### 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\_trans \

--null-string '\\N' --null-non-string '\\N' \

-m 1

Explanation:

--connect describes the connection url

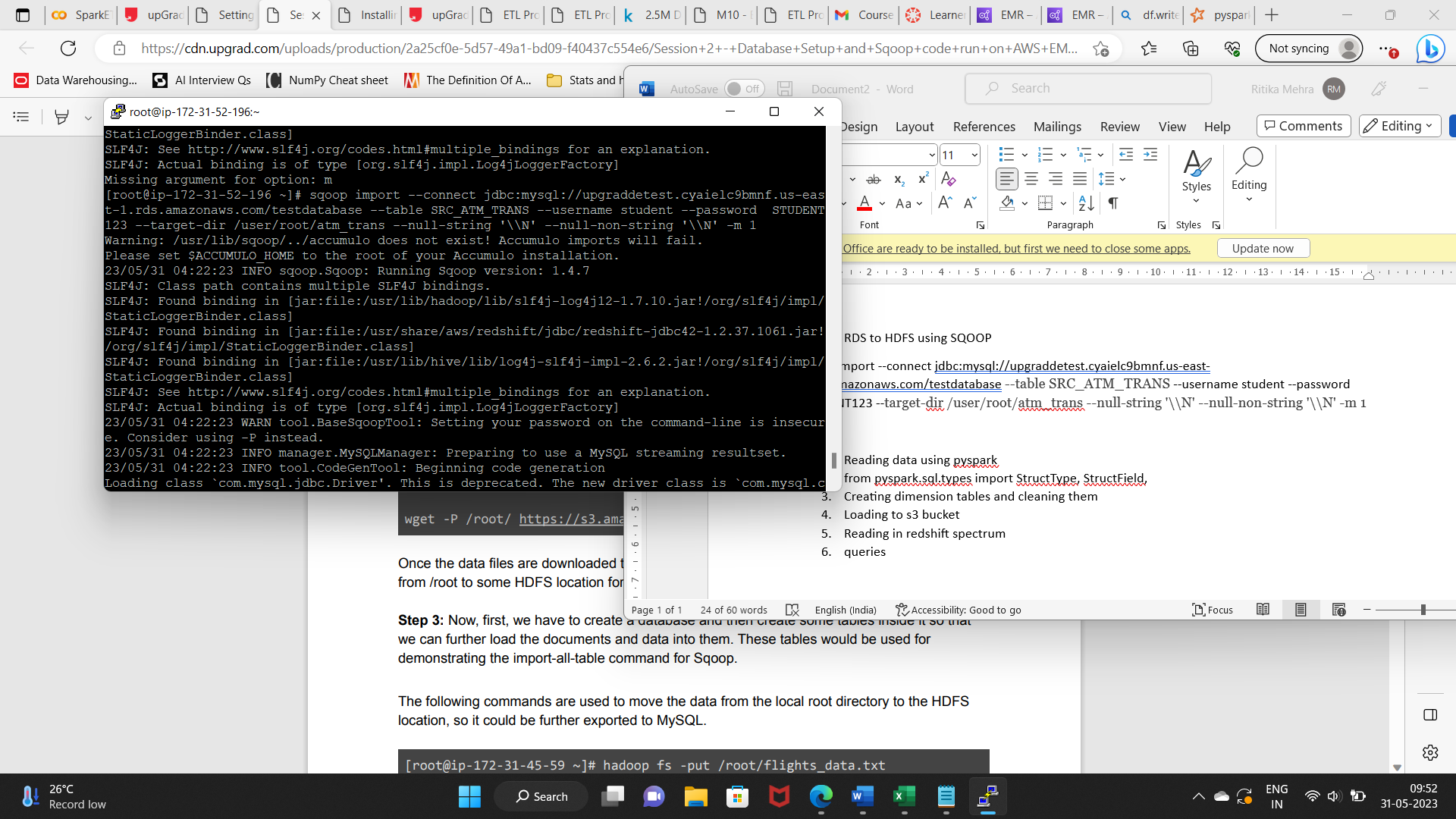
--table is the source table name

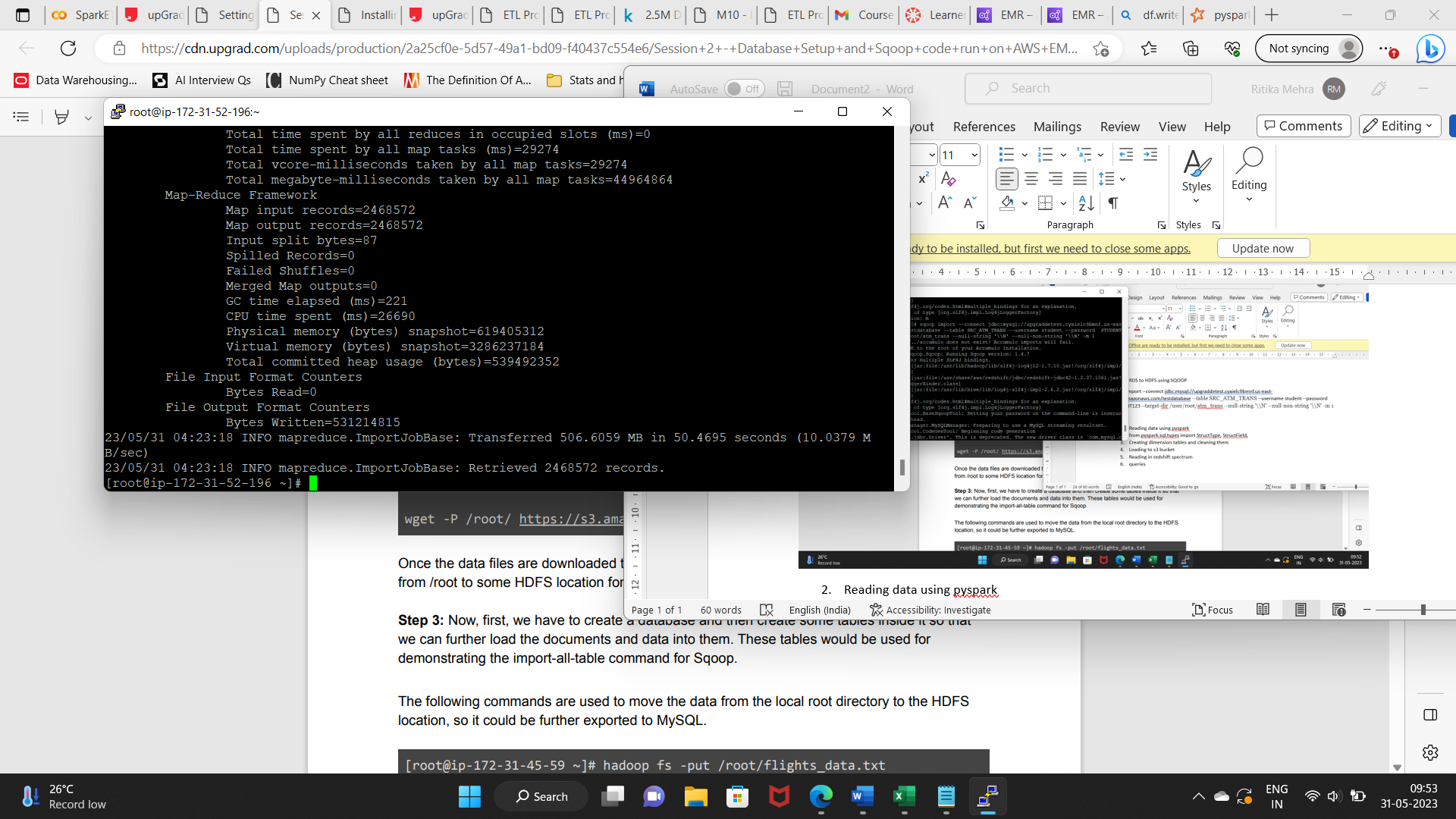
--username and –password are the credentials needed to access the database

--target-dir is the name of the directory to be created to load the data

--null-string and –null-non-string is used to handle null values properly

-m 1 determines that 1 mapper will be used



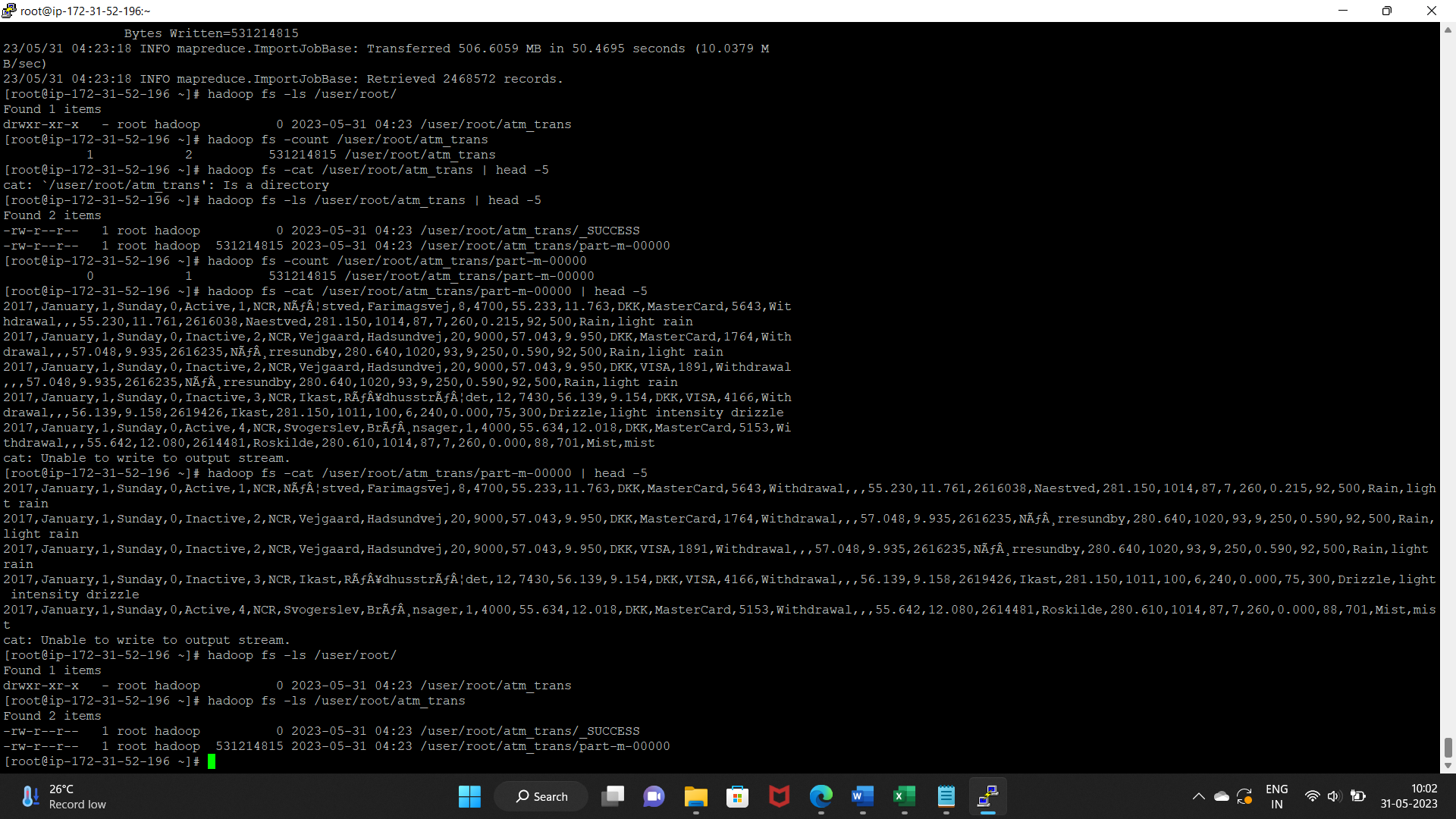


Records loaded: 2468572

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

**Commands used:** hadoop fs -ls /user/root/

hadoop fs -ls /user/root/atm\_trans



**Screenshot of the imported data:**

**Command used:** hadoop fs -cat /user/root/atm\_trans/part-m-00000 | head -5

