Create database driver\_sk\_db;

A black text on a white background

Description automatically generated

CREATE EXTERNAL TABLE driver\_sk\_db.drivers(

driverId INT,

name STRING,

ssn BIGINT,

location STRING,

certified STRING,

wage\_plan STRING)

ROW FORMAT DELIMITED FIELDS TERMINATED BY ','

LINES TERMINATED BY '\n'

STORED As TextFile

LOCATION '/user/mavricbdhnov0116/workspace'

tblproperties(

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

);

A screenshot of a computer program

Description automatically generated

CREATE EXTERNAL TABLE driver\_sk\_db.timesheet(

driverId INT,

week INT,

hours\_logged INT,

miles\_logged INT)

ROW FORMAT DELIMITED FIELDS TERMINATED BY ','

LINES TERMINATED BY '\n'

STORED As TextFile

LOCATION '/user/mavricbdhnov0116/workspace'

tblproperties(

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

);

A screenshot of a computer program

Description automatically generated

A white background with black and white text

Description automatically generated

A white background with black text

Description automatically generated

A screenshot of a computer

Description automatically generated

2. Print the name of the driver in the above query   
select t.driverId,d.name,sum(t.hours-logged) as total\_hours\_logged,sum(t.miles\_logged) as total\_miles\_logged FROM timesheet t JOIN drivers d on t.driverId = d.driverId group by t.driverId, d.name;

3. Write the above query output in to hdfs directory

INSERT OVERWRITE DIRECTORY  /user/mavricbdhnov0116/warehouse SELECT t.driverId,d.name,SUM(t.hours\_logged) AS total\_hours\_logged,SUM(t.miles\_logged) AS total\_miles\_logged FROM timesheet t JOIN drivers d ON t.driverId = d.driverId GROUP BY t.driverId, d.name;

A white screen with black text

Description automatically generated