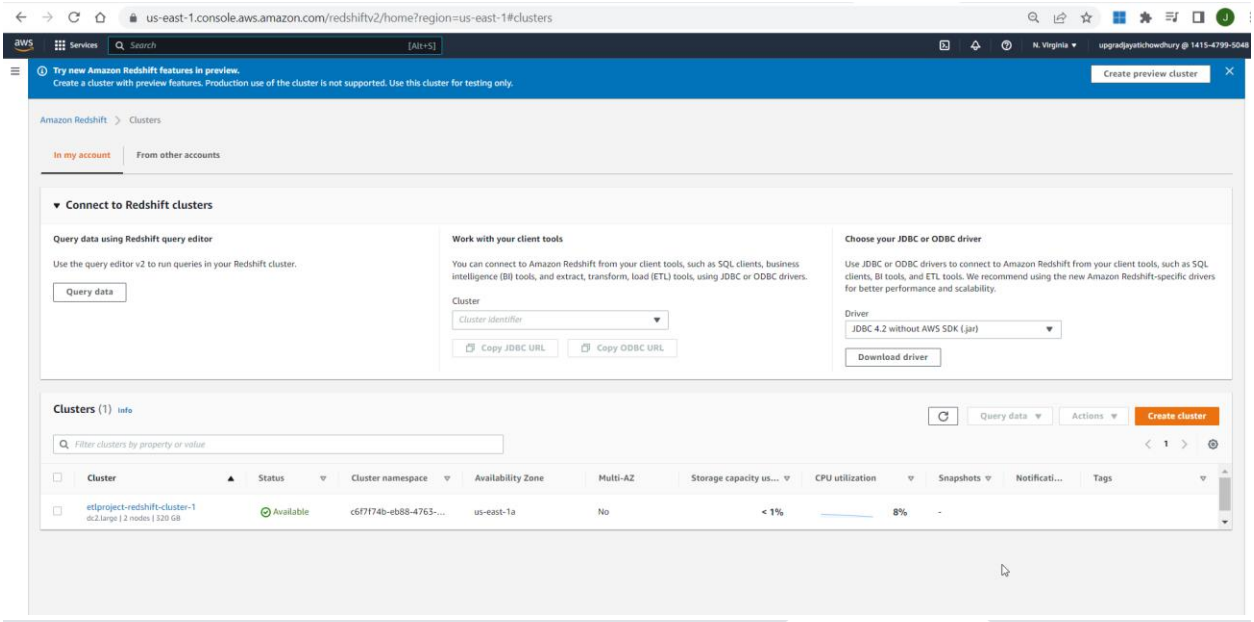
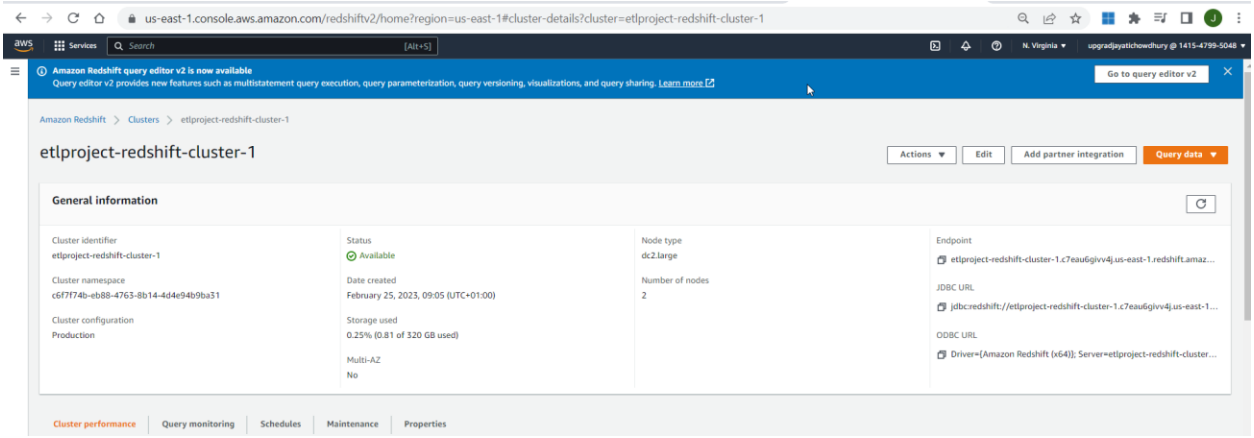
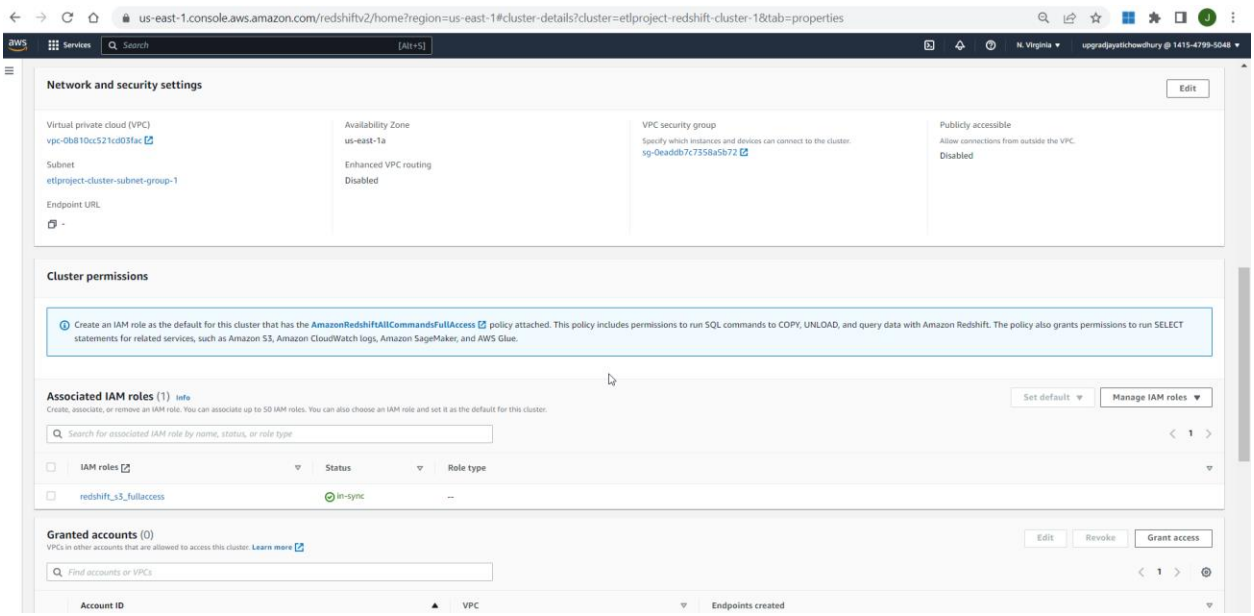
Creation of a Redshift Cluster

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

Screenshot of the type of machine used along with number of nodes

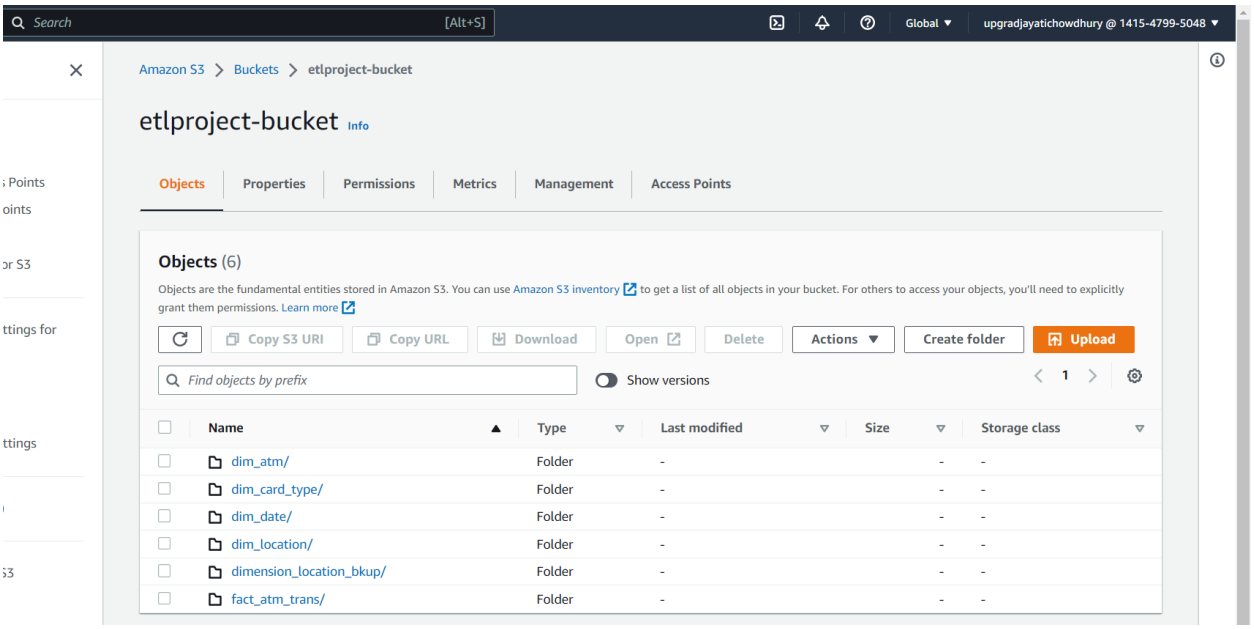






Setting up a database in the Redshift cluster and running queries to create the dimension and fact tables

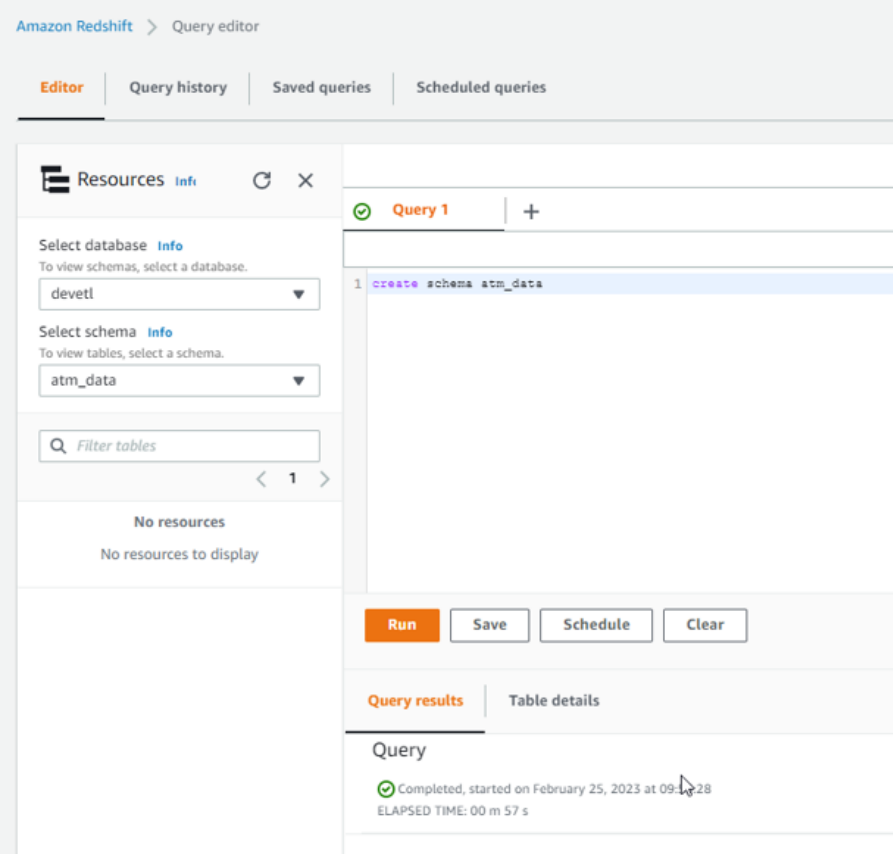
Viewing all data in Amazon S3 Bucket



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

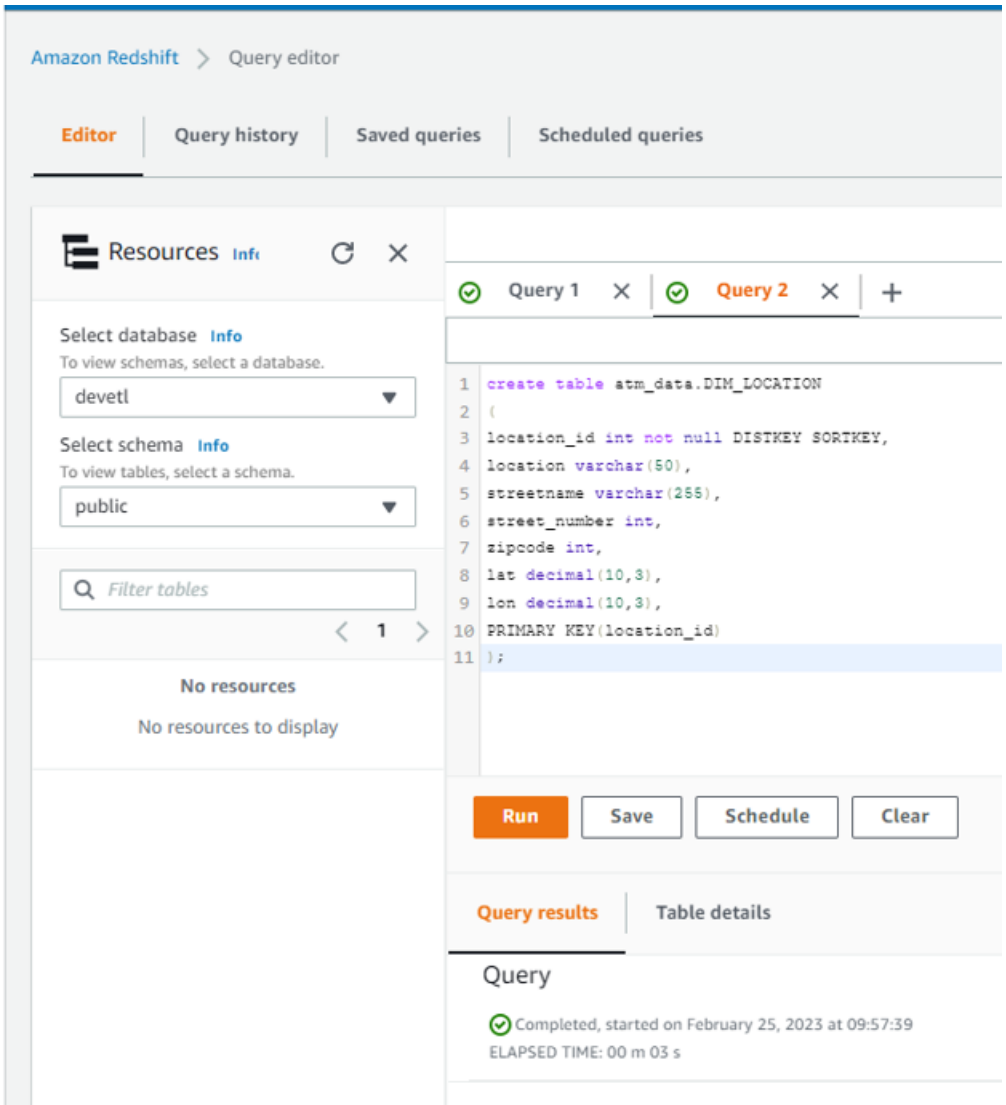
Creating Schema for dimension and fact tables

create schema atm_data;



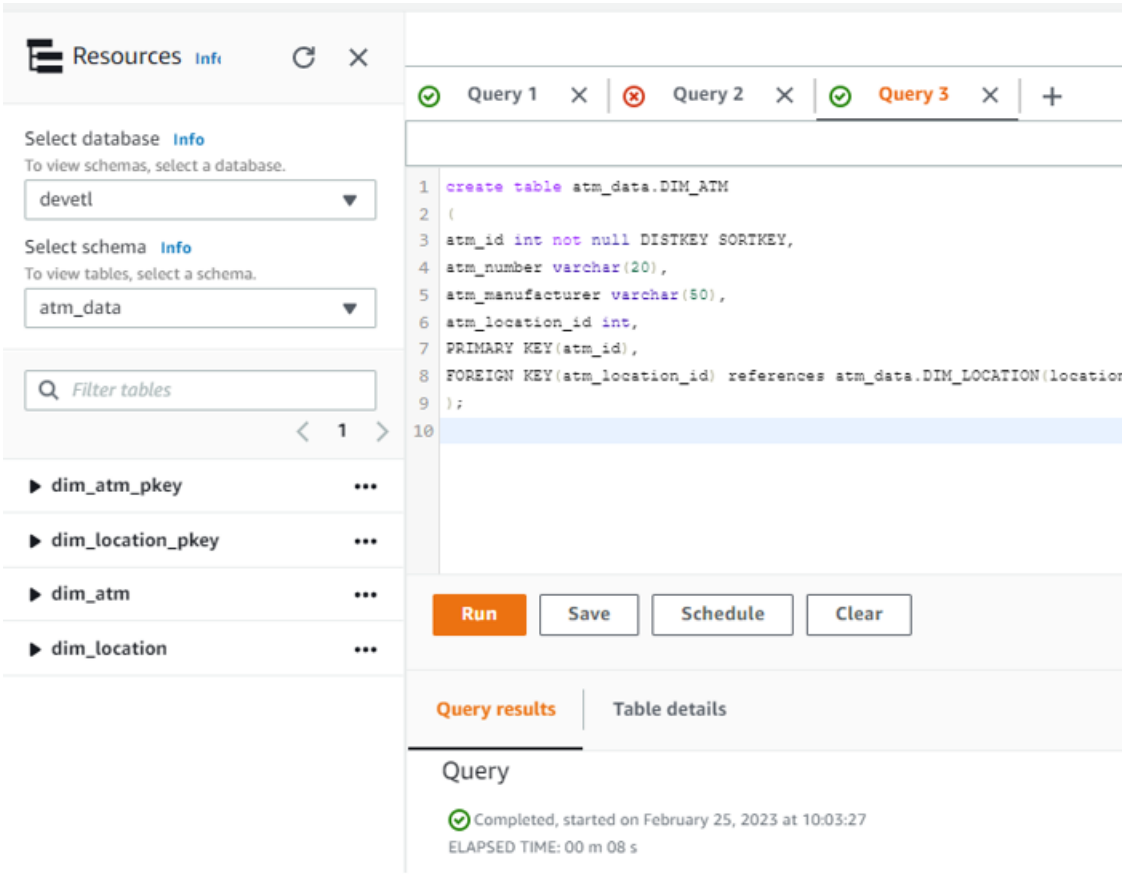
Creating location dimension table

```
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)
);
```



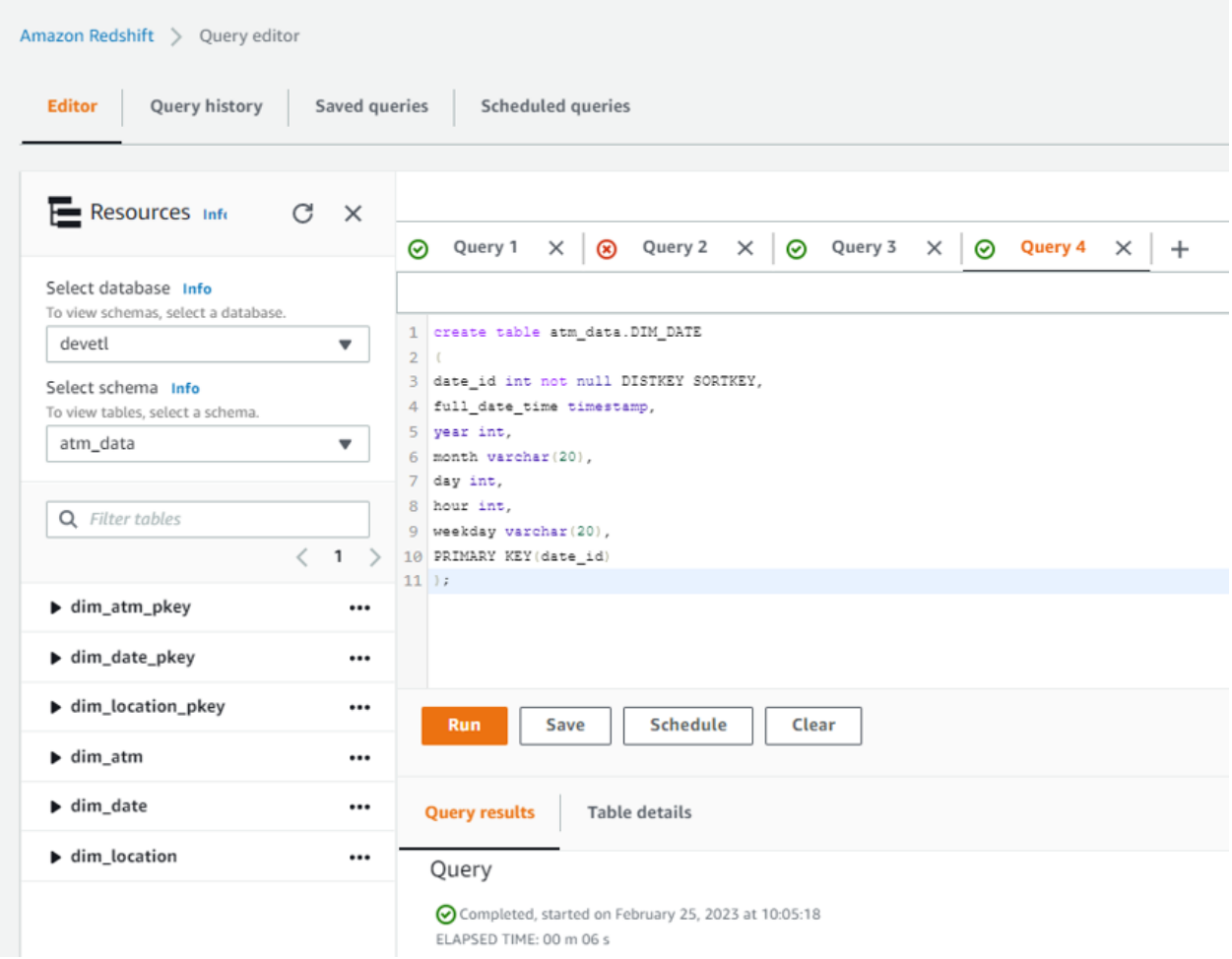
Creating atm dimension table

```
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)
);
```



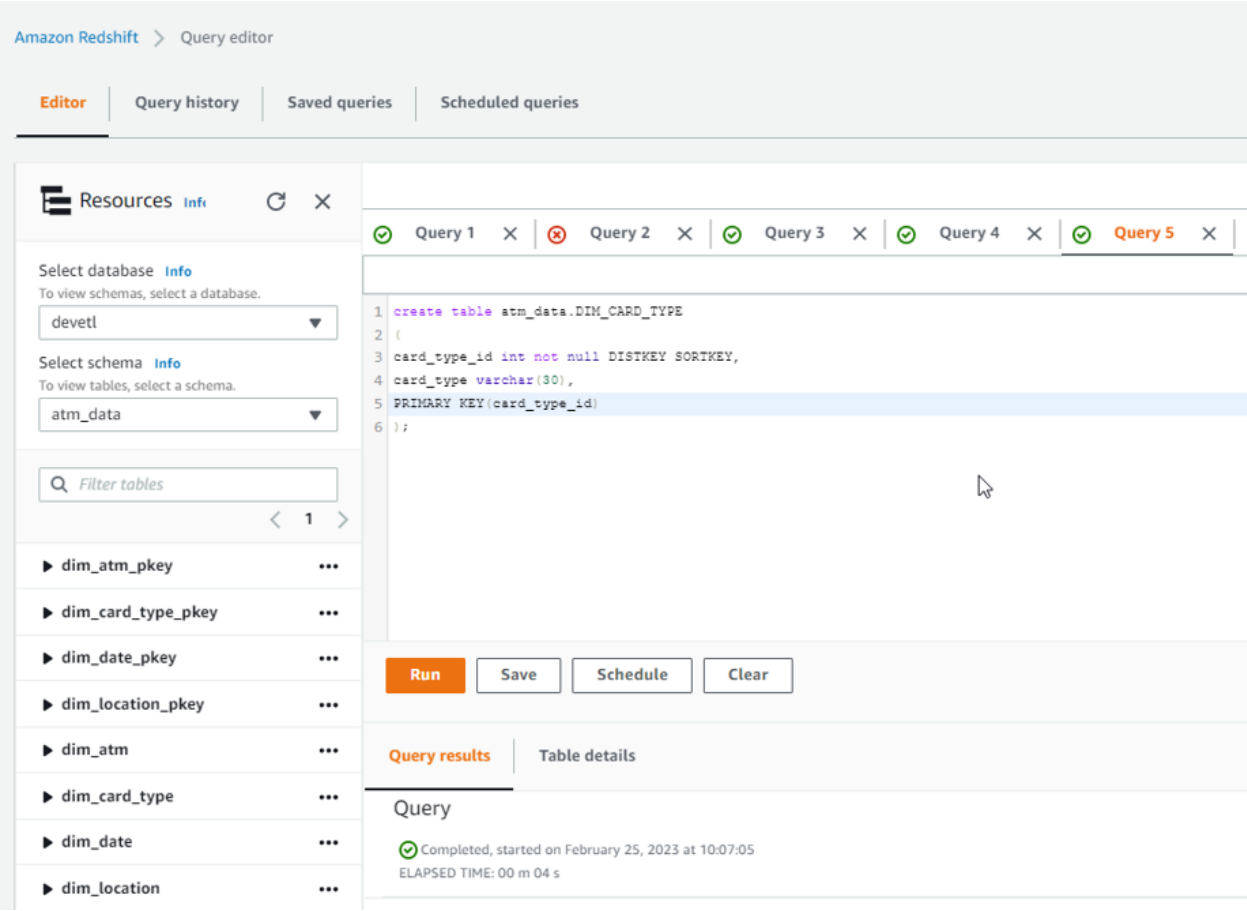
Creating date dimension table

```
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)
);
```



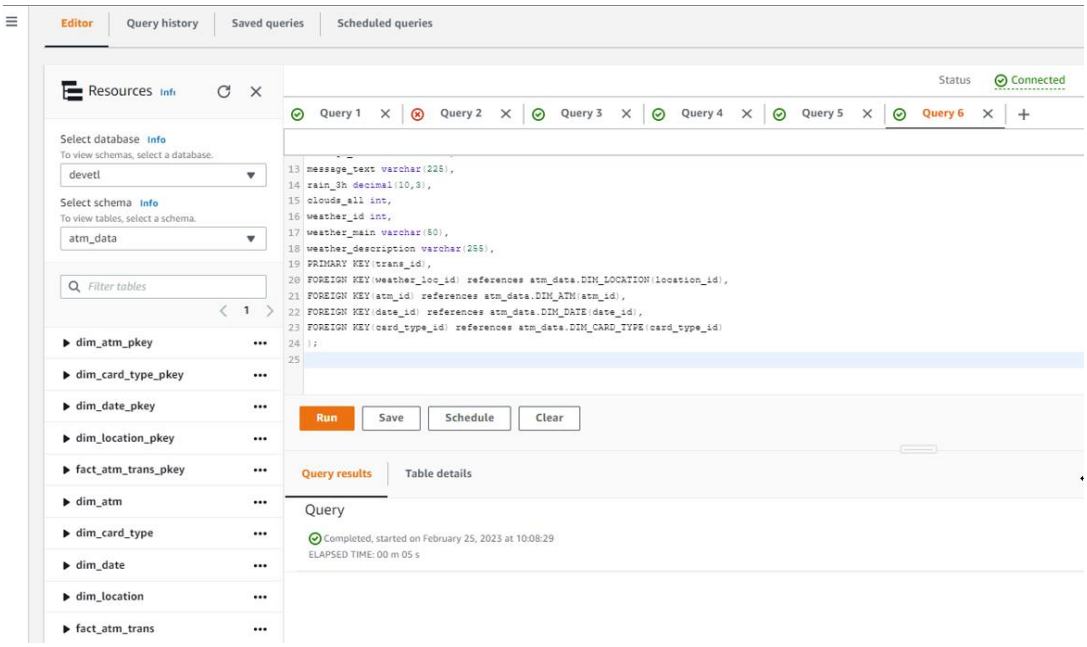
Creating card type dimension table

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



Creating atm transactions 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)
);
```

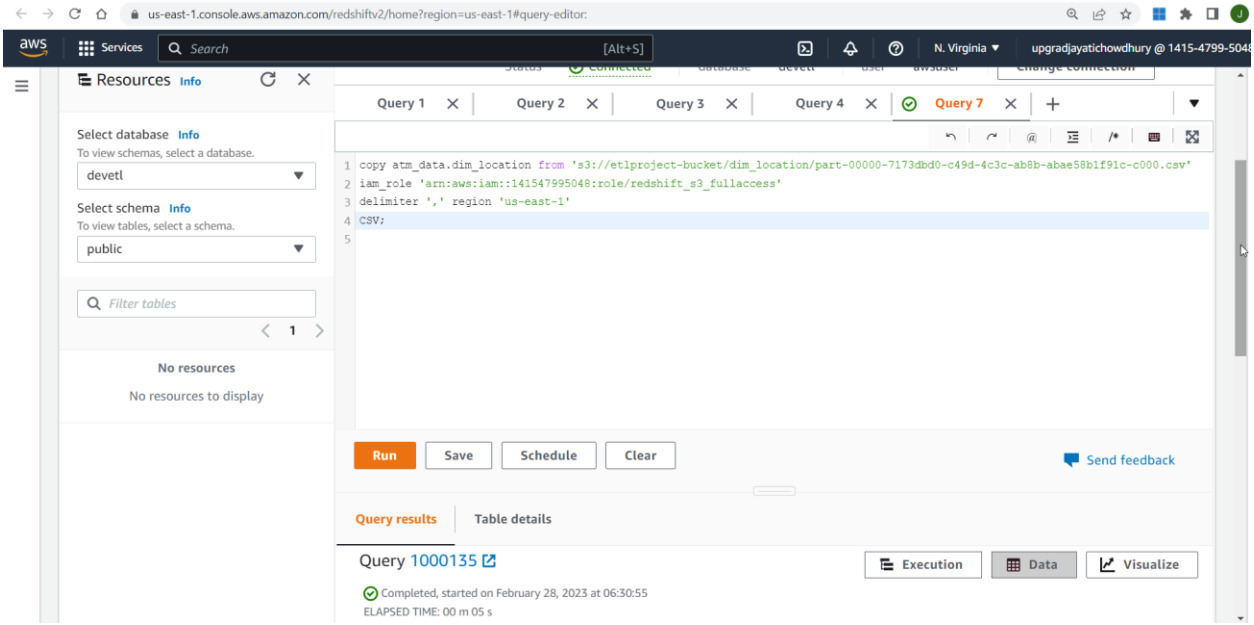


Loading data into a Redshift cluster from Amazon S3 bucket

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

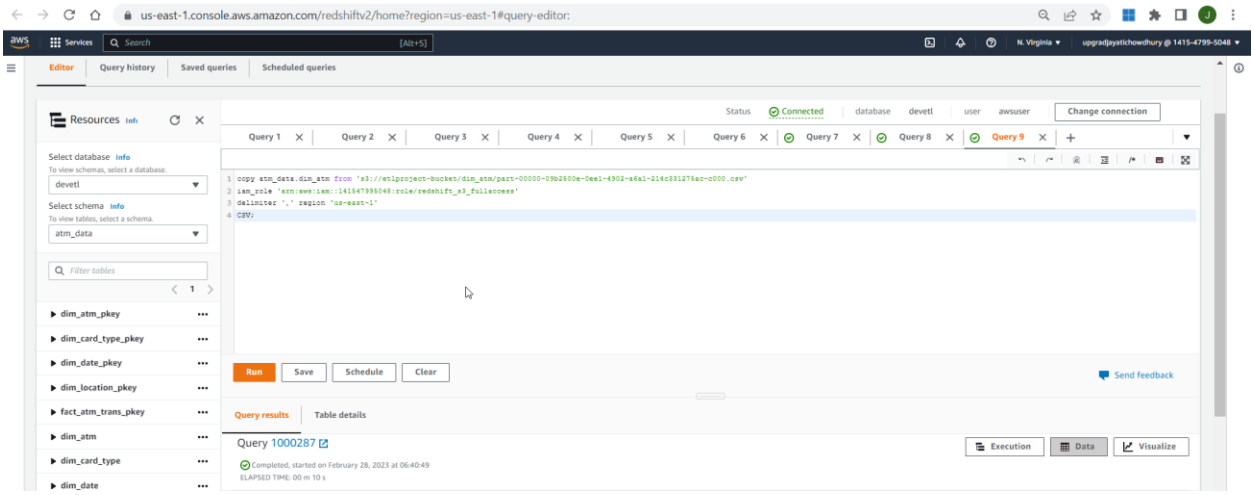
Copying the data to dim_location table

```
copy atm_data.dim_location from 's3://etlproject-bucket/dim_location/part-00000-7173dbd0-c49d-4c3c-ab8b-abae58b1f91c-c000.csv'
iam_role 'arn:aws:iam::141547995048:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
CSV;
```


us-east-1.console.aws.amazon.com/redshiftv2/home?region=us-east-1#query-editor:
aws
Services
Search
[Alt+S]
N. Virginia
upgradjayatichowdhury @ 1415-4799-5048
Resources Info
Select database Info
To view schemas, select a database.
devetl
Select schema Info
To view tables, select a schema.
public
Filter tables
1
No resources
No resources to display
Query 1
Query 2
Query 3
Query 4
Query 7
1 copy atm_data.dim_location from 's3://etlproject-bucket/dim_location/part-00000-7173dbd0-c49d-4c3c-ab8b-abae58b1f91c-c000.csv'
2 iam_role 'arn:aws:iam::141547995048:role/redshift_s3_fullaccess'
3 delimiter ',' region 'us-east-1'
4 CSV;
5
Run
Save
Schedule
Clear
Send feedback
Query results
Table details
Query 1000135
Execution
Data
Visualize
Completed, started on February 28, 2023 at 06:30:55
ELAPSED TIME: 00 m 05 s

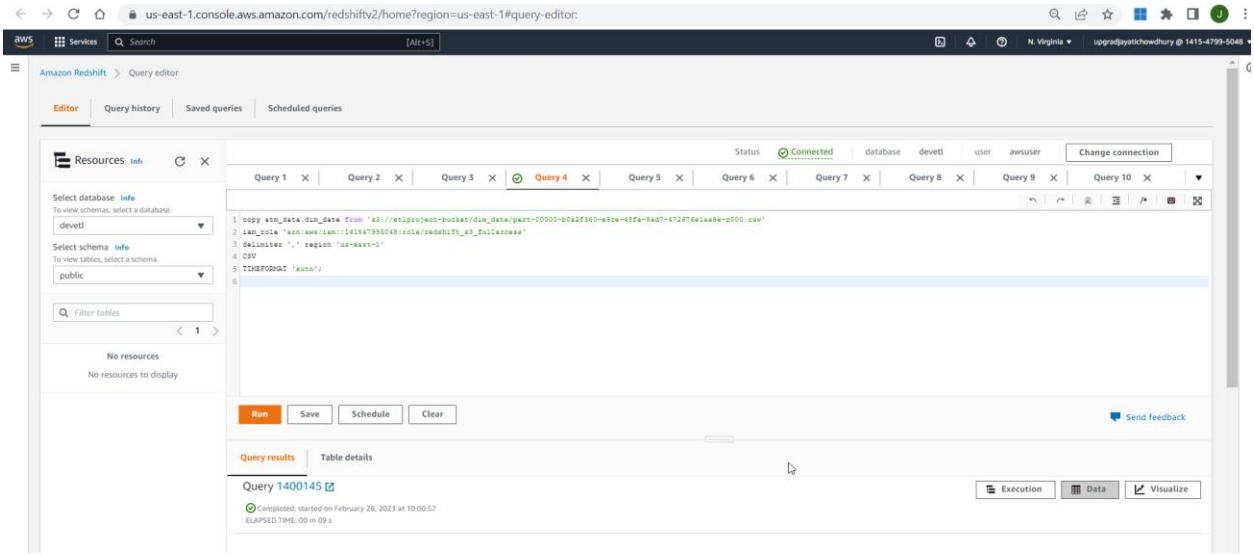
Copying the data to dim_atm table

copy atm_data.dim_atm from 's3://etlproject-bucket/dim_atm/part-00000-09b2500e-0ee1-4902-a6a1-214c331275ac-c000.csv'
iam_role 'arn:aws:iam::141547995048:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
CSV;



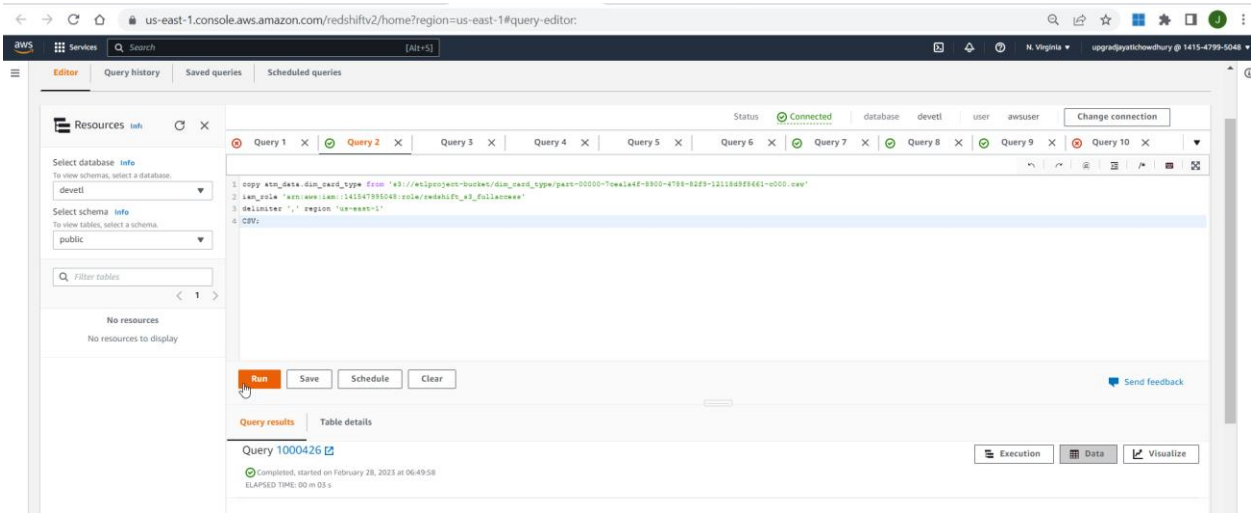
Copying the data to dim_date table

```
copy atm_data.dim_date from 's3://etlproject-bucket/dim_date/part-00000-b0a2f360-e9ce-48fa-8ad7-472676e1aa9e-c000.csv'
iam_role 'arn:aws:iam::141547995048:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
CSV
TIMEFORMAT 'auto';
```



Copying the data to dim_card_type table

copy atm_data.dim_card_type from 's3://etlproject-bucket/dim_card_type/part-00000-7cea1a4f-8900-4788-82f9-12118d9f8661-c000.csv'
iam_role 'arn:aws:iam::141547995048:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
CSV;



Copying the data to fact_atm_trans table

copy atm_data.fact_atm_trans from 's3://etlproject-bucket/fact_atm_trans/part-00000-349988c6-cf3c-4636-8c56-8b0314a2fbc1-c000.csv'
iam_role 'arn:aws:iam::141547995048:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
CSV;