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 \
-m1
```

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

```
hadoop fs -ls /user/root/spar_nord_bank_atm
```

Screenshot of the imported data:

```
[hadoop@ip-172-31-33-249 ~]$ hadoop fs -ls /user/root/spar_nord_bank_atm
Found 2 items
-rw-r--r--   1 hadoop hadoop          0 2023-07-29 19:43 /user/root/spar_nord_ba
nk_atm/_SUCCESS
-rw-r--r--   1 hadoop hadoop  531214815 2023-07-29 19:43 /user/root/spar_nord_ba
nk_atm/part-m-00000
[hadoop@ip-172-31-33-249 ~]$ ^C
[hadoop@ip-172-31-33-249 ~]$
```

As can be seen in the screenshot, there are two files:-
1. The success file
2. The imported file ‘part-m-0000’. There is only one file since only one mapper was used.

A portion of the read data: