**Importing Data from MySQL Database to Hive warehouse directory through sqoop then creating table on top this directory in hive :**

***Retail\_db Database Setup :***

***now connect to mysql database as a root :***

***$ mysql -uroot -pcloudera (no space between -p and password)***

***mysql>show databses;***

***Note : mysql>source /root/data/retail\_db.sql this sql script is used to create retail\_db Database and all the table and inserting records in those tables.This script is available at github/dgadiraju github account .***

**In cloudera VMware retail\_db databases is already exists for exercise , let us create user and assign privileges to access this data base other wise this database is accessible by root user and cloudera as a password only**

**Open the terminal :**

**$mysql –uroot –pcloudera;**

**mysql> show databases;**

**mysql>create user sujit identified by ‘sujit’;**

**mysql>grant all on retail\_db.\* to sujit;**

**mysql>flush privileges;**

**mysql>exit;**

**Now log in again with new user sujit in mysql**

**$mysql –usujit –p**

**Enter password: sujit**

**mysql>show databases;**

**it will display two databases**

**information\_schema**

**retail\_db**

**If sql commands are incorrect on mysql prompt then write ‘ then press enter then write ; press enter button.**

***All above commands creation of data base , tables loading data needs to understand , In cloudera all these are available as part of example***

**Now data are stored in to mysql of hadoop let us connect via sqoop**

**Connecting to databases through sqoop**

List Database Table through Sqoop :

**$sqoop list-databases --connect jdbc:mysql://** **quickstart.cloudera:3306/retail\_db --username sujit --password sujit**

To list databases table:

**$sqoop list-tables --connect jdbc:mysql://** **quickstart.cloudera:3306/retail\_db --username sujit --password sujit**

**So by this we can validate by sqoop that we are able to connect to mysql databases.**

**And able to access Databases and tables**

**Import database tables from mysql to hive directory using sqoop :**

**$sqoop import-all-tables --connect jdbc:mysql://** **quickstart.cloudera:3306/retail\_db --username sujit --password sujit --warehouse-dir=/root/hive/warehouse/retail\_db.db**

**We have created directory in hive but we have not created database on top of this**

**We can check retail\_db into hadoop**

**$hadoop fs –ls /user/hive/warehouse**

**$hadoop fs –ls /user/hive/warehouse.retail\_db.db**

**Create database and table on top of hdfs`s below path /user/hive/warehouse/retail\_db.db**

**this is the retail\_db.db directory that is being created as a result of import from mysql through sqoop.If we go into hive :**

**Hive>show databses**

**It will not show retail\_db as a database**

**Next we have to create database and then table – remember name of database, tables and table fields its datatype all should match then only mapping will be done and we will be able to access record from table via hive.**

**hive> create database retail\_db;**

**hive>use retail\_db;**

**hive>create table orders(order\_id INT,order\_date date,order\_customer\_id INT,order\_status varchar(20)) ROW FORMAT DELIMITED FIELDS TERMINATED BY ‘,’ STORED AS TEXTFILE;**

**TABLE IS CREATED MAPPING OF THIS TABLE AND DATA INTO DIRECTORY OF /USER/HIVE/WAREHOUSE WILL TAKE PLACE.**

**Now we can access record from hive also .**

**hive>select \* from retail\_db.orders limit 10;**

**hive>describe formatted orders;**

**it will show complete detail about location of table Table type , input format, output format etc.**