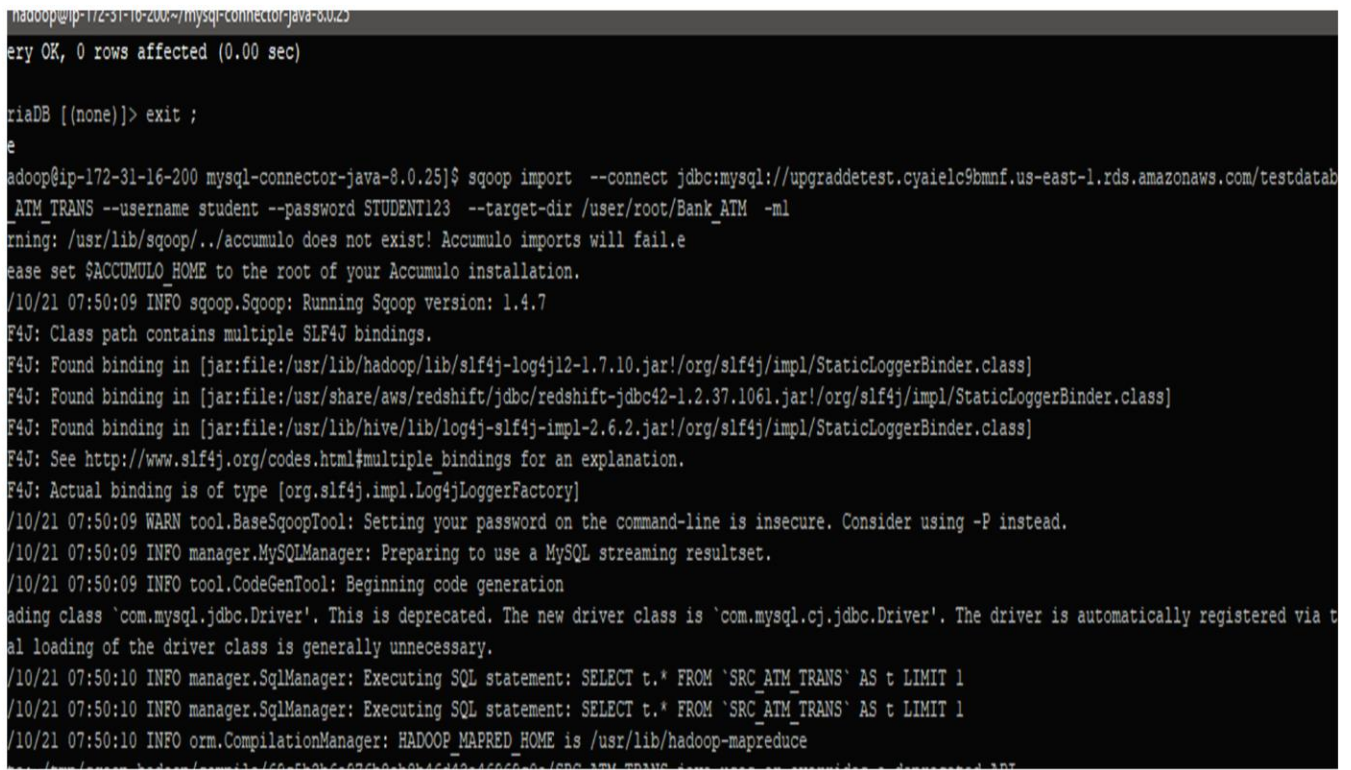# Data Ingestion from the RDS to HDFS using Sqoop

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/Bank_ATM \
 --m 1

This screenshot is command of data ingestion RDS to HDFS from Sqoop .

```
hadoop@ip-172-31-16-200:~/mysql-connector-java-8.0.25                                          –    □
22/10/21 07:51:07 INFO mapreduce.Job: Counters: 30
        File System Counters
                FILE: Number of bytes read=0
                FILE: Number of bytes written=189904
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=87
                HDFS: Number of bytes written=531214815
                HDFS: Number of read operations=4
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=2
        Job Counters
                Launched map tasks=1
                Other local map tasks=1
                Total time spent by all maps in occupied slots (ms)=1187088
                Total time spent by all reduces in occupied slots (ms)=0
                Total time spent by all map tasks (ms)=24731
                Total vcore-milliseconds taken by all map tasks=24731
                Total megabyte-milliseconds taken by all map tasks=37986816
        Map-Reduce Framework
                Map input records=2468572
                Map output records=2468572
                Input split bytes=87
                Spilled Records=0
                Failed Shuffles=0
                Merged Map outputs=0
                GC time elapsed (ms)=305
                CPU time spent (ms)=25580
                Physical memory (bytes) snapshot=629608448
                Virtual memory (bytes) snapshot=3302236160
                Total committed heap usage (bytes)=537395200
        File Input Format Counters
                Bytes Read=0
        File Output Format Counters
                Bytes Written=531214815
22/10/21 07:51:07 INFO mapreduce.ImportJobBase: Transferred 506.6059 MB in 51.3347 seconds (9.8687 MB/sec)
22/10/21 07:51:07 INFO mapreduce.ImportJobBase: Retrieved 2468572 records.
```

I can see the result of sqoop import job 2468572 records have been retrived (same as the checkpoint mentioned in the Validation document):

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

Command used
Hadoop fs -ls /user/root/Bank_ATM

In the screenshot above we can see two items: - 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 I used only one mapper in my import command, thus the data is in a single file.

+Screenshot of the imported data:

Hadoop fs -cat /user/root/Bank_ATM/part-m-00000

<Screenshot of the imported data>