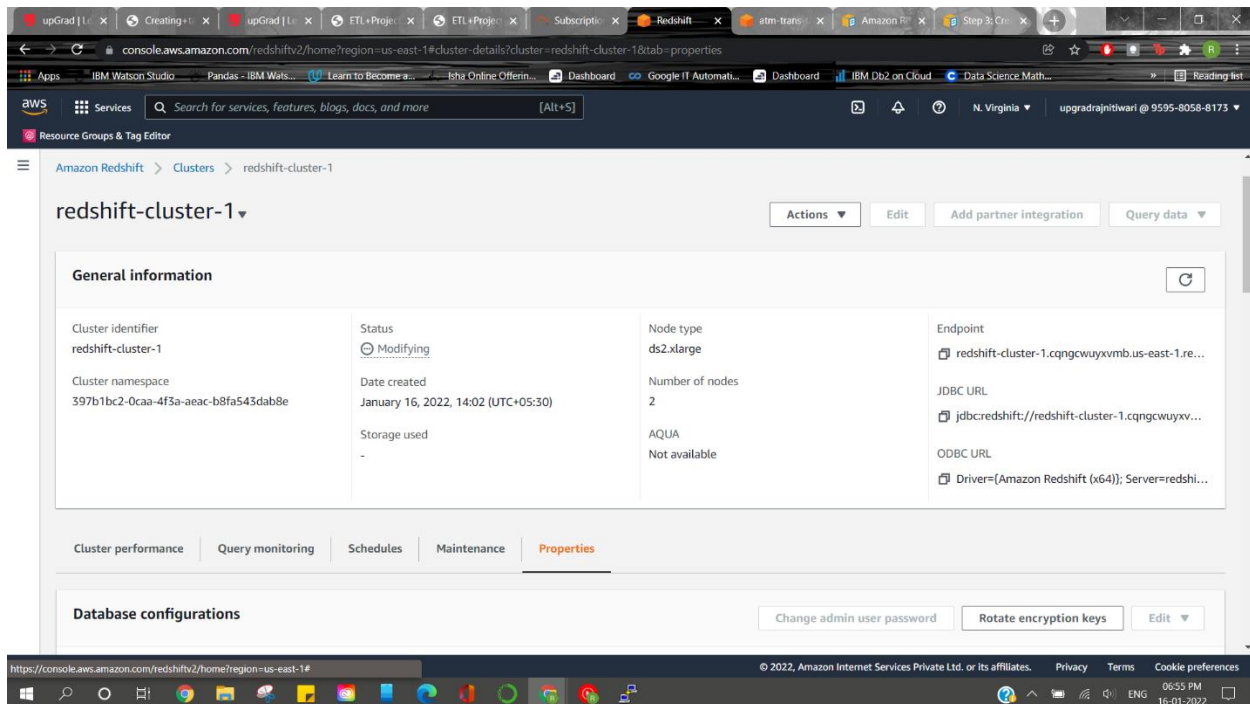


Creation of a RedShift Cluster

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

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



The screenshot displays the AWS Management Console interface for an Amazon Redshift cluster. The cluster is named 'redshift-cluster-1' and is in the 'Modifying' status. The configuration details are as follows:

General information			
Cluster identifier	redshift-cluster-1	Status	Modifying
Cluster namespace	397b1bc2-0caa-4f3a-aeac-b8fa543dab8e	Date created	January 16, 2022, 14:02 (UTC+05:30)
Node type	ds2.xlarge	Number of nodes	2
Endpoint	redshift-cluster-1.cqngcwuyxvmb.us-east-1.re...	AQUA	Not available
JDBC URL	jdbc:redshift://redshift-cluster-1.cqngcwuyxv...	Storage used	-
ODBC URL	Driver=(Amazon Redshift (x64)); Server=redshi...		

Below the general information, there are tabs for 'Cluster performance', 'Query monitoring', 'Schedules', 'Maintenance', and 'Properties'. The 'Properties' tab is currently selected, showing 'Database configurations' with options to 'Change admin user password', 'Rotate encryption keys', and 'Edit'.

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 atm_trans;
```

```
create table atm_trans.dim_location(  
    location_id int,  
    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 atm_trans.dim_atm(  
    atm_id int,  
    atm_number varchar(20),  
    atm_manufacturer varchar(50),  
    atm_location_id int,  
    Primary key(atm_id),  
    Foreign key(atm_location_id) REFERENCES atm_trans.dim_location(location_id)  
);
```

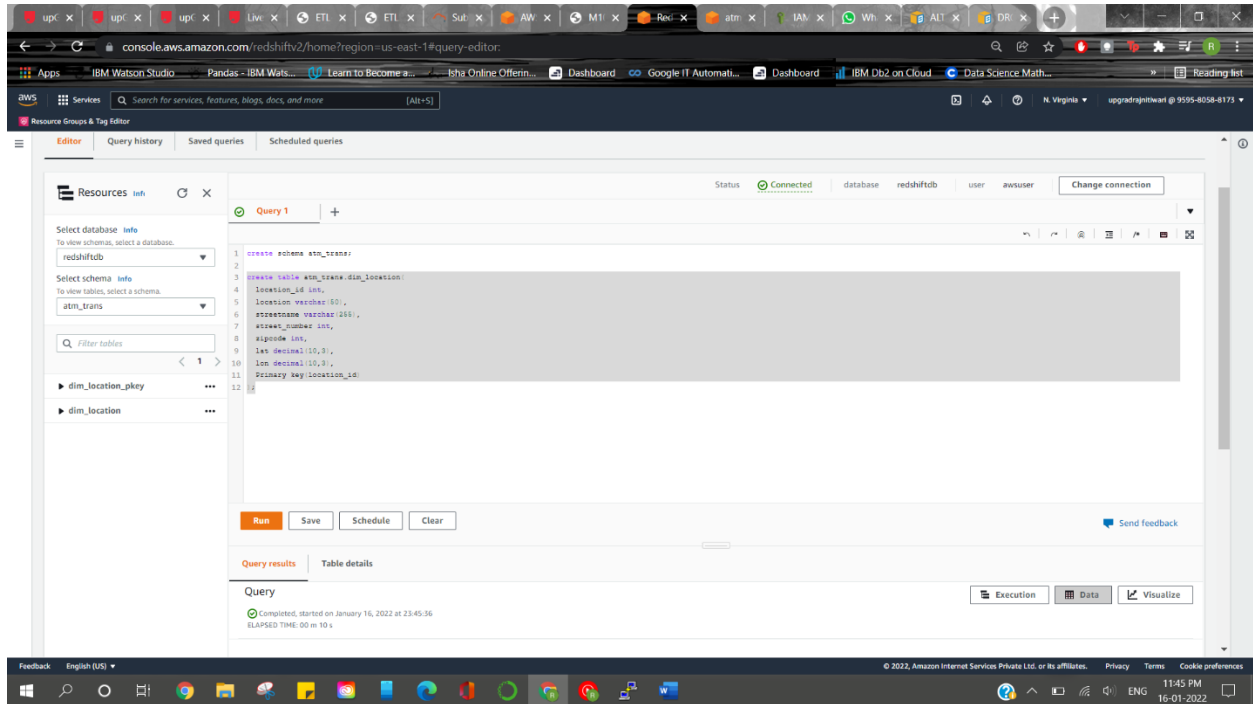
```
create table atm_trans.dim_card_type(  
    card_type_id int,  
    card_type varchar(50),  
    Primary key(card_type_id)  
);
```

```
create table atm_trans.dim_date(  
    date_id int,  
    full_date_time timestamp,  
    year int,  
    month varchar(20),
```

```
day int,  
hour int,  
weekday varchar(20),  
Primary key(date_id)  
);
```

```
create table atm_trans.fact_atm_trans(  
trans_id bigint,  
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(atm_id) REFERENCES atm_trans.dim_atm(atm_id),  
Foreign key(weather_loc_id) REFERENCES atm_trans.dim_location(location_id),  
Foreign key(date_id) REFERENCES atm_trans.dim_date(date_id),  
Foreign key(card_type_id) REFERENCES atm_trans.dim_card_type(card_type_id)  
);
```



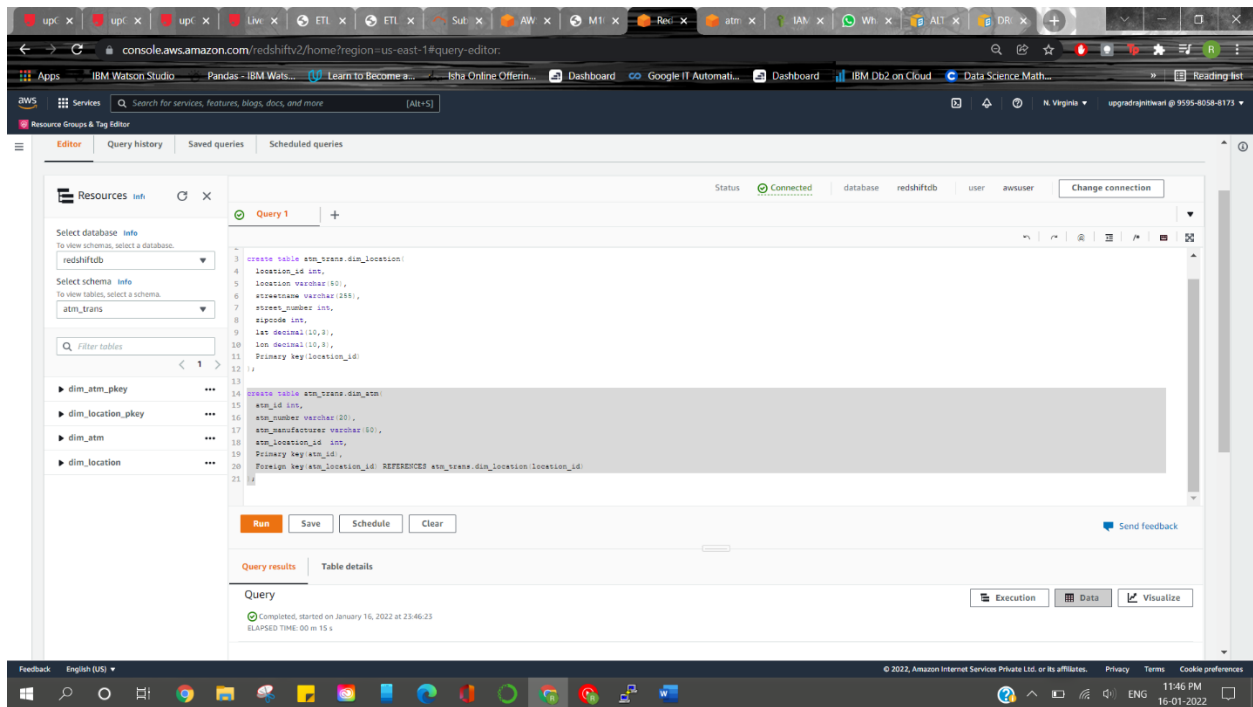
The screenshot shows the AWS Redshift console interface. On the left, the 'Resources' panel is expanded, showing the 'Select database' dropdown set to 'redshiftdb' and the 'Select schema' dropdown set to 'atm_txnns'. The 'Filter tables' search bar is empty. Below the search bar, a list of tables is shown, including 'dim_location_pkey' and 'dim_location'. The main editor area displays a SQL query for creating a table named 'atm_txnns' with the following schema:

```

1 create table atm_txnns (
2   location_id int,
3   location varchar(50),
4   storename varchar(255),
5   store_number int,
6   zipcode int,
7   lat decimal(10,3),
8   lon decimal(10,3),
9   Primary key(location_id)
10 );

```

Below the query editor, the 'Run' button is highlighted. The 'Query results' tab is selected, showing the query execution status: 'Completed, started on January 16, 2022 at 23:45:36' and 'ELAPSED TIME: 00 m 10 s'.



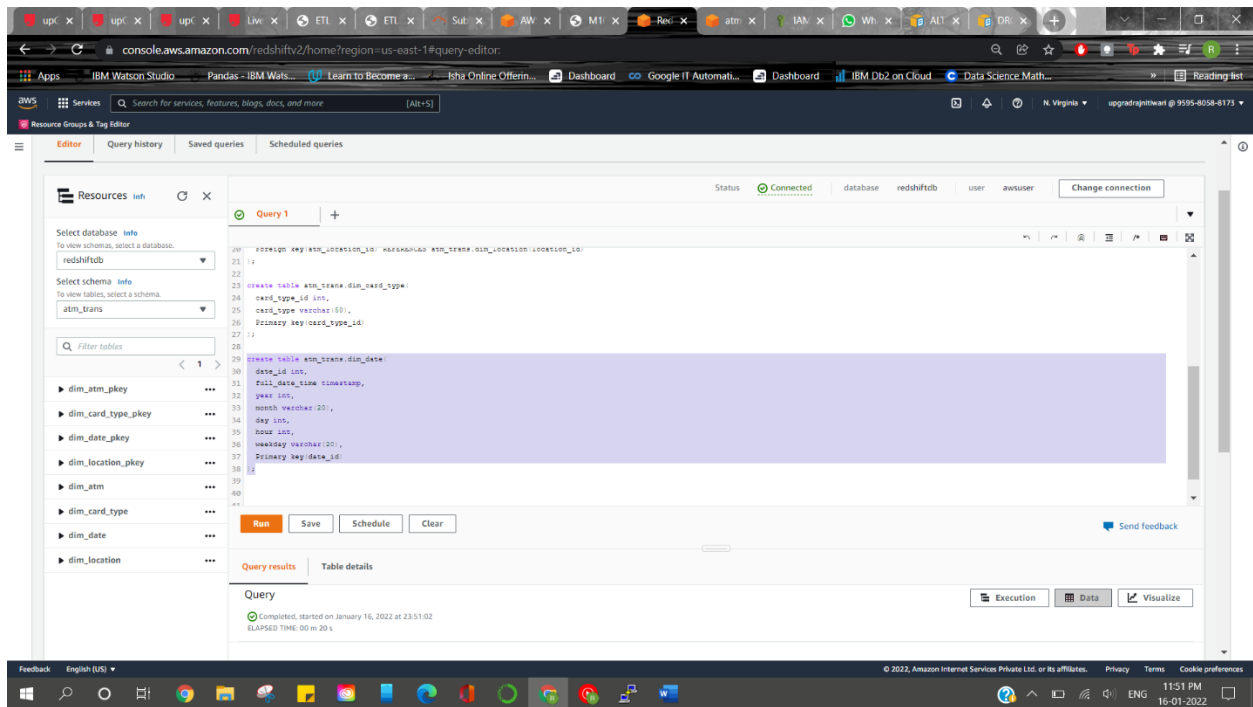
The screenshot shows the AWS Redshift console interface. On the left, the 'Resources' panel is expanded, showing the 'Select database' dropdown set to 'redshiftdb' and the 'Select schema' dropdown set to 'atm_txnns'. The 'Filter tables' search bar is empty. Below the search bar, a list of tables is shown, including 'dim_atm_pkey', 'dim_location_pkey', 'dim_atm', and 'dim_location'. The main editor area displays a SQL query for creating a table named 'atm_txnns' with the following schema:

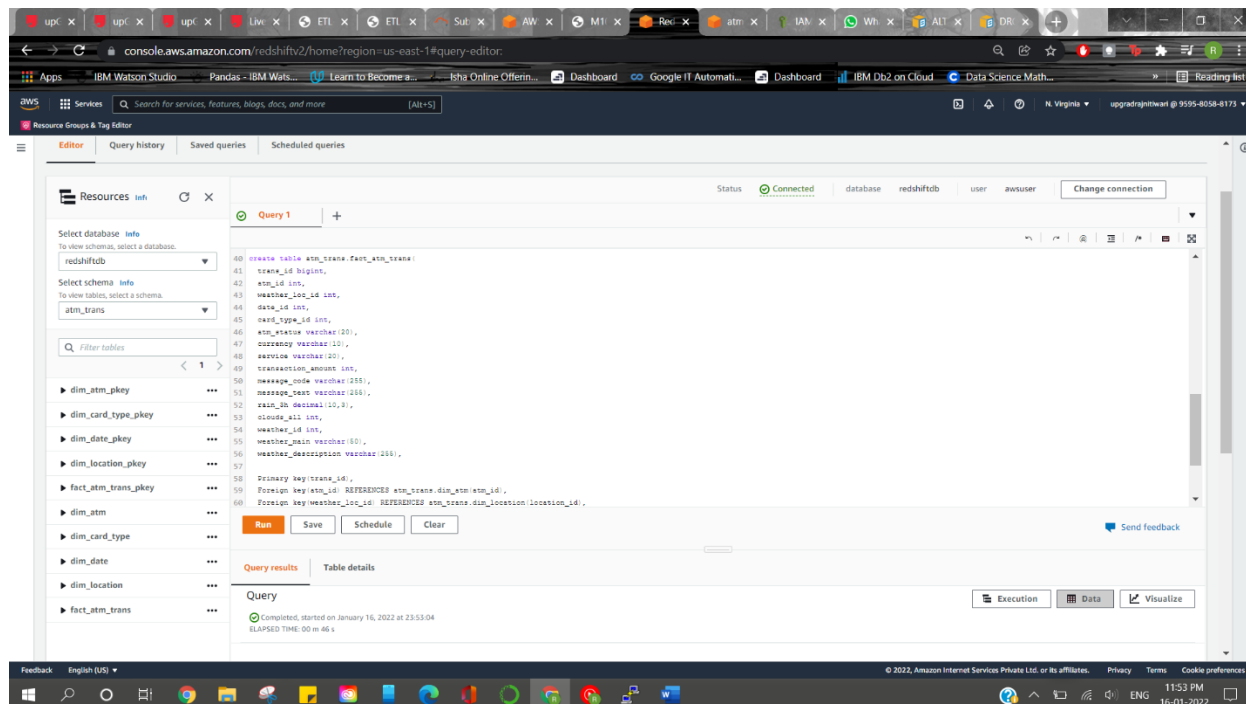
```

14 create table atm_txnns (
15   atm_id int,
16   atm_number varchar(50),
17   atm_manufacturer varchar(50),
18   atm_location_id int,
19   Primary key(atm_id),
20   Foreign key(atm_location_id) REFERENCES atm_txnns.dim_location(location_id)
21 );

```

Below the query editor, the 'Run' button is highlighted. The 'Query results' tab is selected, showing the query execution status: 'Completed, started on January 16, 2022 at 23:46:23' and 'ELAPSED TIME: 00 m 15 s'.





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

<Queries>

copy atm_trans.dim_location from

's3://atm-trans-tables/DIM_LOCATION/part-00000-25a884e6-b153-4803-b433-d183560a1c2f-c000.csv'

iam_role 'arn:aws:iam::959580588173:role/redshift_s3_fullaccess'

delimiter ',' region 'us-east-1';

copy atm_trans.dim_atm from

's3://atm-trans-tables/DIM_ATM/part-00000-55d00002-316c-4096-b260-6e989556c176-c000.csv'

iam_role 'arn:aws:iam::959580588173:role/redshift_s3_fullaccess'

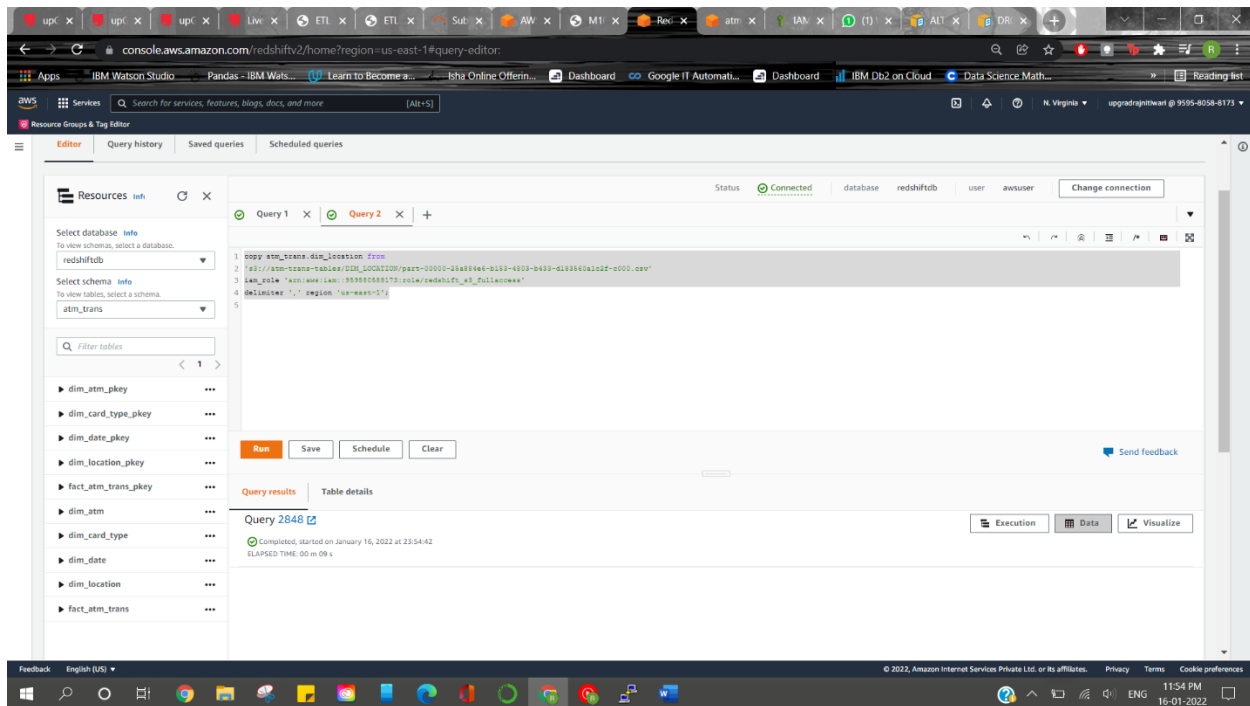
delimiter ',' region 'us-east-1';

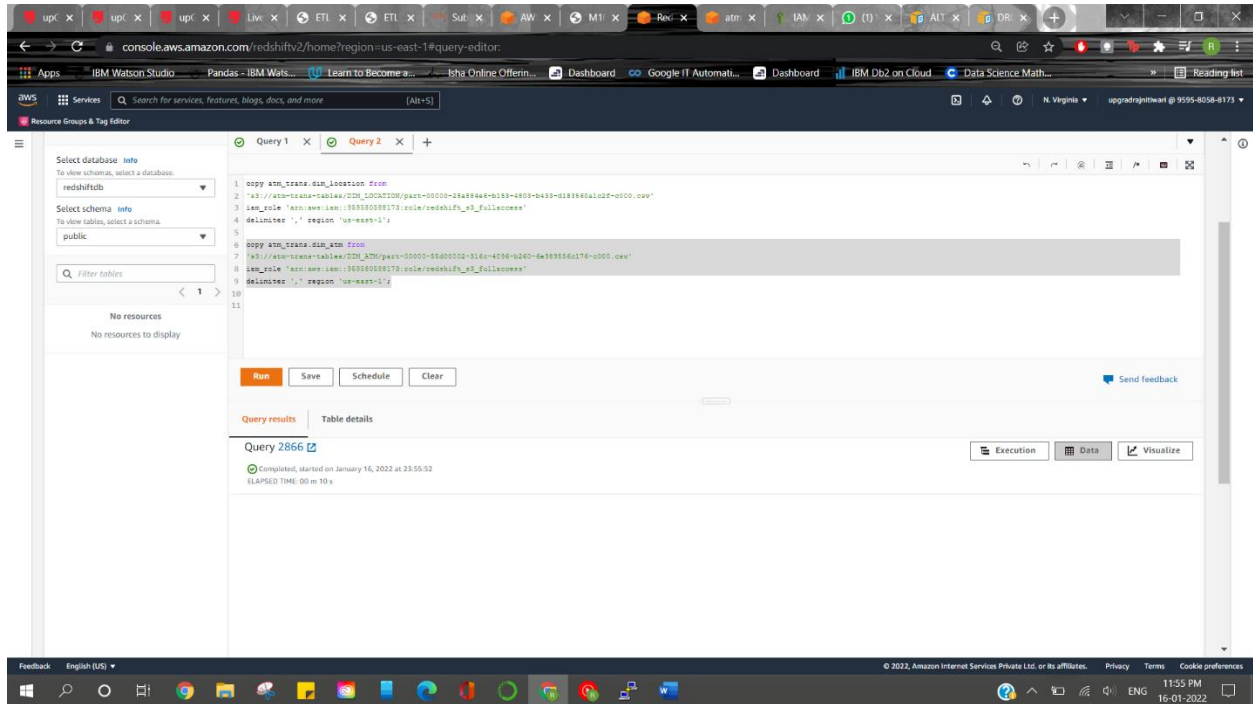
copy atm_trans.dim_card_type from

```
's3://atm-trans-tables/DIM_CARD_TYPE/part-00000-9fda7423-0ed4-49ba-8346-c2185b39028c-
c000.csv'
iam_role 'arn:aws:iam::959580588173:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1';
```

```
copy atm_trans.dim_date from
's3://atm-trans-tables/DIM_DATE/part-00000-944bb259-e8ce-46c9-9b93-ca28164b2de2-
c000.csv'
iam_role 'arn:aws:iam::959580588173:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
timeformat 'YYYY-MM-DDTHH:MI:SS'
CSV;
```

```
copy atm_trans.fact_atm_trans from
's3://atm-trans-tables/FACT_ATM_TRANS/part-00000-2cda0295-66d2-4ee9-bbb8-
5c1f1fe48df0-c000.csv'
iam_role 'arn:aws:iam::959580588173:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
CSV;
```

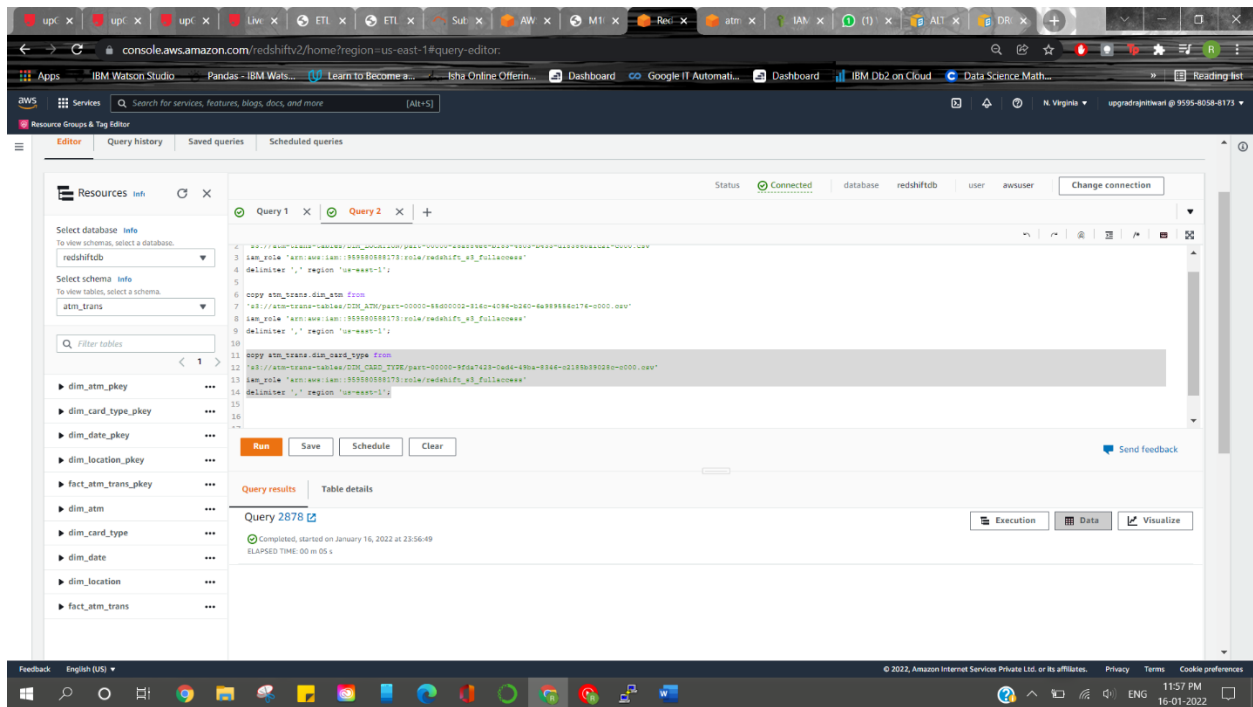




The screenshot shows the AWS Redshift console interface. On the left, the 'Select database' dropdown is set to 'redshiftdb', and the 'Select schema' dropdown is set to 'public'. The 'Filter tables' search bar is empty. The main editor area displays a SQL query (Query 1) with the following content:

```
1 copy atm_trans.dim_location from
2 's3://atm-trans-tables/CDM_LOCATION/part-00000-284844e-0183-0000-6430-6189580a102f-e000.csv'
3 aws_s3a 'arn:aws:s3:::19550008173:role/redshift_s3_fullaccess'
4 delimiter ',' region 'us-east-1';
5
6 copy atm_trans.dim_atm from
7 's3://atm-trans-tables/CDM_ATM/part-00000-8400002-01d0-408d-02d0-6499986d17e-e000.csv'
8 aws_s3a 'arn:aws:s3:::19550008173:role/redshift_s3_fullaccess'
9 delimiter ',' region 'us-east-1';
10
11
```

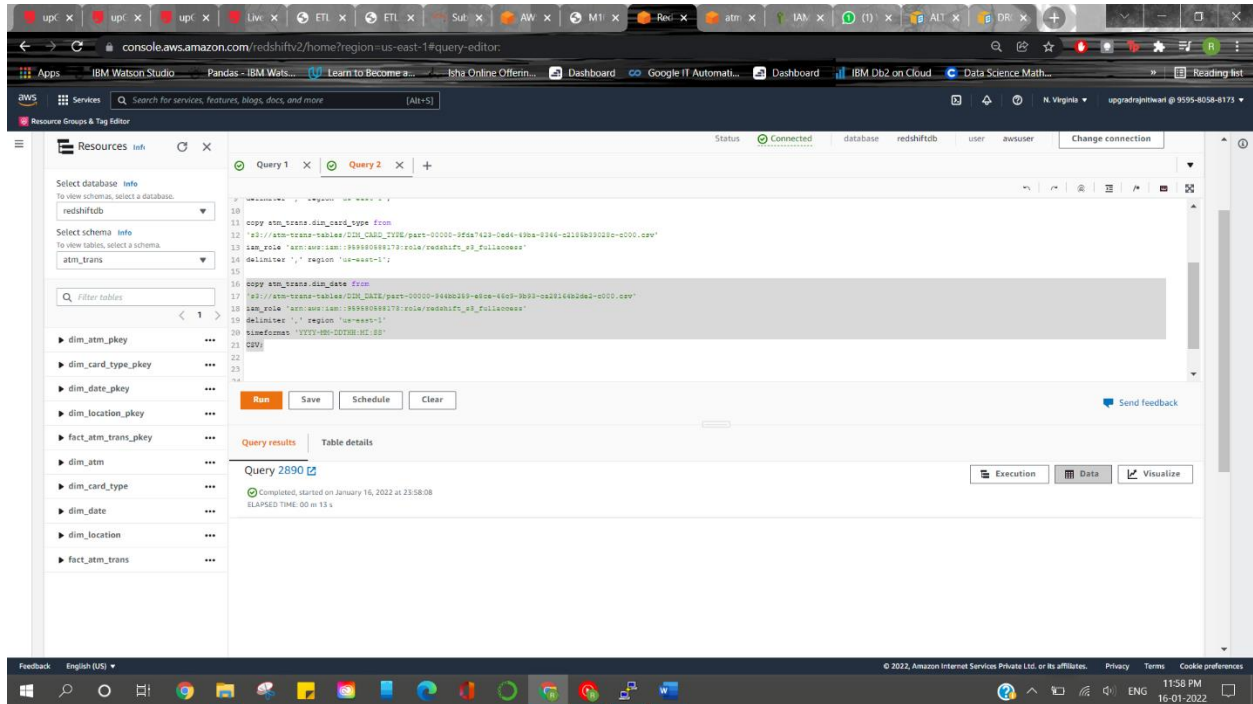
Below the query editor, the 'Run' button is highlighted. The 'Query results' tab is selected, showing 'Query 2866' with a status of 'Completed, started on January 16, 2022 at 23:55:52' and an 'ELAPSED TIME: 00 m 10 s'.



The screenshot shows the AWS Redshift console interface. On the left, the 'Select database' dropdown is set to 'redshiftdb', and the 'Select schema' dropdown is set to 'atm_trans'. The 'Filter tables' search bar is empty. The main editor area displays a SQL query (Query 2) with the following content:

```
1 copy atm_trans.dim_atm_type from
2 's3://atm-trans-tables/CDM_ATM_TYPE/part-00000-8f6a723-0add-40ba-836d-02185b39102b-e000.csv'
3 aws_s3a 'arn:aws:s3:::19550008173:role/redshift_s3_fullaccess'
4 delimiter ',' region 'us-east-1';
5
6 copy atm_trans.dim_atm_type from
7 's3://atm-trans-tables/CDM_ATM_TYPE/part-00000-8f6a723-0add-40ba-836d-02185b39102b-e000.csv'
8 aws_s3a 'arn:aws:s3:::19550008173:role/redshift_s3_fullaccess'
9 delimiter ',' region 'us-east-1';
10
11 copy atm_trans.dim_atm_type from
12 's3://atm-trans-tables/CDM_ATM_TYPE/part-00000-8f6a723-0add-40ba-836d-02185b39102b-e000.csv'
13 aws_s3a 'arn:aws:s3:::19550008173:role/redshift_s3_fullaccess'
14 delimiter ',' region 'us-east-1';
15
16
17
```

Below the query editor, the 'Run' button is highlighted. The 'Query results' tab is selected, showing 'Query 2878' with a status of 'Completed, started on January 16, 2022 at 23:56:49' and an 'ELAPSED TIME: 00 m 05 s'.

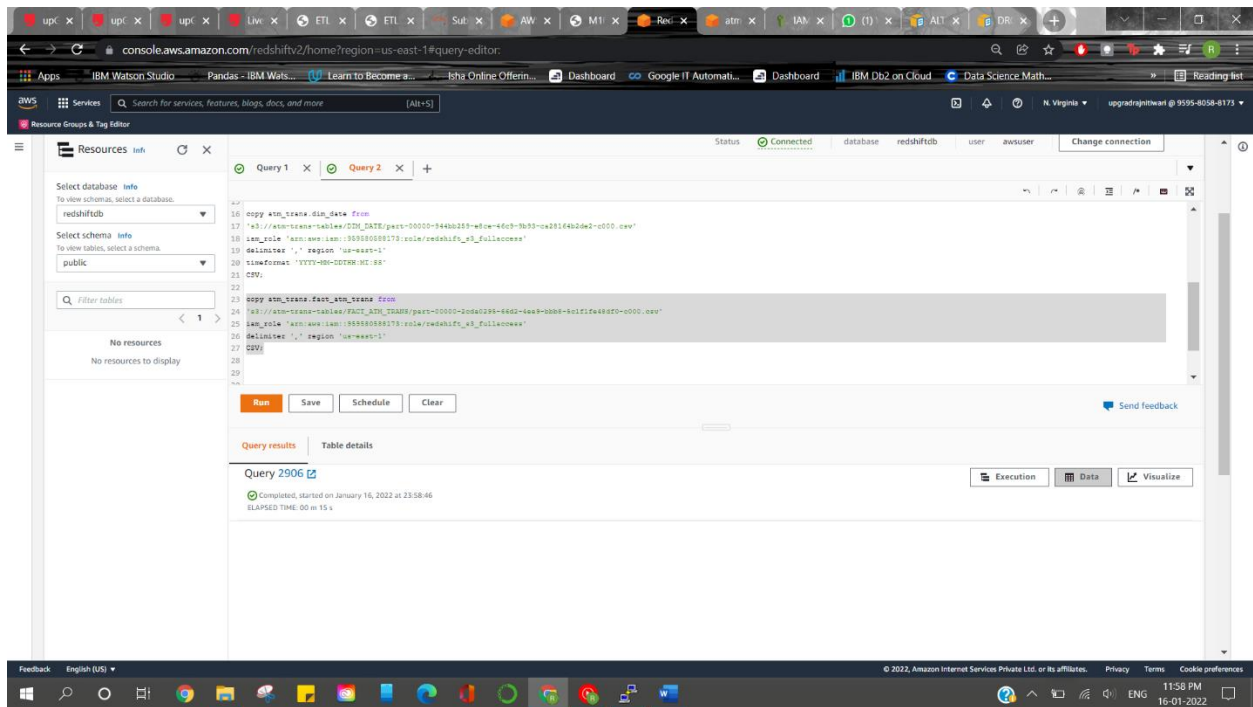


The screenshot shows the AWS Redshift console interface. On the left, the 'Resources' panel is open, showing the 'Select database' dropdown set to 'redshiftdb', the 'Select schema' dropdown set to 'atm_trans', and a list of tables including 'dim_atm_pkey', 'dim_card_type_pkey', 'dim_date_pkey', 'dim_location_pkey', 'fact_atm_trans_pkey', 'dim_atm', 'dim_card_type', 'dim_date', 'dim_location', and 'fact_atm_trans'. The main panel displays 'Query 1' with the following SQL code:

```

10 copy atm_trans_dta_type from
11 's3://atm-trans-tables/DIM_DATA/part-0000-9f6a123-0a6-49a-266-0210839020-e000.csv'
12 iam_role 'arn:aws:iam::39590889173:role/redshift_rol_fullaccess'
13 delimiter ',' region 'us-east-1';
14
15 copy atm_trans_dta_type from
16 's3://atm-trans-tables/DIM_DATA/part-0000-9f6a123-0a6-49a-266-0210839020-e000.csv'
17 iam_role 'arn:aws:iam::39590889173:role/redshift_rol_fullaccess'
18 delimiter ',' region 'us-east-1';
19
20 timeformat 'YYYY-MM-DDTHH:MM:SS'
21 CSV;
  
```

Below the SQL code, the 'Run' button is highlighted. The 'Query results' tab shows 'Query 2890' with a status of 'Completed, started on January 16, 2022 at 23:58:08' and an elapsed time of '00 m 15 s'.



The screenshot shows the AWS Redshift console interface. On the left, the 'Resources' panel is open, showing the 'Select database' dropdown set to 'redshiftdb', the 'Select schema' dropdown set to 'public', and a list of tables including 'dim_atm_pkey', 'dim_card_type_pkey', 'dim_date_pkey', 'dim_location_pkey', 'fact_atm_trans_pkey', 'dim_atm', 'dim_card_type', 'dim_date', 'dim_location', and 'fact_atm_trans'. The main panel displays 'Query 2' with the following SQL code:

```

16 copy atm_trans_dta_type from
17 's3://atm-trans-tables/DIM_DATA/part-0000-9f6a123-0a6-49a-266-0210839020-e000.csv'
18 iam_role 'arn:aws:iam::39590889173:role/redshift_rol_fullaccess'
19 delimiter ',' region 'us-east-1';
20
21 timeformat 'YYYY-MM-DDTHH:MM:SS'
22 CSV;
23
24 copy atm_trans_dta_type from
25 's3://atm-trans-tables/FACT_DATA/part-0000-9f6a123-0a6-49a-266-0210839020-e000.csv'
26 iam_role 'arn:aws:iam::39590889173:role/redshift_rol_fullaccess'
27 delimiter ',' region 'us-east-1';
28
29 timeformat 'YYYY-MM-DDTHH:MM:SS'
30 CSV;
  
```

Below the SQL code, the 'Run' button is highlighted. The 'Query results' tab shows 'Query 2906' with a status of 'Completed, started on January 16, 2022 at 23:58:46' and an elapsed time of '00 m 15 s'.