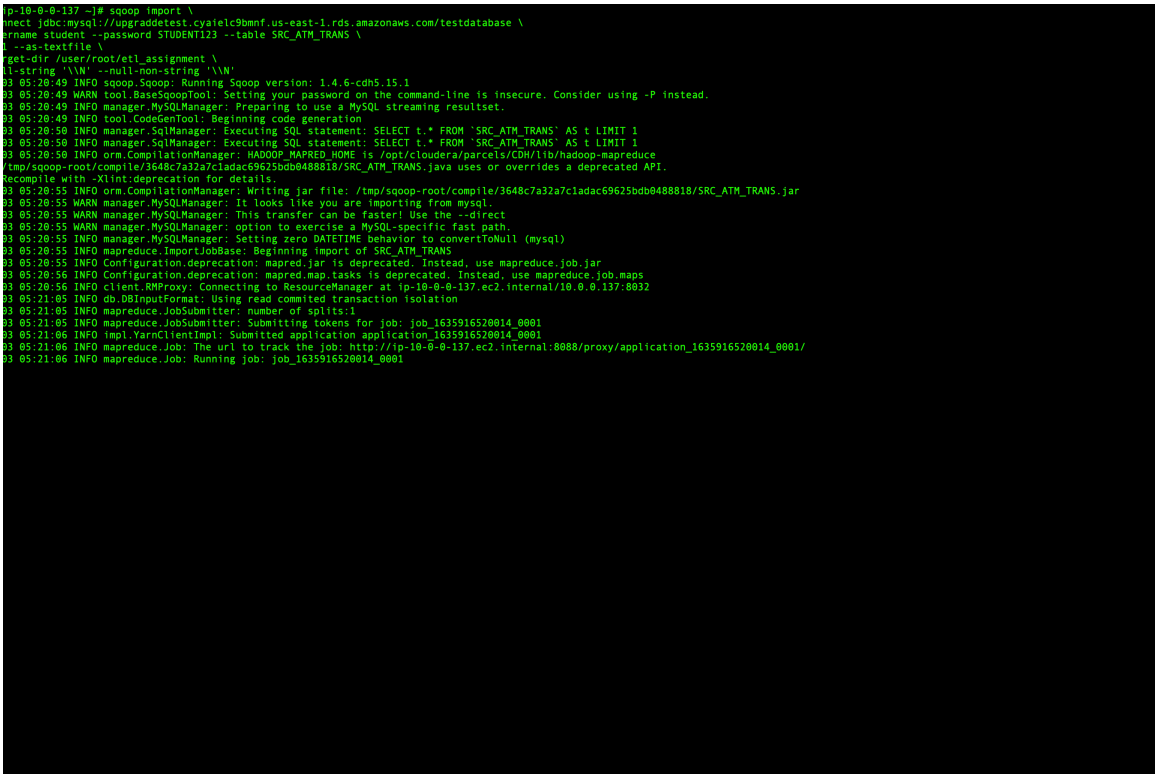
Data Ingestion from the RDS to HDFS using Sqoop

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 \
--username student --password STUDENT123 --table SRC_ATM_TRANS \
--m 1 --as-textfile \
--target-dir /user/root/etl_assignment \
--null-string '\\N' --null-non-string '\\N'
```

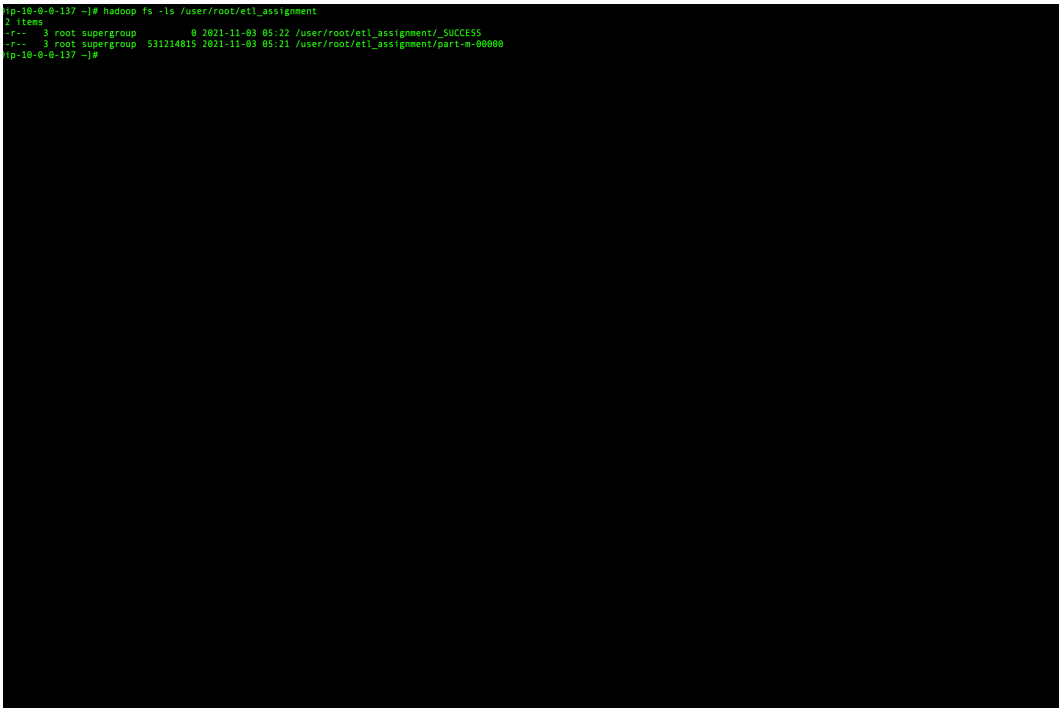


Running the Sqoop Command

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

hadoop fs -ls /user/root/etl_assignment

Screenshot of the imported data:



Data imported to HDFS post Sqoop command

Steps Taken:

1. **Login to the ec2 instance , run the sudo command**
2. **Execute the Sqoop command to import the data in the RDS table SRC_ATM_TRANS to the target directory user/root/etl_assignment**
3. **Once the job completes navigate to the target directory set to find the imported data using the “hadoop fs -ls user/root/etl_assignment” command**
4. **Since we see the _SUCCESS it means the import has gone through successfully**