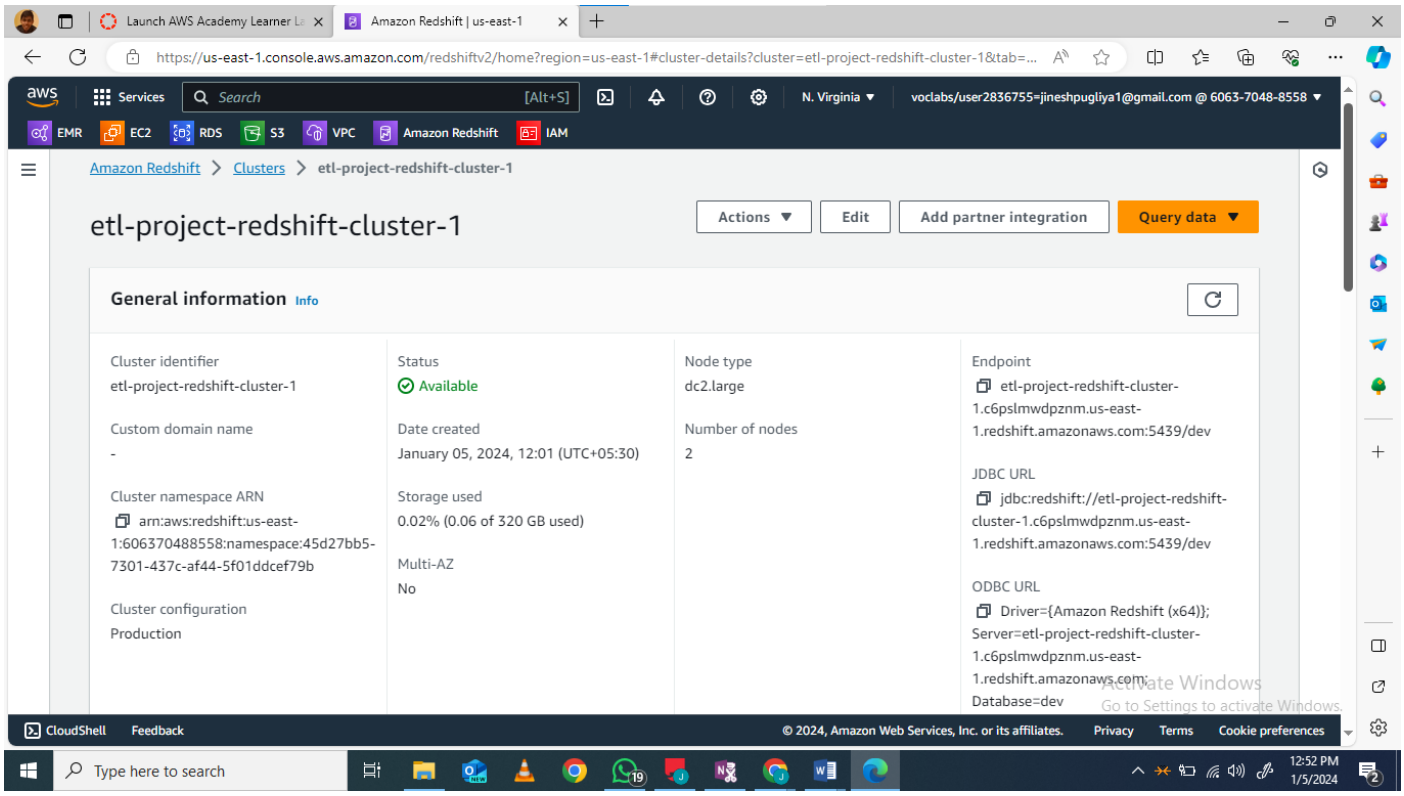
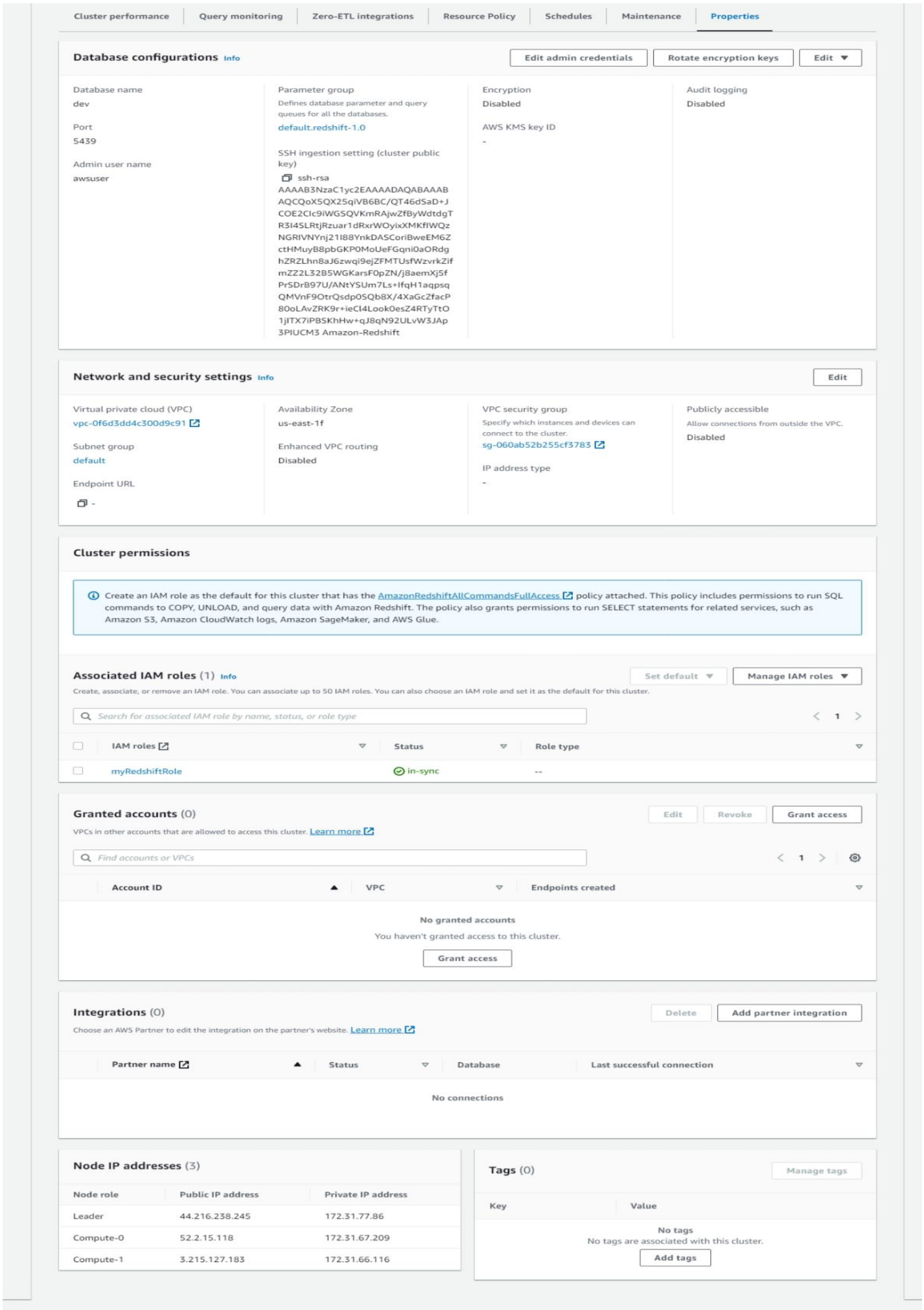
Creation of a Redshift Cluster

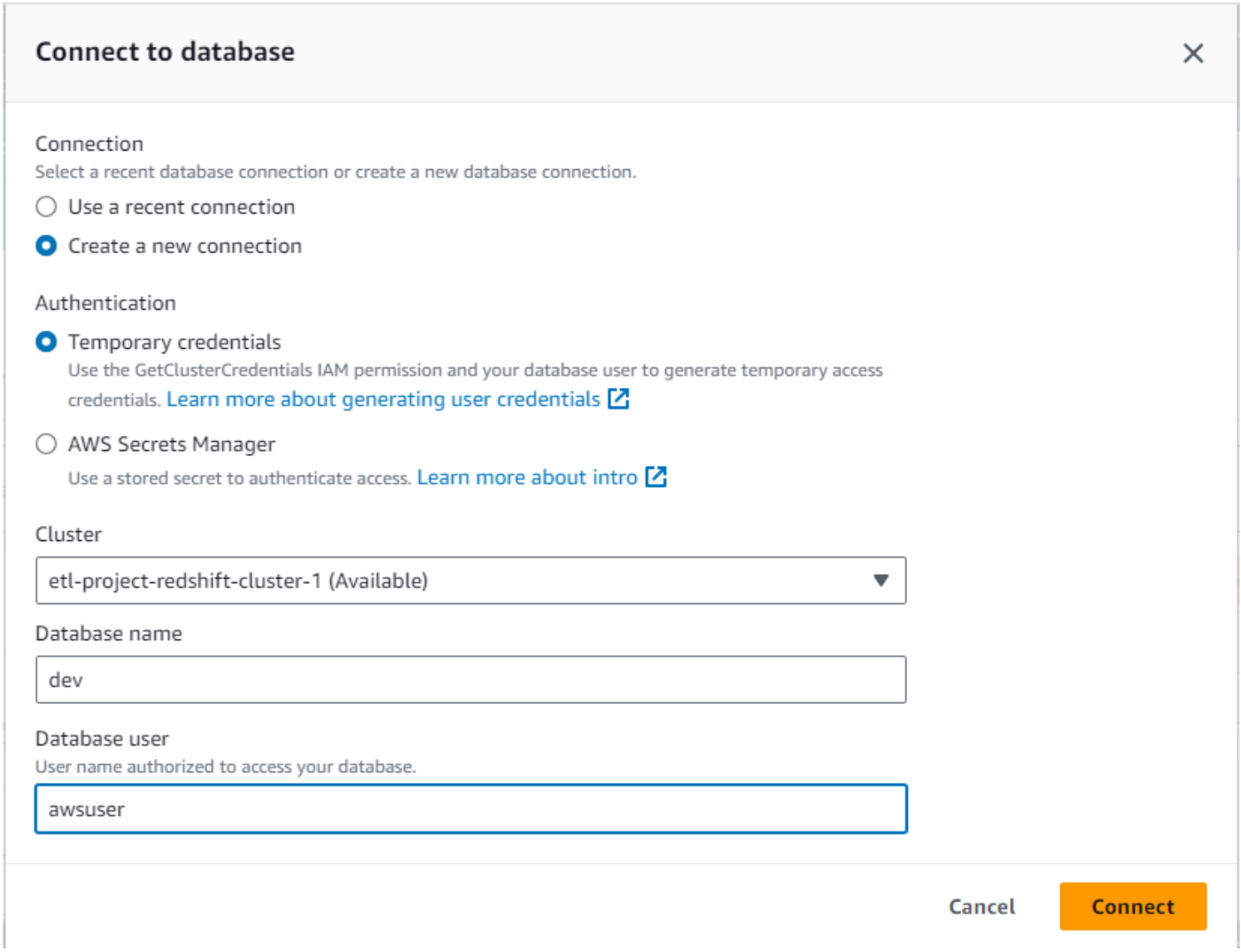
Screenshots of the configuration of the Redshift cluster that you have created:




Cluster performance
Query monitoring
Zero-ETL integrations
Resource Policy
Schedules
Maintenance
Properties
Database configurations Info
Edit admin credentials
Rotate encryption keys
Edit
Database name
dev
Port
5439
Admin user name
awsuser
Parameter group
Defines database parameter and query queues for all the databases.
default.redshift-1.0
SSH ingestion setting (cluster public key)
ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAABAQCQoX5QX25qiVB6BC/QT46dSaD+JCOE2CIc9iWG5QVKmRAjwZfByWdtdgTR3I4SLRtjRzuar1dRxrWOyixXMKfIWQzNGRIVNYnj21I88YnkDASCoriBweEM6ZctHMuyB8pbGKP0MoUeFGqni0aORdghZRZLhn8aJ6zwqi9ejZFMTUsfWzvrkZifmZZ2L32B5WGKarsF0pZN/j8aemXj5fPrSDrB97U/ANtYSUm7Ls+IfqH1aqpsqQMVnF9OtrQsdp0SQb8X/4XaGcZfacP80oLAvZRK9r+ieCl4Look0esZ4RTyTtO1jITX7iPBSKhHw+qJ8qN92ULvW3JAp3PIUCM3 Amazon-Redshift
Encryption
Disabled
AWS KMS key ID
-
Audit logging
Disabled
Network and security settings Info
Edit
Virtual private cloud (VPC)
vpc-0f6d3dd4c300d9c91
Subnet group
default
Endpoint URL
-
Availability Zone
us-east-1f
Enhanced VPC routing
Disabled
VPC security group
Specify which instances and devices can connect to the cluster.
sg-060ab52b255cf3783
IP address type
-
Publicly accessible
Allow connections from outside the VPC.
Disabled
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
IAM roles
Status
Role type
myRedshiftRole
in-sync
--
Granted accounts (0)
Edit
Revoke
Grant access
VPCs in other accounts that are allowed to access this cluster. Learn more
Find accounts or VPCs
1
Account ID
VPC
Endpoints created
No granted accounts
You haven't granted access to this cluster.
Grant access
Integrations (0)
Delete
Add partner integration
Choose an AWS Partner to edit the integration on the partner's website. Learn more
Partner name
Status
Database
Last successful connection
No connections
Node IP addresses (3)
Node role
Public IP address
Private IP address
Leader
44.216.238.245
172.31.77.86
Compute-0
52.2.15.118
172.31.67.209
Compute-1
3.215.127.183
172.31.66.116
Tags (0)
Manage tags
Key
Value
No tags
No tags are associated with this cluster.
Add tags

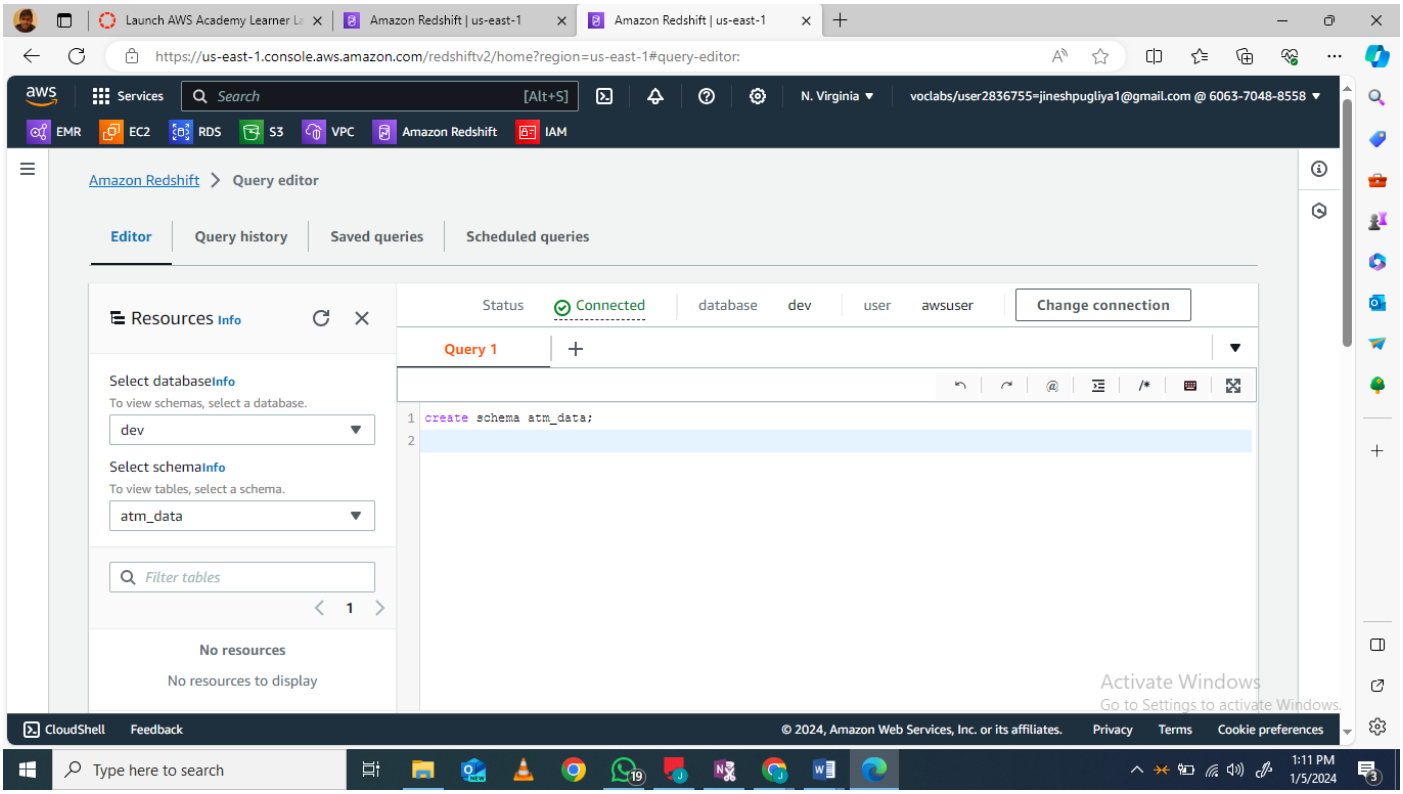
Setting up a database in the Redshift cluster and running queries to create the dimension and fact tables

Connect to database:



Query for creating schema:

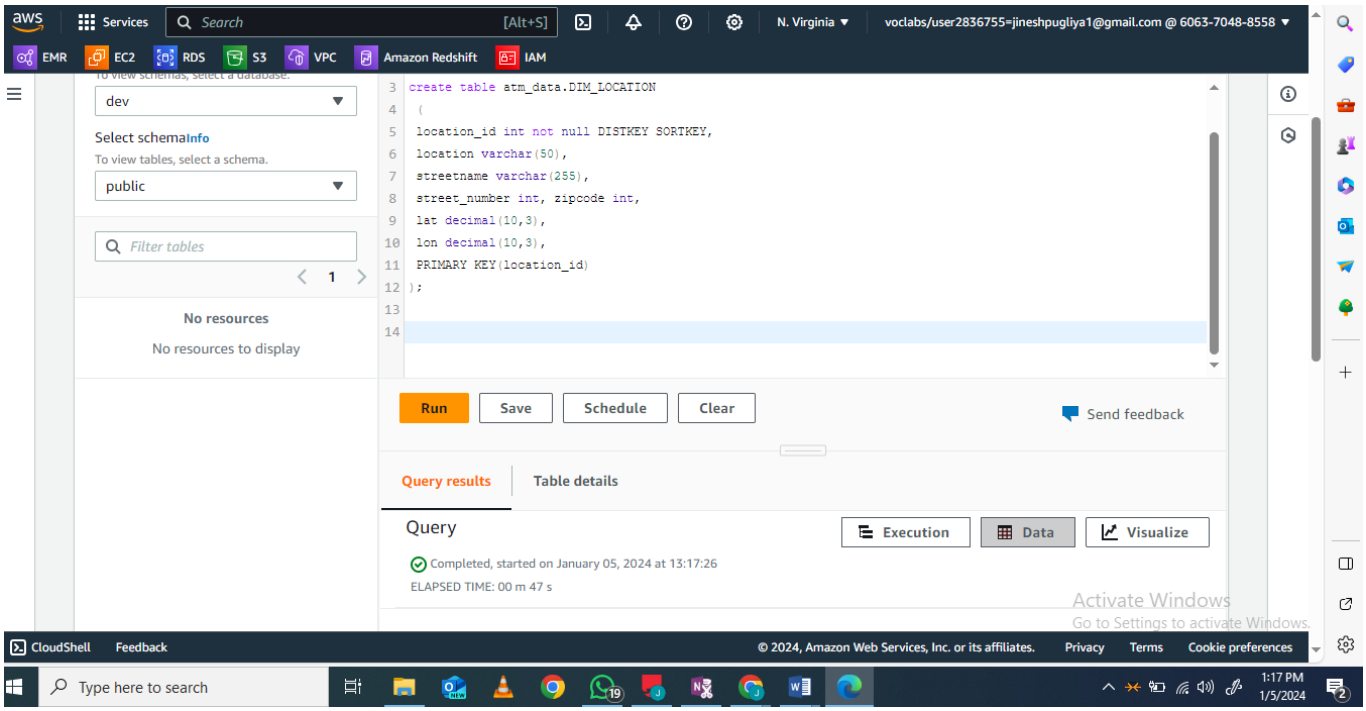
create schema atm_data;



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

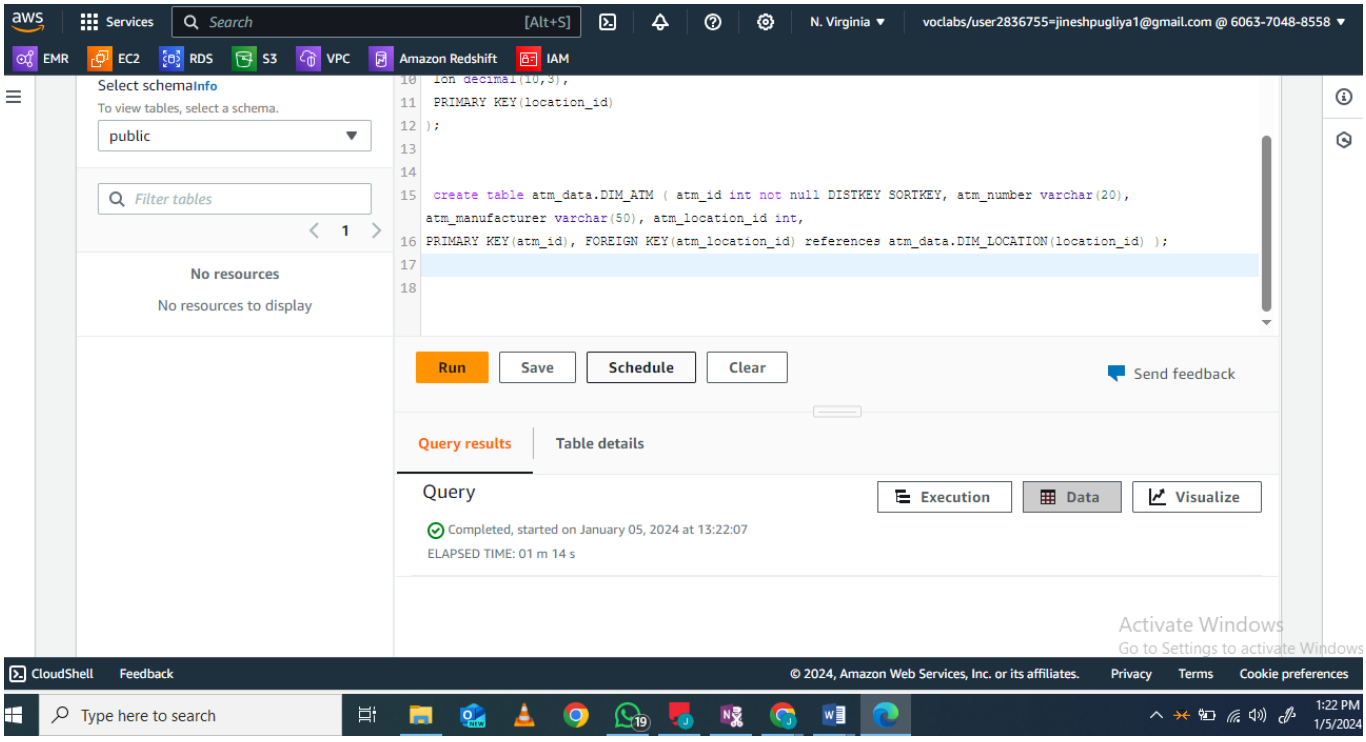
• Creating location dimension table

```
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)
);
```



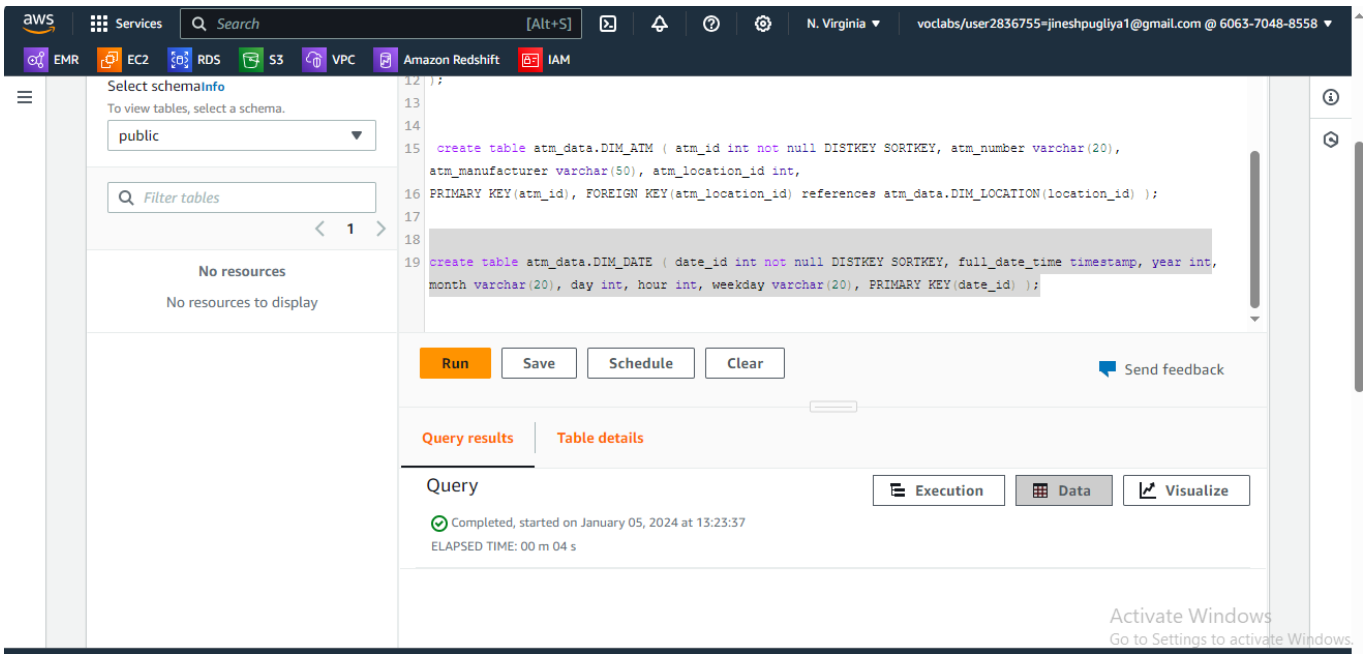
• Creating atm dimension table

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));



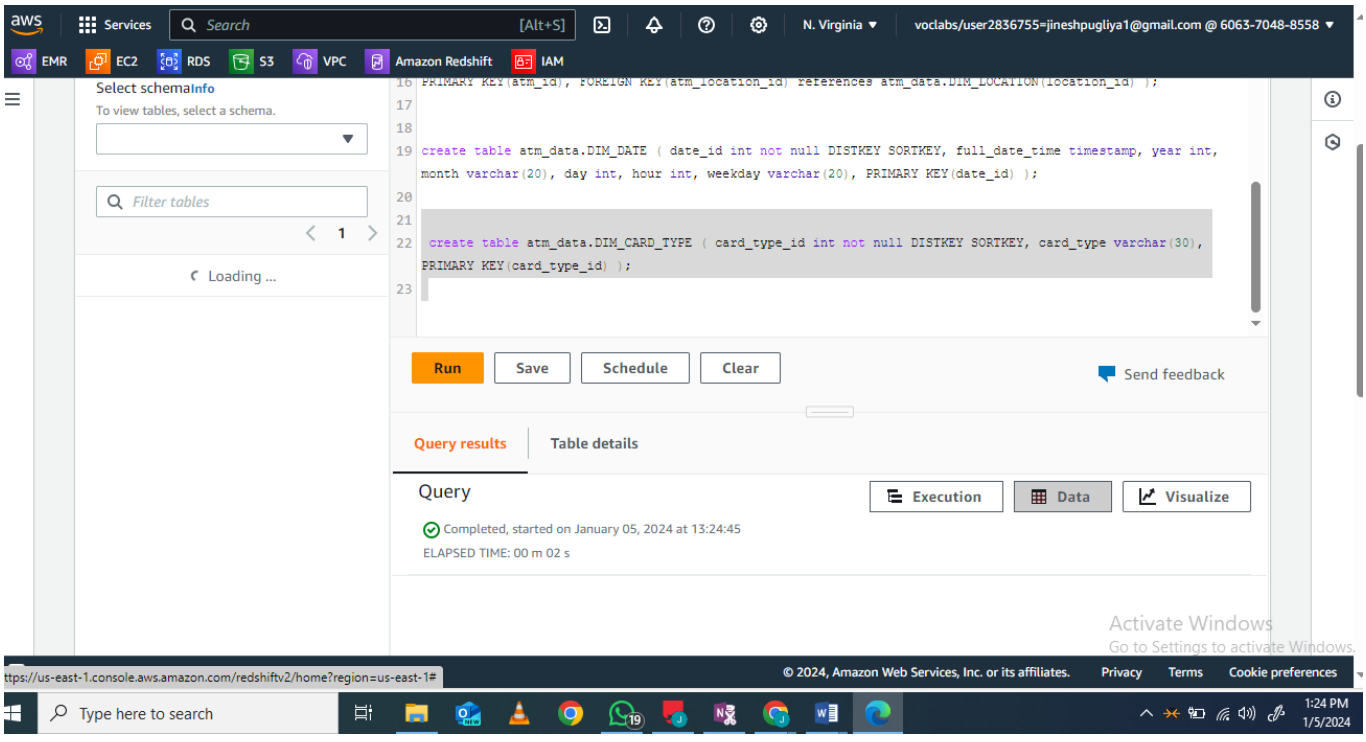
• Creating date dimension table

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));



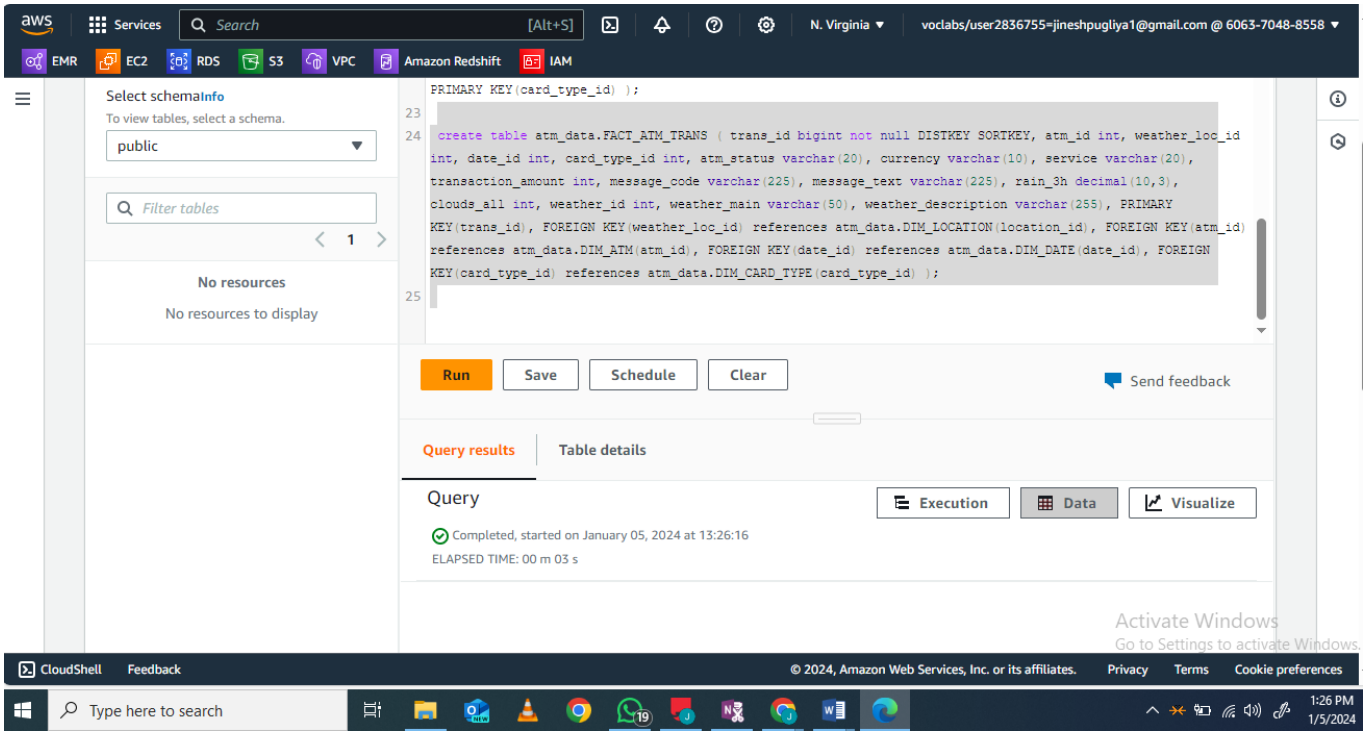
Creating card type dimension table

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));

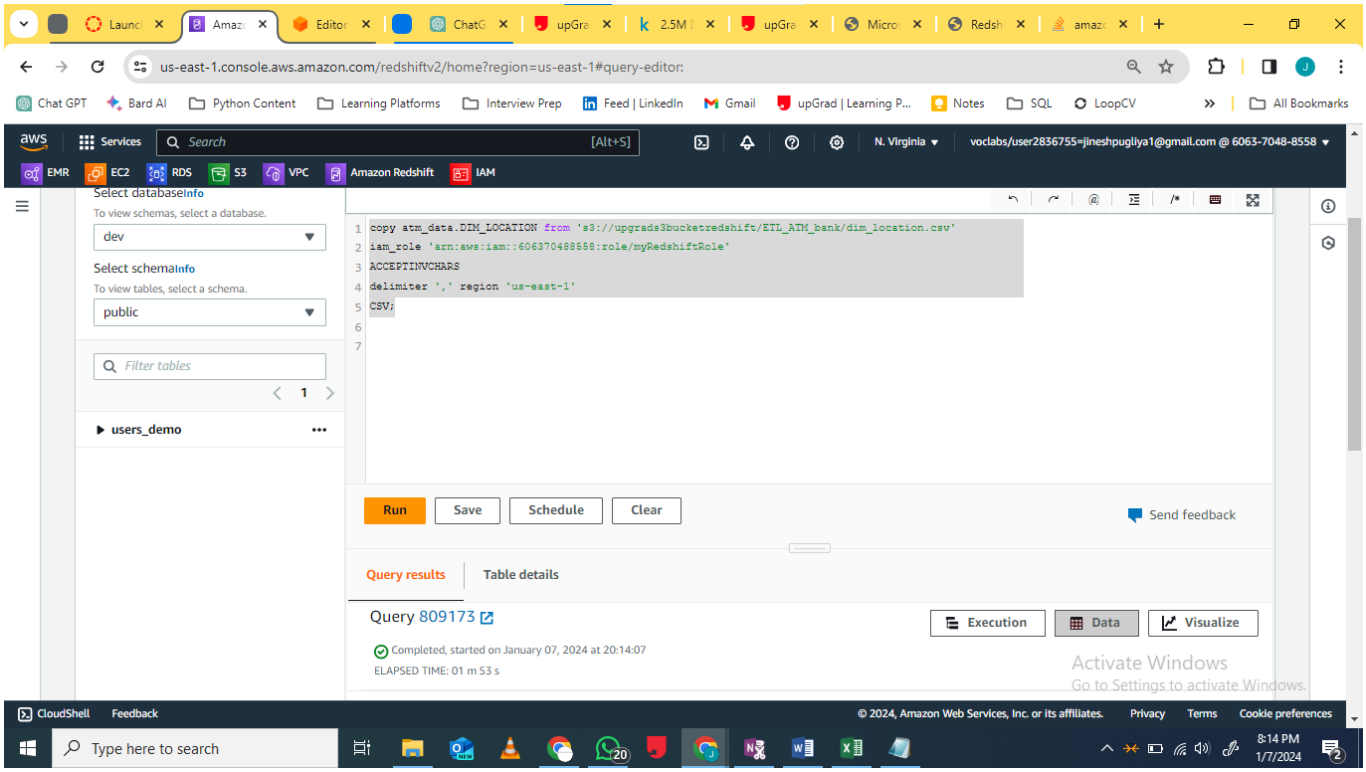


Loading data into a Redshift cluster from Amazon S3 bucket

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

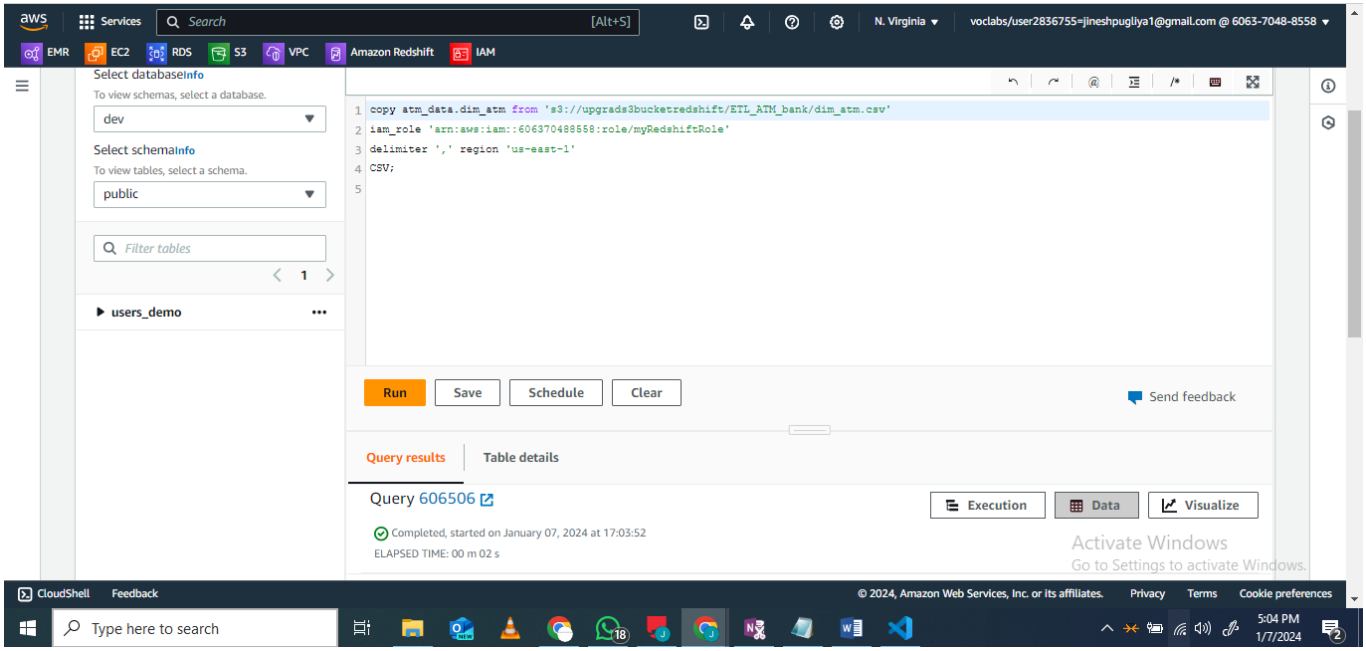
• Copying the data to dim_location table

```
copy atm_data.DIM_LOCATION from
's3://upgrads3bucketredshift/ETL_ATM_bank/dim_location.csv'
iam_role 'arn:aws:iam::606370488558:role/myRedshiftRole'
ACCEPTINVCHARS
delimiter ',' region 'us-east-1'
CSV;
```

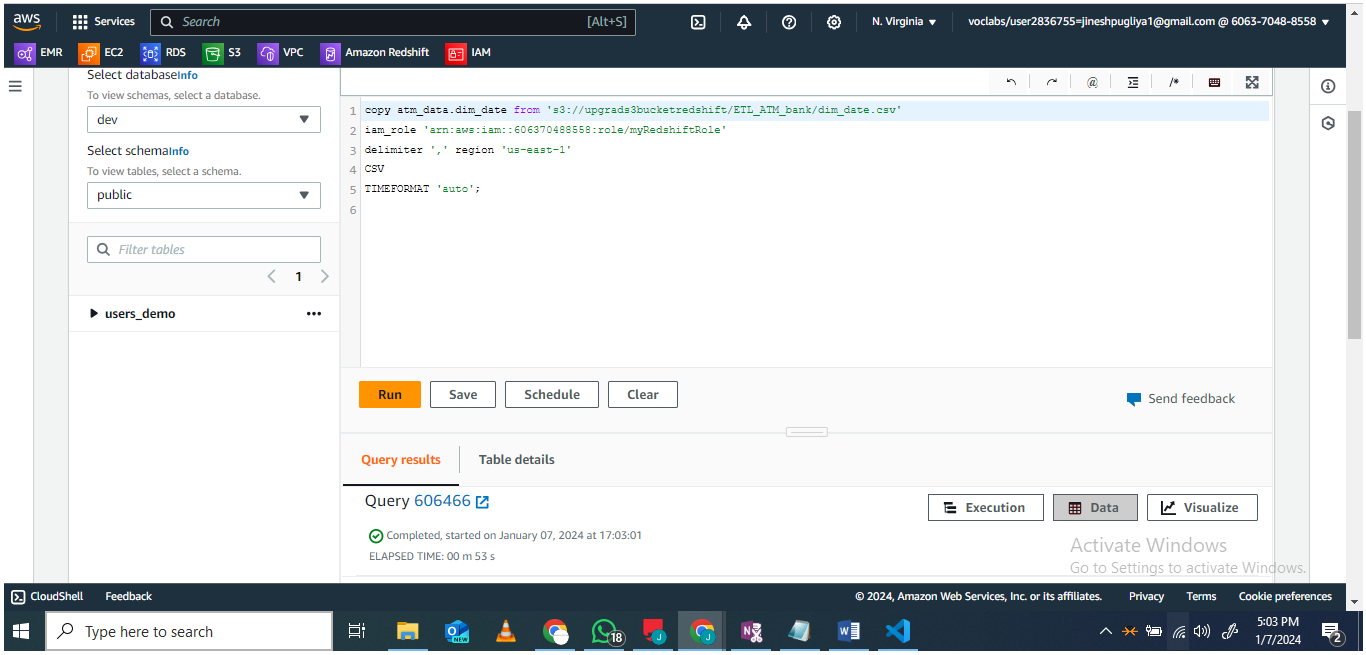


• Copying the data to dim_atm table

copy atm_data.dim_atm from 's3://upgrads3bucketredshift/ETL_ATM_bank/dim_atm.csv'
iam_role 'arn:aws:iam::606370488558:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
CSV;

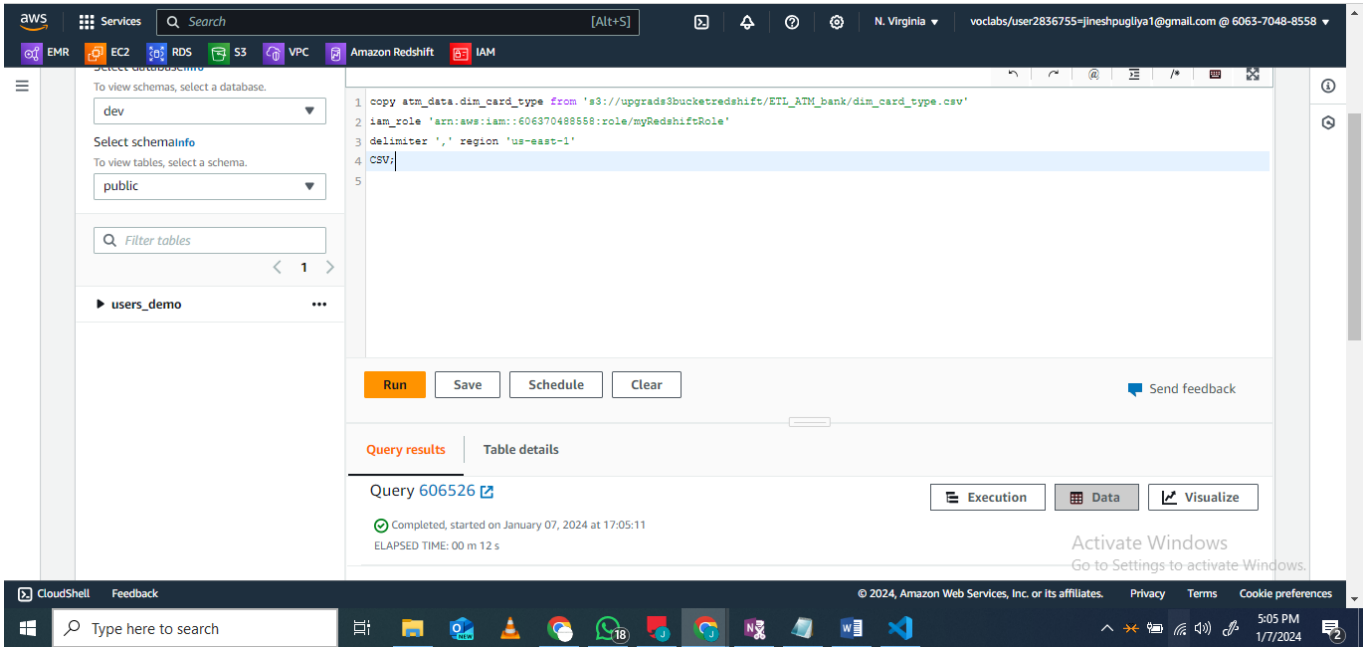


• Copying the data to dim_date table
copy atm_data.dim_date from 's3://upgrads3bucketredshift/ETL_ATM_bank/dim_date.csv'
iam_role 'arn:aws:iam::606370488558:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
CSV
TIMEFORMAT 'auto';



• Copying the data to dim_card_type table

copy atm_data.dim_card_type from
's3://upgrads3bucketredshift/ETL_ATM_bank/dim_card_type.csv'
iam_role 'arn:aws:iam::606370488558:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
CSV;



• Copying the data to fact_atm_trans table

copy atm_data.fact_atm_trans from
's3://upgrads3bucketredshift/ETL_ATM_bank/fact_atm_trans.csv'
iam_role 'arn:aws:iam::606370488558:role/myRedshiftRole'
delimiter ',' region 'us-east-1'
CSV;