**Introduction**

Visualizations have a powerful impact on real-world decisions that consumers, policy-makers, and companies make. To test whether the aspects of a visualization that were chosen are clear and effectively communicate the data without bias or confusion, I created two visualizations based on the Electric Vehicle Population data and presented them to participants. Criteria were established to see if participants were able to clearly understand what the visualizations were communicating and if they could then make a real-world decision based on the trends from these visualizations. In this report, I summarize my data, goals, tasks, and elements of my design, as well as evaluate my approach to determine if my visualizations successfully communicated the data effectively based on my discussion with participants. Finally, I synthesize my findings and communicate future improvements to refine my visualizations.

**Recap**

The dataset that I chose to work with was the Electric Vehicle Population Data which details aspects of Battery Electric Vehicles and Plug-in Hybrid Electric Vehicles that are registered in the state of Washington, United States. Some key attributes of the data show Vehicle Information (Make, Model, Model Year, Electric Range, etc.,), Location Information (County, City, State Vehicle Location, etc.,), and Utility Information (Electric Utility). This dataset was cleaned to remove null values and two columns, ‘Legislative District’ and ‘Electric Utility’ were removed because of the large number of null values in these columns. The key dimensions of the data that could be used for my analysis were Geography (State, County, City, Zip Code, etc., ), Time (Model Year), and Vehicle Attributes (Make, Model, Electric Vehicle Type, Electric Range, and CAFV Eligibility).

**Goals**

I set out to perform my analysis on effective visualizations by attempting to answer two specific questions based on the data. My first question: What is the average electric range between Electric Vehicle Models by Battery Electric Vehicles and Plug-In Hybrid Electric Vehicles? My second question is: How does the EV type (Battery Electric Vehicles or Plug-in Hybrid Electric Vehicles) vary across counties? Using these two questions, I set out to create visualizations that would communicate the results of these questions effectively and test whether these visualizations can directly impact consumer or policy-makers decisions related to Electric Vehicles and infrastructure related to Electric Vehicles.

**Key Elements of the Design**

The first visualization explores the question: What is the average electric range between Electric Vehicle (EV) Models by Battery Electric Vehicles and Plug-In Hybrid Electric Vehicles? Displays a stacked bar chart. The X-Axis is the Average Electric Range measured in miles. The Y-Axis displays the model which is sorted in descending order by range. When the participants hover over the bars, they can see the Make, Model, Average Range, and Type of Electric Vehicle (Battery or Plug-In Hybrid) so that it is easier to read and compare Models by their range. The Type of Vehicle is also differentiated by color, Blue for Battery Electric Vehicles and Orange for Plug-In Hybrid Vehicles. A stacked bar chart was chosen to model the average electric range between different EV Types and Models because it is a simple design that clearly communicates the data and by making the two charts stacked on top of each other, participants are also able to compare the two types of Electric Vehicles between each other in addition to the models within each group. Other visualizations that were considered was a scatter plot but this was ultimately ruled out because it is harder to compare many models and will be harder for participants to interpret and obtain trends.

The second visualization explores the question: How does EV type (Battery Electric Vehicles or Plug-in Hybrid Electric Vehicles) vary across counties? Displays a grouped bar chart. The X-axis for this visualization displays the Number of Vehicles within that county and the Y-axis displays the County. This visualization only shows data for counties that have more than 1000 Electric Vehicles. This was done because of the huge disparity in the number of electric vehicles between some counties compared to others. The Type of Vehicle is also differentiated by color, Blue for Battery Electric Vehicles and Orange for Plug-In Hybrid Vehicles. Participants can interact with the bar chart by hovering over each bar which will display a tool tip that details the exact number of vehicles in that county by Type of Electric Vehicle. A grouped bar chart was chosen for this visualization due to the ease which it offers to participants to directly compare the amount of Battery Electric Vehicles and Plug-In Hybrid Electric Vehicles. Other options that were considered for this visualization were a choropleth map or bubble map but this was ruled out as it is harder for participants to show the two types of Electric Vehicles in one map and for bubble maps, people are worse at distinguishing the size of bubbles, especially as there would be many overlapping bubbles. The grouped bar chart was furthermore chosen over other types of visualizations because we are exploring the distribution differences across counties in Washington and geography is a secondary feature. In addition, a grouped bar chart was chosen over a bar chart that shows each bar per county side by side to reduce clutter and a long graph.

**Evaluation**

To evaluate whether the two visualizations are effective at communicating the trends associated with the two questions, I recruited 3 participants. Due to the difficulty in recruiting experts, these participants were between the ages of 25-30 and were interested in buying an electric vehicle in the near future.

For the first question, I conducted a Think-aloud study where the three participants were given the visualization to interact with and provide feedback on. They were also given three tasks for this visualization which included the following: 1) Identify the Electric Vehicle Model that had the highest and lowest range for both Battery Electric Vehicles and Plug-In Hybrid Electric Vehicles, 2) Find and compare trends between Battery Electric Vehicles and Plug-In Hybrid Electric Vehicles. 3) Make a real-world decision on which Electric Vehicle model they would consider purchasing based on range and explain their reasoning. The measures that were used for this think-aloud study are insight depth, accuracy, and use-case relevance as detailed in the criteria used to determine visualization success. Criteria that would indicate that the visualization successfully communicates the data clearly and effectively were defined as follows: 1) All of the participants were able to identify which EV model had the highest and lowest range in both the Battery Electric Vehicle group and the Plug-In Hybrid Electric Vehicle group, 2) All participants can state at least one trend by comparing the Battery Electric Vehicles and Plug-In Hybrid Electric Vehicles, 3) All participants can reason why they would choose the model that they choose to purchase based on the visualization. Further criteria to indicate that the visualization is successful at communicating the results would be that there is minimal (less than two items identified) to no confusion among the participants.

The results of the think-aloud study indicated that the visualization was a success. All three participants were able to correctly identify the EV model that had the highest and lowest range in each group (the Battery Electric Vehicles and Plug-in Hybrid Electric Vehicles). All three participants were also able to identify a trend by comparing the graphs for the Batter Electric Vehicles and the Plug-In Hybrid Electric Vehicles. A trend that all three participants identified was that the Plug-In Hybrid Vehicles had a dramatically lower Electric Range compared to the Battery Electric Vehicles. Lastly, all three participants were able to decide on an electric vehicle they would consider purchasing based on range. Two participants stated they would consider buying a Battery Electric Vehicle within the top 15 electric vehicles because there is a major difference between the top 15 Battery Electric Vehicles and the other Battery Electric Vehicles for the range they offer. The third participant was interested in a Plug-In Hybrid Electric Vehicle rather than a Battery Electric Vehicle and therefore they considered the range of the top 10 Plug-In Hybrid Electric Vehicles. The third participant noted that they were interested in the Prius Prime before seeing this graph but after seeing the range they could consider a different model that offers a better range that was shown in the top ten of the visualization. There was no confusion they indicated from the stacked bar graph but recommended that the graphs be side by side instead of stacked on top of each other for better comparability. Another useful critique I received was that there seemed to be a larger variety of Plug-In Hybrid models compared with Battery Electric Vehicles making the real-world decision more difficult than if the number of models between the two types of Electric Vehicles were more similar in size. The last suggestion for improvement was to incorporate the price points of each model to make the real-world decision more realistic because, in the real world, people would consider more factors than just range. Overall, this visualization is considered a successful, clear visualization as it effectively communicates what it is supposed to communicate in regards to the range and the question of exploration: What is the average electric range between Electric Vehicle (EV) Models by Battery Electric Vehicles and Plug-In Hybrid Electric Vehicles? But further improvements and additional features can be added.

For the second question, I conducted a semi-structured interview where three participants were recruited to act as policy-makers who were evaluating the visualization to decide on infrastructure within their state. The three participants were asked a series of questions to determine whether the visualization was successful in communicating the data. The first question was ‘What trends between Plug-In Hybrid Vehicles and Battery Electric Vehicles can you identify based on the visualization?’, followed by the question, ‘What trends between counties can you identify based on the visualization?’ Then, the participants were asked, ‘Would this visualization lead you to make any changes to the current infrastructure and explain your reasoning why or why not?’ The next few questions focused on improvements for the visualization, ‘Is there any information that is missing or needed to make a real-world decision based on the data?’ and ‘Was there any confusion from the visualization that could be improved?’ to determine if this visualization is a success the following criteria was used: Participants can correctly identify at least one trend for questions 1 and 2, Participants can make at least 1 recommendation for infrastructure changes and can use features of the graph to justify why they would or would not, and lastly, participants identify less than two items that they believe is missing to make a real-world decision and there are less than two items identified that is deemed confusing in the visualization. From these questions, we can gain insight into depth and identify use-case relevance as measures for determining whether the visualization is successful or not.

The results of this semi-structured interview indicate that the visualization needs improvement. For the first two questions, all three participants were able to correctly identify trends between Plug-In Hybrid Electric Vehicles and Battery Electric Vehicles as well as trends between counties. All three participants noted that there were more Battery Electric Vehicles than Plug-In Hybrid Electric Vehicles and that there is a large difference in the number of Electric Vehicles in King County compared to other counties. Other participants noted that there are more Plug-in Hybrid Electric Vehicles in King County alone than the amount of Electric Vehicles altogether in all counties listed on the graph except for one, Snohomish County. For the third question where participants were asked if they could make a real-world decision and explain why or why not, one participant stated that they would consider adding more charging stations where there are already a lot of Electric Vehicles, another participant stated the opposite, to add more charging stations where there is not a lot of electric vehicles so that people will buy more electric vehicles. Lastly, the last participant said they could not recommend changes in infrastructure because they need more information such as if the counties that have a lot of electric vehicles are close to each other and what kind of area that county is made up of such as an urban area compared with a rural area where electrical vehicles are not efficient. The other two participants also wanted more information to further support their decision such as how many charging stations already exist and if the people in those counties support infrastructure changes related to electric vehicles. The participants also expressed that they were missing information such as the title not including that the graph only shows information for Counties that have more than 1000 Electric Vehicles registered in that county (they were explicitly told this information when presenting the visualization). They also feel that the information was hard to read for counties that had less than 10,000 electric vehicles registered and would have liked that information presented differently than a grouped bar chart. Due to these reasons, I concluded that the visualization was not sufficiently clear enough and did not communicate the data as efficiently as it could. More information needs to be incorporated such as the number of charging stations to make a real-world decision about future infrastructure.

**Synthesis**

The participants that I recruited revealed insightful details about my visualizations that could be used to further improve my design choices as well as what worked well to communicate the data effectively.

For the first visualization, the stacked bar chart which displayed the average electric range by model and Electric Vehicle Type was for the most part successful. The participants were able to accurately identify trends between Battery Electric Vehicles and Plug-In Hybrid Electric Vehicles and make real-world decisions based on the data presented in the visualization. The design elements that were employed, were the simple layout with the data organized in descending order, the color encodings, and the ability to interact with the details from the addition of the tooltip. The stacked design allowed participants to compare the two electric vehicle types effectively and make a clear real-world decision based on only the range factor. However, there are possibilities for improvement because real-world decisions for choosing a car do not typically depend on the range as participants pointed out. In future iterations of the design, I can adjust the comparability and improve it by placing the graphs side by side as well as combining data about the price of the Electric Vehicle which could be more realistic for decision-making.

The second visualization which explored Electric Vehicle type distribution across Counties in Washington, needs further adjustments. This participant semi-structured interview revealed limitations with the visualization such as the need for more information about the Counties to make any real-world decision. Other limitations included the difficulty of reading data in counties that had very few registered Electric Vehicles because the scale was dramatically increased because of one county that had many Electric Vehicles compared to all other counties. In future iterations, it would be beneficial to add additional information to the data such as population size per county, and type of county (Urban versus Rural).

In summary, the design choices were effective in presenting trends, but participant feedback shows the importance of iterations to improve visualizations to be optimal. Future iterations will increase the comparability and add supplemental information to increase the ability to make real-world decisions.