Teamwork assignment

Solve Problems 1 and 2 in a team.
Submit
    (i) SQL script;
    (ii) report - a Word document or pdf with query results.

- Submit via Canvas (one person submits on behalf of a team).
- If the solution is a large table, take a screenshot of just a part of it (make sure column title is visible).
- In the report, include the following note:
- "Prepared by: <list of names of team members who contributed to the solution >."

! Each problem should be solved using a <u>single</u> query.
! Your query should work independent of database instance..

Teamwork assignment

Suggestions on how to approach a teamwork

- Begin by discussing with your team how to approach problem-solving and report preparation.
- Create a shared file (e.g., Google Doc, Overleaf) for collaborative work and submission.
- You may distribute problems to be solved among members maintain open communication and assisting each other as needed. Reach out to team members (not professor or TA) when facing a problem.
- Once you complete your part, review solutions from other members—you may catch typos or errors.

Data

We will work with a <u>sample</u> from this dataset

Data description

Name	Description
VIN	The first 10 characters of each vehicle's Vehicle Identification Number (VIN). Thus, it is not a unique number here.
County	This is the geographic region of a state that a vehicle's owner is listed to reside within. Vehicles registered in Washington state may be located in other states.
City	The city in which the registered owner resides.
State	This is the geographic region of the country associated with the record. These addresses may be located in other states.
Postal Code	The 5-digit zip code in which the registered owner resides.
Model Year	The model year of the vehicle, determined by decoding the Vehicle Identification Number (VIN).

Data description

Name	Description
Make	The manufacturer of the vehicle, determined by decoding the Vehicle Identification Number (VIN).
Model	The model of the vehicle, determined by decoding the Vehicle Identification Number (VIN).
Electric Vehicle Type	2 types: Battery Electric Vehicle (BEV) and Plug-in Hybrid Electric Vehicle (PHEV).
Clean Alternative Fuel Vehicle (CAFV) Eligibility	This categorizes vehicle as Clean Alternative Fuel Vehicles (CAFVs) based on the fuel requirement and electric-only range requirement in House Bill 2042 as passed in the 2019 legislative session.
Electric Range	Describes how far a vehicle can travel purely on its electric charge.
Base MSRP	This is the lowest Manufacturer's Suggested Retail Price (MSRP) for any trim level of the model in question.

Data description

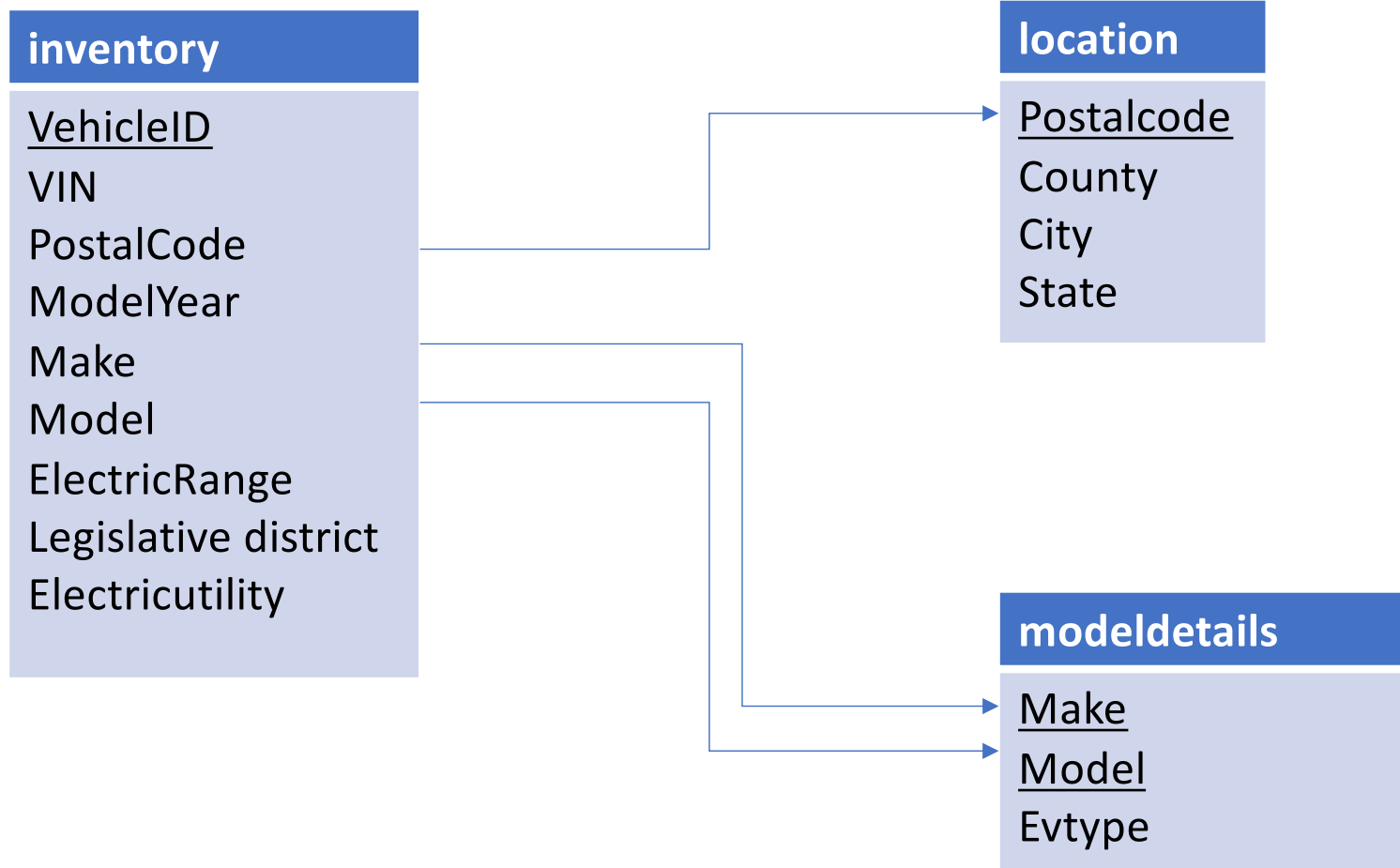| Name	Description
Legislative District	The specific section of Washington State that the vehicle's owner resides in, as represented in the state legislature.
DOL Vehicle ID	Unique number assigned to each vehicle by Department of Licensing for identification purposes.
Vehicle Location	The 5-digit ZIP Code for the registered vehicle.
Electric Utility	This is the electric power retail service territories serving the address of the registered vehicle.
2020 Census Tract	The census tract identifier is a combination of the state, county, and census tract codes as assigned by the United States Census Bureau in the 2020 census, also known as Geographic Identifier (GEOID). More information can be found here

inventory

VehicleID
VIN
PostalCode
ModelYear
Make
Model
ElectricRange
Legislative district
Electricutility

location

Postalcode
County
City
State

modeldetails

Make
Model
Evtype

Figure 1. Relational DB diagram

How to upload csv data

3 options:

1. **Use MySQL Blackboard wizard.** Right-click on the chosen table -> Table Data import Wizard -> select csv file -> use existing table -> match attributes -> Next

2. **Command: LOAD DATA INFILE** (https://dev.mysql.com/doc/refman/8.0/en/load-data.html). Example:
SHOW VARIABLES LIKE "secure_file_priv";
SELECT @@GLOBAL.secure_file_priv;
SHOW GLOBAL VARIABLES LIKE 'local_infile';
SET GLOBAL local_infile = 1;
#
LOAD DATA LOCAL INFILE 'yourpath/Inventory.csv'
  INTO TABLE WSEVdata.inventory
  FIELDS TERMINATED BY ','
  LINES TERMINATED BY '\n'
  IGNORE 1 LINES;

3. Convert CSV to INSERT statements.
Note: You have both csv files and insert statements for each table. Executing INSERT statements is easiest for you now, but when you have time – try other options too.

First, we implement a database in MySQL:

I. Create schema: execute "Schema" file (create schema).

II. Fill in the data: execute files "data_.sql". You can also try one of the above-mentioned approaches (for this you have csv files).

Problem 1

Problem 1. Write SQL queries for the following.

1. How many cars are there in our sample?
2. Which postal code has the largest number of electric vehicles (EV=both BEV and PHEV)? List postal code and associated number of vehicles.
3. Which battery electric vehicle (BEV)-type model is the most popular in our sample? List model and number of cars.
4. What percentage of EV vehicles is BEV and what is plug-in hybrid. Output of SQL query should be a percentage value!*
5. What is the most popular vehicles manufacturer (=make) in our sample (list name and percentage of market)?
6. What are min and max electric range values for each manufacturer (=make)?*
7. What are min and max electric range values for each model? Assume, different manufacturers may produce cars with same model name. For each model, also list associated make and EV type.
8. Which city has the largest number of EVs (EV=both PHEV and BEV)?

*Recall: use single query.

Problem 2

Problem 2. Write SQL queries for the following.

2.1 Create the view called EVNZ that selects those cars (list make, model, electric range, and EVtype) where electric range is not zero.

2.2 From the view EVNZ, find minimum, maximum, and average electric range for each types of electric vehicles (PHEV, BEV).*

2.3 Find minimum, maximum, and average electric range for each types of electric vehicles (PHEV, BEV)* Here, no condition on electric range (thus, do not use EVNZ).

2.4 Compare results from 2.2 and 2.3: at the output create the following table with 5 columns and 4 rows (+title row) as shown in the Table*

Problem	Type	Minimum range	Average range	Maximum range
2.2	PHEV			
2.2	BEV			
2.3	PHEV			
2.3	BEV			

Hint: in 2.4, use select '2.2' as "Problem, and select '2.3, ' to add the left column

*Recall: use single query.