Create external tables for drivers and timesheet data

CREATE EXTERNAL TABLE driver\_details (

    driverId INT,

    name STRING,

    ssn STRING,

    location STRING,

    certified BOOLEAN,

    wage\_plan STRING

)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ','

STORED AS TEXTFILE

tblproperties(

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

);

CREATE EXTERNAL TABLE timesheet\_entry (

    driverId INT,

    week INT,

    hours\_logged DOUBLE,

    miles\_logged DOUBLE

)

ROW FORMAT DELIMITED FIELDS TERMINATED BY ','

LINES TERMINATED BY '\n'

STORED As TextFile

tblproperties(

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

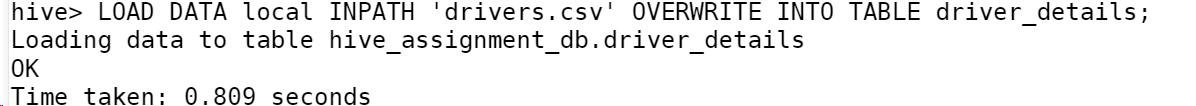
);

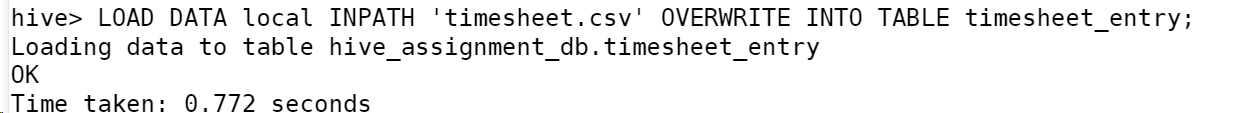
A white background with black text

Description automatically generated

LOAD DATA local INPATH 'drivers.csv' OVERWRITE INTO TABLE driver\_details;

LOAD DATA local INPATH 'timesheet.csv' OVERWRITE INTO TABLE timesheet\_entry;





A close-up of a white background

Description automatically generated

Q1.How many total hours and miles logged by each driver

SELECT

    t.driverId,

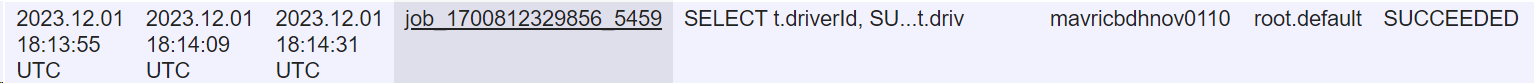
    SUM(t.hours\_logged) AS total\_hours\_logged,

    SUM(t.miles\_logged) AS total\_miles\_logged

FROM timesheet\_entry t

JOIN driver\_details d ON t.driverId = d.driverId

GROUP BY t.driverId;



A computer screen with black text

Description automatically generated

A white background with black text

Description automatically generated

Q2.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\_entry t

JOIN driver\_details d ON t.driverId = d.driverId

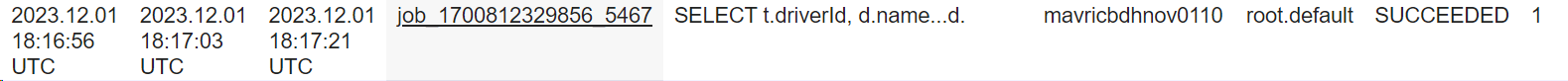
GROUP BY t.driverId, d.name;

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated



Q3.Write the above query output in to hdfs director

INSERT OVERWRITE DIRECTORY '/user/mavricbdhnov0110/driverdetailquery'

SELECT

    t.driverId,

    d.name,

    SUM(t.hours\_logged) AS total\_hours\_logged,

    SUM(t.miles\_logged) AS total\_miles\_logged

FROM timesheet\_entry t

JOIN driver\_details d ON t.driverId = d.driverId

GROUP BY t.driverId, d.name;

A close-up of a computer screen

Description automatically generated

