**Title: Electric Vehicle (EV) Market Segmentation and Analysis**

**Introduction**

**Project Objective:**

The goal of this project is to perform an in-depth analysis of electric vehicles (EVs) using available data. We will analyse various features like price, battery capacity, range, and manufacturer to gain insights into the EV market. The project also aims to segment the market based on vehicle features to identify distinct groups of vehicles. These insights will help understand different market segments, including budget EVs, premium EVs, and mid-range EVs.

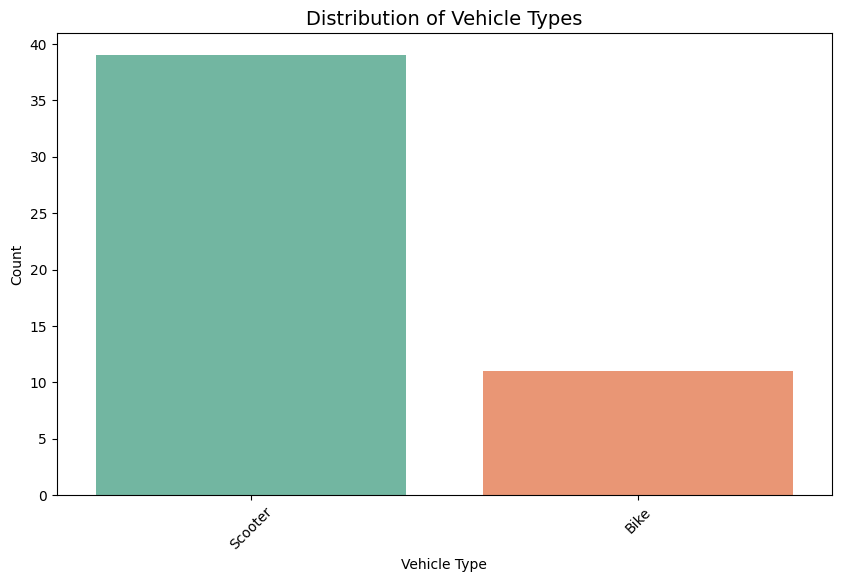
**Data Overview:**

The dataset includes columns such as Vehicle Type, Manufacturer, Battery Capacity (kWh), Range per Charge (km), Charging Time, Price, Power (HP or kW), and Top Speed (km/h).

**Data Analysis and Visualizations:**

**Distribution of Vehicle Types:**

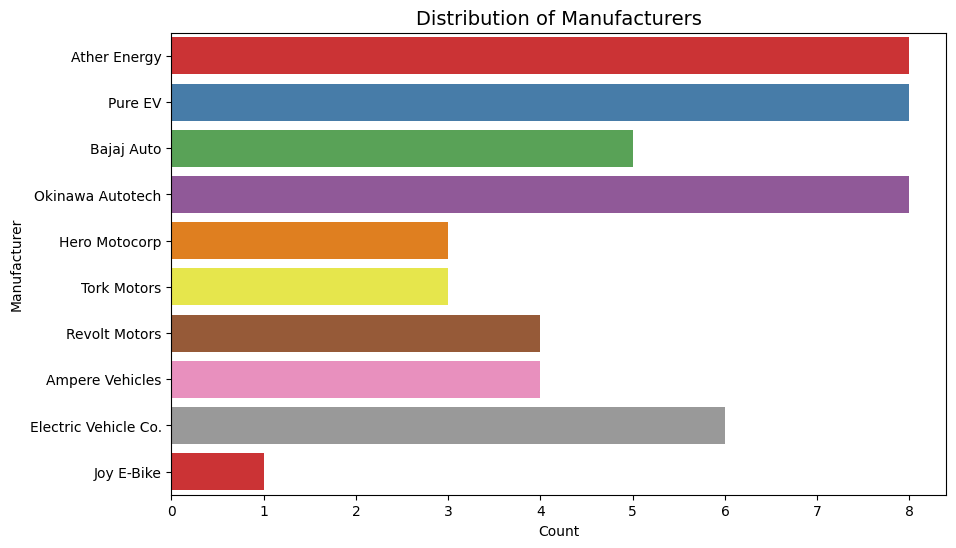
A count plot was used to visualize the distribution of various vehicle types in the dataset.



Conclusion: The majority of the vehicles in this dataset are scooters, followed by bikes. This reflects a trend toward electric two-wheelers in the market.

**Price Distribution by Manufacturer:**

This box plot shows the variation in prices across different manufacturers in the dataset.

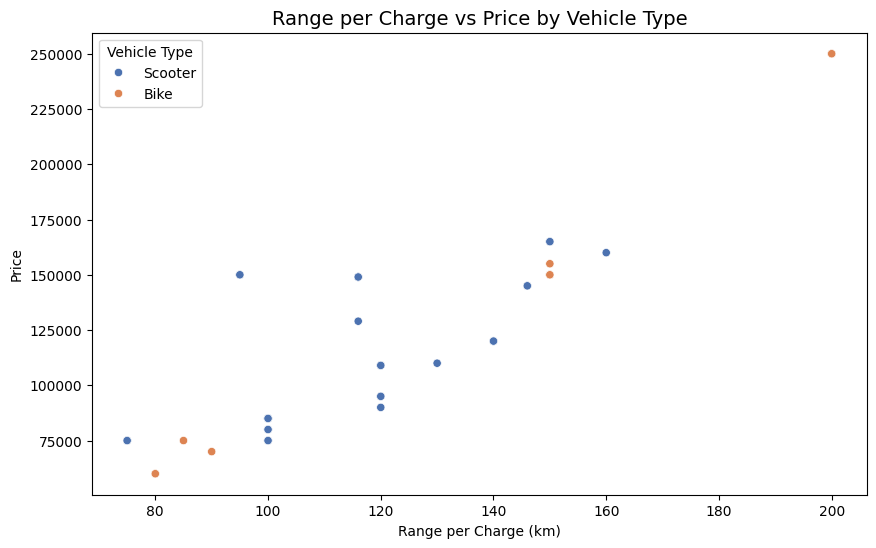


Explanation:

Certain manufacturers like Bajaj and Hero Electric seem to target the budget-conscious segment, while others focus on more premium offerings.

**Range per Charge vs Price:**

A scatter plot was used to analyze the relationship between the range a vehicle can travel on a single charge and its price. The plot is colored by Vehicle Type.



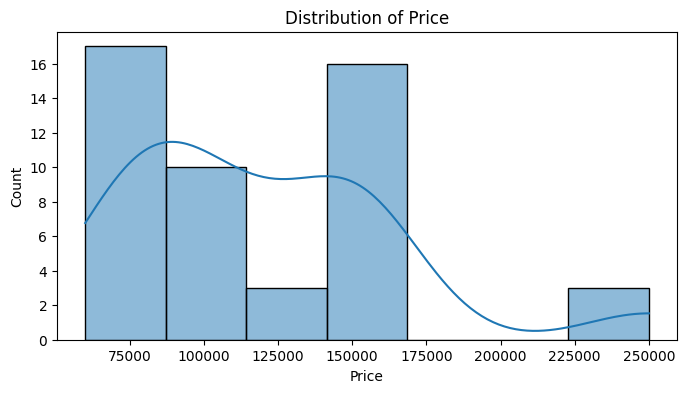
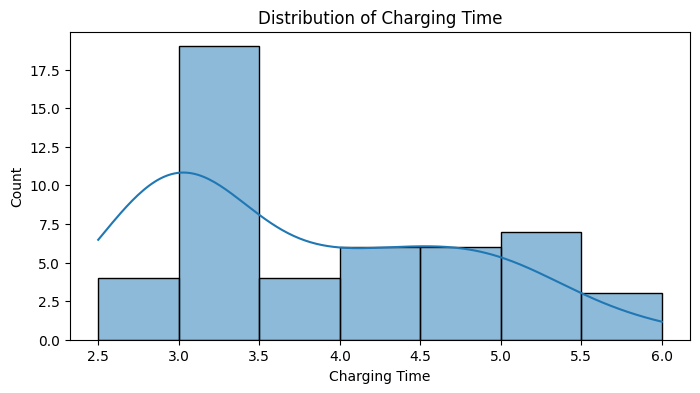
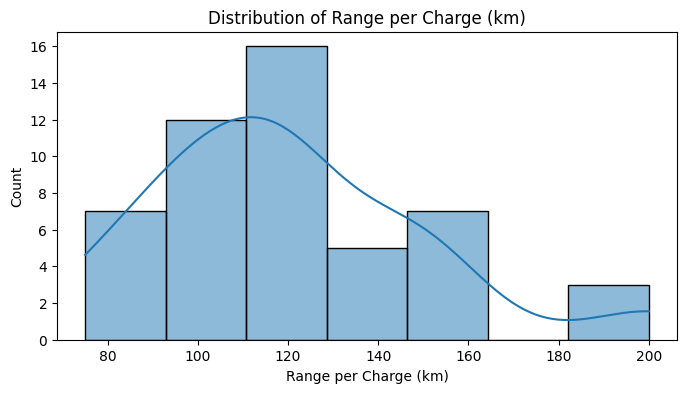
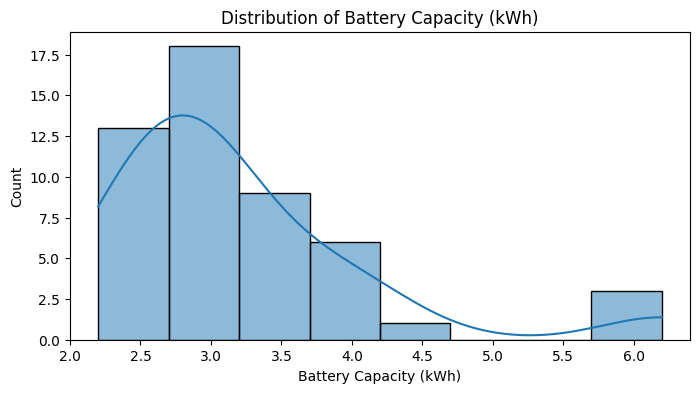
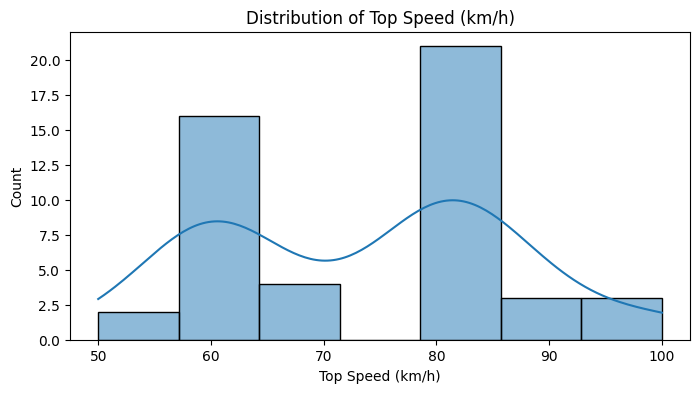
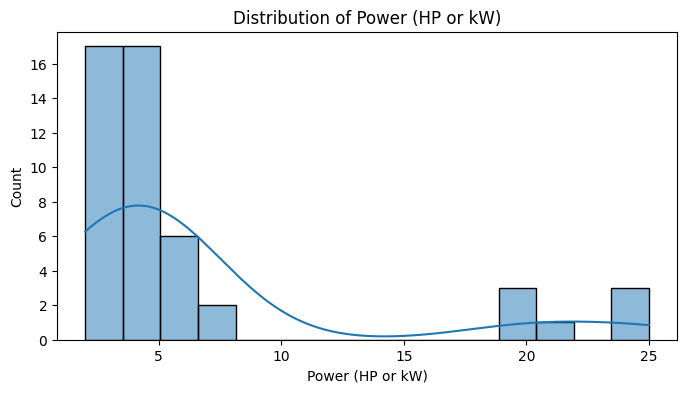
Explanation:

As expected, vehicles with higher range per charge also tend to be more expensive.

Most bikes offer a higher range than scooters but at a premium price.

**Pair plot of Key Features:**

A pair plot was generated to explore the relationships between numeric features such as Battery Capacity, Price, Range per Charge, etc.



Explanation:

A positive correlation can be observed between Battery Capacity and Range per Charge.

Price also correlates with the battery capacity, showing that larger battery sizes are generally associated with more expensive vehicles.

**Market Segmentation**

Using K-Means Clustering, the dataset was segmented into distinct groups. Features like Battery Capacity, Range per Charge, Price, and Power were used to cluster the vehicles into different market segments.

**Segments Identified:**

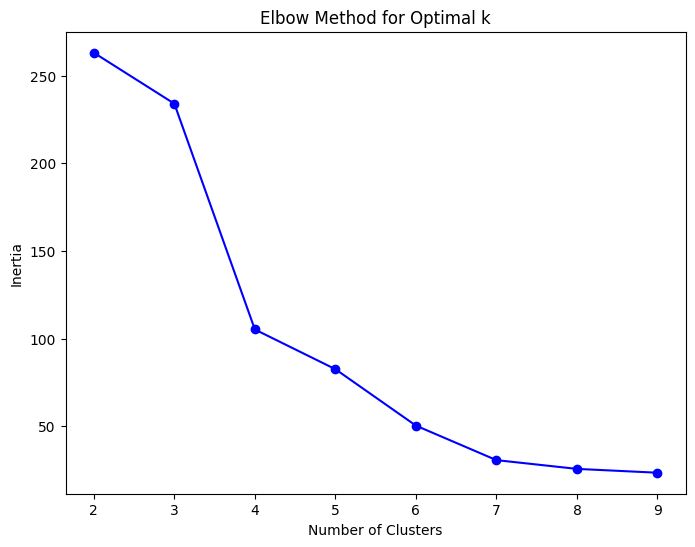
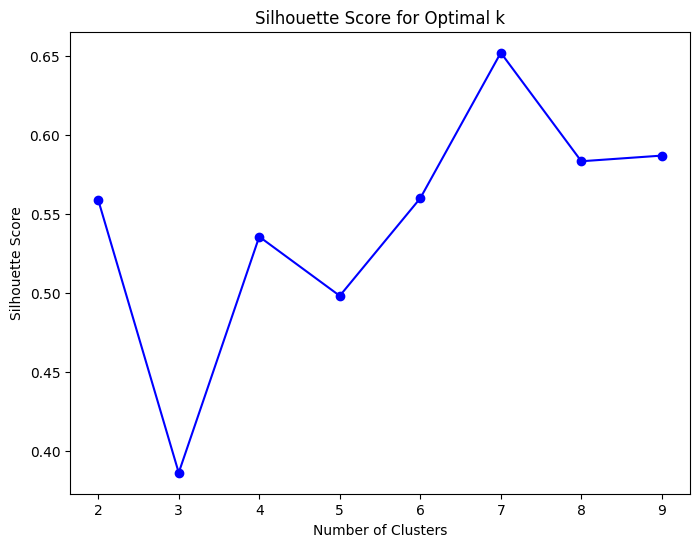
Budget EVs: Vehicles that offer a shorter range and smaller battery capacity but are priced affordably.

Mid-Range EVs: These vehicles provide a balance between price and performance, offering moderate range and power.

Premium EVs: High-performance vehicles with larger battery capacities and higher prices.

Visualizing the Clusters: After applying PCA for dimensionality reduction, the clusters were

visualized

. 

Explanation:

Each cluster represents a different group of vehicles based on price, range, and performance characteristics.

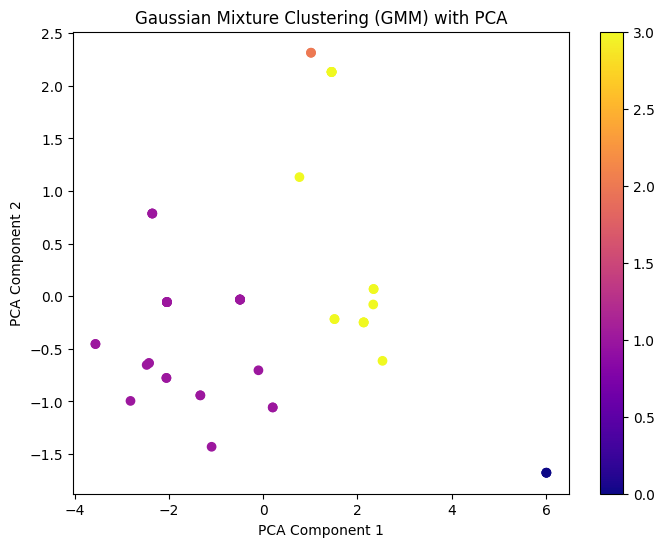
The Premium EVs cluster has vehicles with the highest battery capacity and range, while Budget EVs cluster has more affordable, lower-range vehicles.

**DBSCAN Clusters:**

Some points may be labeled as noise, forming no cluster.

The clusters may have irregular shapes that adapt to the actual distribution of data points.

**GMM Clusters:**

****

Points may belong to multiple clusters with different probabilities.

The clusters can take an elliptical shape, better capturing the complexity of the data.

**Pros : Comparison of Clustering Results**

**K-Means Clustering**: Assumption: Clusters are spherical and equally sized.

Performance: The algorithm may work well when the data is well-separated, but it might not capture clusters with irregular shapes or varying densities.

Result: K-Means segments the EV market into groups based on the selected features (e.g., price, range, battery capacity).

**DBSCAN:**

**Strengths**: DBSCAN is great for discovering clusters with irregular shapes. It doesn’t require specifying the number of clusters beforehand, which is useful if you don’t know how many clusters exist in the data.

It also detects noise, i.e., points that do not belong to any cluster, which can help identify outliers in the EV market.

**Weaknesses**: It may struggle with datasets with highly varying densities.

If the parameters eps and min\_samples are not chosen carefully, it may either split meaningful clusters or group unrelated data points together.

Example Insight:

DBSCAN might reveal small niche segments in the EV market, such as low-production, high-performance vehicles that are priced higher but don’t fit into the typical mass-market categories.

**Gaussian Mixture Models (GMM):**

**Strengths:** GMM allows for more flexibility in the shape and size of clusters compared to K-Means. It can model elliptical clusters, which is useful if the data isn’t perfectly spherical.

GMM also provides a probability distribution for each data point, allowing a point to belong to multiple clusters with different probabilities. This is useful if some vehicles could potentially fall into more than one segment.

**Weaknesses:** It may be computationally more expensive than K-Means.

GMM can struggle if the data has significant noise or if the Gaussian assumption doesn’t hold.

Result from GMM:

Soft clustering: Each vehicle has a probability of belonging to different segments. For instance, a vehicle might have a 70% chance of being in the premium segment and a 30% chance of being in the mid-range segment.

Elliptical clusters: GMM captures more complex relationships in the data, such as cases where price and range form elliptical distributions.

Example Insight:

GMM might reveal overlapping market segments. For example, some vehicles may be priced like budget vehicles but offer the range or performance closer to mid-range vehicles, resulting in soft clustering.

**Conclusion**

Market Segments: The electric vehicle market can be divided into three main segments based on price, range, and power:

Budget EVs: Affordable and popular among cost-conscious buyers.

Mid-Range EVs: Striking a balance between performance and affordability.

Premium EVs: Offering high performance at premium prices.

**Insights:**

Scooters dominate the market, especially in the budget segment, indicating a shift toward affordable electric two-wheelers.

Bikes generally offer better range and performance but at a higher cost.

Practical Use:

Manufacturers can target each segment differently: budget models for mass markets and premium models for niche buyers who prioritize performance.

**Next Steps and Recommendations**

Further Analysis: With more detailed data such as consumer preferences, charging infrastructure, and environmental factors, deeper segmentation could be achieved.

Data Collection: Gathering more data on sales trends, consumer demographics, and real-world performance could further improve market analysis.

Thank You

Pratik Bhoud