Popularity Rise of EVs

Stephen Dong

I will be exploring the two open datasets provided by Data.gov on the population of EVs across the United States. The first dataset will be used to explore how the popularity of EVs have changed since 2017, around when many EVs became affordable as material costs became cheaper. The trend of popularity of EVs will be visualized and analyzed and used to determine the form of regression to be used to predict popularity of EVs in the upcoming years.

The second dataset will be used to analyze the distribution of the actual cars across the United States as of 2023. The first area of exploration on how distributed EV ownership is across the country. This will be visualized first through a regional geo heat map, with each state having only one solid color, to analyze the popularity of EV by state. This is followed by a second hotspot geo heat map visualization utilizing the distribution of counties within the United States to analyze where most purchases are located around. Although the first dataset also contains geographical data, the two cannot be merged as there is no information to guarantee data points in the datasets are not by the same person/respondent, which in the long run could skew the results.

There have also been many brands of EVs produced throughout the decade, making this an interesting opportunity to explore which car the EV population prefers. To do this, I will analyze the top brand of each state and visualize this in a geographical representation. To build off this, I will also analyze the top brand of the entire country. Once a clear analysis of the distribution of EVs across the country has been made, the data can be used to build a predictor via a variation of geospatial clustering to determine which EV brand someone is likely to have given the location coordinate (based on the owner's zip code). The reasoning for this approach is due to the effect of location, and by association the cost of living, on which EV is owned. All predictions are assuming they own an EV. A 95%-5% training test split will be used to measure the accuracy.

Datasets:

1. https://catalog.data.gov/dataset/electric-vehicle-population-size-history-by-county
2. https://catalog.data.gov/dataset/electric-vehicle-population-data