

## ASSIGNMENT 3 - SQOOP

1. Create Movies table in MySQL
2. load data in movie table
3. import movie data from MySQL to HDFS with SQOOB
4. Load rating file in HDFS
5. export rating data from HDFS to MySQL with SQOOB

Induja D  
2319849  
CSDAIA24AZ003

```
Ubuntu (Snapshot 2) [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help

Sun 23:21
ubh01@ubh01: ~

File Edit View Search Terminal Tabs Help

ubh01@ubh01: ~
ubh01@ubh01: ~
ubh01@ubh01: ~

ubh01@ubh01:~$ mysql -u root -p
Enter password:
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 30
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> create database induja
-> create database induja;
ERROR 1064 (42000): You have an error in your SQL syntax; check the manual that corresponds to your MySQL server version for the right syntax to use near 'create database induja' at line 2
mysql> create database induja;
Query OK, 1 row affected (0.00 sec)

mysql> use induja;
Database changed
mysql> create table Movies(movie_id int,movie_title varchar(255),release_date date,video_release date,IMdB_URL varchar(255),unknown int,Action int,Adventure int,Animation int,Childrens int,Comedy int,Crime int,Documentary int,Drama int, Fantasy int,Film_Noir int,Horror int,Musical int,Mystery int,Romance int,Sci_Fci int,Thriller int,War int,Western int);
Query OK, 0 rows affected (0.06 sec)
```

```
Ubuntu (Snapshot 2) [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help

Sun 22:39
ubh01@ubh01: ~

File Edit View Search Terminal Tabs Help

ubh01@ubh01: ~
ubh01@ubh01: ~
ubh01@ubh01: ~

t syntax to use near 'intc,Comedy int,Crime int,Documentary int,Drama int, Fantasy int,Film_Noir int,H' at line 1
mysql> create table Movies(movie_id int,movie_title varchar(255),release_date date,video_release date,IMdB_URL varchar(255),unknown int,Action int,Adventure int,Animation int,Childrens int,Comedy int,Crime int,Documentary int,Drama int, Fantasy int,Film_Noir int,Horror int,Musical int,Mystery int,Romance int,Sci_Fci int,Thriller int,War int,Western int);
Query OK, 0 rows affected (0.05 sec)

mysql> load data local infile '/home/ubh01/movies.item' into table Movies fields terminated by '|'
-> load data local infile '/home/ubh01/movies.item' into table Movies fields terminated by '|';
ERROR 1064 (42000): You have an error in your SQL syntax; check the manual that corresponds to your MySQL server version for the right syntax to use near 'terminated by '|''
load data local infile '/home/ubh01/movies.item' into table Mo' at line 1
mysql> load data local infile '/home/ubh01/movies.item' into table Movies fields terminated by '|';
ERROR 2 (HY000): File '/home/ubh01/movies.item' not found (Errcode: 2 - No such file or directory)
mysql> load data local infile '/home/ubh01/movies.item' into table Movies fields terminated by '|';
Query OK, 1635 rows affected, 39221 warnings (0.05 sec)
Records: 1635 Deleted: 0 Skipped: 0 Warnings: 39221

mysql> quit
Bye
ubh01@ubh01:~$ sqoop import --connect jdbc:mysql://local:3306/indu --username root --password password --table movies --target-dir /t
mported --m 1
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.
24/03/31 22:36:45 INFO sqoop.Sqoop: Running Sqoop version: 1.4.7
```

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

File Machine View Input Devices Help

Sun 23:22

ubh01@ubh01: ~

File Edit View Search Terminal Tabs Help

ubh01@ubh01: ~

```
ubh01@ubh01:~$ mysql -u root -p
Enter password:
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 32
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 induja
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 ratings(user_id int,m_id int,item int,ts int);
Query OK, 0 rows affected (0.03 sec)

mysql> exit
Bye
ubh01@ubh01:~$ sqoop export --connect jdbc:mysql://localhost:3306/induja --table ratings --username sqoop --password password --export-dir /tables/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.
```

Type here to search

ENG 23:22 US 31-03-2024

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

File Machine View Input Devices Help

Sun 23:27

ubh01@ubh01: ~

File Edit View Search Terminal Tabs Help

ubh01@ubh01: ~

```
org.apache.hadoop.mapreduce.lib.input.InvalidInputException: Input path does not exist: hdfs://127.0.0.1:9000/tables/ratings.data
at org.apache.hadoop.mapreduce.lib.input.FileInputFormat.singleThreadedListStatus(FileInputFormat.java:323)
at org.apache.hadoop.mapreduce.lib.input.FileInputFormat.listStatus(FileInputFormat.java:265)
at org.apache.sqoop.mapreduce.ExportInputFormat.getJobSize(ExportInputFormat.java:51)
at org.apache.sqoop.mapreduce.ExportInputFormat.getSplits(ExportInputFormat.java:64)
at org.apache.hadoop.mapreduce.JobSubmitter.writeNewSplits(JobSubmitter.java:301)
at org.apache.hadoop.mapreduce.JobSubmitter.writeSplits(JobSubmitter.java:318)
at org.apache.hadoop.mapreduce.JobSubmitter.submitJobInternal(JobSubmitter.java:196)
at org.apache.hadoop.mapreduce.Job$10.run(Job.java:1290)
at org.apache.hadoop.mapreduce.Job$10.run(Job.java:1287)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:422)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1657)
at org.apache.hadoop.mapreduce.Job.submit(Job.java:1287)
at org.apache.hadoop.mapreduce.Job.waitForCompletion(Job.java:1308)
at org.apache.sqoop.mapreduce.ExportJobBase.doSubmitJob(ExportJobBase.java:324)
at org.apache.sqoop.mapreduce.ExportJobBase.runJob(ExportJobBase.java:301)
at org.apache.sqoop.mapreduce.ExportJobBase.runExport(ExportJobBase.java:442)
at org.apache.sqoop.manager.SqlManager.exportTable(SqlManager.java:931)
at org.apache.sqoop.tool.ExportTool.exportTable(ExportTool.java:80)
at org.apache.sqoop.tool.ExportTool.run(ExportTool.java:99)
at org.apache.sqoop.Sqoop.run(Sqoop.java:147)
at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:70)
at org.apache.sqoop.Sqoop.runSqoop(Sqoop.java:183)
at org.apache.sqoop.Sqoop.runTool(Sqoop.java:234)
at org.apache.sqoop.Sqoop.runTool(Sqoop.java:243)
at org.apache.sqoop.Sqoop.main(Sqoop.java:252)
ubh01@ubh01:~$
```

Type here to search

ENG 23:27 US 31-03-2024

