**HIVE Assignment 2**

**Create external tables for drivers and timesheet data and answer the following queries**

**HIVE QL >>** Create EXTERNAL TABLE drivers\_k (driverId int,name string,ssn string,location string,certified string, wage\_plan string ) ROW FORMAT DELIMITED FIELDS TERMINATED BY ‘,’ tblproperties('skip.header.line.count' = '1');

Create EXTERNAL TABLE timesheet\_n (driverid int,week int,hours\_logged int,miles\_logged int ) ROW FORMAT DELIMITED FIELDS TERMINATED BY ‘,’ tblproperties('skip.header.line.count' = '1');

A close up of a text

Description automatically generated

Loading data into tables:

A screen shot of a computer

Description automatically generated

**Query:**

**How many total hours and miles logged by each driver?**

hive> select driverid , sum(hours\_logged), sum(miles\_logged) from timesheet\_n group by driverid;

A white background with black text

Description automatically generated

A screenshot of a number

Description automatically generated

**Query:**

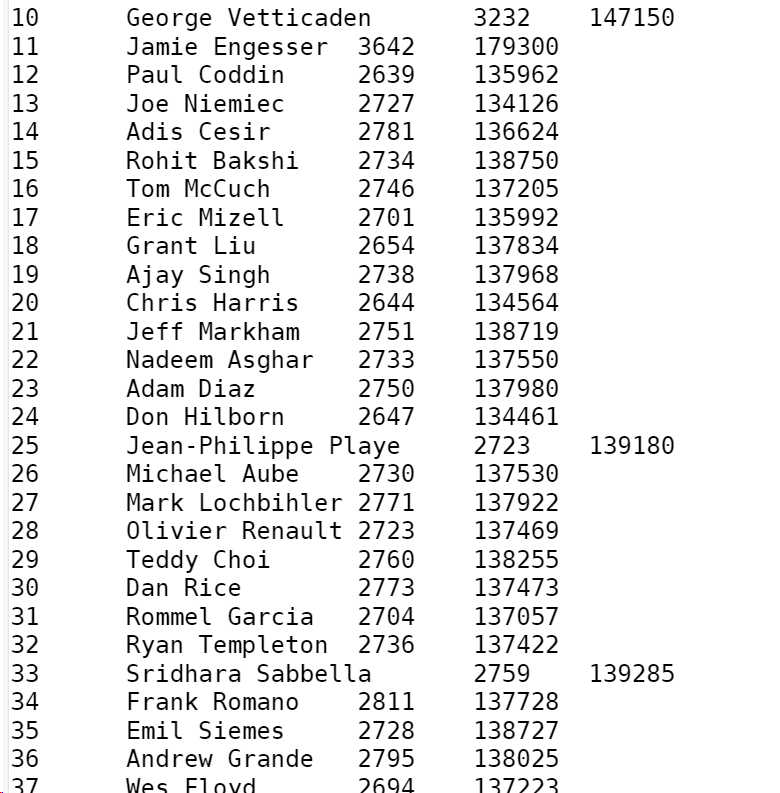
**Print the name of the driver in the above query**

hive> select t.driverid ,d.name, sum(hours\_logged), sum(miles\_logged) from timesheet\_n t join drivers\_n d on t.driverid

= d.driverid group by t.driverid,d.name;

A close up of a document

Description automatically generated



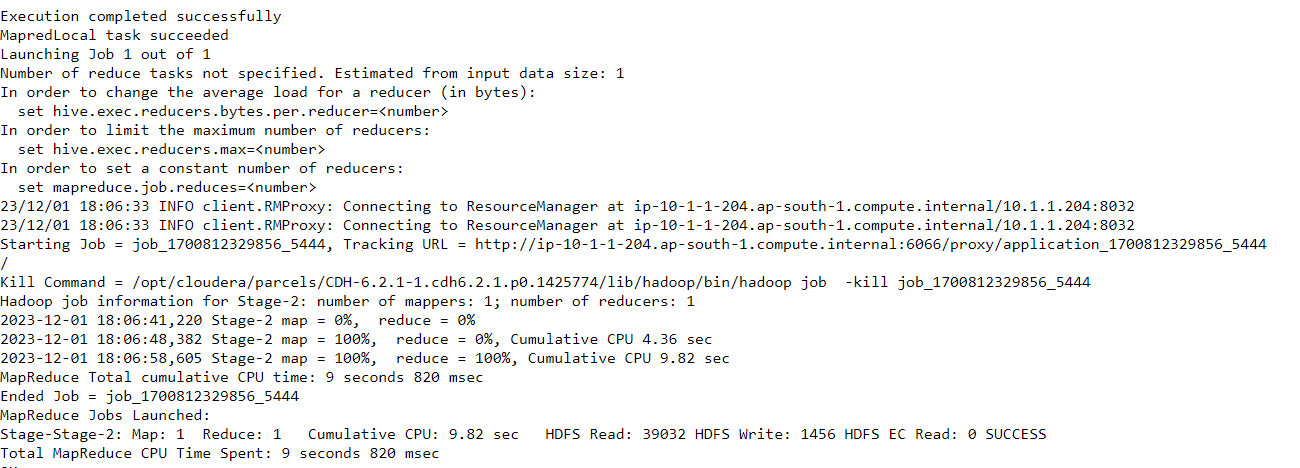
**Query:**

**Write the above query output in to hdfs directory**

hive> insert overwrite directory '/user/$USER/hive\_data' row format delimited fields terminated by ',' stored a

s textfile select t.driverid ,d.name, sum(hours\_logged), sum(miles\_logged) from timesheet\_n t join drivers\_n d on t.driv

erid = d.driverid group by t.driverid ,d.name;



A screenshot of a computer

Description automatically generated

A screenshot of a computer program

Description automatically generated