Electric Vehicle Population Data – Data Cleaning, Analysis Summary and Visualization Techniques

Step by Step Process of Analysis using Python on Jupyter Notebook:

1. Loaded the data on Jupyter Notebook and observed the values of rows and column in raw data.
2. We found that the column 'Legislative District' had 463 null values so we drop the null values. Also after dropping those values, County, City, Postal Code, Electric Utility and 2020 Census Tract also had null values which were also dropped.
3. We performed univariate analysis on each column. Checked if the data type of the columns are in the right format. Columns which were not necessary for the analysis were been dropped.
4. During further observation, we understood that Clean Alternative Fuel Vehicle (CAFV) Eligibility has almost 127000 values, where the eligibility is not known due to battery range was not been researched and Electric Range column also has same 127000 values as 0, due to which the eligibility is not known.
5. After grouping the Electric Vehicle Type, we understood that Battery Electric Vehicle is the only Vehicle type that has 127086 values as zero. Almost 57% of the values are 0. And out of that 57% of zero values, 55% of them are the Make of Tesla. We removed all the 0 values as we did not want redundant info by aggregating the values.
6. Base MSRP was been dropped as it had 216482 rows as 0 value, along with Electric Utility and 2020 Census Tract.
7. Also with the help of feature engineering, we added a new column Category Range with values as Low, Medium, High as per the electric range.
8. We performed bivariate analysis and also checked if there are any outliers remaining in the data.
9. After cleaning the data, the cleaned dataset file was been saved on system for further visualization on Tableau.

Visualization on Tableau:

1. We created a line chart with Count of DOL Vehicle ID and Year of Vehicle registration to understand the trend of electric vehicle adoption over the years.
2. Then we created a vertical bar chart to check Vehicle Type distribution between Battery Electric Vehicle (BEV) and Plug-in Hybrid Electric Vehicle (PHEV) among the total number of vehicles in the data set.
3. We also created another vertical bar chart for spread of electric range across the manufacturing companies listed in the data. Tesla being the highest among all other manufacturing company in terms of electric range.
4. With these 3 visualisations we created a basic Dashboard to give a basic insights of what happened with electric vehicle market over the years.
5. Post Covid, that is after 2020, we saw a dip on the BEV market as compared to the previous years.