# Exercise:

Ref: <https://sqoop.apache.org/docs/1.4.6/SqoopUserGuide.html>

1. **Sqoop Export**

**Create table in Hive and insert data into it:**

CREATE TABlE test(id INT, name STRING) ROW FORMAT delimited fields terminated by ‘,’ LINES TERMINATED BY ‘\n’ STORED AS TEXTFILE;

Insert into test values(1,’name1’);

Insert into test values(2,’name2’);

(describe formatted test) to get location of hive database

**Create table in mysql database (retail)**

**Export Hive table to mysql:**

sqoop export --connect jdbc:mysql://localhost:3306/retail\_db --username retail\_dba --password cloudera --table test --fields-terminated-by ‘,’ --export-dir <HDFS DIRCTORY NAME>

sqoop export --connect jdbc:mysql://localhost:3306/retail\_db --username retail\_dba --password cloudera --table test --fields-terminated-by ',' --export-dir /user/hive/warehouse/test

**See the data into mysql table:**

Select \* from test;

**Try Out**

1. Update records in RDBMS table based on data stored in HDFS
2. **Sqoop Import**

**Import Table from RDBMS**

sqoop import --connect jdbc:mysql://localhost:3306/<DATABASE NAME> --username root –password cloudera --table <TABLE NAME> --m 1 --target-dir <HDFS DIRCTORY NAME>

sqoop import --connect jdbc:mysql://localhost:3306/retail\_db --username retail\_dba --password cloudera --table test --m 1 --target-dir /user/hive/warehouse/import

\* --m 1: one mapper

ALTER TABLE test SET LOCATION '/user/hive/warehouse/import';

**Import data with a condition**

sqoop import --connect jdbc:mysql://localhost:3306/<DATABASE NAME> --username root -p --table <TABLE NAME> --m 1 --where "<CONDITION>" --target-dir <HDFS DIRCTORY NAME>

**Try out**

1. Delete target directory
2. Append imported records to existing data

sqoop import --connect jdbc:mysql://localhost:3306/retail\_db --username retail\_dba --password cloudera --table test --m 1 --where "id>=3" --target-dir /user/hive/warehouse/import **--delete-target-dir**

--append (directly insert into Hive Table if dir known)

--hive-import --hive-table <db.tbname> --split -- fields-terminated by ',' (in case, dir is dynamic)

1. **Sqoop Jobs**

**Create Job**

sqoop job --create myjob -- import --connect jdbc:mysql://localhost:3306/retail\_db --username retail\_dba --password cloudera --table departments --target-dir <HDFS DIRECTORY NAME> --fields-terminated-by ‘,’

**List all created jobs**

sqoop job --list

**Show details of one specific job**

sqoop job --show myjob

**Execute created job**

sqoop job --exec myjob

1. **Eval (modify SQL table)**

**Insert a record into RDBMS table**

sqoop eval --connect jdbc:mysql://localhost:3306/retail\_db --username retail\_dba --password cloudera -e "INSERT INTO test VALUES(999, 'name999')"