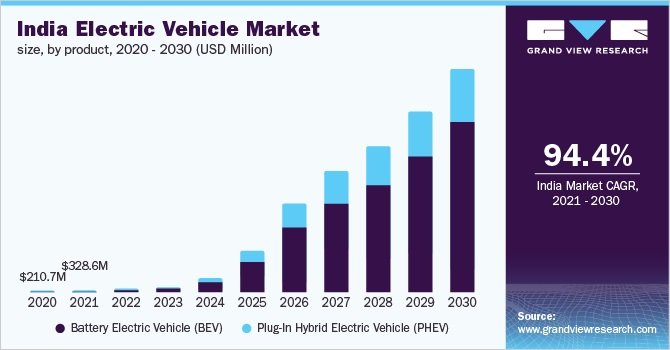
**INDIAN EV MARKET SEGMENTATION**

Himani Dadem



Evolution of EV Market in India

**FERMI ESTIMATION:**

Let's say we want to estimate the number of potential early adopters of electric vehicles in India. To do this, we can break down the problem into smaller pieces and make reasonable assumptions about each piece.

1. First, we can estimate the total population of India. According to the World Bank, the population of India was 1.366 billion in 2020.
2. Next, we can estimate the percentage of the population that is eligible to drive a car. According to the Indian Ministry of Road Transport and Highways, the legal driving age in India is 18 years old. Assuming that people stop driving at the age of 70, we can estimate that the eligible driving population is about 50% of the total population, or 683 million.
3. We can then estimate the percentage of the eligible driving population that owns a car. According to the Society of Indian Automobile Manufacturers, the number of passenger cars sold in India in 2020 was 2.44 million. Assuming that each car has an average lifespan of 10 years, we can estimate that the total number of cars on the road in India is about 24.4 million, or about 3.5% of the eligible driving population.
4. We can then estimate the percentage of the eligible driving population that is aware of electric vehicles. According to a survey conducted by the International Council on Clean Transportation, only about 10% of urban Indians are aware of electric vehicles. Assuming that the awareness rate is similar for the eligible driving population, we can estimate that about 68.3 million people are aware of electric vehicles.
5. Finally, we can estimate the percentage of the aware population that would be interested in purchasing an electric vehicle. According to a survey conducted by Deloitte, about 50% of urban Indians are interested in purchasing an electric vehicle. Assuming that the interest rate is similar for the aware population, we can estimate that there are about 34.15 million potential early adopters of electric vehicles in India.

Of course, these are very rough estimates, and the actual number of potential early adopters of electric vehicles in India may be higher or lower depending on a variety of factors. Nonetheless, Fermi Estimation can be a useful tool for quickly getting a sense of the size of a market segment, and can help inform strategic decisions about product development and marketing.

**DATA SOURCE:**

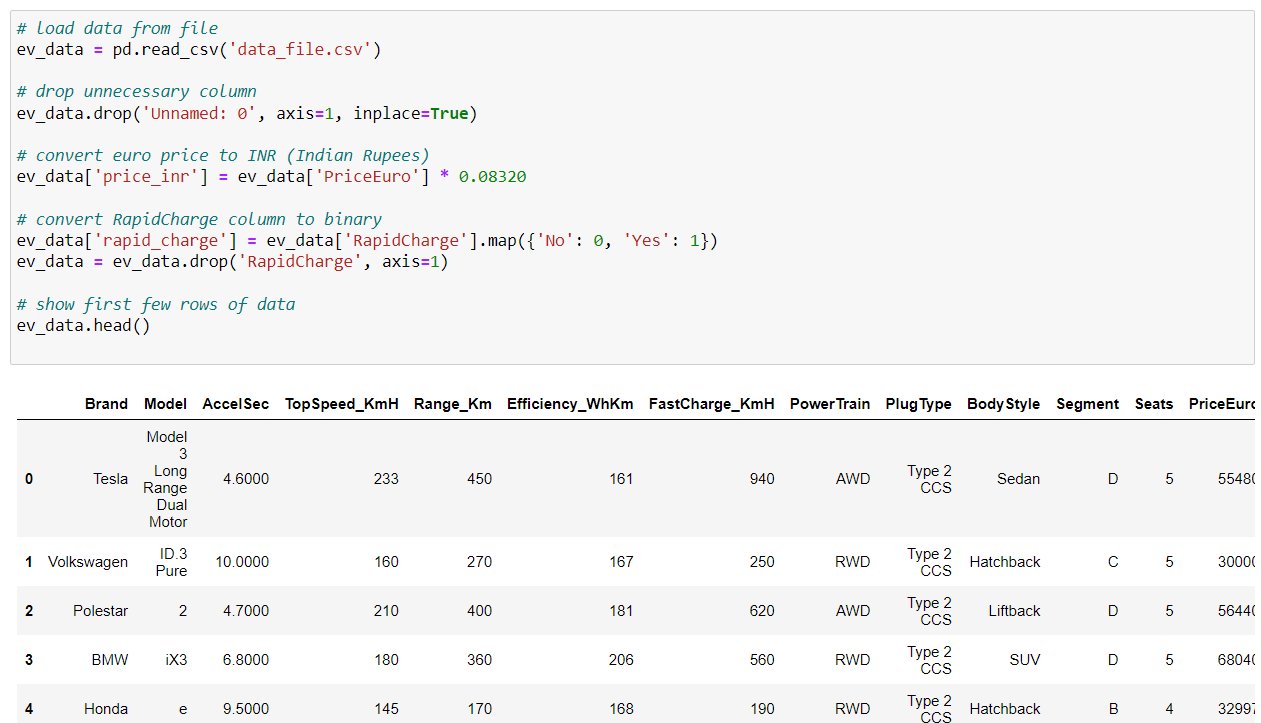
The data has been collected from multiple different sources listed below:

* <https://www.kaggle.com/datasets>
* <https://data.gov.in/>
* <https://www.data.gov/>
* <https://data.worldbank.org/>
* <https://datasetsearch.research.google.com/>

**DATA PRE-PROCESSING:**

**Data Cleaning:**

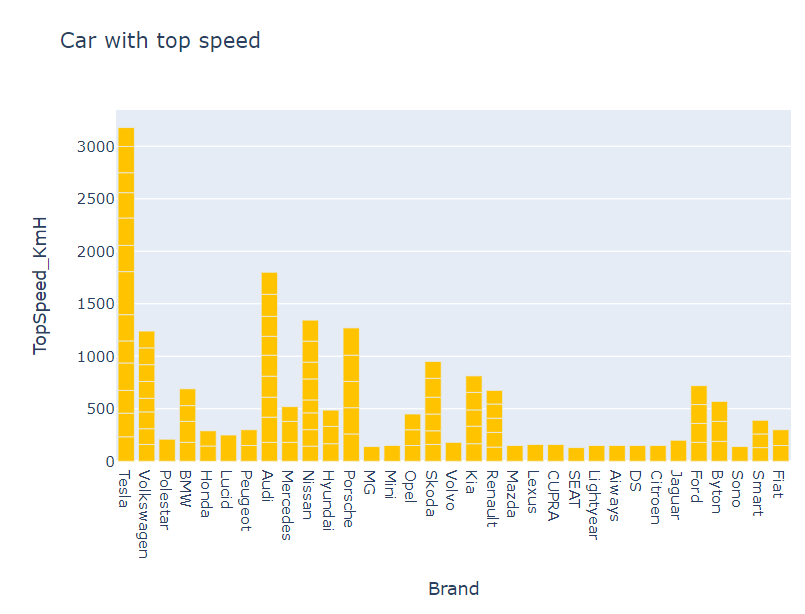
The gathered dataset is concise and serves both visualization and clustering purposes. The analytical process involves utilizing Python libraries including Pandas, NumPy, SciPy, and Scikit-Learn. The outcomes achieved are verifiable and can be replicated consistently.

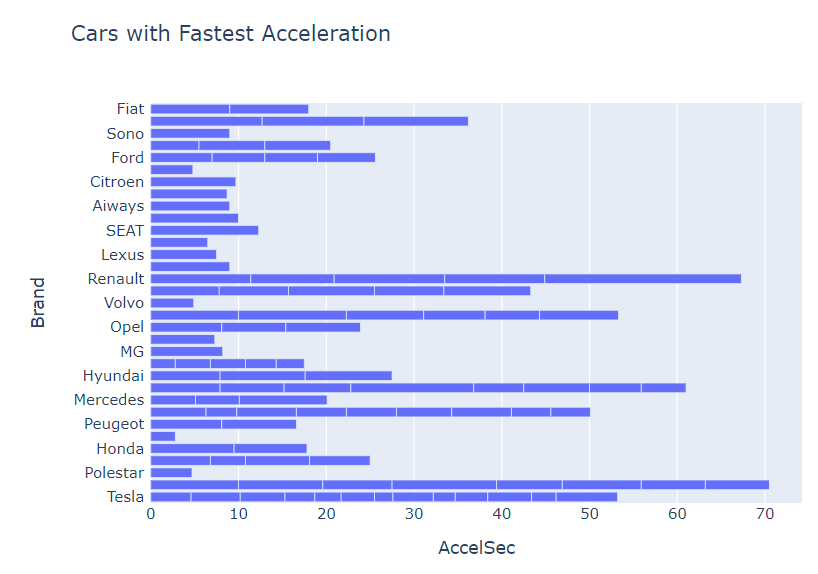


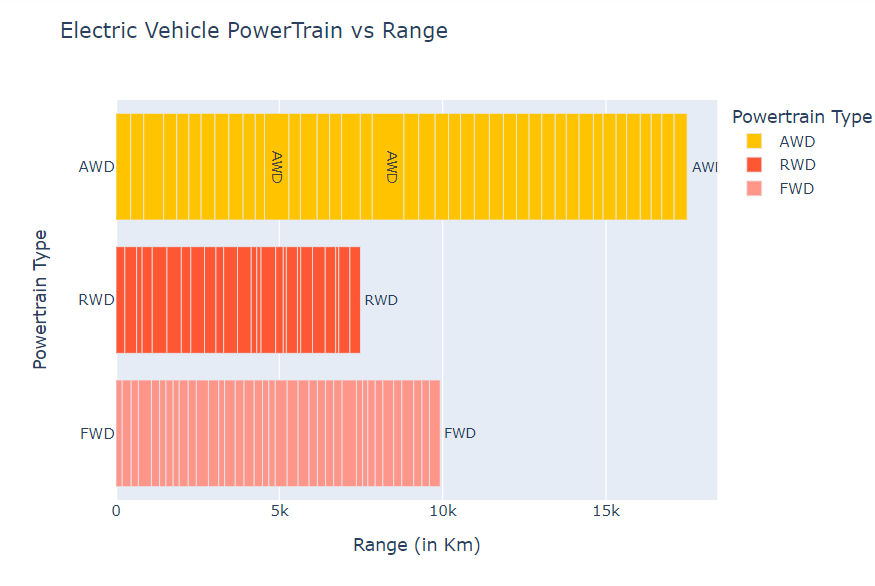
**EDA:**

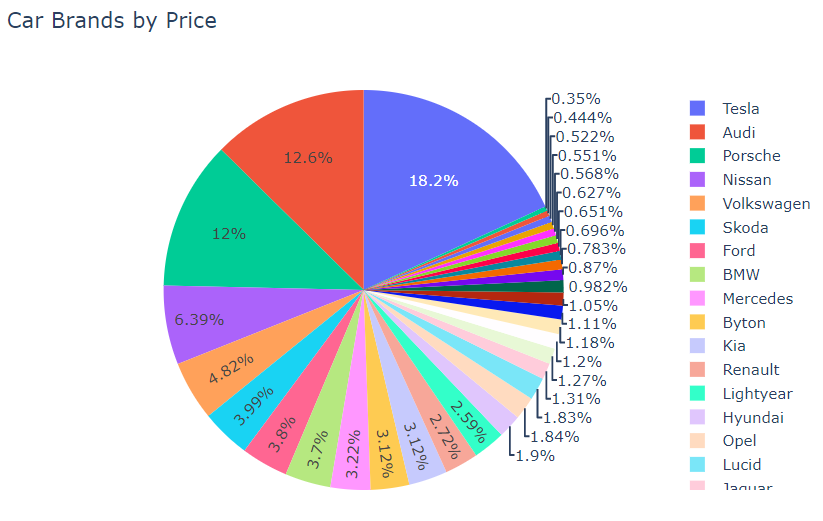
The data that was collected is compact and serves a dual purpose of both visualization and clustering. To carry out the analytical workflow, various Python libraries like NumPy, Pandas, Scikit-Learn, and SciPy are utilized. The results obtained are ensured to be reproducible and can be verified consistently.

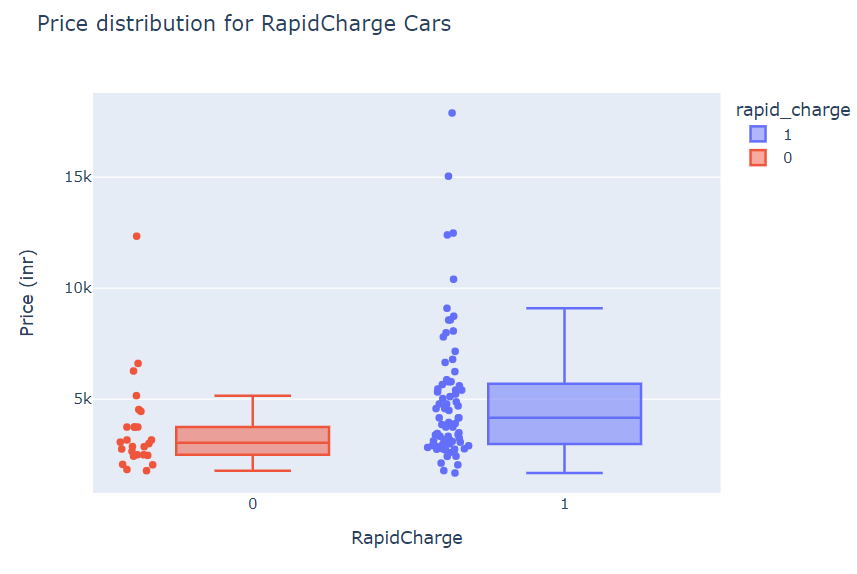
**Comparison plots:**







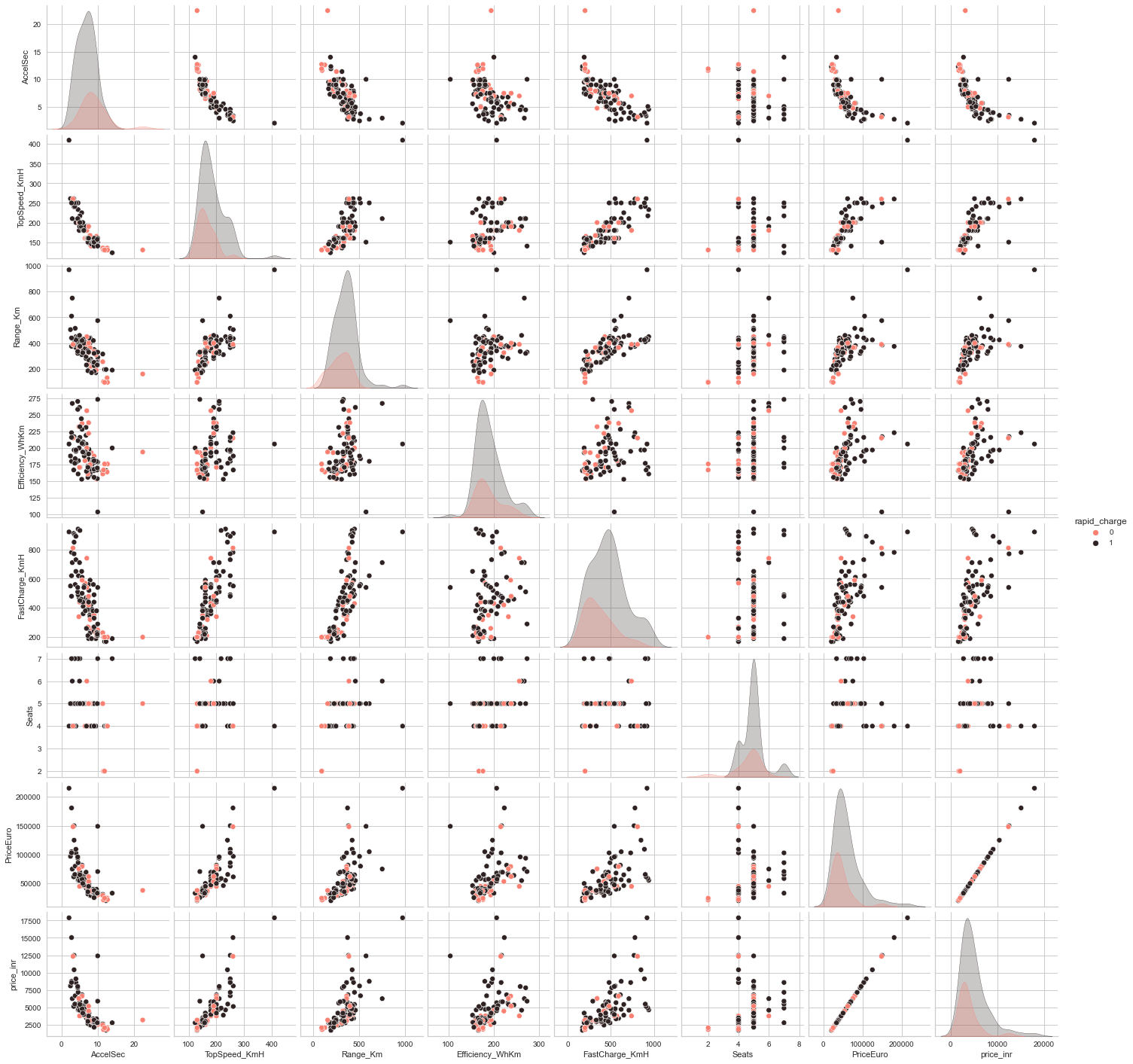




**Pairplots**:

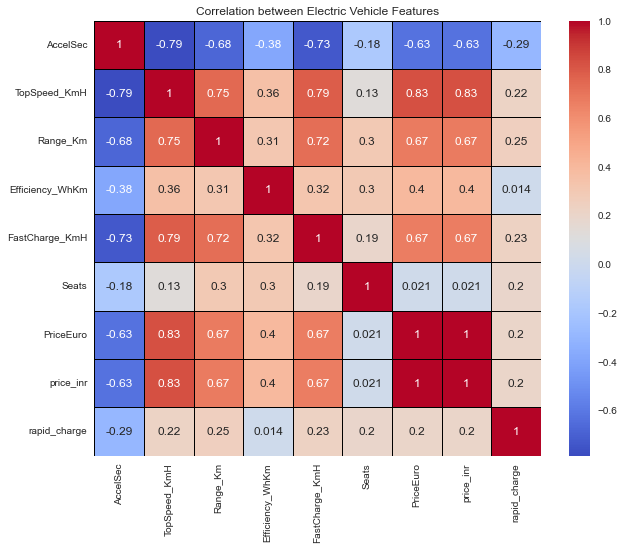
A parallel coordinates plot (parplot) is a data visualization technique used to display multivariate data. It allows the user to visualize how a data point changes with each variable along a parallel axis. The variables are represented by vertical axes, and each data point is plotted as a line that connects the values of the variables. This enables easy identification of trends and patterns in the data.

Parplots are particularly useful for exploring high-dimensional data, as they allow for easy identification of relationships between variables. They also enable the identification of outliers and patterns in the data that may not be immediately obvious with other visualization techniques.

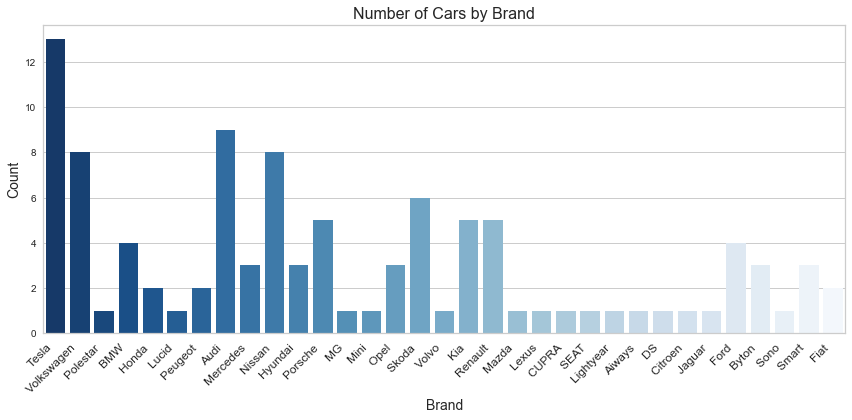


**Correlation Matrix:**

The matrix displays the correlation coefficients for each possible pair of variables in a table format, which can be visualized as a heatmap. A strong relationship between two variables is indicated by a correlation coefficient greater than 0.7. By examining the correlation matrix, researchers can identify which variables are most strongly correlated and use this information to guide further analysis. However, it is important to note that correlation does not necessarily imply causation, and other factors may be influencing the relationship between variables.

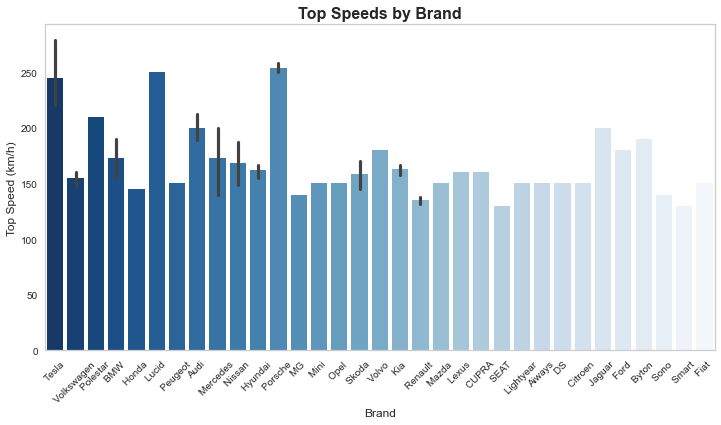
****

**Brands Frequency**:



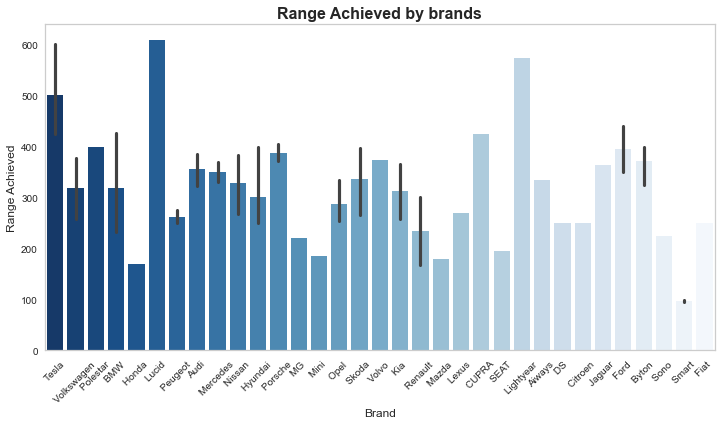
Byton , Fiat and smart are the prominent brands and Polestar being the least

**Top Speed of brands:**



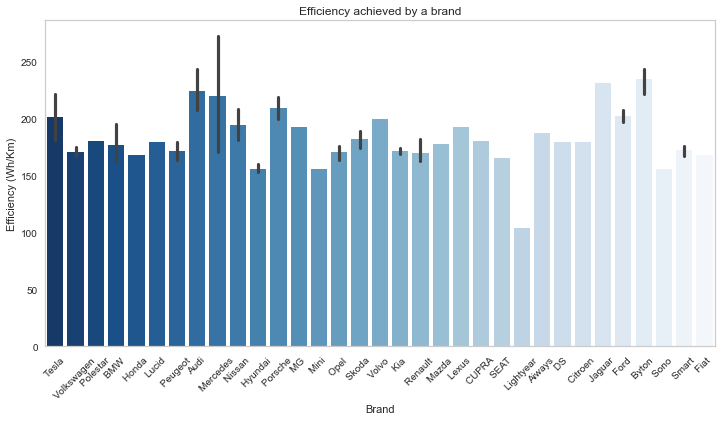
Porsche, Lucid and Tesla produce the fastest cars and Smart the lowest

**Range Achieved by Brands:**



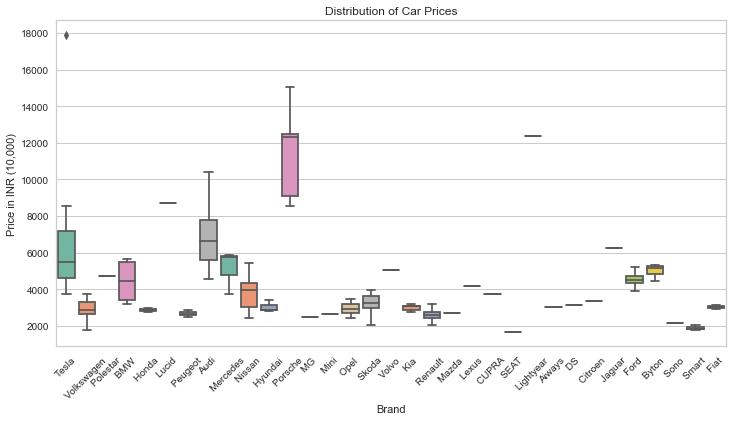
Lucid, Lightyear and Tesla have the highest range and Smart the lowest

**Efficiency of a car:**



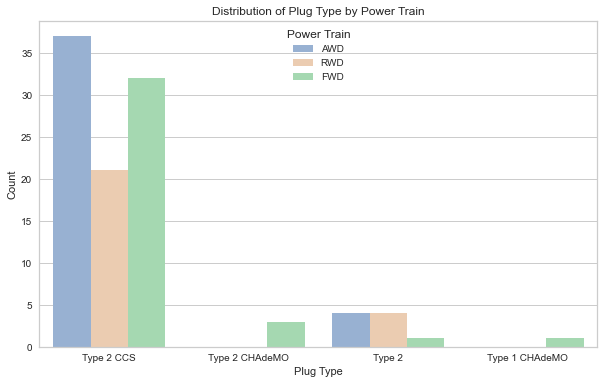
Byton , Jaguar and Audi are the most efficient and Lightyear the least

**Car price:**



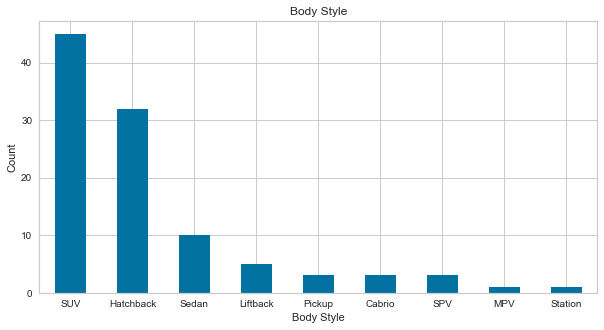
Lightyear, Porsche and Lucid are the most expensive and SEAT and Smart the least

**Types of charging plus:**



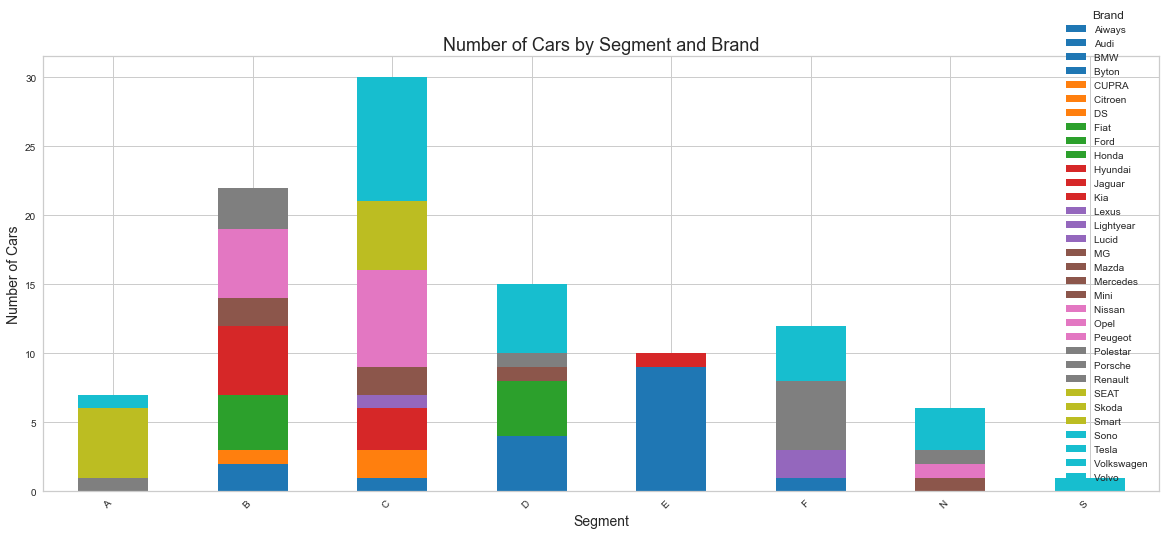
Most companies use Type 2 CCS and Type 1 CHAdeMo the least

**Different body types of the cars:**



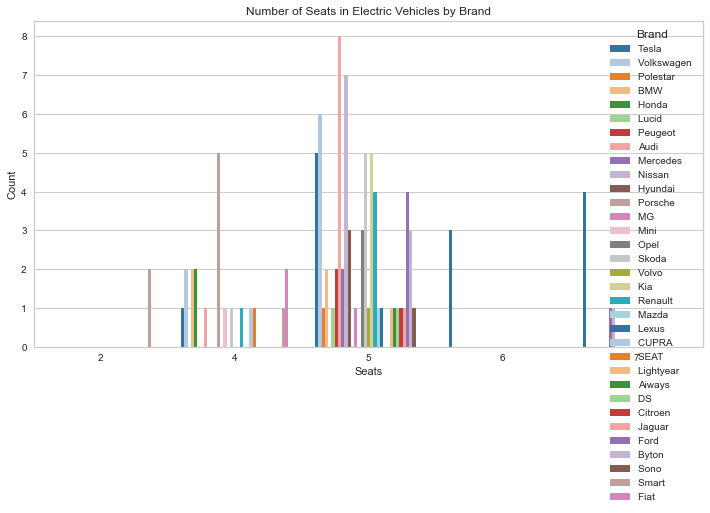
Most cars are either SUV or Hatchback

**Segment in which the cars fall under:**



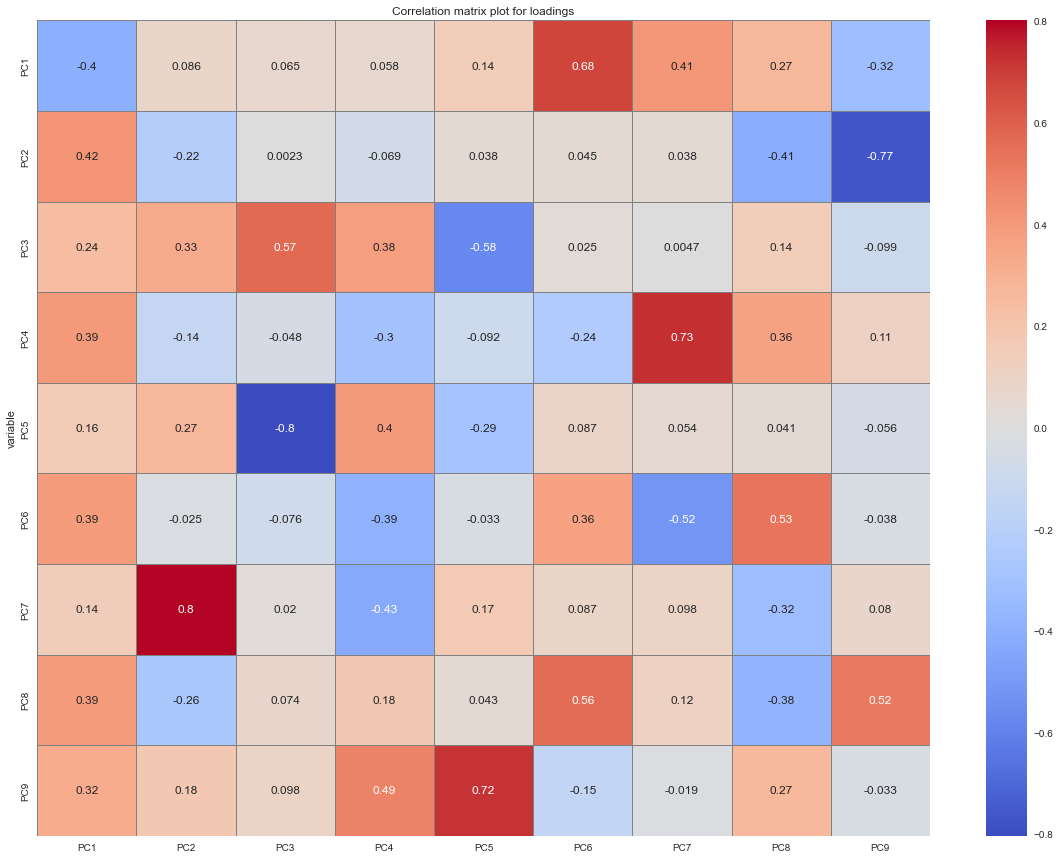
Most cars are either C or B type

**Number of Seats:**



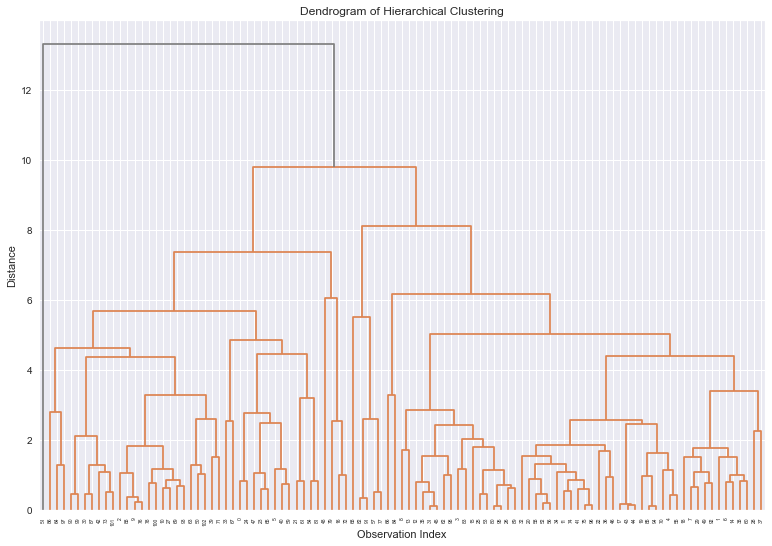
Majority of cars have 5 seats

**Correlation matrix plot for loadings**:



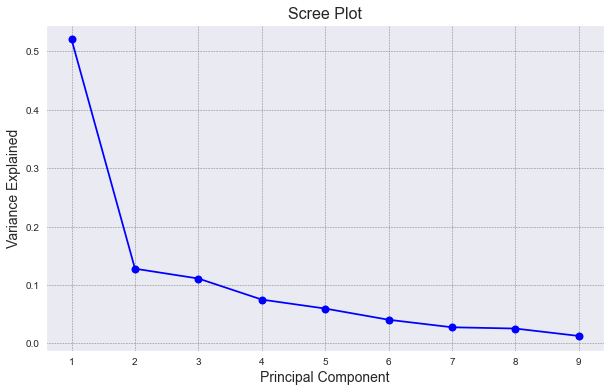
**Dendrogram:**

This technique is specifically tailored to the agglomerative hierarchical method of clustering. This method involves starting with each point as a separate cluster and progressively merging them in a hierarchical manner based on their distances. To determine the optimal number of clusters for hierarchical clustering, a dendrogram is utilized. The dendrogram is a tree-like chart that illustrates the sequences of cluster merges or splits. If two clusters are merged, the dendrogram joins them in a graph, and the height of the join corresponds to the distance between the clusters. As depicted in the diagram, we can determine the optimal number of clusters based on the hierarchical structure of the dendrogram. In line with other cluster validation metrics, the agglomerative hierarchical method can typically be best suited for four to five clusters.



**Scree Plot:**

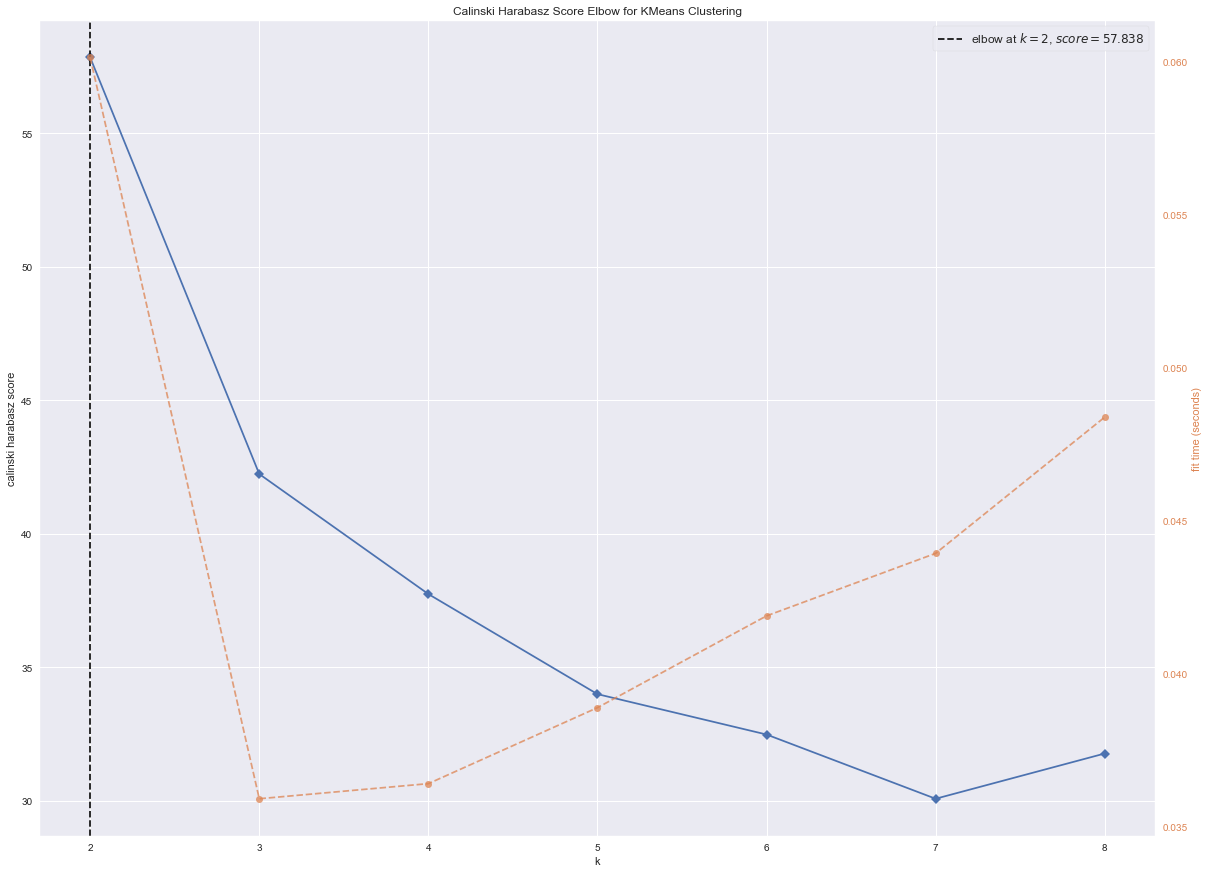
A common approach for determining the number of retained PCs is through the Scree Plot, which employs a graphical representation. This simple line segment plot illustrates the eigenvalues of each PC, with the y-axis showing the eigenvalues and the x-axis displaying the number of factors. Typically, the plot depicts a downward curve, with most Scree Plots sharing a similar shape. The plot usually starts high on the left, falls quickly, and then flattens out. This shape is due to the first component explaining a large portion of the variability, the following few components explaining a moderate amount, and the latter components describing only a small fraction of the variability. The Scree Plot criterion identifies the "elbow" in the curve and selects all components before the line levels off. To ensure that the chosen PCs can account for at least 80% of the variance, the proportion of variance plot is used.



**Elbow Method:**

The Elbow method is a widely used technique for determining the optimal number of clusters. The approach involves calculating the Within-Cluster-Sum of Squared Errors (WSS) for different numbers of clusters (k) and selecting the k for which the WSS changes start to diminish. The concept behind the elbow method is that the explained variation changes quickly for a small number of clusters, then slows down, resulting in the formation of an elbow in the curve. The elbow point represents the number of clusters that can be used for the clustering algorithm.

To facilitate this process, the KElbowVisualizer function fits the KMeans model for a range of cluster values between 2 to 8. The function highlights the elbow point, which is depicted in the diagram. Additionally, the function reports the amount of time required to generate models for different numbers of clusters through the green line.



**Clustering:**

Clustering is a commonly used exploratory data analysis technique to gain insight into the structure of data. It involves identifying subgroups in the data where the data points within the same subgroup, or cluster, are very similar, while those in different clusters are very different. In essence, the objective is to discover homogeneous subgroups within the data, where the data points in each cluster are as similar as possible, as determined by a similarity measure such as euclidean-based distance or correlation-based distance.

The selection of an appropriate similarity measure is dependent on the specific application. Clustering analysis can be performed based on features, where subgroups of samples are identified based on features, or based on samples, where subgroups of features are identified based on samples.

K-Means Algorithm:

K-means clustering is a commonly used unsupervised machine learning algorithm for clustering data. The algorithm works by partitioning the dataset into a fixed number of clusters, K, where each data point is assigned to the nearest centroid (mean) of the cluster. The K-means algorithm has the following steps:

Initialization: Randomly select K points as the initial centroids for the K clusters.

Assignment: Assign each data point to the nearest centroid based on Euclidean distance or other similarity measures.

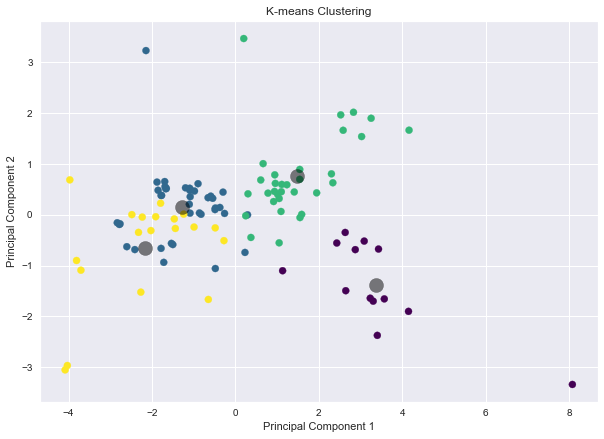
Update: Recalculate the centroid of each cluster by taking the mean of all data points assigned to that cluster.

Repeat: Repeat steps 2 and 3 until the cluster assignments do not change or a maximum number of iterations is reached.

The algorithm seeks to minimize the sum of squared distances between each data point and its assigned centroid. The K-means algorithm is sensitive to the initial centroid selection, so it is common practice to run the algorithm multiple times with different initializations and choose the result with the lowest sum of squared distances.

K-means clustering can be used for a variety of applications, such as image segmentation, customer segmentation, and anomaly detection. However, it may not work well with non-linear data distributions or clusters with varying shapes and sizes.



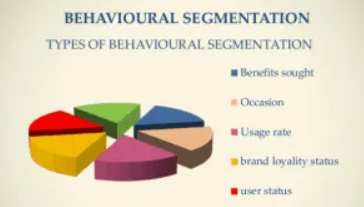


**SEGMENT EXTRACTION:**

Electric Vehicle Market Segmentation aims to categorize the target market into different segments based on their geographic location, socio-demographic factors, behavioral patterns, and psychographic characteristics.

**Behavioral Segmentation:**

Behavioral Segmentation specifically looks for similarities in actual or reported behavior, such as previous experience with the product and the amount spent on the purchase.



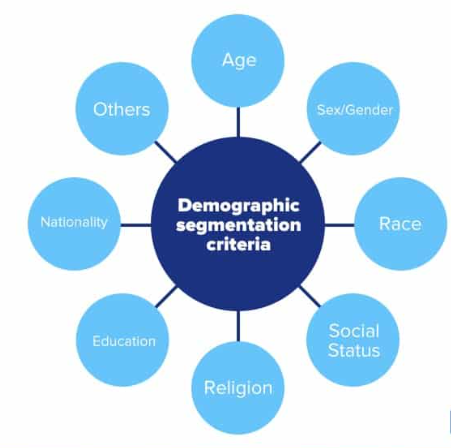
**Psychographic Segmentation:**

Psychographic Segmentation involves grouping individuals based on their beliefs, interests, preferences, aspirations, or the benefits they seek when making a purchase. This type of segmentation is particularly useful for lifestyle segmentation. It involves many segmentation variables to identify and target specific consumer groups with similar psychographic characteristics.



**Demographic segmentation:**

Demographic segmentation is a market segmentation strategy that divides a population into different groups based on demographic variables such as age, gender, income, education, occupation, family size, and ethnicity. This type of segmentation is based on the premise that different groups of people have different needs, wants, and behaviors that can be used to tailor marketing efforts to those groups more effectively.



**Segmentation in the case of Electric Vehicle Market:**

Market segmentation is a marketing strategy that involves dividing customers into homogenous subgroups to which marketers can target with similar marketing strategies. There are two ways of segmenting the market: a-priori and post-hoc. The a-priori approach uses predefined characteristics such as age, gender, income, and education to define segments, followed by profiling based on behavioural, psychographic, or benefit-related variables. On the other hand, the post-hoc approach identifies segments based on the relationship among multiple measured variables. Both approaches use measured variables to determine the segmentation theme. This study uses the a-priori approach to segment potential EV customers into sub-groups.

The use of a blended approach of psychographic and socioeconomic attributes for market segmentation can be beneficial in formulating sub-market strategies that cater to specific consumer tastes and preferences. This approach considers not only demographic and economic factors but also psychographic factors, such as attitudes, values, and lifestyles.

Straughan and Roberts found that psychographic characteristics were more useful than demographic and economic characteristics in explaining environmentally-conscious consumer behavior. This suggests that consumer behavior cannot be fully explained by demographic and economic factors alone, and that psychographic characteristics play an important role in shaping consumer attitudes and preferences.

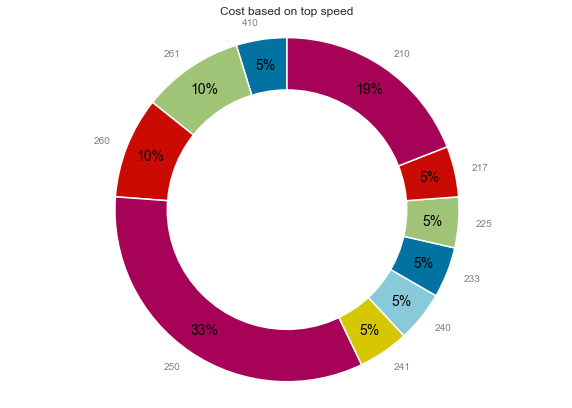
In the present study, perceived-benefit characteristics are added to the blend of psychographic and socioeconomic attributes for segmenting the consumer market. This can help to identify consumer segments with similar perceived benefits from eco-friendly products, which can inform targeted marketing strategies and product development efforts.

Overall, a blended approach of psychographic and socioeconomic attributes can provide a more nuanced understanding of consumer behavior and preferences, which can lead to more effective marketing strategies and product offerings.

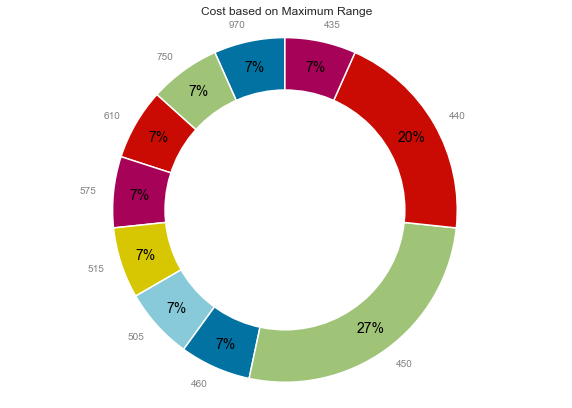
**PROFILING AND DESCRIBING THE SEGMENTS:**

Arranging the prices in ascending or descending order and then matching the corresponding top speeds and maximum range values to them.

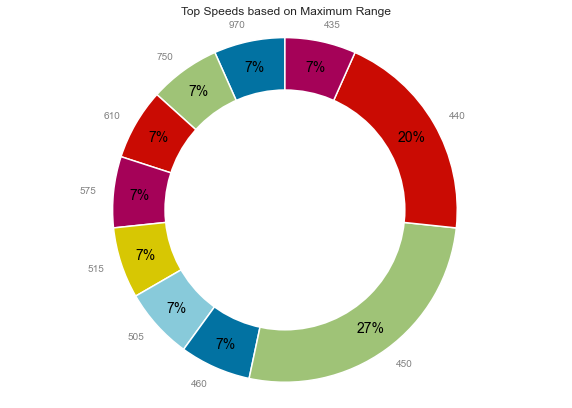
**Cost based on Top speed:**



**Cost based on Maximum range:**



**Top Speeds based on Maximum Range:**



**SELECTION OF TARGET SEGMENT:**

Based on the analysis provided, the optimum target segment for eco-friendly cars would be those that have the following characteristics:

Behavioral:

* Cars with 5 seats

Demographic:

* Top speed and range are important factors, indicating that the target segment is likely to be interested in cars that have a high maximum range and top speed
* Efficiency is also an important factor, suggesting that the target segment is likely to be interested in cars that are highly efficient

Psychographic:

* Price is an important factor, with the target segment likely to be interested in cars that fall within the price range of 1600000 to 18000000.

Overall, the target segment for eco-friendly cars is likely to be a group of consumers who value efficiency, range, and top speed, while also being price-sensitive. Marketing efforts should focus on highlighting these features and benefits, and positioning eco-friendly cars as a practical and affordable alternative to traditional gasoline-powered vehicles.

Yes, based on the analysis provided, the target segment for eco-friendly cars should have cars that are highly efficient, with a focus on top speed and a price range between 16 to 180 lakhs, and preferably with 5 seats. By targeting this segment, eco-friendly car manufacturers can appeal to consumers who are looking for practical and cost-effective alternatives to traditional gasoline-powered cars, while also addressing the concerns of those who prioritize speed and performance.

**CUSTOMIZING THE MARKET MIX:**

Based on the 4Ps of marketing mix, the following strategies can be considered for promoting eco-friendly cars in the target market:

Price: The price of eco-friendly cars should be competitive and reasonable, considering the target segment's ability to pay. To attract more price-sensitive consumers, manufacturers can also offer attractive financing options or discounts.

Product: Eco-friendly cars should be designed to meet the needs and preferences of the target segment. This can include features such as high efficiency, top speed, and spacious interiors with 5 seats. Additionally, eco-friendly cars should be marketed as a practical and cost-effective alternative to traditional gasoline-powered cars.

Place: The distribution strategy should focus on making eco-friendly cars easily accessible to the target segment. This can include partnering with dealerships and online marketplaces to make the cars available in key locations. Additionally, manufacturers can also offer home delivery services and other convenient options to make the buying process easier for consumers.

Promotion: To promote eco-friendly cars, manufacturers can use a mix of advertising, word-of-mouth marketing, and direct marketing tactics. This can include digital marketing campaigns, influencer partnerships, and targeted advertisements in publications that are popular among the target segment. Additionally, manufacturers can also offer test drive opportunities and organize events to showcase the benefits and features of eco-friendly cars.

Overall, a well-designed marketing mix can help manufacturers of eco-friendly cars to effectively target and attract the desired segment, ultimately leading to increased sales and brand recognition.

The marketing mix consists of four elements - Price, Product, Place, and Promotion, which together form a business plan and can lead to great success if executed correctly. These elements are interconnected and depend on each other, and it requires a deep understanding of the target market, consumer behavior, industry trends, and competition to develop a successful marketing mix. This requires market research and consultation with various stakeholders, including consumers, suppliers, and industry experts. The feedback received from these consultations is used to refine the marketing mix to meet the needs of the target market. Regular monitoring and evaluation of the marketing mix's performance can help identify areas of improvement and optimize the marketing efforts.

**REFERENCES:**

1. *Sara Dolnicar , Bettina Grün , Friedrich Leisch (2018) Market Segmentation Analysis Understanding It, Doing It, and Making It Useful*
2. *Deepak Jaiswal,Arun Kumar Deshmukh(2022) Who will adopt electric vehicles? Segmenting and exemplifying potential buyer heterogeneity and forthcoming research, Journal of Retailing and Consumer Services .*
3. *McDonald, M., amp; Dunbar. (2003). Market segmentation. Butterworth Heinemann.*