# Hive Assignment-2-Driver data Analysis

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

CREATE EXTERNAL TABLE drivers

(driverId INT,name STRING,ssn INT,location STRING,certified STRING,wage\_plan STRING)

ROW FORMAT DELIMITED FIELDS TERMINATED BY ','

LINES TERMINATED BY '\n'

STORED As TextFile tblproperties(

"skip.header.line.count"="1"

);

A screenshot of a computer code

Description automatically generated

CREATE EXTERNAL TABLE timesheet

(driverId INT,week INT,hours\_logged INT,miles\_logged INT)

ROW FORMAT DELIMITED FIELDS TERMINATED BY ','

lines terminated by '\n'

stored as TextFile tblproperties("skip.header.line.count"="1");

A screenshot of a computer code

Description automatically generated

Data load into tables

A close-up of a computer code

Description automatically generated

## Query: How many total hours and miles logged by each driver?

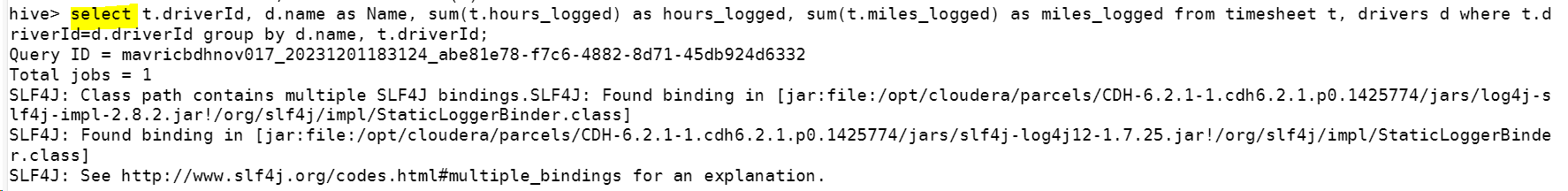
A computer screen with text

Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated

## Query: Print the name of the driver in the above query



A screenshot of a computer

Description automatically generated

## Query: Write the above query output in to hdfs directory

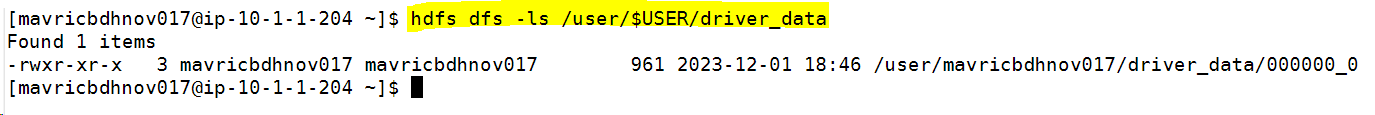
A close up of a document

Description automatically generated

A close-up of a computer screen

Description automatically generated

Query written into file in the hdfs directory



A screenshot of a computer

Description automatically generated