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_Data

\-m 1

Comment: Data has been taken from SRC_ATM_TRANS folder and uploaded to Bank_ATM_Data folder in HDFS under root directory. Total of 2468572 records have been transferred.

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

hadoop fs -ls /user/root/Bank_ATM_Data

Comment: Data transferred is verified within user/root/BANK_ATM_Data Directory.

Screenshot of the imported data:

```
            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)=27426
            Total time spent by all reduces in occupied slots (ms)=0
            Total time spent by all map tasks (ms)=27426
            Total vcore-milliseconds taken by all map tasks=27426
            Total megabyte-milliseconds taken by all map tasks=28084224
        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)=209
            CPU time spent (ms)=28960
            Physical memory (bytes) snapshot=423919616
            Virtual memory (bytes) snapshot=2799554560
            Total committed heap usage (bytes)=385875968
        File Input Format Counters
            Bytes Read=0
        File Output Format Counters
            Bytes Written=531214815
1/07/15 01:35:01 INFO mapreduce.ImportJobBase: Transferred 506.6059 MB in 49.676 seconds (10.1982 MB/sec)
1/07/15 01:35:01 INFO mapreduce.ImportJobBase: Retrieved 2468572 records.
root@ip-10-0-0-17 ~]#
```

```
                    Bytes Written=531214815
21/07/15 01:35:01 INFO mapreduce.ImportJobBase: Transferred 506.6059 MB in 49.676 seconds (10.1982 MB/sec)
21/07/15 01:35:01 INFO mapreduce.ImportJobBase: Retrieved 2468572 records.
[root@ip-10-0-0-17 ~]# hadoop fs -ls /user/root/
Found 3 items
drwx------   - root supergroup          0 2021-07-15 01:32 /user/root/.Trash
drwx------   - root supergroup          0 2021-07-15 01:35 /user/root/.staging
drwxr-xr-x   - root supergroup          0 2021-07-15 01:34 /user/root/Bank_ATM_Data
[root@ip-10-0-0-17 ~]# hadoop fs -ls /user/root/Bank_ATM_Data
Found 2 items
-rw-r--r--   3 root supergroup          0 2021-07-15 01:34 /user/root/Bank_ATM_Data/_SUCCESS
-rw-r--r--   3 root supergroup  531214815 2021-07-15 01:34 /user/root/Bank_ATM_Data/part-m-00000
[root@ip-10-0-0-17 ~]#
```