INST 327 - Project Progress Report

Team 5

Group Members: Melchizedek Ackah-Blay, Nishat Chowdhury, Kassia Tefera, Kibron Tesfatsion, Shamariah Ogu

Final Project Report- Group 5

Introduction:

We designed a database for individual consumers who are interested in purchasing an electric vehicle (EV). Electric vehicles have been gaining significant popularity in recent years, due to factors such as environmental concerns, government incentives, rising gas prices, and an increase in accessible EV charging infrastructure. Our database allows consumers to have a centralized platform for consumers to access information about various electric vehicle features.

The state of Washington has developed programs and incentives for EV owners as a push toward clean and renewable transportation. Washington state residents are eligible to receive tax credits and exemptions (WA, 2019). According to the Washington State Department of Transportation, the state submitted a plan to increase EV infrastructure to 500,000 EV chargers. They state their motivation to accelerate the adoption of EVs, reduce greenhouse gas emissions, and help the U.S. lead global EV transportation efforts (WSDOT, 2023).

Our primary database user groups are Washington residents interested in purchasing an electric vehicle, as well as the state's Department of Transportation to gain insight into which areas of the state have successfully adopted the EV program. We hope that state residents will utilize our database to compare EV options in order to make a purchase that best suits their needs, while also contributing to the state's goal of cleaner transportation. Our database enables consumers to view and filter electric vehicles based on price, features, and the vehicle's eligibility for state-granted incentives.

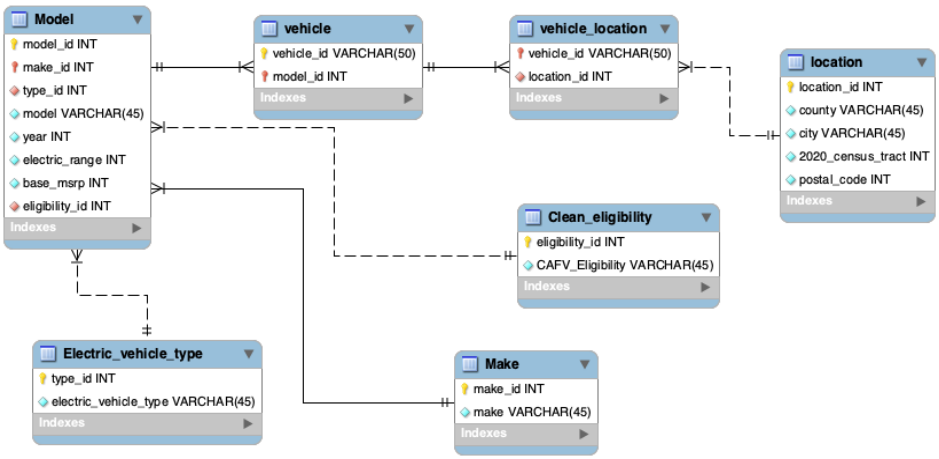
Database Description:

Our goal is to create a database containing all the key and important information related to the Electric Vehicle population in the state of Washington. Our potential database users are electric car consumers, environmental organizations, and even automotive companies seeking

insights. A secondary goal would be to enable access, analysis, and easy data visualization by creating a clean and structured database.

Our main table will be the “model” which includes all the features of a vehicle. The “model” table has a one-to-one relationship with the “vehicle” table, which uses the state-registered vehicle number ID to link each registered vehicle to its model type. The “model” table also shares a one-to-one relationship with the “electric_vehicle_type”, “make”, and “clean_elibility” tables through the model_id. The “model” table has a one-to-many relationship with vehicle_location, using “vehicle_id” in the “vehicle” table as an intermediate. The “model” table also has a many-to-many relationship with the “location” table, using “vehicle_location” as an intermediate through the “vehicle_id” foreign key. We also have the “make” table which has two columns “make_id” and “make.” The function of this table is to access the model of a vehicle in question using the “make_id.” Succeeding this is the “vehicle” table. It has two columns: "vehicle_id” and “model_id.” This table serves as a linking table for the “model” and “vehicle_location” tables. As part of our database, we have the “vehicle_location” table. It has two columns: “vehicle_id” and “location_id.” This again is a linking table for the “vehicle” and “location” tables. We also have the “electric_vehicle_type” table that has two columns, “type_id” and “electric_vehicle_type.” Each ID in the “type_id” column matches an electric vehicle type (Plug-in Hybrid or Battery Electric Vehicle). Moreover, we have a “location” table which has five tables. It is the biggest table in our database. The columns in the “location” table are “location_id,” “county,” “city,” “2020_census_tract,” and “postal_code.” It will be useful in the case that a potential customer wants the full details of where a vehicle is located.

ER Diagram:



Views/Queries

View Name	Req. A	Req. B	Req. C	Req. D	Req. E
model	x	x		x	
make	x	x		x	
electric_vehicle_type	x	x			
clean_eligibility					
vehicle	x	x	x	x	x
vehicle_loc	x	x	x	x	x

ation					
location	**x**	**x**	**x**	**x**	**x**

Changes from the Original Design:

Initially, we wanted our database to have about 500 rows on average per table. A benefit of this was that it would make our database more realistic and informative. However, we quickly realized that it was not feasible due to time constraints and the complexity of it. We have decided on having a maximum of 20 rows on average per table for a few reasons. Foremost, it is a more realistic and achievable goal. Secondly, it is easier to spot an error especially when doing JOIN operations. Additionally, it is more time efficient given that the project is due in about a week.

Secondly, we trimmed down our "potential" tables from nine to seven. We started with vehicle, location, postal code, city, county, clean eligibility, vehicle cost, legislative district, and census. After careful consideration, we amended our tables and ended up with seven, namely: clean eligibility, electric vehicle type, location, make, model, vehicle, and vehicle location. We removed the postal code, city, and county tables because they fall under the location category, and would be a waste of time and space to have them in separate tables. We modified the vehicle table from four columns to two (vehicle_id, model_id). The new vehicle table serves as a linking table for the model and vehicle_location tables.

Finally, we reframed the ERD model of our database to cater to individual consumers interested in buying electric vehicles. We initially were interested in having car manufacturers as our primary users, as they would gain insights into which of their vehicles are most popular in different areas of Washington state and where they might need to focus on marketing or advertising. We originally modeled our ERD with tables that provided information about the popularity of different models and makes based on location, as we were interested in providing business insights about where car manufacturers are most popular. However, after analyzing our dataset, we realized that the information is geared towards car features, which would be best used in the context of a customer looking to compare their options of which vehicle to purchase.

Lessons Learned:

During the process of our project, we were given a data set to extract from. We learned how to pick the most important and most useful information. We strategized which columns we needed to create for our database. We then furthered our understanding of normalization by going through the steps of first normal form, second normal form, and third normal form. Lastly, we learned how to create and use the entity relationship diagram.

Originally, when we created our database ethics section. We largely centered them on taking into consideration the socio-economic, environmental, accessibility, and data privacy factors surrounding electric vehicles (EVs). Due to the high prices that often come with EVs, we wanted to ensure that our database users would be able to see the prices of the cars so they could weigh which options were more financially affordable for them. Since then, we learned that creating a price-id table or column would seem unnecessary due to being able to view them through creating queries and that having a Base_msrp would better satisfy that need. Surrounding environmental factors, we also learned that creating entities/columns specifically for the climate, terrain, and elevation would make our database over-complicated and confusing to the users. This is why we decided to include columns like CAFV_Eligibility to consider environmental factors. Finally, we further learned the importance of centering data on accessibility (ex: helping users know which cars are most accessible to them based on location) and data privacy of users. This is why we stuck with including location-related entities/columns and excluding the DOL vehicle ID from our database.

We learned how to clean our CSV file by removing unnecessary columns and rows with the use of Python code. Our original data set had over 100,000 rows of data with many unnecessary columns that did not fit our normalized data. We utilized Python code to drop columns not needed and reduced the number of rows to a number more manageable, we then created a new CSV file titled “clean_electric_vehicle.csv”

One of the most important things we learned was how to create our database through Mysql, we learned how to add tables by creating Excel documents that represented our normalized data and how to import and convert that Excel document into an SQL database table.

Throughout this project, we learned how to work together and better collaborate to make sure we were on the same page. We worked together on completing the ERD as well as all of the tables. We learned that collaboration makes projects much easier and enjoyable.

Database Ethics Considerations:

There are socio-economic factors that should be taken into account such as the high costs of adopting EVs. Consumers from lower-income backgrounds who are interested in adopting EVs face financial barriers and it is important to consider this. We consider this by allowing consumers to view cheaper models by including the base MSRP in the "model" table. Users would be able to filter out more expensive car models by using the "base_msrp" column. Alternatively, they can decide to sort the vehicles from lowest to highest base MSRP and choose the right fit for them.

The database might also not consider the lack of charging infrastructure in more rural areas of the state. EVs in more rural areas may not have as much access to charging infrastructure. This could lead to their electrical range data being biased and the electrical range being lower than EVs in less rural areas. Additionally, our data may not include potential factors that can impact electrical range in EVs such as climate, terrain, and elevation.

When creating a user-oriented database, it is important to keep data privacy and other ethical considerations in mind. We will be removing the DOL vehicle ID which directly traces vehicles to their registration information. We feel that the removal of this data column would help enhance data privacy and safeguard the personal information of individuals associated with these vehicles. Other information such as vehicle models are included in the database. It is highly unlikely that a user can be tracked just by knowing their car model and no other attribute about that individual or their vehicle in question.

Potential Future Work:

In the future, there are multiple ways our group wants to move forward concerning our database. Some of the changes we want to continue to make are expanding our database to include more car specs and cars in general. This would give database users the option to view a more diverse range of cars and allow them to gain more information on features such as cost, performance, and comfortability. We also want to make changes such as classifying cars by

whether they are luxury or not luxury. This would allow users to gain more information on the features of each car and help them better distinguish which cars they prefer. It would also help database users see which cars are financially affordable versus cars that aren't as affordable. Additionally, we want to create a renting option for buyers. This would be more convenient for buyers who are in a specific location for a short time and need a car temporarily. Renting cars is also more affordable than buying cars. This would make it a much more financially affordable option for users who aren't able to afford to buy a car and want a cheaper option.

References

WA, S. (2019). *Electric Vehicle Incentives*. Puget Sound Energy EV Buyers Guide. https://ev.pse.com/incentives#:~:text=Electric%20Vehicle%20Incentives,including%20rebates%20and%20tax%20credits.&text=Individuals%20who%20purchase%20qualified%20residential,credit%20of%20up%20to%20%241%2C000.

WSDOT, W. (2023, November 21). *Washington State Plan for Electric Vehicle Infrastructure Deployment*. https://wsdot.wa.gov/construction-planning/statewide-plans/washington-state-plan-electric-vehicle-infrastructure-deployment