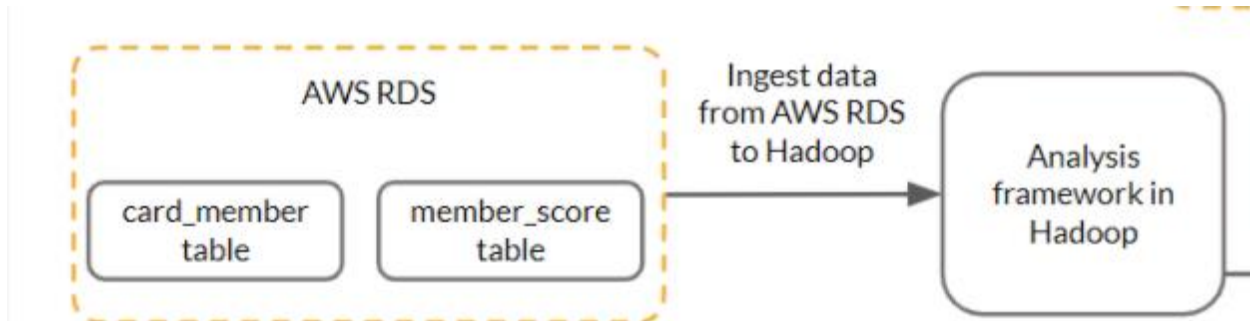


## Data Ingestion from the RDS to HDFS using Sqoop

Here we are connecting to RDS connection string for importing member\_score and card\_member tables and Ingesting the relevant data from AWS RDS to Hadoop.



**Connect string:** jdbc:mysql://upgradawsrds1.cyaielc9bmnf.us-east1.rds.amazonaws.com/cred\_financials\_data

**Source tables** are member\_score and card\_member

**Target directory:** /user/root/capstone\_project

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

```
sqoop import \
--connect jdbc:mysql://upgradawsrds1.cyaielc9bmnf.us-east1.rds.amazonaws.com/cred_financials_data \
--table member_score \
--username upgraduser --password upgraduser \
--target-dir /user/root/capstone_project/member_score \
-m 1
```

```
sqoop import \
--connect jdbc:mysql://upgradawsrds1.cyaielc9bmnf.us-east1.rds.amazonaws.com/cred_financials_data \
--table card_member \
--username upgraduser --password upgraduser \
--target-dir /user/root/capstone_project/card_member \
-m 1
```

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

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

```
[root@ip-10-0-0-149 ~]# hadoop fs -ls /user/root/capstone_project
Found 2 items
drwxr-xr-x - root supergroup          0 2021-10-16 07:29 /user/root/capstone_project/card_member
drwxr-xr-x - root supergroup          0 2021-10-16 07:23 /user/root/capstone_project/member_score
```

## Screenshot of the imported data

`hadoop fs -ls /user/root/capstone_project/card_member`

`hadoop fs -ls /user/root/capstone_project/member_score`

```
[root@ip-10-0-0-149 ~]# hadoop fs -ls /user/root/capstone_project/card_member
Found 2 items
-rw-r--r--  3 root supergroup          0 2021-10-16 07:29 /user/root/capstone_project/card_member/_SUCCESS
-rw-r--r--  3 root supergroup    85082 2021-10-16 07:29 /user/root/capstone_project/card_member/part-m-00000
[root@ip-10-0-0-149 ~]#
```

```
[root@ip-10-0-0-149 ~]# hadoop fs -ls /user/root/capstone_project/member_score
Found 2 items
-rw-r--r--  3 root supergroup          0 2021-10-16 07:23 /user/root/capstone_project/member_score/_SUCCESS
-rw-r--r--  3 root supergroup    19980 2021-10-16 07:23 /user/root/capstone_project/member_score/part-m-00000
[root@ip-10-0-0-149 ~]#
```

## For viewing data used cat command

`hadoop fs -cat /user/root/capstone_project/member_score/part-m-00000`

```
[root@ip-10-0-0-149 ~]# hadoop fs -cat /user/root/capstone_project/member_score/part-m-00000
000037495066290,339
000117826301530,289
001147922084344,393
001314074991813,225
001739553947511,642
003761426295463,413
004494068832701,217
006836124210484,504
006991872634058,697
007955566230397,372
008732267588672,213
008765307152821,399
009136568025042,308
009190444424572,559
009250698176266,233
009873334520465,298
011716573646690,249
011877954983420,497
012390918683920,407
012731668664932,612
012871894714576,689
015235737327750,244
015509173543086,176
015582765997171,345
015777120121252,464
019834203768714,670
019999364464243,604
023329987813357,268
```

hadoop fs -cat /user/root/capstone\_project/card\_member/part-m-00000

```
[root@ip-10-0-0-149 ~]# hadoop fs -cat /user/root/capstone_project/card_member/part-m-00000
340028465709212,009250698176266,2012-02-08 06:04:13.0,05/13,United States,Barberton
340054675199675,835873341185231,2017-03-10 09:24:44.0,03/17,United States,Fort Dodge
340082915339645,512969555857346,2014-02-15 06:30:30.0,07/14,United States,Graham
340134186926007,887711945571282,2012-02-05 01:21:58.0,02/13,United States,Dix Hills
340265728490548,680324265406190,2014-03-29 07:49:14.0,11/14,United States,Rancho Cucamonga
340268219434811,929799084911715,2012-07-08 02:46:08.0,08/12,United States,San Francisco
340379737226464,089615510858348,2010-03-10 00:06:42.0,09/10,United States,Clinton
340383645652108,181180599313885,2012-02-24 05:32:44.0,10/16,United States,West New York
340803866934451,417664728506297,2015-05-21 04:30:45.0,08/17,United States,Beaverton
340889618969736,459292914761635,2013-04-23 08:40:11.0,11/15,United States,West Palm Beach
340924125838453,188119365574843,2011-04-12 04:28:47.0,12/13,United States,Scottsbluff
341005627432127,872138964937565,2013-09-08 03:16:50.0,02/17,United States,Chillum
341029651579925,974087224071871,2011-01-14 00:20:25.0,08/12,United States,Valley Station
341311317050937,561687420200207,2014-03-18 06:23:23.0,02/15,United States,Vincennes
341344252914274,695906467918552,2012-03-02 03:21:01.0,03/13,United States,Columbine
341363858179050,009190444424572,2012-02-19 05:16:44.0,04/14,United States,Cheektowaga
341519629171378,533670008048847,2013-05-13 07:59:32.0,01/15,United States,Centennial
341641153427489,230523184584316,2013-03-25 08:51:18.0,11/15,United States,Colchester
341719092861087,304847505155781,2015-12-06 08:06:35.0,11/17,United States,Vernon Hills
341722035429601,979218131207765,2015-12-22 10:46:23.0,01/17,United States,Elk Grove Village
341724964458347,210778177559185,2011-02-07 01:43:13.0,03/14,United States,Fond du Lac
341920010925925,185524819205415,2010-11-15 00:57:15.0,10/16,United States,Roosevelt
342016727481601,329679840450294,2015-02-15 06:56:32.0,10/16,United States,Oakdale
342048672911139,409346767454057,2011-03-03 10:41:02.0,08/12,United States,Winter Springs
342106756753728,439669104078192,2013-11-25 02:05:46.0,03/14,United States,Pittsburg
342136378319429,756971727303005,2016-01-09 01:19:50.0,10/17,United States,Superior
342341653931950,569294535026544,2014-09-06 12:16:48.0,10/14,United States,Woodstock
342400733369183,557238536043272,2016-03-01 07:30:53.0,08/17,United States,Derry
342409040326783,093291208532654,2015-05-02 10:06:45.0,04/16,United States,Key West
342541202178377,228647326620185,2013-12-28 05:34:26.0,04/15,United States,Taylors
342659393657862,075642075917789,2016-02-10 07:27:55.0,07/17,United States,Norman
342679959520647,328980632419725,2017-06-20 06:30:36.0,07/17,United States,Oklahoma City
```