Electric Vehicles Market Analysis

Introduction

The surging demand for electric vehicles (EVs) is driven by their efficiency, reduced environmental impact, and declining production costs. As a result, consumer preferences are shifting towards EVs, prompting substantial investments in research and development to enhance their performance and durability. To foster wider adoption of eco-friendly vehicles and shape the future of transportation, it is essential to comprehend market dynamics and identify opportunities for improving product features, expanding distribution channels, and minimizing production expenses.

Market analysis is a crucial skillset in the data analysis career, which is in high demand due to the rapidly evolving automotive industry landscape. Undertaking this project will enable me to gain a comprehensive understanding of market analysis and learn various processes involved in it. Through this project, I can apply analytical techniques and tools to evaluate the demand and supply of EVs, identify consumer preferences and behaviours, and determine market trends and opportunities. This data-driven approach will help me make informed decisions that can increase the affordability and accessibility of EVs to a broader consumer base.

The growth of the EV market presents an exciting opportunity for data analysts to apply their skills and knowledge to facilitate the adoption of eco-friendly transportation. By undertaking this project, I can develop my market analysis capabilities, enhance my career prospects, and contribute to shaping a sustainable future.

Languages used: Python, R, SQL and pyspark.

Literature Survey

The rising market share of electric vehicles (EVs) is a well-documented phenomenon, with various factors contributing to its growth, including government incentives, declining battery costs, and heightened consumer awareness of the environmental advantages of EVs. In 2021, global sales of EVs surged to an unprecedented 6.5 million units, a substantial increase from a mere 500,000 in 2015, with China being the largest market, accounting for over 50% of all EV sales. The United States and Europe are also significant markets for EVs, with sales increasing rapidly in recent years.

Upon conducting a literature survey, I came across a study conducted in Germany that aimed to identify the possible market share of EVs in the new car fleet, considering various scenarios from 2010 to 2030. The researchers found that the success of alternative powertrain vehicles was contingent on external factors such as purchase price incentives, rising oil prices, and low energy costs, and suggested that original equipment manufacturers (OEMs) could stimulate the market by accepting negative mark-ups during the first ten years after market introduction.[1]

Another relevant work I encountered was an investigation of electric vehicles' economic perspective for the future electric market, which revealed that advancements in smart grid technology and charging station flexibility have significantly contributed to the growth of the EV market. However, the availability of charging infrastructure remains a challenge, which is being addressed through vehicle-to-grid (V2G) and vehicle-to-infrastructure (V2I) solutions. The study also emphasized the importance of policy measures to educate consumers on the total cost of ownership (TCO) of EVs based on individual preferences and improve charging infrastructure and battery costs.[2]

Furthermore, I discovered a study that employed a probabilistic simulation model to compare the total cost of ownership (TCO) of EVs with conventional vehicles. The results showed that EVs are more cost-efficient for smaller vehicles and longer driving distances, with the TCO being subject to uncertain vehicle and operating costs. The authors suggested that even if the TCO of EVs becomes comparable to conventional vehicles by 2025, consumers' purchasing decisions are not solely based on TCO, necessitating policy measures to improve charging infrastructure and battery costs and educate consumers on TCO based on individual preferences.[3]

Overall, these works provide valuable insights into the various factors affecting the EV market's growth and the challenges that need to be addressed to increase its adoption further. They also offer suggestions for policy measures and strategies that could be implemented to improve the infrastructure and promote the use of EVs in different market segments.

Dataset

Electric Vehicle Population Data is a monthly-updated dataset provided by the US Department of Energy's Alternative Fuels Data Centre. It provides information on the number of registered electric vehicles (both BEVs and PHEVs) in the US, along with their make, model, year of manufacture, and location. This dataset is a valuable resource for researchers, policymakers, and industry stakeholders to understand the growth and adoption of electric vehicles in the US and inform future strategies to promote their use.[4]

Metadata: The dataset consists of 124,716 observation and 17 columns.

VIN (1-10): The 1st 10 characters of each vehicle's Vehicle Identification Number (VIN).

County: The county in which the registered owner resides.

City: City in which the registered owner resides.

State: State in which the registered owner resides.

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).

Make: Manufacturer of the vehicle.

Model: Model Name of the mentioned vehicle.

Electric Vehicle Type: Type of the vehicle present in the dataset.

Clean Alternative Fuel Vehicle (CAFV) Eligibility: Clean Alternative for the data present in this dataset.

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.

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**.**

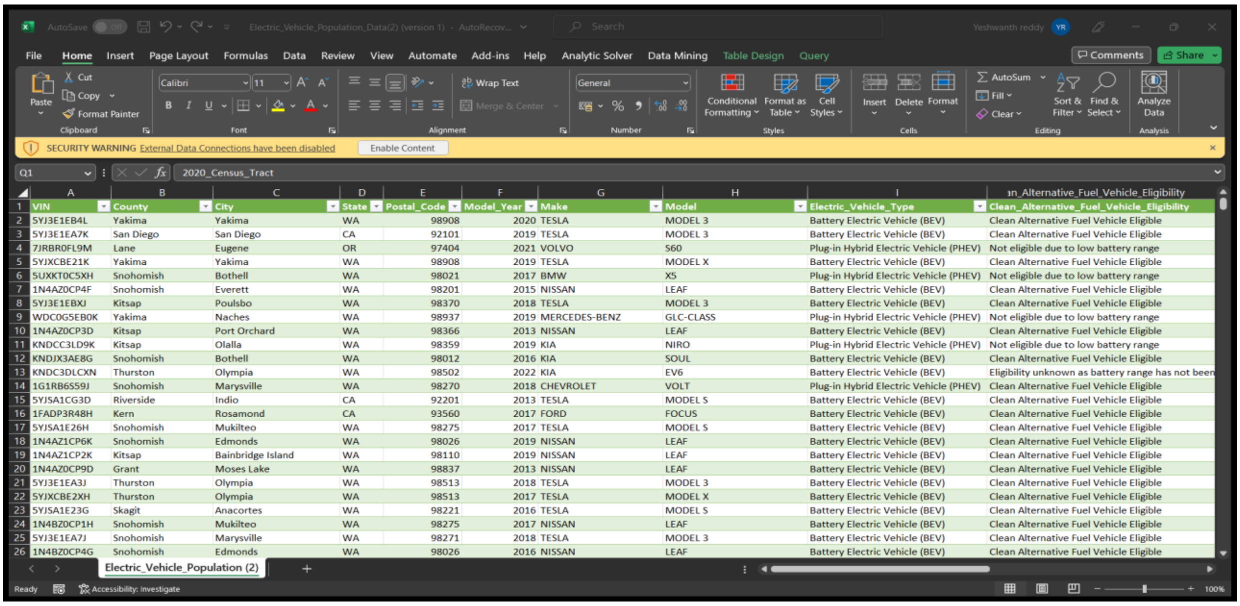
Vehicle Location: The centre of the ZIP Code for the registered vehicle.

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.

Data Cleaning

The First step before performing operation on any dataset is to make sure the dataset is clean if not, we must clean the dataset. Before loading the dataset into any programming language, I have loaded the dataset into Excel and have changed the names of the columns in the dataset. I have added an underscore when there is a space in the column name. For example, Electric Utility to Electric_Utility to avoid further Issue with the dataset.



	A	B	C	D	E	F	G	H	I	an_Alternative_Fuel_Vehicle_Eligibility
1	VIN	County	City	State	Postal_Code	Model_Year	Make	Model	Electric_Vehicle_Type	Clean_Alternative_Fuel_Vehicle_Eligibility
2	5YJ3E1EB4L	Yakima	Yakima	WA	98908	2020	TESLA	MODEL 3	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
3	5YJ3E1EA7K	San Diego	San Diego	CA	92101	2019	TESLA	MODEL 3	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
4	7JRBR0FL9M	Lane	Eugene	OR	97404	2021	VOLVO	S60	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range
5	5YJXCBE21K	Yakima	Yakima	WA	98908	2019	TESLA	MODEL X	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
6	5UXKT0C5XH	Snohomish	Bothell	WA	98021	2017	BMW	X5	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range
7	1N4AZ0CP4F	Snohomish	Everett	WA	98201	2015	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
8	5YJ3E1EBXJ	Kitsap	Poulsbo	WA	98370	2018	TESLA	MODEL 3	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
9	WDC0G5EB0K	Yakima	Naches	WA	98937	2019	MERCEDES-BENZ	GLC-CLASS	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range
10	1N4AZ0CP3D	Kitsap	Port Orchard	WA	98366	2013	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
11	KNDCC3LD9K	Kitsap	Olalla	WA	98359	2019	KIA	NIRO	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range
12	KNDJX3AE8G	Snohomish	Bothell	WA	98012	2016	KIA	SOUL	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
13	KNDC3DLCXN	Thurston	Olympia	WA	98502	2022	KIA	EV6	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not been
14	1G1RB6S59J	Snohomish	Marysville	WA	98270	2018	CHEVROLET	VOLT	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible
15	5YJSA1CG3D	Riverside	Indio	CA	92201	2013	TESLA	MODEL S	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
16	1FADP3R48H	Kern	Rosamond	CA	93560	2017	FORD	FOCUS	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
17	5YJSA1E26H	Snohomish	Mukilteo	WA	98275	2017	TESLA	MODEL S	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
18	1N4AZ1CP6K	Snohomish	Edmonds	WA	98026	2019	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
19	1N4AZ1CP2K	Kitsap	Bainbridge Island	WA	98110	2019	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
20	1N4AZ0CP9D	Grant	Moses Lake	WA	98837	2013	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
21	5YJ3E1EA3J	Thurston	Olympia	WA	98513	2018	TESLA	MODEL 3	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
22	5YJXCBE2XH	Thurston	Olympia	WA	98513	2017	TESLA	MODEL X	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
23	5YJSA1E23G	Skagit	Anacortes	WA	98221	2016	TESLA	MODEL S	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
24	1N4BZ0CP1H	Snohomish	Mukilteo	WA	98275	2017	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
25	5YJ3E1EA7J	Snohomish	Marysville	WA	98271	2018	TESLA	MODEL 3	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
26	1N4BZ0CP4G	Snohomish	Edmonds	WA	98026	2016	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible

After Changing the column names using Excel, I have loaded the dataset into jupyter Notebook for further cleaning using Python. I have dropped DOL_Vehicle_ID, Electric_Utility, 2020_Census_Tract and Legislative_District columns as I will not be using these columns in my analysis. After dropping the columns, I have checked for NULL values in the dataset.

```
VIN                                        0
County                                     2
City                                       2
State                                      0
Postal_Code                                2
Model_Year                                 0
Make                                       0
Model                                    181
Electric_Vehicle_Type                      0
Clean_Alternative_Fuel_Vehicle_Eligibility 0
Electric_Range                             0
Base_MSRP                                  0
Vehicle_Location                          29
dtype: int64
```

I have found that vehicle model has more NULL values, I cannot impute data as we have to take into consideration the company name for the imputing and the popular model in that may not have a model in that model_Year so I have decided to drop those 181 rows as it is less than 0.01% of whole dataset. I then used the info command to we find the details about the dataset.

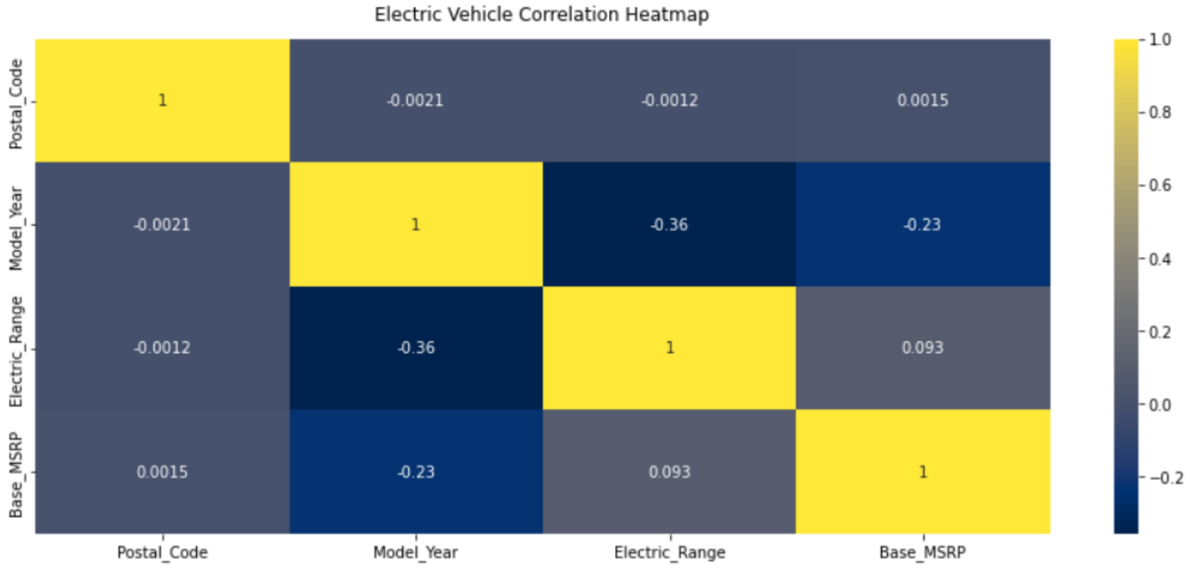
```
 #   Column                                      Non-Null Count   Dtype
---  ------                                      --------------   -----
 0   VIN                                         124506 non-null  object
 1   County                                      124506 non-null  object
 2   City                                        124506 non-null  object
 3   State                                       124506 non-null  object
 4   Postal_Code                                 124506 non-null  int32
 5   Model_Year                                  124506 non-null  int64
 6   Make                                        124506 non-null  object
 7   Model                                       124506 non-null  object
 8   Electric_Vehicle_Type                       124506 non-null  object
 9   Clean_Alternative_Fuel_Vehicle_Eligibility  124506 non-null  object
 10  Electric_Range                              124506 non-null  int64
 11  Base_MSRP                                   124506 non-null  int64
 12  Vehicle_Location                            124506 non-null  object
```

Now I have 12 columns with 124,506 observations to do my data analysis.

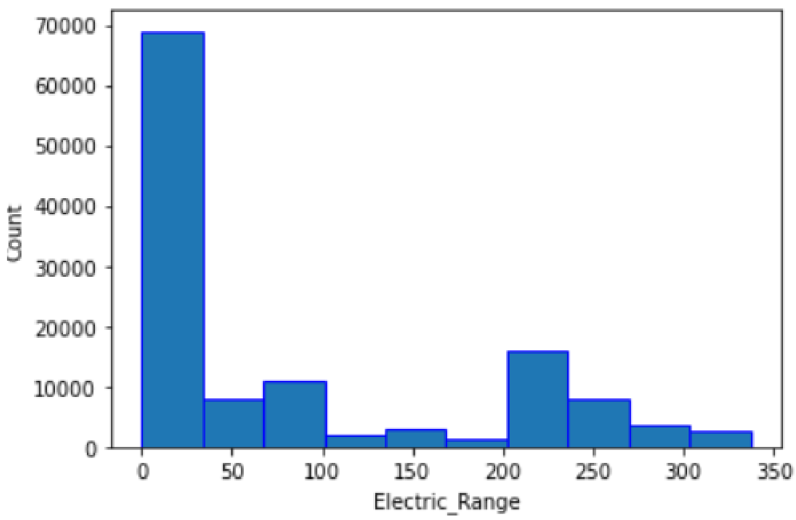
Exploratory Data Analysis

EDA is an essential process in data science that involves analysing data through summary statistics and graphical representations. Its purpose is to identify patterns, outliers, test hypotheses and assumptions. EDA provides valuable insights that help data scientists make informed decisions in subsequent stages of data analysis.

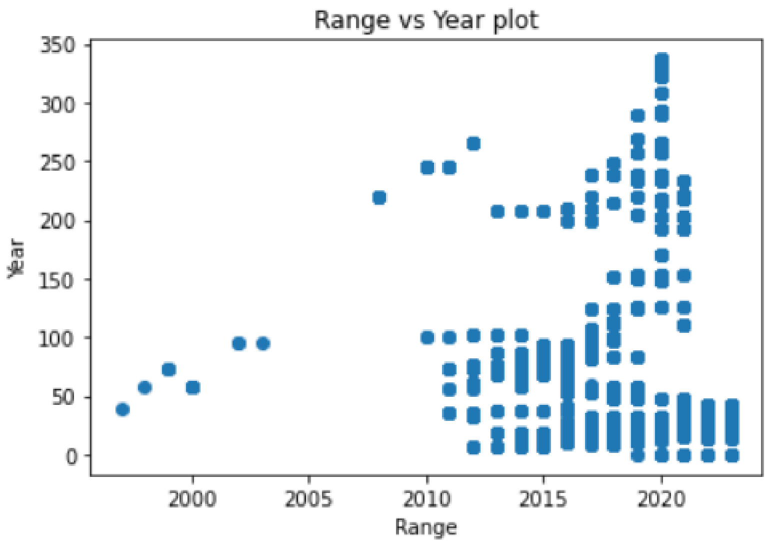
First, I have plotted the corelation map to see if there is any patterns or linearity between any columns. As we can see from the below graph that no two variables are linear to each other. So, this would be good for doing the data analysis



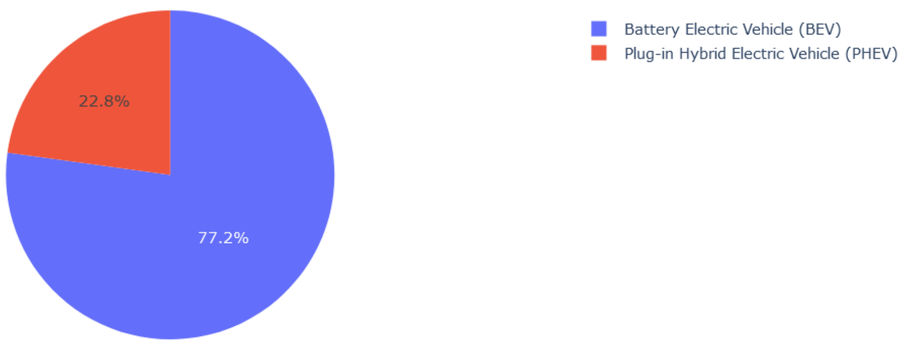
To understand the Electric vehicles and their popularity I have chose to plot the electric Ranges of the vehicles in the dataset. We can see from the below graph that most of the vehicles in the dataset have very low range i.e., between 0-50.

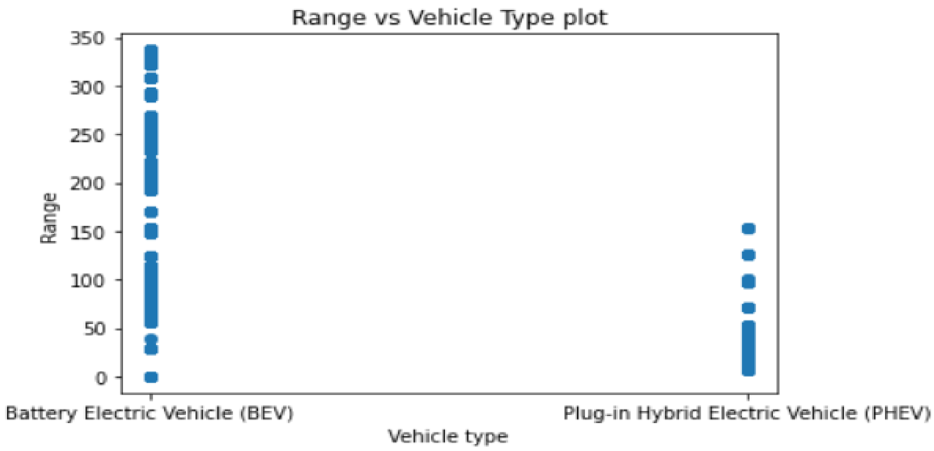


To better understand the evolution of the electric vehicles I have plotted a scatter plot between the electric range and Model years, from the graph we can see that the range has been increasing with the years and the electric vehicles are becoming more efficient.



To check the distribution of Vehicles in the dataset I have plotted a plotly graph between Vehicle Types i.e., Battery Electric Vehicle and Plug-In Hybrid Vehicle and I have found most of the Vehicles in the dataset come under Battery Electric Vehicles. We also we from the Vehicle type and range graph that Hybrid Electric Vehicles has lower ranges than the Battery Electric Vehicle.



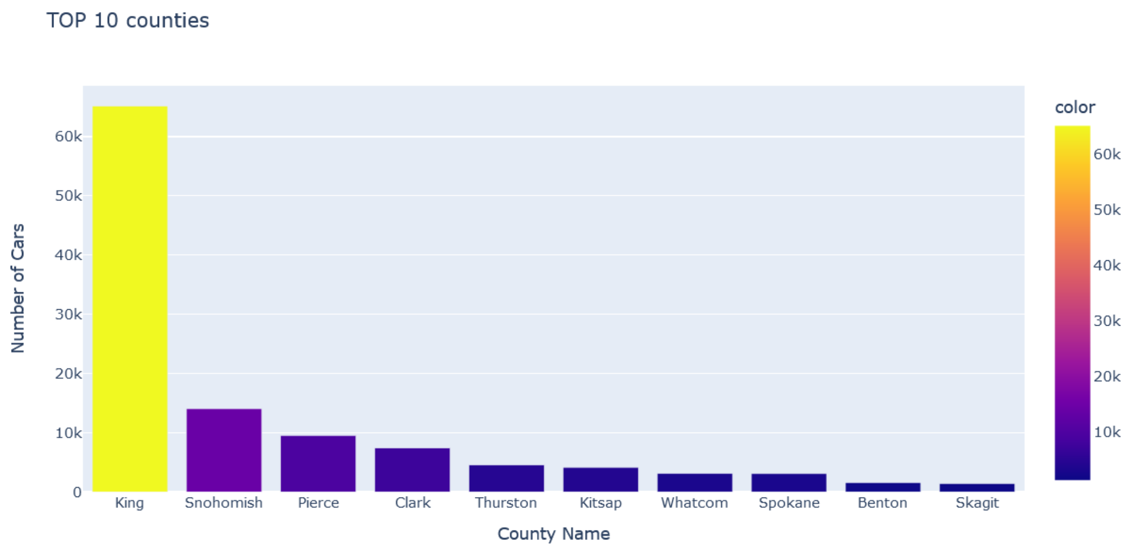


Research Questions

1. What is the electric vehicles distribution across Washington State?

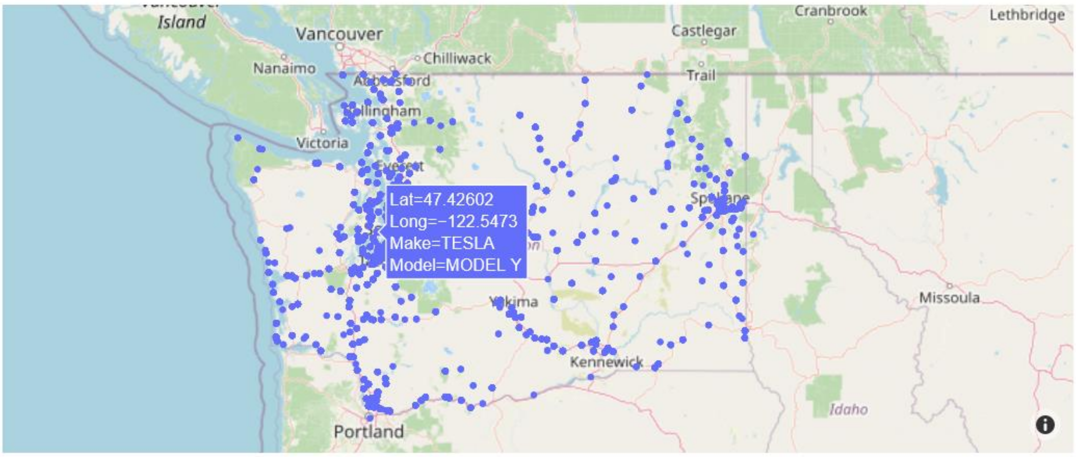
This is an important research question because it helps to understand how many electric vehicles are in the state. Knowing this information can help policymakers, industry stakeholders, and automobile manufacturers identify areas with high and low adoption rates, target policies and incentives to promote adoption in low adoption areas and adapt production and marketing strategies to meet regional demands.

To better understand he research question, I have first plotted the Top 10 counties with the greatest number of EV's. The below graph is plotly graph that give the name of the county and its count when pointed on it. We can see from the graph that most the vehicles registered in the Washington state below to king county. The king county has the highest number of EV's with a number of 65 thousand EV's registered in that county which is about 50% of the values in the dataset. From the below graph we can say that the people in the king county are using EV's multiple times of that of another county. This would be the case because that king county is near Seattle which has more population density compared to other counties.



To get a better understanding of the distribution of the EV I have first extracted the longitude and latitude values from the vehicle location column and have plotted the values in the world map using plotly to get a better understanding of the distribution. The map would also tell the company name and model of the that particular data point if we point on it.

Distribution of EVs



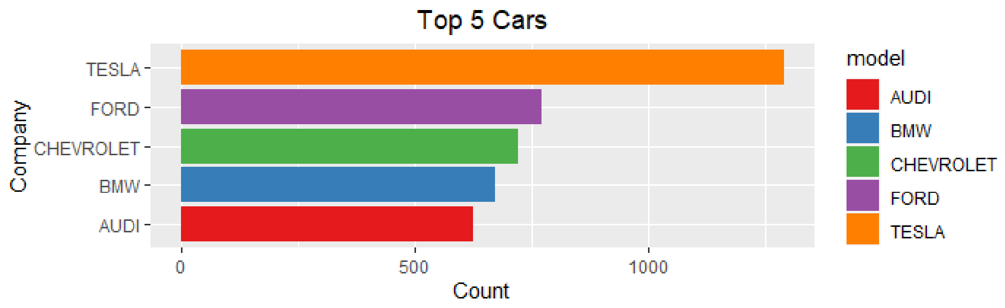
We can see that there are more vehicles around the Seattle city or along the west coastline that the towards the centre. This might be the case because the population there is higher, and the government may take some action to increase the awareness among the people.

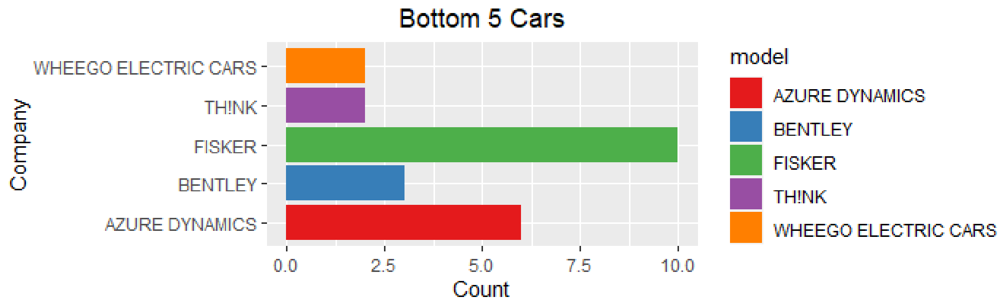
2.What brand electric vehicles are more popular in Washington State and what model is more common among them?

This research question is important because it can help automobile manufacturers understand their market share in the state and which models are more popular. This information can be used to optimize production and marketing efforts, making them more targeted and effective.

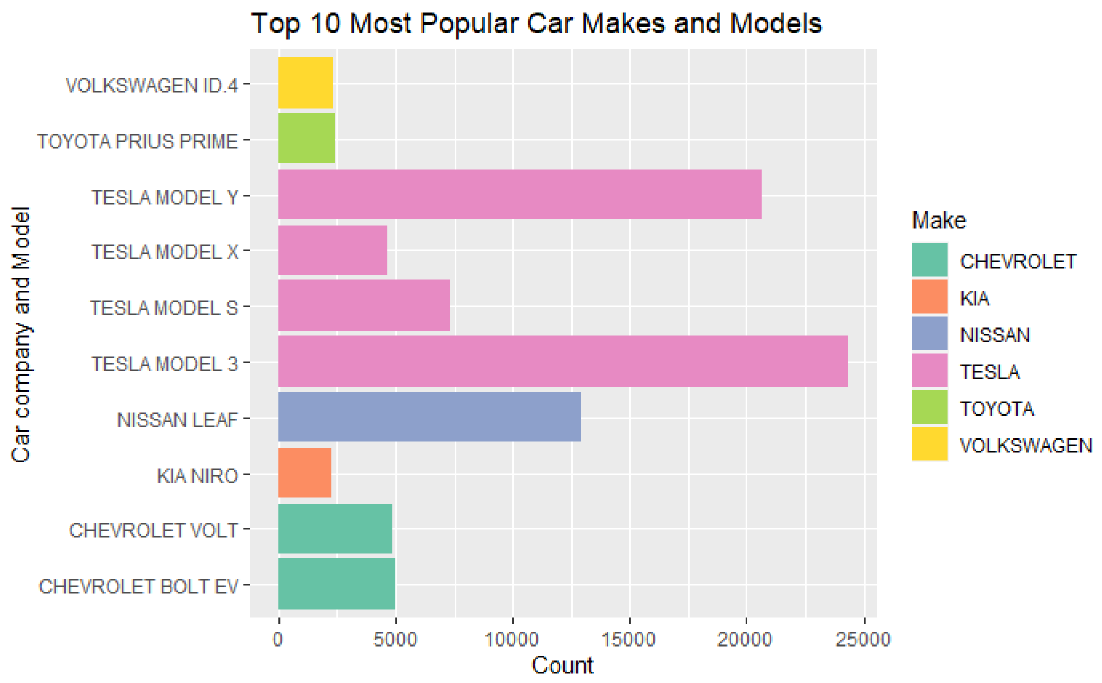
To find the most popular electric vehicles I have used R to answer this research question. I have extracted the dataframe from python to csv file and loaded them into the r studio to get the cleaned dataset.

To find the most popular vehicle I have plotted the most popular and least popular vehicles so this information will help the companies to make necessary actions. The top 5 companies with most vehicles are Tesla, Ford, Chevrolet, BMW and AUDI. And bottom 5 companies are Wheego electric cars, Think, Bentley, Azure dynamics and Fisher these company cars are very few with single digit numbers.





To find the most popular car model I have counted the car model with the greatest number of counts. The most popular car models are Tesla model Y and Tesla model 3 and most of the tesla model are popular and the next company car model that is popular which is non tesla is Nissan leaf car.



I have used MySQL to find the most popular car model. I got the number as the tesla model 3 with 24,409 counts. And I have used Pyspark to the values for top10 models,

```
+-----------+-----+
|      model|count|
+-----------+-----+
|    MODEL 3|24409|
|    MODEL Y|20673|
|       LEAF|12924|
|    MODEL S| 7345|
|    BOLT EV| 4988|
|       VOLT| 4888|
|    MODEL X| 4647|
|PRIUS PRIME| 2435|
|       ID.4| 2329|
|       NIRO| 2295|
+-----------+-----+
```

From the above analysis I have found that Tesla is the most successful EV's company with model 3 and model Y as it most sold vehicles. Though Nissan Leaf is the most popular non-tesla model, Nissan company is not present is in the top 5 companies who sold most cars. Likewise, though Ford have sold many EV's it does have any a model in top 10 sold cars. Nissan should work on more models in EV's, and Ford should try to have a popular EV's so that it stays in competition.

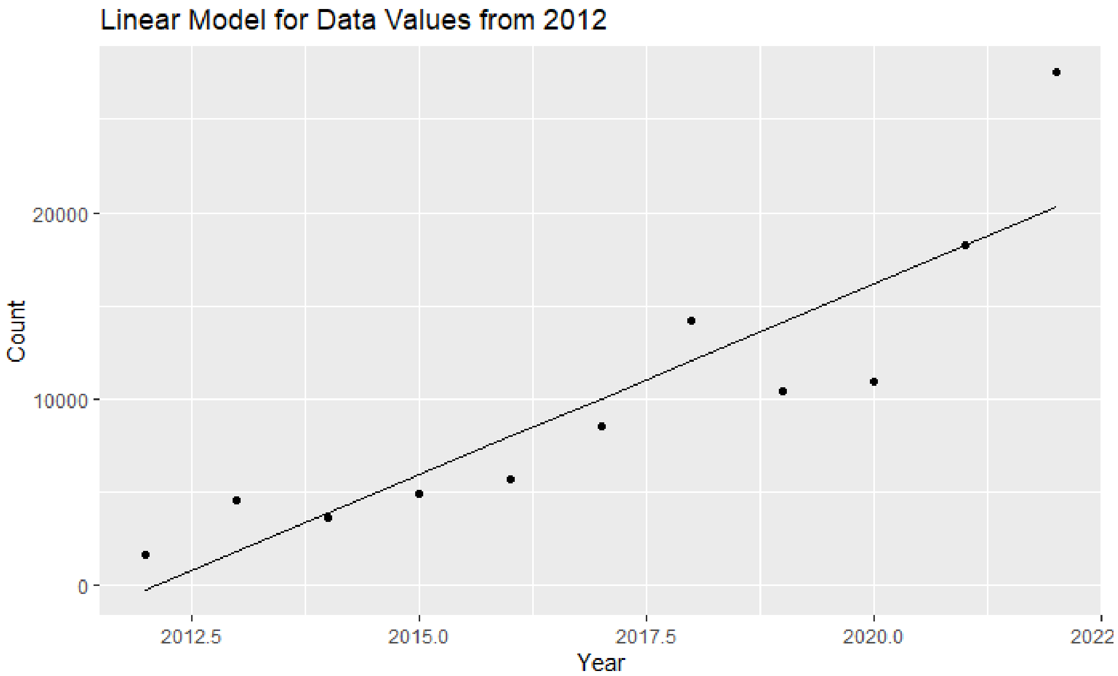
3.What is the growth of Electric vehicles across the years and what its future?

This research question is important because it helps to understand how many electric vehicles have been sold over the years and how much they are expected to grow in the future. This information is important for industry stakeholders and automobile manufacturers to make informed decisions.

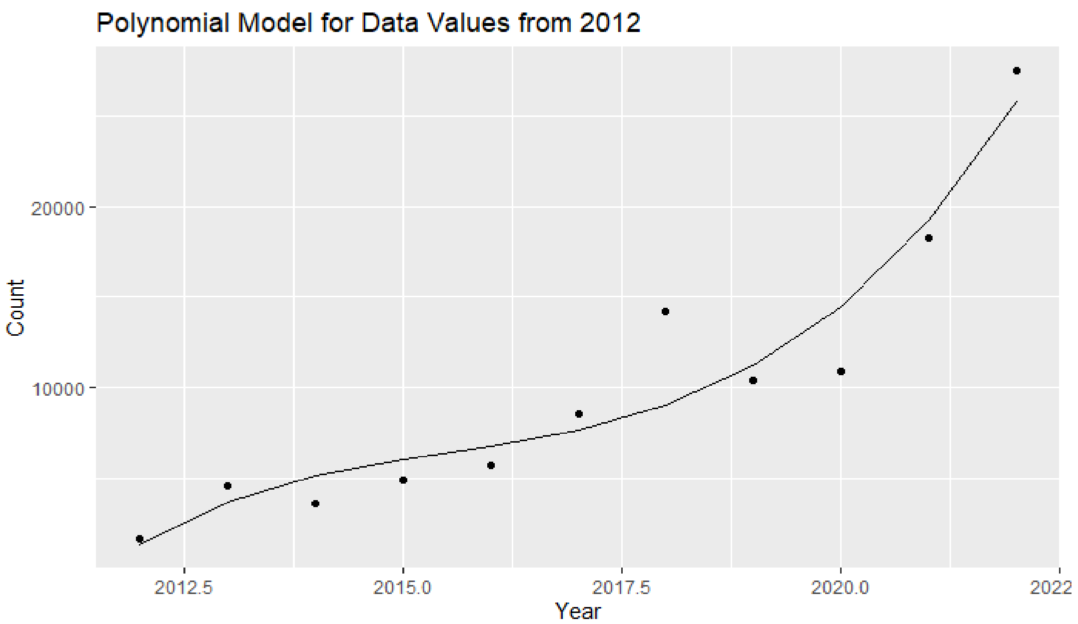
To get the number of vehicles sold across the years I have used MySQL to get table with number of vehicles sold across the years.

	Model_Year	count
4	2000	9
5	2002	2
6	2003	1
7	2008	21
8	2010	24
9	2011	826
10	2012	1666
11	2013	4577
12	2014	3608
13	2015	4935
14	2016	5699
15	2017	8557
16	2018	14223
17	2019	10448
18	2020	10925
19	2021	18293
20	2022	27514
21	2023	13172

We can see from the count table that the number of vehicles sold are increasing across the years, we can also see that it has decreased from 2022 to 2023 as the data is till 2023 April only, so for the prediction model I have chose to remove the year 2023 to get accurate results and have also only selected the from 2012 as that is when the electric have stated increasing exponentially. So, I have plotted a linear model for the count and model_year.



From the above linear model, we can see that the model fails to recognize the trend exponential increase in the recent years, so linear model would not be a great fit for prediction of the increase in Electric vehicles. So, I have chose to make a polynomial model which gives use a better change of predicting the values. I have chosen to use a 3rd degree polynomial model to get better results.



As we can see from the graph that 3rd polynomial model almost finds the correct line for the increase in the electric vehicles. I have tried to predict the values in the year 2023 and year 2025 to check my predictions. For the year 2023 the model has predicted the value as 34571.8 which looks to be accurate as for 4 months the EV's sold in 2023 is 12 thousand in the 8 month it would be around 36K which is near to the 34.5k the model suggest. According to the model the number of EV's in 2025 would be 59774.07 which is a huge claim as the number should almost be doubled in next 2 years this is not bad claim considering the increasing awareness of electric vehicles among people.

This model suggests that EV's would be a huge market in the coming years in the Washington state as these predictions are made based on the vehicles registered in the Washington state. This is a good sign moving forward into the future.

Future Work

To enhance the depth and accuracy of our market analysis, I aim to conduct a comparative analysis between the electric vehicle population data and the electric charging station data. The purpose of this analysis is to derive meaningful insights that can aid in providing more informed recommendations to stakeholders and automobile manufacturers regarding the establishment of electric charging stations. By establishing this correlation, we can provide more precise suggestions for reducing the capital cost associated with establishing charging stations. This approach aims to leverage data-driven insights to enhance the decision-making process and ensure that stakeholders and manufacturers can make informed decisions that align with the evolving landscape of the electric vehicle market.

Conclusion

The market analysis conducted on the electric vehicle (EV) population in Washington State indicates that the distribution of EVs is geographically disparate, with an augmented density of EVs in urban areas. Tesla has emerged as the most popular electric vehicle brand in the state, followed by Ford and Chevrolet. The Model 3 is the prevailing electric vehicle model in Washington State. The growth rate of EVs in the state has been noteworthy in recent years, with a marked escalation in EV registrations annually. The projections for the future indicate that this trend will continue as an increasing number of EV models become available and charging infrastructure improves.

Reference:

[1] B. Propfe, D. Kreyenberg, J. Wind, and S. Schmid, “Market penetration analysis of electric vehicles in the German passenger car market towards 2030,” International Journal of Hydrogen Energy, 16-Mar-2013. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0360319913004151. [Accessed: 08-May-2023].

[2] S. R. S, P. Venugopal, R. V, H. Haes Alhelou, A. Al-Hinai, and P. Siano, “Analysis of electric vehicles with an economic perspective for the future electric market,” MDPI, 31-May-2022. [Online]. Available: https://www.mdpi.com/1999-5903/14/6/172. [Accessed: 08-May-2023].

[3] A. Zubaryeva, X. Wu, O. P. R. V. Vliet, H.-K. Tseng, C. Thiel, S. Shafiee, R. Prud’homme, M. Peters, G. J. Offer, S. G. Nurre, T. E. Lipman, C. Lin, T. Lieven, J. I. Lewis, F. Kley, A. Hoen, M. Eppstein, Á. Driscoll, M. A. Delucchi, M. Catenacci, L. Bunce, R. Bointner, B. Battke, E. Baker, J. Axsen, P. Arocena, B. M. Al-Alawi, and M. Contestabile, “Total cost of ownership of electric vehicles compared to conventional vehicles: A probabilistic analysis and projection across market segments,” Energy Policy, 14-Feb-2015. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0301421515000671. [Accessed: 08-May-2023].

[4] “Electric Vehicle Population Data,” Catalog, 18-Mar-2023. [Online]. Available: https://catalog.data.gov/dataset/electric-vehicle-population-data. [Accessed: 08-May-2023].