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.

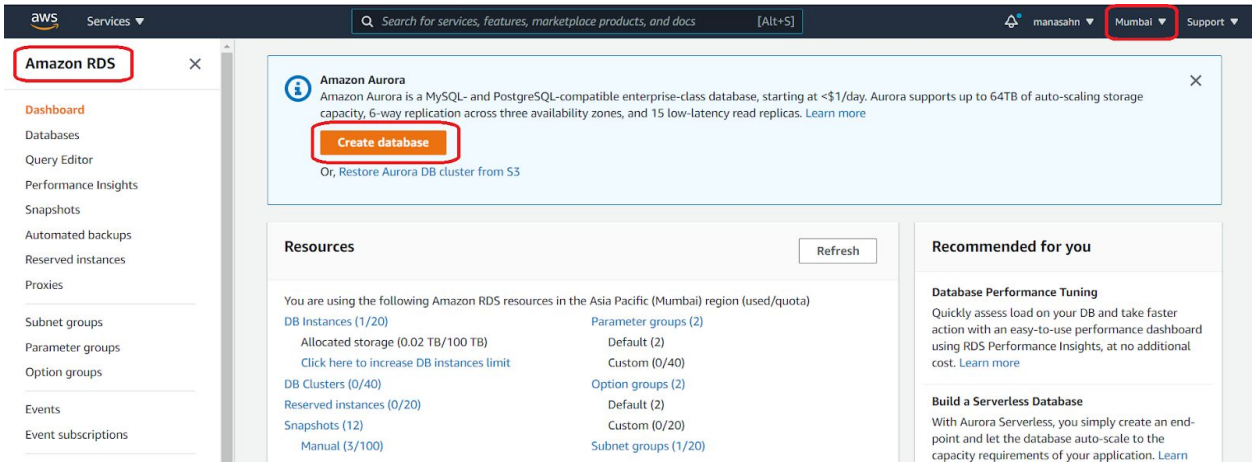
In order to mimic this setup for our project practice,

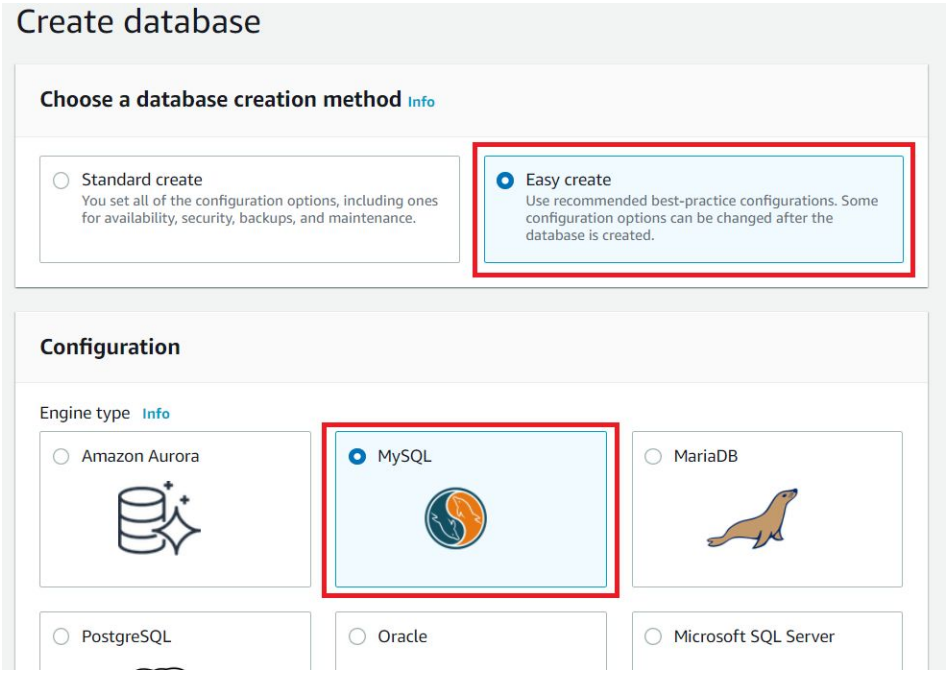
Download the datasets - **card_member.csv** and **member_score.csv** from the learning portal (courses.bigdatabysumit.com) , in the course “Real Time project on Bigdata” under the chapter **Data** and under the lesson **Resources.**

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 :

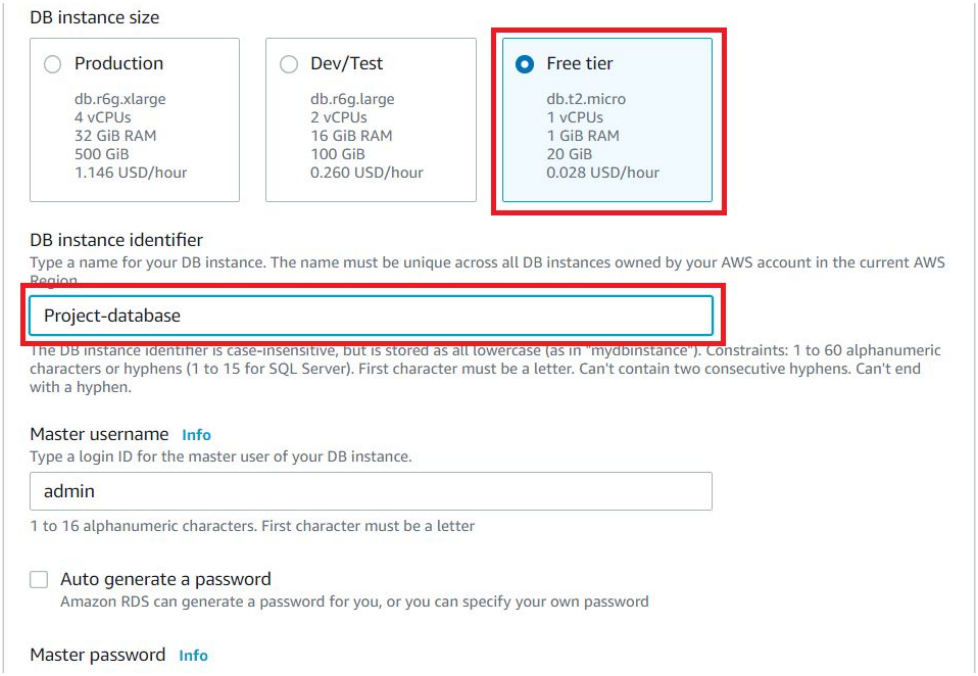
1. Sign in to the AWS Management Console and open the Amazon RDS console at https://console.aws.amazon.com/rds/
2. In the upper-right corner of the Amazon RDS console, choose the AWS Region in which you want to create the DB instance.
3. In the navigation pane, choose **Databases**.
4. Choose **Create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



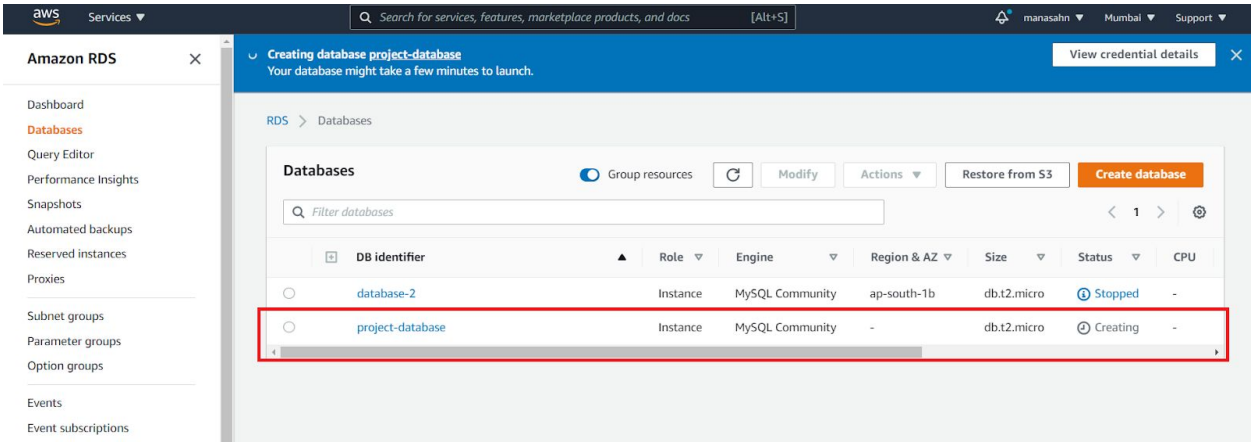


6. For **DB instance size**, choose **Free tier**.
7. For **DB instance identifier**, enter a name for the DB instance, or leave the default name.
8. To use an automatically generated master password for the DB instance, enable **Auto generate a password**.
9. To enter your master password, disable **Auto 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.



Database instance being created.



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.

Mater 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.