

Market Segmentation of Electric Vehicle

Tushar Bhakat

06.09.2024

Data collection: <https://www.kaggle.com/datasets/geoffnel/evs-one-electric-vehicle-dataset>

1. ML Model for Segmentation

In this project, we leveraged **K-Means Clustering** to perform the market segmentation analysis. K-Means is a widely-used machine learning algorithm that excels at partitioning a dataset into distinct groups based on shared characteristics. The decision to use K-Means was based on its efficiency in handling larger datasets and its interpretability, which made it ideal for our analysis.

- **Why K-Means?:** We chose this algorithm because it's simple yet effective for clustering tasks, and it provides clear boundaries between segments. It's also fast, making it a good fit for our dataset size.
- **Approach:**
 1. **Data Preprocessing:** We began by standardizing the data using StandardScaler. This step was crucial to ensure that features such as price, range, and speed were on the same scale, preventing any single feature from dominating the clustering process.
 2. **Feature Selection:** We focused on key numerical attributes like price, top speed, acceleration, and range, as these metrics directly influence consumer preferences in the electric vehicle market.
 3. **Clustering:** After experimenting with different values of 'K', we used the **Elbow Method** to determine the optimal number of clusters. This method helped us visualize the trade-off between the number of clusters and the variance within each cluster.
 4. **Dimensionality Reduction (PCA):** Once the clusters were formed, we applied Principal Component Analysis (PCA) to reduce the data to two dimensions, allowing us to better visualize the segments and understand their separation.

2. Conclusion & Insights

- a) **Variety in Car Brands:** The dataset contains a wide range of car brands, with some brands standing out in terms of model frequency. Brands like Tesla and BMW are particularly dominant, reflecting their significant footprint in the electric vehicle (EV) market. This diversity in brand presence highlights the competition and innovation driving the EV industry.
- b) **Top Speed Performance:** When looking at top speeds, brands like Tesla and Porsche clearly stand out. These brands cater to consumers who prioritize performance, especially in the luxury and sports car segments. Their focus on delivering high-speed EVs is evident, positioning them as leaders for customers seeking speed and performance.
- c) **Range of EVs:** The distance a car can travel on a single charge varies significantly across brands. Tesla, in particular, offers models with some of the longest ranges, making them attractive to buyers who prioritize practical and long-distance driving. This variation in range plays a major role in how cars are perceived, with longer-range vehicles generally more appealing to a wider audience.
- d) **Charging Plug Types:** From the pie chart on charging plugs, it's clear there's some level of standardization within the EV market. Popular plug types dominate, suggesting that manufacturers are working toward compatibility across charging stations. This is a crucial factor for customer satisfaction and convenience.
- e) **Body Styles:** SUVs and sedans are the most common body styles, aligning with broader automotive trends. These body types are popular for their versatility and comfort, appealing to a broad demographic of consumers. The data reflects that the EV market is following similar trends as traditional vehicles, with a focus on spacious and practical body styles.
- f) **Seating Capacity:** Most EVs in the dataset offer 4 or 5 seats, making them suitable for families or small groups. There are niche segments for 2-seater sports cars and 7-seater SUVs, catering to more specific market demands. This balance shows that EV manufacturers are addressing a range of consumer needs, from sporty models to family-oriented vehicles.
- g) **Pricing Analysis:** There is significant variation in car prices across the dataset. High-end brands are naturally more expensive, while some manufacturers are targeting the budget-conscious market. A clear relationship exists between price and features like rapid charging – more expensive models tend to have more advanced features. This range in pricing offers consumers options depending on their budget and desired features.

- h) **Clustering and Market Segmentation:** The K-Means clustering revealed clear market segments, from luxury to budget cars and high-performance vehicles. This segmentation is valuable for manufacturers and marketers as it allows for a more focused approach to product development and marketing strategies. Understanding these segments enables better-targeted offerings.
- i) **Principal Component Analysis (PCA):** PCA was useful in reducing the complexity of the dataset and visualizing the key relationships between variables. The separation of clusters in the PCA plot demonstrates that the chosen features for segmentation were effective in defining distinct market segments. This reinforces the idea that the features selected provide valuable insights into the EV market.

Insights and Future Improvements:

- **Current Market Trends:** The analysis shows that the EV market is strongly focused on range, practicality, and charging convenience, while performance and luxury still have a significant place, particularly among premium brands. SUVs and sedans continue to be popular body styles, likely due to their practicality.
- **Growth Potential:** There is room for growth in segments like affordable EVs and performance-focused models. Advances in battery technology could increase the range and reduce charging times, making EVs even more attractive to a wider range of consumers.
- **Opportunities for Improvement:** With more time and resources, this market segmentation project could be improved by incorporating more data, such as consumer preferences or environmental data. Additionally, experimenting with other machine learning techniques, like hierarchical clustering or decision trees, could yield more refined insights into niche market segments.

3. Improvements for Future Segmentation Projects

- **Data Collection:** With more time and resources, we could significantly enhance our segmentation analysis by collecting additional data. Key features to include would be:
 - i. **Demographic Data:** Consumer demographics such as age, income, and location could provide a deeper understanding of target customer groups.
 - ii. **Charging Infrastructure:** Information on charging station availability in different regions would help refine segmentation for electric vehicle customers.

- iii. **Brand Loyalty:** Data on consumer loyalty and repeat purchases could improve segmentation by identifying long-term brand preferences.
- iv. **Geographic Data:** Adding regional availability and sales data for different car models would refine the analysis by considering geographic differences.
- **New Models to Explore:**
 - i. **Hierarchical Clustering:** This method would allow for more granular segmentation, potentially uncovering sub-clusters within the broader market segments.
 - ii. **Gaussian Mixture Models (GMM):** By considering the probabilities of belonging to multiple clusters, GMM could offer more flexibility in segmenting ambiguous data points.
 - iii. **DBSCAN:** This density-based algorithm could be used to identify outliers or niche segments that are not well-suited for K-Means.

4. Estimated Market Size

The global electric vehicle market is rapidly expanding, with an estimated market size of approximately **\$800 billion** by 2030. This figure is for the entire EV industry, covering both luxury and budget segments. It includes all potential electric vehicle categories, ranging from small city cars to high-end performance vehicles.

5. Key Variables for Optimal Market Segmentation

Based on our analysis, the following four variables emerged as the most critical for creating optimal market segments:

- i. **Price (PriceEuro):** This feature is essential in separating luxury brands from budget-conscious options.
- ii. **Range (Range_Km):** Driving range is a crucial factor for electric vehicle buyers, and it helps to differentiate between practical daily drivers and performance-oriented vehicles.
- iii. **Acceleration (AccelSec):** Acceleration performance is a key consideration for buyers interested in high-performance vehicles, and it further distinguishes the sportier options from the rest.

- iv. **Top Speed (TopSpeed_KmH):** Similar to acceleration, top speed is an important feature for performance-focused buyers, allowing us to segment vehicles based on their maximum speed capabilities.

These variables were selected because they provide a comprehensive view of customer priorities, enabling more accurate segmentation based on consumer preferences.

GitHub Link: https://github.com/Tushar264/Electric-Vehicle_Market-Segmentation