

Load data from AWS RDS to Hadoop

Prior to running the sqoop command we need to install mysql connectors so that we can pull the data from RDS to Hadoop

The code:

```
wget https://de-mysql-connector.s3.amazonaws.com/mysql-connector-java-8.0.25.tar.gz
tar -xvf mysql-connector-java-8.0.25.tar.gz
cd mysql-connector-java-8.0.25/
sudo cp mysql-connector-java-8.0.25.jar /usr/lib/sqoop/lib/
```

<Command to import data from AWS RDS to Hadoop>

```
sqoop import \

--connect jdbc:mysql://upgraddetest.cyaieic9bmnf.us-east-1.rds.amazonaws.com/testdatabase \

--table bookings \

--username student --password STUDENT123 \

--target-dir /user/root/bookings_1 \

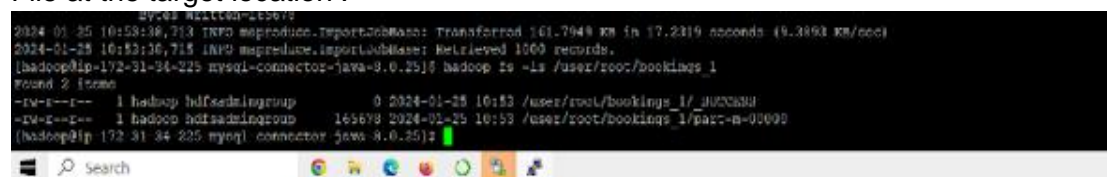
--m 1
```

<Command to view the imported data>

```
Hadoop fs -ls /user/root/bookings_1
```

<Screenshot of the data>

File at the target location :



```
2024-01-25 10:53:38,713 INFO mapreduce.jobreporter: Transferred 161.7949 KB in 17.2319 seconds (9.3893 KB/sec)
2024-01-25 10:53:38,715 INFO mapreduce.jobreporter: Retrieved 1000 records.
[hadoop@ip-172-31-34-225 mysql-connector-java-8.0.25]$ hadoop fs -ls /user/root/bookings_1
Found 2 items
-rw-r--r-- 1 hadoop hdfsadmin group 0 2024-01-25 10:53 /user/root/bookings_1/_SUCCESS
-rw-r--r-- 1 hadoop hdfsadmin group 165678 2024-01-25 10:53 /user/root/bookings_1/part-m-00000
[hadoop@ip-172-31-34-225 mysql-connector-java-8.0.25]$
```

Data inside the part file :

 hadoop@ip-172-31-34-223:~/mysql-connector-java-8.0.23[illegible]