

## Data Ingestion from the RDS to HDFS using Sqoop

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

#### Card Member: -

```
sqoop import \  
--connect jdbc:mysql://${rds_connection}/${database} \  
--username ${username} \  
--password ${password} \  
--table card_member \  
--warehouse-dir /user/hadoop/cred_financials_data \  
--delete-target-dir \  
--num-mappers 1 \  
--compress
```

Where `${rds_connection}` = upgradawsrds1.cyaielec9bmnf.us-east-1.rds.amazonaws.com  
`${database}` = cred\_financials\_data  
`${username}` = upgraduser  
`${password}` = upgraduser

#### Member Score: -

```
sqoop import \  
--connect jdbc:mysql://${rds_connection}/${database} \  
--username ${username} \  
--password ${password} \  
--table member_score \  
--warehouse-dir /user/hadoop/cred_financials_data \  
--delete-target-dir \  
--num-mappers 1 \  
--compress
```

Where `${rds_connection}` = upgradawsrds1.cyaielec9bmnf.us-east-1.rds.amazonaws.com  
`${database}` = cred\_financials\_data  
`${username}` = upgraduser  
`${password}` = upgraduser

The above two commands are wrapped into a single shell script with name **data\_ingestion.sh** as shown below:

```
[hadoop@ip-172-31-80-45 script]$ cat data_ingestion.sh
#!/bin/bash
# Sqoop Import Card Member and Member Score data from AWS RDS to HDFS
# Execution Command: ./sqoop_data_ingestion.sh <AWS RDS Connection String> <database> <username> <password>

# AWS RDS Credentials
rds_connection=$1
database=$2
username=$3
password=$4

# Sqoop Import - Card Member table
sqoop import \
--connect jdbc:mysql://${rds_connection}/${database} \
--username ${username} \
--password ${password} \
--table card_member \
--warehouse-dir /user/hadoop/cred_financials_data \
--delete-target-dir \
--num-mappers 1 \
--compress

# Sqoop Import - Member Score table
sqoop import \
--connect jdbc:mysql://${rds_connection}/${database} \
--username ${username} \
--password ${password} \
--table member_score \
--warehouse-dir /user/hadoop/cred_financials_data \
--delete-target-dir \
--num-mappers 1 \
--compress
```

Execute the data ingestion script by running the below command:

```
/home/hadoop/cred_financials_data/script/data_ingestion.sh
upgradawsrds1.cyaiehc9bmnf.us-east-1.rds.amazonaws.com cred_financials_data
upgraduser upgraduser
```

## Command to see the list of imported data in HDFS

### Card Member: -

```
hadoop fs -ls /user/hadoop/cred_financials_data/card_member
```

### Member Score: -

```
hadoop fs -ls /user/hadoop/cred_financials_data/member_score
```

## Screenshot of the imported data

### Card Member: -

```
[hadoop@ip-172-31-80-45 ~]$ hadoop fs -ls /user/hadoop/cred_financials_data/card_member
Found 2 items
-rw-r--r-- 1 hadoop hadoop          0 2022-12-17 18:30 /user/hadoop/cred_financials_data/card_member/_SUCCESS
-rw-r--r-- 1 hadoop hadoop    34628 2022-12-17 18:30 /user/hadoop/cred_financials_data/card_member/part-m-00000.gz
```

### Member Score: -

```
[hadoop@ip-172-31-80-45 ~]$ hadoop fs -ls /user/hadoop/cred_financials_data/member_score
Found 2 items
-rw-r--r-- 1 hadoop hadoop          0 2022-12-17 18:30 /user/hadoop/cred_financials_data/member_score/_SUCCESS
-rw-r--r-- 1 hadoop hadoop    10186 2022-12-17 18:30 /user/hadoop/cred_financials_data/member_score/part-m-00000.gz
```