## Load data from AWS RDS to Hadoop

**<Command to import data from AWS RDS to Hadoop>**

**1. First we need to setup MySQL Connector on AWS EMR**

**a. Run the following command to install the MySQL connector jar file:**

**wget https://de-mysql-connector.s3.amazonaws.com/mysql-connector-java-8.0.25.tar.gz**

**b. Run the following step to extract the MySQL connector tar file**

**tar -xvf mysql-connector-java-8.0.25.tar.gz**

**c. go to the MySQL Connector directory and then copy it to the Sqoop library to complete the installation.**

**cd mysql-connector-java-8.0.25/**

**sudo cp mysql-connector-java-8.0.25.jar /usr/lib/sqoop/lib/**

**2. Now we run the Sqoop import command to import data from AWS RDS to Hadoop**

sqoop import \

--connect jdbc:mysql://upgraddetest.cyaielc9bmnf.us-east-1.rds.amazonaws.com/testdatabase \ --table bookings \

--username student \

--password STUDENT123 \

--target-dir /user/root/project/bookings \

-m1 \

**<Command to view the imported data>**

hadoop fs –ls /user/root/project/bookings

hadoop fs -cat /user/root/project/bookings/part-m-00000 | wc -l

**<Screenshot of the data>**





