



INNOVATION. AUTOMATION. ANALYTICS

PROJECT ON

Exploratory Data Analysis (Electric vehicles Dataset)

Submitted By:
ARUNRAJ DAARA



About me

- Hello! My name is Arunraj Daara, and I've completed a B.Tech in Computer Science Engineering.
- I want to learn Data Science because it combines my interest in technology and problem-solving. Data Science allows me to analyze data to uncover insights and make informed decisions. I'm excited about the potential to work on real-world problems and contribute to innovative solutions using data!
- Completed a Datascience Certificaton from Innomatics Research Labs
- Feel free to reach out! You can do so by following the links below:
- www.linkedin.com/in/arunrajdaara
- <https://github.com/ARUNRAJDAARA>

Objective of the Project

- 1. Analyze Trends:** Identify trends in electric vehicle (EV) registrations over time, focusing on model years and makes.
- 2. Examine Adoption Patterns:** Explore how different regions (states and counties) are adopting electric vehicles and the factors influencing this adoption.
- 3. Assess Vehicle Types:** Analyze the distribution and characteristics of Battery Electric Vehicles (BEVs) and Plug-in Hybrid Electric Vehicles (PHEVs) within the dataset.
- 4. Investigate Pricing and Range:** Evaluate the relationship between the base MSRP of electric vehicles and their electric range to understand pricing dynamics.
- 5. Visualize Data:** Create visual representations (like bar charts and choropleth maps) to effectively communicate findings and trends in the electric vehicle market.
- 6. Inform Stakeholders:** Provide insights that can help policymakers, manufacturers, and consumers make informed decisions regarding electric vehicles.
- 7. Support Sustainable Transportation:** Contribute to the understanding of electric vehicles as a sustainable transportation option and identify areas for future growth and development.

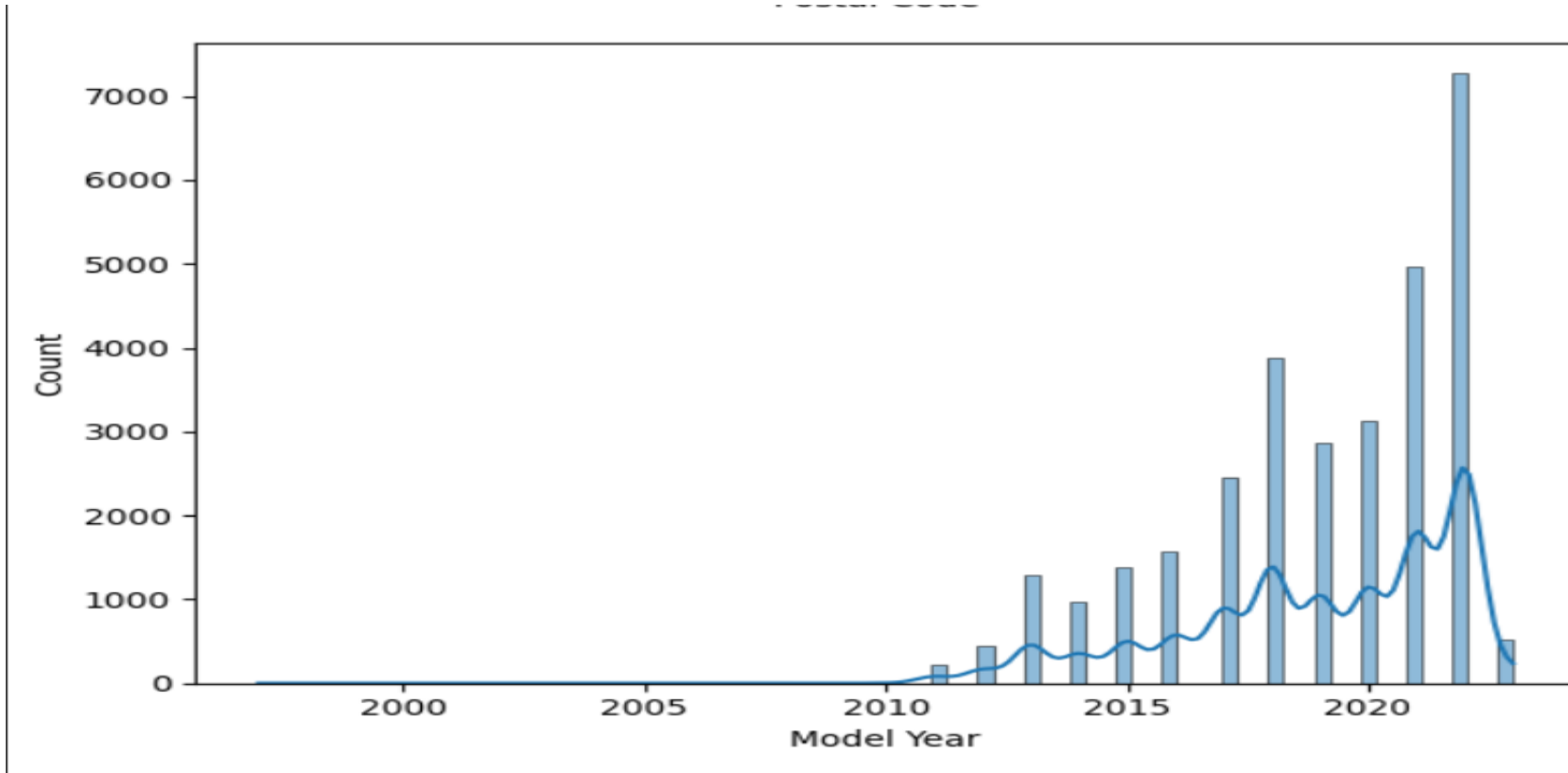
Summary of the Dataset

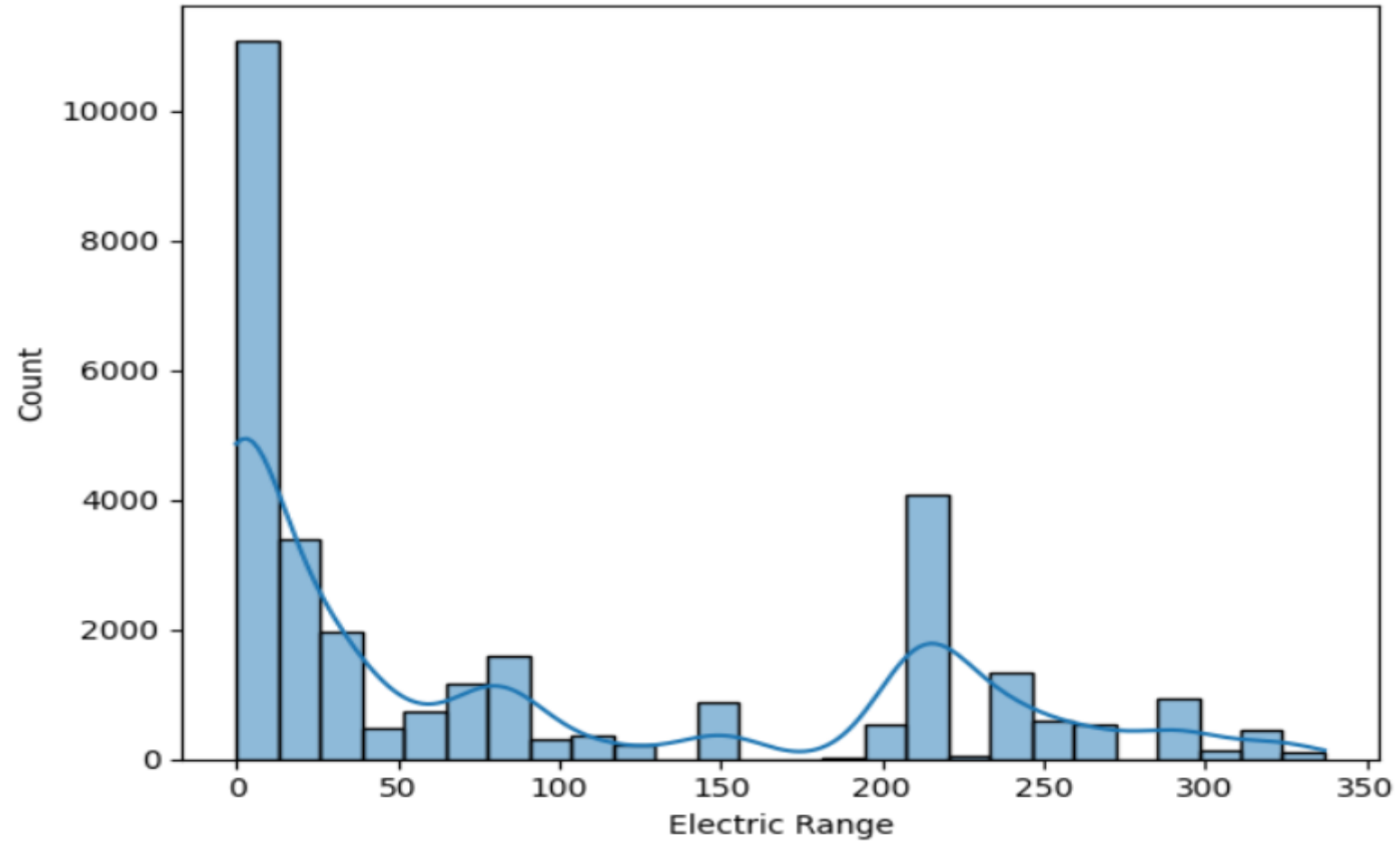
- The electric vehicle dataset comprises “30,957 data points”, offering a detailed view of electric vehicle registrations across various parameters. It includes “10 categorical columns”, such as VIN (1-10), County, City, State, Postal Code, Model Year, Make, Model, Electric Vehicle Type, and Clean Alternative Fuel Vehicle (CAFV) Eligibility.
- These columns provide essential information about the vehicle's identity, geographic registration, and classification. Additionally, the dataset contains “7 numerical columns”, including Electric Range, Base MSRP, Legislative District, DOL Vehicle ID, Vehicle Location, Electric Utility, and 2020 Census Tract, which quantify critical aspects of the vehicles, such as their price, range, and utility provider.
- This comprehensive dataset serves as a valuable resource for analyzing trends and adoption patterns in the electric vehicle market, facilitating insights that can guide policymakers, manufacturers, and consumers in their efforts to promote sustainable transportation solutions.

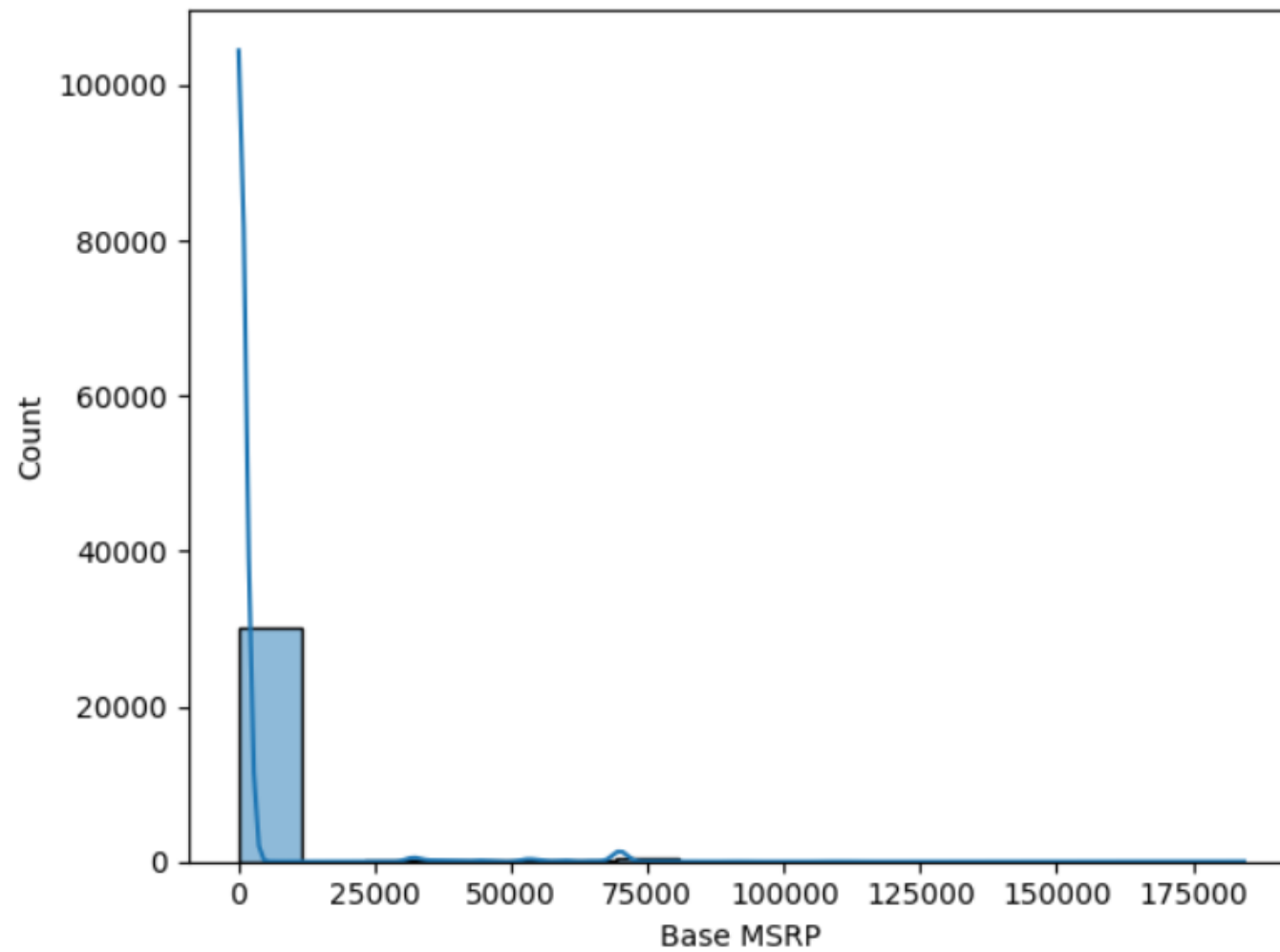
Comprehensive Data preparation

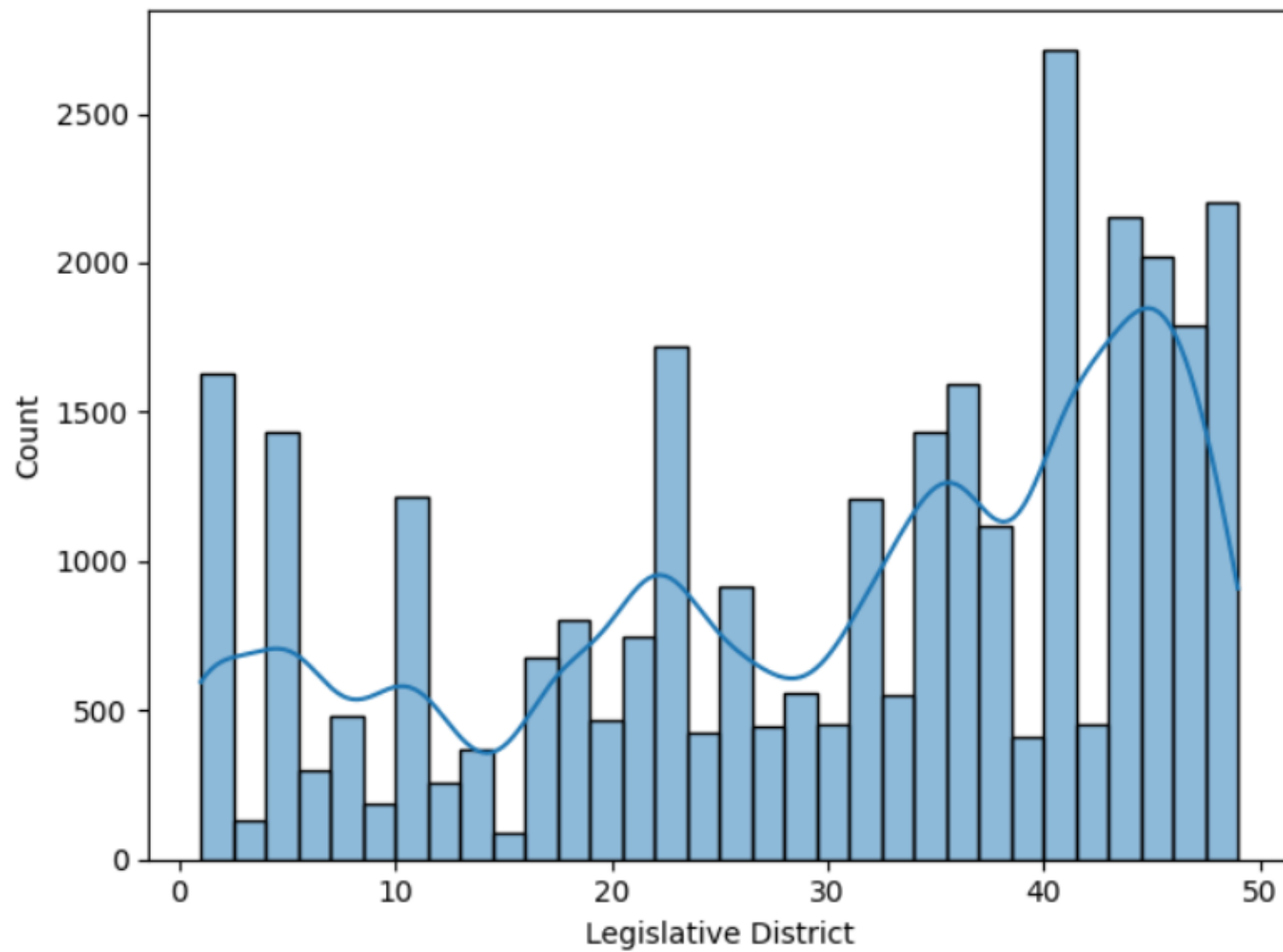
- **Examine Missing Values:** Let's dive into the dataset to identify any missing values that may affect our analysis. Understanding where data is lacking is crucial for ensuring the integrity of our insights.
- **Identify Duplicated Values:** We'll check for any duplicate entries that could skew our results. By ensuring each data point is unique, we can maintain the accuracy of our analysis.
- **Explore Different Columns:** We'll take a closer look at the various columns in the dataset to understand their significance and how they relate to our salary analysis.
- **Check Data Types:** It's essential to verify the data types of each column to ensure they are appropriate for our analysis. This helps us apply the right statistical methods.
- **Analyze Unique Values:** Let's investigate the unique values in each column to uncover any interesting patterns or anomalies that could inform our understanding.
- **Review Statistics:** We'll examine the descriptive statistics of the dataset, providing us with a summary of key metrics that can highlight trends and insights.
- **Categorical Analysis:** We'll explore the various categories present in the different categorical columns to understand their distribution and impact on salary levels.
- **Drop Unnecessary Columns:** Finally, we'll clean the dataset by removing any unnecessary columns that do not contribute to our analysis, ensuring that we focus on the most relevant information.

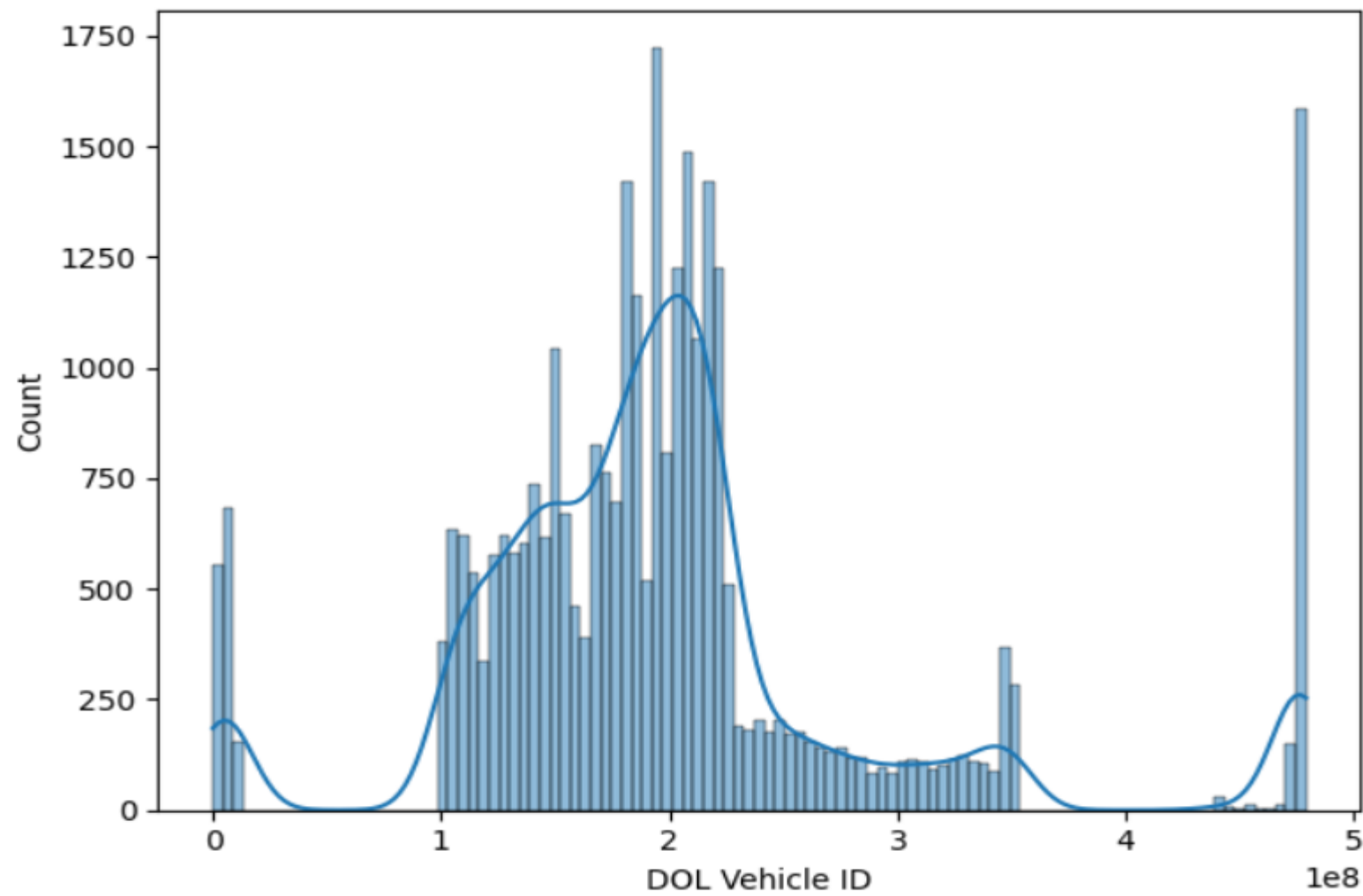
Univariate Analysis

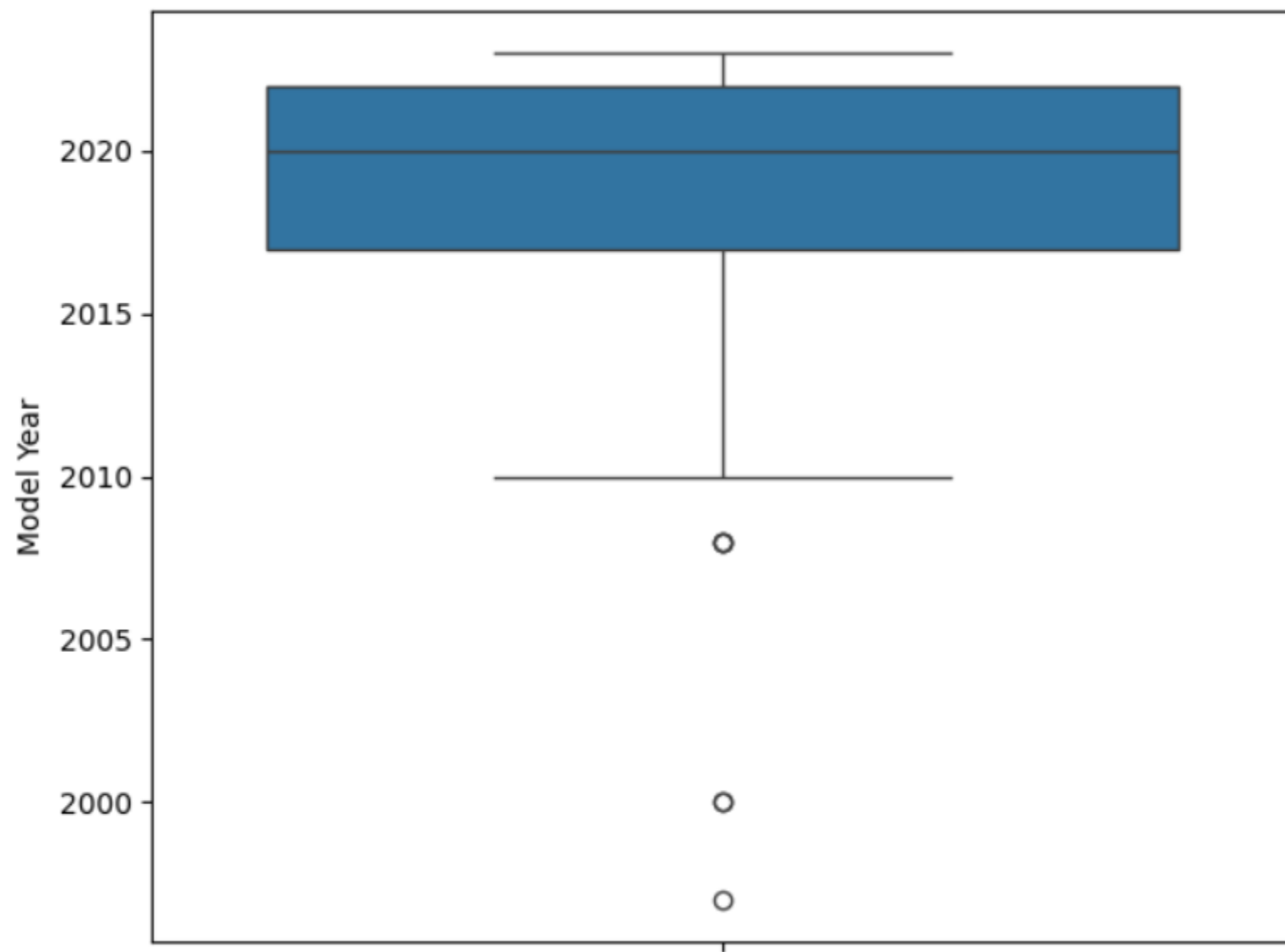




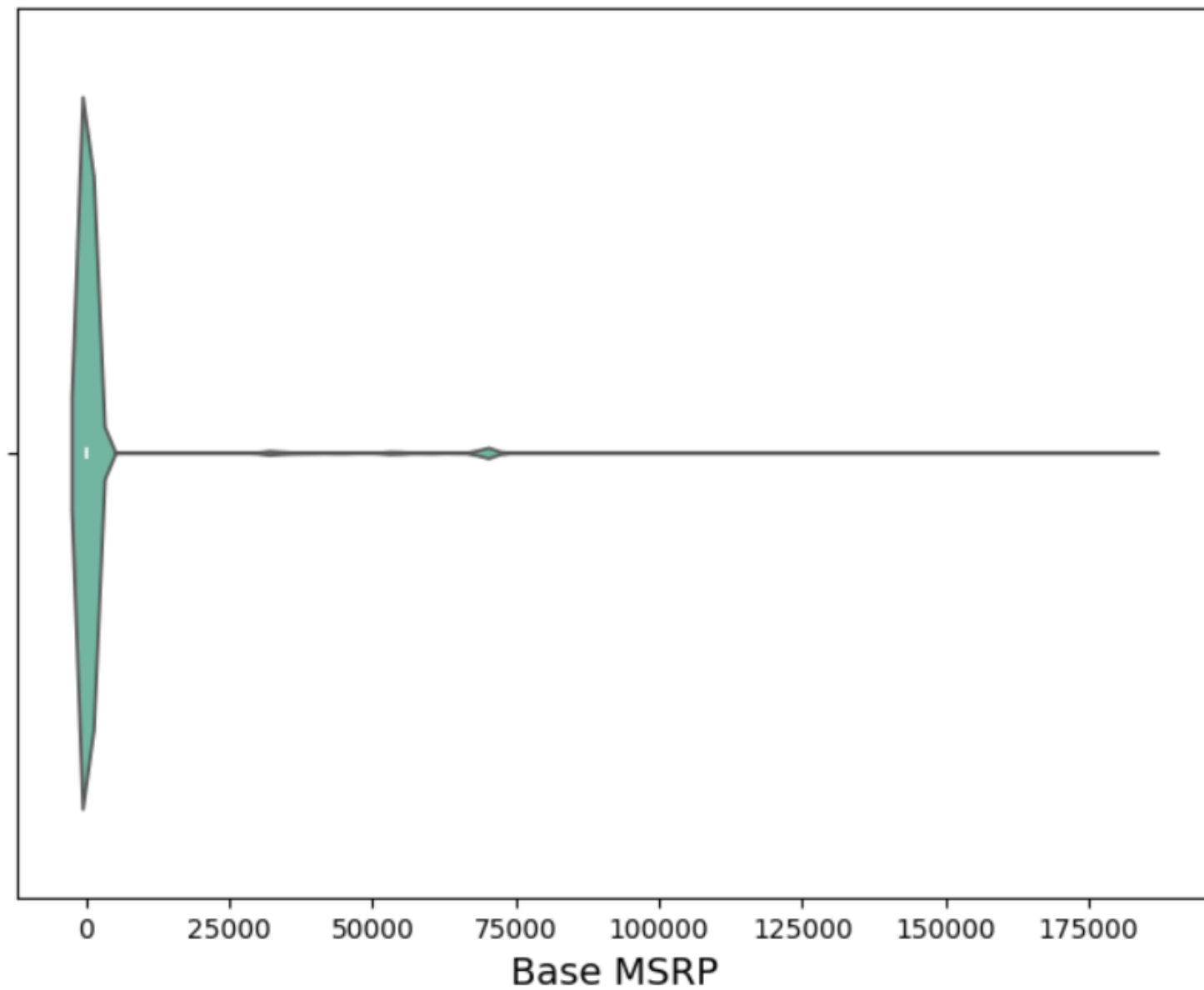


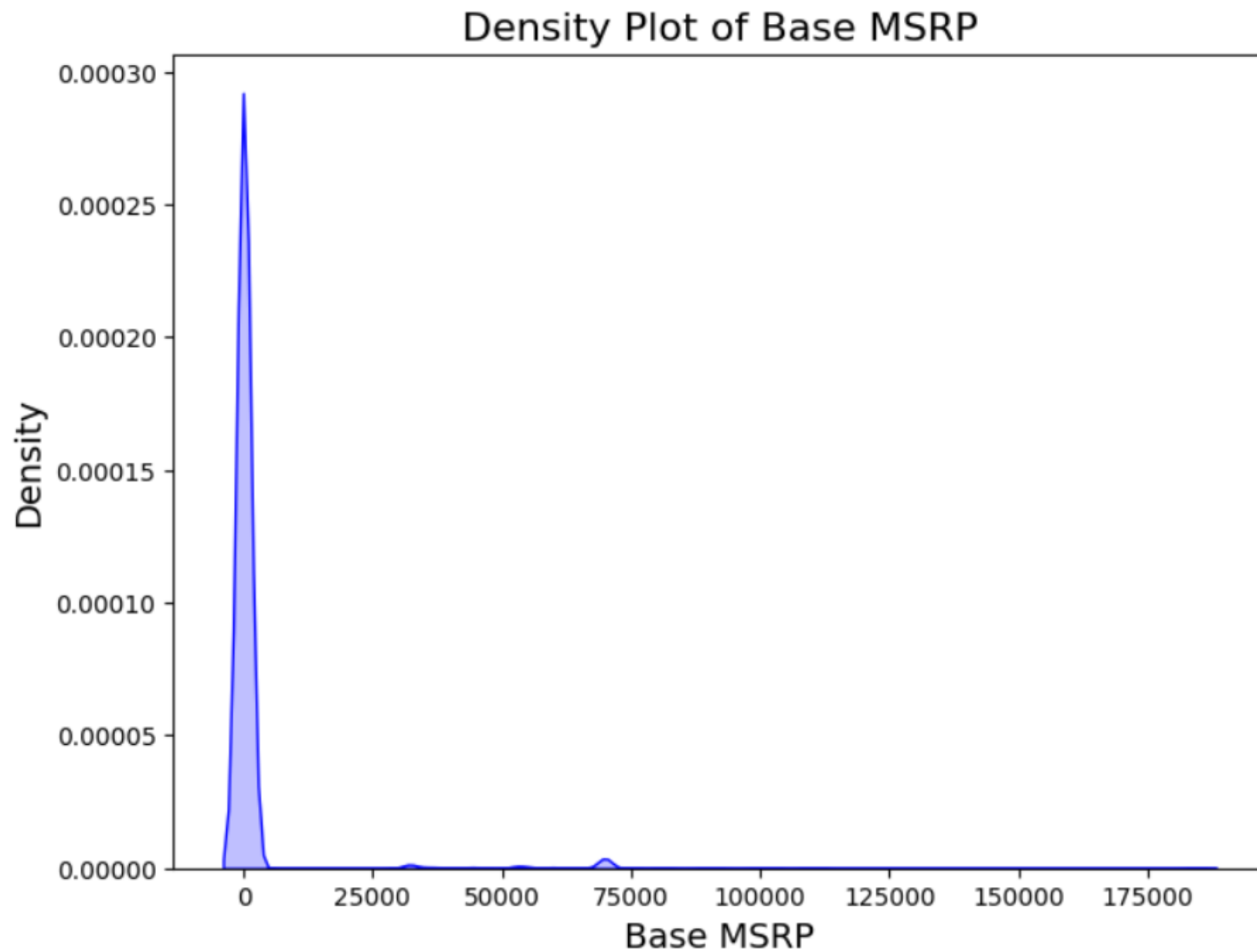




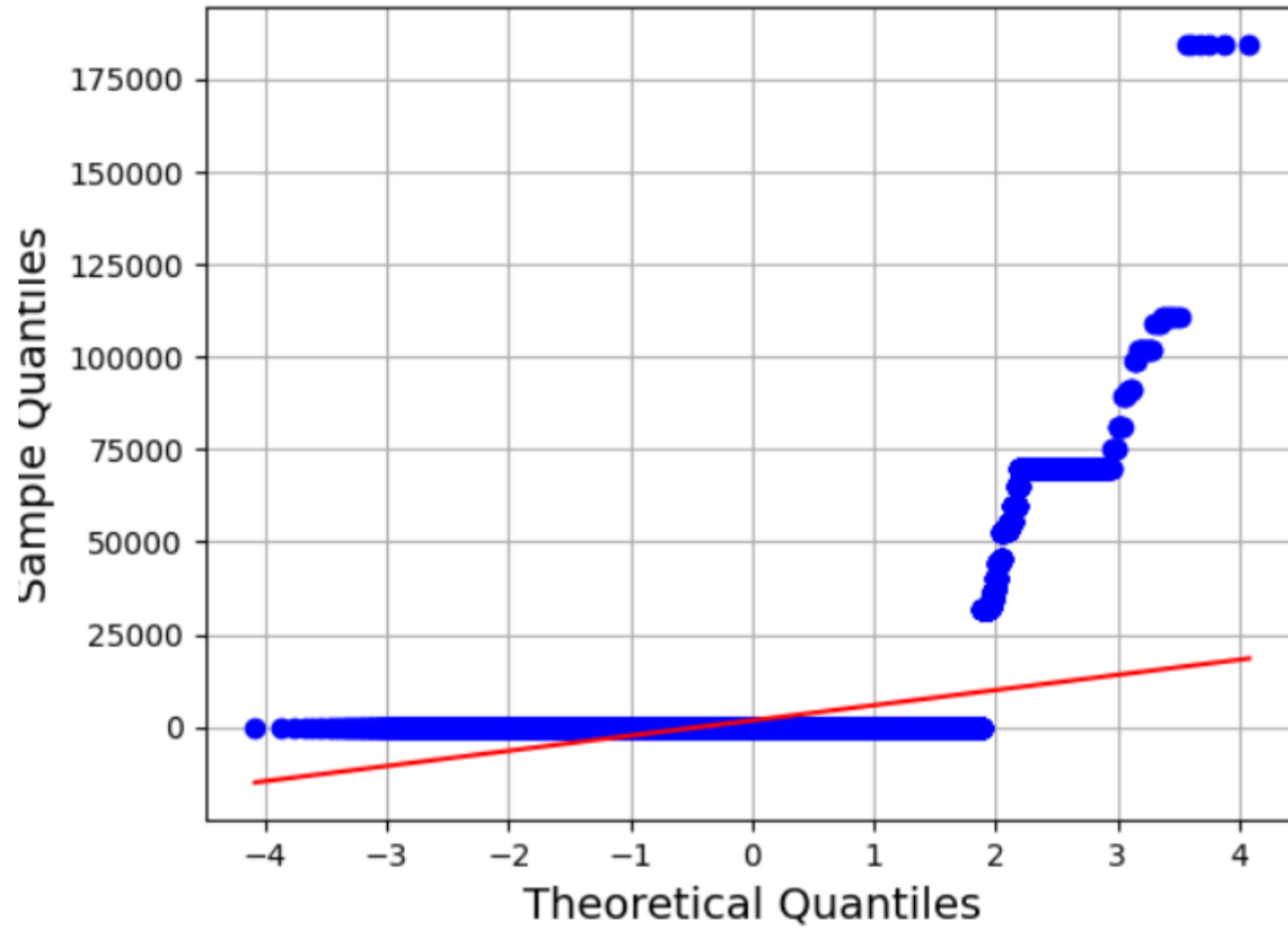


Violin Plot of Base MSRP





Q-Q Plot of Base MSRP

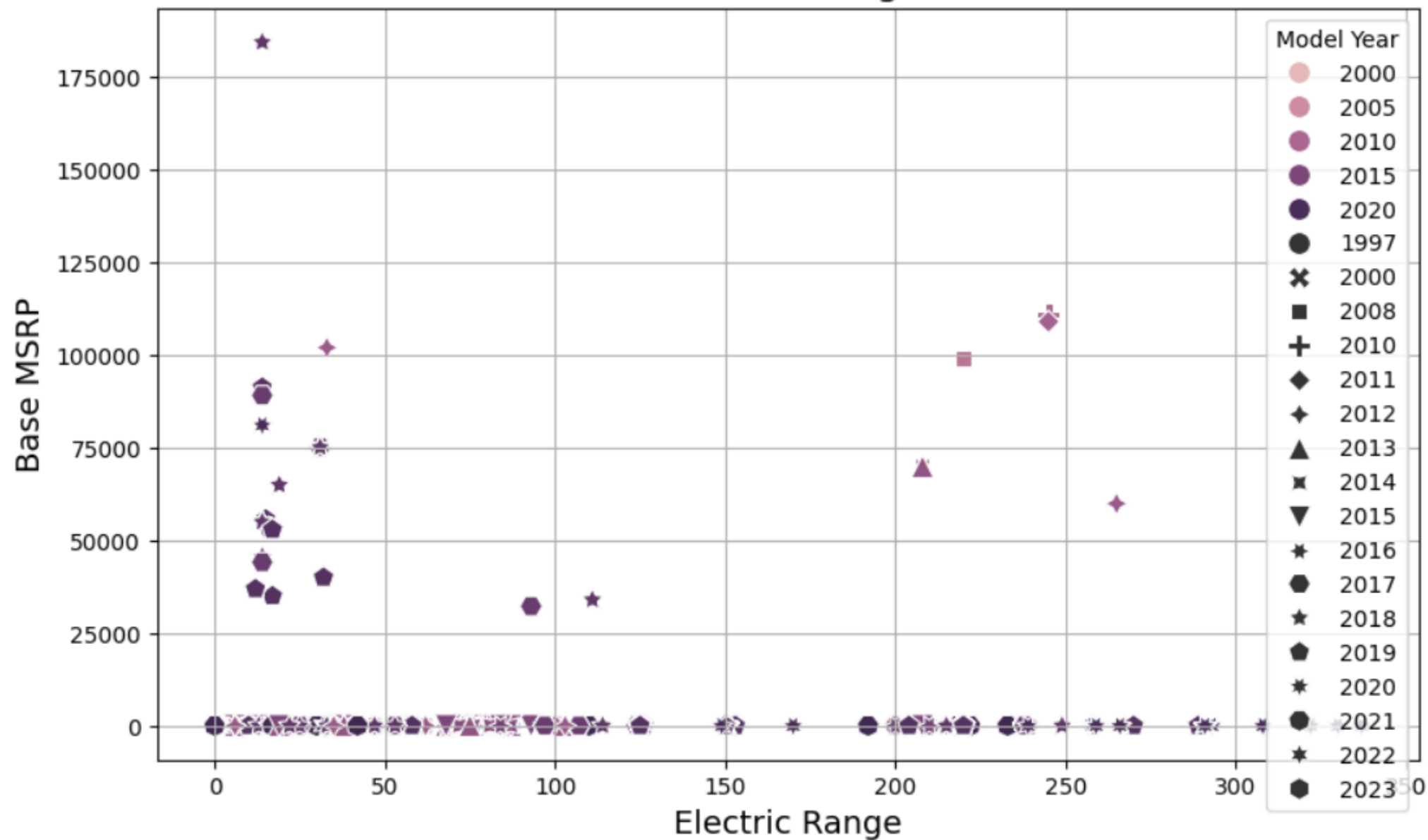


Univariate analysis involves examining a single variable to summarize and understand its distribution and characteristics. In your electric vehicle dataset, you can perform univariate analysis on various columns, such as:

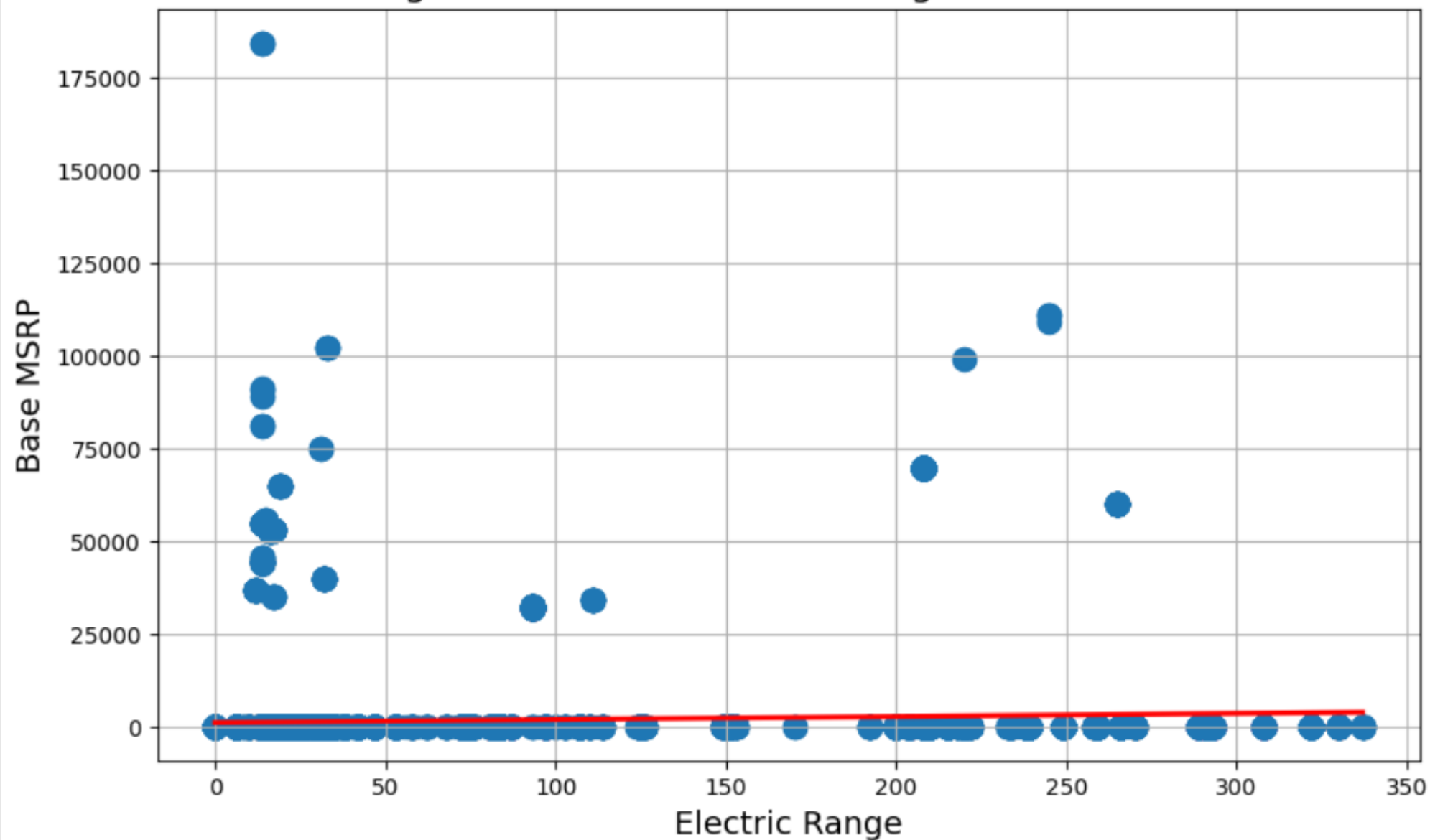
- **Model Year:** Analyze the distribution of vehicle registrations over different years. You can create histograms or bar charts to show how the number of registered electric vehicles changes annually.
- **Make:** Count the frequency of different manufacturers. A bar chart can illustrate which brands dominate the market.
- **Electric Range:** Assess the range of electric vehicles by plotting a histogram to understand the typical range offered by different models.
- **Base MSRP:** Analyze the price distribution of electric vehicles to determine pricing trends and affordability.

This type of analysis helps identify patterns, trends, and potential outliers within individual variables.

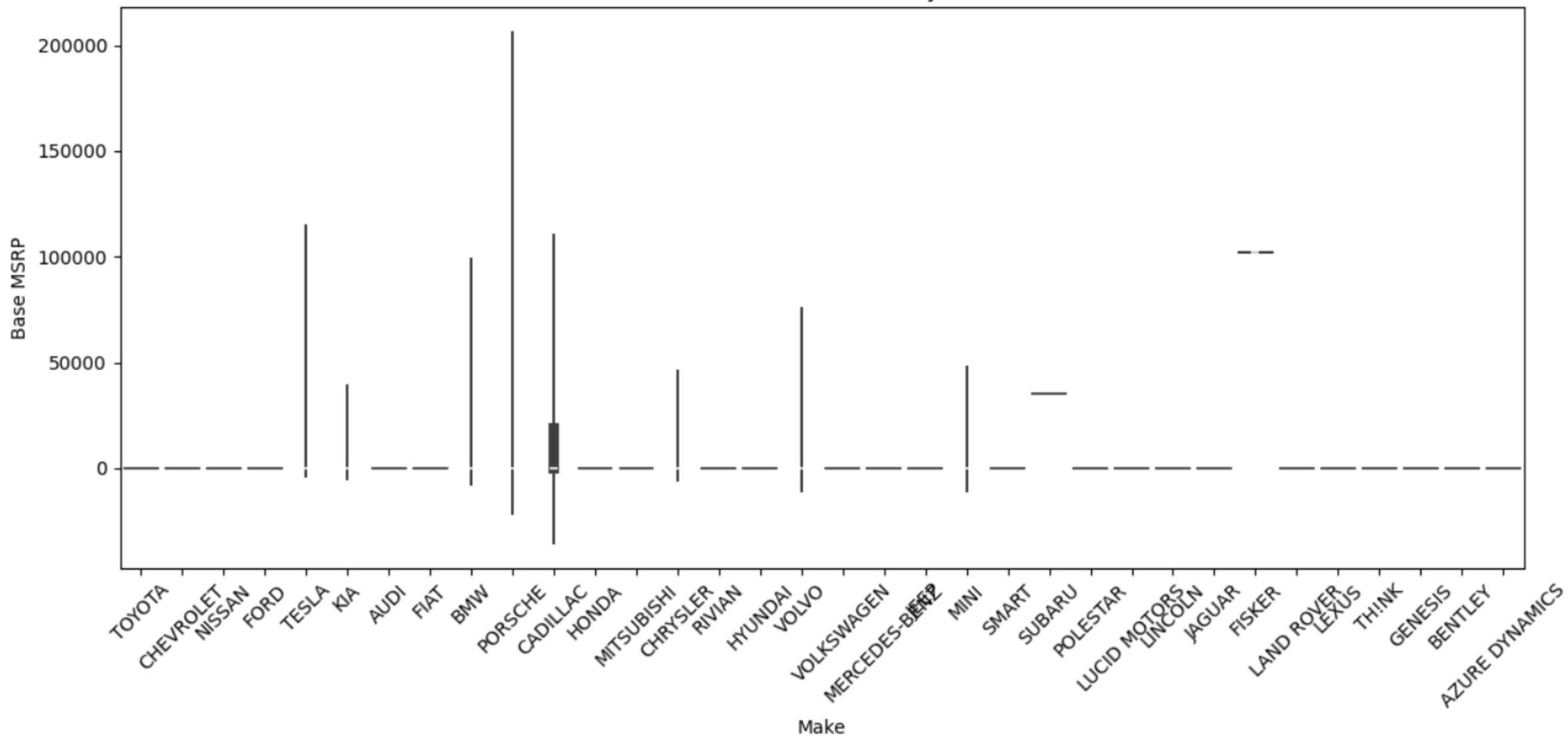
Scatter Plot of Electric Range vs Base MSRP

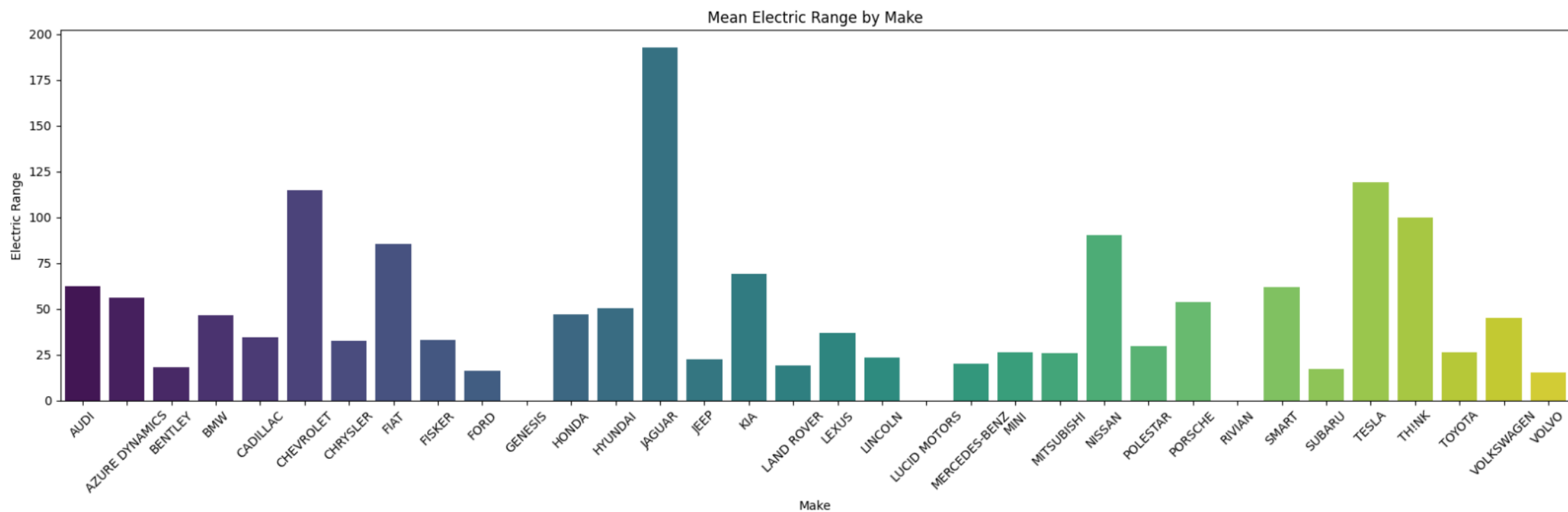


Regression Plot of Electric Range vs Base MSRP



Violin Plot of Base MSRP by Make

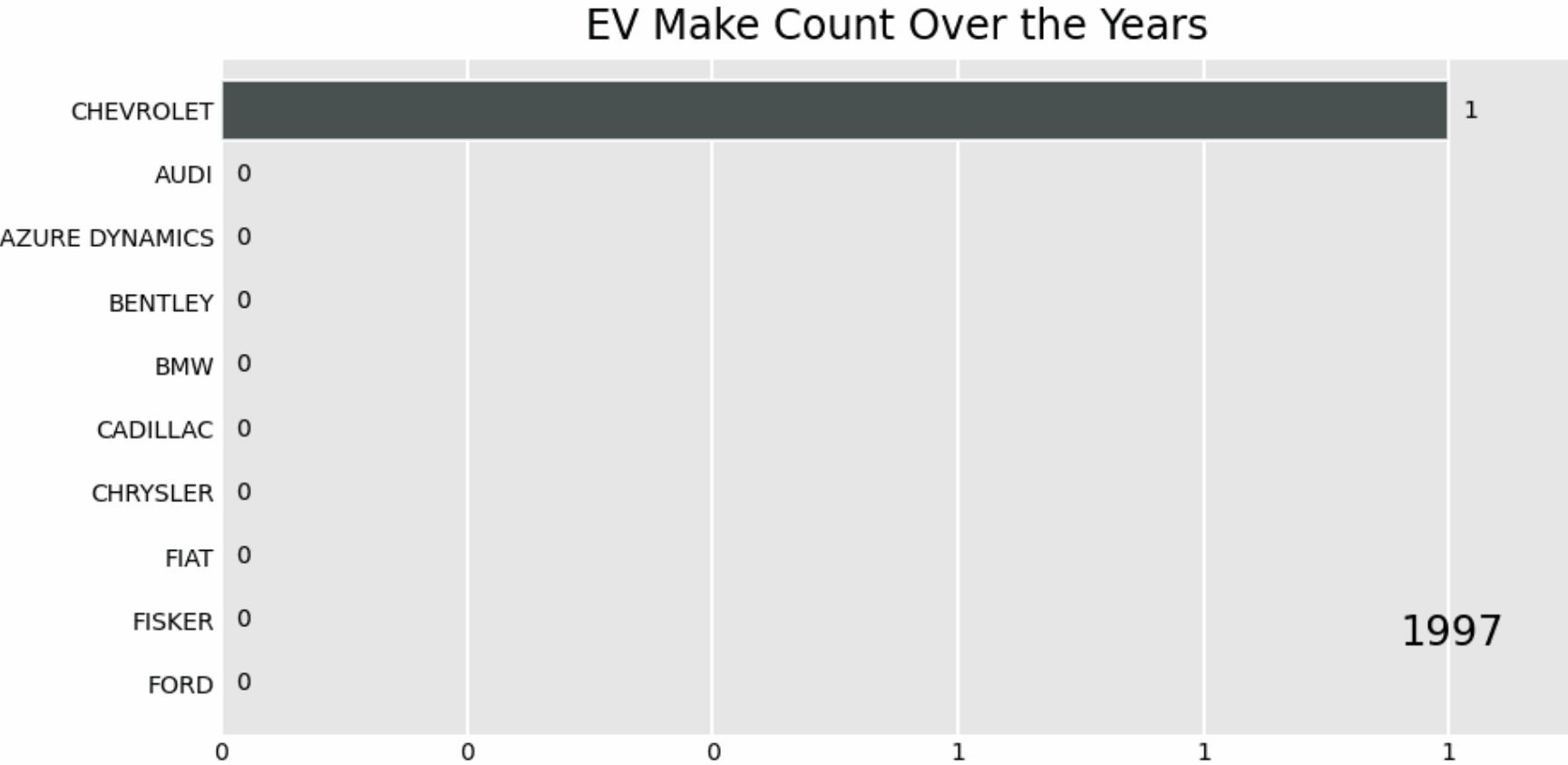




Task-2:

- This code effectively visualizes the distribution and evolution of vehicle registrations across the United States by state and model year. The resulting choropleth map allows viewers to easily identify trends in vehicle counts, highlighting areas of growth and decline over time.
- The animation feature provides an engaging way to observe changes year by year, making it clear how different states are contributing to the overall vehicle landscape. This visualization serves as a powerful tool for understanding regional differences in vehicle adoption and can inform stakeholders about potential markets for electric vehicles, policy implications, and infrastructure planning to support sustainable transportation initiatives.

Task-3: Create a Racing Bar Plot to display the animation of EV Make and its count each year.



- This analysis leveraged a bar chart race to visually represent the trends in Electric Vehicle (EV) registrations by make over the years. By transforming the dataset, we first converted the 'Model Year' into a string format and grouped the data by year and make to count the number of registrations for each vehicle type. The resulting pivot table allowed us to create a dynamic visualization that highlights the evolution of EV adoption across different manufacturers.
- The bar chart race effectively showcases the increasing competition among EV manufacturers and the shifting landscape of the market. It reveals how various brands have gained or lost prominence over the years, providing valuable insights into consumer preferences and industry trends. This visual representation not only enhances the understanding of EV adoption patterns but also serves as a tool for stakeholders to assess the growth of electric mobility and make informed decisions moving forward.
- Ultimately, this analysis emphasizes the importance of data visualization in comprehending complex datasets, facilitating a clearer interpretation of the transition towards sustainable transportation.

Use Cases:

- **Market Research:** Automakers can use the analysis to identify trends in consumer preferences for specific makes and models, helping them tailor their offerings and marketing strategies.
- **Policy Development:** Policymakers can leverage insights into electric vehicle adoption patterns by state to craft incentives and infrastructure investments that promote sustainable transportation.
- **Consumer Insights:** Potential buyers can benefit from understanding which electric vehicles offer the best range and pricing, aiding their decision-making process.
- **Utility Planning:** Electric utility companies can analyze the distribution of electric vehicles by region to better plan for energy demand and charging infrastructure.

By applying these findings, stakeholders can foster a more informed approach to the growing electric vehicle market, ultimately supporting the transition to sustainable transportation solutions.

Conclusion

The analysis of the electric vehicle dataset reveals significant trends and insights into the adoption of electric vehicles. The data shows a clear increase in registrations over the years, reflecting a growing interest in sustainable transportation. Key factors such as vehicle make, model year, and pricing contribute to understanding consumer preferences and market dynamics. The findings highlight the importance of supporting policies and infrastructure to further promote electric vehicle adoption. Overall, this analysis underscores the potential for electric vehicles to play a crucial role in the transition to cleaner, more sustainable transportation options.

THANK
YOU

