

Data Ingestion from the RDS to HDFS using Sqoop

Sqoop Import command used for importing table from RDS to HDFS:

Made sure root directory has no folders with the name SRC_ATM_TRANS before running the Sqoop command. Otherwise, the Sqoop command will fail.

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

```
[root@ip-172-31-47-28 ~]# hadoop fs -ls /user/root/SRC_ATM_TRANS
ls: `/user/root/SRC_ATM_TRANS': No such file or directory
[root@ip-172-31-47-28 ~]#
```

Extra steps to download mysql-connector-java-8.0.30.jar

Ran below command to download:

```
wget https://repo1.maven.org/maven2/mysql/mysql-connector-java/8.0.30/mysql-connector-java-8.0.30.jar
```

Ran below command to copy to sqoop library:

```
mv mysql-connector-java/8.0.30/mysql-connector-java-8.0.30.jar /usr/lib/sqoop/lib/
```

Ran below sqoop command to import the data from RDS table SRC_ATM_TRANS to hdfs folder /user/root/SRC_ATM_TRANS

```
sqoop import \  
--connect jdbc:mysql://upgradtest.cyaieic9bmnf.us-east-1.rds.amazonaws.com/testdatabase \  
--username student \  
--password STUDENT123 \  
--table SRC_ATM_TRANS \  
--target-dir /user/root/SRC_ATM_TRANS \  
--driver com.mysql.cj.jdbc.Driver \  
--m 1
```

Screenshot of result of Sqoop import job. **2468572** records have been imported.

```

2024-11-06 05:10:32,320 INFO mapreduce.JobSubmitter: Number of splits:1
2024-11-06 05:10:32,322 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1730868995268_0001
2024-11-06 05:10:32,322 INFO mapreduce.JobSubmitter: Executing with tokens: []
2024-11-06 05:10:32,493 INFO conf.Configuration: resource-types.xml not found
2024-11-06 05:10:32,494 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2024-11-06 05:10:32,878 INFO impl.YarnClientImpl: Submitted application application_1730868995268_0001
2024-11-06 05:10:32,936 INFO mapreduce.Job: The url to track the job: http://ip-172-31-47-28.ec2.internal:20888/proxy/application_1730868995268_0001/
2024-11-06 05:10:32,936 INFO mapreduce.Job: Running job: job_1730868995268_0001
2024-11-06 05:10:41,132 INFO mapreduce.Job: Job job_1730868995268_0001 running in uber mode : false
2024-11-06 05:10:41,134 INFO mapreduce.Job: map 0% reduce 0%
2024-11-06 05:11:11,350 INFO mapreduce.Job: map 100% reduce 0%
2024-11-06 05:11:11,356 INFO mapreduce.Job: Job job_1730868995268_0001 completed successfully
2024-11-06 05:11:11,434 INFO mapreduce.Job: Counters: 33
    File System Counters
        FILE: Number of bytes read=0
        FILE: Number of bytes written=299241
        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)=1321488
        Total time spent by all reduces in occupied slots (ms)=0
        Total time spent by all map tasks (ms)=27531
        Total vcore-milliseconds taken by all map tasks=27531
        Total megabyte-milliseconds taken by all map tasks=42287616
    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)=258
        CPU time spent (ms)=23320
        Physical memory (bytes) snapshot=429944832
        Virtual memory (bytes) snapshot=3195990016
        Total committed heap usage (bytes)=306184192
        Peak Map Physical memory (bytes)=429944832
        Peak Map Virtual memory (bytes)=3195990016
    File Input Format Counters
        Bytes Read=0
    File Output Format Counters
        Bytes Written=531214815
2024-11-06 05:11:11,442 INFO mapreduce.ImportJobBase: Transferred 506.6059 MB in 47.2143 seconds (10.7299 MB/sec)
2024-11-06 05:11:11,447 INFO mapreduce.ImportJobBase: Retrieved 2468572 records.
[root@ip-172-31-47-28 ~]#

```

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

hadoop fs -ls /user/root/SRC_ATM_TRANS

Screenshot of target directory which contains part-m-00000 file with imported data from RDS.

```
[root@ip-172-31-47-28 ~]# hadoop fs -ls /user/root/SRC_ATM_TRANS
Found 2 items
-rw-r--r-- 1 root hdfsadmingroup 0 2024-11-06 05:11 /user/root/SRC_ATM_TRANS/_SUCCESS
-rw-r--r-- 1 root hdfsadmingroup 531214815 2024-11-06 05:11 /user/root/SRC_ATM_TRANS/part-m-00000
[root@ip-172-31-47-28 ~]#
```

Screenshot of the imported data:

Command to see top 100 records from the imported file.

hadoop fs -cat /user/root/SRC_ATM_TRANS/part-m-00000 | head -n 100

Screenshot of imported data in hdfs:

```
[root@ip-172-31-47-28 ~]# hadoop fs -cat /user/root/SRC_ATM_TRANS/part-m-00000 | head -n 100
2017,January,1,Sunday,0,Active,1,NCR,NÅFÅstved,Farimagvej,8,4700,55.233,11.763,DKK,MasterCard,5643,Withdrawal,,,55.230,11.761,2616038,Maestved,281.150,1014,87,7,260,0.215,92,500,Rain,light rain
2017,January,1,Sunday,0,Inactive,2,NCR,Vejgaard,Hadsundvej,20,9000,57.043,9.950,DKK,MasterCard,1764,Withdrawal,,,57.048,9.935,2616235,NÅFÅ,rrsundby,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Inactive,2,NCR,Vejgaard,Hadsundvej,20,9000,57.043,9.950,DKK,VISA,1591,Withdrawal,,,57.048,9.935,2616235,NÅFÅ,rrsundby,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Inactive,3,NCR,Ikast,FÅFÅrhusvej,12,7400,56.139,9.154,DKK,VISA,1466,Withdrawal,,,56.139,9.152,2619426,Ikast,281.150,1011,100,6,240,0.000,75,300,Drizzle,light intensity drizzle
2017,January,1,Sunday,0,Active,4,NCR,Svejestev,NÅFÅ,asseger,1,4000,55.631,12.018,DKK,MasterCard,5153,Withdrawal,,,55.642,12.080,2614481,Roskilde,280.610,1014,87,7,260,0.000,88,701,Mist,mist
2017,January,1,Sunday,0,Active,5,NCR,Nibe,Torvet,1,9240,56.863,9.639,DKK,MasterCard,3263,Withdrawal,,,56.981,9.639,2616483,Nibe,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,6,NCR,Federicia,SjÅFÅllandsgade,33,7000,55.564,9.757,DKK,MasterCard,887,Withdrawal,,,55.566,9.753,2621951,Federicia,281.150,1014,93,7,230,0.290,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,7,Diebold Nixdorf,Hjallerup,Hjallerup Centret,18,9320,57.168,10.148,DKK,MasterCard - on-us,4626,Withdrawal,,,57.165,10.146,2620275,Hjallerup,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,8,NCR,GlyngÅFÅ, re,FÅFÅrgevej,1,7870,56.762,8.867,DKK,MasterCard,470,Withdrawal,,,56.793,8.853,2615964,Nykobing Mors,281.150,1011,100,6,240,0.000,75,300,Drizzle,light intensity drizzle
2017,January,1,Sunday,0,Active,9,Diebold Nixdorf,Hadsund,Storegade,12,9560,56.716,10.114,DKK,VISA,8473,Withdrawal,,,56.715,10.117,2620952,Hadsund,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,10,NCR,NÅFÅ, rrsundby,Torvet,6,9400,57.059,9.922,DKK,Dankort,953,Withdrawal,,,57.048,9.919,2624886,Aalborg,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Inactive,12,NCR,Svejestev,Fridtjof Mansens Vej,2,9210,57.023,9.940,DKK,Visa Dankort,9346,Withdrawal,,,57.048,9.935,2616235,NÅFÅ, rrsundby,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Inactive,2,NCR,Vejgaard,Hadsundvej,20,9000,57.043,9.950,DKK,MasterCard - on-us,3874,Withdrawal,,,57.048,9.935,2616235,NÅFÅ, rrsundby,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Inactive,12,NCR,ÅFÅesterÅFÅV Duns,ÅFÅesterÅFÅV,12,9000,57.049,9.922,DKK,MasterCard - on-us,1329,Withdrawal,,,57.048,9.919,2624886,Aalborg,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,13,NCR,SÅFÅby,Vestergade,3,9300,57.334,10.515,DKK,MasterCard - on-us,5024,Withdrawal,,,57.441,10.537,2621527,Frederikshavn,281.140,1015,94,12,251,1.275,92,500,Rain,light rain
2017,January,1,Sunday,0,Inactive,14,NCR,NÅFÅ, rning,NÅFÅ, rrealle,12,8362,56.086,10.037,DKK,Visa Dankort - on-us,1133,Withdrawal,,,56.157,10.211,2624652,Århus,281.150,1012,87,5,250,0.000,92,300,Drizzle,light intensity drizzle
2017,January,1,Sunday,0,Active,15,NCR,Vestre,Mastervej,36,9000,57.053,9.905,DKK,MasterCard,584,Withdrawal,,,57.048,9.919,2624886,Aalborg,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,12,NCR,ÅFÅesterÅFÅV Duns,ÅFÅesterÅFÅV,12,9000,57.049,9.922,DKK,MasterCard - on-us,9570,Withdrawal,,,57.048,9.919,2624886,Aalborg,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Inactive,16,NCR,Skive,Adelgade,6,7800,56.567,9.037,DKK,Visa Dankort,6068,Withdrawal,,,56.567,9.033,2613731,Skive,281.150,1012,100,7,240,0.000,75,300,Drizzle,light intensity drizzle
2017,January,1,Sunday,0,Active,17,NCR,Banders,ÅFÅesterovold,16,8900,56.462,10.038,DKK,VISA,1630,Withdrawal,,,56.715,10.117,2620952,Hadsund,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,17,NCR,Banders,ÅFÅesterovold,16,8900,56.462,10.038,DKK,MasterCard,9948,Withdrawal,,,56.715,10.117,2620952,Hadsund,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,6,NCR,Federicia,SjÅFÅllandsgade,33,7000,55.564,9.757,DKK,MasterCard,4849,Withdrawal,,,55.566,9.753,2621951,Federicia,281.150,1014,93,7,230,0.290,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,10,NCR,NÅFÅ, rrsundby,Torvet,6,9400,57.059,9.922,DKK,MasterCard,4400,Withdrawal,,,57.048,9.919,2624886,Aalborg,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,6,NCR,Federicia,SjÅFÅllandsgade,33,7000,55.564,9.757,DKK,VISA,7452,Withdrawal,,,55.566,9.753,2621951,Federicia,281.150,1014,93,7,230,0.290,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,18,Diebold Nixdorf,Viborg,Toldboden,3,8800,56.448,9.401,DKK,VISA,4062,Withdrawal,,,56.453,9.402,2610319,Viborg,281.150,1012,100,7,240,0.000,75,300,Drizzle,light intensity drizzle
2017,January,1,Sunday,0,Active,19,NCR,Kolding,Vejlevej,135,6000,55.505,9.457,DKK,MasterCard,7951,Withdrawal,,,55.490,9.472,2618528,Kolding,281.150,1015,93,8,240,0.250,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,20,NCR,Bispensgade,Bispensgade,35,9800,57.453,9.996,DKK,MasterCard,7573,Withdrawal,,,57.464,9.982,2620214,Hjorring,281.265,1018,96,11,256,1.260,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,8,NCR,GlyngÅFÅ, re,FÅFÅrgevej,1,7870,56.762,8.867,DKK,Visa Dankort - on-us,4009,Withdrawal,,,56.793,8.853,2615964,Nykobing Mors,281.150,1011,100,6,240,0.000,75,300,Drizzle,light intensity drizzle
2017,January,1,Sunday,0,Active,21,NCR,ÅFÅesterÅFÅV NÅFÅ, ller,ÅFÅesterÅFÅV,12,9000,57.049,9.922,DKK,Visa Dankort,7327,Withdrawal,,,57.048,9.919,2624886,Aalborg,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,21,NCR,ÅFÅesterÅFÅV NÅFÅ, ller,ÅFÅesterÅFÅV,12,9000,57.049,9.922,DKK,MasterCard - on-us,4607,Withdrawal,,,57.048,9.919,2624886,Aalborg,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,22,NCR,Roskilde,FÅFÅ, bebmhvnvej,65,4000,55.642,12.106,DKK,Visa Dankort,1053,Withdrawal,,,55.642,12.080,2614481,Roskilde,280.610,1014,87,7,260,0.000,88,701,Mist,mist
2017,January,1,Sunday,0,Inactive,16,NCR,Skive,Adelgade,6,7800,56.567,9.037,DKK,Visa Dankort,1445,Withdrawal,,,56.567,9.033,2613731,Skive,281.150,1012,100,7,240,0.000,75,300,Drizzle,light intensity drizzle
2017,January,1,Sunday,0,Active,6,NCR,Federicia,SjÅFÅllandsgade,33,7000,55.564,9.757,DKK,MasterCard,4063,Withdrawal,,,55.566,9.753,2621951,Federicia,281.150,1014,93,7,230,0.290,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,23,Diebold Nixdorf,Vodskov,Vodskovvej,27,9310,57.104,10.027,DKK,Visa Dankort - on-us,5993,Withdrawal,,,57.109,10.022,2610095,Vodskov,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,24,NCR,Hobro,Adelgade,31,9500,56.638,9.794,DKK,MasterCard - on-us,7768,Withdrawal,,,56.643,9.790,2620167,Hobro,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,15,NCR,Vestre,Mastervej,36,9000,57.053,9.905,DKK,Visa Dankort,862,Withdrawal,,,57.048,9.919,2624886,Aalborg,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,25,Diebold Nixdorf,Odense,FÅFÅ, lldvej,3,5000,55.394,10.370,DKK,VISA,1005,Withdrawal,,,55.396,10.389,261876,Odense,281.150,1016,87,5,240,0.190,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,21,NCR,ÅFÅesterÅFÅV NÅFÅ, ller,ÅFÅesterÅFÅV,12,9000,57.049,9.922,DKK,MasterCard - on-us,2438,Withdrawal,,,57.048,9.919,2624886,Aalborg,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,26,NCR,HolbÅFÅk,Sløssvolden,7,4300,55.718,11.704,DKK,Visa Dankort - on-us,9175,Withdrawal,,,55.700,11.700,2620141,HolbÅFÅk Kommune,280.180,1014,93,7,230,0.000,75,701,Mist,mist
2017,January,1,Sunday,0,Active,9,Diebold Nixdorf,Hadsund,Storegade,12,9560,56.716,10.114,DKK,VISA,8562,Withdrawal,,,56.715,10.117,2620952,Hadsund,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,0,Active,27,NCR,Herning,Dalsgade,30,7400,56.135,8.971,DKK,MasterCard,5283,Withdrawal,,,56.138,8.974,2620425,Herning,281.150,1011,100,7,240,0.000,75,300,Drizzle,light intensity drizzle
2017,January,1,Sunday,1,Inactive,2,NCR,Vejgaard,Hadsundvej,20,9000,57.043,9.950,DKK,MasterCard,730,Withdrawal,,,57.048,9.935,2616235,NÅFÅ, rrsundby,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,1,Active,22,NCR,Roskilde,FÅFÅ, bebmhvnvej,65,4000,55.642,12.106,DKK,VISA,7795,Withdrawal,,,55.642,12.080,2614481,Roskilde,279.620,1013,93,6,250,0.000,88,701,Mist,mist
2017,January,1,Sunday,1,Active,26,NCR,Federicia,SjÅFÅllandsgade,33,7000,55.564,9.757,DKK,VISA,7087,Withdrawal,,,55.566,9.753,2621951,Federicia,281.150,1013,93,7,230,0.290,92,500,Rain,light rain
2017,January,1,Sunday,1,Active,25,NCR,Roskilde,FÅFÅ, bebmhvnvej,65,4000,55.642,12.106,DKK,VISA,641,Withdrawal,,,55.642,12.080,2614481,Roskilde,279.620,1013,93,6,250,0.000,88,701,Mist,mist
2017,January,1,Sunday,1,Active,24,NCR,Hobro,Adelgade,31,9500,56.638,9.794,DKK,MasterCard - on-us,2780,Withdrawal,,,56.643,9.790,2620167,Hobro,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,1,Inactive,2,NCR,NÅFÅ, rsted,Farimagvej,8,4700,55.233,11.763,DKK,VISA,1975,Withdrawal,,,55.230,11.761,2616038,Maestved,281.150,1014,87,7,250,0.215,92,500,Rain,light rain
2017,January,1,Sunday,1,Inactive,12,NCR,ÅFÅesterÅFÅV Duns,ÅFÅesterÅFÅV,12,9000,57.049,9.922,DKK,MasterCard - on-us,1555,Withdrawal,,,57.048,9.919,2624886,Aalborg,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,1,Active,29,NCR,Skelagervej,15,Skelagervej,15,9000,57.023,9.891,DKK,Visa Dankort - on-us,1837,Withdrawal,,,57.048,9.919,2624886,Aalborg,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,1,Inactive,30,NCR,NyÅFÅ, bing Mors,Kirkevej,1,7900,56.795,8.860,DKK,VISA,4519,Withdrawal,,,56.793,8.853,2615964,Nykobing Mors,281.150,1011,100,6,240,0.000,75,310,Drizzle,light intensity drizzle rain
2017,January,1,Sunday,1,Active,31,NCR,Slagelse,Mariendals Alle,29,4200,55.398,11.342,DKK,VISA,7087,Withdrawal,,,55.403,11.355,2613460,Slagelse,280.150,1013,93,6,250,0.205,92,500,Rain,light rain
2017,January,1,Sunday,1,Active,32,NCR,SjÅFÅdding,Tarpvejvej,59,6710,55.498,8.408,DKK,MasterCard,1876,Withdrawal,,,55.467,8.450,2622447,Ålborg,281.150,1014,93,8,240,0.220,92,500,Rain,light rain
2017,January,1,Sunday,1,Active,3,NCR,Vadum,Ellehømsvej,43,5430,57.118,9.861,DKK,MasterCard - on-us,8120,Withdrawal,,,57.118,9.857,2610823,Vadum,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,1,Active,28,NCR,LÅFÅ, gsfÅFÅ, e,ÅFÅesterbrogade,9,9670,56.967,9.253,DKK,MasterCard - on-us,4972,Withdrawal,,,56.962,9.258,2617467,Logtor,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
2017,January,1,Sunday,1,Active,6,NCR,Federicia,SjÅFÅllandsgade,33,7000,55.564,9.757,DKK,Visa Dankort,500,Withdrawal,,,55.566,9.753,2621951,Federicia,281.150,1013,93,7,230,0.290,92,500,Rain,light rain
2017,January,1,Sunday,1,Active,1,NCR,NÅFÅ, rsted,Farimagvej,8,4700,55.233,11.763,DKK,MasterCard,7583,Withdrawal,,,55.230,11.761,2616038,Maestved,281.150,1014,87,7,250,0.215,92,500,Rain,light rain
2017,January,1,Sunday,1,Active,6,NCR,Federicia,SjÅFÅllandsgade,33,7000,55.564,9.757,DKK,Visa Dankort,616,Withdrawal,,,55.566,9.753,2621951,Federicia,281.150,1013,93,7,230,0.290,92,500,Rain,light rain
2017,January,1,Sunday,1,Inactive,30,NCR,NyÅFÅ, bing Mors,Kirkevej,1,7900,56.795,8.860,DKK,MasterCard,931,Withdrawal,,,56.793,8.853,2615964,Nykobing Mors,281.150,1011,100,6,240,0.000,75,310,Drizzle,light intensity drizzle rain
2017,January,1,Sunday,1,Active,34,NCR,Skippersvej,Vestre Alle,2,9000,57.034,9.908,DKK,Visa Dankort,7401,Withdrawal,,,57.048,9.919,2624886,Aalborg,280.640,1020,93,9,250,0.590,92,500,Rain,light rain
```