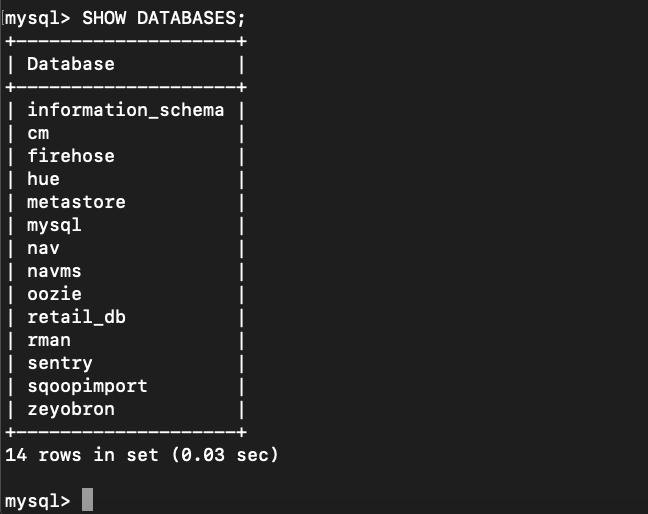
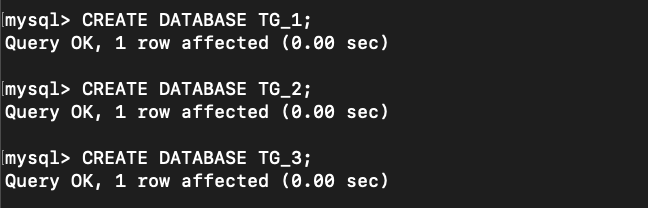
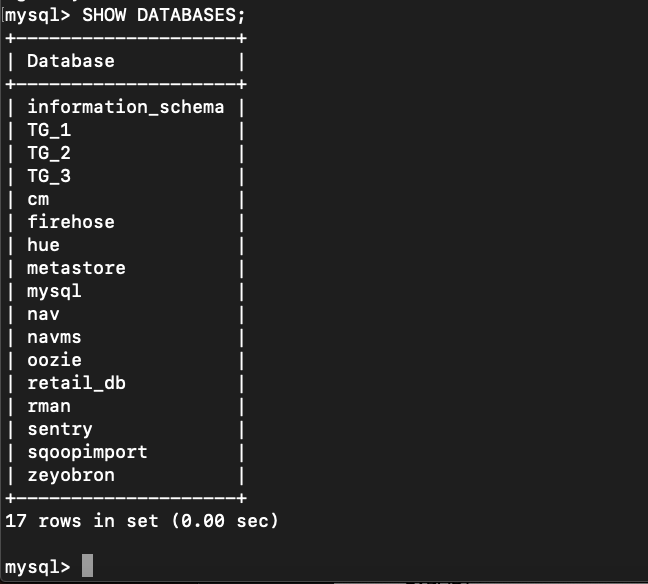
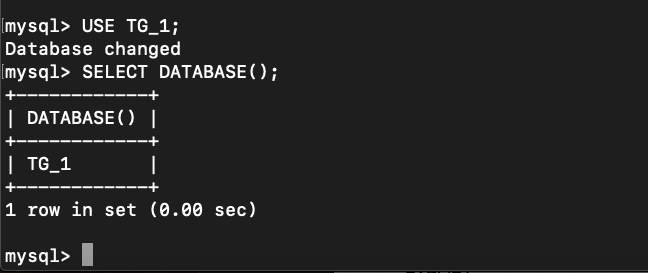
APACHE SQOOP L1

MYSQL :

1. List out all the databases available in your environment.

2. Create 3 databases and name it as TG\_1, TG\_2, TG\_3 in mysql.



3. make TG\_1 as active database and create following tables avaliable.

empdata:

ename,empno,sal,deptno

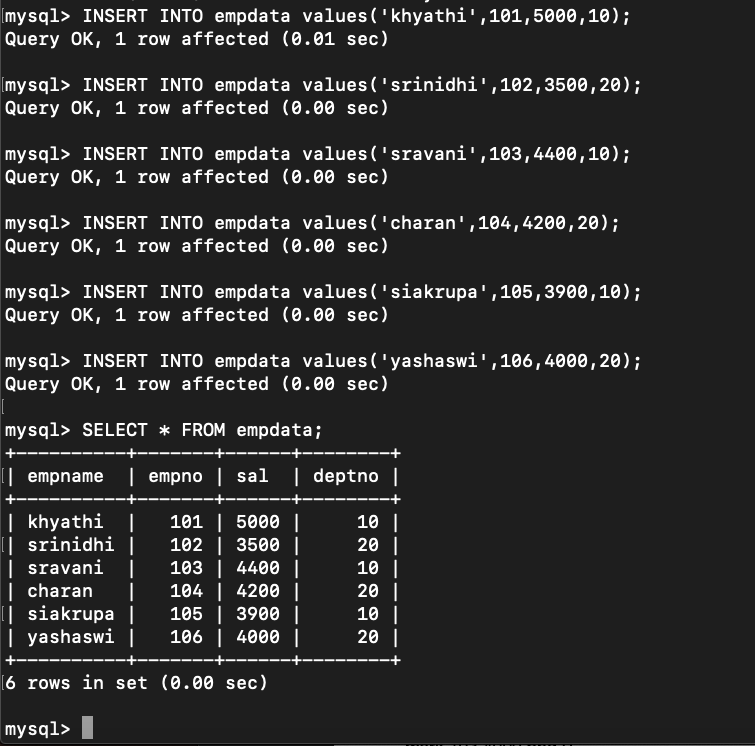
khyathi,101,5000,10

srinidhi,102,3500,20

sravani,103,4400,10

charan,104,4200,20

saikrupa,105,3900,10

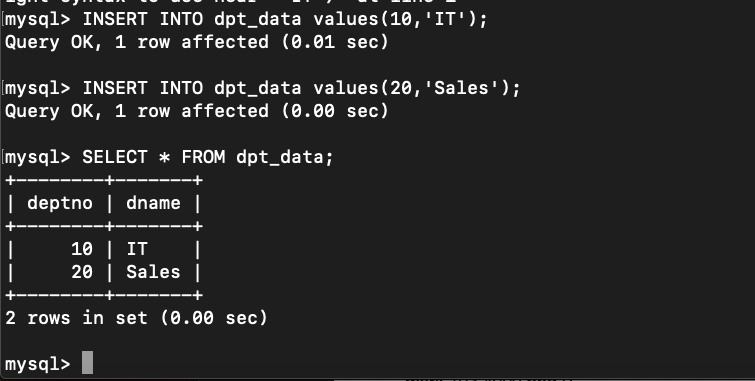
yashaswi,106,4000,20

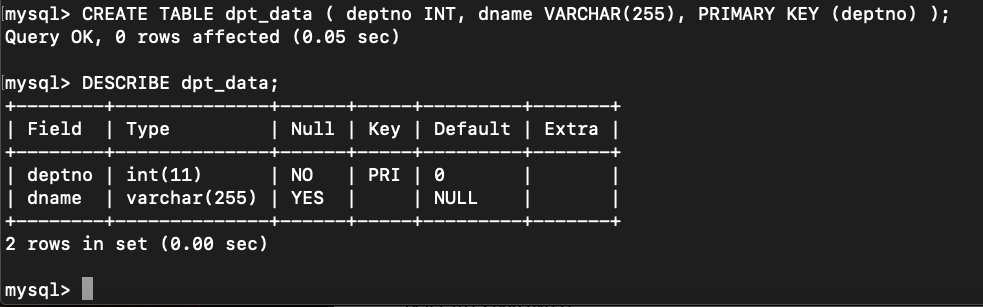
dpt\_data :

deptno,dname

10, IT

20,Sales

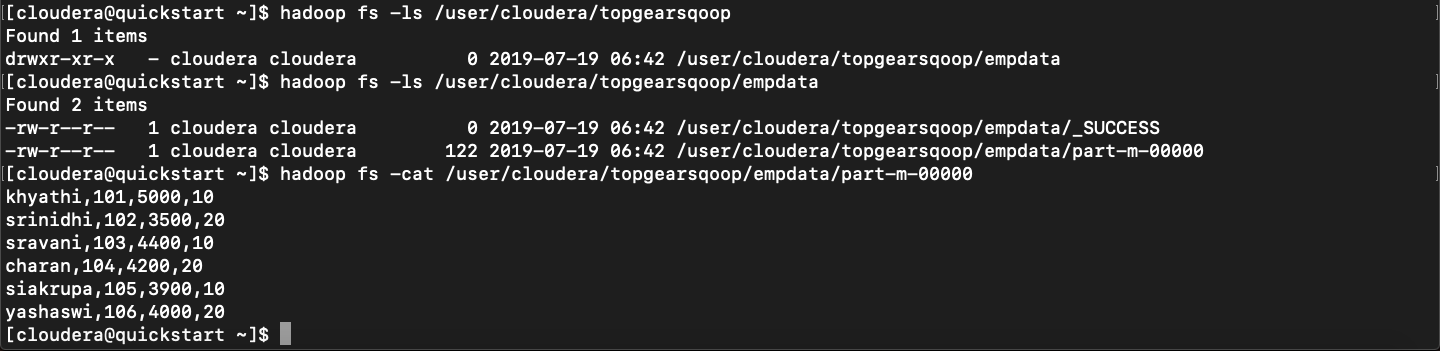


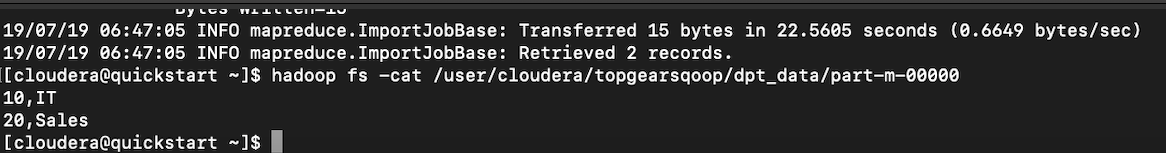
3. Display metadata of above tables you created.

Import in sqoop :

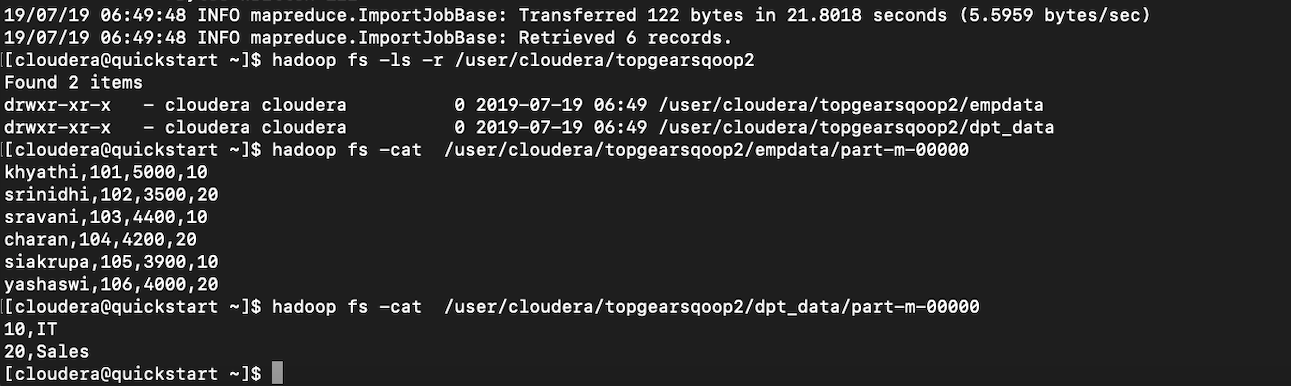
4. import above two tables one by one into TG\_1.

sqoop import --connect jdbc:mysql://localhost/TG\_1 --username root --password cloudera --table empdata --m 1 --warehouse-dir /user/cloudera/topgearsqoop



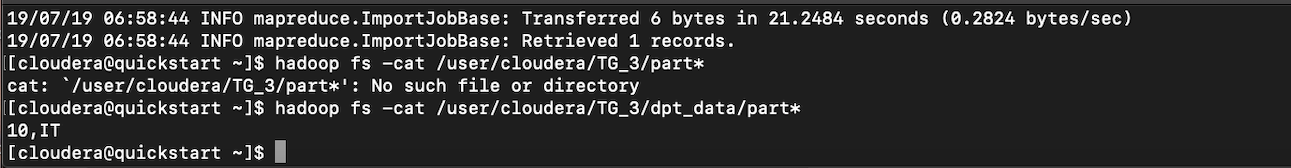
sqoop import --connect jdbc:mysql://localhost/TG\_1 --username root --password cloudera --table dpt\_data --m 1 --warehouse-dir /user/cloudera/topgearsqoop

5. import above two tables at a time into TG\_2.

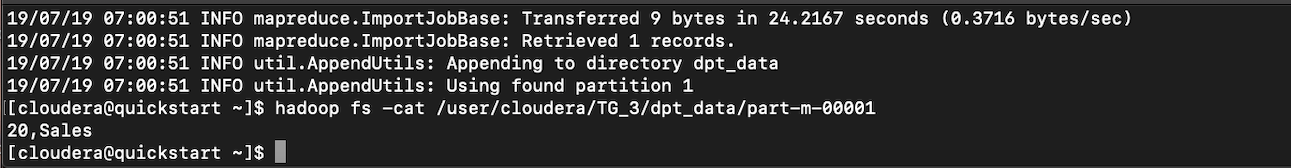
sqoop import-all-tables --connect jdbc:mysql://localhost/TG\_1 --username root --password cloudera --m 1 --warehouse-dir /user/cloudera/topgearsqoop2

6. import first table data whose deptno is 10 into TG\_3

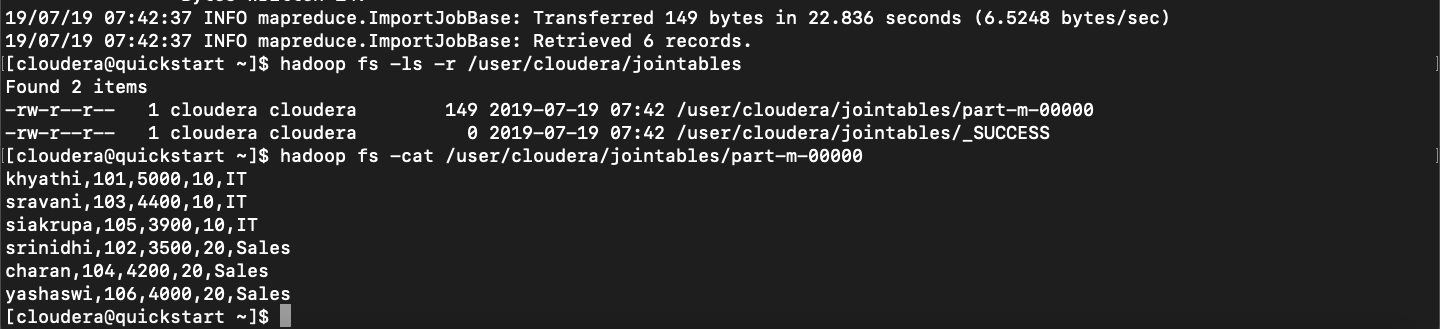
sqoop import --connect jdbc:mysql://localhost/TG\_1 --username root --password cloudera --table dpt\_data --where "deptno=10" --m 1 --warehouse-dir /user/cloudera/TG\_3



7. import first table data whose deptno is 20 appending into above table from TG\_3.

sqoop import --connect jdbc:mysql://localhost/TG\_1 --username root --password cloudera --table dpt\_data --where “deptno=20” —append —m 1 --warehouse-dir /user/cloudera/TG\_3

8. import join result of above two tables using a query and load into hdfs.

sqoop import --connect jdbc:mysql://localhost/TG\_1 --username root --password cloudera --query 'SELECT empdata.empname, empdata.empno,empdata.sal,empdata.deptno, dpt\_data.dname FROM empdata INNER JOIN dpt\_data ON empdata.deptno = dpt\_data.deptno WHERE $CONDITIONS' --m 1 --target-dir \user\cloudera\jointables

EXPORT :

Consider following hdfs file.

empdtls:

ename,empno,sal,loc,dept

peter,101,2000,hyd,IT

sam,102,3900,pune,SALES

mark,103,4000,bng,IT

blake,104,5400,chn,mrktng

smith,105,3200,delhi,sales

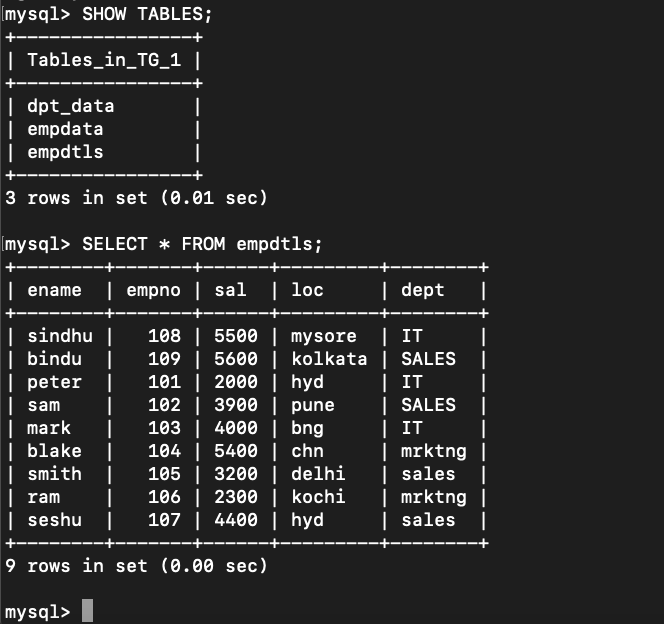
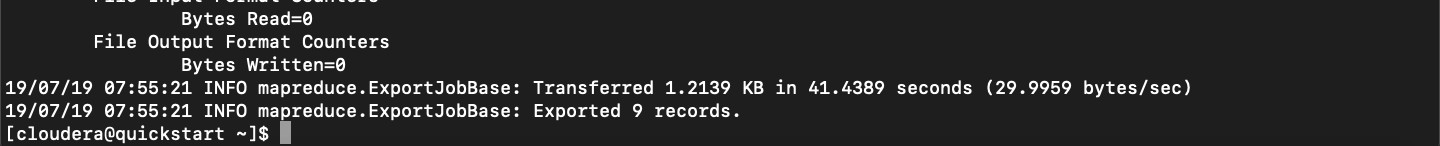
ram,106,2300,kochi,mrktng

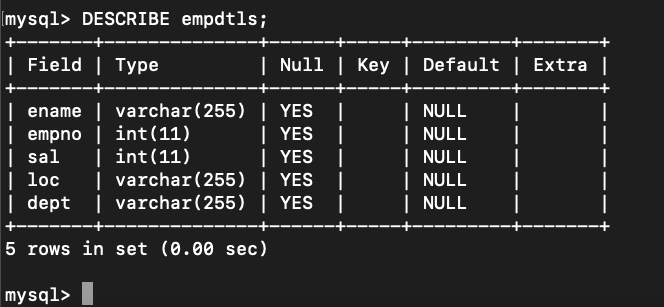
seshu,107,4400,hyd,sales

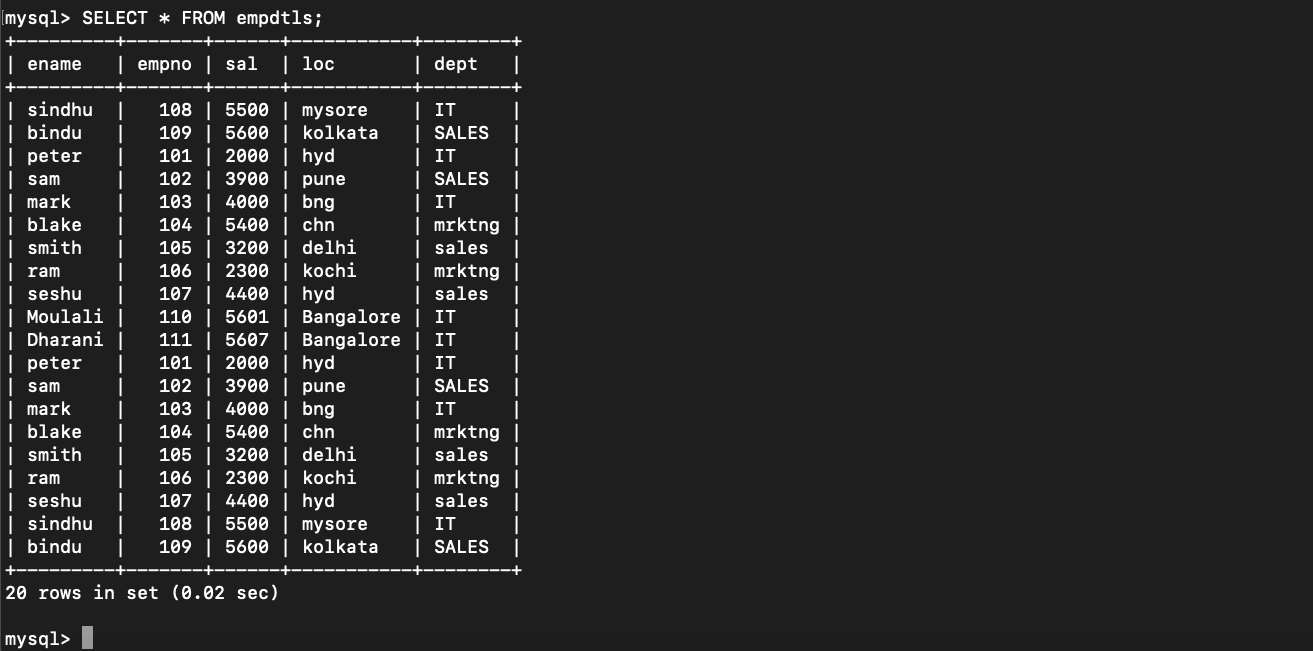
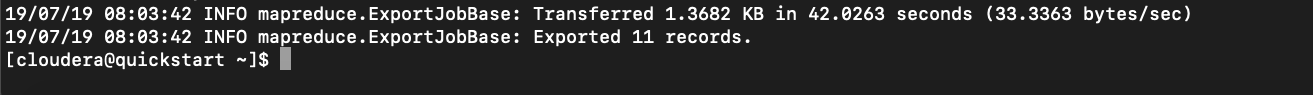
sindhu,108,5500,mysore,IT

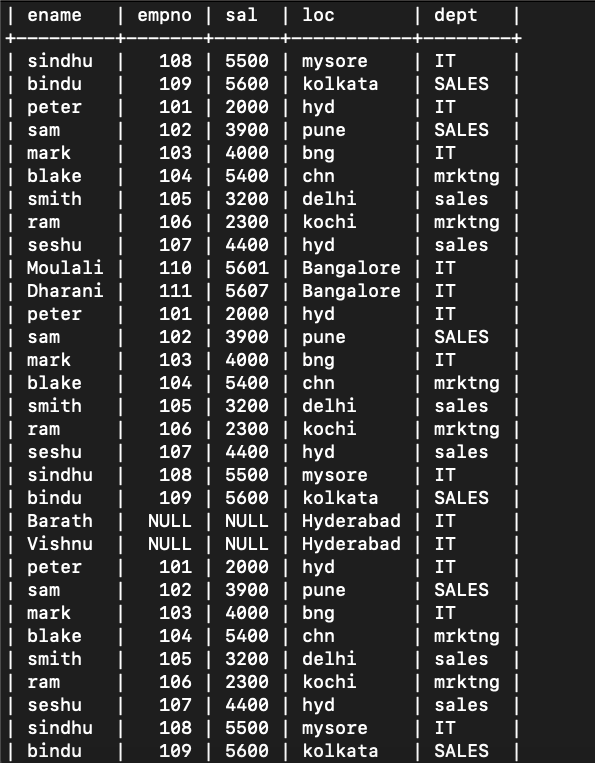
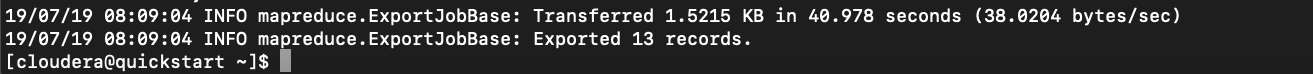
bindu,109,5600,kolkata,SALES

9. export this file from hdfs to mysql.

sqoop export --connect jdbc:mysql://localhost/TG\_1 --username root --password cloudera --table empdtls --export-dir /user/cloudera/TG\_1

10. display structure of this exported file and display contents.

11. add few records in hdfs file and export them to mysql.

12. add couple of records into hdfs file with nulls(numeric) and export them to mysql.

13. verify data from hdfs file and mysql file. update all nulls to -1

sqoop export --connect jdbc:mysql://localhost/TG\_1 --username root --password cloudera --table empdtls —input-null-non-string '-1' --export-dir /user/cloudera/TG\_1

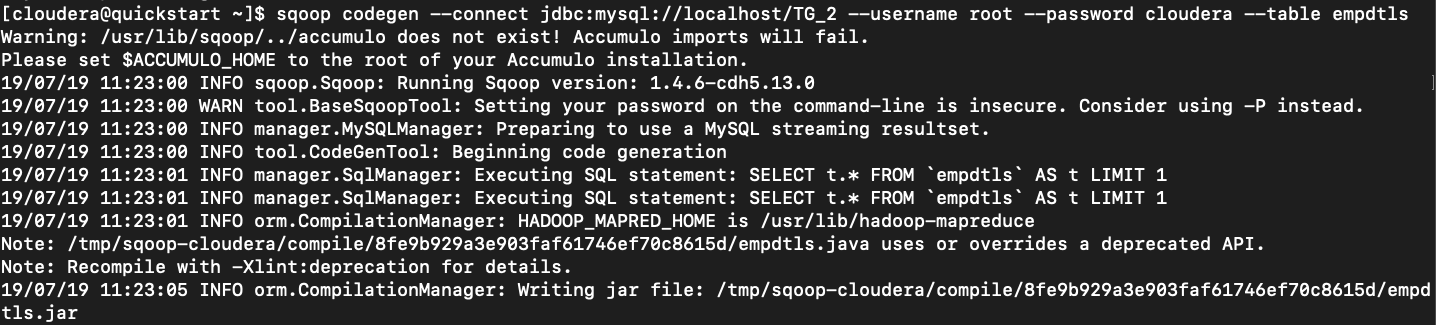
Handling NULL values in snoop export is not possible…

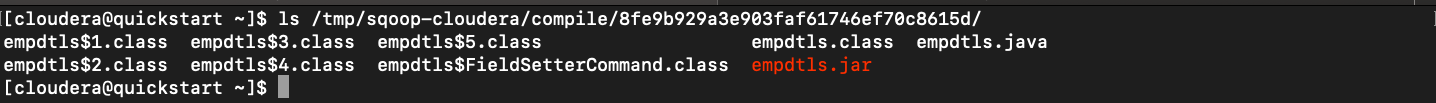
CODEGEN :

14. Create a database in mysql , name it as TG\_4.

15. create above table empdtls into TG\_4.

16. Develop a javacode for importing empdtls table to hdfs. ( Attach or copy .java file while submiting)





17. Develop a javacode for exporting hdfs file to a table in mysql. ( Attach or copy .java file while submiting)

Not able to find…

Unable to differentiate between the sqoop import and sqoop export codegen