------------------------------------------------------------------------------------------------------------

NO (“) and we need to place =,,,,,,,,, DONT FORGET TO IMPORT IN HIVE/WAREHOUSE dir

**SQOOP EVAL**

------------------------------------------------------------------------------------------------------------

sqoop eval \

--connect "jdbc:mysql://quickstart.cloudera:3306/retail\_db" \

--username retail\_dba \

--password cloudera \

--query "select count(1) from order\_items

**EXAM**

**-----------------------------------------------------------------------------------------------------**

**sqoop import**

**--connect "jdbc:mysql://quickstart.cloudera:3306/retail\_db"**

**--username=retail\_dba**

**--password=cloudera**

**--table departments**

**--fields-terminated-by '|'**

**--lines-terminated-by '\n'**

**--as-textfile**

**--target-dir=/user/cloudera/cloudera1**

**=========================================================**

**IMORT ALL TABLES**

--------------------------------------------------------------------

sqoop import-all-tables \

-m 4 \

--connect "jdbc:mysql://quickstart.cloudera:3306/retail\_db" \

--username=retail\_dba \

--password=cloudera \

--as-avrodatafile \ (**DONT FORGET TO CREATE DIR IN LOCAL**)

--warehouse-dir=/user/hive/warehouse/retail\_stage.db

--warehouse-dir=/user/cloudera/sqoop\_import

--target-dir=/user/cloudera/departments **for single table**

--as-textfile

--as-sequencefile \

===============================================================

**NOTE: use target dir only for a single table , for importing all tables use warehouse dir**

===============================================================

**IMPORT SINGLE TABLE (HDFS)**

------------------------------------------------------------------------------------------------------

sqoop import \

-m 2 \

--connect "jdbc:mysql://quickstart.cloudera:3306/retail\_db" \

--username=retail\_dba \

--password=cloudera \

**--table departments \**

--as-sequencefile \

**--target-dir=/user/cloudera/**

**DELIMITERS**

-----------------------------------------------------------------------------------------------------

-m 2 \

--as-sequencefile \

--append \ **---------------------means overwrite(update in hive)**

--fields-terminated-by '|' \

--lines-terminated-by '\n' \

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

--null-string hi \

--null-non-string 1 \ **----only for special cases**

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

**BOUNDARY QUERY and COULUMNS**

------------------------------------------------------------------------------------------------------------

sqoop import \

-m 2 \

--connect "jdbc:mysql://quickstart.cloudera:3306/retail\_db" \

--username=retail\_dba \

--password=cloudera \

--table departments \

--boundary-query "select 2, 8 from departments limit 1" \

--columns department\_id,department\_name \

--as-avrodatafile

--target-dir =/use/cloudera/table1

=============================================================

**QUERY and SPLIT-BY and WHERE**

------------------------------------------------------------------------------------------------------------

sqoop import \

-m 2 \

--connect "jdbc:mysql://quickstart.cloudera:3306/retail\_db" \

--username=retail\_dba \

--password=cloudera \

--table departments \

--query="select \* from orders join order\_items on orders.order\_id = order\_items.order\_item\_order\_id where \$CONDITIONS" \

{{{{ A join B on A.id=B.id}}}}

--split-by order\_id \ ----------------- is to provide primary key

--where "department\_id > 7" \ ---------------------- ids with greater than 7 will be imported

-m 4 \

--as-avrodatafile

===============================================================

**INCRIMENTAL LOAD ==>> will create new files by updating data**

------------------------------------------------------------------------------------------------------------

sqoop import \

-m 2 \

--connect "jdbc:mysql://quickstart.cloudera:3306/retail\_db" \

--username=retail\_dba \

--password=cloudera \

--table departments \

--target-dir /user/hive/warehouse/retail\_stage.db/departments \

--check-column "department\_id" \

--incremental append \

--last-value 7 \

--outdir java\_files

**SQOOP JOB**

------------------------------------------------------------------------------------------------------------

sqoop job --create sqoop\_job \

--(space) import \ {{there should be space after -- dont forget}}

--connect "jdbc:mysql://quickstart.cloudera:3306/retail\_db" \

--username=retail\_dba \

--password=cloudera \

--table departments \

--target-dir /user/hive/warehouse/retail\_stage.db/departments \

--append \

--check-column "department\_id" \

--incremental append \

--last-value 7 \

--outdir java\_files

sqoop job --list || sqoop job --show sqoop\_job || sqoop job --exec sqoop\_job

**HIVE IMPORT ALL TABLES**

------------------------------------------------------------------------------------------------------------

sqoop import-all-tables \

--num-mappers 1 \

--connect "jdbc:mysql://quickstart.cloudera:3306/retail\_db" \

--username=retail\_dba \

--password=cloudera \

--hive-import \

--hive-overwrite \

--create-hive-table \

--compress \

--compression-codec org.apache.hadoop.io.compress.SnappyCodec \

--outdir java\_files

=============================================================

**OVERWRITE**

------------------------------------------------------------------------------------------------------------

sqoop import \

--table departments \

--hive-home /user/hive/warehouse/retail\_ods.db \

--hive-import \

--hive-overwrite \

--hive-table departments \

--outdir java\_files

==============================================================

**Create hive table**

-----------------------------------------------------------------------------------------------------------

sqoop import \

--connect "jdbc:mysql://quickstart.cloudera:3306/retail\_db" \

--username=retail\_dba \

--password=cloudera \

--table departments \

--fields-terminated-by '|' \

--lines-terminated-by '\n' \

--hive-home /user/hive/warehouse \

--hive-import \

--hive-table departments\_test \

--create-hive-table \

--outdir java\_files