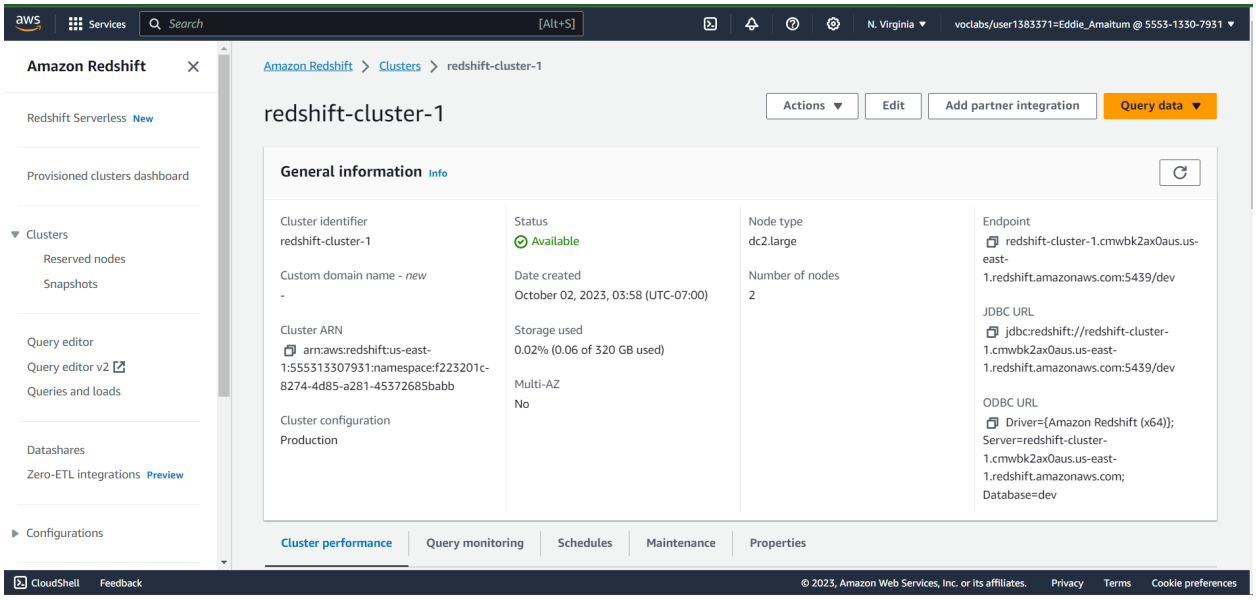
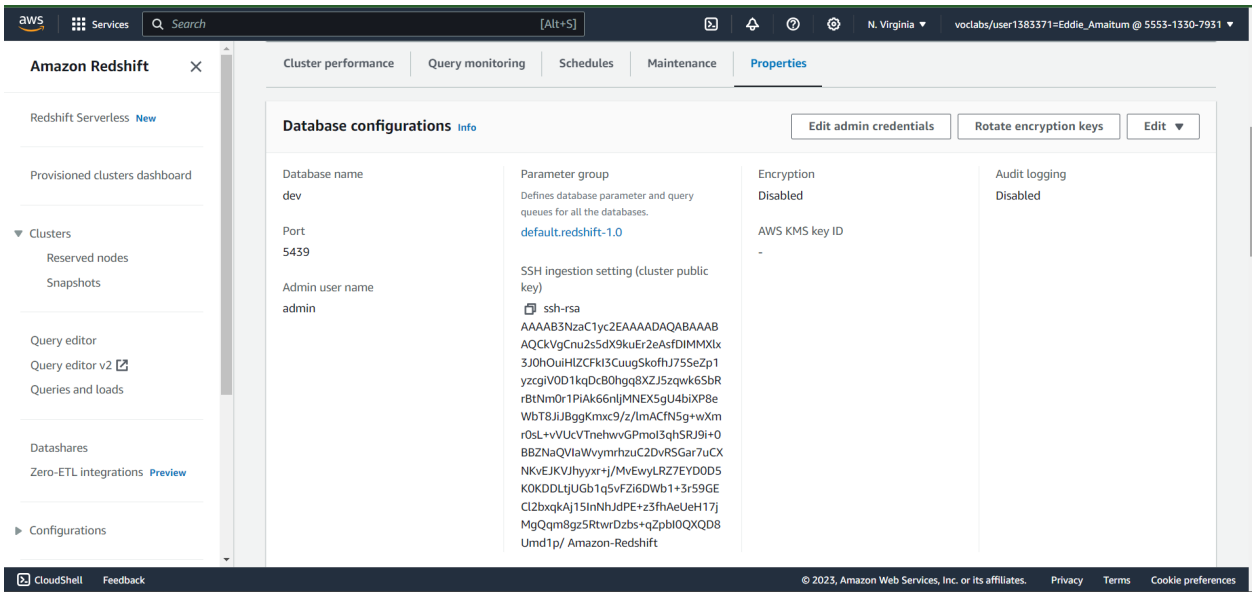
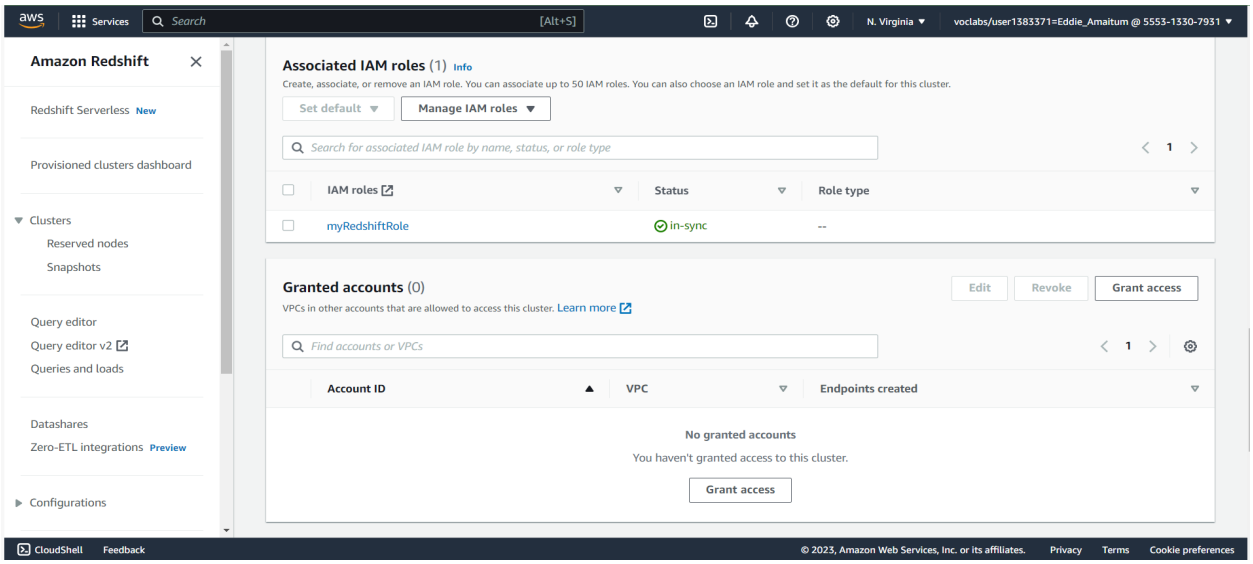
Creation of a Redshift Cluster

Screenshots of the configuration of the Redshift cluster

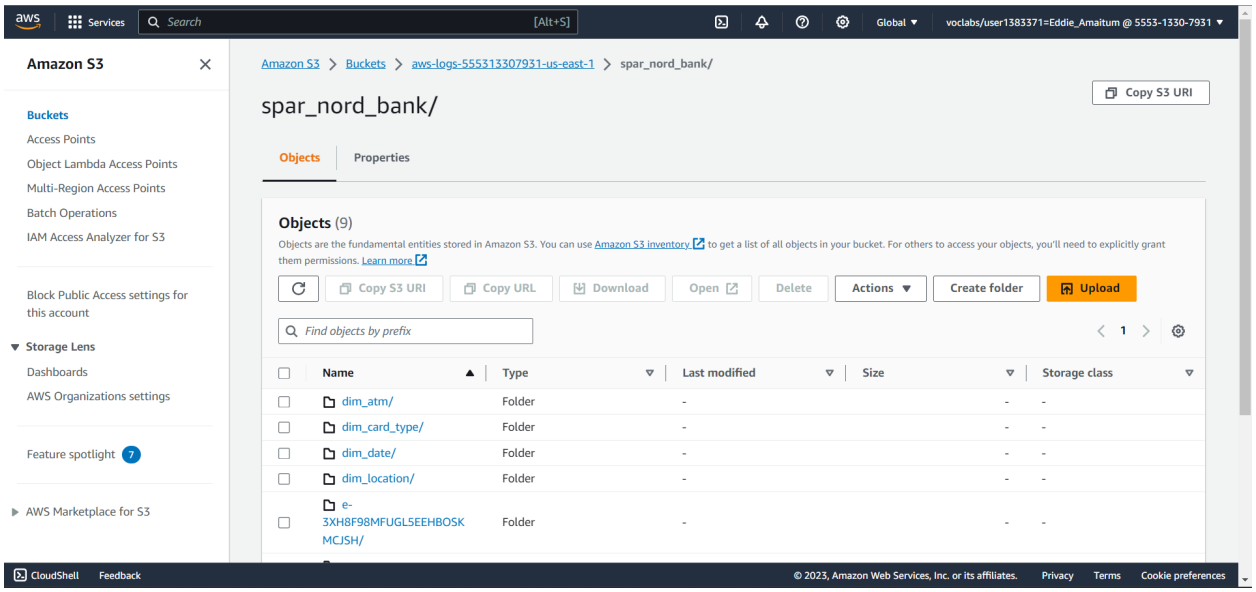






Setting up a database in the Redshift cluster and running queries to create the dimension and fact tables

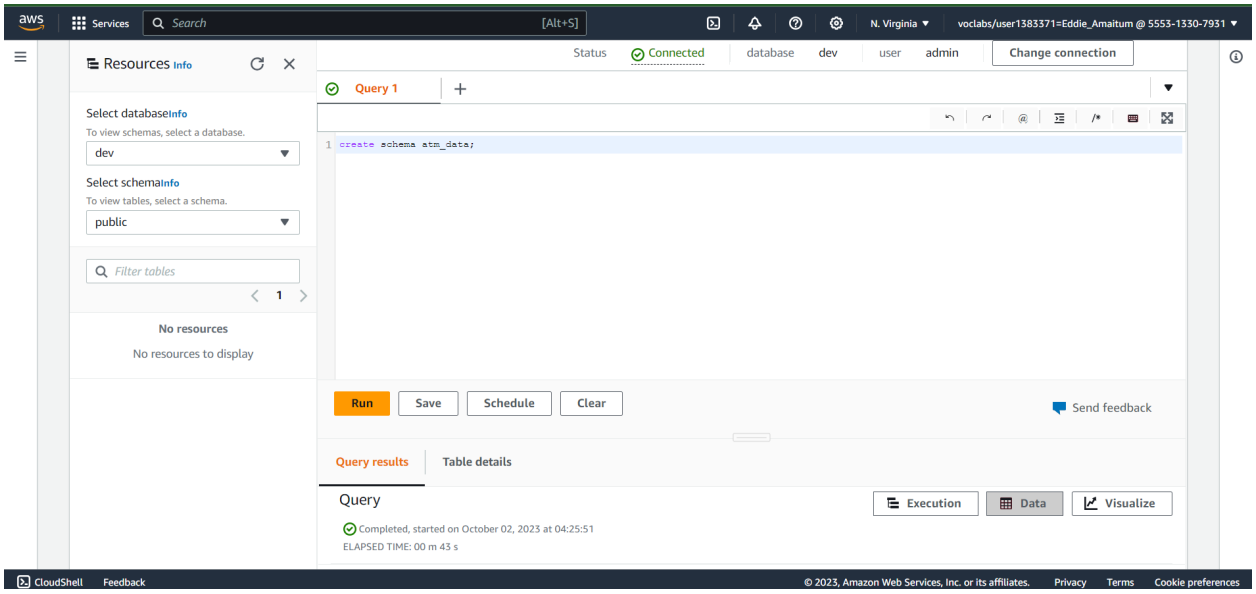
S3 bucket containing files:



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

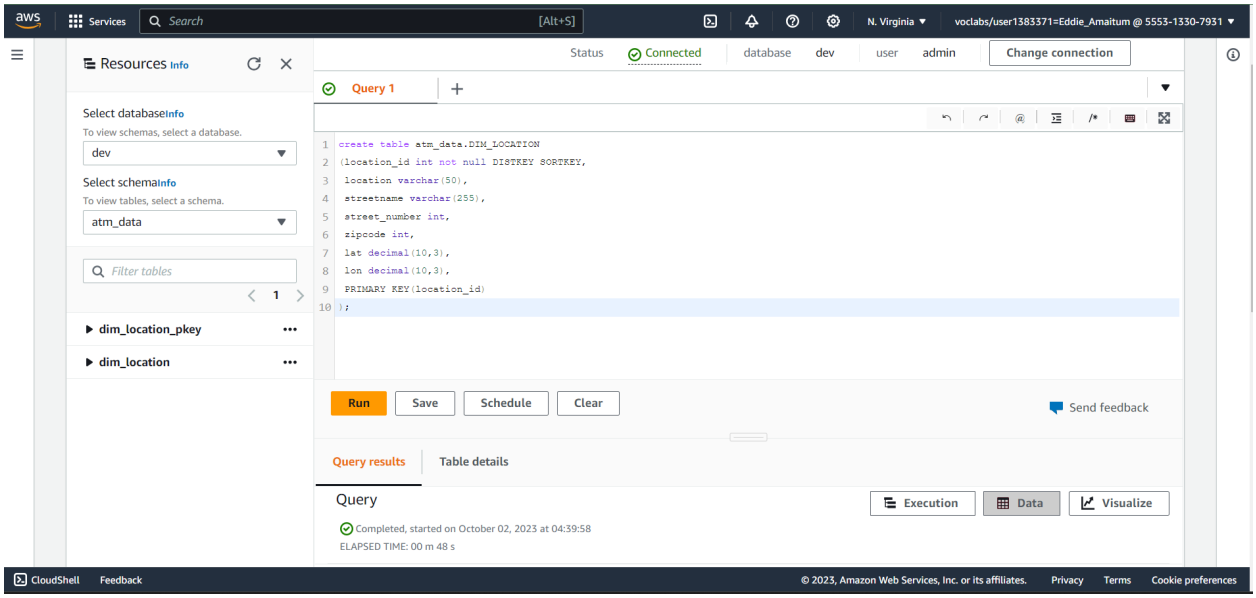
Query for creating schema:

```
create schema atm_data;
```



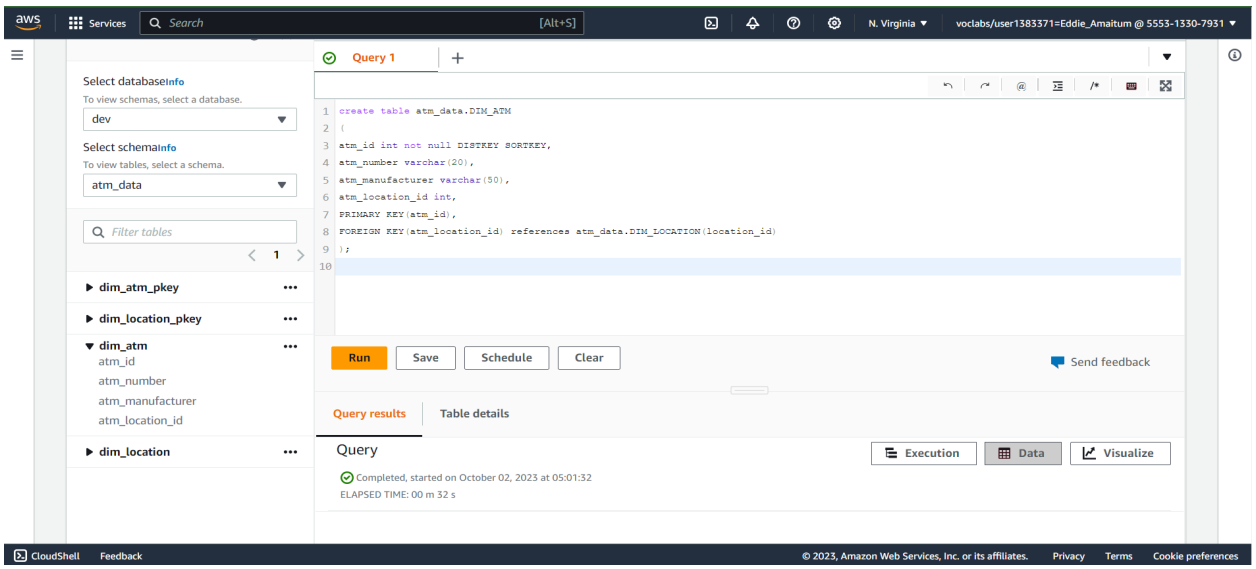
Query for creating location dimension table:

```
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)
);
```



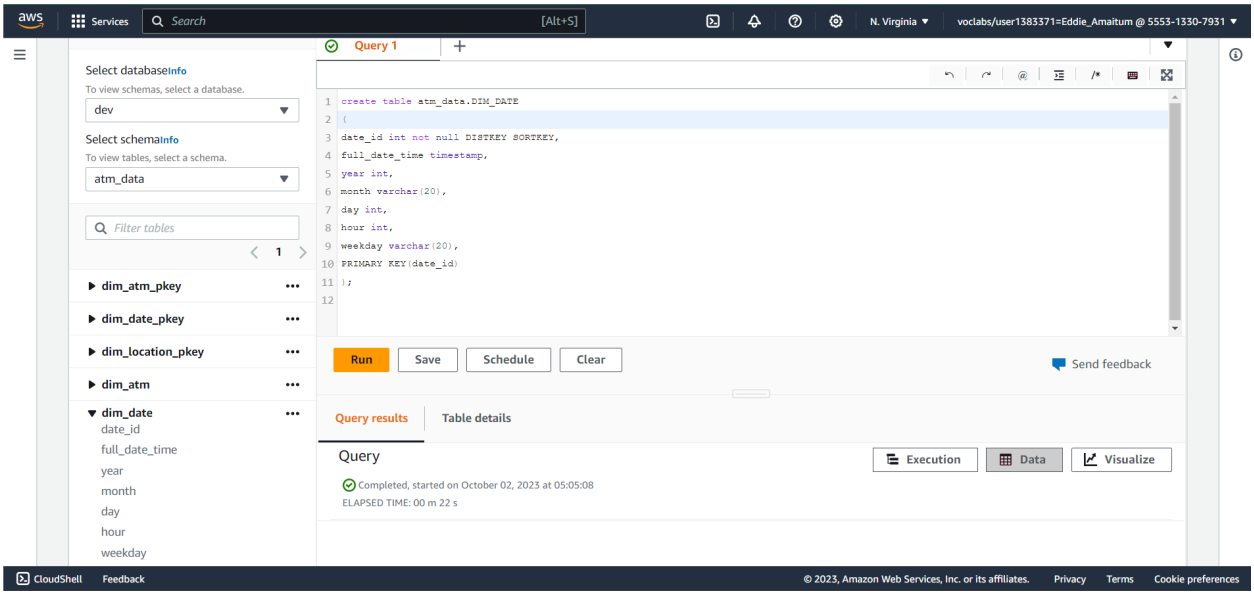
Query for creating atm dimension table :

```
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)
);
```



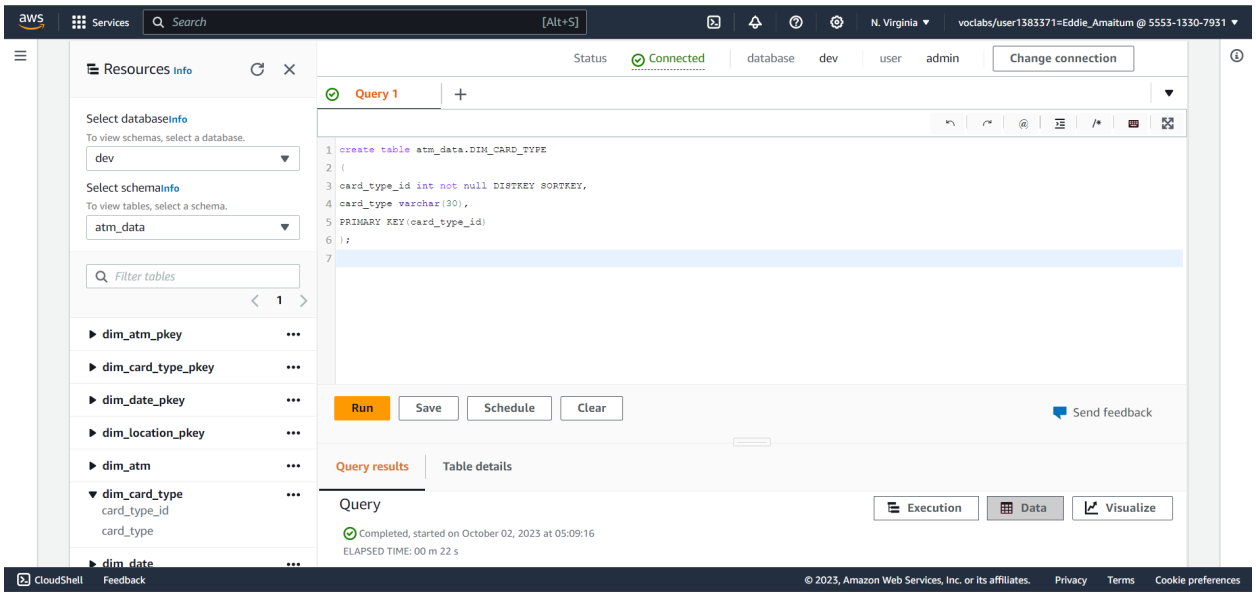
Query for creating date dimension table:

```
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)
);
```



Query for creating card type dimension table:

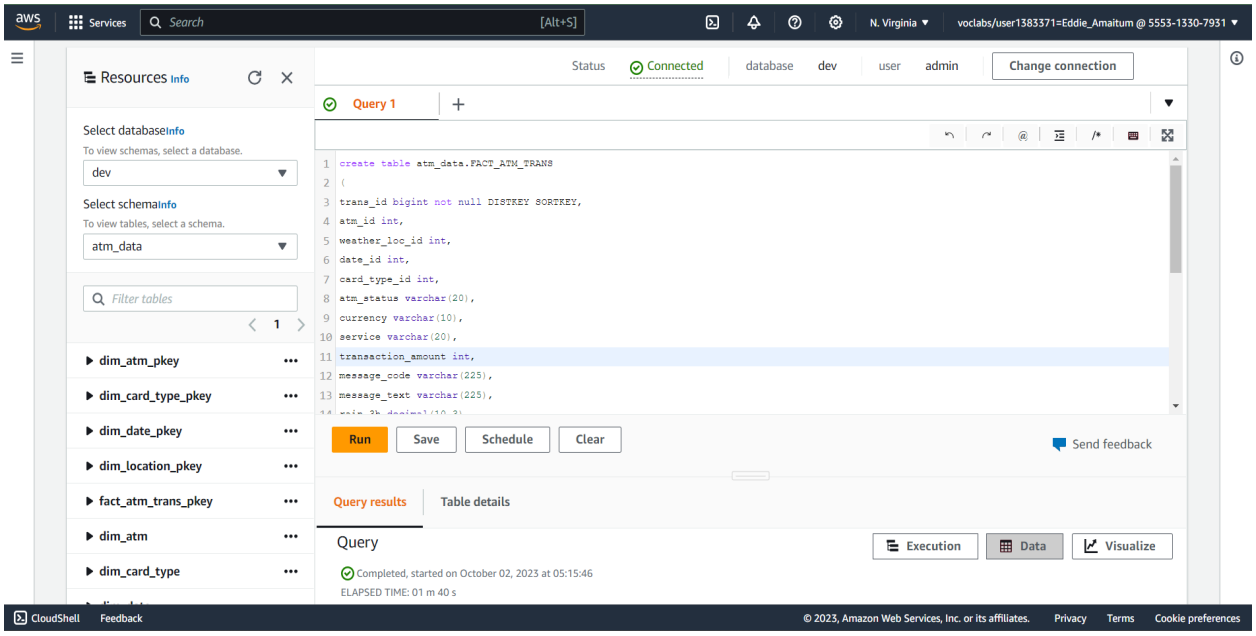
```
create table atm_data.DIM_CARD_TYPE
(
card_type_id int not null DISTKEY SORTKEY,
card_type varchar(30),
PRIMARY KEY(card_type_id)
);
```



Query for creating atm transaction 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)
);

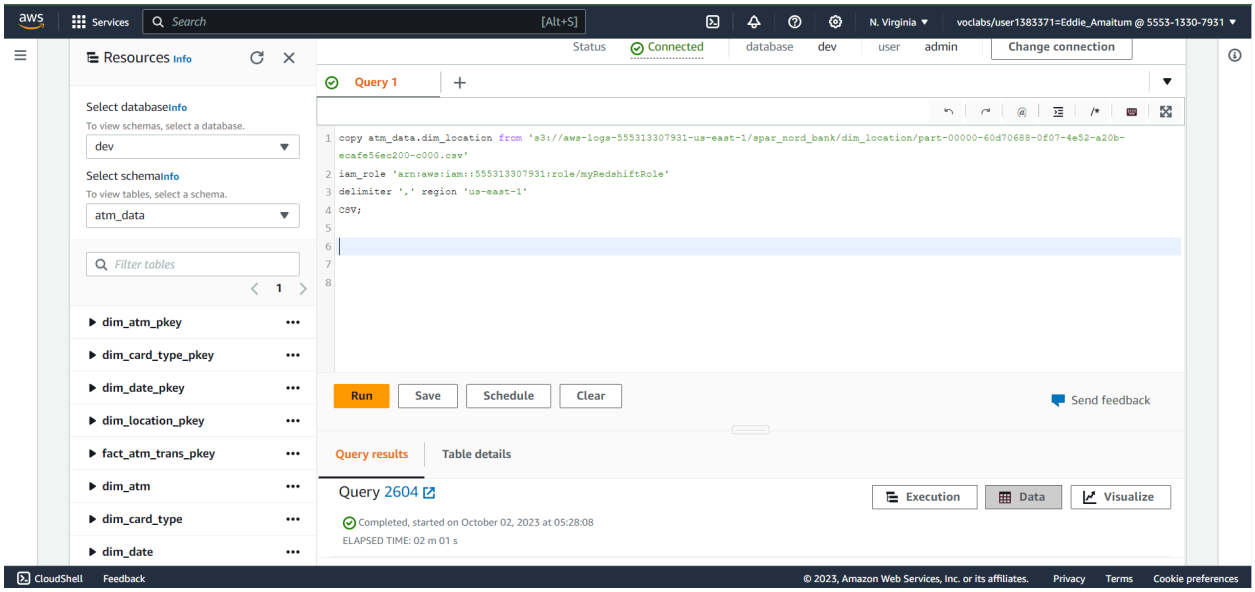


Loading data into a Redshift cluster from Amazon S3 bucket

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

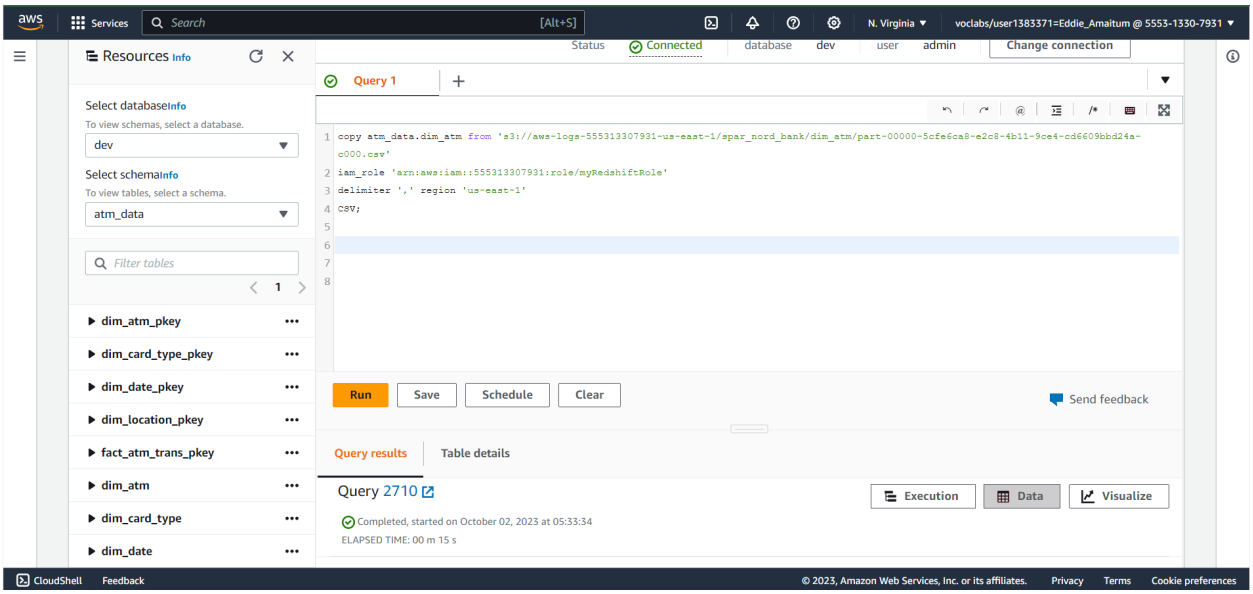
Query for copying data to dim_location table:

copy atm_data.dim_location from
's3://aws-logs-555313307931-us-east-1/spar_nord_bank/dim_location/part-00000-60d70688-0f07-4e52-a20b-ecafe56ec200-c000.csv'
iam_role 'arn:aws:iam::555313307931:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
CSV;



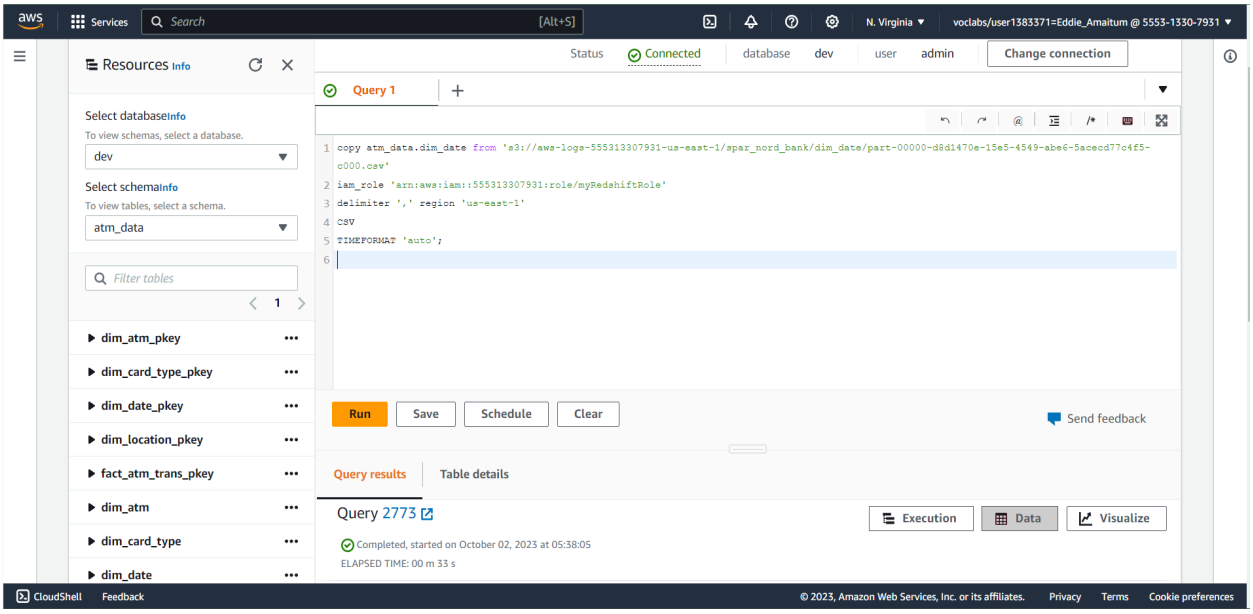
Query for copying data to dim_atm table:

copy atm_data.dim_atm from
's3://aws-logs-555313307931-us-east-1/spar_nord_bank/dim_atm/part-00000-5cfe6ca8-e2c8-4b11-9ce4-cd6609bbd24a-c000.csv'
iam_role 'arn:aws:iam::555313307931:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
CSV;



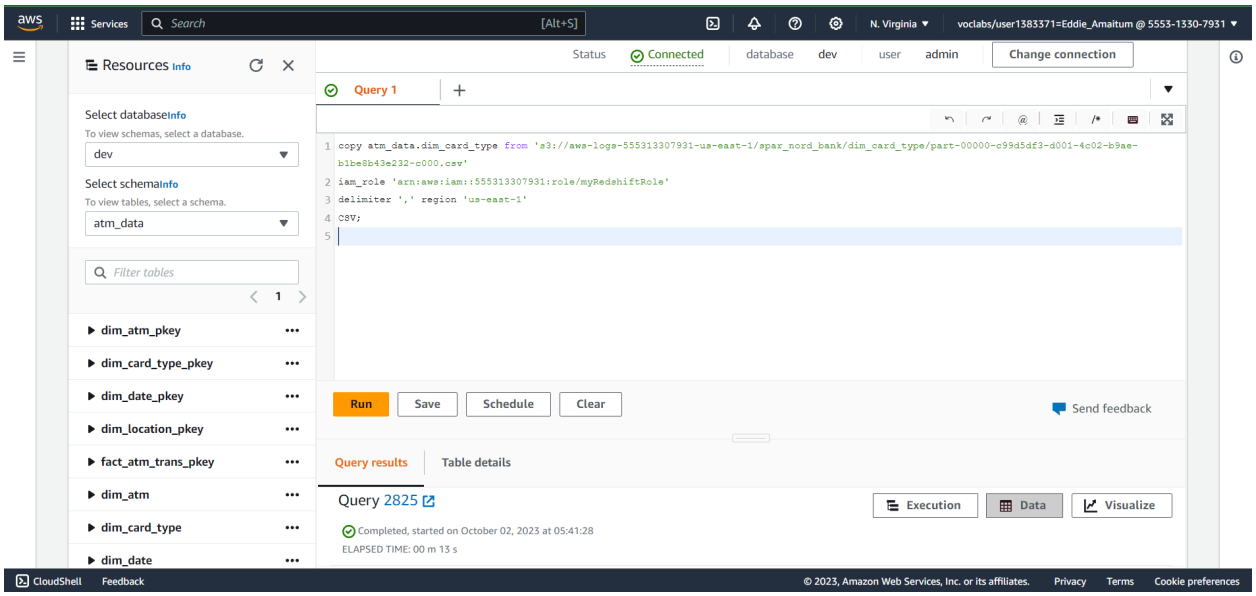
Query for copying data to dim_date table:

```
copy atm_data.dim_date from
's3://aws-logs-555313307931-us-east-1/spar_nord_bank/dim_date/part-00000-d8
d1470e-15e5-4549-abe6-5acecd77c4f5-c000.csv'
iam_role 'arn:aws:iam::555313307931:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
CSV
TIMEFORMAT 'auto';
```



Query for copying data to dim_card_type table:

```
copy atm_data.dim_card_type from
's3://aws-logs-555313307931-us-east-1/spar_nord_bank/dim_card_type/part-0000
0-c99d5df3-d001-4c02-b9ae-b1be8b43e232-c000.csv'
iam_role 'arn:aws:iam::555313307931:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
CSV;
```



Query for copying data to fact_atm_trans table:

copy atm_data.fact_atm_trans from
's3://aws-logs-555313307931-us-east-1/spar_nord_bank/fact_atm_trans/part-00000-7eb3fb69-77aa-4fa5-8702-c05b7d236790-c000.csv'
iam_role 'arn:aws:iam::555313307931:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
CSV;