Analyzing Trends in Electric Vehicle Usage: US and Global Perspectives

Contributors:

Kavya Moharana:

- netID: kavyam3
- ORCID: 0009-0005-8035-8567
- contributions:
  - Data Integration, Python data cleaning and preprocessing for Data.gov's Electric Vehicle Population Dataset
  - Data Dictionaries for datasets, formatting of python notebook structure, summary/descriptive statistics
  - Python analysis and visualizations for answering Research Question 1
  - Report: Summary, Written analysis for findings of Research Question 1, Future Work, Metadata
  - Creation of GitHub Repository and committing/releasing final work to GitHub

Mahnur Khalid:

- netID: mkhal2
- contributions: 0009-0001-2651-4208
  - Data Acquisition from source
  - Data Integration, Python data cleaning and preprocessing for Kaggle's Electric Vehicle Population History by US County Dataset
  - Python analysis and visualizations for answering Research Question 2
  - Report: Data Profile, Written analysis for findings of Research Question 2, Future Work
  - Creation of archival record through Zenodo

Aisha Ziad:

- netID: aziad2
- ORCID: 0009-0002-2881-5589
- contributions:
  - Data Acquisition from source
  - Data Integration, Python data cleaning and preprocessing for Kaggle/IEA's Global EV Sales Dataset
  - Python analysis and visualizations for answering Research Question 3
  - Report: Reproducing Section, Written analysis for findings of Research Question 3, Future Work, Citations

Archival Record:

DOI 10.5072/zenodo.141771

Link: https://sandbox.zenodo.org/records/141771

Summary

The overall goal or motivation of this project is to conduct comprehensive analysis on Electric Vehicle (EV) data to understand potential trends in the use of fuel efficient vehicles both in the United States and globally. The US Environmental Protection Agency categorizes EVs as 'Green' vehicles which are an alternative to conventional gasoline or diesel (fossil fuels) powered motor vehicles because they reduce emissions of CO2 and other harmful greenhouse gasses into the atmosphere. As the production and sales of such environmentally friendly vehicles have been growing in recent years, it may be indicative of a shift in perspective towards climate change or changes in affordability of EVs. Out aim is to analyze data from the Washington State Department of Licensing and the International Energy Agency to examine the differences between EV types, understand EV population differences between certain US states, and finally to see if these insights are reflected within global electric vehicle sales.

Research Questions

The questions we intend to address with our project are:

- What is the growth rate of electric vehicle registrations over time and are certain vehicle types (BEVs vs. PHEVs) growing faster in specific areas?
- How do electric vehicle registrations differ between states and how might this relate to factors like population density, income, or urbanization?
- How do trends in electric vehicle registration or sales in the USA compare to other countries globally and do they indicate a change in environmental consciousness?

Battery Electric Vehicles or BEVs are powered solely by an electric battery and have no gas engine parts. Plug-In Hybrid Vehicles or PHEVs are the same as hybrids but have a larger battery and electric motor along with a gas tank. The purpose of the first research question is to further understand the distribution of each type of EV and if there are differences in the growth rate of EVs based on whether it is BEV or PHEV. Answering this research question would allow us to gain insight into electric vehicle preferences and observe regional patterns based on type of EV. The second question intends to explore the potential geographic factors that influence the adoption of EVs across different states. Analyzing how the rate of EV registrations vary from state to state would allow us to understand which US counties have people that are more likely to use an EV and whether this is correlated with factors such as proximity to major cities. After analyzing EV trends within counties in the United States, we thought it would be interesting to compare and contrast these insights with the EV sales of other countries around the world. This question would have us assess rates of alternative fuel usage and make potential inferences on the levels of environmental awareness internationally based on how EV rates differ by region of the world.

Analysis Process and Findings

We conducted research of these questions through Data acquisition, Data integration, Cleaning and Quality assessment, Python Pandas Analysis, and Data Visualization which are all documented within the main Python Jupyter Notebook file (titled ProjectAnalysis).

For data acquisition and integration we downloaded the datasets locally and performed preprocessing, cleaning, and exploratory analysis (summary or descriptive statistics) as shown in the Jupyter Notebook. We provided formal Data Dictionaries for each dataset in order to further understand the dataset and become more familiar with the types of information we were working with. Each team member formatted and cleaned the data in a way that would make it easier to anlyze trends between specific variables of the

datasets. The subsequent data analysis consisted of calculating metrics (calculating growth rate, percentages, aggregating data) for variables specific to each research question and inputting them into new dataframes or time-series to produce insightful visualizations. These visualizations allowed us to identify specific trends for growth rate of BEVs versus PHEVs, regions of the US and Washington State, and countries around the world.

Our findings show that there is significant growth in electric vehicle (EV) adoption in the United States over time, with Battery Electric Vehicles (BEVs) outpacing Plug-in Hybrid Electric Vehicles (PHEVs) in most regions. Densely populated counties like King, Snohomish, and Pierce in Washington State show the highest EV registration counts, while counties like Thurston and Spokane show faster BEV growth advantage rates. Comparatively, states like Arkansas and Iowa stand out with high EV density, which may be influenced by their lower population densities and other potential socio-economic factors. On a global scale, we found that the U.S. and China dominate the EV market. Countries such as Norway and Iceland are also leading, potentially due to their focus on environmental policies. The general upward trend in EV registrations and supporting infrastructure, both nationally and internationally, indicates growing environmental consciousness and an overall shift toward sustainable transportation.

All data visualizations related to the findings can also be found in the visualization folder of the repository.

Data Profile

Electric Vehicle Population Data from Data.gov

This dataset is originally sourced from Data.gov's federal data portal through the Washington State Department of Licensing. We can assure that it has been acquired from a credible source as it is from an official US government website, which is important for reproducibility purposes as well as ensuring the integrity of the data. Assessing the licensing and terms of use, this dataset is listed as open access and intended for general public use. The terms of use for this dataset are further clarified by the Open Data Commons, where it is clearly established that the data is meant to be publicly accessible information. The Washington State Department of Licensing data specifically focuses on characteristics of the overall registered Electric Vehicle (EV) population within the state of Washington, covering both Battery Electric Vehicles (BEVs) and Plug-In Hybrid Vehicles (PHEVs) types. The format of the dataset is a JSON file and includes registration information that was collected starting from November 2020 to the most recent update of October 2024. It includes vital EV related categorical attributes such as County, City, Model Year, and Make as well as numeric values like Electric Range, Base MSRP, or Vehicle ID. The purpose for including this dataset in our analysis is to further understand how different types of electric vehicles (BEV versus PHEV) are distributed across different regions or counties of Washington state, providing direct insights for the variations in EV rates on a geographic level.

Electric Vehicle Population by US County from Kaggle

This dataset regarding Electric Vehicle population history is sourced from the popular open-source data science platform, Kaggle. The dataset particularly focuses on data of vehicles that were registered by Washington State Department of Licensing (DOL) every month from 2017 to 2024. For data integrity and reproducibility purposes, we can assure that it has been acquired from a credible source as it is from an official US government website and is regularly updated. When searching for the license and terms of use of this dataset, we found that its terms of use are outlined in and covered by the Apache License 2.0. The license also clearly states that the dataset is open to the public and is allowed to be accessed and modified for general public use. Although the dataset was acquired from Kaggle, it is originally sourced from the

State Department of Washington, confirming the reliability and accuracy of information. Despite also being collected by the state of Washington, the EV data captures information beyond just the State of Washington (California, Texas, Florida, etc), and breaks it down by county level for each state along with distinction of passenger vehicles versus trucks. The format of this data set is a CSV file, and it is expected to be updated on a monthly basis. The dataset includes categorical attributes such as Date, County, State and more. The numeric data includes information such as the number of total vehicles, number of electric vehicles, number of plug-in hybrid electric vehicles, and total non-electric vehicles.

Global Electric Vehicle Sales: 2010 - 2024 from Kaggle

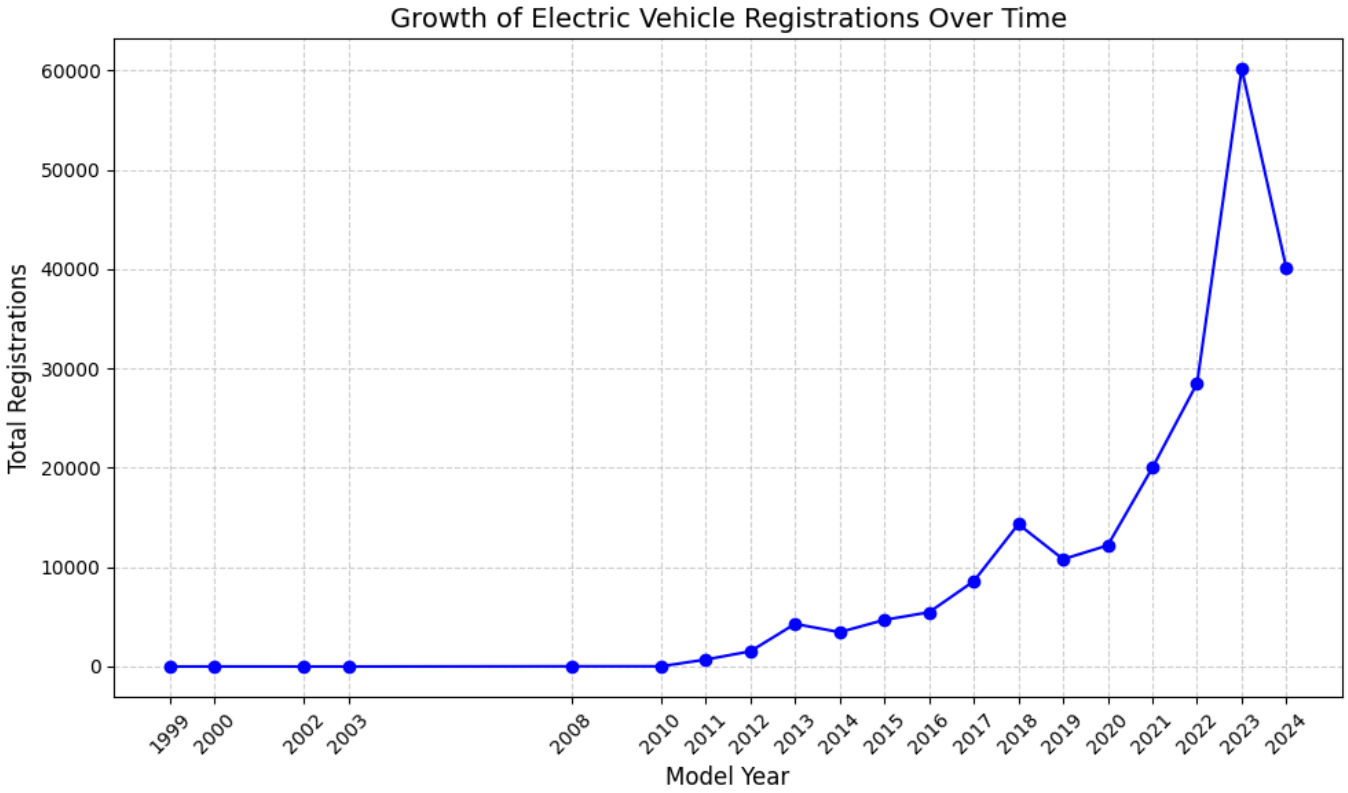
This dataset is also sourced from Kaggle.com but is originally from International Energy Agency's (IEA) 2024 Global EV Outlook report. It has an Attribution 4.0 International license from the nonprofit Creative Commons organization, outlining the terms of use. There, it is stated that the general public is allowed to access, share and modify this dataset, as long as credit is given appropriately and the licensing of the dataset is addressed. Since this dataset originates from the International Energy Agency (IEA), we can confidently confirm that it is sourced from a reputable organization. This ensures the dataset is credible for reproducible research. The dataset can be accessed and acquired directly from the IEA's official website, where it's available to the public for research and analysis purposes. The scope of this dataset is from December 31st 2010 to December 30th 2024, at a global scale covering regions like North America, Asia, and multiple other countries. The format of this data set is a CSV file and this data is never updated. This data set includes a variety of attributes and information ranging from Country, category, parameter, year, unit, value, mode, and powertrain.

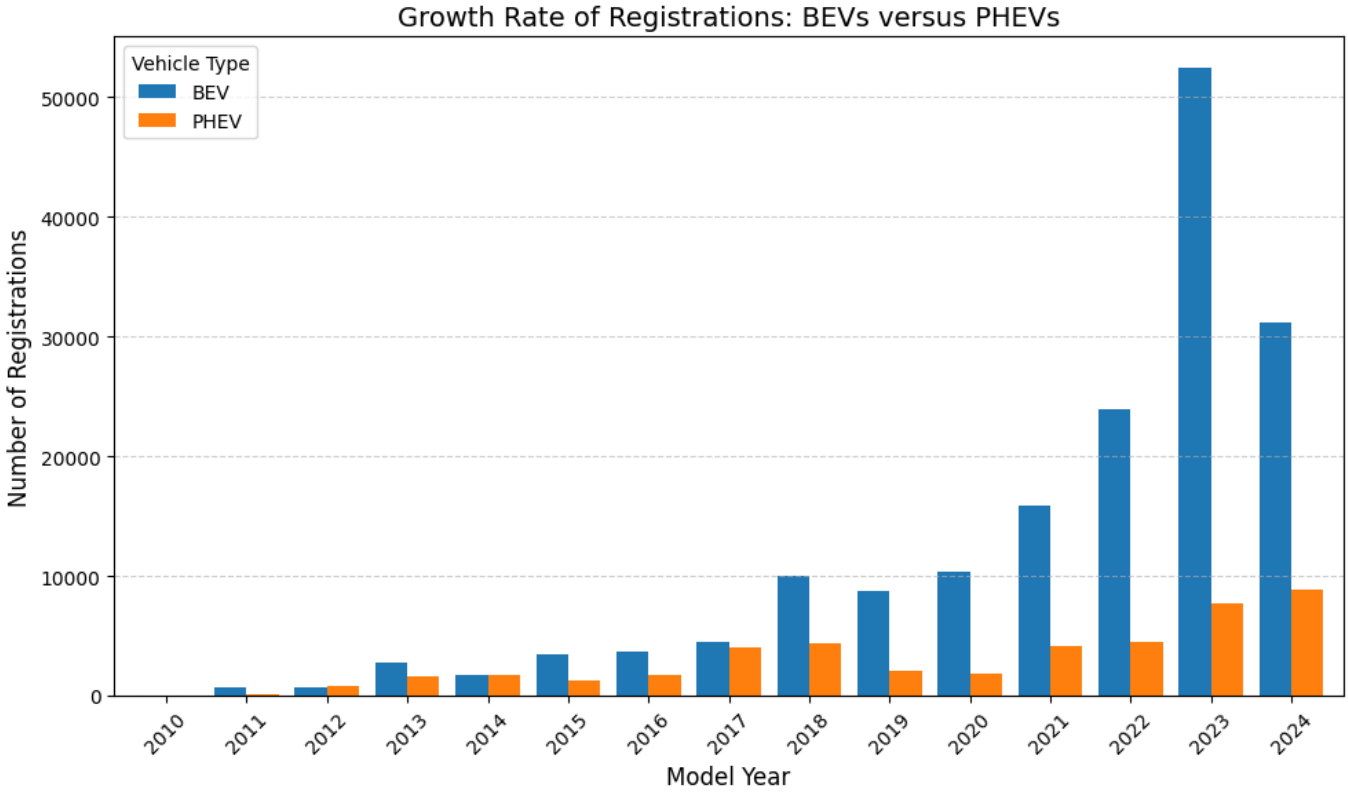
Findings

1. What is the growth rate of electric vehicle registrations over time and are certain vehicle types (BEVs vs. PHEVs) growing faster in specific areas?

Analysis by: Kavya Moharana

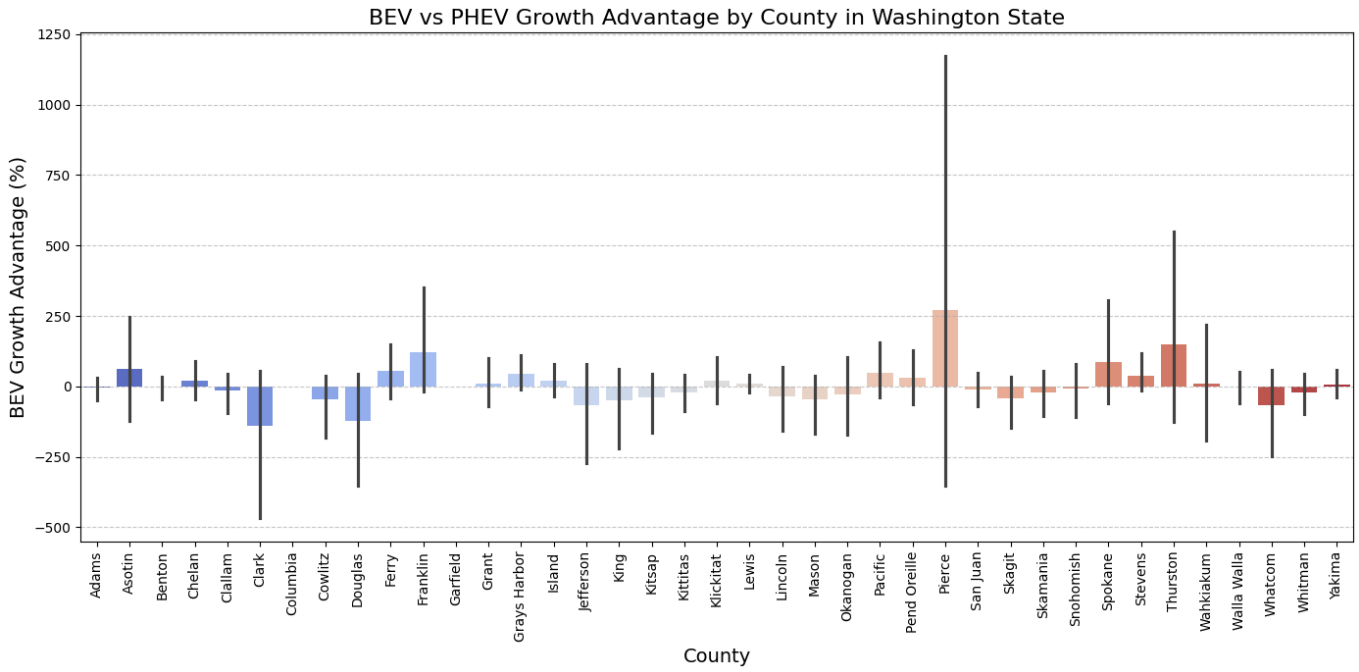
In order to analyze the differences between the BEV and PHEV electric vehicle types and how they have been growing in specific areas of Washington, I first looked at the growth rate of overall EV registrations over time. Grouping by Model Year of registered EVs in relation to the total number of registrations and visualizing it with a line plot, there was clear upward trend in EV registrations from 1999 to 2023 and a decrease in registrations for 2024. This decrease may also be due to the lack of data for the months of November and December as we acquired the dataset when the latest update was October 2024. The second grouped bar chart displays the differences in growth rate of registrations between Battery and Plug-In Hybrid EVs, by which we can see that there is significantly higher number of BEVs than PHEVs in the registration data. This bar chart also shows a slight increase in PHEV registrations between 2019 and 2024.

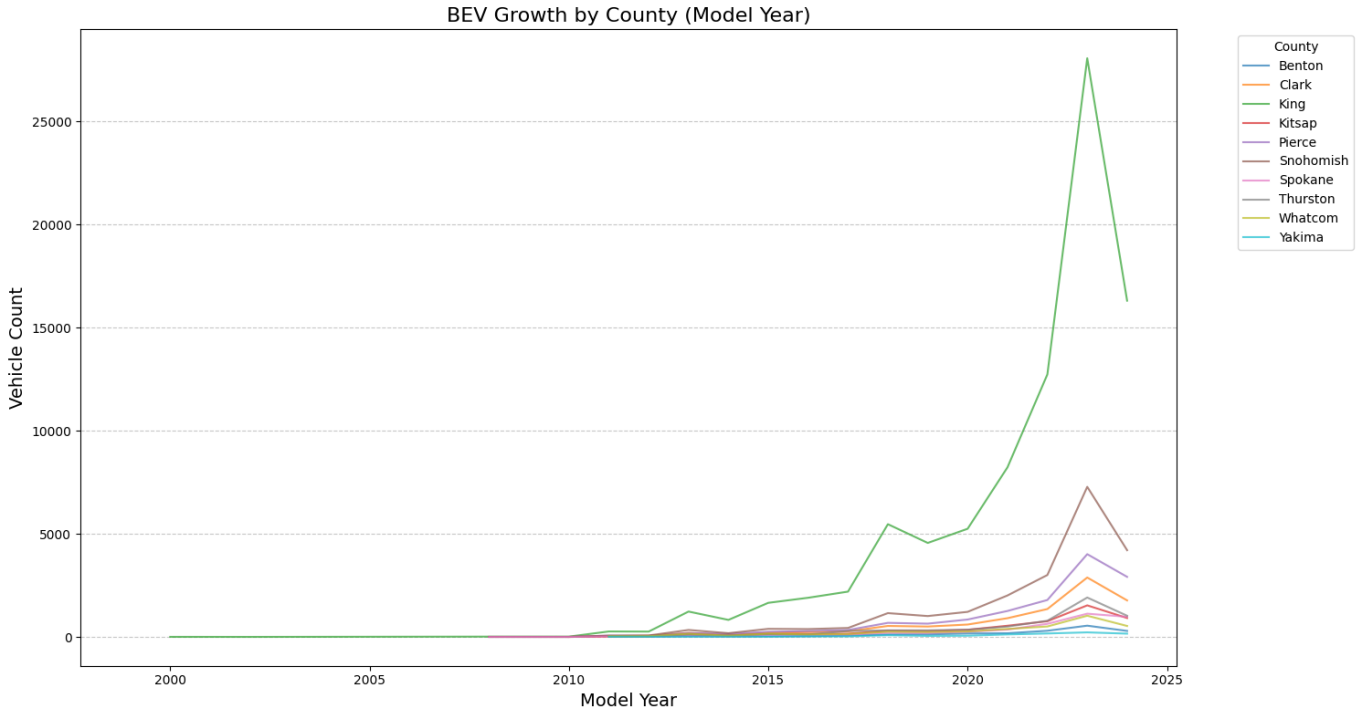




For regional trends of BEVs and EV adoption, I looked at registration counts and percentage for major counties and cities in the state of Washington. Counties like King, Snohomish, and Pierce along with the Seattle metropolitan area expectedly had the most EV registrations which aligns with population density. Then, I grouped and aggregated vehicle counts by county, vehicle type, and model year, before calculating the year-over-year growth rates for each EV type. This analysis compares BEV and PHEV growth rates to identify counties where BEVs are growing faster. The resulting barplot visualization lists the top 15 counties with significant BEV growth advantages. We see that Pierce, Thurston, and Spokane counties had the largest BEV Growth Advantage percentages over time. Finally, the last figure shows the growth trends of

BEVs over different model years in the top 10 most populated counties of Washington State. The line chart highlights variations in Battery EV adoption across densely populated regions of Washington. From this visualization it is clear that King County, the 12th-most populous county in the United States, had the highest BEV count along with Snohomish and Pierce. Most counties show a gradual upward trend from 2010 to 2020 with a sharp decline around 2023.



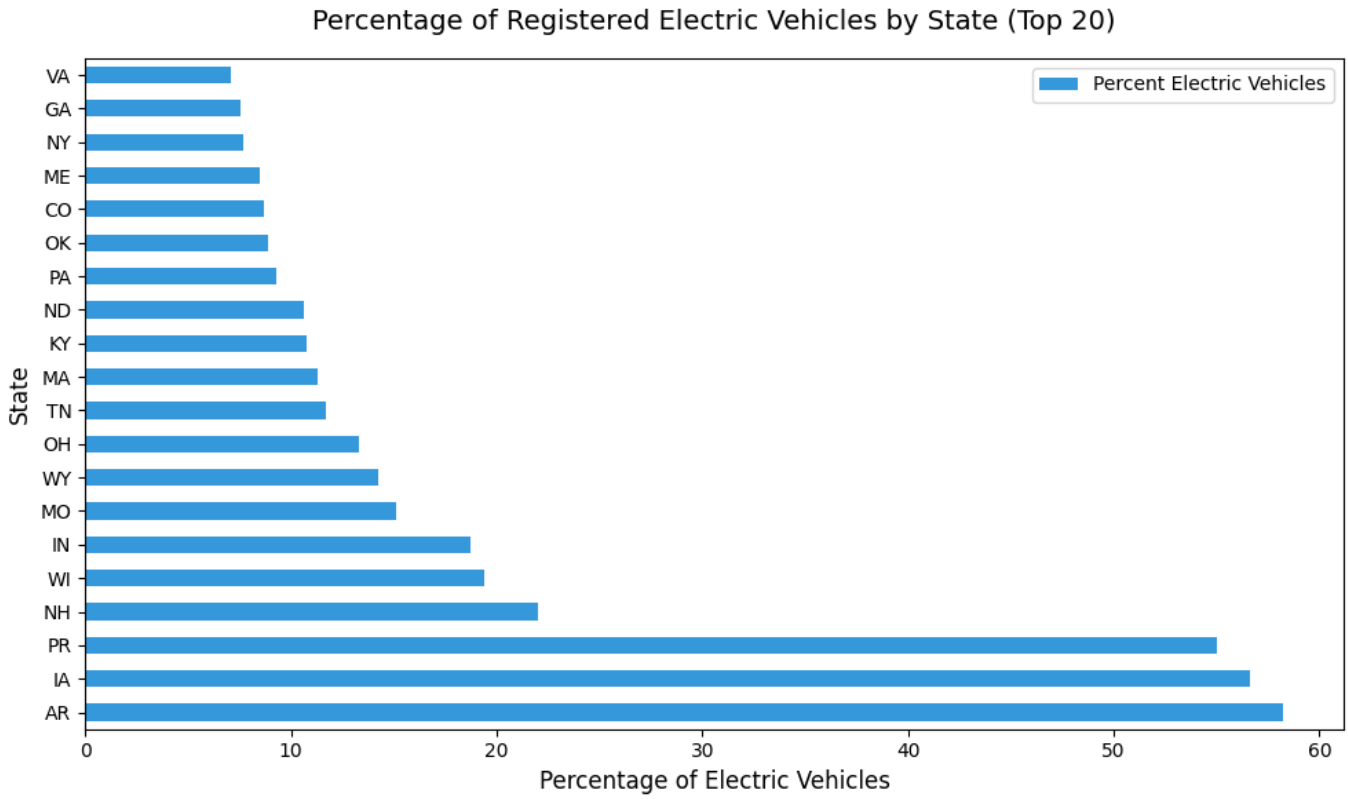


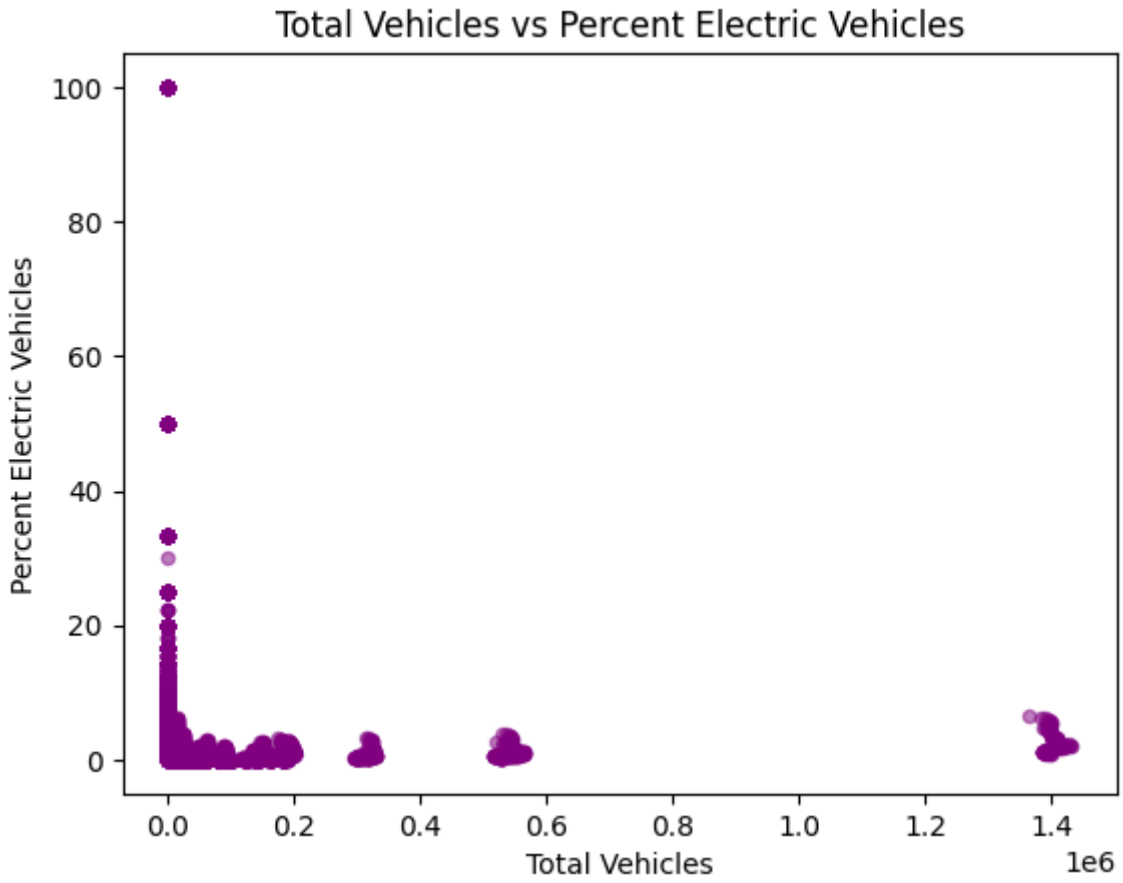
In summary, the growth rate of electric vehicle (EV) registrations in Washington State has shown a clear upward trend from 1999 to 2023, with a slight decline in 2023 and 2024. Battery Electric Vehicles (BEVs) significantly outnumber Plug-in Hybrid Electric Vehicles (PHEVs) in registrations, with BEVs consistently showing higher growth rates. Regional analysis revealed that areas with high population density, such as King, Snohomish, and Pierce counties, have the largest EV registration counts. Pierce, Thurston, and Spokane counties also have the highest BEV growth advantage. This indicates that BEVs are growing faster in these counties compared to PHEVs, highlighting upward regional trends in EV adoption.

2. How do electric vehicle registrations differ between states and how might this relate to factors like population density, income, or urbanization?

Analysis by: Mahnur Khalid

To analyze how electric vehicle registrations differ between states and how this might relate to factors like population density, income, or urbanization, I created a bar plot to visualize the distribution of EVs within the dataset. After analyzing the plot, we observe that Arkansas and Iowa emerge as the top two states with a significant proportion of electric vehicles, each approaching 60%. This suggests that these states are leading in EV density compared to others in the U.S. These trends could possibly be influenced by factors such as population density, income, and urbanization. Both Arkansas and Iowa are less densely populated and lack highly urbanized cities as compared to other US states, which might impact the overall percentage of EVs in these states. The lower population density could make it easier for EV presence, as fewer vehicles may be on the road. Additionally, it would be interesting to explore the role of income in these states, especially considering their high EV percentages. Understanding how these factors contribute can provide insights into the driving forces behind EV prevalence in these two states.





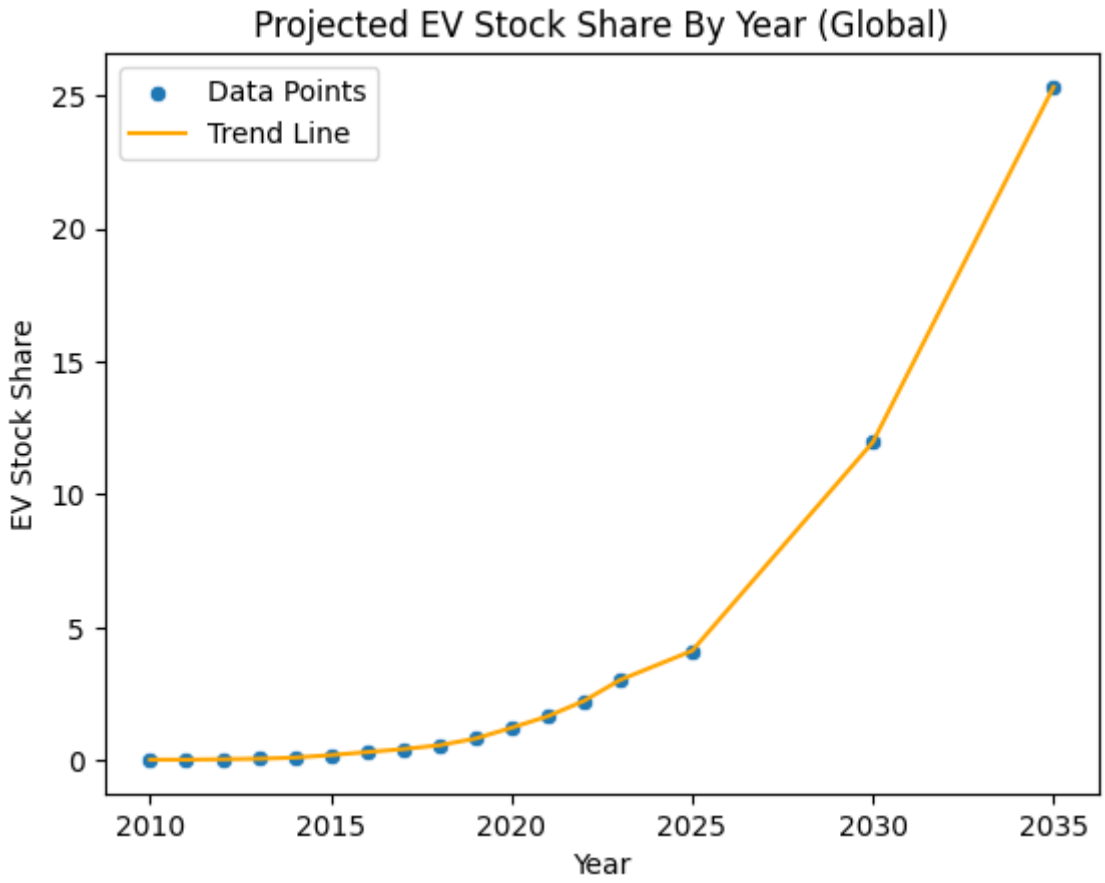
This analysis suggests in states with lower population densities, the fewer total vehicles contribute to a higher relative share of electric vehicles. In Addition, income levels and local policies could influence EV adoption, as higher incomes might facilitate EV purchases as well as infrastructure development. This also underscores the importance of examining regional socio-economic and geographic factors to further understand EV adoption trends across US states.

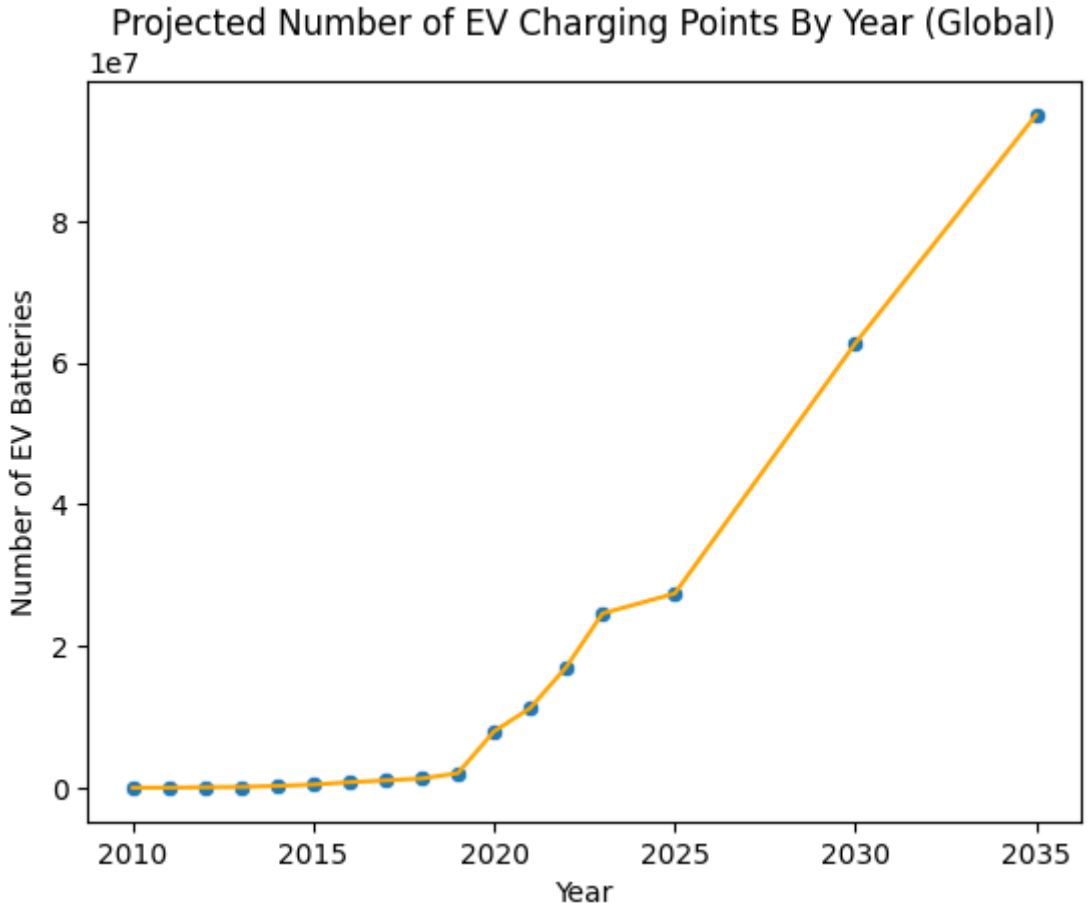
3. How do trends in electric vehicle registration or sales in the USA compare to other countries globally and do they indicate a change in environmental consciousness?

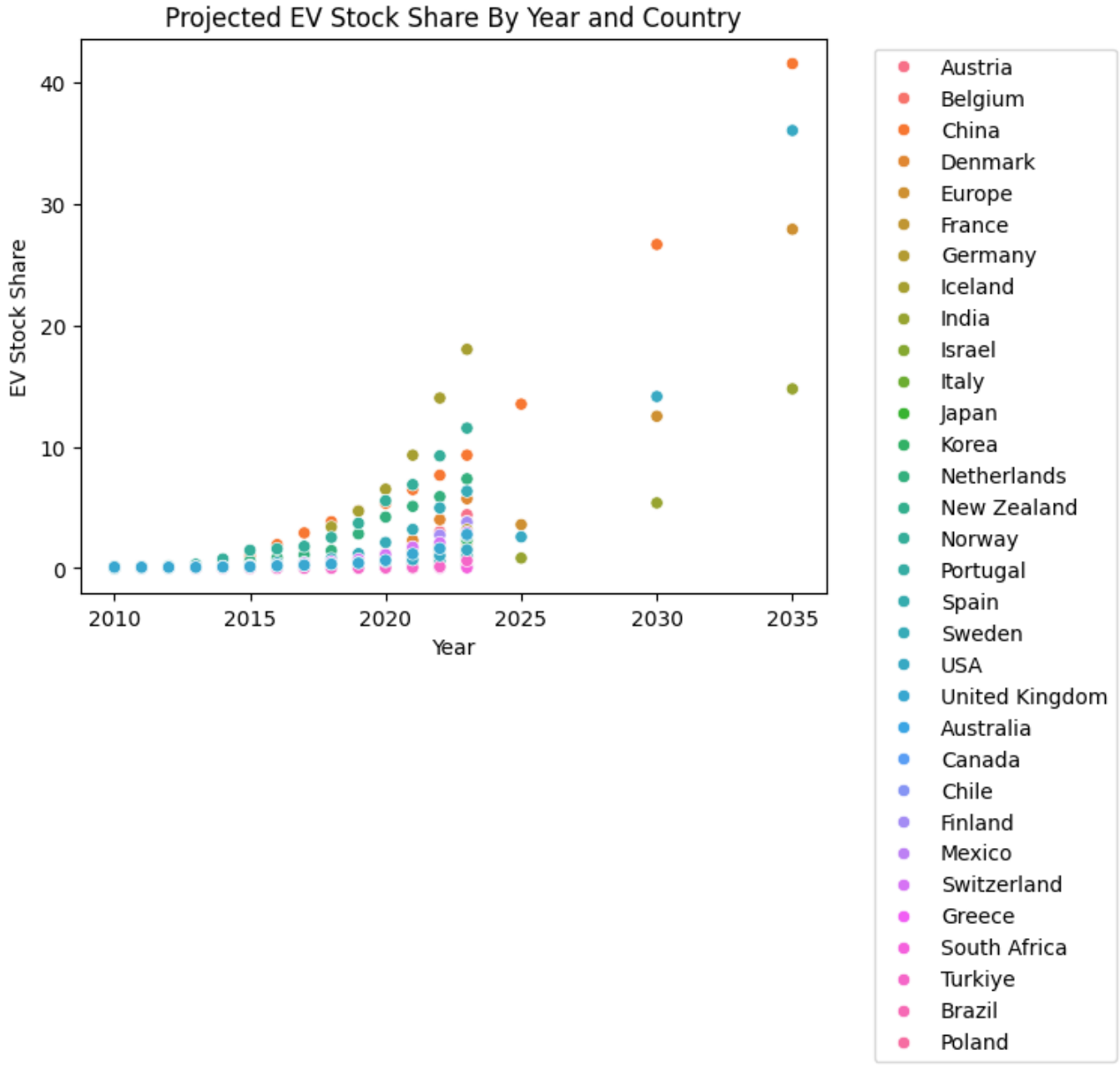
Analysis by: Aisha Ziad

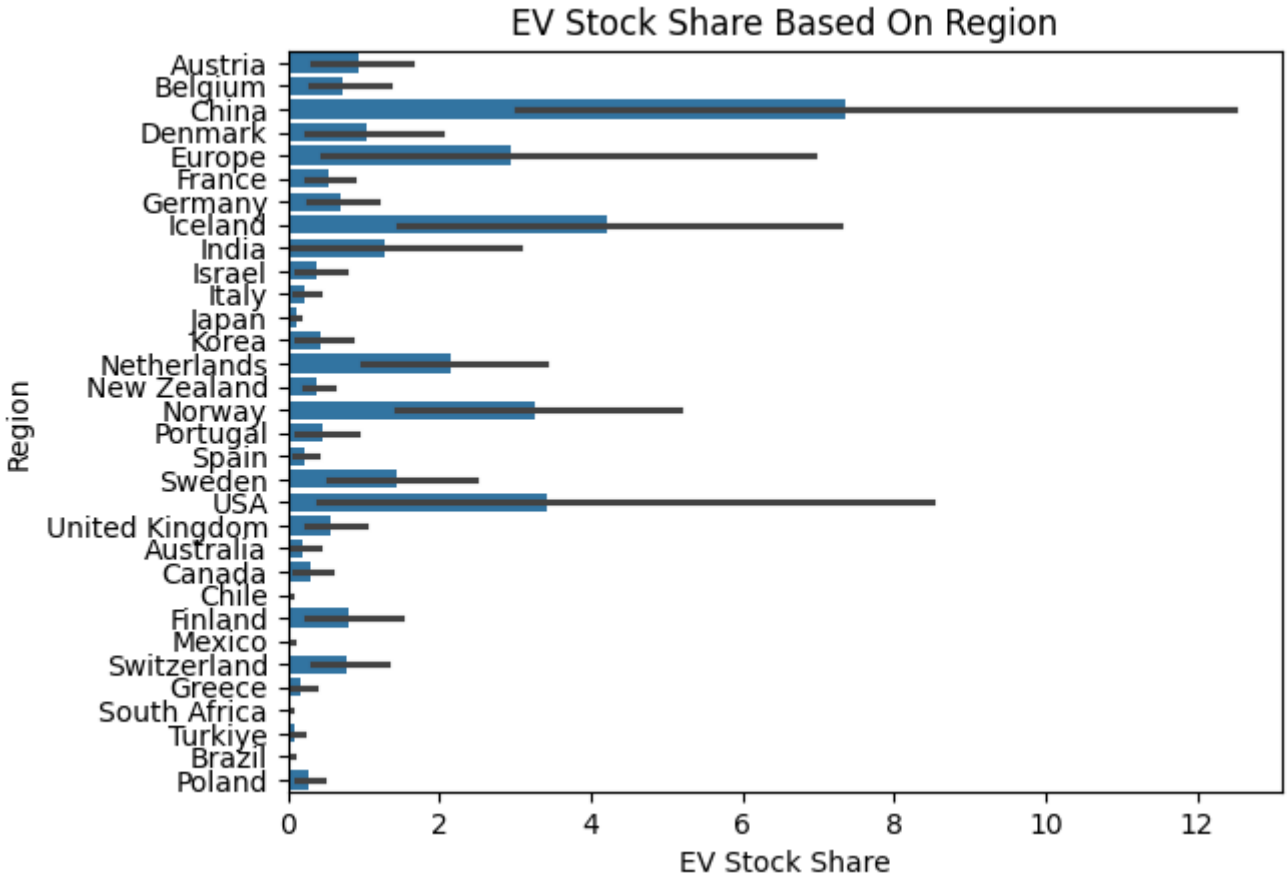
Our third research question is to understand how electric vehicles trends differ across the United States and other regions in the world. I created a bar plot to see how valuable the EV Stock Share is across different parts of the world. We see that the United States and China both are leaders in the electric vehicle market, indicated by their high stock values. Interestingly, other countries such as Iceland and Norway also have higher EV stock values. While we did drop three data points when we were cleaning the dataframe, this still doesn't change the fact that both the United States and China are leaders in the electric vehicle industry.

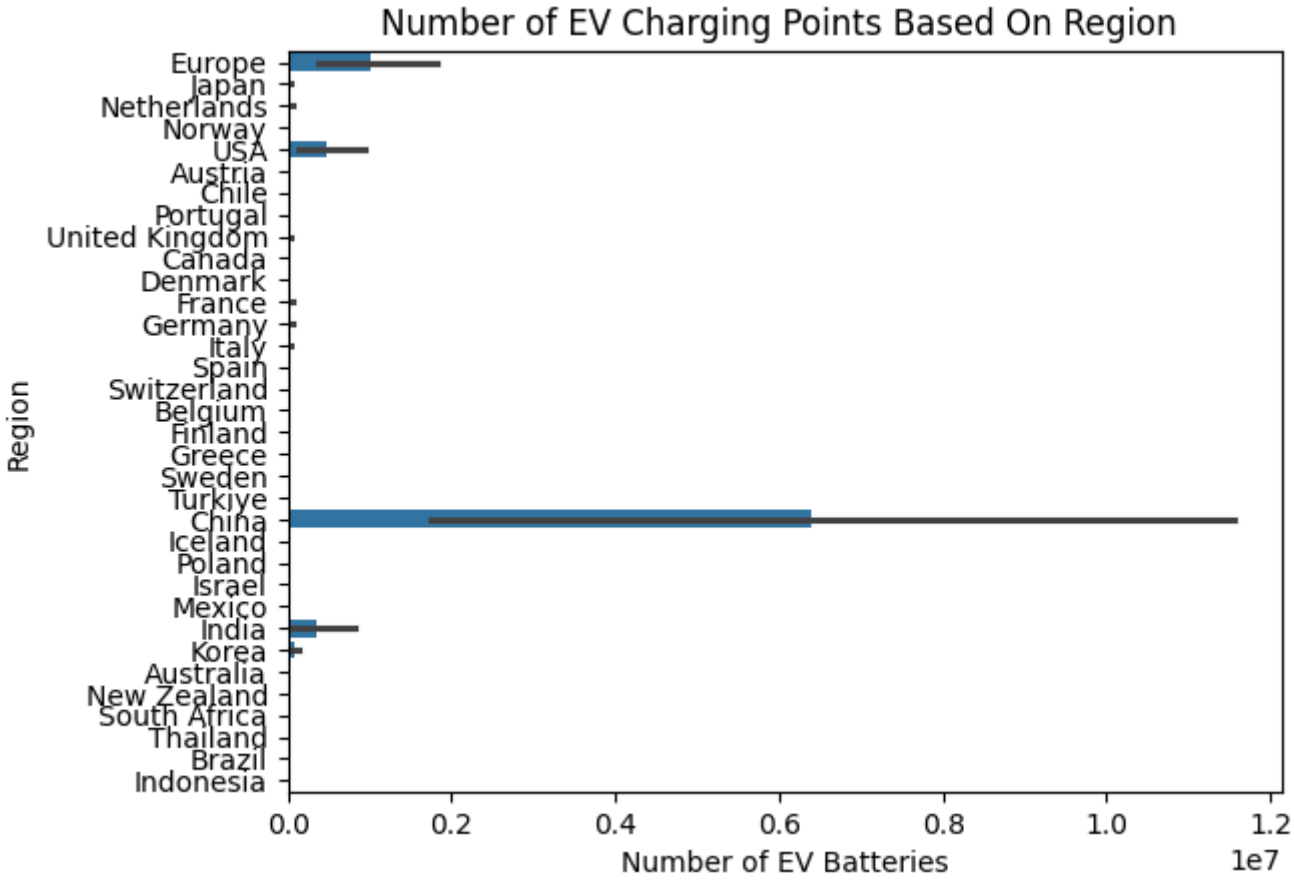
Our second and third visualizations further emphasizes this. In the plot comparing EV Stock Share Based on Year, we see that there is an upward trend in EV stock value for nearly all regions as the years pass. This implies that there is, in fact, a change in environmental consciousness. There are more EV purchases being made, and it has a positive effect on its overall stock value. The number of EV charging points also proves that more electric vehicles are being purchased. If there wasn't an uptick in the number of EVs, the number of charging ports would not increase either.


Projected EV Stock Share By Year (Global)
Data Points
Trend Line
EV Stock Share
Year
0
5
10
15
20
25
2010
2015
2020
2025
2030
2035


Projected Number of EV Charging Points By Year (Global)
1e7
Number of EV Batteries
Year
0
2
4
6
8
2010
2015
2020
2025
2030
2035


Projected EV Stock Share By Year and Country
EV Stock Share
Year
0
10
20
30
40
2010
2015
2020
2025
2030
2035
Austria
Belgium
China
Denmark
Europe
France
Germany
Iceland
India
Israel
Italy
Japan
Korea
Netherlands
New Zealand
Norway
Portugal
Spain
Sweden
USA
United Kingdom
Australia
Canada
Chile
Finland
Mexico
Switzerland
Greece
South Africa
Turkiye
Brazil
Poland





Future Work

Kavya

Having conducted analysis on the first research question, there are several possible directions for future work on this topic. A deeper analysis of what factors contribute to the higher BEV growth rates in certain counties (such as Pierce and Spokane in Washington State) could further highlight the impact of local policies and infrastructure on EV adoption. Furthermore, exploring temporal gaps or anomalies, like the decrease in EV registrations in 2023-2024, could help validate whether this trend is due to actual market fluctuations or simply just data limitations. In the future, merging the Washington State and US County datasets could provide better insights on the regional trends and census tracts as well. It is also important to note that while the data from the Washington State Department of Licensing provides information about other US States, it does not give us the complete picture on national trends as it only looks at vehicle *registration* data.

Mahnur

After analyzing the second research question on how electric vehicle (EV) registrations differ between states, we can now use the findings from this analysis to direct further steps for the second part of our research question. By understanding which states have a high prevalence of electric vehicles and which have extremely low rates, we can now look into the external factors influencing these rates. For future work, I am interested in integrating external data, such as sourcing population and income information, to analyze how these factors might correlate with EV rates. By visualizing and analyzing income and population trends within these states, we can explore whether these factors explain the patterns we observed. Additionally, I would like to examine Arkansas and Iowa in more detail because I'm curious to investigate whether the higher rates of EV registrations are concentrated in major cities or distributed more evenly across the state. Taking a deeper analysis within these states would help us better understand the role of urbanization in influencing EV rates.

Aisha

When answering our third research question, it is important to consider how we have data points that have been dropped, and regions that are vague. For example, in our data cleaning steps, we dropped regions with the values 'EU27,' 'Rest of the world', and 'World.' Due to this, we must consider if our analysis would change, and if we would see any other countries taking the lead for highest EV Stock Value. However, since both China and USA were untouched regions (when conducting preliminary data analysis and cleaning), we can assume that they are still very prominent in the electric vehicle industry. Another important thing to note is that while some of our regions are continents, such as Europe and United States, we also have individual countries noted in the dataframe. For example, we find that Norway and Iceland also have high EV Stock Values. In the future, for more accurate and identifiable results, we could work to group together countries into continents. However, we don't believe that it heavily impacts our results at this moment. Once again, we see that USA and China are strong leaders in the EV industry, with a few other countries close behind in terms of EV Stock Share Value.

Reproducing

To reproduce the results, one must conduct data cleaning and analysis for each of our datasets, by running all steps documented in the Python Jupyter Notebook titled 'ProjectAnalysis.ipynb' after acquiring the datasets from the provided resources.

The second and third datasets can be accessed through the U of I box folder link provided below. These two datasets must be taken out and placed in the same directory as the analysis notebook. After doing so,

the project steps can also be reproduced by running the Snakefile which provides an HTML version of the Jupyter Notebook and its results.

Box Link: https://uofi.box.com/s/n96dlzp6odintyehxzxnthggxpu20qza

Brief Overview for Reproducing Results:

- **Electric Vehicle Population Dataset from Data.gov:**
    - Read in the JSON file from the data.gov website using the appropriate steps.
    - Drop any unnecessary columns that do not have importance or significance when looking at our research question.
    - Check for duplicates and drop any NA values in the dataframe that may affect our results
    - To better understand our dataframe, we will observe the summary statistics and also begin some preliminary data analysis. For example, we will single out our numeric columns, make subsets of our dataframes (such as EV registrations by County), and group by certain parameters.
    - Next, we will plot our data (derived from grouping certain values together), and observe the overall trend. We will then use our plots to answer our research question and back up our findings.
- **Electric Vehicle Population by US County Dataset from Kaggle:**
    - Read in the CSV file using the appropriate steps. We opt not to drop any columns for this dataset.
    - Check for duplicates and drop any NA values in the dataframe that may affect our results
    - As we did for the first dataset, we will observe summary statistics and perform preliminary data analysis such as grouping together important variables.
    - Next, we will plot our data (derived from grouping certain values together), and observe the overall trend.
    - We will then use our plots to answer our research question and back up our findings.
- **Global Electric Vehicle Sales Dataset from Kaggle:**
    - Read in the CSV file using the appropriate steps. We find that there are a few vague values in the 'region' column, and choose to drop them. Specifically, these are the 'EU27,' 'Rest of the world,' and 'World' values.
    - Check for duplicates and drop any NA values in the dataframe that may affect our results
    - Once again, we will observe summary statistics and perform preliminary data analysis such as grouping together important variables.
    - Next, we will plot our data (derived from grouping certain values together), and observe the overall trend.
    - We will then use our plots to answer our research question and back up our findings.

References

Citations for Datasets Used:

> "Electric Vehicle Population Data." Data.gov, Data.wa.gov, 10 Nov. 2020,

> catalog.data.gov/dataset/electric-vehicle-population-data.

> Ford, Patrick L. “Global EV Sales: 2010-2024.” Kaggle, 2024,
>
> www.kaggle.com/datasets/patricklford/global-ev-sales-2010-2024/data.

> Maharaj, Sahir. “Electric Vehicle Population by Country (2024).” Kaggle, 2024,
>
> www.kaggle.com/datasets/sahirmaharajj/electric-vehicle-population-size-2024/data?select=Electric_Vehicle_Population_Size_History_By_County_.csv.

Metadata

> Khalid, M., Ziad, A., & Moharana, K. (2024). IS477 FA24 Final Project. Zenodo.
>
> https://doi.org/10.5072/zenodo.141771