1. Import the table called “**webpage**” from “**devsh\_loudacre**” database in Mysql to hdfs.

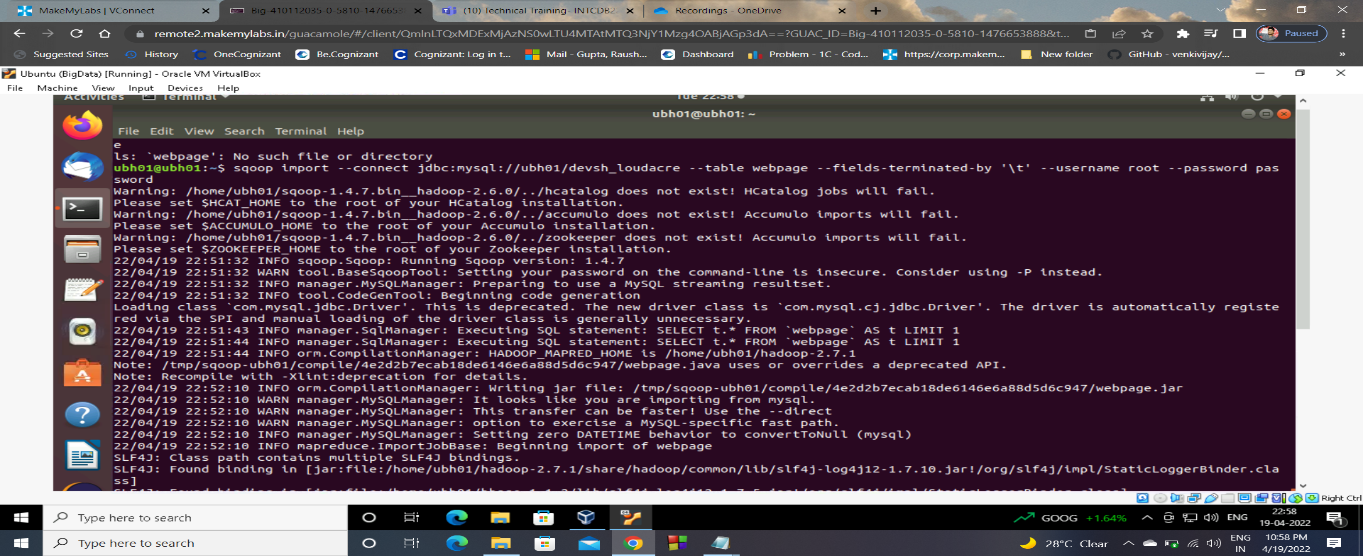
$ sqoop list-databases –-connect jdbc:mysql://ubh01/ –-username root –password password

$ 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)



2.Create a table on Hive called “**webpage\_titanic**” which must contain only the data related to “**titanic**” under the column “**associated\_files**”.

$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).

**Hive>create table webpage(web\_page\_num int,web\_page\_file\_name string,associated\_file string)**

>row format delimited

>fields terminated by ‘,’

>location ‘/user/ubh01/mydata’;

Hive> **create table webpage\_titanic(web\_page\_num int,web\_page\_file\_name string,associated\_file string)**

**>row format delimited**

**>fields terminated by ‘,’;**

**Hive>insert overwrite table webpage\_titanic select \* from webpage where associated\_file like ‘%titanic%’;**

3.Make the table “**webpage\_titanic**” transactional.

hive>

SET hive.support.concurrency=true;

SET hive.txn.manager=org.apache.hadoop.hive.ql.lockmgr.DbTxnManager;

SET hive.exec.dynamic.partition.mode=nonstrict;

CREATE TABLE webpage\_titanic(**web\_page\_num int,web\_page\_file\_name string,associated\_file string**)

TBLPROPERTIES("transactional"="true");

4.

Create partitioned table based on **depts** (dept10, dept20, and dept30).

hdfs dfs -mkdir /user/root/depts

hdfs dfs -mkdir /user/ubh01/depts/dept10

hdfs dfs -mkdir /user/ubh01/depts/dept20

hdfs dfs -mkdir /user/ubh01/depts/dept30

hdfs dfs -put dept10.txt /user/ubh01/depts/dept10

hdfs dfs -put dept20.txt /user/ubh01/depts/dept20

hdfs dfs -put dept30.txt /user/ubh01/depts/dept30

$hive

Hive>create external table deptsnew(empno int,ename string,job string,mgr int,hiredate string,salary int,comm int,deptno int)

>row format delimited

>fields terminated by '/'

>location '/user/root/depts';

Hive> select \* from depts;

(you must not see any rows displayed)

Hive >create external table deptspartempno int,ename string,job string,mgr int,hiredate string,salary int,comm int)

>partitioned by (deptno string)

>row format delimited

>fields terminated by '/'

>location '/user/root/depts';

Hive>ALTER TABLE deptspart ADD PARTITION (deptno='dept10') LOCATION '/user/root/depts/dept10';

Hive>ALTER TABLE deptspart ADD PARTITION (deptno='dept20') LOCATION '/user/root/depts/dept20';

Hive>ALTER TABLE deptspart ADD PARTITION (deptno='dept30') LOCATION '/user/root/depts/dept30';

