

DATA INGESTION USING SQOOP

1. In the given screenshot the following code is executed :
sqoop import \
--connect jdbc:mysql://upgraddetest.cyaieic9bmnf.us-east-1.rds.amazonaws.com/testdatabase\
--table SRC_ATM_TRANS \
--username student --password STUDENT123 \
--target-dir /user/root/spar_nord_bank_atm \
-m 1

```
root@ip-172-31-4-225:~#
E::::E      EEEEE M::::::::::M      M::::::::::M RR::::R      R::::R
E::::E      M::::M:M::M      M::M::::M      R::::R      R::::R
E:::::EEEEEEEEEE M::::M M::M M::M M::M M::M M::M R::RRRRRR::::R
E:::::EEEEEEEEEE M::::M M::M:M::M M::::M      R::::RRRRRR::::R
E:::::EEEEEEEEEE M::::M M::M:M::M M::::M      R::RRRRRR::::R
E::::E      M::::M M::M      M::::M      R::::R      R::::R
E::::E      EEEEE M::::M      MMM      M::::M      R::::R      R::::R
EE:::::EEEEEEEEEE M::::M      M::::M      R::::R      R::::R
E:::::EEEEEEEEEE M::::M      M::::M      R::::R      R::::R
EEEEEEEEEEEEEEEEEE MMMMMM      MMMMMM RRRRRRR      RRRRRR

[hadoop@ip-172-31-4-225 ~]$ sudo -i

EEEEEEEEEEEEEEEEEE MMMMMMMM      MMMMMMMM RRRRRRRRRRRRRRRR
E:::::EEEEEEEEEE M::::M      M::::M      R::::RRRRRR::::R
EE:::::EEEEEEEEEE M::::M      M::::M      R::RRRRRR::::R
E::::E      EEEEE M::::M      M::::M      RR::::R      R::::R
E::::E      M::::M:M::M      M::M::::M      R::::R      R::::R
E:::::EEEEEEEEEE M::::M M::M M::M M::M M::M M::M R::RRRRRR::::R
E:::::EEEEEEEEEE M::::M M::M:M::M M::::M      R::::RRRRRR::::R
E:::::EEEEEEEEEE M::::M M::M:M::M M::::M      R::RRRRRR::::R
E::::E      M::::M M::M      M::::M      R::::R      R::::R
E::::E      EEEEE M::::M      MMM      M::::M      R::::R      R::::R
EE:::::EEEEEEEEEE M::::M      M::::M      R::::R      R::::R
E:::::EEEEEEEEEE M::::M      M::::M      R::::R      R::::R
EEEEEEEEEEEEEEEEEE MMMMMMMM      MMMMMMMM RRRRRRR      RRRRRR

[root@ip-172-31-4-225 ~]# jdbc:mysql://upgraddetest.cyaieic9bmnf.us-east-1.rds.amazonaws.com/testdatabase
-bash: jdbc:mysql://upgraddetest.cyaieic9bmnf.us-east-1.rds.amazonaws.com/testdatabase: No such file or directory
[root@ip-172-31-4-225 ~]# sqoop import \
> --connect jdbc:mysql://upgraddetest.cyaieic9bmnf.us-east-1.rds.amazonaws.com/testdatabase \
> --table SRC_ATM_TRANS \
> --username student --password STUDENT123 \
> --target-dir /user/root/spar_nord_bank_atm \
> -m 1
Warning: /usr/lib/sqoop/./accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
22/09/11 21:40:10 INFO sqoop.Sqoop: Running Sqoop version: 1.4.7
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/lib/hadoop/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/share/aws/redshift/jdbc/redshift-jdbc42-1.2.37.1061.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/lib/hive/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
```

DATA INGESTION USING SQOOP

2. Specifications used by the system for running the data ingestion job:

```
22/09/11 21:40:59 INFO mapreduce.job: Counters: 39
File System Counters
  FILE: Number of bytes read=0
  FILE: Number of bytes written=189605
  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)=1166400
  Total time spent by all reduces in occupied slots (ms)=0
  Total time spent by all map tasks (ms)=24300
  Total vcore-milliseconds taken by all map tasks=24300
  Total megabyte-milliseconds taken by all map tasks=37324800
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)=276
  CPU time spent (ms)=27910
  Physical memory (bytes) snapshot=618344448
  Virtual memory (bytes) snapshot=3298091008
  Total committed heap usage (bytes)=505413632
File Input Format Counters
  Bytes Read=0
File Output Format Counters
  Bytes Written=531214815
```

3. AS per the source we imported 2468572 records

```
File Input Format Counters
  Bytes Read=0
File Output Format Counters
  Bytes Written=531214815
22/09/11 21:40:59 INFO mapreduce.ImportJobBase: Transferred 506.6059 MB in 41.6895 seconds (12.1519 MB/sec)
22/09/11 21:40:59 INFO mapreduce.ImportJobBase: Retrieved 2468572 records.
```

4. On using `hadoop fs -ls /user/root/spar_nord_bank_atm`

```
[root@ip-172-31-4-225 ~]# hadoop fs -ls
Found 1 items
drwxr-xr-x  - root hadoop          0 2022-09-11 21:40 spar_nord_bank_atm
[root@ip-172-31-4-225 ~]# hadoop fs -ls /user/root/spar_nord_bank_atm
Found 2 items
-rw-r--r--  1 root hadoop          0 2022-09-11 21:40 /user/root/spar_nord_bank_atm/ SUCCESS
-rw-r--r--  1 root hadoop 531214815 2022-09-11 21:40 /user/root/spar_nord_bank_atm/part-m-00000
```

5. Screenshot showing the imported data

DATA INGESTION USING SQOOP

```
17,February,23,Thursday,19,Inactive,88,MCR,Storcenter indp. A,Hobrovej,452,9200,57.005,9.876,DKK,VISA,4776,Withdrawal,,,56.972,9.848,2612021,Svenstrup,274.704,994,
66,0.000,92,804,Clouds,overcast clouds
17,February,23,Thursday,19,Active,39,MCR,Svenstrup,Godthåfsvej,14,9230,56.973,9.851,DKK,Visa Dankort,7290,Withdrawal,,,56.972,9.848,2612021,Svenstrup,274.704,99
8,66,0.000,92,804,Clouds,overcast clouds
17,February,23,Thursday,19,Inactive,52,MCR,Farsøvej,8,9640,56.771,9.340,DKK,Mastercard - on-us,2120,Withdrawal,,,56.773,9.339,2622310,Farsø,274.704,994,88,8
.000,92,804,Clouds,overcast clouds
17,February,23,Thursday,19,Active,64,MCR,Horsens,Gråvej,5,8700,55.859,9.854,DKK,Mastercard - on-us,8729,Withdrawal,,,55.861,9.850,2619771,Horsens,273.830,
00,6.70,0.000,75,600,Snow,light snow
17,February,23,Thursday,19,Active,19,MCR,Næsby,Torvet,6,9400,57.059,9.822,DKK,Mastercard - on-us,7287,Withdrawal,,,57.048,9.919,2624886,Aalborg,274.704,99
8,66,0.000,92,804,Clouds,overcast clouds
17,February,23,Thursday,19,Inactive,52,MCR,Farsøvej,8,9640,56.771,9.340,DKK,VISA,249,Withdrawal,,,56.773,9.339,2622310,Farsø,274.704,994,88,8,66,0.000,92,80
4,overcast clouds
17,February,23,Thursday,19,Active,19,MCR,Næsby,Torvet,6,9400,57.059,9.822,DKK,Mastercard - on-us,8383,Withdrawal,,,57.048,9.919,2624886,Aalborg,274.704,99
8,66,0.000,92,804,Clouds,overcast clouds
17,February,23,Thursday,19,Active,40,Diebold Nixdorf,Børneshøj,Algade,4,9700,57.269,9.945,DKK,Mastercard - on-us,6167,Withdrawal,,,57.118,9.857,2610823,Vadum,
24,994,88,8,66,0.000,92,804,Clouds,overcast clouds
17,February,23,Thursday,19,Active,33,MCR,Vadum,Ellehøjvej,43,9430,57.118,9.861,DKK,Visa Dankort - on-us,2685,Withdrawal,,,57.118,9.857,2610823,Vadum,274.704,99
8,66,0.000,92,804,Clouds,overcast clouds
17,February,23,Thursday,19,Active,34,MCR,Skipperen,Vestre Alle,2,9000,57.034,9.908,DKK,Næstvedkort - on-us,4926,Withdrawal,,,57.048,9.919,2624886,Aalborg,274.704,9
8,66,0.000,92,804,Clouds,overcast clouds
17,February,23,Thursday,19,Active,63,MCR,Brugsen i Breum,Åkvej,1,7870,56.608,9.069,DKK,Visa Dankort,6576,Withdrawal,,,56.586,9.136,2619908,Højlev,274.704,
9,8,66,0.000,92,804,Clouds,overcast clouds
17,February,23,Thursday,19,Active,34,MCR,Skipperen,Vestre Alle,2,9000,57.034,9.908,DKK,Næstvedkort - on-us,8103,Withdrawal,,,57.048,9.919,2624886,Aalborg,274.704,9
8,66,0.000,92,804,Clouds,overcast clouds
17,February,23,Thursday,19,Inactive,102,MCR,Aalborg Storcenter Afd,Hobrovej,452,9200,57.005,9.876,EUR,Visa Dankort - on-us,785,Withdrawal,,,56.972,9.848,2612021,S
venstrup,274.704,994,88,8,66,0.000,92,804,Clouds,overcast clouds
17,February,23,Thursday,19,Active,34,MCR,Skipperen,Vestre Alle,2,9000,57.034,9.908,DKK,Næstvedkort - on-us,9617,Withdrawal,,,57.048,9.919,2624886,Aalborg,274.704,9
8,66,0.000,92,804,Clouds,overcast clouds
17,February,23,Thursday,19,Active,13,MCR,Søby,Vestergade,3,9300,57.334,10.515,DKK,Visa Dankort - on-us,5730,Withdrawal,,,57.441,10.537,2621927,Frederikshavn,275
997,99,9,59,0.000,92,804,Clouds,overcast clouds
17,February,23,Thursday,19,Active,75,MCR,Nordkraft,Kjellerups Torv,1,9000,57.047,9.932,DKK,CIRKUS,9800,Withdrawal,,,57.048,9.935,2616235,Næsby,274.704,994
8,66,0.000,92,804,Clouds,overcast clouds
17,February,23,Thursday,19,Inactive,52,MCR,Farsøvej,8,9640,56.771,9.340,DKK,Visa Dankort - on-us,1810,Withdrawal,,,56.773,9.339,2622310,Farsø,274.704,994,88
8,000,92,804,Clouds,overcast clouds
```