**ACD\_BDD\_Session\_15\_Assignment\_1\_Main**

**Problem Statement**

Link: https://acadgild.com/blog/file-formats-in-apache-hive/

Refer the above given link for File formats in Hive and implement all the file formats in Hive by using any publicly available data set of your choice.

**Solution**

**TEXTFILE FORMAT**

CREATE TABLE olympic\_textfile(

Athlete STRING,

Age INT,

Country STRING,

Year INT,

Closing\_Date STRING,

Sport STRING,

Gold\_Medal INT,

Silver\_Medal INT,

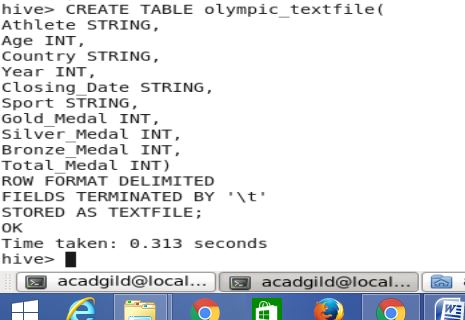
Bronze\_Medal INT,

Total\_Medal INT)

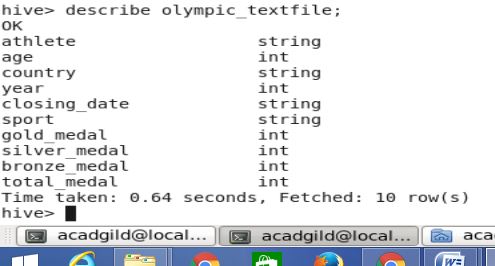
ROW FORMAT DELIMITED

FIELDS TERMINATED BY '\t'

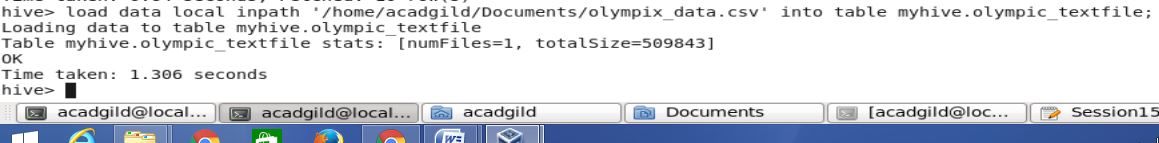
STORED AS TEXTFILE;



describe olympic\_textfile;



load data local inpath '/home/acadgild/Documents/olympix\_data.csv' into table myhive.olympic\_textfile;



Select athlete from olympic\_textfile;

C:\Users\RAKESH\Desktop\Screenshot\Session15_Assign_1\S1.4.JPG



**SEQUENCEFILE FORMAT**

CREATE TABLE olympic\_sequencefile(

Athlete STRING,

Age INT,

Country STRING,

Year INT,

Closing\_Date STRING,

Sport STRING,

Gold\_Medal INT,

Silver\_Medal INT,

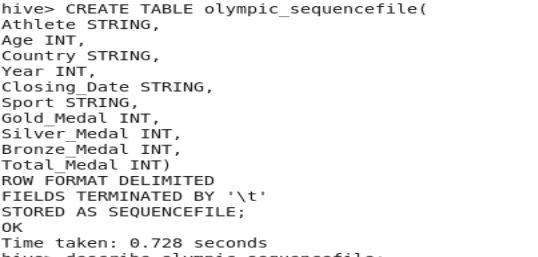
Bronze\_Medal INT,

Total\_Medal INT)

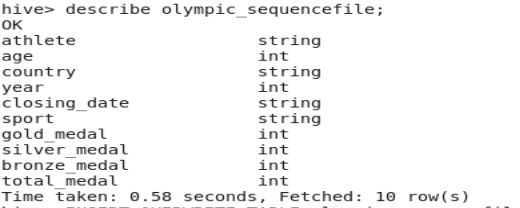
ROW FORMAT DELIMITED

FIELDS TERMINATED BY '\t'

STORED AS SEQUENCEFILE;

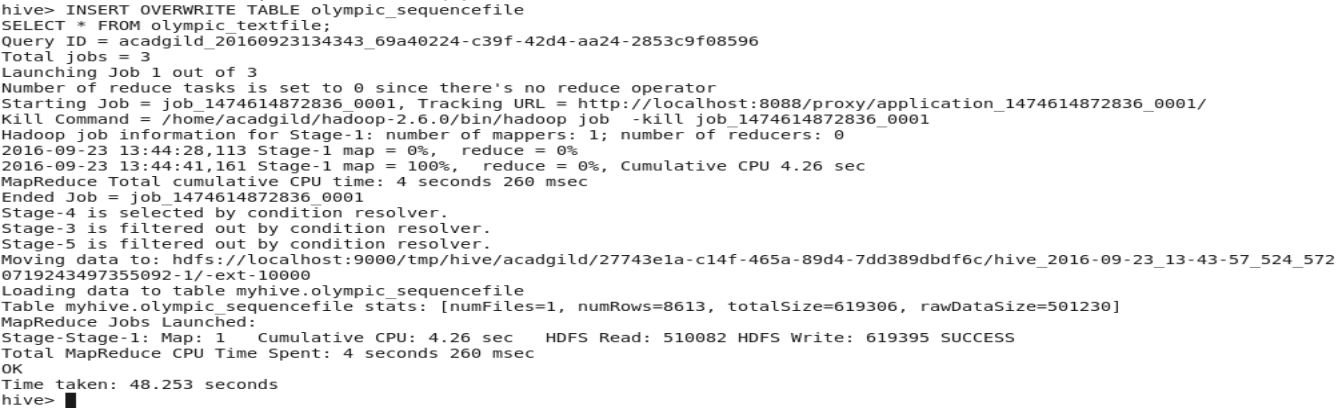


describe olympic\_sequencefile;



INSERT OVERWRITE TABLE olympic\_sequencefile

SELECT \* FROM olympic\_textfile;



SELECT athlete FROM olympic\_sequencefile;



**RCFILE FORMAT**

CREATE TABLE olympic\_rcfile(

Athlete STRING,

Age INT,

Country STRING,

Year INT,

Closing\_Date STRING,

Sport STRING,

Gold\_Medal INT,

Silver\_Medal INT,

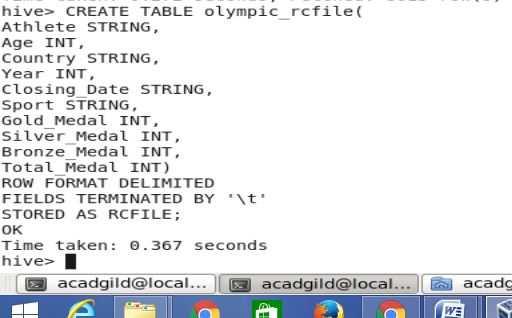
Bronze\_Medal INT,

Total\_Medal INT)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY '\t'

STORED AS RCFILE;

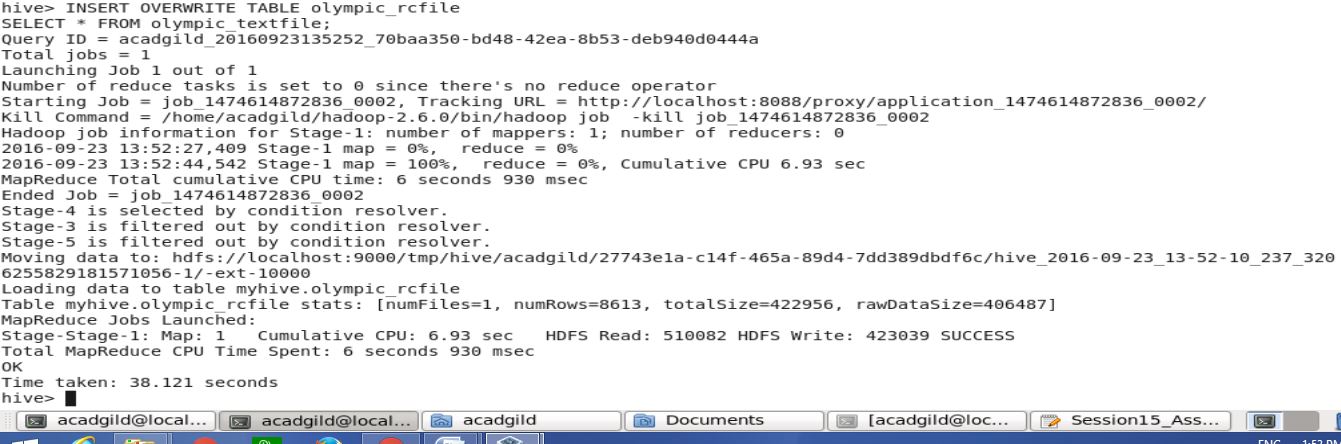


describe olympic\_rcfile;



INSERT OVERWRITE TABLE olympic\_rcfile

SELECT \* FROM olympic\_textfile;



SELECT athlete FROM olympic\_rcfile;



**ORCFILE FORMAT**

CREATE TABLE olympic\_orcfile(

Athlete STRING,

Age INT,

Country STRING,

Year INT,

Closing\_Date STRING,

Sport STRING,

Gold\_Medal INT,

Silver\_Medal INT,

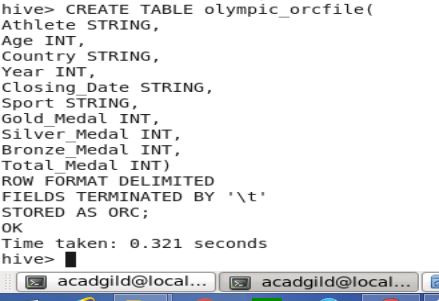
Bronze\_Medal INT,

Total\_Medal INT)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY '\t'

STORED AS ORC;

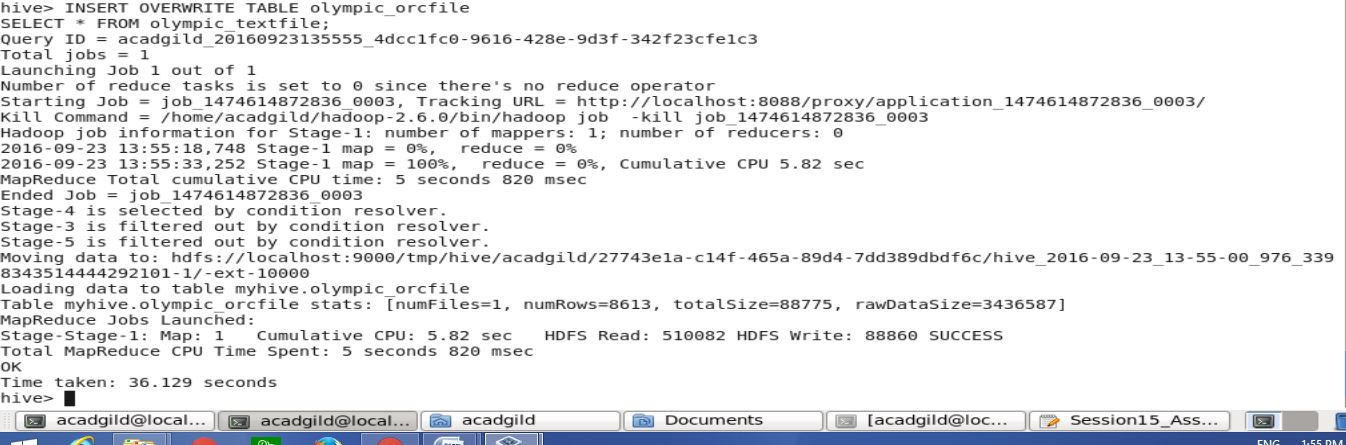


describe olympic\_orcfile;



INSERT OVERWRITE TABLE olympic\_orcfile

SELECT \* FROM olympic\_textfile;



SELECT athlete FROM olympic\_orcfile;

