## Data Ingestion from the RDS to HDFS using Sqoop

Below are the steps performed to ingest the data from RDS table into Hive table using the sqoop command:

1. Log in into EMR instance and complete the initial setup by running the below commands. This is to download the MySQL connector, which will be used by sqoop to set up jdbc connection with table in MySQL RDS:

wget https://de-mysql-connector.s3.amazonaws.com/mysql-connector-java-8.0.25.tar.gz

tar -xvf mysql-connector-java-8.0.25.tar.gz

cd mysql-connector-java-8.0.25/

sudo cp mysql-connector-java-8.0.25.jar /usr/lib/sqoop/lib/

**Sqoop command used for importing table from RDS to HDFS**

1. **Sqoop command to import card\_member table from RDS to Hive**

sqoop import \

--connect jdbc:mysql://upgradawsrds1.cyaielc9bmnf.us-east-1.rds.amazonaws.com/cred\_financials\_data \

--table card\_member \

--username upgraduser --password upgraduser \

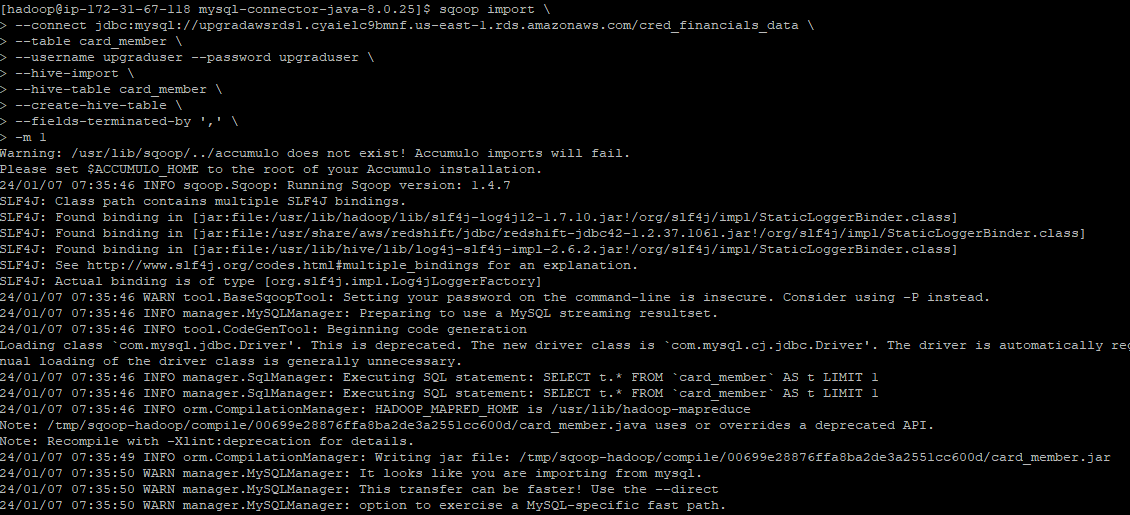
--hive-import \

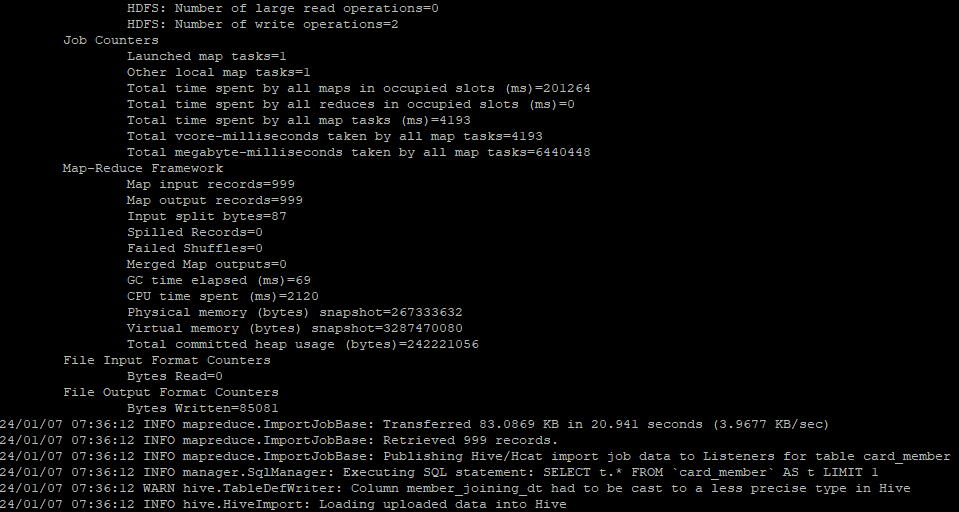
--hive-table card\_member \

--create-hive-table \

--fields-terminated-by ',' \

-m 1





1. **Sqoop command to import member\_score table from RDS to Hive**

sqoop import \

--connect jdbc:mysql://upgradawsrds1.cyaielc9bmnf.us-east-1.rds.amazonaws.com/cred\_financials\_data \

--table member\_score \

--username upgraduser --password upgraduser \

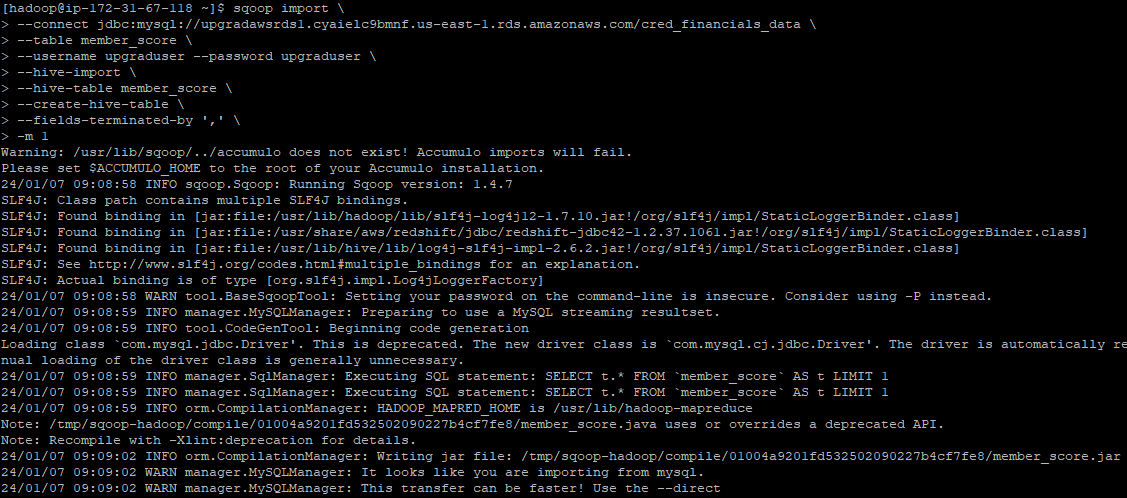
--hive-import \

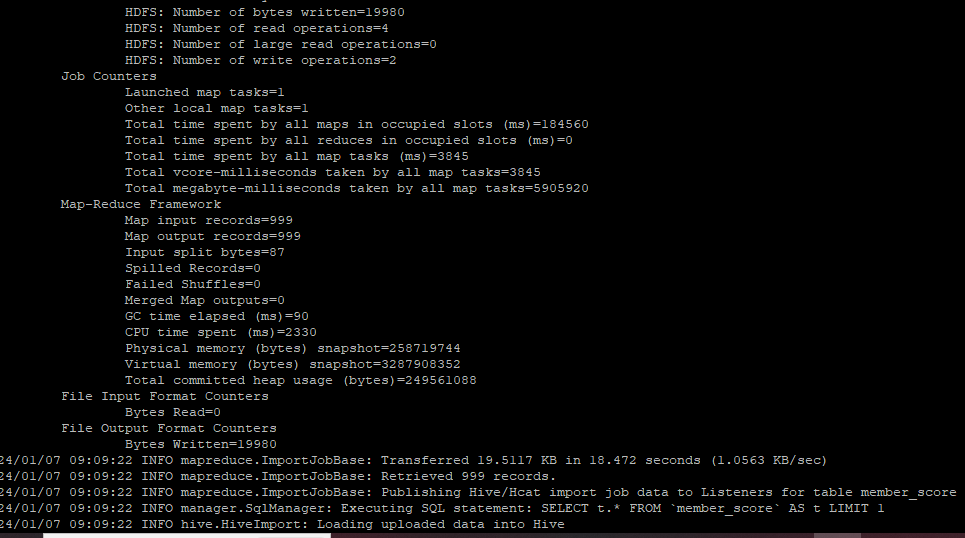
--hive-table member\_score \

--create-hive-table \

--fields-terminated-by ',' \

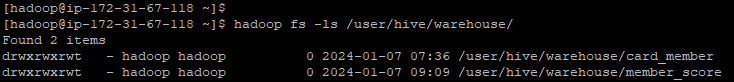
-m 1



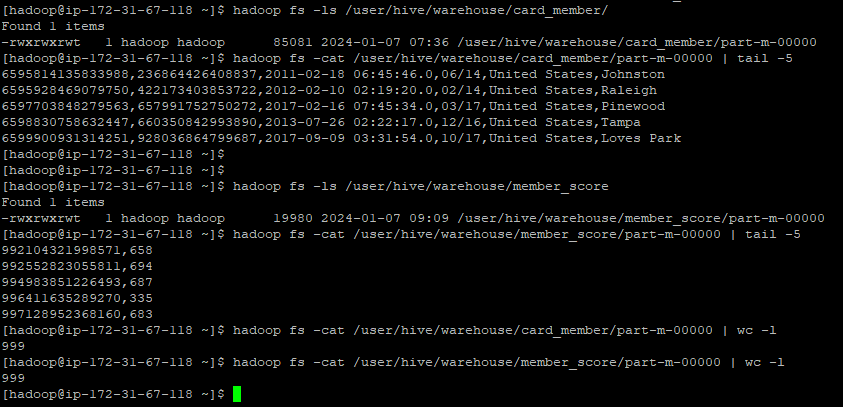


**<Command to see the list of imported data in HDFS>**

hadoop fs -ls /user/hive/warehouse/

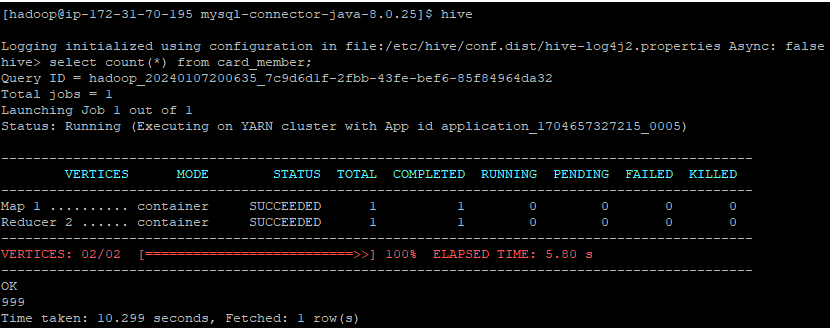


**Screenshot of the imported data**

****

**Screenshot of the imported data in Hive:**

1. select count(\*) from card\_member;



1. select count(\*) from member\_score;

