**Abstract**

My exploratory data analysis (EDA) project focuses on the "Electric Vehicle Population Data" This rich dataset serves as a vital resource for understanding the landscape of electric vehicles in the United States. Through My analysis, I aimed to uncover geographic distribution patterns, vehicle demographics, charging infrastructure readiness, temporal trends, and environmental policy implications. This EDA project contributes to informed decision-making, sustainable transportation planning, and the broader effort to reduce emissions and promote eco-friendly transportation solutions, making it a valuable resource for researchers, policymakers, and environmental enthusiasts alike.

**Scope of the project**

This project involves a meticulous analysis of the "Electric Vehicle Population Data" dataset, sourced from data.gov, with a focus on comprehensive exploratory data analysis (EDA). The scope encompasses data collection and preprocessing, geospatial analysis, demographic exploration, charging infrastructure assessment, temporal trends analysis, and an evaluation of environmental and policy implications. Findings will be presented through data visualizations and detailed reports, culminating in actionable recommendations for stakeholders and policymakers. Acknowledging dataset limitations, the project will consider potential areas for future research. This EDA project aims to provide valuable insights into the adoption and impact of electric vehicles in the United States, supporting informed decision-making in sustainable transportation.

**Introduction**

Brief Summary: The chosen dataset comprises electric vehicle data sourced from the State of Washington, accessible through Data.gov. This dataset encompasses a wide range of information related to electric vehicles, including their distribution, adoption trends, and charging infrastructure within the state. The data will be used to explore patterns of electric vehicle adoption and usage, with a focus on understanding the factors driving the growth of electric vehicles in Washington. The research aims to answer questions related to the state's electric vehicle landscape, including geographic distribution, types of electric vehicles in use, and the availability of charging infrastructure, thus contributing to the broader understanding of sustainable transportation trends in the region.

**Data Summary**

The data that I acquired from is from the website Data.gov

[**https://catalog.data.gov/dataset/electric-vehicle-population-data**](https://catalog.data.gov/dataset/electric-vehicle-population-data)

**Potentially numerous use cases**

Use Case 1: Vehicle Make Analysis

Objective: Determine the distribution of electric vehicles by make in the dataset.

Attributes: VIN (1-10), Make

Description: In this use case, we aimed to understand the distribution of electric vehicles by their make. The VIN (1-10) column represents the Vehicle Identification Number, while the Make column contains the manufacturer's name. By analyzing this data, we can identify which electric vehicle makes are present in the dataset and gain insights into the diversity of manufacturers.

Use Case 2: Sales Trends by Model Year

Objective: Analyze the trends in electric vehicle sales over different model years.

Attributes: VIN (1-10), Model Year

Description: In this use case, we focused on understanding the trends in electric vehicle sales by model year. The VIN (1-10) represents the Vehicle Identification Number, and the Model Year column contains information about the year in which the vehicles were manufactured. By visualizing and analyzing this data, we can identify patterns and trends in sales over time, which may reveal important insights into the evolution of electric vehicle adoption.

**Use Case 3**: Manufacturer-Utility Associations

**Objective:** Determine the associations between electric vehicle makes and the electric utilities they are associated with.

**Attributes:** VIN (1-10), Make, Electric Utility

**Description:** In this use case, we explored the relationships between electric vehicle makes, the first part of the Vehicle Identification Number (VIN), and the electric utilities they are associated with. The Make column represents the manufacturer's name, and the Electric Utility column contains information about the utility provider. By analyzing this data, we can understand which electric utilities are associated with specific manufacturers and gain insights into the distribution of electric vehicle makes across different utilities in Washington State.

VIN (1-10) County City State Postal Code Model Year Make Model Electric Vehicle Type Clean Alternative Fuel Vehicle (CAFV) Eligibility Electric Range Base MSRP Legislative District DOL Vehicle ID Vehicle Location Electric Utility 2020 Census Tract

**Python Code:**

dataframe = pandas.read\_csv('Electric\_Vehicle\_Population\_Data.csv')

dataframe.dtypes

**Output**

**VIN (1-10) object**

**County object**

**City object**

**State object**

**Postal Code float64**

**Model Year int64**

**Make object**

**Model object**

**Electric Vehicle Type object**

**Clean Alternative Fuel Vehicle (CAFV) Eligibility object**

**Electric Range int64**

**Base MSRP int64**

**Legislative District float64**

**DOL Vehicle ID int64**

**Vehicle Location object**

**Electric Utility object**

**2020 Census Tract float64**

**dtype: object**

[**https://catalog.data.gov/dataset/electric-vehicle-population-data**](https://catalog.data.gov/dataset/electric-vehicle-population-data)

**DATA ANALYSIS**

Data analysis for the "Electric Vehicle Population Data" involves dissecting and structuring datasets to gain insights into the adoption and distribution of electric vehicles (EVs) in Washington. Leveraging Python libraries such as NumPy and pandas for data manipulation, the analysis aims to identify growth patterns, geographic hotspots, and prevalent EV types (BEVs and PHEVs) in the region. Additionally, the assessment includes an exploration of the availability and accessibility of charging infrastructure. To present these insights effectively, the analysis employs seaborn and matplotlib data visualization libraries. Seaborn, built upon matplotlib, offers a high-level interface for crafting visually appealing and informative statistical graphics. The analysis provides valuable information for policymakers, industry stakeholders, and EV enthusiasts, facilitating informed decision-making and the promotion of sustainable transportation in Washington.

# Inference Research Question 1: Were there more Tesla’s sold than any other kind of vehicle in Washington State?

# Analysis:

# To answer this question, I started by examining the dataset, which contains information about electric vehicle sales in Washington State. I wanted to determine if Tesla’s dominated the electric vehicle market in the state. The first step was to tally the number of vehicles sold by each make. I used Python code to generate a bar chart that visually represented the count of electric vehicles sold for each make. This analysis helped me understand the distribution of electric vehicle makes in the state.

# Code:

# tally = dataframe['Make'].value\_counts()

# tally.plot(kind='bar', title='Electric Vehicles Sold in Washington State')

# Findings:

# Based on the analysis, it's evident that Tesla is the dominant electric vehicle manufacturer in Washington State, with the highest number of electric vehicles sold compared to any other make. Chevrolet, Nissan, and other manufacturers follow Tesla in terms of sales, but Tesla stands out as the most popular choice among electric vehicle buyers.

**Research Question 2: How many vehicles were sold in each year? What was the trend in sales over time?**

# Analysis:

# For this question, I aimed to understand the trend in electric vehicle sales over time. I was curious to see how the sales numbers varied by model year. I used Python code to create a bar chart illustrating the count of electric vehicles sold in each model year. This visual representation allowed me to identify trends and patterns in electric vehicle sales over different model years.

# Code:

# tally\_model\_year = dataframe['Model Year'].value\_counts()

# tally\_model\_year.plot(kind='bar', title='EVs Sold in each Model Year')

# Findings:

# The analysis reveals a positive trend in electric vehicle sales in Washington State. There is a notable increase in sales in the 2018 and 2019 model years. This suggests that electric vehicles have been gaining popularity in the state over the years, with a peak in sales during those specific model years.

# Research Question 3: Which make is using which electric utility?

# Analysis:

# To address this question, I wanted to understand the associations between electric vehicle makes and electric utilities. I began by grouping the data by electric utility and examining the top manufacturers using each utility. The Python code allowed me to generate a bar chart displaying the top manufacturers for each electric utility. This analysis provided insights into which electric utilities were associated with specific electric vehicle makes.

# Code:

# result\_dataframe = dataframe.groupby(['Electric Utility'])

# counts = result\_dataframe['Make'].value\_counts().groupby(level=0).nlargest(10)

# top\_n = 10

# top\_counts.plot(kind='bar')

# Findings:

# The analysis provides a clear view of which electric utilities are associated with particular electric vehicle manufacturers. For example, Tesla, Chevrolet, and Nissan are among the top manufacturers using specific electric utilities in Washington State. This information can be valuable for understanding the distribution of electric vehicle makes across different utilities in the state.

# Conclusion

The Python code and data analysis helped answer these research questions, providing valuable insights into the electric vehicle market in Washington State, including the dominance of Tesla, the positive sales trend over the years, and the associations between electric vehicle makes and utilities. These findings are essential for understanding the electric vehicle landscape in the state and can inform various stakeholders in the industry.

The dataset on electric vehicle population data provides valuable information for understanding the growth and adoption of electric vehicles. It offers insights into trends, geographical distribution, and various aspects of electric vehicle ownership. This data is crucial for policymakers, researchers, and companies in the automotive and energy sectors to make informed decisions and develop strategies to promote sustainable transportation.

By analyzing this dataset, one can gain a deeper understanding of the electric vehicle market, monitor its growth over time, and identify potential areas for investment and improvement. The data also highlights the shift towards cleaner and more environmentally friendly modes of transportation, which is an essential step in reducing the environmental impact of the automotive industry.

Overall, this dataset is a valuable resource for stakeholders interested in the future of transportation and sustainable energy solutions.

**References**

<https://catalog.data.gov/dataset/electric-vehicle-population-data>