

## 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 \  
--username student \  
--password STUDENT123 \  
--table SRC_ATM_TRANS \  
--target-dir /user/hdfs/SRC_ATM_TRANS \  
--m 1
```

```
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)=239  
  CPU time spent (ms)=27880  
  Physical memory (bytes) snapshot=628543488  
  Virtual memory (bytes) snapshot=3299880960  
  Total committed heap usage (bytes)=527433728  
File Input Format Counters  
  Bytes Read=0  
File Output Format Counters  
  Bytes Written=531214815  
24/07/28 21:01:58 INFO mapreduce.ImportJobBase: Transferred 506.6059 MB in 43.4474 seconds (11.6602 MB/sec)  
24/07/28 21:01:58 INFO mapreduce.ImportJobBase: Retrieved 2468572 records.  
[hadoop@ip-172-31-50-227 ~]$ |
```

The above screenshot shows Retrieved 2468572 records

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

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

```
[hadoop@ip-172-31-50-227 ~]$ hadoop fs -ls /user/root/SRC_ATM_TRANS  
Found 2 items  
-rw-r--r--  1 hadoop hadoop          0 2024-07-28 21:01 /user/root/SRC_ATM_TRANS/_SUCCESS  
-rw-r--r--  1 hadoop hadoop 531214815 2024-07-28 21:01 /user/root/SRC_ATM_TRANS/part-m-000000  
[hadoop@ip-172-31-50-227 ~]$ |
```

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:

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

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

```
2017,December,31,Sunday,23,Active,57,NCR,HillerÅfÅ,d,KÅfÅ,benhavnsvej,31,3400,55.933,12.314,DKK,VISA,2690,Withdrawal,,55.767,12.433,2621394,Gladsakse Kommune,280.150,992,100,3,190,0.000,90,804,Clouds,overcast clouds
2017,December,31,Sunday,23,Active,6,NCR,Fredricia,SjÅfÅllandsgade,33,7000,55.564,9.787,DKK,MasterCard,8122,Withdrawal,,55.566,9.753,2621951,Fredricia,280.150,990,100,6,220,0.000,92,701,Mist,mist
2017,December,31,Sunday,23,Active,21,NCR,ÅfEasterÅfÅv,NÅfÅ,ller,ÅfEasterÅfÅv,12,9000,57.049,9.922,DKK,MasterCard - on-us,2542,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,6,NCR,Fredricia,SjÅfÅllandsgade,33,7000,55.564,9.757,DKK,Visa Dankort,8341,Withdrawal,,55.566,9.753,2621951,Fredricia,280.150,990,100,6,220,0.000,92,701,Mist,mist
2017,December,31,Sunday,23,Active,51,NCR,Gistrup,Hadsundvej,346,9260,56.997,9.993,DKK,Visa Dankort,4079,Withdrawal,,56.994,9.991,2621449,Gistrup,277.589,999,87,6,208,0.000,76,803,Clouds,broken clouds
2017,December,31,Sunday,23,Active,19,NCR,Holding,Vejlevej,135,6000,55.585,9.457,DKK,MasterCard,5370,Withdrawal,,55.490,9.472,2618528,Holding,280.150,990,100,6,220,0.000,92,701,Mist,mist
2017,December,31,Sunday,23,Active,12,NCR,Bispensgade,Bispensgade,35,9800,57.453,9.996,DKK,MasterCard - on-us,4630,Withdrawal,,57.464,9.982,2620214,Hjorring,278.589,998,99,8,210,0.000,44,802,Clouds,scattered clouds
2017,December,31,Sunday,23,Active,20,NCR,Bispensgade,Bispensgade,35,9800,57.453,9.996,DKK,MasterCard - on-us,7031,Withdrawal,,57.464,9.982,2620214,Hjorring,278.589,998,99,8,210,0.000,44,802,Clouds,scattered clouds
2017,December,31,Sunday,23,Active,80,NCR,Menu,KÅfÅ,bmand,Klarup,Klarupvej,52,9270,57.013,10.046,DKK,MasterCard,1263,Withdrawal,,57.004,10.101,2612301,Storvorde,277.589,999,87,6,208,0.000,76,803,Clouds,broken clouds
2017,December,31,Sunday,23,Active,32,NCR,SÅfÅdding,Tarpbagevej,59,6710,55.498,8.488,DKK,Visa Dankort,5221,Withdrawal,,55.407,8.450,2622447,Esbjerg,280.150,989,93,4,180,0.000,92,701,Mist,mist
2017,December,31,Sunday,23,Active,63,NCR,Brugsen i Breum,AakjÅfÅrsvej,1,7870,56.688,9.869,DKK,MasterCard,2317,Withdrawal,,56.586,9.136,2619908,Hojlslev,277.589,999,87,6,208,0.000,76,803,Clouds,broken clouds
2017,December,31,Sunday,23,Inactive,12,NCR,ÅfEasterÅfÅv,Duus,ÅfEasterÅfÅv,12,9000,57.049,9.922,DKK,Visa Dankort - on-us,5924,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,11,NCR,Sauersvej,Fridtjof Nansens Vej,2,0210,57.023,9.940,DKK,VISA,2667,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,46,Diebold Nixdorf,HelsingÅfÅ,r,Sct. Olai Gade,39,3000,56.036,12.612,DKK,MasterCard - on-us,5563,Withdrawal,,56.036,12.614,2620473,Helsingor,279.830,991,93,2,190,0.000,88,701,Mist,mist
2017,December,31,Sunday,23,Active,47,NCR,Fredriksberg,Gammel Kongevej,157,1850,55.677,12.537,DKK,MasterCard,9815,Withdrawal,,55.679,12.535,2621902,Fredriksberg,279.890,992,100,3,190,0.000,90,701,Mist,mist
2017,December,31,Sunday,23,Active,46,Diebold Nixdorf,HelsingÅfÅ,r,Sct. Olai Gade,39,3000,56.036,12.612,DKK,MasterCard - on-us,2384,Withdrawal,,56.036,12.614,2620473,Helsingor,279.830,991,93,2,190,0.000,88,701,Mist,mist
2017,December,31,Sunday,23,Active,64,NCR,Horsens,GrÅfÅ,nlandsvej,5,8700,55.859,9.854,DKK,Visa Dankort - on-us,2075,Withdrawal,,55.861,9.850,2619771,Horsens,279.900,989,100,4,200,0.138,92,500,Rain,light rain
2017,December,31,Sunday,23,Active,64,NCR,Horsens,GrÅfÅ,nlandsvej,5,8700,55.859,9.854,DKK,Visa Dankort - on-us,5223,Withdrawal,,55.861,9.850,2619771,Horsens,279.900,989,100,4,200,0.138,92,500,Rain,light rain
2017,December,31,Sunday,23,Active,64,NCR,Horsens,GrÅfÅ,nlandsvej,5,8700,55.859,9.854,DKK,Visa Dankort,0690,Withdrawal,,55.861,9.850,2619771,Horsens,279.900,989,100,4,200,0.138,92,500,Rain,light rain
2017,December,31,Sunday,23,Active,24,NCR,Hobro,Adelgade,31,9500,56.638,9.794,DKK,MasterCard - on-us,7789,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,Hjorring,278.589,998,99,8,210,0.000,44,802,Clouds,scattered clouds
2017,December,31,Sunday,23,Active,28,NCR,LÅfÅ,gtÅfÅ,r,ÅfEasterbrogade,8,9070,56.907,9.283,DKK,VISA,812,Withdrawal,,56.902,9.283,2617407,Logstor,277.589,999,87,6,208,0.000,76,803,Clouds,broken clouds
2017,December,31,Sunday,23,Active,28,NCR,LÅfÅ,gtÅfÅ,r,ÅfEasterbrogade,8,9070,56.907,9.283,DKK,VISA,5660,Withdrawal,,56.902,9.283,2617407,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,1,4000,55.454,12.181,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,istved,Farimagvej,8,4700,55.233,11.763,DKK,Visa Dankort,3049,Withdrawal,,55.230,11.761,2616038,Naestved,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,35,9800,57.403,9.996,DKK,MasterCard,5023,Withdrawal,,57.464,9.982,2620214,Hjorring,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,Hostrupvej,6,7500,56.373,8.625,DKK,MasterCard,5666,Withdrawal,,56.360,8.616,2620046,Holstebro,280.150,988,93,4,210,0.000,92,804,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,ÅfEasterÅfÅv,Duus,ÅfEasterÅfÅv,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,ÅfEasterÅfÅv,Duus,ÅfEasterÅfÅv,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,36,0600,56.179,9.552,DKK,VISA,076,Withdrawal,,56.170,9.545,2610030,Silkeborg,279.800,988,93,4,210,0.000,88,804,Clouds,overcast clouds
[hadoop@ip-172-31-60-75 ~]$
```