

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.cyaieic9bmnf.us-east
1.rds.amazonaws.com/testdatabase
--table SRC_ATM_TRANS
--username student
--password STUDENT123
--target-dir /home/hadoop/SRC_ATM_TRANS/
-m 1
--as-textfile
```

```
hadoop@ip-172-31-86-195:~$
22/11/08 04:03:13 INFO mapreduce.Job: Counters: 30
File System Counters
  FILE: Number of bytes read=0
  FILE: Number of bytes written=189654
  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)=1149120
  Total time spent by all reduces in occupied slots (ms)=0
  Total time spent by all map tasks (ms)=23940
  Total vcore-milliseconds taken by all map tasks=23940
  Total megabyte-milliseconds taken by all map tasks=36771840
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)=290
  CPU time spent (ms)=28240
  Physical memory (bytes) snapshot=636735488
  Virtual memory (bytes) snapshot=3316473856
  Total committed heap usage (bytes)=536870912
File Input Format Counters
  Bytes Read=0
File Output Format Counters
  Bytes Written=531214815
22/11/08 04:03:13 INFO mapreduce.ImportJobBase: Transferred 506.6059 MB in 39.7642 seconds (12.7403 MB/sec)
22/11/08 04:03:13 INFO mapreduce.ImportJobBase: Retrieved 2468572 records.
[hadoop@ip-172-31-86-195 ~]$ hadoop fs -cp /home/hadoop/SRC_ATM_TRANS /user/livy/
[hadoop@ip-172-31-86-195 ~]$ hadoop fs -ls /user/livy
Found 2 items
drwxr-xr-x - livy livy 0 2022-11-08 03:00 /user/livy/.sparkStaging
drwxr-xr-x - hadoop livy 0 2022-11-08 04:03 /user/livy/SRC_ATM_TRANS
[hadoop@ip-172-31-86-195 ~]$ hadoop fs -ls /user/livy/SRC_ATM_TRANS
Found 2 items
-rw-r--r-- 1 hadoop livy 0 2022-11-08 04:03 /user/livy/SRC_ATM_TRANS/_SUCCESS
-rw-r--r-- 1 hadoop livy 531214815 2022-11-08 04:03 /user/livy/SRC_ATM_TRANS/part-m-00000
```

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

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

Screenshot of the imported data:



```
hadoop@ip-172-31-86-195:~$  
2017, January, 3, Tuesday, 10, Active, 35, NCR, Aabybro, ÅfEastergade, 6,9440, 57.162, 9.730, DKK, Mastercard - on-us, 99, Withdrawal, , , 57.158, 9.737, 2625037, Aabybro, 277.150, 1  
003, 86, 4, 260, 0.000, 75, 803, Clouds, broken clouds  
2017, January, 3, Tuesday, 10, Inactive, 88, NCR, Storcenter indg. A, Hobrovej, 452, 9200, 57.005, 9.876, DKK, MasterCard, 2530, Withdrawal, , , 56.972, 9.848, 2612021, Svenstrup, 2  
77.150, 1003, 86, 4, 260, 0.000, 75, 803, Clouds, broken clouds  
2017, January, 3, Tuesday, 10, Active, 62, Diebold Nixdorf, Terndrup, Bymidten, 2, 9575, 56.815, 10.057, DKK, Visa Dankort - on-us, 2350, Withdrawal, , , 56.715, 10.117, 2620952, H  
adsund, 277.150, 1003, 86, 4, 260, 0.000, 75, 803, Clouds, broken clouds  
2017, January, 3, Tuesday, 10, Active, 60, NCR, Hirtshale, JÅfÅ rgen Fibigersgade, 4, 9850, 57.591, 9.957, DKK, Visa Dankort - on-us, 4160, Withdrawal, , , 57.588, 9.959, 2620279, H  
irtshale, 277.150, 1003, 86, 4, 260, 0.000, 75, 803, Clouds, broken clouds  
2017, January, 3, Tuesday, 10, Active, 35, NCR, Aabybro, ÅfEastergade, 6,9440, 57.162, 9.730, DKK, Mastercard - on-us, 3750, Withdrawal, , , 57.158, 9.737, 2625037, Aabybro, 277.150  
1, 1003, 86, 4, 260, 0.000, 75, 803, Clouds, broken clouds  
2017, January, 3, Tuesday, 10, Active, 91, NCR, Skive Lobby, Adelgade, 8, 7800, 56.567, 9.027, DKK, HÅfÅ!vekort - on-us, 6269, Withdrawal, , , 56.567, 9.033, 2613731, Skive, 279.150  
1, 1006, 93, 4, 250, 0.000, 90, 310, Drizzle, light intensity drizzle rain  
2017, January, 3, Tuesday, 10, Active, 52, NCR, Intern HJÅfÅ rring, ÅfEastergade, 8, 9800, 57.459, 9.988, DKK, Visa Dankort - on-us, 96, Withdrawal, , , 57.464, 9.982, 2620214, HJ  
orring, 277.150, 1003, 86, 4, 260, 0.000, 75, 803, Clouds, broken clouds  
2017, January, 3, Tuesday, 10, Active, 87, NCR, Aalborg Storcenter indg. D, Hobrovej, 452, 9200, 57.005, 9.876, DKK, Visa Dankort, 1673, Withdrawal, , , 56.972, 9.848, 2612021, Sve  
nstrup, 277.150, 1003, 86, 4, 260, 0.000, 75, 803, Clouds, broken clouds  
2017, January, 3, Tuesday, 10, Active, 38, NCR, Hasserisvej, 113, 9000, 57.044, 9.898, DKK, Dankort - on-us, 8076, Withdrawal, , , 57.048, 9.919, 2624886, Aalborg, 277.150  
1, 1003, 86, 4, 260, 0.000, 75, 803, Clouds, broken clouds  
2017, January, 3, Tuesday, 10, Active, 49, NCR, Bindslev, NÅfÅ rrebro, 18, 9881, 57.541, 10.200, DKK, HÅfÅ!vekort, 5361, Withdrawal, , , 57.471, 10.203, 2614010, Sindal, 277.150, 100  
3, 86, 4, 260, 0.000, 75, 803, Clouds, broken clouds  
2017, January, 3, Tuesday, 10, Active, 80, NCR, Menu KÅfÅ bmand Klarup, Klarupvej, 52, 9270, 57.013, 10.046, DKK, HÅfÅ!vekort - on-us, 2576, Withdrawal, , , 57.004, 10.101, 261230  
1, Storvoerde, 277.150, 1003, 86, 4, 260, 0.000, 75, 803, Clouds, broken clouds  
2017, January, 3, Tuesday, 10, Active, 20, NCR, Bispenegade, Bispenegade, 35, 9800, 57.453, 9.996, DKK, Mastercard - on-us, 6798, Withdrawal, , , 57.464, 9.982, 2620214, Hjørring, 2  
77.150, 1003, 86, 4, 260, 0.000, 75, 803, Clouds, broken clouds  
2017, January, 3, Tuesday, 10, Active, 91, NCR, Skive Lobby, Adelgade, 8, 7800, 56.567, 9.027, DKK, HÅfÅ!vekort - on-us, 6263, Withdrawal, , , 56.567, 9.033, 2613731, Skive, 279.150  
1, 1006, 93, 4, 250, 0.000, 90, 310, Drizzle, light intensity drizzle rain  
2017, January, 3, Tuesday, 10, Active, 35, NCR, Aabybro, ÅfEastergade, 6,9440, 57.162, 9.730, DKK, Mastercard - on-us, 920, Withdrawal, , , 57.158, 9.737, 2625037, Aabybro, 277.150,  
1003, 86, 4, 260, 0.000, 75, 803, Clouds, broken clouds  
2017, January, 3, Tuesday, 10, Active, 87, NCR, Aalborg Storcenter indg. D, Hobrovej, 452, 9200, 57.005, 9.876, DKK, VISA, 5810, Withdrawal, , , 56.972, 9.848, 2612021, Svenstrup, 2  
77.150, 1003, 86, 4, 260, 0.000, 75, 803, Clouds, broken clouds  
2017, January, 3, Tuesday, 10, Active, 57, NCR, HillerÅfÅ d, KÅfÅ benhavnsvej, 31, 3400, 55.933, 12.314, DKK, MasterCard, 6289, Withdrawal, , , 55.767, 12.433, 2621394, Gladsaxe K  
ommune, 276.150, 1008, 100, 6, 240, 0.000, 75, 300, Drizzle, light intensity drizzle  
2017, January, 3, Tuesday, 10, Active, 71, NCR, AalbÅfÅ!k, Centralvej, 5, 9982, 57.593, 10.412, DKK, Visa Dankort, 4017, Withdrawal, , , 57.492, 10.496, 2612274, Strandby, 277.150, 1  
003, 86, 4, 260, 0.000, 75, 803, Clouds, broken clouds  
2017, January, 3, Tuesday, 10, Active, 79, NCR, Middelfart, Brogade, 9, 5500, 55.507, 9.727, DKK, MasterCard, 1219, Withdrawal, , , 55.506, 9.731, 2616933, Middelfart, 278.540, 1012,  
93, 7, 260, 0.000, 75, 310, Drizzle, light intensity drizzle rain  
2017, January, 3, Tuesday, 10, Active, 59, Diebold Nixdorf, MylÅfÅ bing Mors Lobby, Kirketorvet, 1, 7900, 56.795, 8.860, DKK, Dankort - on-us, 4043, Withdrawal, , , 56.793, 8.853  
1, 2615964, Nykobing Mors, 279.150, 1006, 93, 4, 250, 0.000, 90, 310, Drizzle, light intensity drizzle rain  
2017, January, 3, Tuesday, 10, Inactive, 88, NCR, Storcenter indg. A, Hobrovej, 452, 9200, 57.005, 9.876, DKK, Visa Dankort, 6559, Withdrawal, , , 56.972, 9.848, 2612021, Svenstrup  
1, 277.150, 1003, 86, 4, 260, 0.000, 75, 803, Clouds, broken clouds  
2017, January, 3, Tuesday, 10, Active, 92, NCR, Intern HJÅfÅ rring, ÅfEastergade, 8, 9800, 57.459, 9.988, DKK, Visa Dankort - on-us, 667, Withdrawal, , , 57.464, 9.982, 2620214, H  
jørring, 277.150, 1003, 86, 4, 260, 0.000, 75, 803, Clouds, broken clouds  
2017, January, 3, Tuesday, 10, Active, 45, NCR, Abildgaard, HJÅfÅ rringvej, 144, 9900, 57.447, 10.506, DKK, Visa Dankort - on-us, 3658, Withdrawal, , , 57.441, 10.537, 2621927, Fre  
derikshavn, 277.150, 1003, 86, 4, 260, 0.000, 75, 803, Clouds, broken clouds  
2017, January, 3, Tuesday, 10, Active, 60, NCR, Hirtshale, JÅfÅ rgen Fibigersgade, 4, 9850, 57.591, 9.957, DKK, Visa Dankort, 6286, Withdrawal, , , 57.588, 9.959, 2620279, Hirtshal  
s, 277.150, 1003, 86, 4, 260, 0.000, 75, 803, Clouds, broken clouds  
2017, January, 3, Tuesday, 10, Active, 35, NCR, Aabybro, ÅfEastergade, 6,9440, 57.162, 9.730, DKK, Mastercard - on-us, 458, Withdrawal, , , 57.158, 9.737, 2625037, Aabybro, 277.150,
```