Data Ingestion from the RDS to HDFS using Sqoop

Sqoop Import command used for importing table from RDS to HDFS:

```
sqoop import \
--connect jdbc:mysql://upgraddetest.cyaielc9bmnf.us-east-1.rds.amazonaws.com/
testdatabase  \
--table SRC_ATM_TRANS \
--username student --password STUDENT123 \
--target-dir /user/root/spar_nord_bank_atm \
-m 1
```


Total of **2468572** rows have been retrieved.

Command used to see the list of imported data in HDFS:

hadoop fs -ls /user/root/spar_nord_bank_atm

```
[hadoop@ip-172-31-26-34 ~]$ hadoop fs -ls /user/root/spar_nord_bank_atm
Found 2 items
-rw-r--r--   1 hadoop hadoop          0 2022-12-19 10:37 /user/root/spar_nord_bank_atm/_SUCCESS
-rw-r--r--   1 hadoop hadoop  531214815 2022-12-19 10:37 /user/root/spar_nord_bank_atm/part-m-00000
[hadoop@ip-172-31-26-34 ~]$
```

There are two items in the screenshot :

- The first file is the success file, indicating that the MapReduce job was successful.
- The second file 'part-m-00000' is the one that I imported.
 Since we used only one mapper in my import command that is why the data is in a single file.

hadoop fs -cat /user/root/spar_nord_bank_atm/part-m-00000