

CODE :

1. Creating Table in RDS :

```
REATE TABLE yellow_trip(  
  VendorID INT  
  ,tpep_pickup_datetime DATETIME  
  ,tpep_dropoff_datetime DATETIME  
  ,passenger_count INT  
  ,trip_distance FLOAT  
  ,RatecodeID INT  
  ,store_and_fwd_flag VARCHAR(225)  
  ,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  
  ,congestion_surcharge FLOAT  
  ,Airport_fee FLOAT  
);
```

2. Loding Data to RDS Table :

```
LOAD DATA LOCAL INFILE '/home/ec2-user/yellow_tripdata_2017-01.csv'  
INTO TABLE yellow_trip  
FIELDS TERMINATED BY','  
LINES TERMINATED BY'\n'  
IGNORE 1 LINES;
```

SCREENSHOT :

✔ Your cluster "UPGRADE_EMR_2" has been successfully created.

✕

Amazon EMR > EMR on EC2: Clusters > UPGRADE_EMR_2

UPGRADE_EMR_2

Updated 10 minutes ago

Actions

▼ Summary

Cluster info

Cluster ID
j-2BMXFRK5V5L8T

Cluster configuration
Instance groups

Capacity
1 Primary | 0 Core | 0 Task

Applications

Amazon EMR version
emr-5.30.1

Installed applications
Hadoop 2.8.5, Hive 2.3.6, Hue 4.6.0,
Pig 0.17.0

Cluster management

Log destination in Amazon S3
[aws-logs-465161302776-us-east-1/elasticmapreduce](#)

Persistent application UIs
[YARN timeline server](#)

[Tez UI](#)

Primary node public DNS
 `ec2-54-227-124-63.compute-1.amazonaws.com`
[Connect to the Primary Node using SSH](#)

Status and time

Status
 [Waiting](#)

Creation time
September 09, 2023, 22:09 (UTC+05:30)

Elapsed time
20 minutes, 35 seconds

[illegible]

Creating RDS Instance

upgraderds1

Modify

Actions ▼

Summary

DB identifier upgraderds1	CPU <div><div></div></div> 3.30%	Status ✔ Available	Class db.t3.micro
Role Instance	Current activity <div><div></div></div> 0 Connections	Engine MySQL Community	Region & AZ us-east-1f

Endpoint & port

Endpoint
upgraderds1.cknzwrjzbilx.us-east-1.rds.amazonaws.com

Port
3306

Networking

Availability Zone
us-east-1f

VPC
[vpc-0f7d1dcfea0bcfeaf](#)

Subnet group
default-vpc-0f7d1dcfea0bcfeaf

Subnets
[subnet-0c6351658f5c1977a](#)
[subnet-088f5a2abf5e3bf5f](#)
[subnet-07f27c197e26587d0](#)
[subnet-078534716307ece59](#)
[subnet-086c6f5830a654f5d](#)
[subnet-0616da78a369f98ca](#)

Network type
IPv4

Security

VPC security groups
[default \(sg-089642e45f9bc3b99\)](#)
✔ Active

Publicly accessible
Yes

Certificate authority [Info](#)
rds-ca-2019

Certificate authority date
August 22, 2024, 22:38 (UTC+05:30)

DB instance certificate expiration date

⚠ [August 22, 2024, 22:38 \(UTC+05:30\)](#)

EMR connected to RDS

Select EC2 instance

Database

upgraderds1

EC2 instance

Choose the EC2 instance to connect to this database. Only EC2 instances in the same VPC as the database are shown. If no EC2 instances in the same VPC are available, you can create a new EC2 instance.

i-06574c85e96229f0b
RAJANECC1 us-east-1b



[Create EC2 instance](#)



Connection already exists

A connection between RDS database [upgraderds1](#) (security group [rds-ec2-1](#)) and EC2 instance [i-06574c85e96229f0b](#) (security group [ec2-rds-1](#)) already exists.

Cancel

Continue

Creating Table

```

MySQL [(none)]> show databases;
+-----+
| Database          |
+-----+
| information_schema |
| mysql              |
| performance_schema |
| sys                |
| upgrade            |
+-----+
5 rows in set (0.01 sec)

MySQL [(none)]> use upgrade;
Database changed
MySQL [upgrade]> CREATE TABLE yellow_trip(
  -> VendorID INT
  -> ,tpep_pickup_datetime DATETIME
  -> ,tpep_dropoff_datetime DATETIME
  -> ,passenger_count INT
  -> ,trip_distance FLOAT
  -> ,RatecodeID INT
  -> ,store_and_fwd_flag VARCHAR(225)
  -> ,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
  -> ,congestion_surcharge FLOAT
  -> ,Airport_fee FLOAT
  -> );
Query OK, 0 rows affected (0.03 sec)

```

Inserting data to RDS

```

mysql> use upgrade;
Reading table information for completion of table and column names
You can turn off this feature to get a quicker startup with -A

Database changed
mysql>
mysql>
mysql>
mysql> LOAD DATA LOCAL INFILE '/home/ec2-user/yellow_tripdata_2017-01.csv'
  -> INTO TABLE yellow_trip
  -> FIELDS TERMINATED BY','
  -> LINES TERMINATED BY'\n'
  -> IGNORE 1 LINES;
Query OK, 9710820 rows affected, 65535 warnings (2 min 28.32 sec)
Records: 9710820 Deleted: 0 Skipped: 0 Warnings: 19421640

mysql> LOAD DATA LOCAL INFILE '/home/ec2-user/yellow_tripdata_2017-02.csv'
  -> INTO TABLE yellow_trip
  -> FIELDS TERMINATED BY','
  -> LINES TERMINATED BY'\n'
  -> IGNORE 1 LINES;
Query OK, 9169775 rows affected, 65535 warnings (2 min 23.22 sec)
Records: 9169775 Deleted: 0 Skipped: 0 Warnings: 18339550

```