

## Creating an environment to simulate the third party setup with AWS RDS-

card\_member.csv and member\_score.csv are the two data sets that would be made available by the third party providers in real time.

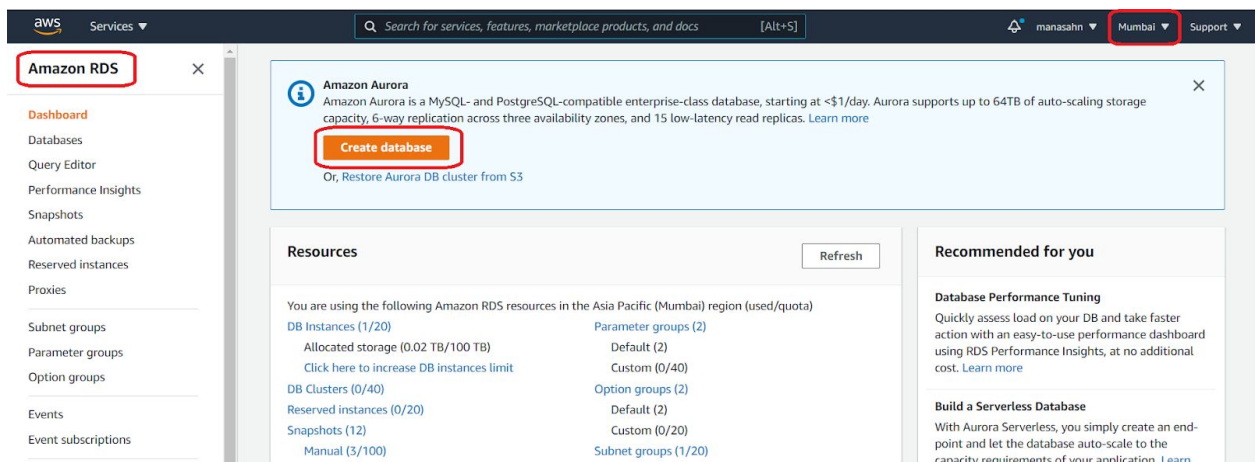
In order to mimic this setup for our project practice,

Download the datasets - **card\_member.csv** and **member\_score.csv** from the learning portal (courses.bigdatabysumit.com) , in the course “Real Time project on Bigdata” under the chapter **Data** and under the lesson **Resources**.

After downloading the datasets, copy the datasets to HDFS using -put/ -copyFromLocal command.

### Steps to create tables in the AWS RDS database and load the two datasets :

1. Sign in to the AWS Management Console and open the Amazon RDS console at <https://console.aws.amazon.com/rds/>
2. In the upper-right corner of the Amazon RDS console, choose the AWS Region in which you want to create the DB instance.
3. In the navigation pane, choose **Databases**.
4. Choose **Create database** and make sure that **Easy Create** is chosen.
5. In **Configuration**, choose **MySQL**.



## Create database


**Choose a database creation method** [Info](#)


☐ **Standard create**  
You set all of the configuration options, including ones for availability, security, backups, and maintenance.


☒ **Easy create**  
Use recommended best-practice configurations. Some configuration options can be changed after the database is created.


**Configuration**


**Engine type** [Info](#)


☐ Amazon Aurora  


☒ **MySQL**  


☐ MariaDB  


☐ PostgreSQL  


☐ Oracle  


☐ Microsoft SQL Server  


6. For **DB instance size**, choose **Free tier**.
7. For **DB instance identifier**, enter a name for the DB instance, or leave the default name.
8. To use an automatically generated master password for the DB instance, enable **Auto generate a password**.
9. To enter your master password, disable **Auto generate a password**, and then enter the same password in **Master password** and **Confirm password**.
10. Choose **Create database**.

If you chose to use an automatically generated password, the **View credential details** button appears on the **Databases** page. To view the master username and password for the DB instance, choose **View credential details**. You can use the username and password that appears to connect to the DB instance.

**DB instance size**

☐ **Production**  
db.r6g.xlarge  
4 vCPUs  
32 GiB RAM  
500 GiB  
1.146 USD/hour

☐ **Dev/Test**  
db.r6g.large  
2 vCPUs  
16 GiB RAM  
100 GiB  
0.260 USD/hour

☒ **Free tier**  
db.t2.micro  
1 vCPUs  
1 GiB RAM  
20 GiB  
0.028 USD/hour

**DB instance identifier**  
Type a name for your DB instance. The name must be unique across all DB instances owned by your AWS account in the current AWS Region.

**Project-database**

The DB instance identifier is case-insensitive, but is stored as all lowercase (as in 'mydbinstance'). Constraints: 1 to 60 alphanumeric characters or hyphens (1 to 15 for SQL Server). First character must be a letter. Can't contain two consecutive hyphens. Can't end with a hyphen.

**Master username** [Info](#)  
Type a login ID for the master user of your DB instance.

**admin**

1 to 16 alphanumeric characters. First character must be a letter

☐ **Auto generate a password**  
Amazon RDS can generate a password for you, or you can specify your own password

**Master password** [Info](#)

Database instance being created.

The screenshot shows the AWS Management Console interface. At the top, a blue banner indicates 'Creating database project-database' with a progress bar and the message 'Your database might take a few minutes to launch.' Below this, the 'Databases' table is visible, listing two database instances: 'database-2' (Status: Stopped) and 'project-database' (Status: Creating). The 'project-database' row is highlighted with a red box. The table columns include DB identifier, Role, Engine, Region & AZ, Size, Status, and CPU.

DB identifier	Role	Engine	Region & AZ	Size	Status	CPU
database-2	Instance	MySQL Community	ap-south-1b	db.t2.micro	Stopped	-
project-database	Instance	MySQL Community	-	db.t2.micro	Creating	-

Once the status of the database instance changes to **Available**, you can connect to the database instance using the following command from the MySQL terminal. On clicking the database instance created, and navigating to **Connectivity and Security** tab we can get the **Endpoint and Port** details. **Master Username and Password** are the once while creating the DB instance.

Now, connect to the RDS DB Instance using the following :

**mysql -h <Endpoint> -P 3306 -u <Master Username> -p**

Once connected,

1. **create the card\_member and member\_score tables**
2. **sqoop export** the data from HDFS to the tables.