



## **ASSIGNMENT-3**

**(SQOOP)**

Employee ID : 2320074

CSDAIA24AZ003

```
Ubuntu (Snapshot 2) [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Activities Terminal
Fri 23:08
ubh01@ubh01: ~

File Edit View Search Terminal Help
ubh01@ubh01:~$ ./hadoop-2.7.1/sbin/start-all.sh
This script is deprecated. Instead use start-dfs.sh and start-yarn.sh
24/03/29 23:07:14 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes wh
ere applicable
Starting namenodes on [localhost]
localhost: starting namenode, logging to /home/ubh01/hadoop-2.7.1/logs/hadoop-ubh01-namenode-ubh01.out
localhost: starting datanode, logging to /home/ubh01/hadoop-2.7.1/logs/hadoop-ubh01-datanode-ubh01.out
Starting secondary namenodes [0.0.0.0]
0.0.0.0: starting secondarynamenode, logging to /home/ubh01/hadoop-2.7.1/logs/hadoop-ubh01-secondarynamenode-ubh01.out
24/03/29 23:07:31 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes wh
ere applicable
starting yarn daemons
starting resourcemanager, logging to /home/ubh01/hadoop-2.7.1/logs/yarn-ubh01-resourcemanager-ubh01.out
localhost: starting nodemanager, logging to /home/ubh01/hadoop-2.7.1/logs/yarn-ubh01-nodemanager-ubh01.out
ubh01@ubh01:~$ mysql -u root -p
Enter password:
Welcome to the MySQL monitor. Commands end with ; or \g.
Your MySQL connection id is 4
Server version: 5.7.42-0ubuntu0.18.04.1 (Ubuntu)

Copyright (c) 2000, 2023, Oracle and/or its affiliates.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql>
```

```
Ubuntu (Snapshot 2) [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Activities Terminal
Fri 23:10
ubh01@ubh01: ~

File Edit View Search Terminal Help
ubh01@ubh01:~$ mysql -u root -p
Enter password:
Welcome to the MySQL monitor. Commands end with ; or \g.
Your MySQL connection id is 4
Server version: 5.7.42-0ubuntu0.18.04.1 (Ubuntu)

Copyright (c) 2000, 2023, Oracle and/or its affiliates.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> show databases;
+-----+
| Database |
+-----+
| information_schema |
| metastore |
| metastore_db |
| mysql |
| performance_schema |
| sumit |
| sys |
+-----+
7 rows in set (0.01 sec)

mysql>
```

Ubuntu (Snapshot 2) [Running] - Oracle VM VirtualBox

File Machine View Input Devices Help

Activities Fri 23:11 ubh01@ubh01: ~

File Edit View Search Terminal Help

Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

```
mysql> show databases;
+-----+
| Database |
+-----+
| information_schema |
| metastore |
| metastore_db |
| mysql |
| performance_schema |
| sumit |
| sys |
+-----+
7 rows in set (0.01 sec)
```

```
mysql> create database cognizant;
Query OK, 1 row affected (0.00 sec)
```

```
mysql> use cognizant;
Database changed
mysql> show tables;
Empty set (0.01 sec)
```

```
mysql>
```

Type here to search

ENG 23:11 29-03-2024

Ubuntu (Snapshot 2) [Running] - Oracle VM VirtualBox

File Machine View Input Devices Help

Activities Fri 23:12 ubh01@ubh01: ~

File Edit View Search Terminal Help

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

```
mysql> show databases;
+-----+
| Database |
+-----+
| information_schema |
| metastore |
| metastore_db |
| mysql |
| performance_schema |
| sumit |
| sys |
+-----+
7 rows in set (0.01 sec)
```

```
mysql> create database cognizant;
Query OK, 1 row affected (0.00 sec)
```

```
mysql> use cognizant;
Database changed
mysql> show tables;
Empty set (0.01 sec)
```

```
mysql> create table inportttt(uid int,mid int,rating int,ts int);
Query OK, 0 rows affected (0.05 sec)
```

```
mysql>
```

Type here to search

ENG 23:12 29-03-2024

```
Ubuntu (Snapshot 2) [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Activities Terminal
Fri 23:17
ubh01@ubh01: ~

mysql> quit;
Bye
ubh01@ubh01:~$ wget https://raw.githubusercontent.com/deeksharm/DP203/main/ratings.data
--2024-03-29 23:16:25-- https://raw.githubusercontent.com/deeksharm/DP203/main/ratings.data
Resolving raw.githubusercontent.com (raw.githubusercontent.com)... 185.199.111.133, 185.199.110.133, 185.199.109.133, ...
Connecting to raw.githubusercontent.com (raw.githubusercontent.com)|185.199.111.133|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 2079229 (2.0M) [text/plain]
Saving to: 'ratings.data'

ratings.data          100%[=====] 1.98M --.-KB/s in 0.1s

2024-03-29 23:16:25 (17.1 MB/s) - 'ratings.data' saved [2079229/2079229]

ubh01@ubh01:~$ mysql -u root -p
Enter password:
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 5
Server version: 5.7.42-0ubuntu0.18.04.1 (Ubuntu)

Copyright (c) 2000, 2023, Oracle and/or its affiliates.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql>
```

```
Ubuntu (Snapshot 2) [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Activities Terminal
Fri 23:19
ubh01@ubh01: ~

2024-03-29 23:16:25 (17.1 MB/s) - 'ratings.data' saved [2079229/2079229]

ubh01@ubh01:~$ mysql -u root -p
Enter password:
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 5
Server version: 5.7.42-0ubuntu0.18.04.1 (Ubuntu)

Copyright (c) 2000, 2023, Oracle and/or its affiliates.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> load data local infile 'ratings.data' into table importttt;
ERROR 1046 (3D000): No database selected
mysql> use cognizant;
Reading table information for completion of table and column names
You can turn off this feature to get a quicker startup with -A

Database changed
mysql> load data local infile 'ratings.data' into table importttt;
Query OK, 100003 rows affected (0.69 sec)
Records: 100003 Deleted: 0 Skipped: 0 Warnings: 0

mysql>
```

```
Ubuntu (Snapshot 2) [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Activities Terminal
Fri 23:19
ubh01@ubh01: ~

File Edit View Search Terminal Help
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> load data local infile 'ratings.data' into table importttt;
ERROR 1046 (3D000): No database selected
mysql> use cognizant;
Reading table information for completion of table and column names
You can turn off this feature to get a quicker startup with -A

Database changed
mysql> load data local infile 'ratings.data' into table importttt;
Query OK, 100003 rows affected (0.69 sec)
Records: 100003 Deleted: 0 Skipped: 0 Warnings: 0

mysql> select * from importttt limit 5;
+-----+
| uid | mid | rating | ts |
+-----+
| 0 | 50 | 5 | 881250949 |
| 0 | 172 | 5 | 881250949 |
| 0 | 133 | 1 | 881250949 |
| 196 | 242 | 3 | 881250949 |
| 186 | 302 | 3 | 891717742 |
+-----+
5 rows in set (0.00 sec)

mysql>
```

```
Ubuntu (Snapshot 2) [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Activities Terminal
Fri 23:33
ubh01@ubh01: ~

File Edit View Search Terminal Help
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)=3350
Total time spent by all reduces in occupied slots (ms)=0
Total time spent by all map tasks (ms)=3350
Total vcore-seconds taken by all map tasks=3350
Total megabyte-seconds taken by all map tasks=3430400

Map-Reduce Framework
Map input records=100003
Map output records=100003
Input split bytes=87
Spilled Records=0
Failed Shuffles=0
Merged Map outputs=0
GC time elapsed (ms)=123
CPU time spent (ms)=2800
Physical memory (bytes) snapshot=194748416
Virtual memory (bytes) snapshot=1931104256
Total committed heap usage (bytes)=139460608

File Input Format Counters
Bytes Read=0
File Output Format Counters
Bytes Written=1979226

24/03/29 23:33:17 INFO mapreduce.ImportJobBase: Transferred 1.8875 MB in 20.4079 seconds (94.7102 KB/sec)
24/03/29 23:33:17 INFO mapreduce.ImportJobBase: Retrieved 100003 records.
ubh01@ubh01:~$
```

```
Ubuntu (Snapshot 2) [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Activities Terminal Sat 00:16
ubh01@ubh01: ~

File Edit View Search Terminal Help
ubh01@ubh01:~$ hdfs dfs -ls /CTS
24/03/30 00:15:31 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r-- 1 ubh01 supergroup 0 2024-03-29 23:33 /CTS/_SUCCESS
-rw-r--r-- 1 ubh01 supergroup 1979226 2024-03-29 23:33 /CTS/part-m-000000
ubh01@ubh01:~$ hdfs dfs -put /home/ubh01/ratings.data /
24/03/30 00:16:09 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
ubh01@ubh01:~$ hdfs dfs -ls /
24/03/30 00:16:19 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 4 items
drwxr-xr-x - ubh01 supergroup 0 2024-03-29 23:33 /CTS
-rw-r--r-- 1 ubh01 supergroup 2079229 2024-03-30 00:16 /ratings.data
drwx-wx-wx - ubh01 supergroup 0 2024-03-29 23:32 /tmp
drwxr-xr-x - ubh01 supergroup 0 2022-01-20 17:54 /user
ubh01@ubh01:~$
```

```
Ubuntu (Snapshot 2) [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Activities Terminal Sat 00:18
ubh01@ubh01: ~

File Edit View Search Terminal Help
drwxr-xr-x - ubh01 supergroup 0 2024-03-29 23:33 /CTS
-rw-r--r-- 1 ubh01 supergroup 2079229 2024-03-30 00:16 /ratings.data
drwx-wx-wx - ubh01 supergroup 0 2024-03-29 23:32 /tmp
drwxr-xr-x - ubh01 supergroup 0 2022-01-20 17:54 /user
ubh01@ubh01:~$ mysql -u root -p
Enter password:
Welcome to the MySQL monitor. Commands end with ; or \g.
Your MySQL connection id is 14
Server version: 5.7.42-0ubuntu0.18.04.1 (Ubuntu)

Copyright (c) 2000, 2023, Oracle and/or its affiliates.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> use cognizant;
Reading table information for completion of table and column names
You can turn off this feature to get a quicker startup with -A

Database changed
mysql> create table exportttt (uid int,mid int,rating int ,ts varchar(255));
Query OK, 0 rows affected (0.03 sec)

mysql> quit;
Bye
ubh01@ubh01:~$
```



```
Ubuntu (Snapshot 2) [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Activities Terminal Sat 00:22 ubh01@ubh01: ~

File Edit View Search Terminal Help
Reading table information for completion of table and column names
You can turn off this feature to get a quicker startup with -A

Database changed
mysql> create table exportttt (uid int,mid int,rating int ,ts varchar(255));
Query OK, 0 rows affected (0.03 sec)

mysql> quit;
Bye
ubh01@ubh01:~$ sqoop export --connect jdbc:mysql://localhost:3306/cognizant --table exportttt --username root --password password --ex
port-dir /ratings.data --fields-terminated-by '\t' --lines-terminated-by '\n'
Warning: /home/ubh01/sqoop-1.4.7.bin__hadoop-2.6.0/./hcatalog does not exist! HCatalog jobs will fail.
Please set $HCAT_HOME to the root of your HCatalog installation.
Warning: /home/ubh01/sqoop-1.4.7.bin__hadoop-2.6.0/./accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
Warning: /home/ubh01/sqoop-1.4.7.bin__hadoop-2.6.0/./zookeeper does not exist! Accumulo imports will fail.
Please set $ZOOKEEPER_HOME to the root of your Zookeeper installation.
24/03/30 00:22:08 INFO sqoop.Sqoop: Running Sqoop version: 1.4.7
24/03/30 00:22:08 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
24/03/30 00:22:08 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
24/03/30 00:22:08 INFO tool.CodeGenTool: Beginning code generation
Loading class 'com.mysql.jdbc.Driver'. This is deprecated. The new driver class is 'com.mysql.cj.jdbc.Driver'. The driver is automati
cally registered via the SPI and manual loading of the driver class is generally unnecessary.
24/03/30 00:22:09 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM 'exportttt' AS t LIMIT 1
24/03/30 00:22:09 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM 'exportttt' AS t LIMIT 1
24/03/30 00:22:09 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /home/ubh01/hadoop-2.7.1
Note: /tmp/sqoop-ubh01/compile/b906aaba12bd5b5dcd288d612ceaf668/exportttt.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
24/03/30 00:22:10 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-ubh01/compile/b906aaba12bd5b5dcd288d612ceaf668/exportttt.j
ar
```

```
Ubuntu (Snapshot 2) [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Activities Terminal Sat 00:25 ubh01@ubh01: ~

File Edit View Search Terminal Help
File Output Format Counters
Bytes Written=0
24/03/30 00:22:31 INFO mapreduce.ExportJobBase: Transferred 1.9859 MB in 20.091 seconds (101.2187 KB/sec)
24/03/30 00:22:31 INFO mapreduce.ExportJobBase: Exported 100003 records.
ubh01@ubh01:~$ mysql -u root -p
Enter password:
Welcome to the MySQL monitor. Commands end with ; or \g.
Your MySQL connection id is 20
Server version: 5.7.42-0ubuntu0.18.04.1 (Ubuntu)

Copyright (c) 2000, 2023, Oracle and/or its affiliates.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> select * from exportttt limit 5;
ERROR 1046 (3D000): No database selected
mysql> use cognizant;
Reading table information for completion of table and column names
You can turn off this feature to get a quicker startup with -A

Database changed
mysql> select * from exportttt limit 5;
+-----+-----+-----+-----+
| uid | mid | rating | ts |
+-----+-----+-----+-----+
| 0 | 50 | 5 | 881250949 |
+-----+-----+-----+-----+
```

Ubuntu (Snapshot 2) [Running] - Oracle VM VirtualBox

File Machine View Input Devices Help

Activities Terminal Sat 00:24 ubh01@ubh01: ~

File Edit View Search Terminal Help

Copyright (c) 2000, 2023, Oracle and/or its affiliates.

Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> select \* from exportttt limit 5;  
ERROR 1046 (3D000): No database selected  
mysql> use cognizant;  
Reading table information for completion of table and column names  
You can turn off this feature to get a quicker startup with -A

Database changed  
mysql> select \* from exportttt limit 5;

uid	mid	rating	ts
0	50	5	881250949
0	172	5	881250949
0	133	1	881250949
196	242	3	881250949
186	302	3	891717742

5 rows in set (0.00 sec)

mysql>

ENG US 12:26 AM 3/30/2024