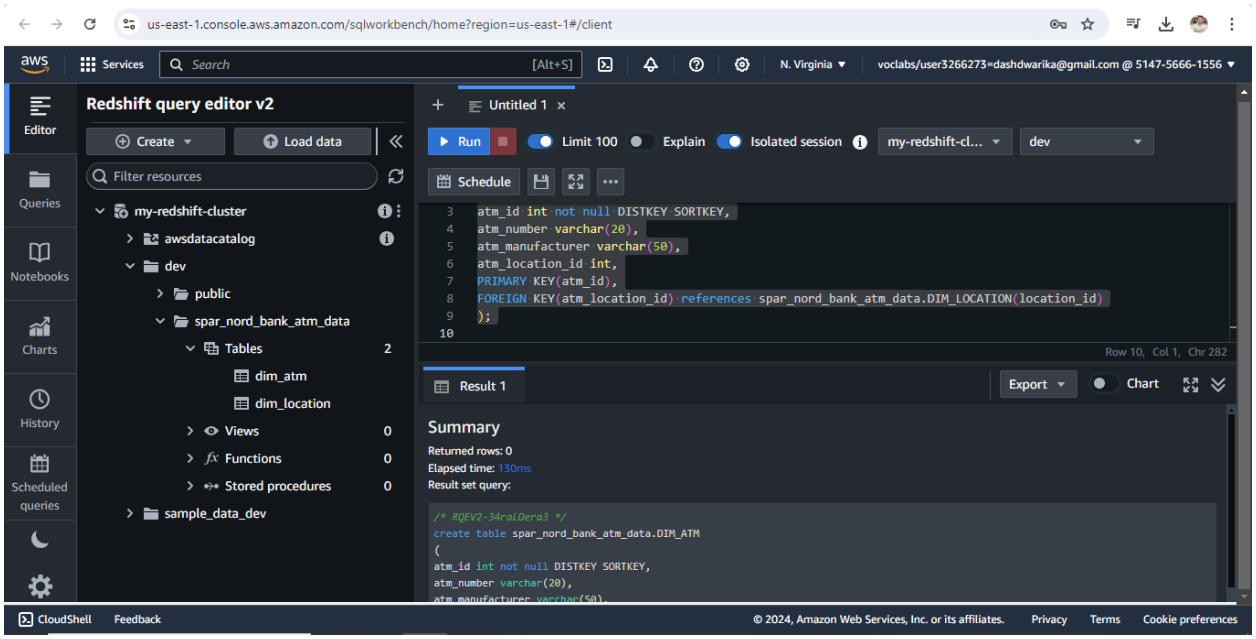
```
create schema spar_nord_bank_atm_data;
```



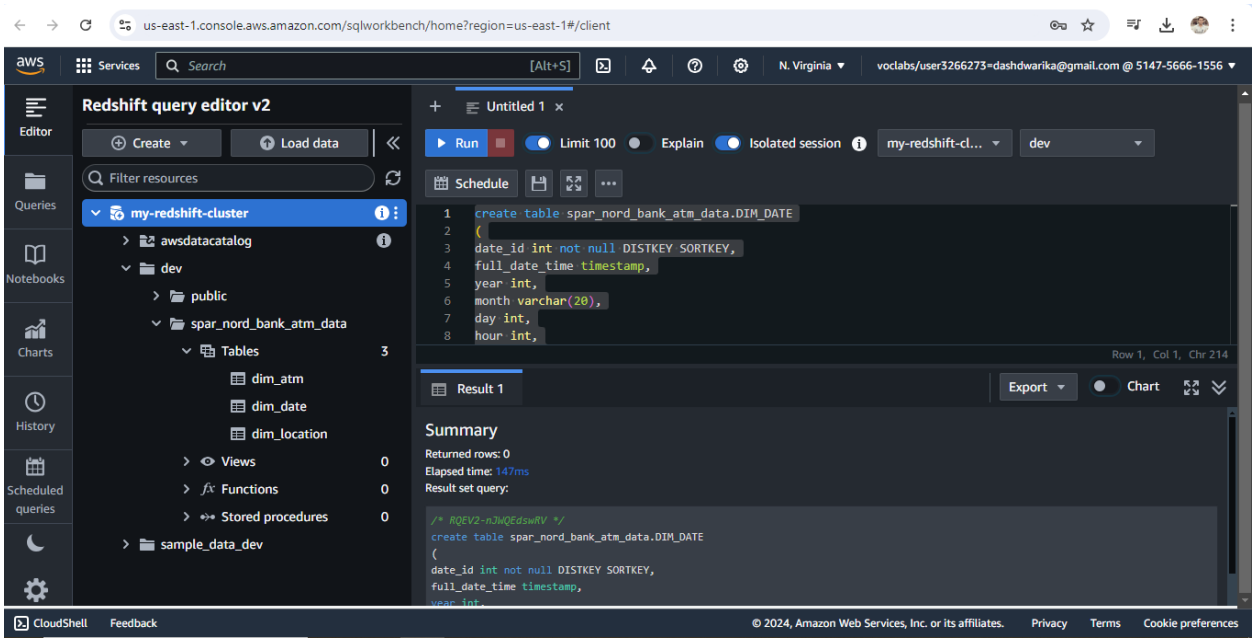

```
create table spar_nord_bank_atm_data.DIM_LOCATION
(
location_id int not null DISTKEY SORTKEY,
location varchar(50),
streetname varchar(255),
street_number int,
zipcode int,
lat decimal(10,3),
lon decimal(10,3),
PRIMARY KEY(location_id)
);
```



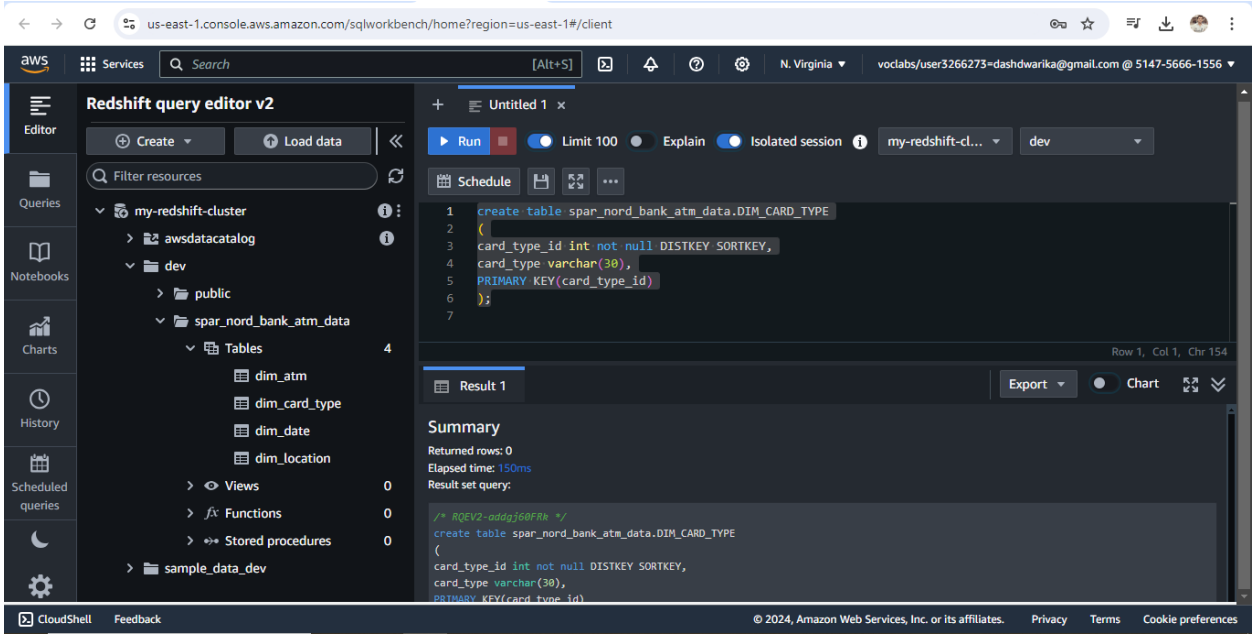

```
create table spar_nord_bank_atm_data.DIM_ATM
(
atm_id int not null DISTKEY SORTKEY,
atm_number varchar(20),
atm_manufacturer varchar(50),
atm_location_id int,
PRIMARY KEY(atm_id),
FOREIGN KEY(atm_location_id) references
spar_nord_bank_atm_data.DIM_LOCATION(location_id)
);
```



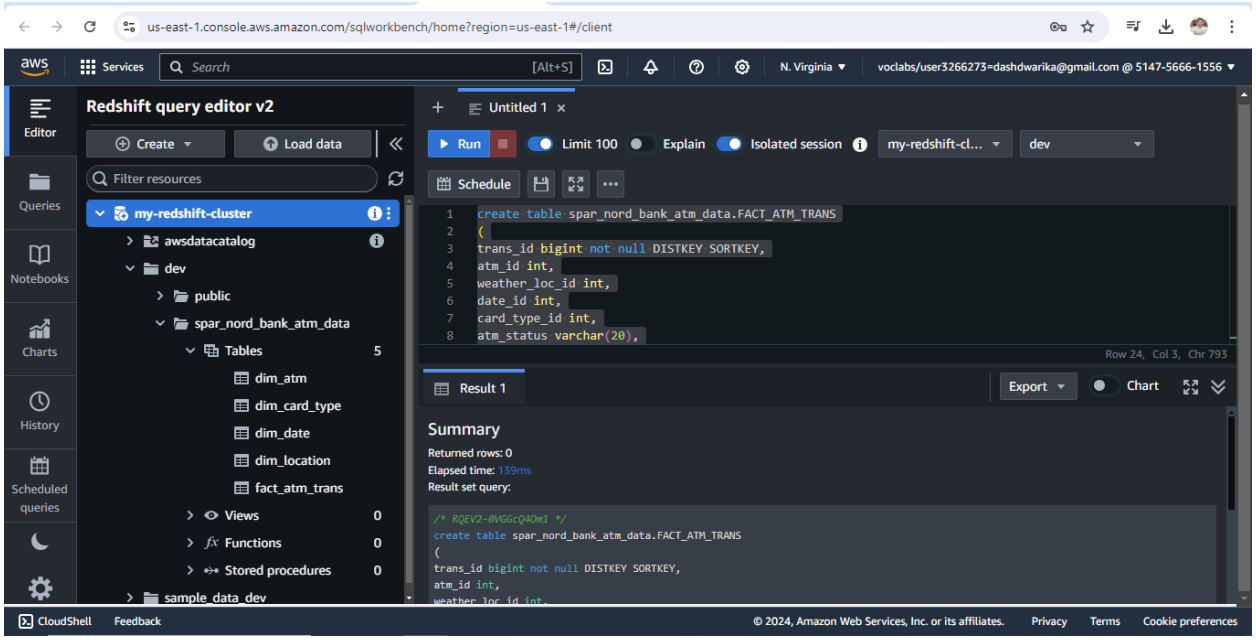

```
create table spar_nord_bank_atm_data.DIM_DATE
(
date_id int not null DISTKEY SORTKEY,
full_date_time timestamp,
year int,
month varchar(20),
day int,
hour int,
weekday varchar(20),
PRIMARY KEY(date_id)
);
```




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



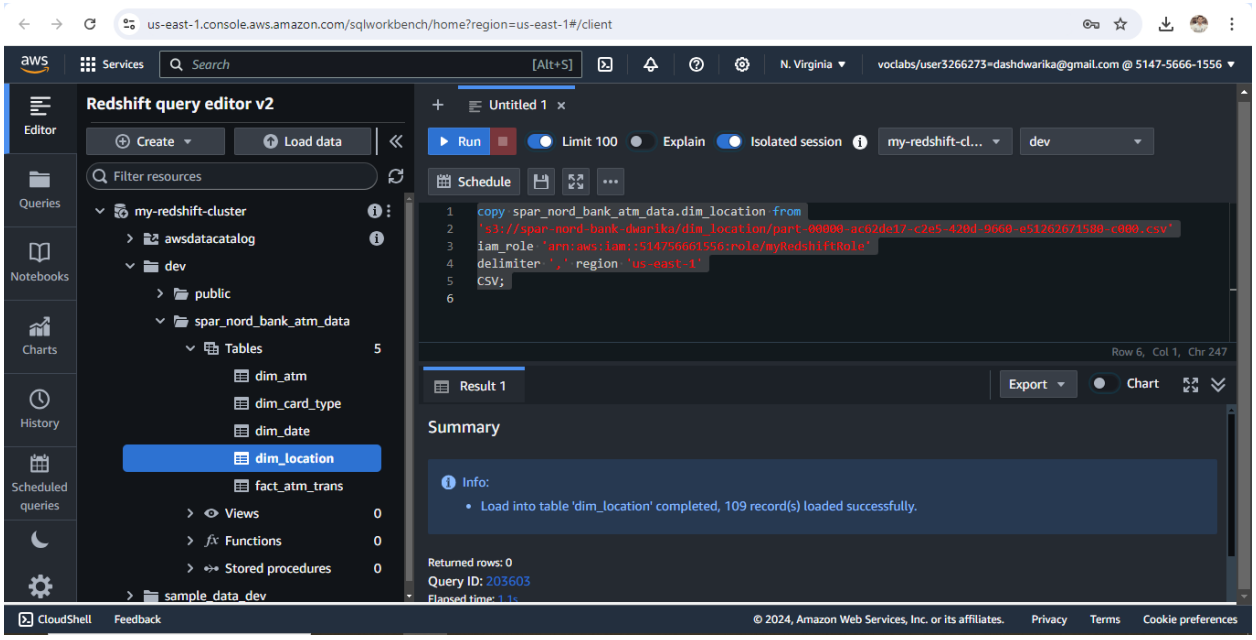
```
create table spar_nord_bank_atm_data.FACT_ATM_TRANS
(
trans_id bigint not null DISTKEY SORTKEY,
atm_id int,
weather_loc_id int,
date_id int,
card_type_id int,
atm_status varchar(20),
currency varchar(10),
service varchar(20),
transaction_amount int,
message_code varchar(225),
message_text varchar(225),
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
spar_nord_bank_atm_data.DIM_LOCATION(location_id),
FOREIGN KEY(atm_id) references spar_nord_bank_atm_data.DIM_ATM(atm_id),
FOREIGN KEY(date_id) references spar_nord_bank_atm_data.DIM_DATE(date_id),
FOREIGN KEY(card_type_id) references
spar_nord_bank_atm_data.DIM_CARD_TYPE(card_type_id)
);
```



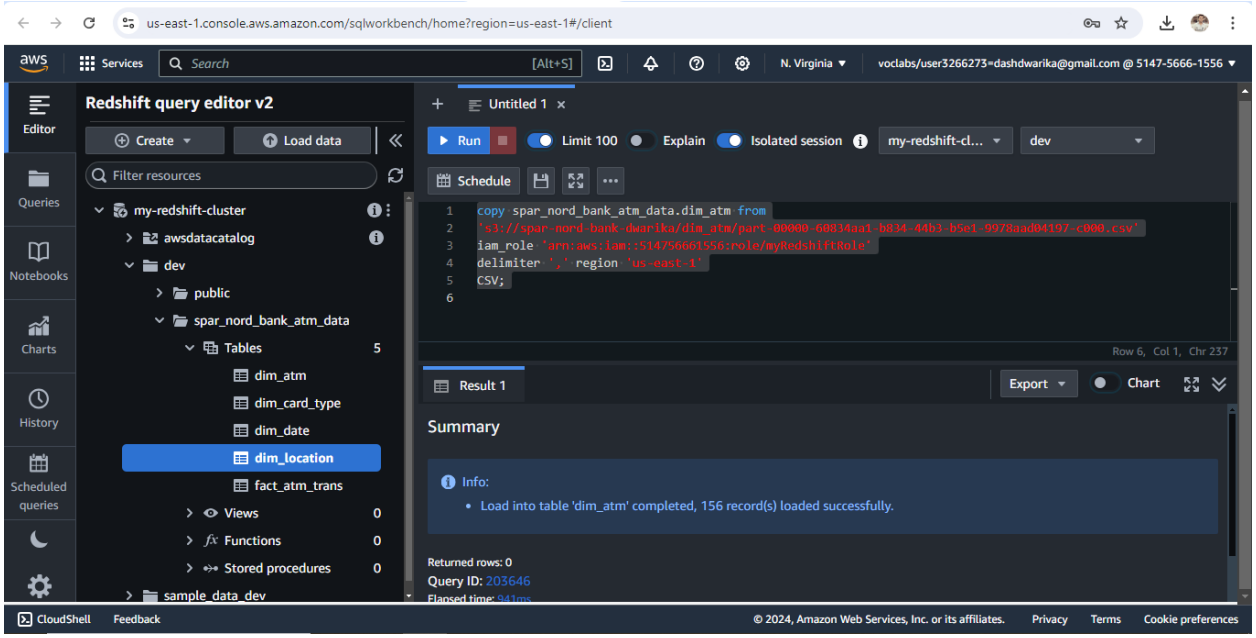
Loading data into a Redshift cluster from Amazon S3 bucket

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

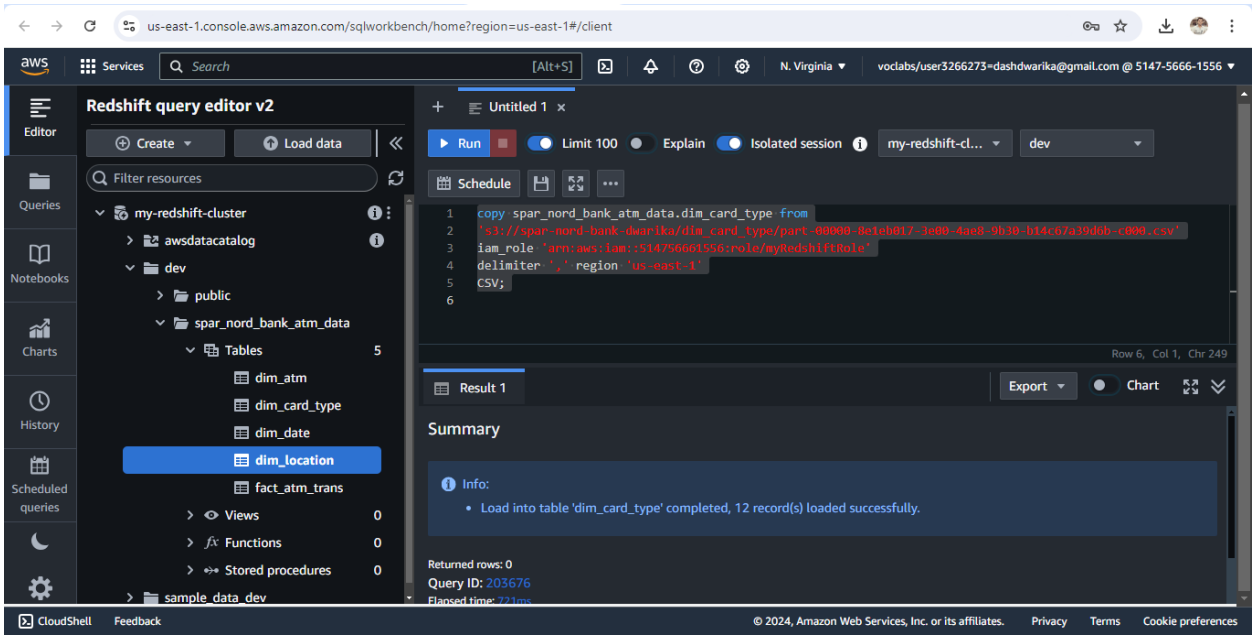
```
copy spar_nord_bank_atm_data.dim_location from
's3://spar-nord-bank-dwarika/dim_location/part-00000-ac62de17-c2e5-420d-9660-
e51262671580-c000.csv'
iam_role 'arn:aws:iam::514756661556:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
CSV;
```




```
copy spar_nord_bank_atm_data.dim_atm from
's3://spar-nord-bank-dwarika/dim_atm/part-00000-60834aa1-b834-44b3-b5e1-
9978aad04197-c000.csv'
iam_role 'arn:aws:iam::514756661556:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
CSV;
```

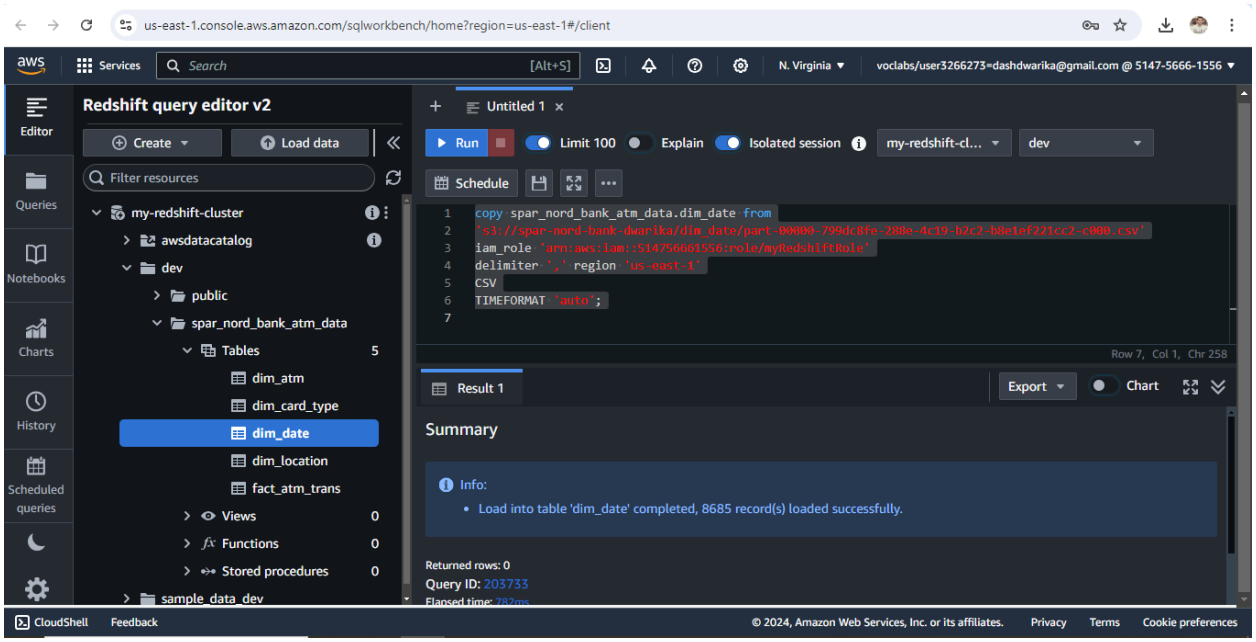



```
copy spar_nord_bank_atm_data.dim_card_type from
's3://spar-nord-bank-dwarika/dim_card_type/part-00000-8e1eb017-3e00-4ae8-9b30-
b14c67a39d6b-c000.csv'
iam_role 'arn:aws:iam::514756661556:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
CSV;
```




```
copy spar_nord_bank_atm_data.dim_date from
```

```
's3://spar-nord-bank-dwarika/dim_date/part-00000-799dc8fe-288e-4c19-b2c2-
b8e1ef221cc2-c000.csv'
iam_role 'arn:aws:iam::514756661556:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
CSV
TIMEFORMAT 'auto';
```




```
copy spar_nord_bank_atm_data.fact_atm_trans from
's3://spar-nord-bank-dwarika/fact_atm_trans/part-00000-fd716829-86aa-4836-b5c6-
49443a3cd2f7-c000.csv'
iam_role 'arn:aws:iam::514756661556:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
CSV;
```