

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://upgradtest.cyaieic9bmnf.us-east-1.rds.amazonaws.com/testdatabase \  
--table SRC_ATM_TRANS \  
--username student --password STUDENT123 \  
--target-dir /user/root/etl_assignment \  
-m 1 \  
--direct
```

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

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

Screenshot of the imported data:

```
05 INFO mapreduce.ImportJobBase: Transferred 506.6059 MB in 53.1321 seconds (9.5348 MB/sec)  
05 INFO mapreduce.ImportJobBase: Retrieved 2468572 records.  
-rw-r--r--  1 hadoop hadoop          0 2022-02-06 12:35 /user/root/etl_assignment/_SUCCESS  
-rw-r--r--  1 hadoop hadoop 531214815 2022-02-06 12:35 /user/root/etl_assignment/part-m-00000
```