

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/atm_data \  
-m 1
```

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

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

Screenshot of the imported data:

```
22/07/30 10:59:07 INFO mapreduce.ImportJobBase: Transferred 506.6059 MB in 44.5311 seconds (11.3764 MB/sec)  
22/07/30 10:59:07 INFO mapreduce.ImportJobBase: Retrieved 2468572 records.  
[root@ip-172-31-65-124 ~]# hadoop fs -ls /user/root/atm_data/  
Found 2 items  
-rw-r--r--  1 root hdfsadmin 0 2022-07-30 10:59 /user/root/atm_data/_SUCCESS  
-rw-r--r--  1 root hdfsadmin 531214815 2022-07-30 10:59 /user/root/atm_data/part-m-00000  
[root@ip-172-31-65-124 ~]#
```