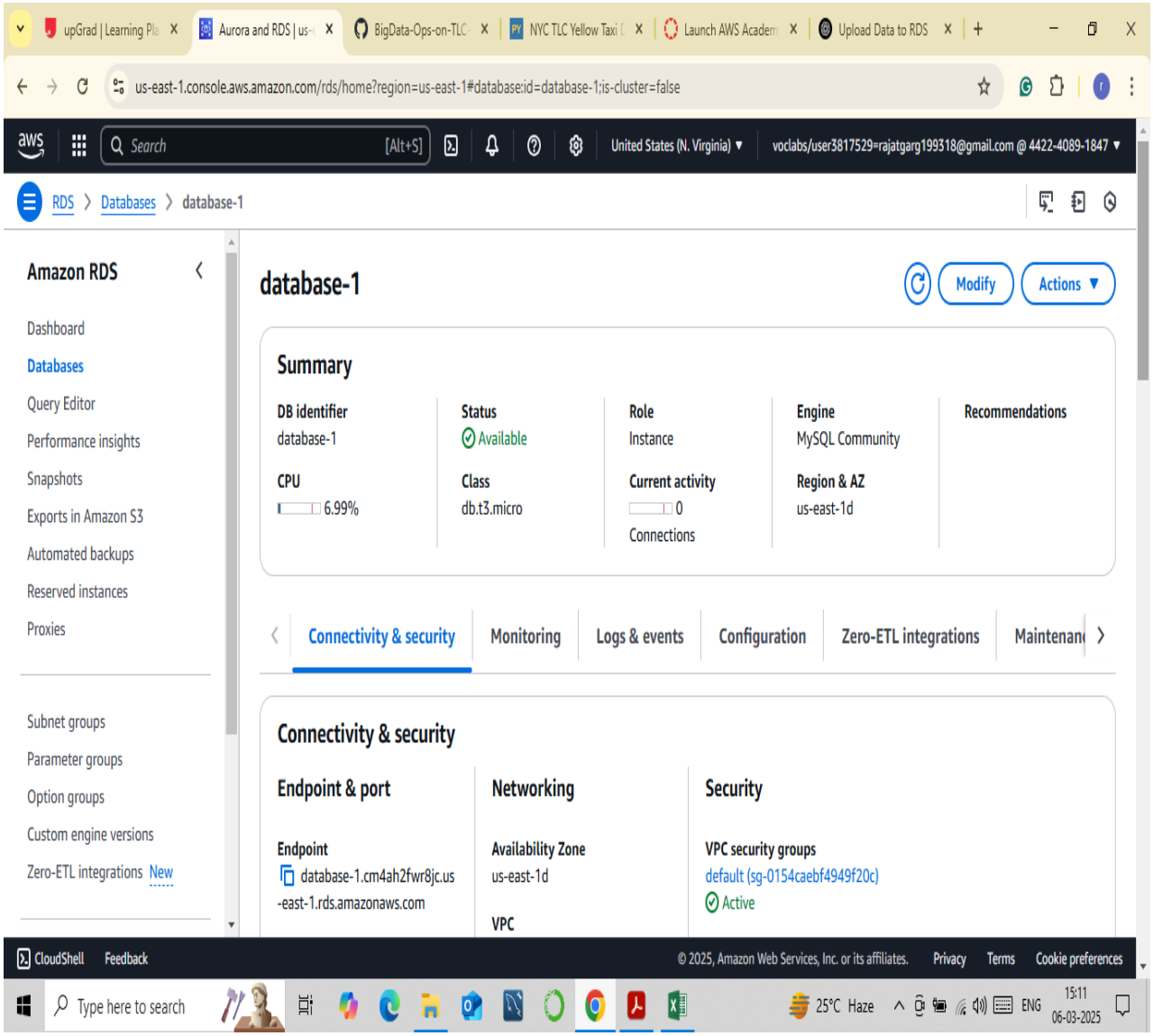
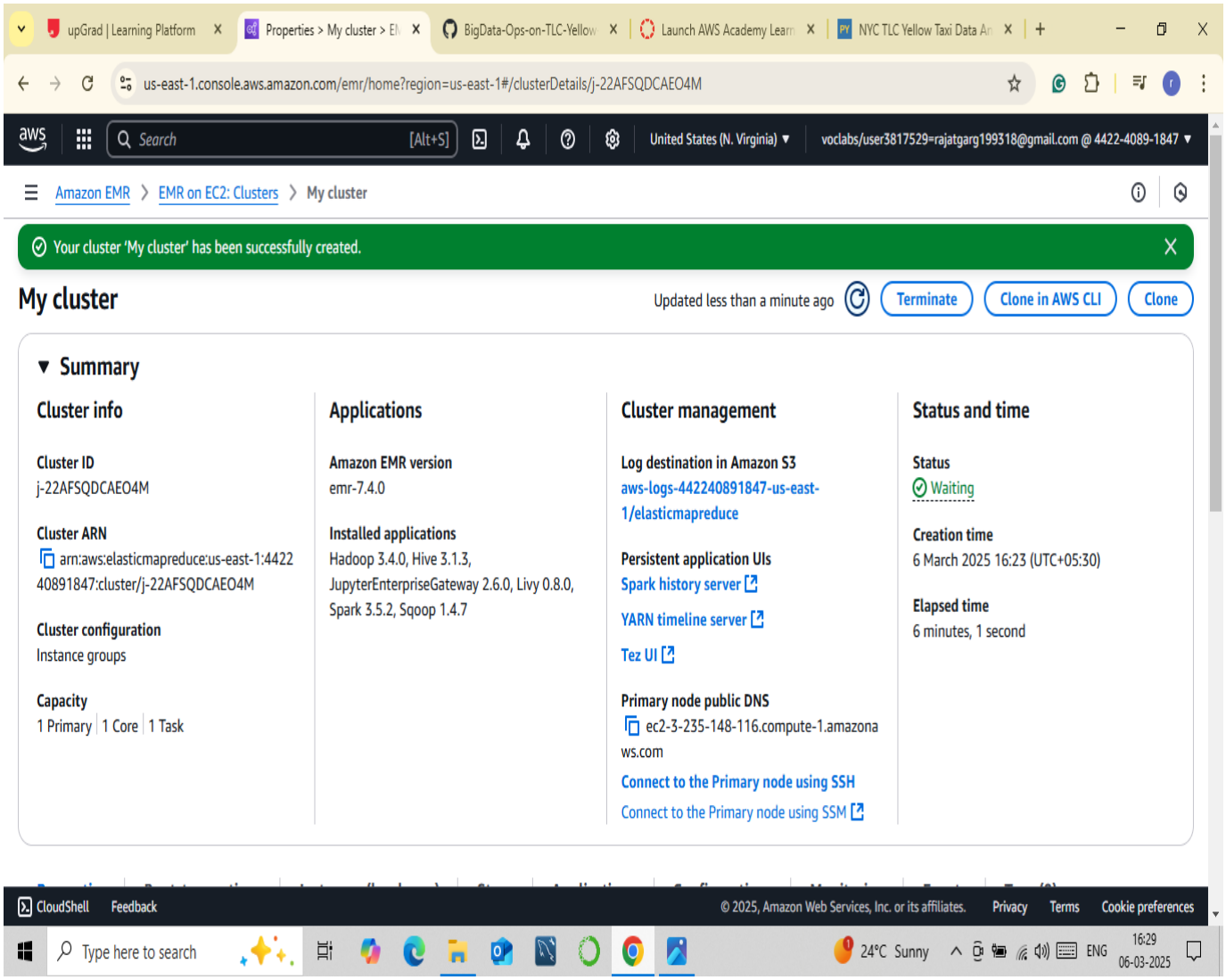
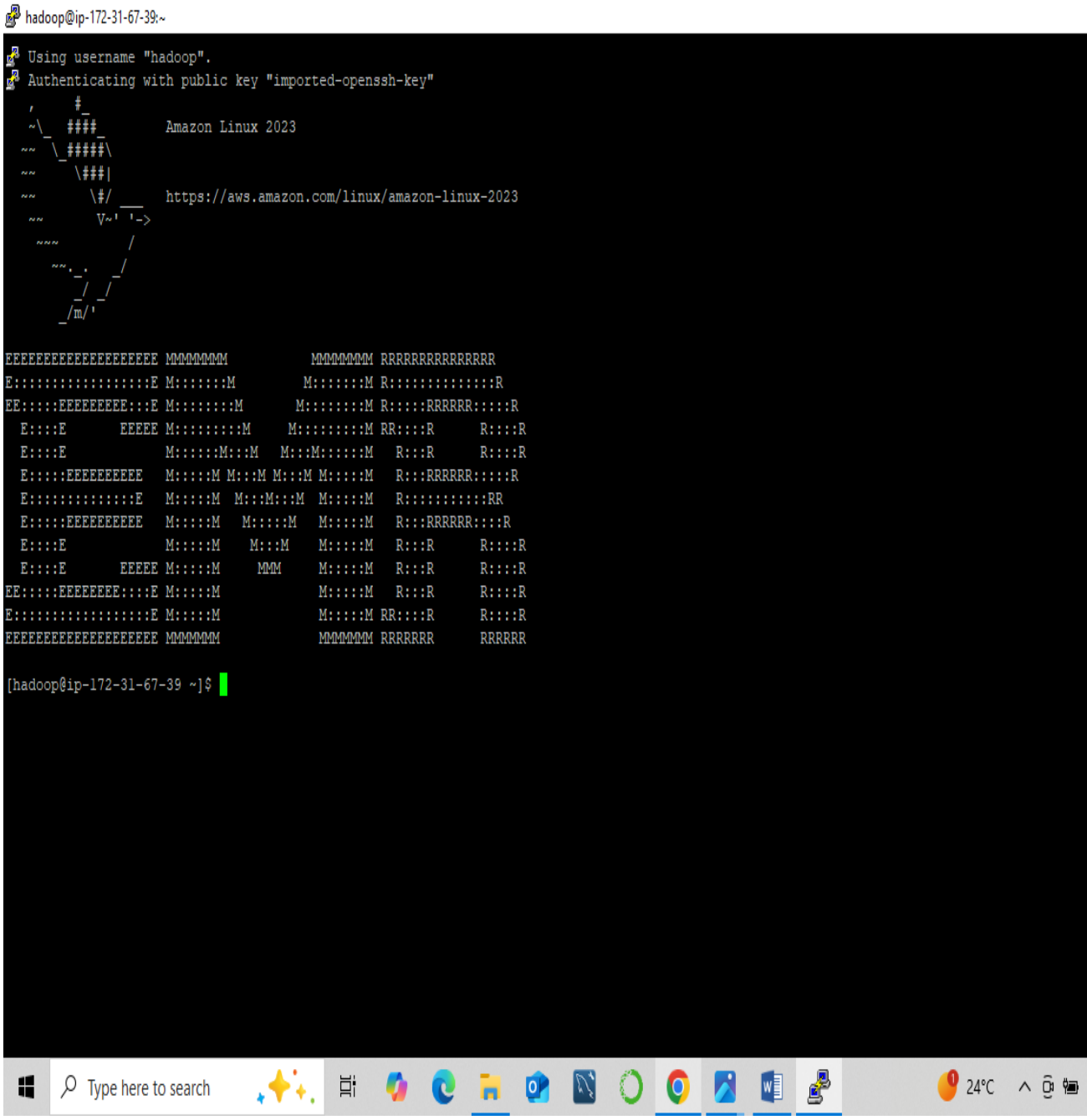
1. **Creation of RDS Instance In AWS**

2. 

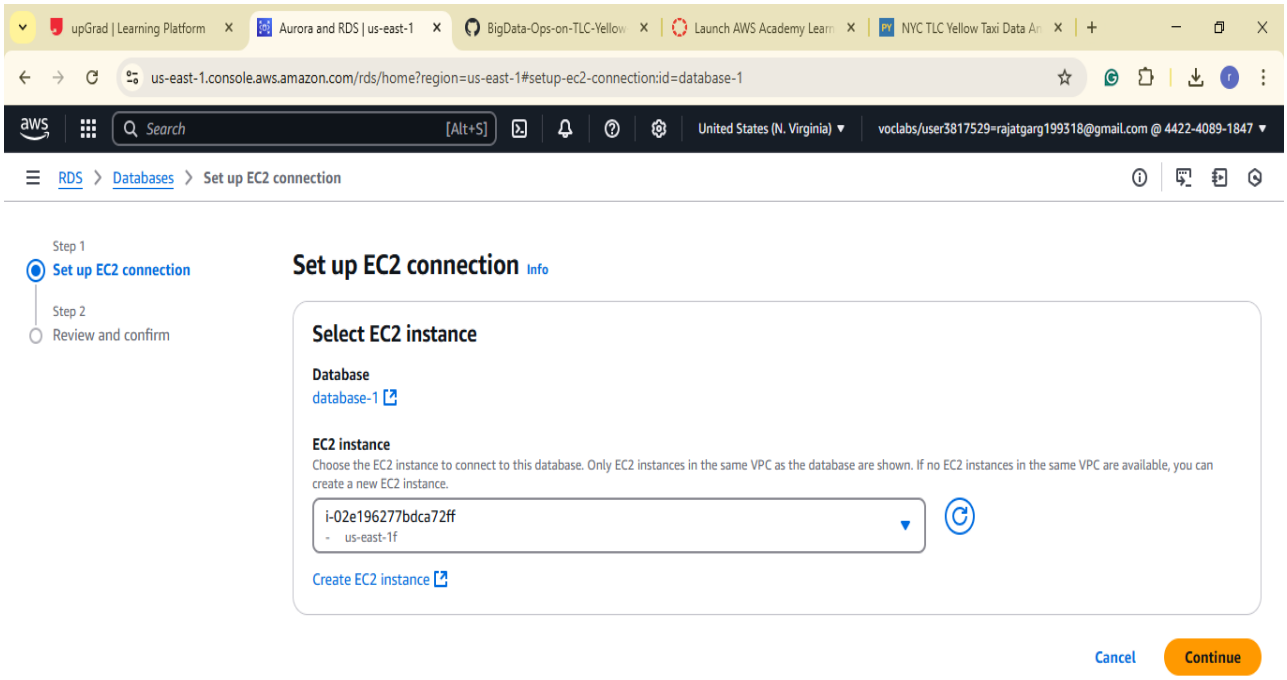
3. **Creation Of EMR Instance with bundled applications such as Hadoop,hbase ,sqoop**



4. **EMR Instance is running through putty**

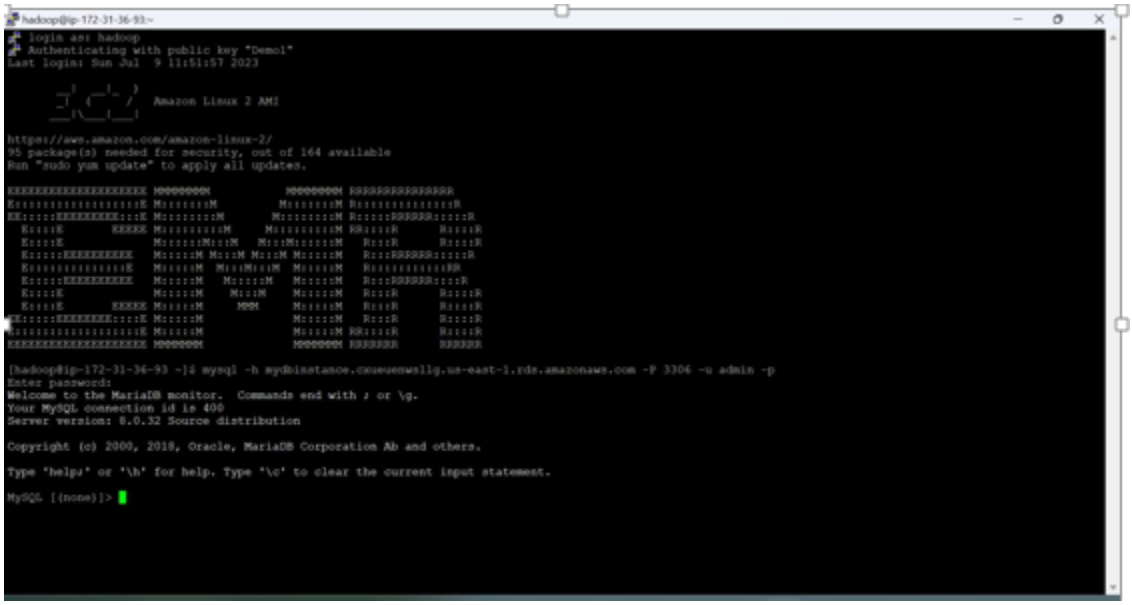
5. 

6. **To connect RDS with EMR instance, we have to click on "Action" button on RDS instance menu and then "Set up an EC2 Instance".**



7. 

8. **Login to RDS through EMR instance using command: `mysql –h**



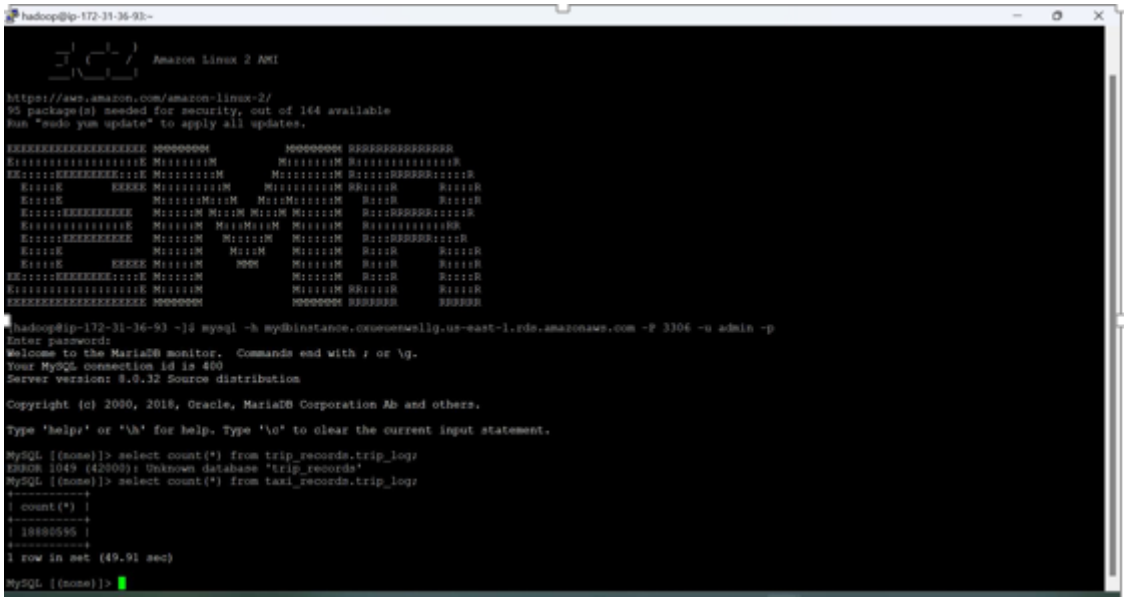
9. **Creation of Database “taxi_records” and table “trip_log”**
   `create database taxi_records’
   ‘CREATE TABLE trip_log
   (VendorID INT, tpep_pickup_datetime VARCHAR(50), tpep_dropoff_datetime

VARCHAR(50), Passenger_count INT, Trip_distance FLOAT, RatecodeID INT, store_and_fwd_flag VARCHAR(2), PULocationID INT, DOLocationID INT, payment_type INT, fare_amount FLOAT, extra FLOAT, mta_tax FLOAT, tip_amount FLOAT, tolls_amount FLOAT, improvement_surcharge FLOAT, total_amount FLOAT, Airport_fee FLOAT);'

1. **Downloading required csv files from internet in local using command**
   ` wget https://nyc-tlc-upgrad.s3.amazonaws.com/yellow_tripdata_2017-01.csv`
   ` wget https://nyc-tlc-upgrad.s3.amazonaws.com/yellow_tripdata_2017-02.csv`
2. **To load data in mysql table we have to login and then run sql command:**

LOAD DATA LOCAL INFILE '/home/hadoop/yellow_tripdata_2017-01.csv'
INTO TABLE trip_log FIELDS TERMINATED BY ',' LINES TERMINATED BY '\n'
IGNORE 1 LINES;

LOAD DATA LOCAL INFILE '/home/hadoop/yellow_tripdata_2017-02.csv'
INTO TABLE trip_log FIELDS TERMINATED BY ',' LINES TERMINATED BY '\n'
IGNORE 1 LINES;



SELECT COUNT(*) FROM taxi_records.trip_log;