1. Create hdfs directory by name “mydata” under “/user/ubh01”

STEPS:

$ $HADOOP\_HOME/sbin/start-dfs.sh (To Start the hdfs services).

$ $HADOOP\_HOME/sbin/start-yarn.sh (To Start Yarn Service).

$ sudo jps (To check all services are running or not).

$ hdfs dfs –mkdir mydata /user/ubh01/ (Creating a directory in directory(/user/ubh01).

2. Import table called “device” under devsh\_loudacre database from Mysql into hdfs under /user/ubh01/mydata location.

STEPS:

1st we have to log in in toMySql and create databse devsh\_loudacre.sql

Log in to MySql.

$ mysql –u root –p

Enter the password – password

Mysql>create database devsh\_loudacre; (creating database).

Mysql> quit; (to quit from MySql).

$ mysql –u root devsh\_loudacre < devsh\_loudacre.sql (To Run The

Script.) (Press Enter)

Enter the password: password

$mysql –u root –p (press ENter)

Enter the password: password

Mysql> use devsh\_loudacre; (Test the table).

Mysql> show tables; (it will all the tables of devsh\_loudacre database).

We have to imort the file using sqoop.

$ sqoop list-databases –-connect jdbc:mysql://ubh01/ –-username root –password password

$ sqoop list-tables –-connect jdbc:mysql://ubh01/devsh\_loudacre –-username root –-password password

$sqoop import –connect jdbc:mysql://ubh01/devsh\_loudacre -- table device –- fields-terminated-by ‘/t’ –username root –-password password --target-dir /user/ubh01/mydata/ (importing a table device in sqoop).

$ hdfs dfs –ls /user/ubh01/mydata (show files)

$ hdfs dfs –tail /user/ubh01/mydata /part-m-00000(show data)

3.Check from namenode WebUI how many blocks got created.

->In name node WEbUi the total 5 block(\_SUCCESS,part- m0,m1,m2,m3) will be created and where all have Replication factor will be 1.

4. Create hive table called “mydevices” on hive based on below sample data of the devices table imported under /user/ubh01/mydata.

-> First we have start hive service by typing hive.

$ hive (It will start the hive services).

hive> create external table mydevice(device\_num int, release\_dt string, device\_name string, device\_type string)

>row format delimited

>fields terminated by ‘/t’

>location ‘/user/ubh01/mydata’; (Table creation done)

Hive>describe mydevice; (To Describe the mydevice table).

Hive>describe formatted mydevice; (To describe more details like delimited char , tupples etc.).

Hive>select \* from mydevice; (show all the rows of the table mydevice)

5. Filter the records from mydevices with device\_name as “Sorrento…”

Hive>select \* from mydevice where device\_name like ‘Sorrento%’;

6. Count the number of each device\_name.

Hive> select device\_name,count(\*) as Total from mydevice group by device\_name;



