

## 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.cyaie1c9bmnf.us-east-1.rds.amazonaws.com/  
testdatabase \  
--table SRC_ATM_TRANS \  
--username student --password STUDENT123 \  
--target-dir /user/root/SRC_ATM_TRANS \  
-m 1
```

**In the screenshot below, we can see the result of the Sqoop Import Job, 2468572 records have been retrieved**

```
22/12/13 18:16:44 INFO mapreduce.Job: Running job: job_1670954330129_0001  
22/12/13 18:16:53 INFO mapreduce.Job: Job job_1670954330129_0001 running in uber mode : false  
22/12/13 18:16:53 INFO mapreduce.Job: map 0% reduce 0%  
22/12/13 18:17:19 INFO mapreduce.Job: map 100% reduce 0%  
22/12/13 18:17:20 INFO mapreduce.Job: Job job_1670954330129_0001 completed successfully  
22/12/13 18:17:21 INFO mapreduce.Job: Counters: 30  
File System Counters  
FILE: Number of bytes read=0  
FILE: Number of bytes written=189664  
FILE: Number of read operations=0  
FILE: Number of large read operations=0  
FILE: Number of write operations=0  
HDFS: Number of bytes read=87  
HDFS: Number of bytes written=531214815  
HDFS: Number of read operations=4  
HDFS: Number of large read operations=0  
HDFS: Number of write operations=2  
Job Counters  
Launched map tasks=1  
Other local map tasks=1  
Total time spent by all maps in occupied slots (ms)=1170288  
Total time spent by all reduces in occupied slots (ms)=0  
Total time spent by all map tasks (ms)=24381  
Total vcore-milliseconds taken by all map tasks=24381  
Total megabyte-milliseconds taken by all map tasks=37449216  
Map-Reduce Framework  
Map input records=2468572  
Map output records=2468572  
Input split bytes=87  
Spilled Records=0  
Failed Shuffles=0  
Merged Map outputs=0  
GC time elapsed (ms)=234  
CPU time spent (ms)=27260  
Physical memory (bytes) snapshot=613580800  
Virtual memory (bytes) snapshot=3290480640  
Total committed heap usage (bytes)=537395200  
File Input Format Counters  
Bytes Read=0  
File Output Format Counters  
Bytes Written=531214815  
22/12/13 18:17:21 INFO mapreduce.ImportJobBase: Transferred 506.6059 MB in 42.9881 seconds (11.7848 MB/sec)  
22/12/13 18:17:21 INFO mapreduce.ImportJobBase: Retrieved 2468572 records.  
[root@ip-172-31-48-160 ~]#
```

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

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

```
[root@ip-172-31-48-160 ~]# hadoop fs -ls /user/root/SRC_ATM_TRANS
Found 2 items
-rw-r--r--  1 root hadoop          0 2022-12-13 18:17 /user/root/SRC_ATM_TRANS/_SUCCESS
-rw-r--r--  1 root hadoop 531214815 2022-12-13 18:17 /user/root/SRC_ATM_TRANS/part-m-00000
[root@ip-172-31-48-160 ~]#
```

In the screenshot above we can see two items:

- The first file is the success file, indicating that the MapReduce job was successful.
- The second file 'part-m-00000' is the one that I imported. Since I used only one mapper in my import command, thus the data is in a single file.

**Screenshot of the imported data:**

A portion of data read from part-m-00000 file:

```

2017, December, 31, Sunday, 23, Active, 24, NCR, Hobro, Adelgade, 31, 9500, 56.638, 9.794, DKK, Mastercard - on-us, 7780, Withdrawal, , 56.643, 9.790, 2620167, Hobro, 277.589, 999, 87, 6, 208, 0.000, 76, 803, Clouds, broken clouds
2017, December, 31, Sunday, 23, Active, 20, NCR, Bispensgade, Bispensgade, 35, 9800, 57.453, 9.996, DKK, Mastercard - on-us, 2830, Withdrawal, , 57.464, 9.982, 2620214, Hjørring, 278.589, 998, 99, 8, 210, 0.000, 44, 802, Clouds, scattered clouds
2017, December, 31, Sunday, 23, Active, 28, NCR, Låfå, gståfå, r, Åfæsterbrogade, 8, 9670, 56.967, 9.253, DKK, VISA, 812, Withdrawal, , 56.962, 9.258, 2617467, Logstor, 277.589, 999, 87, 6, 208, 0.000, 76, 803, Clouds, broken clouds
2017, December, 31, Sunday, 23, Active, 28, NCR, Låfå, gståfå, r, Åfæsterbrogade, 8, 9670, 56.967, 9.253, DKK, VISA, 5568, Withdrawal, , 56.962, 9.258, 2617467, Logstor, 277.589, 999, 87, 6, 208, 0.000, 76, 803, Clouds, broken clouds
2017, December, 31, Sunday, 23, Active, 49, NCR, Bindslev, NÅfå, rrebro, 18, 9881, 57.541, 10.200, DKK, MasterCard, 5403, Withdrawal, , 57.471, 10.203, 2614010, Sindal, 277.589, 999, 87, 6, 208, 0.000, 76, 803, Clouds, broken clouds
2017, December, 31, Sunday, 23, Active, 49, NCR, Bindslev, NÅfå, rrebro, 18, 9881, 57.541, 10.200, DKK, MasterCard, 312, Withdrawal, , 57.471, 10.203, 2614010, Sindal, 277.589, 999, 87, 6, 208, 0.000, 76, 803, Clouds, broken clouds
2017, December, 31, Sunday, 23, Active, 56, Diebold Nixdorf, KÅfå, ge, SÅfå, ndre Alle, 1, 4600, 55.454, 12.181, DKK, VISA, Dankort - on-us, 5351, Withdrawal, , 55.588, 12.251, 2621218, Greve Kommune, 280.150, 992, 100, 3, 190, 0.000, 90, 804, Clouds, overcast clouds
2017, December, 31, Sunday, 23, Active, 103, Diebold Nixdorf, Vejgaard, Hadsundvej, 20, 9000, 57.043, 9.950, DKK, Mastercard - on-us, 5821, Withdrawal, , 57.048, 9.935, 2616235, NÅfå, rresundby, 277.589, 999, 87, 6, 208, 0.000, 76, 803, Clouds, broken clouds
2017, December, 31, Sunday, 23, Active, 1, NCR, NÅfå, stved, Farinagsvej, 8, 4700, 55.233, 11.763, DKK, VISA, Dankort, 3049, Withdrawal, , 55.230, 11.761, 2616038, Næstved, 280.150, 991, 100, 5, 210, 1.175, 92, 500, Rain, light rain
2017, December, 31, Sunday, 23, Active, 85, Diebold Nixdorf, KÅfå, benhavn, Regnbuepladsen, 5, 1550, 55.676, 12.571, DKK, MasterCard, 7785, Withdrawal, , 55.676, 12.566, 2618425, Copenhagen, 280.150, 992, 100, 3, 190, 0.000, 90, 804, Clouds, overcast clouds
2017, December, 31, Sunday, 23, Active, 15, NCR, Vestre, Kastetvej, 36, 9000, 57.053, 9.905, DKK, Mastercard - on-us, 9777, Withdrawal, , 57.048, 9.919, 2624886, Aalborg, 277.589, 999, 87, 6, 208, 0.000, 76, 803, Clouds, broken clouds
2017, December, 31, Sunday, 23, Active, 15, NCR, Vestre, Kastetvej, 36, 9000, 57.053, 9.905, DKK, MasterCard, 5531, Withdrawal, , 57.048, 9.919, 2624886, Aalborg, 277.589, 999, 87, 6, 208, 0.000, 76, 803, Clouds, broken clouds
2017, December, 31, Sunday, 23, Active, 34, NCR, Skipperen, Vestre Alle, 2, 9000, 57.034, 9.908, DKK, MasterCard, 8323, Withdrawal, , 57.048, 9.919, 2624886, Aalborg, 277.589, 999, 87, 6, 208, 0.000, 76, 803, Clouds, broken clouds
2017, December, 31, Sunday, 23, Active, 20, NCR, Bispensgade, Bispensgade, 35, 9800, 57.453, 9.996, DKK, MasterCard, 5023, Withdrawal, , 57.464, 9.982, 2620214, Hjørring, 278.589, 998, 99, 8, 210, 0.000, 44, 802, Clouds, scattered clouds
2017, December, 31, Sunday, 23, Active, 34, NCR, Skipperen, Vestre Alle, 2, 9000, 57.034, 9.908, DKK, MasterCard, 147, Withdrawal, , 57.048, 9.919, 2624886, Aalborg, 277.589, 999, 87, 6, 208, 0.000, 76, 803, Clouds, broken clouds
2017, December, 31, Sunday, 23, Active, 70, Diebold Nixdorf, Holstebro, Høstrupsvej, 6, 7500, 56.373, 8.625, DKK, MasterCard, 5666, Withdrawal, , 56.360, 8.616, 2620046, Holstebro, 280.150, 998, 88, 93, 4, 210, 0.000, 92, 804, Clouds, overcast clouds
2017, December, 31, Sunday, 23, Active, 49, NCR, Bindslev, NÅfå, rrebro, 18, 9881, 57.541, 10.200, DKK, MasterCard, 7886, Withdrawal, , 57.471, 10.203, 2614010, Sindal, 277.589, 999, 87, 6, 208, 0.000, 76, 803, Clouds, broken clouds
2017, December, 31, Sunday, 23, Inactive, 12, NCR, ÅfæsterÅfÅV Duus, ÅfæsterÅfÅV, 12, 9000, 57.049, 9.922, DKK, Mastercard - on-us, 2436, Withdrawal, , 57.048, 9.919, 2624886, Aalborg, 277.589, 999, 87, 6, 208, 0.000, 76, 803, Clouds, broken clouds
2017, December, 31, Sunday, 23, Inactive, 12, NCR, ÅfæsterÅfÅV Duus, ÅfæsterÅfÅV, 12, 9000, 57.049, 9.922, DKK, Mastercard - on-us, 8519, Withdrawal, , 57.048, 9.919, 2624886, Aalborg, 277.589, 999, 87, 6, 208, 0.000, 76, 803, Clouds, broken clouds
2017, December, 31, Sunday, 23, Active, 10, NCR, NÅfå, rresundby, Torvet, 6, 9400, 57.059, 9.922, DKK, Mastercard - on-us, 5286, Withdrawal, , 57.048, 9.919, 2624886, Aalborg, 277.589, 999, 87, 6, 208, 0.000, 76, 803, Clouds, broken clouds
2017, December, 31, Sunday, 23, Active, 37, NCR, Silkeborg, Borgergade, 36, 8600, 56.179, 9.552, DKK, VISA, 876, Withdrawal, , 56.170, 9.545, 2614030, Silkeborg, 279.800, 988, 93, 4, 210, 0.000, 8, 804, Clouds, overcast clouds
[root@ip-172-31-73-111 ~]#

```