Electric Vehicles in Washington State

Isabel Heard

Introduction

In recent years, the electric vehicle (EV) landscape has undergone a transformative shift. With more people and companies taking interest in adopting a more sustainable way of driving. Washington state, known for their commitment to environmental initiatives, served as the focal point of my analysis. The analysis seeks to understand the popularity of different makes and models, and how companies have improved their mileage on the cars over time. Furthermore, look into the correlation between EV prevalence and median household income across counties. Based on this information, this work seeks to determine the occurrence of EV cars in Washington state, and their usage amongst specific counties.

Dataset

Two datasets are used in this analysis. The first dataset is Electric Vehicle Population from the open data portal for the state of Washington. This dataset includes information on Battery Electric Vehicles (BEVs) and Plug-in Hybrid Electric Vehicles (PHEVs) that are currently registered through Washington's Department of Licensing. The dataset contains 159K rows and is updated as of October 31, 2023. It includes the variables: VIN, County, City, State, Zip Code, Model Year, Make, Model, EV Type, CAFV Eligibility, Electric Range, Base MSRP, Legislative District, DOL Vehicle ID, Vehicle Location, Electric Utility, and 2020 Census Tract.

The second dataset used is Median Household Income from the open data portal for Pierce County, WA. This dataset contains all of the median household income information by county in Washington State, from 2011 to 2021. There are 1,965 rows in this dataset. It includes the variables: Name (county name), Variable (median household income), Value, Year, Geography ID, Geography Name, Geography Type, ACS Year Estimate, Location, and Date.

Methods

Before diving into how each question was answered, it is important to discuss the data cleaning and exploratory data analysis portion of this report. When addressing the issue of missing data, I luckily only have a few missing values, and they ended up being in variables that went unused in analysis. For the EV dataset, I had to subset for only cars in Washington, because there were a few cars that were registered in different states, and I was only concerned with the ones registered in WA. For the income dataset, I had to clean up the county names so that I

would be able to merge them with the EV dataset. This way, county names were identical in both.

EV Popularity. Firstly, I wanted to find out which make, and model were most popular. To do this, I created a table for both variables and then counted how many there were of each unique make and model. I then sorted by the top ten of each, as to not have an overwhelming graph. Next, using ggplot, I made a bar chart, where the y-axis would be the count frequency, and the x-axis would be the type of make/model.

County EV Population. I was interested in which counties had the most EVs. To do this, I created a table with only county and number of cars. Then I counted each row in the new data set and sorted it by county. Like before, I only wanted to look at the top ten counties, as to not overwhelm the eye. With this I took the top ten counties and used ggplot to create a bar chart, where the y-axis was the count frequency of cars, and the x-axis would be the counties.

Proportion of People who have EVs. I felt that it would be interesting to look into what proportion of people who have EVs in the top three counties I found in the previous graph. With that, I used census data to look into the total population of those three counties as well as the whole state for 2021. The reason for 2021, is that the year I am using for median household income is 2021, and I want to keep things consistent. I used the subset function to get each county/state in their own data frame. Then I counted each row to find the number of cars in each. I then calculated the proportion of people to electric car ownership. After some thought, I decided it would make more sense to calculate the ratio of EVs to the total number of registered vehicles in WA. Obtaining an accurate total was challenging due to variations in ownership (private, public), and vehicle types (buses, trucks, motorcycles). Nevertheless, I managed to arrive at a rough estimate. I think both proportion calculations can be meaningful. It depends if you are interested in understanding the market share of EVs among all vehicles, or EV adoption rates among the entire population.

Median Income vs Counties. I wanted to create an interactive plot that shows median household income in the top ten most EV prevalent counties, along with the EV count. To do this I had to load in the plotly library to make it interactive. First, I selected my data to only have the top ten counties I wanted to look at. Then I filtered my data by grouping by county and getting the EV count. I then filtered the income data to only look at the year 2021. I needed to use a left join to merge the data by county name. Then, made the plot, with median income on the y-axis,

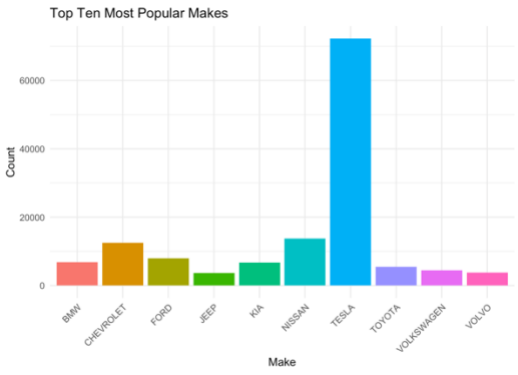
county on the x-axis, and when you click on the bar associated with a county, it will tell you how many registered EVs there are, along with the specific dollar amount for the income.

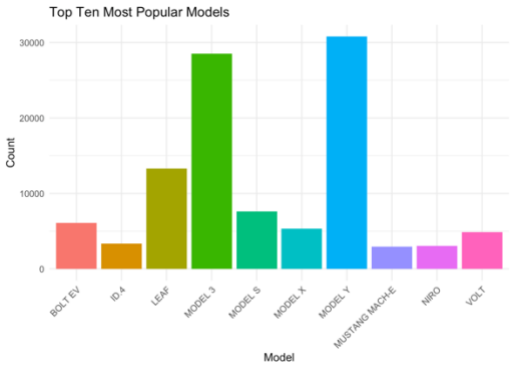
Electric Range Over Model Years by Make. From my previous analysis, I was able to find the top five companies that produced the most EV cars. I wanted to make an interactive plot to show how their mileage span has changed over the years. I had my x-axis be the model year, and the y-axis be electric range. Each company was given a different color to tell them apart on the plot. Then, when you select a data point, it will tell you what model that car is.

Bonus: Map of Nissans in Pierce County. Lastly, I wanted to see if I could plot any of these cars on a map. The data set came with longitude and latitude points, but I had to manipulate the Vehicle.Location variable in order to separate the two. I was able to complete this task by including a regular expression operation. Now that I had my longitude and latitude variables, I was able to load the library leaflet, which helps create interactive web maps. I had my data only include Nissan EVs in Pierce County, as any more data would take quite a while to run. Overall, for this, I just wanted to play around with the package and see what the output would look like. I figured since I had longitude and latitude data, I would give it a shot. When you click on a data point on the map, it will tell you the car model, year made, city it is registered in, and electric range.

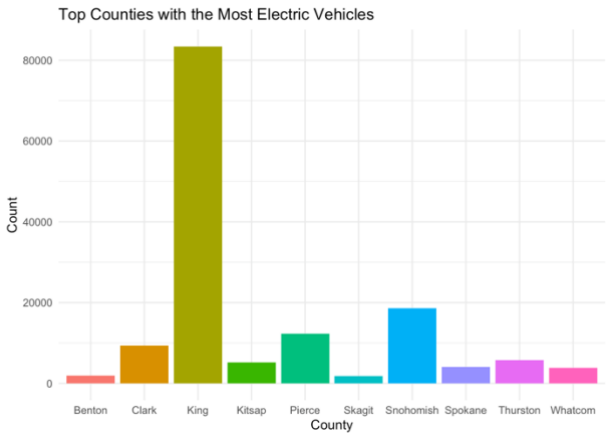
Results

EV Popularity. Through this bar graph, we are clearly able to see that Tesla, Nissan, and Chevrolet have the three highest number of cars registered in Washington. Tesla, has roughly 72K, taking up most of the market, while the next highest is Nissan with 13K. We are also able to see that the most popular models mostly include either Tesla's (Model Y and Model 3) or Nissan's (Leaf) cars. In future work, it would be interesting to see if the most popular models were the cheapest on the market.



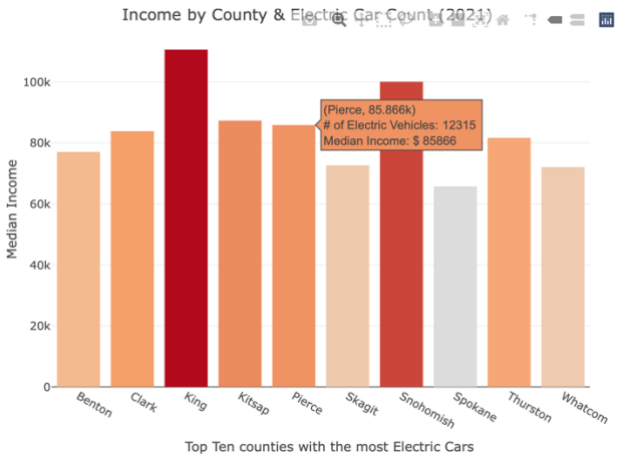


County EV Population. Looking into which counties have the most registered EVs. King County clearly outnumbers the rest with 83K. Snohomish at 18K and Pierce County at 12K. I think that the obvious answer for this is that King County is the most populated county, as well as the wealthiest. Also, being that it is heavily populated, there is a lot of accessibility to charging stations/not worrying about whether or not you have enough miles to get somewhere. Some of the lower counties in the visualization contain more remote land, and therefore a higher chance of your EV running out of electricity.



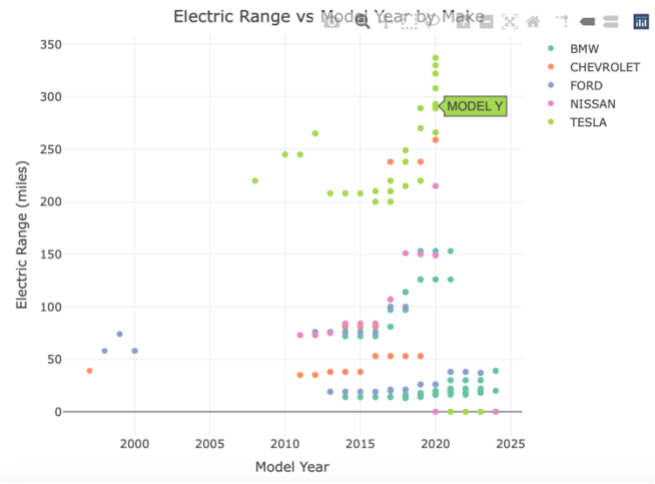
Proportion of People who have EVs. Based on the total number of people who live in King County, the proportion of those people who have EVs is 3.7%, meaning there are 83,413 EVs. Now this could mean that 83K people drive EVs, only if you assume that for every one car, there is only one person driving it. So, technically, a larger part of the population could "have" an EV if it is shared amongst a household/more than one person. For Snohomish, 2.2%, Pierce, 1.3%, and Washington as a whole, 2.0%. Based on how many cars are registered in Washington, 1.99% of those cars are electric. Meaning that there are roughly 160K electric cars registered in Washington.

Median Income vs Counties. Next, we have the interactive plot showing median household income by top ten counties with EVs. When you interact with each bar chart, it will tell you the number of electric vehicles and the specific median income. Here we can see that King and Snohomish County have the highest median household income, which also follows the other bar chart for most EV's. Kitsap County has the next highest income at 87K, yet they have significantly less EVs than Pierce County, who is fourth for median income. I think from this

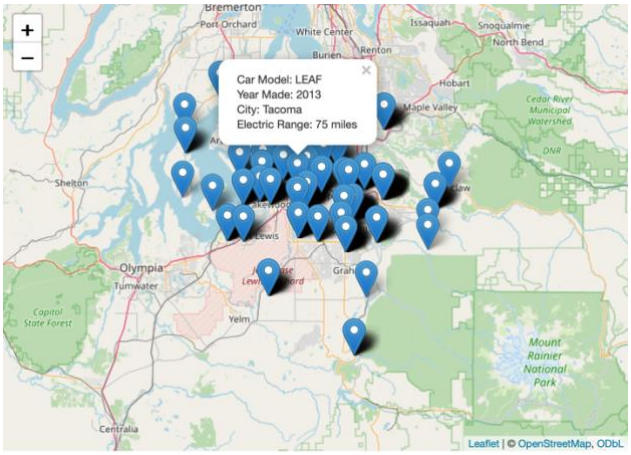


plot, you can conclude that median household income does have some correlation with number of EVs.

Electric Range Over Model Years by Make. Looking at the top five manufacturers, I wanted to see how their electric mileage improved over time. Each dot represents a different model type, and the color of the dot represents the manufacturer. We can see that Tesla has the best mileage by far, hovering between the 200 and 350 range. For all of the companies you can see an upward trend for the mile range. Which makes sense, as you want to improve the range of your car over time. Some of the newer cars have zero for mileage range. According to the data source, this is because new BEVs that have an electric range of 30 miles or more will be entered as a zero if their range has not been researched. Also notable in this visualization is that many of the cars in the 0 – 80-mile range are hybrid cars. So, many of the BMW, Ford, and Chevrolet cars are in this range because they produce mainly hybrid cars.



Bonus: Map of Nissans in Pierce County. Lastly, this graph is more for just showing myself that I could plot some points, there is not much analysis that the map can give us at this point. Yet, it can show us that more cars are, as expected, in more populated areas. As you get further away from the cities, the odds of there being electric vehicles are lower. I would have liked to add more types of cars or spread my map for the whole state, but my computer would not allow. So, if I had more time, I would potentially consider using the sf package in R to see if that gives me better results.



Conclusions & Future Work

According to the results of the electric vehicle analysis in Washington, Tesla makes up most of the EV population. This makes sense as they were the first company to mass produce electric cars with a large electric range on them. Yet, there are many companies who have started producing EVs as well. So, I would imagine that soon enough, Tesla will not have as large of a market share in a few years. When we look at how companies have improved their mileage range, it is clear that other companies are working to improve their cars.

With the county analysis, it shows us that most people who own an electric car live in a populated area, usually near cities. This makes sense, as having an EV in the countryside, where there may not be as many charging options, and you have to drive longer to get from point A to point B, and EV may not be your best option. And the counties that have a higher median household income, will tend to have more electric cars. This makes sense, as EVs are not going to be the cheapest option on the market.

If I had more time with this project, I would spend more time working on the map and making it have more functionalities. Such as, different colors for different models. Or figuring out how to have the capacity to model across more than one county and car type. Also, I would have liked to create a shiny dashboard where the user could change the make, model, to see how many cars were in a specific county or city. Or create a place for the user to look for models that have more than say a 200-mile range on the car. The dashboard would sort of pose as a search bar for EVs.

Furthermore, I would have liked to get data on how much money these cars cost when they were bought, so that I could look into the trend of spending on EVs. My guess is that they would become a bit more affordable as more companies are making EVs, but it would be interesting to test that theory. I would also like to expand this analysis to the country as a whole, so that way analysis could be done on a state level. It would be interesting to see which states are at the forefront of adopting EVs and what percentage of the people have an EV compared to a regular car. With additional data, I think it would also be noteworthy to see if EVs really are better for the environment, and if the adoption of EVs will help with climate change.

Appendix

Code and output can be found in the html file turned in.