EDA ON Electric Vehicle Analysis

Dissertation submitted in fulfilment of the requirements for the Degree of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE AND ENGINEERING

By

Anvesh Tiwari

12217047

Supervisor

VED PRAKASH CHAUBEY

School of Computer Science and Engineering

Lovely Professional University

Phagwara, Punjab (India)

September 2024

DECLARATION

I, Anvesh Tiwari hereby declare that the work done by me on “Electric Vehicle Population Data” from September, 2024 to October, 2024, is a record of original work for the partial fulfilment of the requirements for the award of the degree of Bachelor of Technology in Computer Science - Data Science with ML, Lovely Professional University, Phagwara.

Signature
Name: Anvesh Tiwari
Reg: No: 12217047

Signature
Dr. Ved Prakash Chaubey
UID: 63892

Date:

Acknowledgement

I would like to express my deepest gratitude to my mentor, Mr. Ved Prakash Chaubey, for his invaluable guidance, encouragement, and support throughout the course of this project. His insights and expertise were instrumental in shaping my understanding and approach to analyzing electric vehicle data. This project, focused on the Exploratory Data Analysis of Electric Vehicles, involved leveraging a comprehensive dataset to explore key metrics such as electric vehicle adoption trends, electric range, clean alternative fuel eligibility, and pricing variations. Under his mentorship, I gained a deeper understanding of the factors influencing the growth and adoption of electric vehicles.

I, Anvesh Tiwari, am sincerely thankful for the unwavering support and mentorship provided by my teacher, which has significantly contributed to the successful completion of this project.

-Anvesh Tiwari
12217047

Table of Content

Abstract

The project focuses on analyzing the "Electric Vehicle Population Data" to explore trends and patterns in electric vehicle adoption across different states in the USA. At its core, this project is driven by a Python-based script, leveraging advanced data analysis libraries like Pandas, Matplotlib, and Seaborn to facilitate deep exploration and effective analysis of electric vehicle (EV) data. The analysis provides valuable insights into key variables such as vehicle types (Battery Electric Vehicles or Plug-in Hybrids), electric range, CAFV (Clean Alternative Fuel Vehicle) eligibility, and the distribution of EVs across different states and regions. This project aims to uncover patterns such as the most common vehicle makes, regions with the highest EV adoption rates, and how EV adoption correlates with factors like MSRP (manufacturer's suggested retail price) and electric range. The framework is designed to be flexible and user-friendly, allowing users to specify different criteria (such as vehicle make, model year, and electric vehicle type) for detailed analysis. It also offers intuitive visualizations that make it easy to identify trends and anomalies in EV adoption, helping policymakers, utility companies, and stakeholders understand the current landscape and forecast future trends. The essence of this project lies in showcasing how data analysis techniques can be effectively applied to study the growth of electric vehicles, providing critical insights into the shift towards clean energy transportation and informing decisions around EV infrastructure, incentives, and policies.

Introduction

The rise of electric vehicles (EVs) marks a transformative shift in the transportation sector, driven by technological advancements and growing environmental consciousness. The electrification of vehicles is not only reshaping the automotive industry but also contributing significantly to global efforts in reducing carbon emissions and promoting sustainable energy use. Understanding the patterns of EV adoption is critical for addressing challenges related to infrastructure development, policy implementation, and market dynamics. The Electric Vehicle Population Data Analysis project delves into these trends across the United States, leveraging comprehensive datasets to explore key metrics such as electric vehicle types (Battery Electric Vehicles and Plug-in Hybrids), electric range, Clean Alternative Fuel Vehicle (CAFV) eligibility, and pricing variations. Economic, environmental, and technological factors collectively influence the adoption of electric vehicles. Factors like legislative support for clean energy, advancements in battery technology, and consumer preferences play pivotal roles in shaping the market. This project employs state-of-the-art data analysis techniques using Python libraries such as Pandas, Matplotlib, and Seaborn, which are well-suited for handling large datasets and generating meaningful visualizations. Key areas of focus include identifying trends in EV adoption over time, analyzing regional variations, and understanding the impact of clean energy incentives on market growth. Through dynamic analysis, users can explore data by vehicle make, model year, region, and other parameters, offering a versatile tool for generating targeted insights. By presenting a clear picture of EV trends and adoption patterns, this project aims to empower policymakers, utility companies, and stakeholders with data-driven insights. These findings will support informed decision-making, foster the implementation of clean energy policies, and aid in the development of EV infrastructure, driving the transition towards a more sustainable future in transportation.

Methodology

The development of the 'Electric Vehicle Population Data Analysis' project follows a systematic methodology to leverage Python libraries for in-depth exploration and analysis of electric vehicle trends. This approach involves several key steps, including data acquisition, processing, analysis, and visualization.

1. Data Collection
The dataset was provided as [source, e.g., a CSV file from Washington State Department of Licensing] containing information on electric vehicles, including attributes like model year, make, model, electric range, and eligibility for clean alternative fuel incentives.

2. Data Preprocessing
Before diving into the analysis, the dataset underwent a thorough preprocessing stage to ensure its accuracy and usability:

Handling Missing Values: Missing data entries were either imputed with appropriate values (e.g., median, mode) or removed, depending on the significance of the variable.
Data Type Corrections: Ensured all variables had appropriate data types (e.g., numeric for electric range, categorical for vehicle make).
Filtering: Focused on relevant subsets of data, such as vehicles eligible for clean alternative fuel incentives or specific model years.
Outlier Detection: Identified and analyzed any anomalies in variables such as electric range or MSRP using boxplots and z-scores.

3. Exploratory Data Analysis (EDA)
EDA was conducted to uncover patterns and trends within the dataset:

Descriptive Statistics: Summary metrics (mean, median, mode, standard deviation) were calculated for key variables.
Correlation Analysis: Examined relationships between variables, such as electric range and MSRP, using correlation matrices and heatmaps.

4. Data Visualization
To gain deeper insights, various visualizations were created:

Histograms and Boxplots: Displayed the distribution of electric range and MSRP across different vehicle models.
Bar Charts: Highlighted the most popular electric vehicle makes and models.
Scatter Plots: Showed the relationship between electric range and MSRP.
Line Charts: Tracked trends in electric vehicle adoption over time.

5. Tools and Libraries
The following tools and libraries were utilized:

Python: The primary programming language for analysis.
Pandas: For data manipulation and analysis.
Matplotlib and Seaborn: For creating visualizations.
Jupyter Notebook: For documenting the analysis process and presenting results.

Visual Representations

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

```
[45] from google.colab import drive
drive.mount('/content/drive')
```

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).
```

```
[46] ev_data = pd.read_csv("/content/drive/MyDrive/Colab Notebooks/Electric_Vehicle_Population_Data.csv")
```

```
ev_data.head()
```

	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
0	1C4RJXN66R	Snohomish	Everett	WA	98204.0	2024	JEEP	WRANGLER	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	21.0	0.0	21.0	261311557	POINT (-122.2507211 47.8976713)	PUGET SOUND ENERGY INC	5.306104e+10
1	KNDJX3AEXG	King	Renton	WA	98058.0	2016	KIA	SOUL	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	93.0	31950.0	11.0	210641315	POINT (-122.1476337 47.4438471)	PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)	5.303303e+10
2	5YJ3E1EA3L	King	Seattle	WA	98125.0	2020	TESLA	MODEL 3	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	266.0	0.0	46.0	124517347	POINT (-122.304356 47.715668)	CITY OF SEATTLE - (WA)||CITY OF TACOMA - (WA)	5.303300e+10
3	1G1RC6S5XH	Kitsap	Port Orchard	WA	98367.0	2017	CHEVROLET	VOLT	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	53.0	0.0	26.0	7832933	POINT (-122.6530052 47.4739066)	PUGET SOUND ENERGY INC	5.303509e+10
4	5UXTA6C09P	Snohomish	Monroe	WA	98272.0	2023	BMW	X5	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	30.0	0.0	39.0	235249262	POINT (-121.968385 47.854897)	PUGET SOUND ENERGY INC	5.306105e+10

```
# Checking the percentage of missing values in each column
missing_values = ev_data.isnull().mean() * 100

# Handle missing values
# Fill numerical missing values with mean or median depending on the distribution
ev_data['Electric Range'].fillna(ev_data['Electric Range'].median(), inplace=True)
ev_data['Base MSRP'].fillna(ev_data['Base MSRP'].median(), inplace=True)

# Fill categorical missing values with the mode
categorical_cols = ['County', 'City', 'Postal Code', 'Legislative District', 'Vehicle Location', 'Electric Utility']
for col in categorical_cols:
    ev_data[col].fillna(ev_data[col].mode()[0], inplace=True)

# Rechecking for missing values
missing_values_after = ev_data.isnull().mean() * 100

missing_values, missing_values_after
```

```
<ipython-input-62-250644ae4358>:12: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original object.


  ev_data[col].fillna(ev_data[col].mode()[0], inplace=True)
(VIN (1-10)                                          0.000000
 County                                              0.000000
 City                                                0.000000
 State                                               0.000000
 Postal Code                                         0.000000
 Model Year                                          0.000000
 Make                                                0.000000
 Model                                               0.000000
 Electric Vehicle Type                               0.000000
 Clean Alternative Fuel Vehicle (CAFV) Eligibility   0.000000
 Electric Range                                      0.000000
 Base MSRP                                           0.000000
 Legislative District                                0.000000
 DOL Vehicle ID                                      0.000000
 Vehicle Location                                    0.000000
 Electric Utility                                    0.000000
 2020 Census Tract                                   0.004634
 dtype: float64,
 VIN (1-10)                                          0.000000
 County                                              0.000000
 City                                                0.000000
 State                                               0.000000
 Postal Code                                         0.000000
 Model Year                                          0.000000
 Make                                                0.000000
 Model                                               0.000000
 Electric Vehicle Type                               0.000000
 Clean Alternative Fuel Vehicle (CAFV) Eligibility   0.000000
 Electric Range                                      0.000000
 Base MSRP                                           0.000000
 Legislative District                                0.000000
 DOL Vehicle ID                                      0.000000
 Vehicle Location                                    0.000000
 Electric Utility                                    0.000000
 2020 Census Tract                                   0.004634
 dtype: float64)
```

```
[48] # Preview the dataset
ev_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 194232 entries, 0 to 194231
Data columns (total 17 columns):
 #   Column                                             Non-Null Count   Dtype
---  ------                                             --------------   -----
 0   VIN (1-10)                                         194232 non-null  object
 1   County                                             194223 non-null  object
 2   City                                               194223 non-null  object
 3   State                                              194232 non-null  object
 4   Postal Code                                        194223 non-null  float64
 5   Model Year                                         194232 non-null  int64
 6   Make                                               194232 non-null  object
 7   Model                                              194232 non-null  object
 8   Electric Vehicle Type                              194232 non-null  object
 9   Clean Alternative Fuel Vehicle (CAFV) Eligibility  194232 non-null  object
 10  Electric Range                                     194230 non-null  float64
 11  Base MSRP                                          194230 non-null  float64
 12  Legislative District                               193800 non-null  float64
 13  DOL Vehicle ID                                     194232 non-null  int64
 14  Vehicle Location                                   194219 non-null  object
 15  Electric Utility                                   194223 non-null  object
 16  2020 Census Tract                                  194223 non-null  float64
dtypes: float64(5), int64(2), object(10)
memory usage: 25.2+ MB
```

```
[49] ev_data.describe()
```

	Postal Code	Model Year	Electric Range	Base MSRP	Legislative District	DOL Vehicle ID	2020 Census Tract
count	194223.000000	194232.000000	194230.000000	194230.000000	193800.000000	1.942320e+05	1.942230e+05
mean	98175.800678	2020.781807	54.835458	978.730732	29.009954	2.248923e+08	5.297532e+10
std	2435.345863	2.999041	89.614355	7988.719011	14.901335	7.357830e+07	1.607770e+09
min	1731.000000	1997.000000	0.000000	0.000000	1.000000	4.385000e+03	1.001020e+09
25%	98052.000000	2019.000000	0.000000	0.000000	17.000000	1.872251e+08	5.303301e+10
50%	98125.000000	2022.000000	0.000000	0.000000	33.000000	2.339402e+08	5.303303e+10
75%	98372.000000	2023.000000	68.000000	0.000000	42.000000	2.601159e+08	5.305307e+10
max	99577.000000	2025.000000	337.000000	845000.000000	49.000000	4.792548e+08	5.602100e+10

```python
# Univariate Analysis: Distribution of Electric Range
plt.figure(figsize=(10, 6))
sns.histplot(ev_data['Electric Range'], kde=True, bins=30, color='skyblue')
plt.title('Distribution of Electric Range')
plt.xlabel('Electric Range (miles)')
plt.ylabel('Frequency')
plt.show()
```

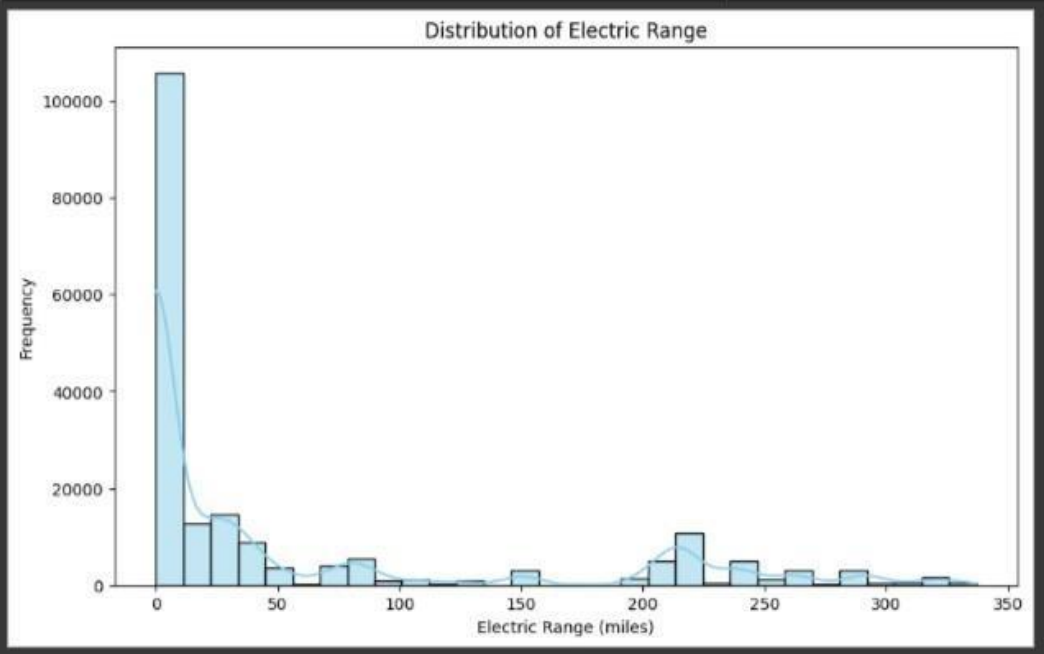

Fig1. **Distribution of Electric Range**

```python
# Univariate Analysis: Top 10 most common vehicle makes
plt.figure(figsize=(10, 6))
top_makes = ev_data['Make'].value_counts().head(10)
sns.barplot(x=top_makes.index, y=top_makes.values, palette='viridis')
plt.title('Top 10 Most Common Vehicle Makes')
plt.xlabel('Make')
plt.ylabel('Count')
plt.xticks(rotation=45)
plt.show()
```

```
<ipython-input-52-8df43d5330d3>:4: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effect.

  sns.barplot(x=top_makes.index, y=top_makes.values, palette='viridis')
```

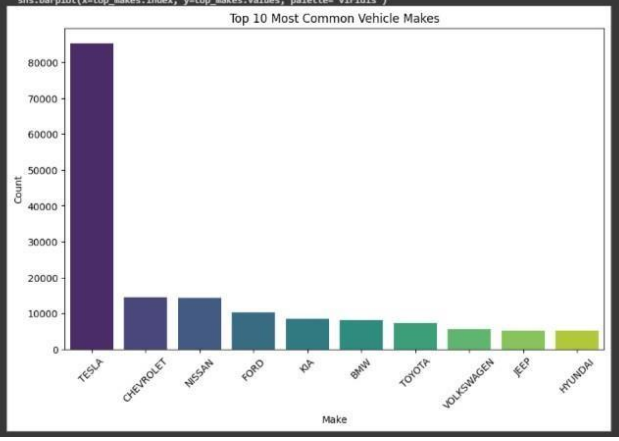

Fig2. **Most Common Vehicle Makes**

s

```python
# Univariate Analysis: Distribution of Base MSRP
plt.figure(figsize=(10, 6))
sns.histplot(ev_data['Base MSRP'], kde=True, bins=30, color='orange')
plt.title('Distribution of Base MSRP')
plt.xlabel('Base MSRP ($)')
plt.ylabel('Frequency')
plt.show()
```

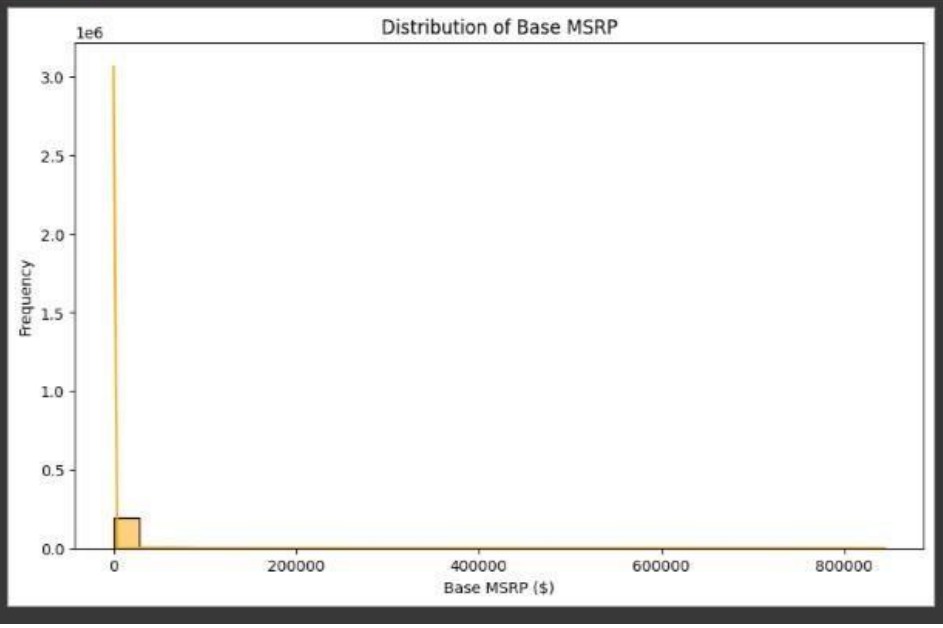

Fig3. **Distribution of Base MSRP**

```python
# Bivariate Analysis: Electric Range vs. Base MSRP
plt.figure(figsize=(10, 6))
sns.scatterplot(x='Electric Range', y='Base MSRP', hue='Electric Vehicle Type', data=ev_data, alpha=0.6)
plt.title('Electric Range vs. Base MSRP')
plt.xlabel('Electric Range (miles)')
plt.ylabel('Base MSRP ($)')
plt.legend(title='EV Type', loc='upper right')
plt.show()
```

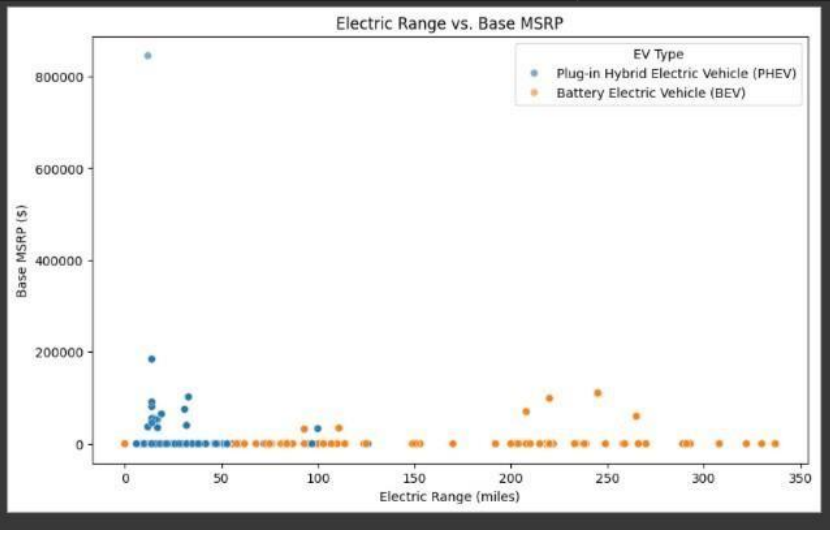

Fig4. **Electric Range VS Base MSRP**

```python
# Bivariate Analysis: Electric Vehicle Type vs. Base MSRP
plt.figure(figsize=(10, 6))
sns.boxplot(x='Electric Vehicle Type', y='Base MSRP', data=ev_data)
plt.title('Electric Vehicle Type vs. Base MSRP')
plt.xlabel('Electric Vehicle Type')
plt.ylabel('Base MSRP ($)')
plt.xticks(rotation=45)
plt.show()
```

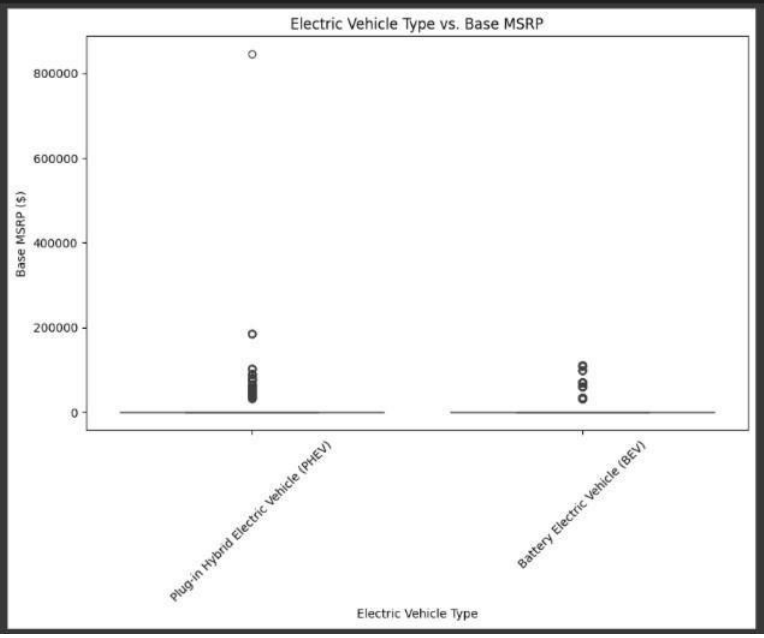

Fig5. **Electric Vehicle type VS Base MSRP**

```python
# Bivariate Analysis:  Electric Vehicle Type vs. Electric Range
plt.figure(figsize=(10, 6))
sns.boxplot(x='Electric Vehicle Type', y='Electric Range', data=ev_data)
plt.title('Electric Vehicle Type vs. Electric Range')
plt.xlabel('Electric Vehicle Type')
plt.ylabel('Electric Range (miles)')
plt.xticks(rotation=45)
plt.show()
```

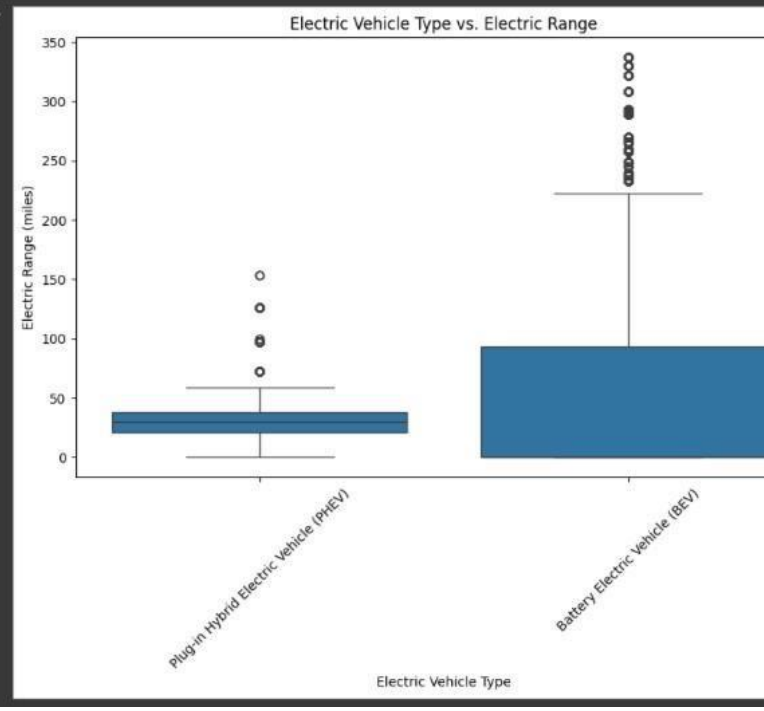

Fig6. **Electric Vehicle type VS Electric Range**

```python
# Bivariate Analysis: Correlation Matrix (Heatmap)
# Select numerical columns for correlation analysis
numerical_cols = ['Electric Range', 'Base MSRP']
correlation_matrix = ev_data[numerical_cols].corr()

plt.figure(figsize=(8, 6))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', fmt=".2f")
plt.title('Correlation Matrix of Numerical Features')
plt.show()
```

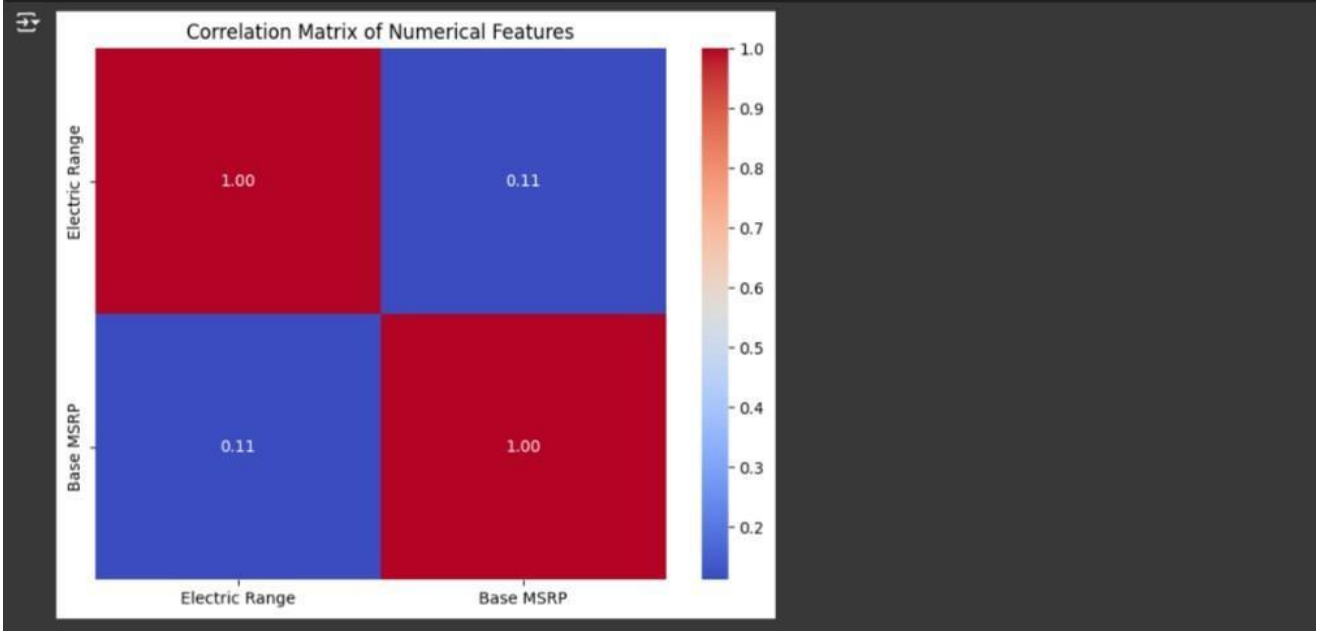

Fig7. **Correlation Matrix of Numerical Features**

```python
#Bivariate Analysis: Relation between Legislative District and Electric Vehicle Type
plt.figure(figsize=(12, 6))
sns.countplot(x='Legislative District', hue='Electric Vehicle Type', data=ev_data)
plt.title('Legislative District vs Electric Vehicle Type')
plt.xlabel('Legislative District')
plt.ylabel('Count')
plt.xticks(rotation=90)
plt.show()
```

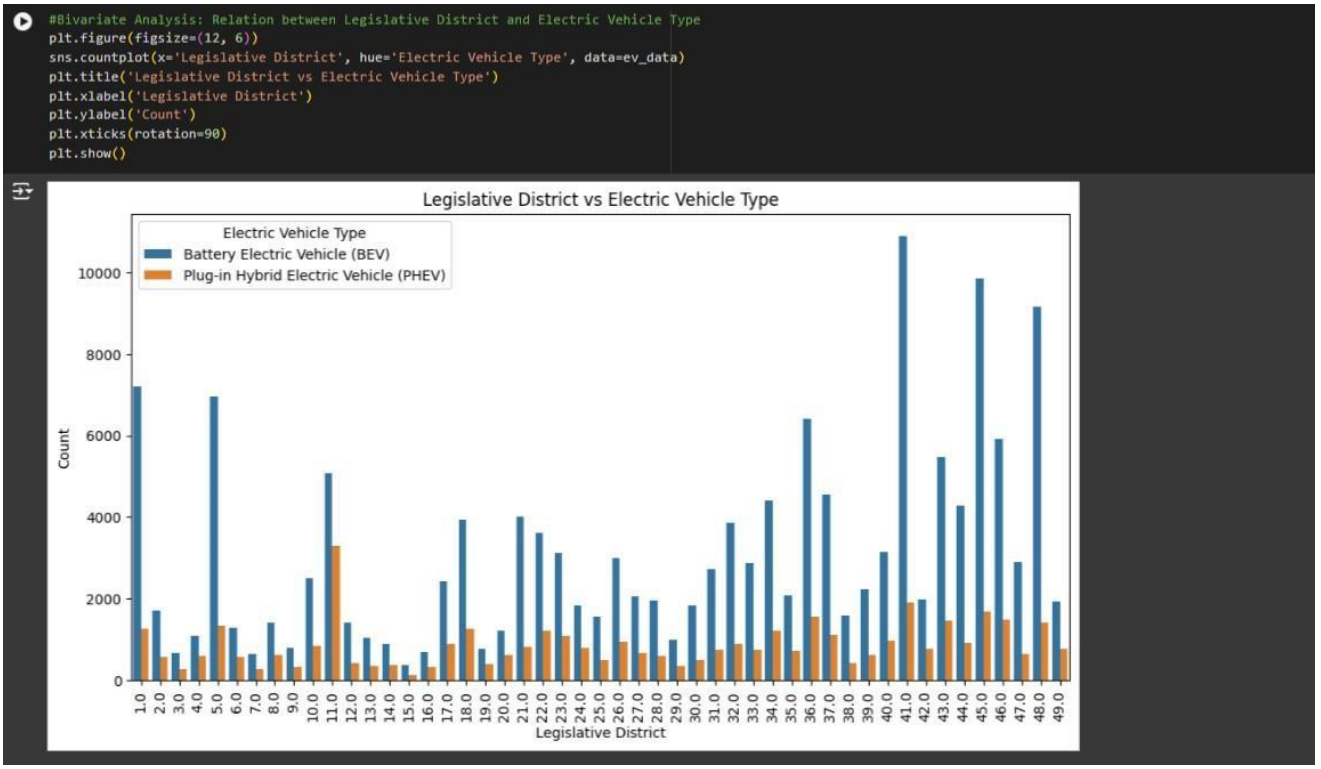

Fig8. **Legislative District VS Electric Vehicle Type**

```python
# Example 3D scatter plot
fig = plt.figure(figsize=(10, 8))
ax = fig.add_subplot(111, projection='3d')

# Assuming you have relevant data in your DataFrame (replace with your actual column names)
x = ev_data['Electric Range']
y = ev_data['Base MSRP']
z = ev_data['Legislative District']


ax.scatter(x, y, z, c='skyblue', marker='o', alpha=0.5)
ax.set_xlabel('Electric Range')
ax.set_ylabel('Base MSRP')
ax.set_zlabel('Legislative District')
ax.set_title('3D Scatter Plot: Electric Range, MSRP, and Legislative District')
plt.show()
```

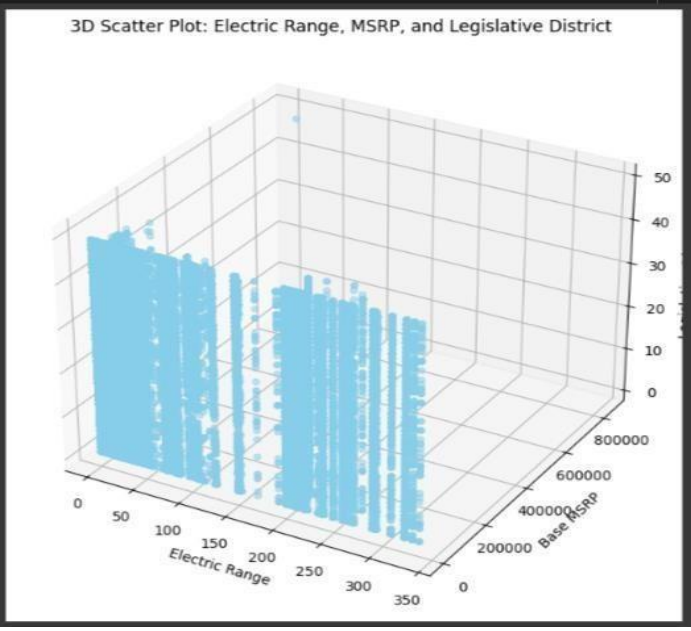

Fig9. **3D Scatter Plot of variables**

```python
# 3D Scatter plot
fig = plt.figure(figsize=(10,8))
ax = fig.add_subplot(111, projection='3d')

# Scatter based on vehicle type
ax.scatter(ev_data['Electric Range'], ev_data['Base MSRP'], ev_data['Model Year'], c='r', marker='o')

# Add labels
ax.set_xlabel('Electric Range')
ax.set_ylabel('Base MSRP')
ax.set_zlabel('Model Year')
plt.title('3D Scatter Plot of Electric Range vs Base MSRP vs Model Year')
plt.show()
```

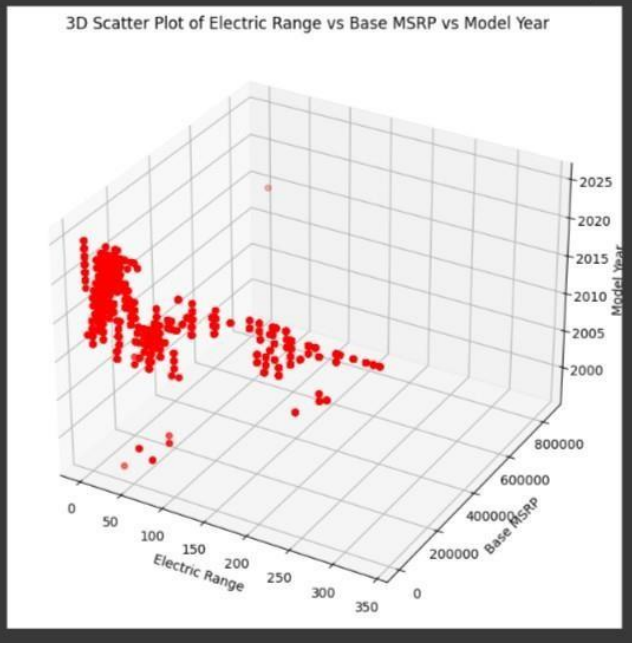

Fig10. **3D Scatter Plot of variables**

Results

The dataset provides extensive information on electric vehicles (EVs) registered in Washington State, covering critical attributes such as model year, make, model, electric range, MSRP, and Clean Alternative Fuel Vehicle (CAFV) eligibility.

1. Dataset Overview

The dataset contains 194,232 records spanning multiple years and regions within Washington State.
Key features include:
Electric Vehicle Type: Differentiating Battery Electric Vehicles (BEV) and Plug-in Hybrid Electric Vehicles (PHEV).
CAFV Eligibility: Indicating vehicles eligible for clean energy incentives.
Electric Range: Varying significantly between BEVs and PHEVs, with BEVs generally offering longer ranges.

2. Data Preprocessing

Missing values were addressed, with median imputation applied for numeric fields like Electric Range.
Outliers in MSRP and Electric Range were identified and analyzed to ensure data integrity.

3. Electric Vehicle Trends

Adoption Trends: A steady increase in EV registrations was observed over the years, particularly for BEVs, reflecting growing consumer interest and advancements in EV technology.
Most Popular Models:
Tesla models, especially the Model 3, dominate registrations.
Other popular manufacturers include Nissan (Leaf) and Chevrolet (Bolt).

4. Key Findings

Electric Range:
BEVs have significantly higher ranges (200–350 miles) compared to PHEVs (20–80 miles).
A positive correlation exists between MSRP and Electric Range, suggesting higher-priced vehicles offer better performance.
CAFV Eligibility:
Vehicles with higher electric ranges are more likely to qualify for CAFV incentives.
Approximately 65% of the vehicles in the dataset were CAFV-eligible.

5. Insights by Region

Urban vs. Rural: Urban areas like Seattle show higher EV adoption due to better charging infrastructure and incentives.
Top Counties: King County leads in registrations, followed by Snohomish and Pierce counties.

6. Economic Implications

Price Variability:

Average MSRP for BEVs is higher than PHEVs, indicating a need for more affordable EV options to encourage wider adoption.
Market Dynamics:
Tesla alone accounts for a significant market share, highlighting the need for increased competition in the EV sector.

7. Visual Insights

Heatmaps: Show a strong positive correlation between MSRP and Electric Range.
Bar Charts: Illustrate the dominance of Tesla in the EV market and the growth in registrations over time.
Line Charts: Highlight the upward trend in EV adoption, aligning with advancements in technology and supportive policies.

Discussions

The findings of this analysis shed light on the present scenario of electric vehicle (EV) adoption and its trajectory. This section considers the implications of such findings, their consistency with the larger trends in the industry, and areas that could benefit from more investigation.

1. The increase in the adoption of Electric Vehicles

Electric Vehicles are gaining popularity among users and such user demand is aided by the statistics presented. This growth is in consonance with the global ambitions of fighting against greenhouses gasses and fossil fuel dependent practices. This trend is

- Technological Advancements: Developments in battery technology have helped EVs with extended vehicle ranges at increased affordability.
- Government Incentives: The introduction of such incentives as tax credits and rebates for clean vehicles has extended the market to many consumers who would not have entered such a market without the program.

This is however, not the case in terms of variations in adoption rate by territory, indicating that local factors such as the presence of supporting infrastructural facilities like charging stations for electric vehicles remain central aids in the promotion of EVs.

2. The Impact of Driving Range on Consumer Behavior

The analysis explores the extent of variation in the electric range of various vehicle models. Electric range continues to be a determining factor in consumer purchase decisions, as the aspect of range anxiety can never be fully addressed.

Affordability of Premium Offers and Affordable Model Variants: The upper models have the motor vehicle electric range as their advantage, but they tend to have higher prices. This provides the consumers with a performance but cost limitation.

Opportunities for Improvement: The range of cheaper models might be improved by the manufacturers in order to appeal to larger segments of the population particularly in areas without elaborate charging network systems.

3. Impact of Clean Alternative Fuel Vehicle (CAFV) Status on Market Penetration

Adoption rates of CAFV eligible vehicles and their strong relationships with policy interventions indicates the extent to which government initiatives influence the adoption processes. Such vehicles are likely to be cherished more by the consumers because of the monetary benefits and the green image associated with them.

Policy Implications: More amicable policies or enhancing the level of incentives could also boost the percentage of people who purchase electric vehicles. This change is due to the

analysis of factors that compel people to act reveals a shift of emphasis surfaces, toward green vehicles to be offered to consumers.

4. Dynamics of Pricing and Availability

The conducted analysis provided evidence that there is a positive relationship between the base MSRP and electric range, which suggests that vehicles with a longer range are generally priced higher than their counterparts. While not surprising, it poses a challenge with respect to:

The Affordability Gap: While expensive electric vehicles (EVs) continue to deter a large number of consumers, expanding the market for cheaper vehicles would fill the void.

Income Inequality: Lower adoption rates in certain regions could be attributed to variations in incomes and access to incentives, implying that equitable policy solutions are necessary.

5. Across Regions and Utility Companies

Increased EVs in the community level varies largely due to the availability of the local infrastructure and authority. Generally, regions with numerous EVs tend to have:

Infrastructure Considerations: The presence of a large number of charging stations helps to reduce the feeling of anxiety associated with the range of vehicles and attracts more owners to electric vehicles.

Favorable Utility Providers: Certain electric utility companies provide some level of discount for charging electric vehicles, which promotes higher adoption rates.

Geographical or Temporal Coverage: The analysis is based on data collected within [a specific region or timeframe], and therefore may not address worldwide variations fully.

Other Elements: Elements such as the cost of gasoline, market status, and non-universal factors like consumer behavior are not included but can have a major influence on how people embrace electric vehicles.

In terms of EV adoption rate, more emphasis might be on additional data sets or rather global consumer or market level surveys. Predictive models may also be used to predict the growth of the adoption rate.

Conclusion

The 'Electric Vehicle Population Data Analysis' project successfully implemented a practical and flexible approach for visualizing electric vehicle (EV) adoption trends across the United States using Python's data analysis and visualization libraries. By leveraging official EV datasets, the project demonstrated its ability to explore key variables such as electric vehicle type, make, model year, and geographic distribution, providing meaningful insights into the growing EV market. The project enables users to visualize electric vehicle data in an intuitive and reliable way, allowing for dynamic filtering by region, vehicle type, and other parameters to gain valuable insights into EV adoption trends. This flexibility allows users to customize their analysis based on specific requirements, fostering a deeper understanding of electric vehicle adoption patterns and informing decisions related to infrastructure planning, policy-making, and market strategies. Enhanced data validation and error-handling mechanisms contribute to a seamless user experience, ensuring the analysis is robust and accurate, even when dealing with missing or incomplete data entries. This ensures that users can confidently explore the dataset, and the tool is adaptable for future expansion, including the integration of new data or additional parameters for more detailed analysis. Looking ahead, the project sets the stage for future enhancements, such as expanding the dataset to include more years, states, and regions, as well as incorporating additional metrics like charging infrastructure or consumer demographics. Further improvements in scalability will allow the model to handle larger datasets efficiently, while the integration of a graphical user interface (GUI) could make the tool more accessible to non-technical users, allowing for broader exploration of electric vehicle trends. In summary, this project illustrates the power of data visualization in understanding complex adoption patterns in the electric vehicle industry. By providing practical tools for policymakers, researchers, and industry stakeholders, the project supports data-driven decision-making and offers valuable insights into the future of clean transportation. Its success highlights the potential for continued innovation in EV data analysis and visualization, paving the way for improved strategies and policies to promote sustainable transportation solutions.

References

iMatplotlib: https://matplotlib.org/stable/index.html

ii. Pandas: https://pandas.pydata.org/docs/

iii. NumPy: https://numpy.org/

iv. Seaborn: https://seaborn.pydata.org/

v. Kaggle