

Creation of a Redshift Cluster

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

← →

https://us-east-1.console.aws.amazon.com/redshiftv2/home?region=us-east-1#create-cluster

☆ 📄 📄 📄 📄

aws

Services 🔍 Search [Alt+S]

📧 🔄 🛑 🌐

N. Virginia

vodlabs/user3050060=ducran2306@gmail.com @ 6374-2349-8116

Amazon Redshift > Clusters > Create cluster

Create cluster info

📘 Looking for free trial? Try Redshift Serverless. First-time Redshift Serverless customers receive a \$300 credit to use in their account. Launch Redshift Serverless ✕

Cluster configuration

Cluster identifier

This is the unique key that identifies a cluster.

redshift-cluster-1

The identifier must be from 1-63 characters. Valid characters are a-z (lowercase only) and - (hyphen).

Choose the size of the cluster

☒ I'll choose

☐ Help me choose

Node type info

Choose a node type that meets your CPU, RAM, storage capacity, and drive type requirements.

dc2.large ▾

Number of nodes

Enter the number of nodes that you need.

1

Range (1-32)

Configuration summary info

dc2.large | 1 node

\$182.50/month

Estimated on-demand compute price

160 GB

Total compressed storage

The total storage capacity for the

Database configurations

Admin user name

Enter a login ID for the admin user of your DB instance.

awsuser

The name must be 1-128 alphanumeric characters, and it can't be a reserved word 📄

Admin password

Select an option to manage your admin password.

☐ Manage admin credentials in AWS Secrets Manager info

AWS manages a KMS key that encrypts your data.

☐ Generate a password

Amazon Redshift generates an admin password.

☒ Manually add the admin password



Manually enter the admin password.

Admin user password

Must be 8-64 characters long. Must contain at least one uppercase letter, one lowercase letter and one number. Can be any printable ASCII character except "/", "", or "gp".

☐ Show password

Cluster permissions


 Create an IAM role as the default for this cluster that has the [AmazonRedshiftAllCommandsFullAccess](#)  policy attached. This policy includes permissions to run SQL commands to COPY, UNLOAD, and query data with Amazon Redshift. The policy also grants permissions to run SELECT statements for related services, such as Amazon S3, Amazon CloudWatch logs, Amazon SageMaker, and AWS Glue.

Associated IAM roles (1) [Info](#)


Set default ▼

Manage IAM roles ▼

Create, associate, or remove an IAM role. You can associate up to 50 IAM roles. You can also choose an IAM role and set it as the default for this cluster.

 Search for associated IAM role by name, status, or role type

< 1 >

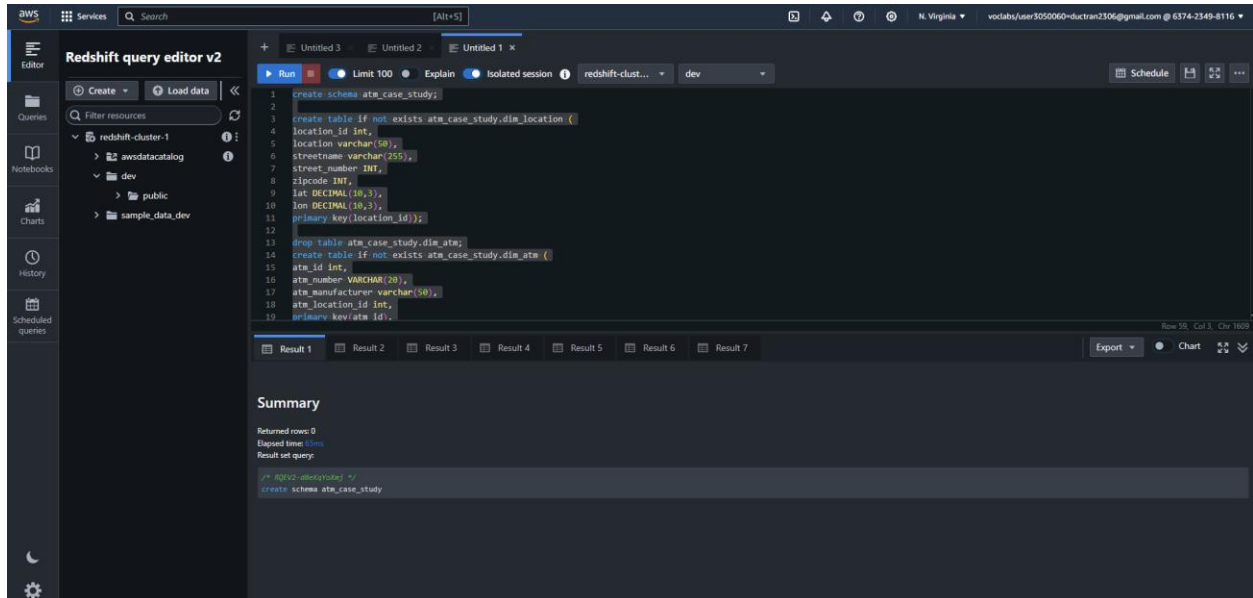
<input type="checkbox"/>	IAM roles 	▼	Status	▼	Role type	▼
<input type="checkbox"/>	myRedshiftRole		Not applied		--	

IAM role with read only on S3.

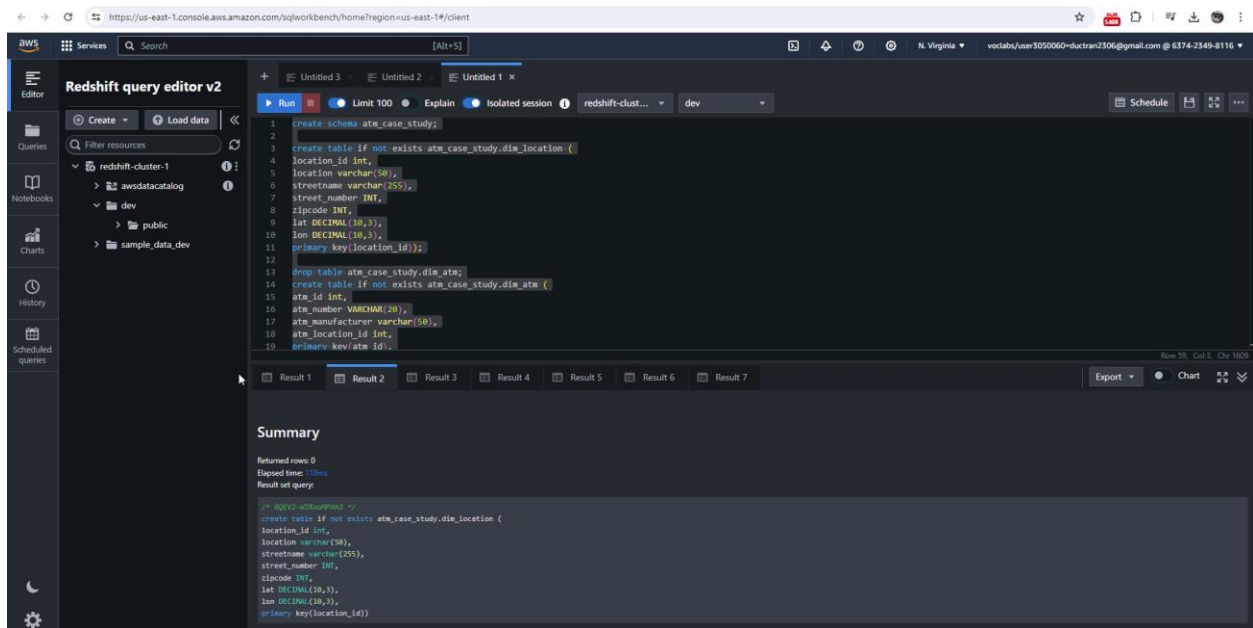
Setting up a database in the Redshift cluster and running queries to create the dimension and fact tables

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

create schema atm_case_study;



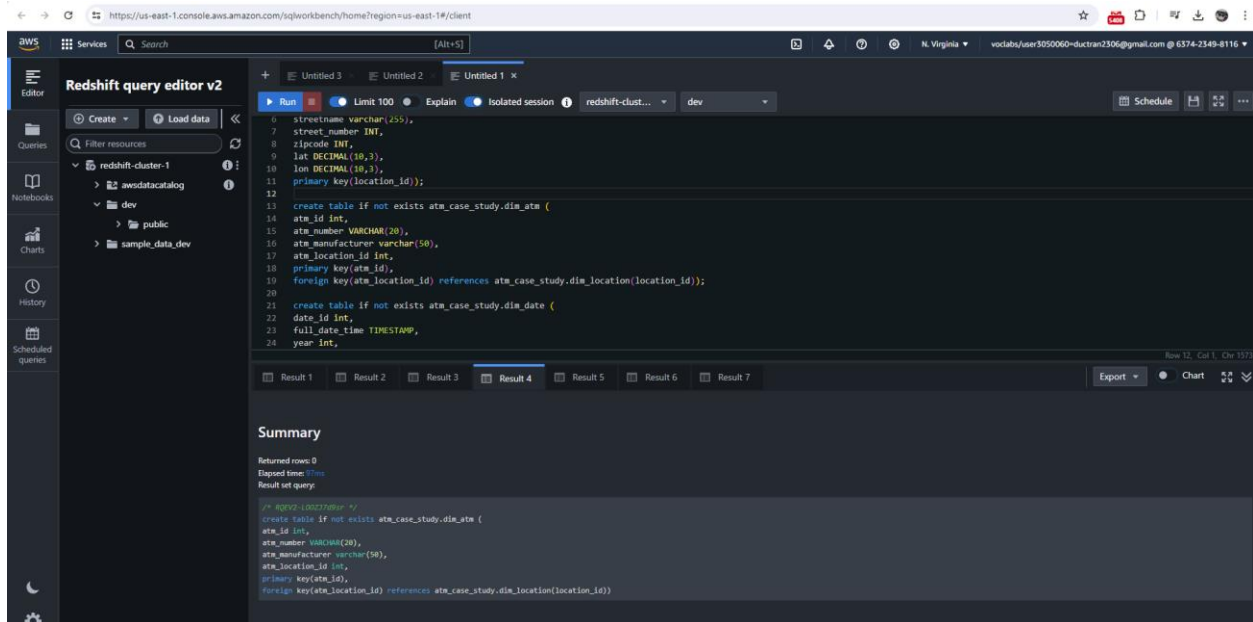
```
create table if not exists atm_case_study.dim_location (  
location_id int,  
location varchar(50),  
streetname varchar(255),  
street_number INT,  
zipcode INT,  
lat DECIMAL(10,3),  
lon DECIMAL(10,3),  
primary key(location_id));
```



```

create table if not exists atm_case_study.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_case_study.dim_location(location_id));

```

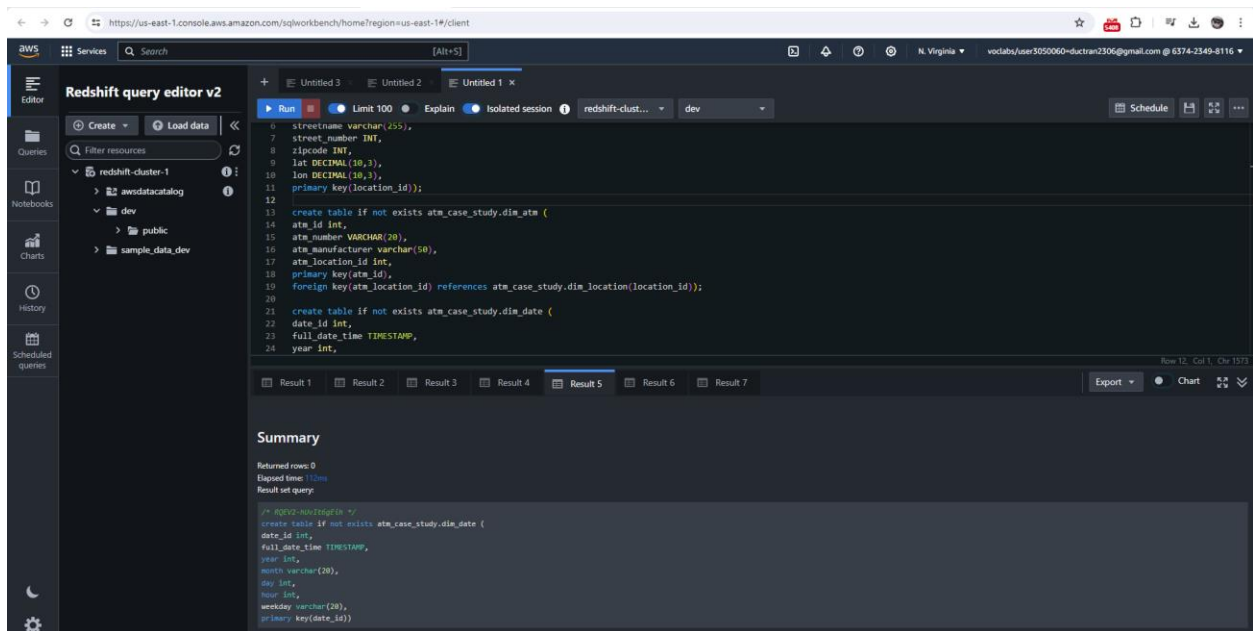


```

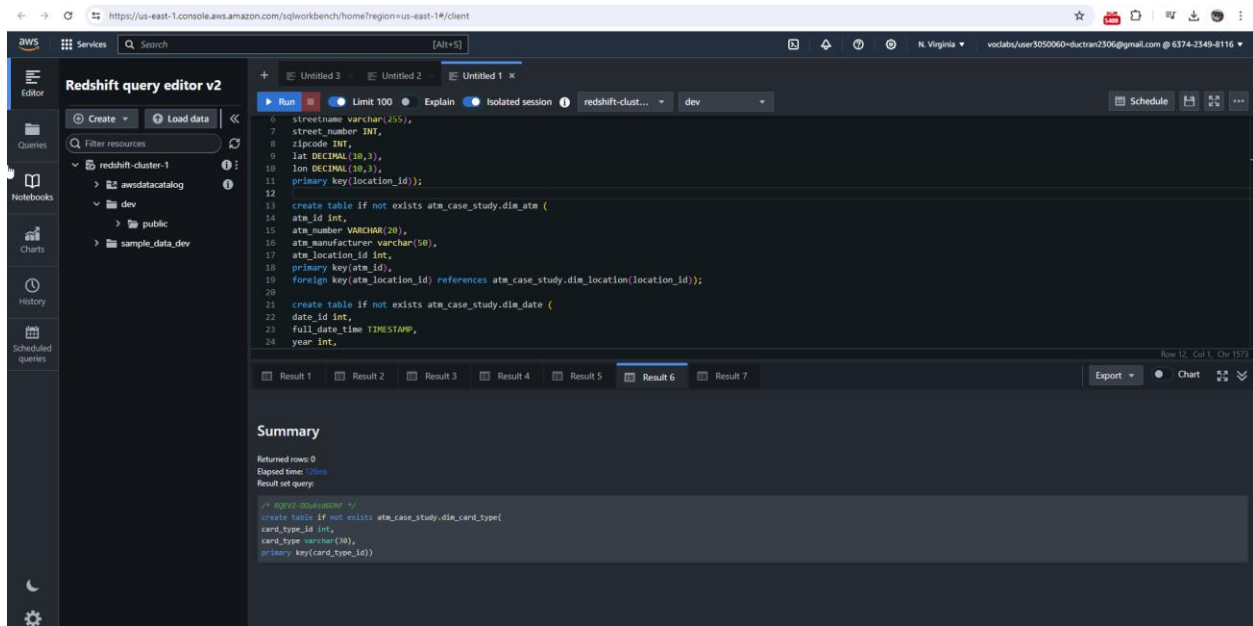
create table if not exists atm_case_study.dim_date (
  date_id int,

```

```
full_date_time TIMESTAMP,  
year int,  
month varchar(20),  
day int,  
hour int,  
weekday varchar(20),  
primary key(date_id));
```



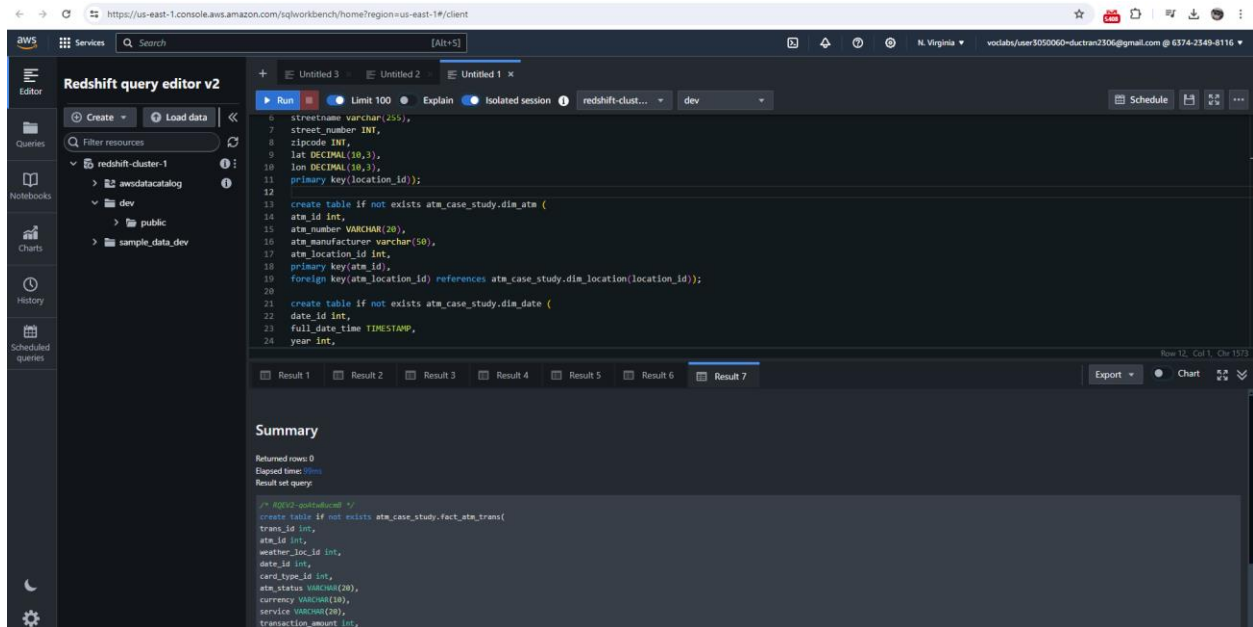
```
create table if not exists atm_case_study.dim_card_type(  
card_type_id int,  
card_type varchar(30),  
primary key(card_type_id));
```



```

create table if not exists atm_case_study.fact_atm_trans(
trans_id int,
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_case_study.dim_atm(atm_id),
foreign key(weather_loc_id) references atm_case_study.dim_location(location_id),
foreign key(date_id) references atm_case_study.dim_date(date_id),
foreign key(card_type_id) references atm_case_study.dim_card_type(card_type_id)
);

```



Loading data into a Redshift cluster from Amazon S3 bucket

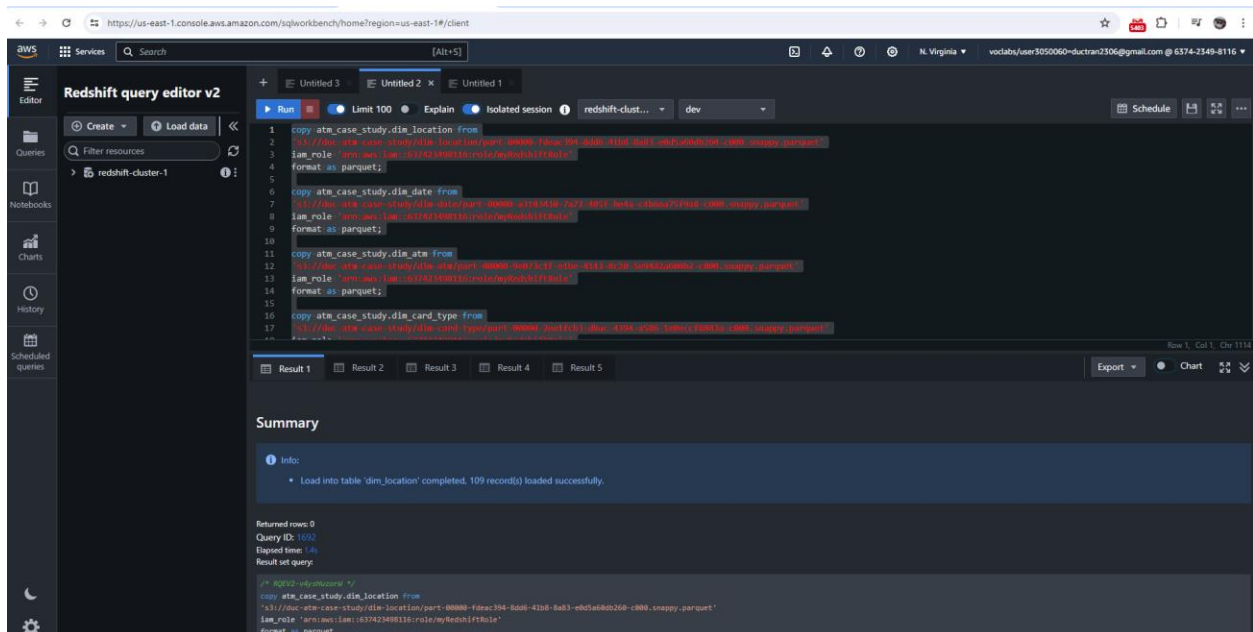
Queries to copy the data from S3 buckets to the Redshift cluster in the appropriate tables

copy atm_case_study.dim_location from

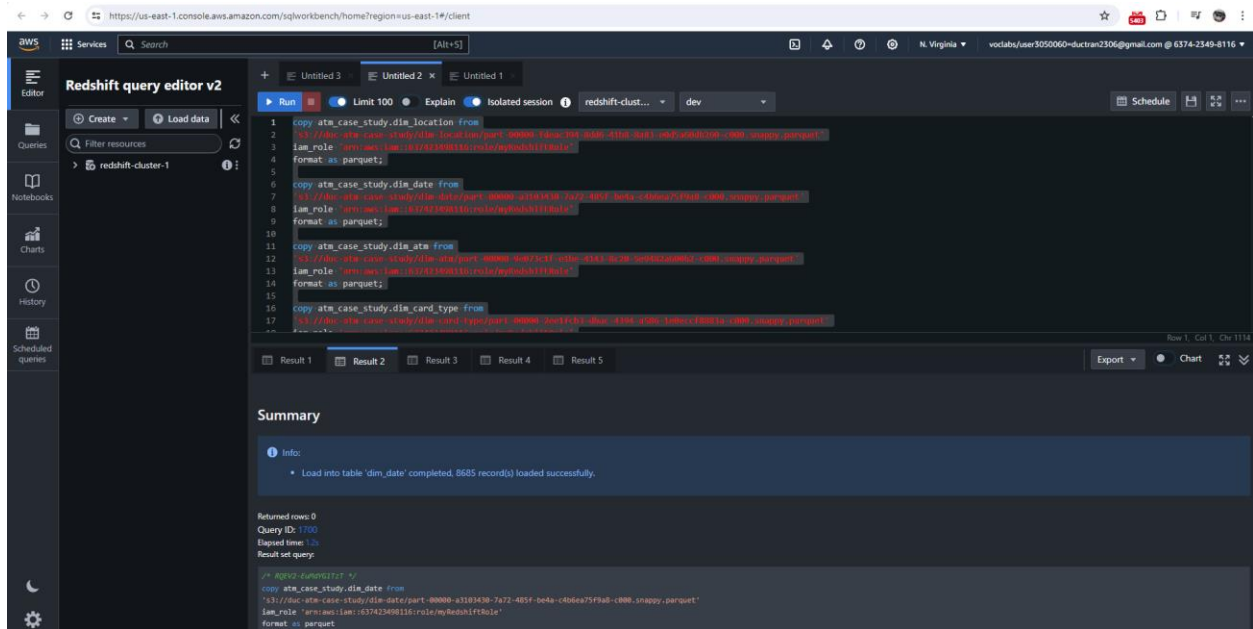
's3://duc-atm-case-study/dim-location/part-00000-6edc4220-c1b7-46aa-b65b-c341364c9f2d-c000.snappy.parquet'

iam_role 'arn:aws:iam::637423498116:role/myRedshiftRole'

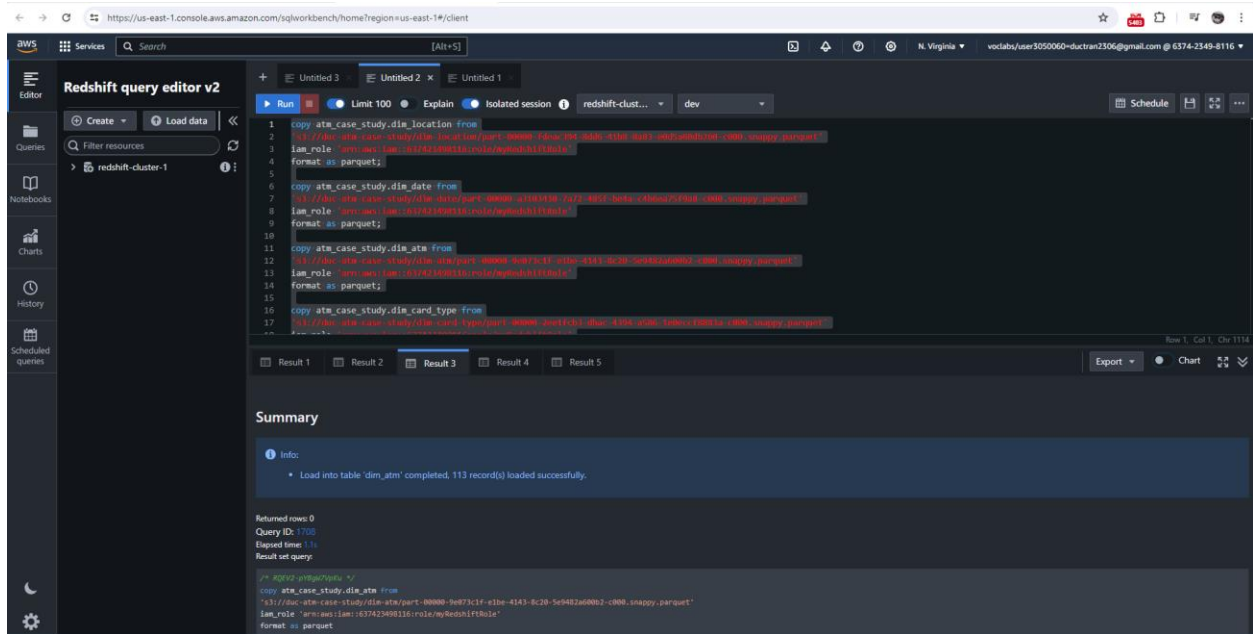
format as parquet;



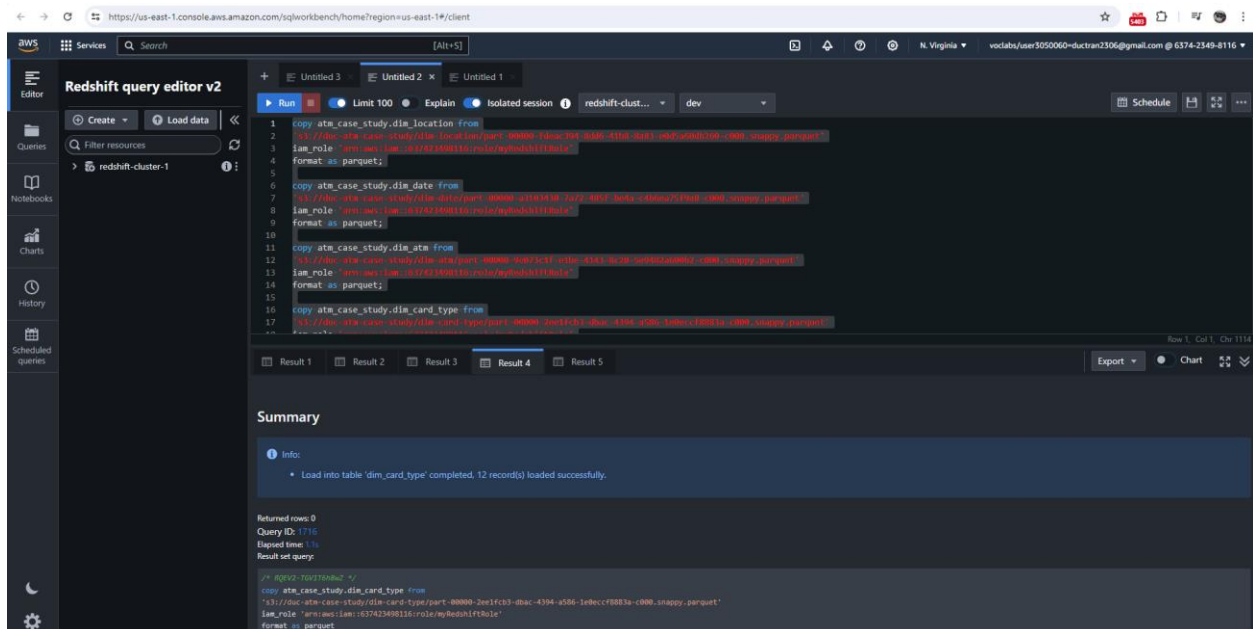

```
copy atm_case_study.dim_date from
's3://duc-atm-case-study/dim-date/part-00000-f7e92d9e-50d0-4dc3-8140-82f4195db2f7-
c000.snappy.parquet'
iam_role 'arn:aws:iam::637423498116:role/myRedshiftRole'
format as parquet;
```



```
copy atm_case_study.dim_atm from
's3://duc-atm-case-study/dim-atm/part-00000-cc1004a8-68fd-4e49-9180-8895aa9d4a5b-
c000.snappy.parquet'
iam_role 'arn:aws:iam::637423498116:role/myRedshiftRole'
format as parquet;
```



copy atm_case_study.dim_card_type from
's3://duc-atm-case-study/dim-card-type/part-00000-eba3ac17-33eb-4b80-a9ce-b791a3735287-c000.snappy.parquet'
iam_role 'arn:aws:iam::637423498116:role/myRedshiftRole'
format as parquet;



copy atm_case_study.fact_atm_trans from
's3://duc-atm-case-study/trans-atm/part-00000-58d17713-5b3b-4de0-ba59-6dc416929cf1-c000.snappy.parquet'

The screenshot displays the AWS Redshift Query Editor v2 interface. The top navigation bar includes the AWS logo, a 'Services' menu, a search bar, and the user's profile 'N. Virginia'. The left sidebar contains icons for Editor, Queries, Query Editor, Charts, History, and Scheduled queries. The main workspace is titled 'Redshift query editor v2' and shows a SQL query being executed. The query is a multi-part INSERT statement into the 'fact_atm_trans' table. The bottom section displays the 'Summary' of the query execution, indicating it was successful and loaded 2468572 records. The query ID is 17104 and the elapsed time is 5/s. The result set query is also shown.