Data Bootcamp 041524

2024 June 10

Data Analysis of Electric Vehicles Registered in Washington State

Project 1

PRESENTED BY

Andrew Montemayor

Will Parker

Fatima Rodriguez

Hannah Lashway

Ben Rosensweig

THESIS

Electric vehicles are popping up all over the place. People often cite their affordability, efficiency, popularity and environmental friendliness as reasons for purchase. Even the government has joined in, incentivizing the purchase of both electric and plug-in hybrids by offering tax credits.

We were inspired by the cultural relevancy of electric vehicles to select our dataset which looks at all of “the Battery Electric Vehicles (BEVs) and Plug-in Hybrid Electric Vehicles (PHEVs) that are currently registered through the Washington State Department of Licensing (DOL).”

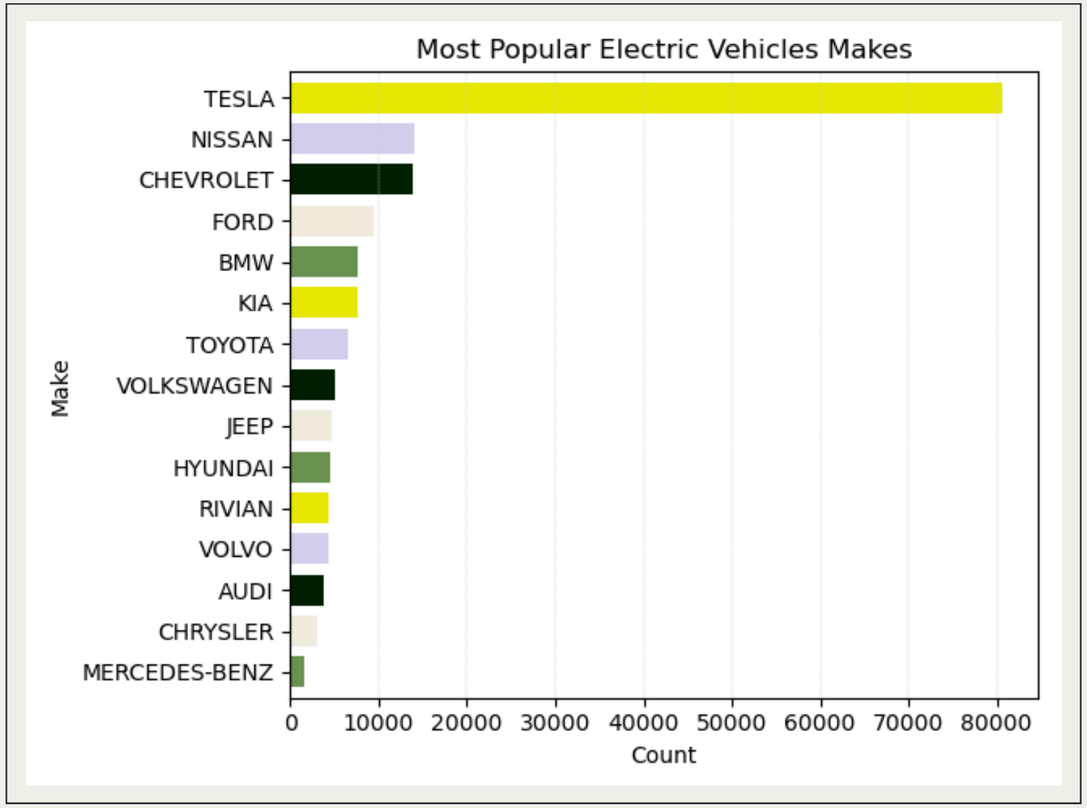
Through utilizing geographic data such as zip codes, consumer data like model popularity, and technical data like battery range, we aim to gain a better understanding of why people are choosing to purchase electric vehicles and what are the specific trends we can observe in terms of vehicle selection and location.

DATA CLEANING

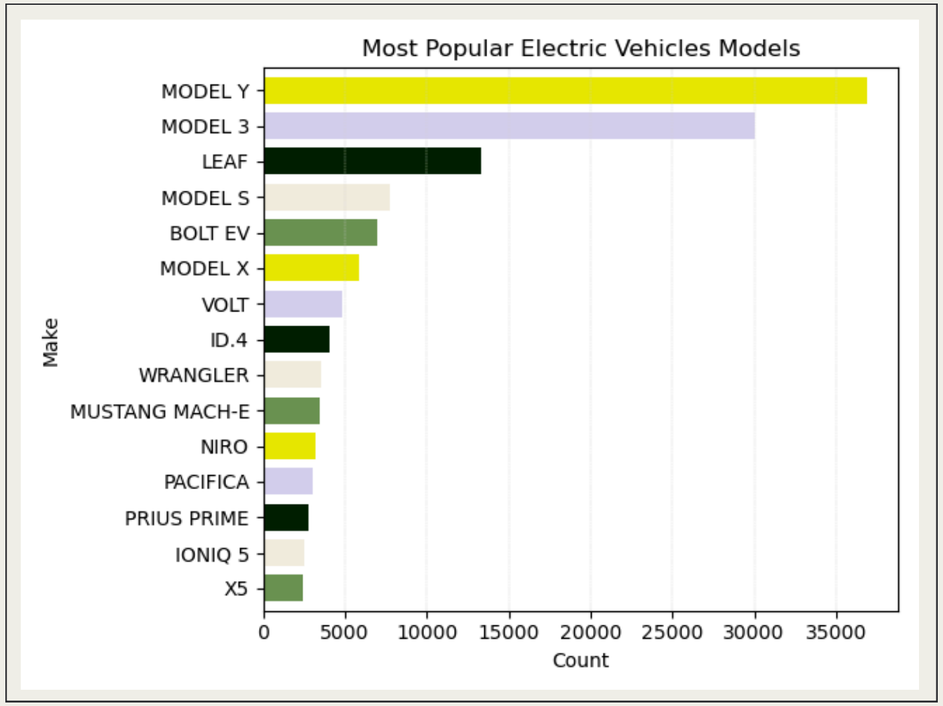
1. Dropped rows with missing location data
2. Isolated our data to only include vehicles in Washington State
3. Converted our “Vehicle Location” column into usable data:
    a. removed the “POINT” characters from the column
    b. once left with just the coordinates, we split each string into separate Longitude and Latitude columns
    c. Converted the data from strings to float
5. Converted the data in in “Postal Code” column from float to int
6. Dataframe was updated to only include columns relevant to analysis
7. Filtered out duplicate rows by removing duplicate VIN ID’s
8. Removed vehicles that had a value of 0 in the “Range” column

THESIS 1:

What are the most and least popular car make and models for each year based off of annual car registrations?



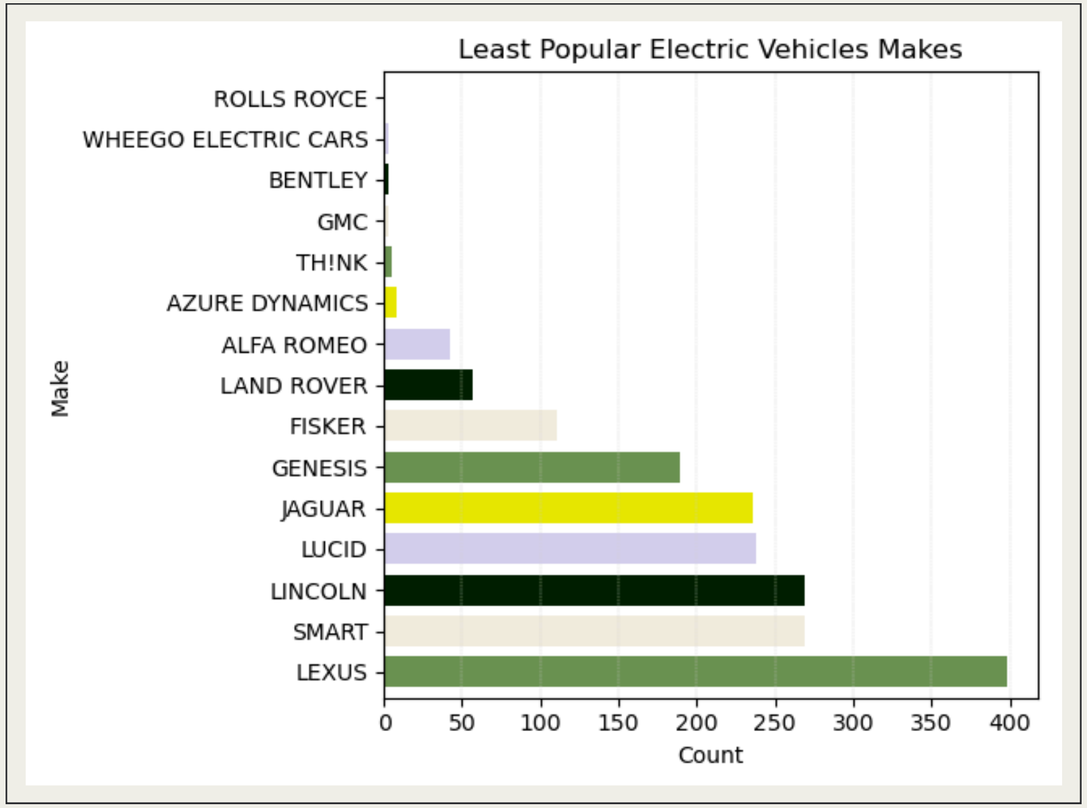
Based off of our analysis, Tesla was was the most popular EV brand. Three of their cars were in our top five most popular models when comparing the number of registrations per vehicle model with the Model Y being our overall most popular model.



This data proves what we had suspected in terms of Tesla's popularity and overall hold on the EV market.

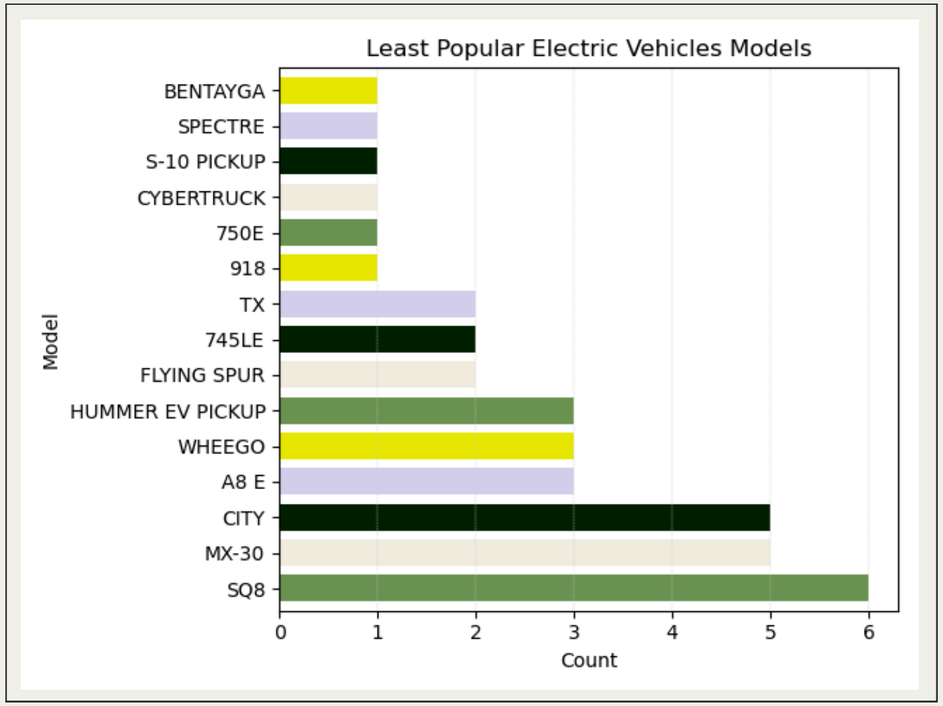
THESIS 1:

What are the most and least popular car make and models for each year based off of annual car registrations?



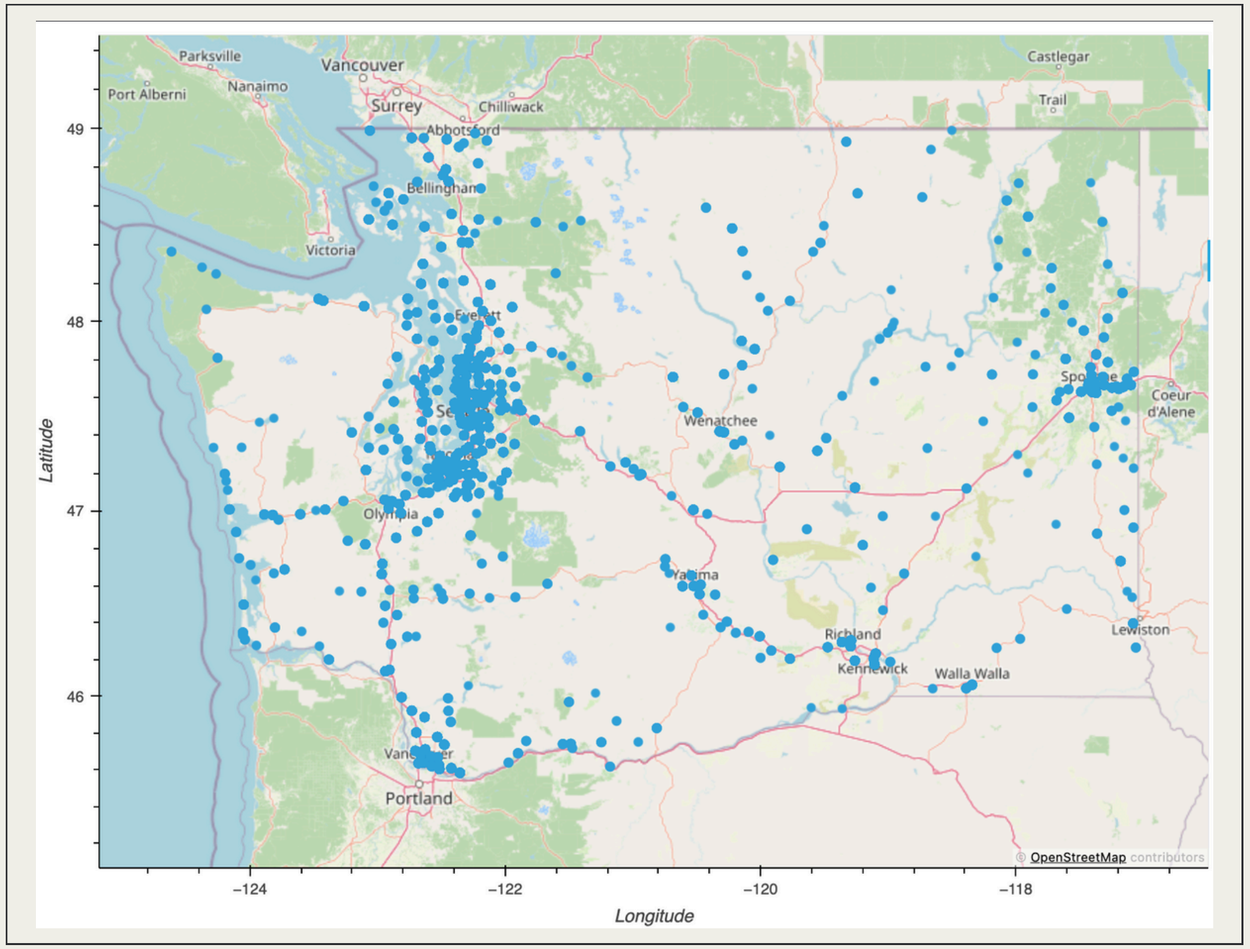
The trend we observed when looking at the least popular makes and models in terms of vehicle registrations is that most of our least popular makes only have one EV option on the market and thus those models ended up on our least popular list as well.

- Rolls Royce - Spectre
- Bentley - Bentayga

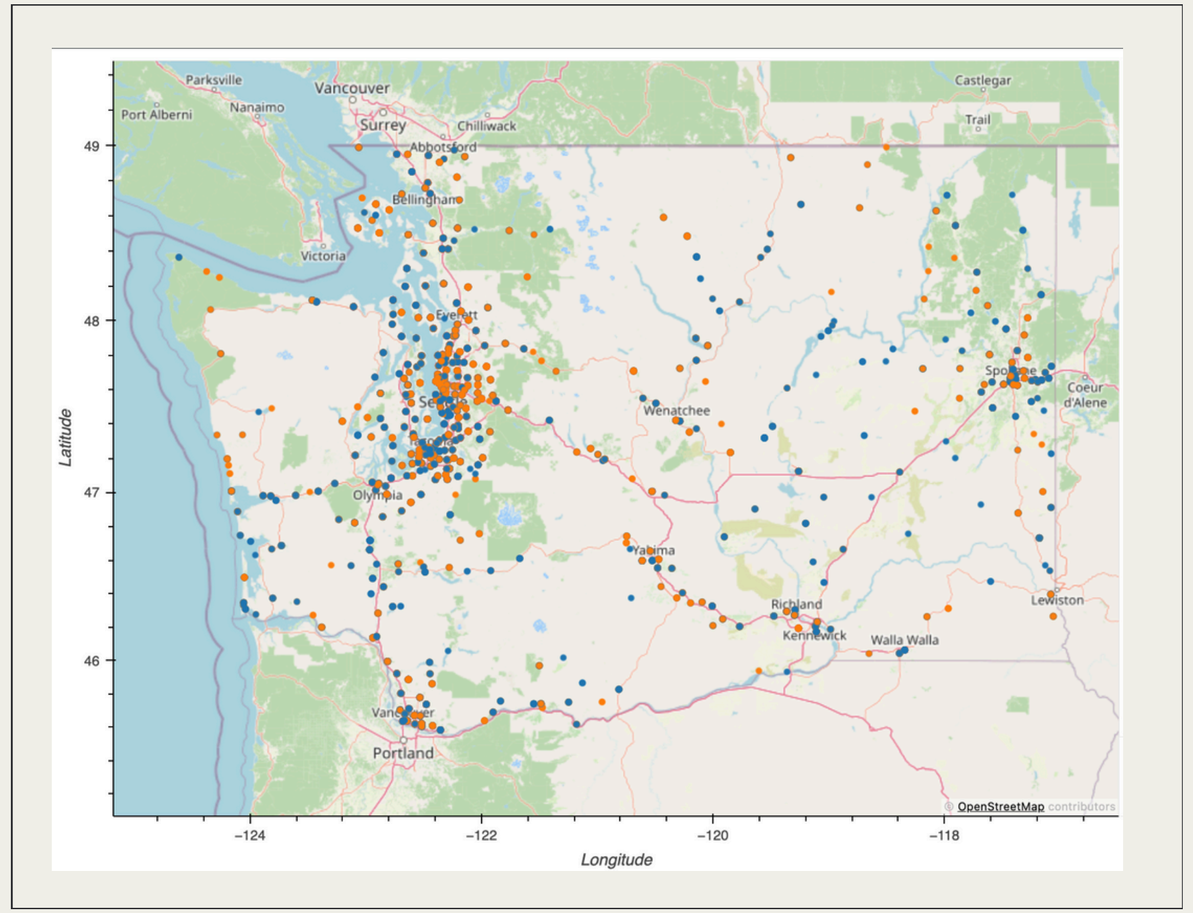


We also had a tie for our least popular EV models, with all of these only having one registration each in the state of Washington. The outlier here is the Tesla Cybertruck.

THESIS 2: Is there a relationship between geographic location and type of Electric Vehicle (i.e. BEV vs. Hybrid EV)?



Distr. of EV



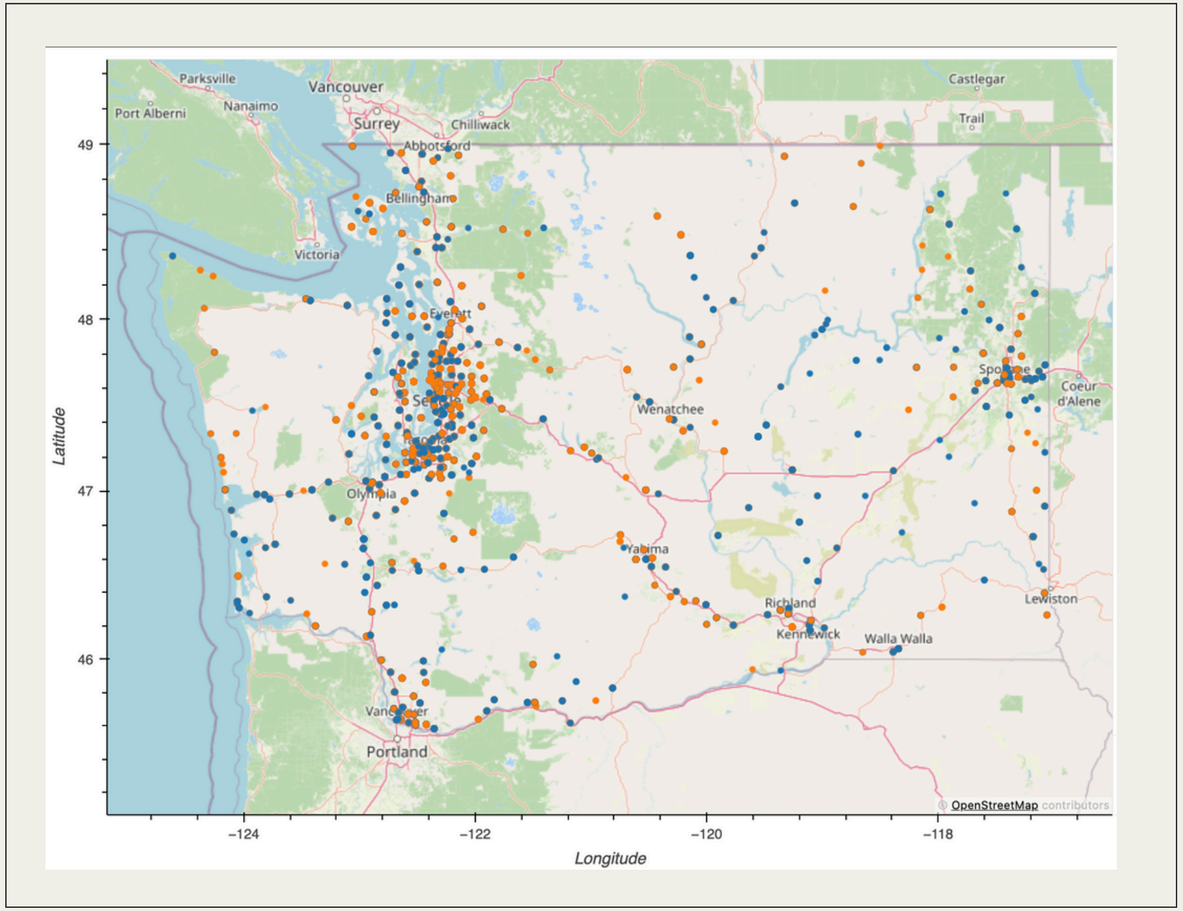
Distr. of EV by Vehicle Type

Plug-in Hybrid Electric Vehicle (PHEV)
Battery Electric Vehicle (BEV)

THESIS 2: Is there a relationship between geographic location and type of Electric Vehicle (i.e. BEV vs. Hybrid EV)?

Inconclusive:

- Based off our current dataset, our findings were inconclusive
- “Somewhat” even distribution of both BEVs and Hybrids in the same geographic clusters/locations
- Need population, economic, and vehicle MSRP/value data to form a more conclusive answer

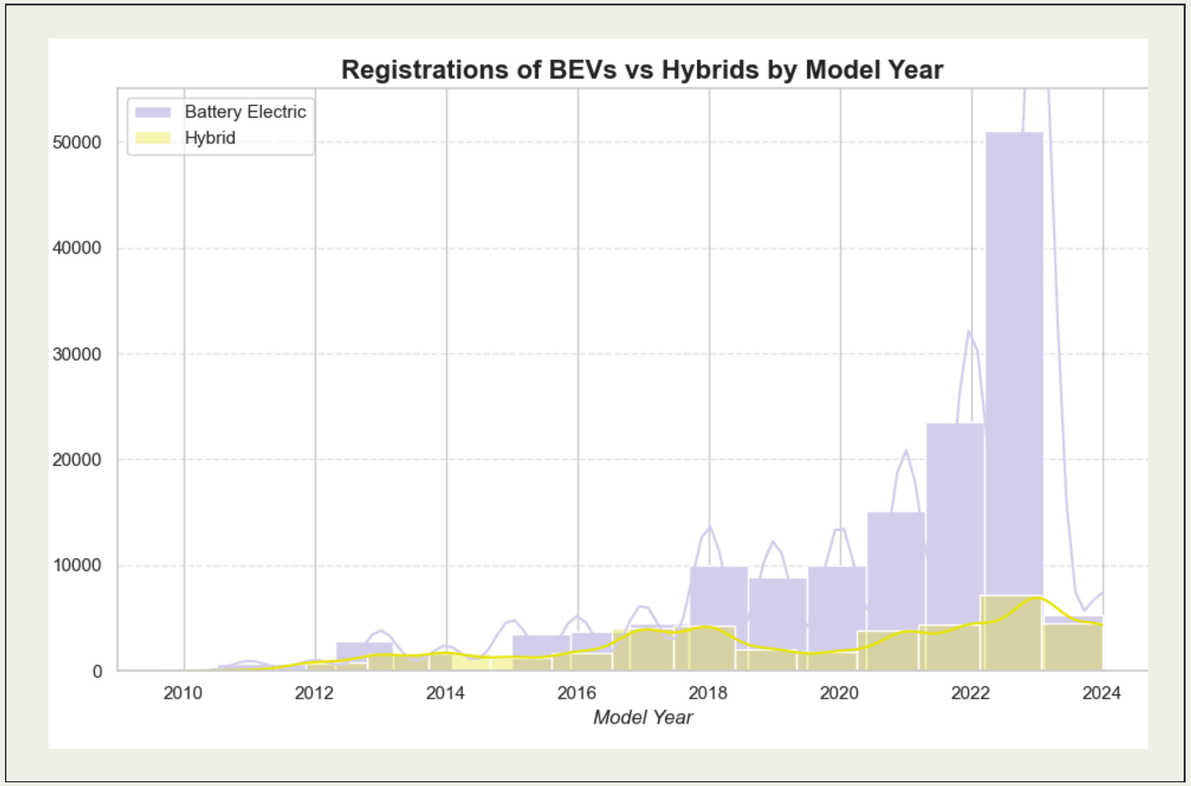


Distr. of EV by Vehicle Type

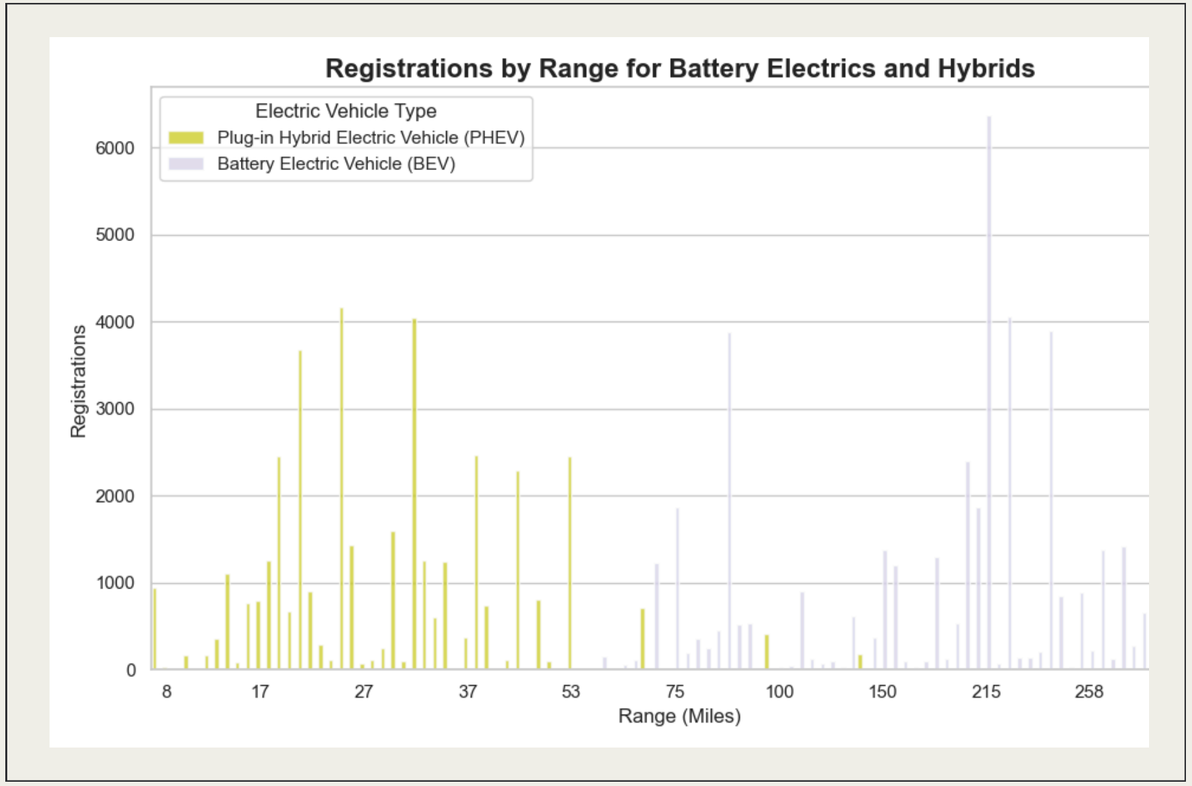
THESIS 3: Is there a correlation between increase in battery range and annual car registration based off of the model year?

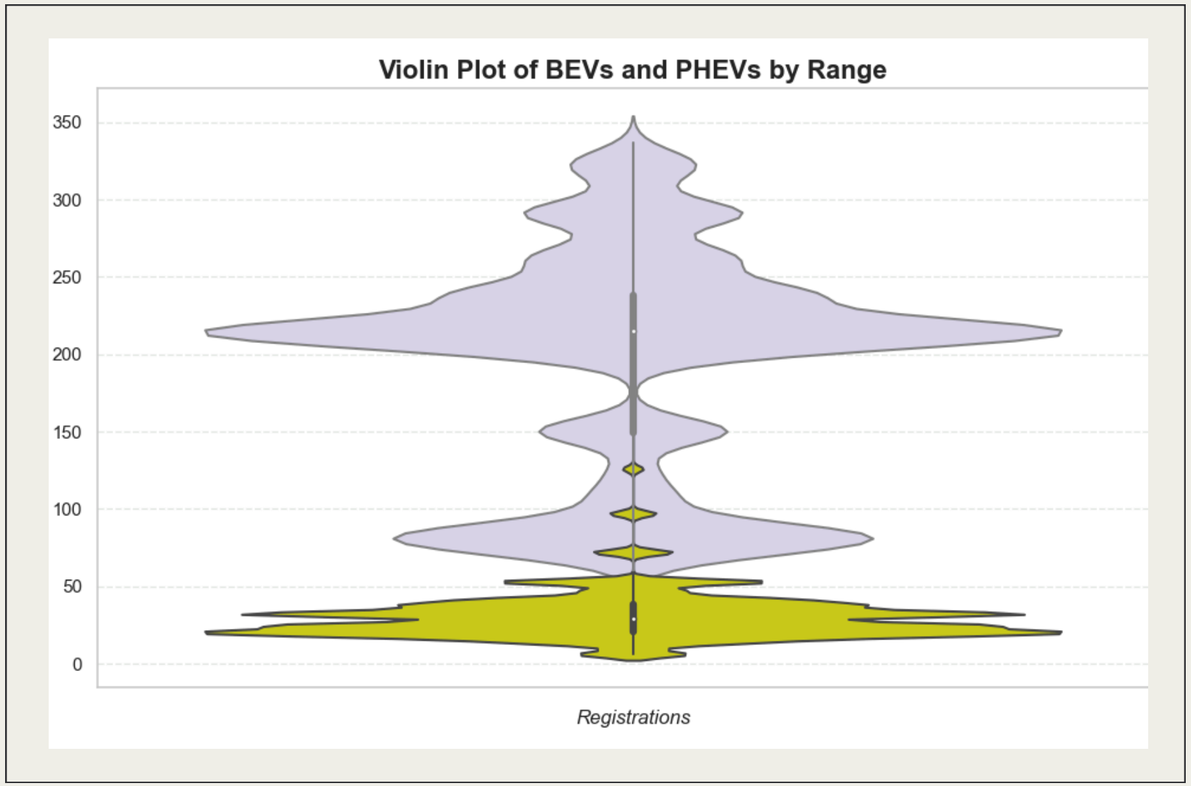
Relative:

- At this point we cannot gauge an answer for our overall data, inclusive of both BEVs and Hybrids
- There is a clear correlation in the relationship between battery range and model year for BEVs
    - Significant increase in annual car registrations as the average battery range increases
- Hybrids show no clear relationship between car registrations and battery ranges
    - The lack of improvement in battery range could be explanatory of this



- BEVs and Hybrids have similar registration counts at certain points
- Hybrids have a much smaller range of 10 to 53 miles with most falling between 20-35 miles
- Battery EVs have a much wider range of 75 to 330 miles

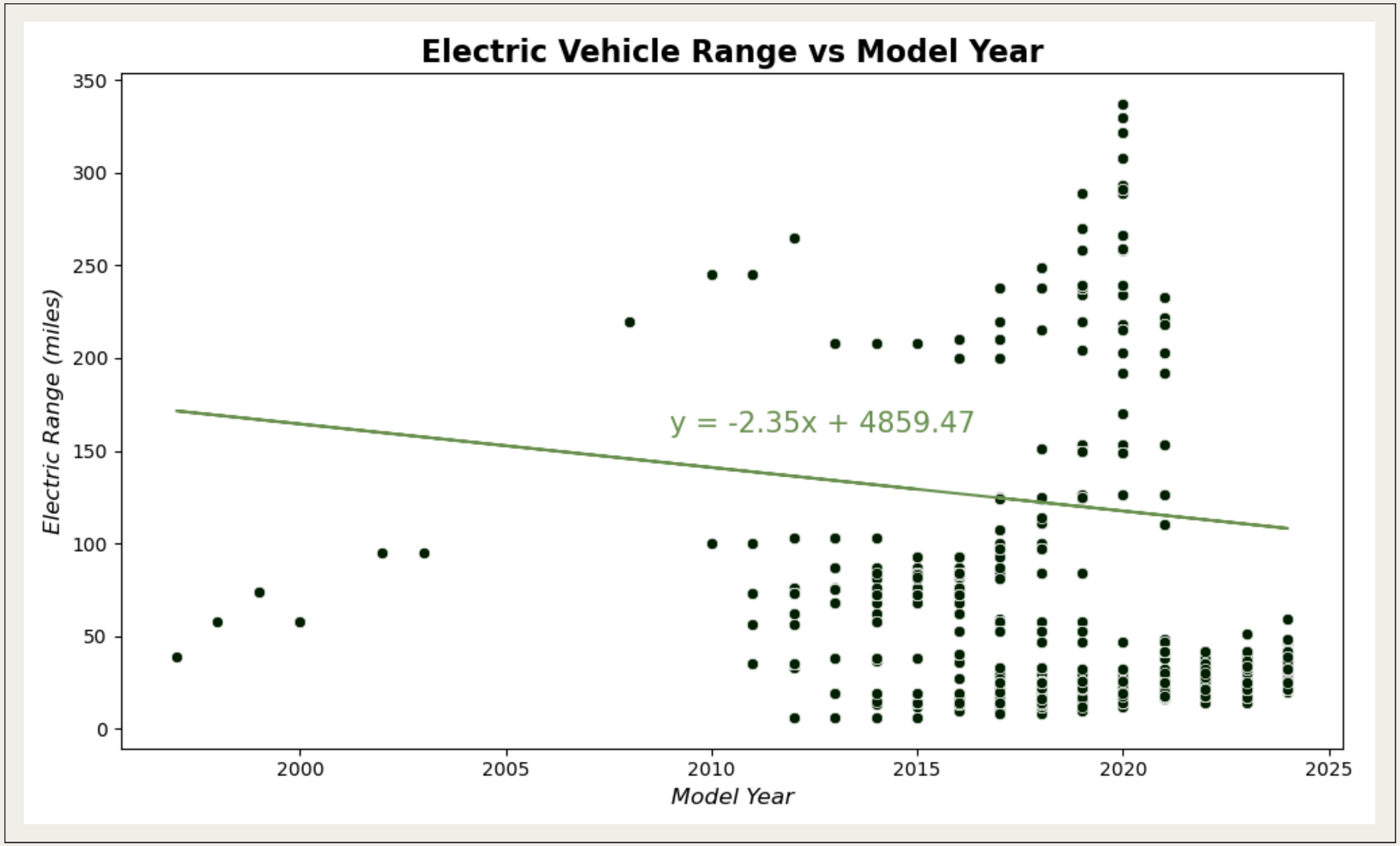


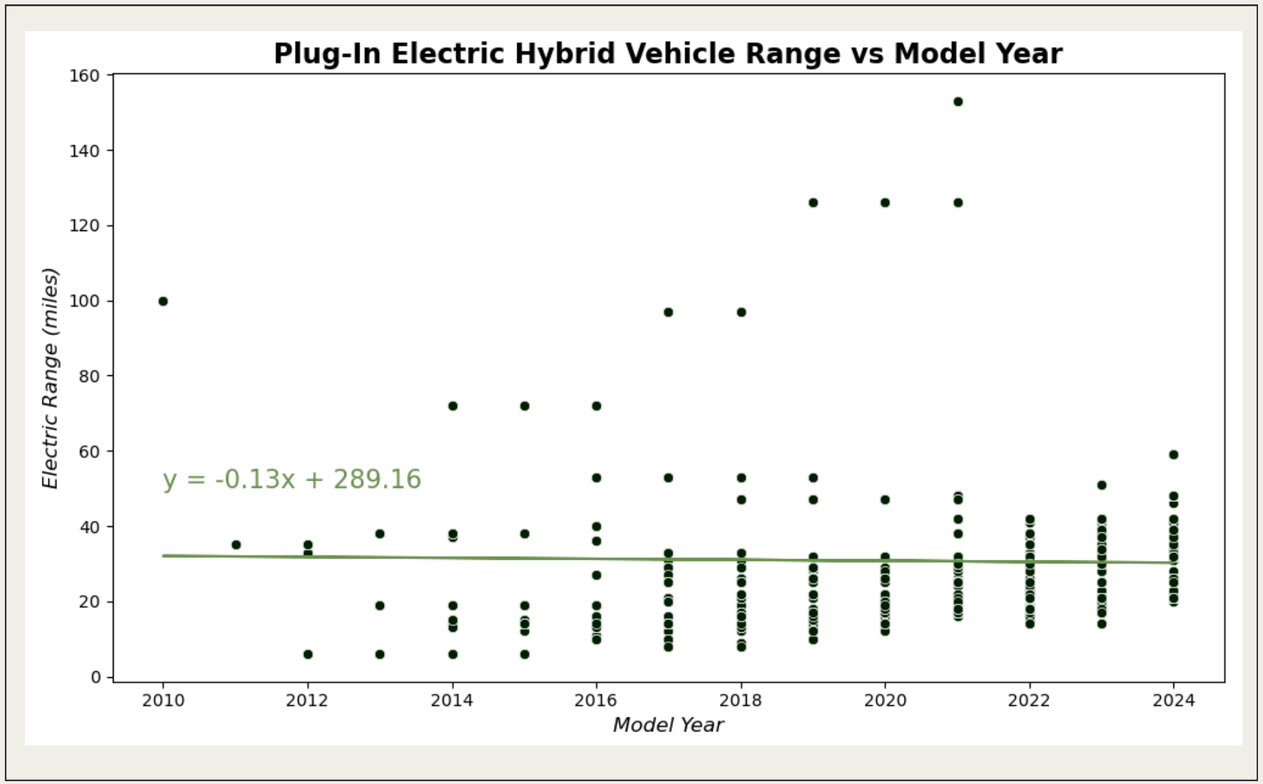


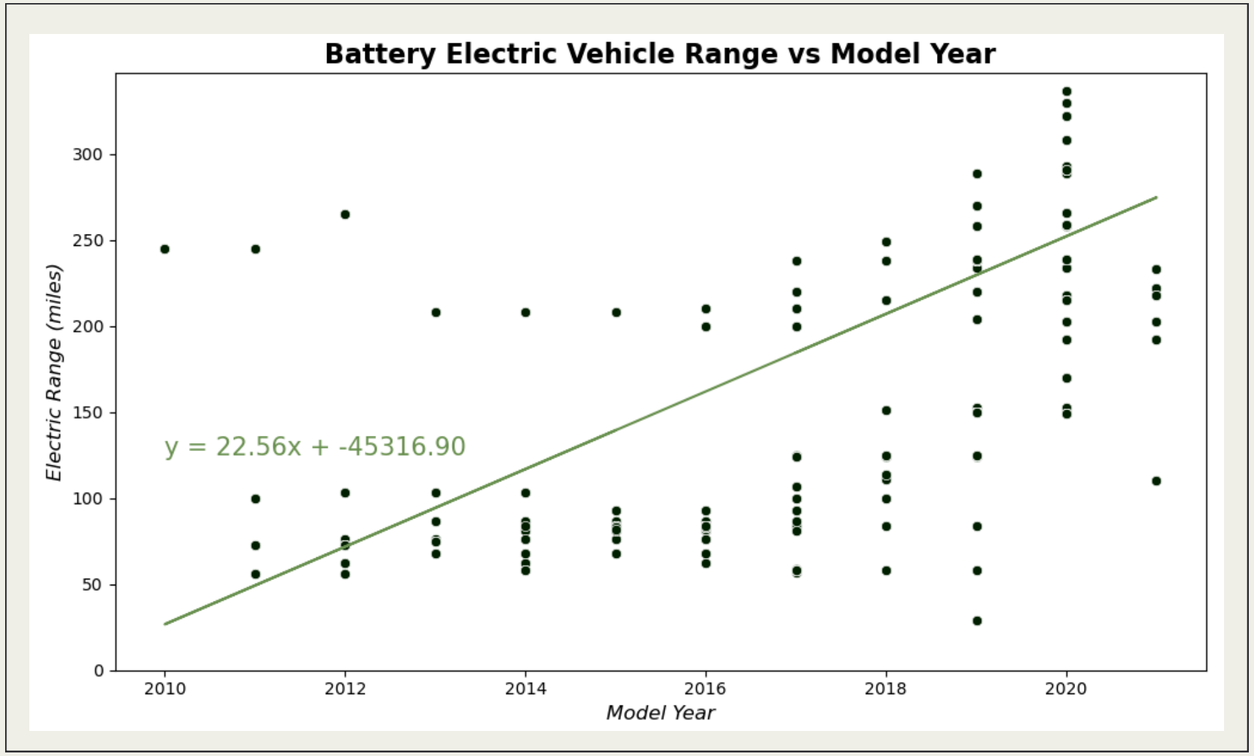
- There are three distinct points of registration increase within our BEV range with the largest being around when BEVs reached a range of 200 miles
  - This goes to show the positive relationship between range increase and car registrations for BEVs

THESIS 3: Is there a correlation between increase in battery range and annual car registration based off of the model year?

Yes:




Plug-In Electric Hybrid Vehicle Range vs Model Year
Electric Range (miles)
Model Year
y = -0.13x + 289.16
0
20
40
60
80
100
120
140
160
2010
2012
2014
2016
2018
2020
2022
2024


Battery Electric Vehicle Range vs Model Year
Electric Range (miles)
Model Year
y = 22.56x + -45316.90
0
50
100
150
200
250
300
2010
2012
2014
2016
2018
2020

SUMMARY PAGE

- Currently BEVs make up the majority of car registrations
- Our data shows that BEVs are continuing to take up a higher percentage of the overall car registrations annually as their ranges increase
- Our top 5 models were all BEVS
- As predicted, Tesla was the most popular make and held three of the most popular models in terms of vehicle registrations
- We weren’t able to provide a conclusive and definitive answer with our second thesis as a result of dataset limitations
- Our next steps would be to incorporate APIs with socio-economic data to more clearly determine if there is a correlation between location and the type of electric vehicle people are selecting.

BIASES + LIMITATIONS

Limitation
No Population Data

- We could sort our data by zip code, county or city but did not have information such as the population or average income for those cities, counties and zip codes.
- This made it difficult to understand if people were selecting EV's based off of their income or location.

Limitation
Incomplete Columns

- Our MSRP column had some values but for the majority of our rows, we only had '0' values
- As a result, we had to drop this column during our data cleaning and revise one of our thesis questions.
- Without this data we were unable to address or incorporate any financial questions

Limitation/Bias
Limited BEV Data

- During data cleaning, we removed all rows that had a battery range of '0' to have a more accurate reflection of the actual battery ranges
- Our cleaned dataset did not include any ranges for BEV's from 2022 to 2024.
- As a result, our data for BEV's is very limited.
    - Challenging to gauge the actual improvement in battery range for BEVs
    - Limited data may explain the bias that Hybrids hold in our linear regression resulting in a negative graph

Limitation
Incomplete Registration

- Our Registration Year column was only partially completed
- As a result we had to remove rows that didn't include a registration year
- We had registration data for 2023 for all vehicles but not for 2022 or 2024
- This meant we couldn't compare year over year change

FUTURE WORK PAGE

- Open Charge Map API
  - Could test hypothesis on vehicle registration locations based on the proximity of nearest charging stations
- VinAudit.com
- Vehicle Market Value API
  - Could test hypothesis on vehicle registration locations based on vehicle MSRP/value.
- Census.gov
- Economic Census API & Poverty Statistics API
  - Could test hypothesis on vehicle registration locations based on household income.
- Quarterly Workforce Indicators API
  - Could test hypothesis on vehicle registration locations based on county workforce financial statistics

WORKS CITED

- dataset
  - https://www.kaggle.com/datasets/jainaru/electric-vehicle-population
- Stack Overflow
- Pandas Documentation
  - https://pandas.pydata.org/pandas-docs/stable/index.html
- Previous Classwork + Assignments
- EV statistic sources:
  - https://cars.usnews.com/cars-trucks/advice/how-does-the-electric-car-tax-credit-work#:~:text=Since%202008%2C%20the%20federal%20government,hybrid%20electric%20vehicle%20(PHEV)
  - https://evadoption.com/2019-us-ev-sales-decreased-an-estimated-7-to-9-6-reasons-why/
  - https://en.wikipedia.org/wiki/Government_incentives_for_plug-in_electric_vehicles

Questions