

# 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://upgradtest.cyaiehc9bmnf.us-east-1.rds.amazonaws.com/testdatabase \  
stdatabase \  
--username student \  
--password STUDENT123 \  
--table SRC_ATM_TRANS \  
--target-dir /user/root/SRC_ELT_ATM_TRANS \  
-m 1
```

```
[hadoop@ip-172-31-43-169 ~]$ sqoop import \  
> --connect jdbc:mysql://upgradtest.cyaiehc9bmnf.us-east-1.rds.amazonaws.com/testdatabase \  
> --username student \  
> --password STUDENT123 \  
> --table SRC_ATM_TRANS \  
> --target-dir /user/root/SRC_ELT_ATM_TRANS \  
> -m 1
```

hadoop@ip-172-31-43-169:~

```
FILE: Number of bytes read=0  
FILE: Number of bytes written=298603  
FILE: Number of read operations=0  
FILE: Number of large read operations=0  
FILE: Number of write operations=0  
HDFS: Number of bytes read=85  
HDFS: Number of bytes written=531214815  
HDFS: Number of read operations=6  
HDFS: Number of large read operations=0  
HDFS: Number of write operations=2  
HDFS: Number of bytes read erasure-coded=0  
Job Counters  
  Launched map tasks=1  
  Other local map tasks=1  
  Total time spent by all maps in occupied slots (ms)=2165856  
  Total time spent by all reduces in occupied slots (ms)=0  
  Total time spent by all map tasks (ms)=22561  
  Total vcore-milliseconds taken by all map tasks=22561  
  Total megabyte-milliseconds taken by all map tasks=69307392  
Map-Reduce Framework  
  Map input records=2468572  
  Map output records=2468572  
  Input split bytes=85  
  Spilled Records=0  
  Failed Shuffles=0  
  Merged Map outputs=0  
  GC time elapsed (ms)=143  
  CPU time spent (ms)=20910  
  Physical memory (bytes) snapshot=1051475968  
  Virtual memory (bytes) snapshot=4425187328  
  Total committed heap usage (bytes)=968359936  
  Peak Map Physical memory (bytes)=1051475968  
  Peak Map Virtual memory (bytes)=4433252352  
File Input Format Counters  
  Bytes Read=0  
File Output Format Counters  
  Bytes Written=531214815  
2024-02-04 07:10:17,280 INFO mapreduce.ImportJobBase: Transferred 506.6059 MB in 35.2826 seconds (14.3585 MB/sec)  
2024-02-04 07:10:17,282 INFO mapreduce.ImportJobBase: Retrieved 2468572 records.
```

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

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

```
[hadoop@ip-172-31-43-169 ~]$ hadoop fs -ls /user/root/SRC_ELT_ATM_TRANS
Found 2 items
-rw-r--r-- 1 hadoop hdfsadmingroup 0 2024-02-04 07:10 /user/root/SRC_ELT_ATM_TRANS/ SUCCESS
-rw-r--r-- 1 hadoop hdfsadmingroup 531214815 2024-02-04 07:10 /user/root/SRC_ELT_ATM_TRANS/part-m-00000
```

**Screenshot of the imported data:**

```
hadoop fs -cat /user/root/SRC_ATM_TRANS/part-m-00000
```

```
root@ip-172-31-34-38~
2017, March, 31, Friday, 15, Active, 44, NCR, Lyngby, Jernbanevej, 6, 2800, 55.772, 12.500, DKK, Visa Dankort - on-us, 6469, Withdrawal, , , 55.770, 12.504, 2618468, Kongens Lyngby, 290.180, 10
13, 71, 4, 150, 0.000, 0.800, Clear, Sky is Clear
2017, March, 31, Friday, 15, Active, 33, NCR, Vadum, Ellehammersvej, 43, 9430, 57.118, 9.861, DKK, Visa Dankort - on-us, 239, Withdrawal, , , 57.118, 9.857, 2610823, Vadum, 288.150, 1010, 77, 4, 1
50, 0.000, 48, 802, Clouds, scattered clouds
2017, March, 31, Friday, 15, Active, 34, NCR, Skipperen, Vestre Alle, 2, 9000, 57.034, 9.908, DKK, Visa Dankort, 1786, Withdrawal, , , 57.048, 9.919, 2624886, Aalborg, 288.150, 1010, 77, 4, 150, 0.
000, 48, 802, Clouds, scattered clouds
2017, March, 31, Friday, 15, Active, 9, Diebold Nixdorf, Hadsund, Storegade, 12, 9560, 56.716, 10.114, DKK, Visa Dankort - on-us, 8215, Withdrawal, , , 56.715, 10.117, 2620952, Hadsund, 288.15
0, 1010, 77, 4, 150, 0.000, 48, 802, Clouds, scattered clouds
2017, March, 31, Friday, 15, Active, 22, NCR, Roskilde, K  f  , benhavnsvej, 65, 4000, 55.642, 12.106, DKK, Visa Dankort, 5718, Withdrawal, , , 55.642, 12.080, 2614481, Roskilde, 288.920, 1012, 59,
4, 160, 0.000, 32, 802, Clouds, scattered clouds
2017, March, 31, Friday, 15, Active, 46, Diebold Nixdorf, Helsing  f  , r, Sct. Olai Gade, 39, 3000, 56.036, 12.612, DKK, Visa Dankort - on-us, 3946, Withdrawal, , , 56.036, 12.614, 2620473, Hel
singor, 290.150, 1013, 55, 3, 190, 0.000, 44, 802, Clouds, scattered clouds
2017, March, 31, Friday, 15, Active, 39, NCR, Svenstrup, Godth  f  bsvej, 14, 9230, 56.973, 9.851, DKK, Visa Dankort, 2573, Withdrawal, , , 56.972, 9.848, 2612021, Svenstrup, 288.150, 1010, 77, 4,
150, 0.000, 48, 802, Clouds, scattered clouds
2017, March, 31, Friday, 15, Inactive, 14, NCR, H  f  , rning, N  f  , rrealte, 12, 8362, 56.086, 10.037, DKK, MasterCard, 1027, Withdrawal, , , 56.157, 10.211, 2624652, Aarhus, 292.390, 1011, 55, 3, 160
, 0.000, 48, 802, Clouds, scattered clouds
2017, March, 31, Friday, 15, Active, 31, NCR, Slagelse, Mariendals Alle, 29, 4200, 55.398, 11.342, DKK, Visa Dankort, 7416, Withdrawal, , , 55.403, 11.355, 2613460, Slagelse, 292.150, 1012, 59, 4
, 160, 0.000, 44, 802, Clouds, scattered clouds
2017, March, 31, Friday, 15, Active, 20, NCR, Bispensgade, Bispensgade, 35, 9800, 57.453, 9.996, DKK, Visa Dankort - on-us, 4001, Withdrawal, , , 57.464, 9.982, 2620214, H  r  ring, 288.150, 1010
, 77, 4, 150, 0.000, 64, 803, Clouds, broken clouds
2017, March, 31, Friday, 15, Active, 80, NCR, Menu K  f  , bmand Klarup, Klarupvej, 52, 9270, 57.013, 10.046, DKK, Visa Dankort - on-us, 7758, Withdrawal, , , 57.004, 10.101, 2612301, St  r  rde,
288.150, 1010, 77, 4, 150, 0.000, 48, 802, Clouds, scattered clouds
2017, March, 31, Friday, 15, Active, 53, NCR, S  f  , by Syd, Trafikcenter S  f  , by Syd, 1, 9300, 57.313, 10.450, DKK, Mastercard - on-us, 6785, Withdrawal, , , 57.441, 10.537, 2621927, Frederiks
havn, 288.150, 1010, 77, 4, 150, 0.000, 64, 803, Clouds, broken clouds
2017, March, 31, Friday, 15, Active, 19, NCR, Kolding, Vejlevej, 135, 6000, 55.505, 9.457, DKK, VISA, 652, Withdrawal, , , 55.490, 9.472, 2618528, Kolding, 292.130, 1010, 49, 5, 180, 0.000, 0, 800, Cl
ear, Sky is Clear
2017, March, 31, Friday, 15, Active, 1, NCR, N  f  , stved, Farimagsgvej, 8, 4700, 55.233, 11.763, DKK, MasterCard, 2905, Withdrawal, , , 55.230, 11.761, 2616038, N  r  stved, 292.150, 1012, 59, 4, 160, 0
.000, 56, 803, Clouds, broken clouds
2017, March, 31, Friday, 15, Active, 105, Diebold Nixdorf, N  f  , rresundby, Torvet, 6, 9400, 57.059, 9.922, DKK, Dankort - on-us, 2570, Withdrawal, , , 57.048, 9.919, 2624886, Aalborg, 288.150,
1010, 77, 4, 150, 0.000, 48, 802, Clouds, scattered clouds
2017, March, 31, Friday, 15, Active, 78, Diebold Nixdorf, Nyborg, Vestergade, 35, 5800, 55.318, 10.781, DKK, Visa Dankort, 4133, Withdrawal, , , 55.060, 10.607, 2612045, Svendborg, 290.900, 101
0, 67, 4, 130, 0.000, 0, 800, Clear, Sky is Clear
2017, March, 31, Friday, 15, Active, 103, Diebold Nixdorf, Vejgaard, Hadsundvej, 20, 9000, 57.043, 9.950, DKK, Visa Dankort - on-us, 2956, Withdrawal, , , 57.048, 9.935, 2616235, N  f  , rresund
by, 288.150, 1010, 77, 4, 150, 0.000, 48, 802, Clouds, scattered clouds
2017, March, 31, Friday, 15, Active, 27, NCR, Herning, Dalgaasgade, 30, 7400, 56.135, 8.971, DKK, Mastercard - on-us, 2380, Withdrawal, , , 56.139, 8.974, 2620425, Herning, 294.150, 1009, 49, 5, 19
0, 0.000, 56, 803, Clouds, broken clouds
2017, March, 31, Friday, 15, Active, 26, NCR, Holb  f  , k, Slotsvolden, 7, 4300, 55.718, 11.704, DKK, Visa Dankort - on-us, 3032, Withdrawal, , , 55.700, 11.700, 2620141, Holb  f  , k Kommune, 290.
210, 1013, 71, 4, 150, 0.000, 0, 800, Clear, Sky is Clear
2017, March, 31, Friday, 15, Inactive, 16, NCR, Skive, Adelgade, 8, 7800, 56.567, 9.027, DKK, MasterCard, 8019, Withdrawal, , , 56.567, 9.033, 2613731, Skive, 294.150, 1009, 49, 5, 190, 0.000, 56, 80
3, Clouds, broken clouds
2017, March, 31, Friday, 15, Active, 104, NCR, Intern   f  wester  f  ,   f  wester  f  , 12, 9000, 57.049, 9.922, DKK, Mastercard - on-us, 1001, Withdrawal, , , 57.048, 9.919, 2624886, Aalborg, 288.
150, 1010, 77, 4, 150, 0.000, 48, 802, Clouds, scattered clouds
2017, March, 31, Friday, 15, Inactive, 52, NCR, Fars  f  , Torvet, 8, 9640, 56.771, 9.340, DKK, MasterCard, 949, Withdrawal, , , 56.773, 9.339, 2622310, Fars  , 288.150, 1010, 77, 4, 150, 0.000, 48, 80
2, Clouds, scattered clouds
```