EV Segmentation

Spencer Dsheel

# **Market Segmentation**

Market segmentation is the process of dividing a larger market into smaller groups of consumers with similar characteristics, needs, or behaviours. Businesses can optimise their products, services, and marketing strategies by identifying and focusing on specific segments of their audience. This enables them to cater for the unique needs and preferences of their target customers more effectively.

Offering the correct product to the right consumer at the appropriate time is the key to a business's success. To accomplish this, marketers must have a thorough understanding of their customers. Market segmentation is a critical tool for companies to examine and classify consumer behaviour, which can yield enormous benefits. Since every consumer has distinct needs, a marketing strategy that caters to everyone will not be sufficient to engage users effectively. Marketing segmentation enables organizations to divide users into distinct groups, analyze their requirements, and communicate with them proficiently.

## **Types of Segmentation**

### **Demographic Segmentation**



Segmentation based on demographic characteristics is a broad and straightforward categorization. It involves grouping individuals based on factors such as age, gender, religion, location, or occupation. This segmentation method is used to create large groups of people with similar demographic characteristics.

For example, beauty cream is more relevant to women aged 16 to 35 than men in their 80s.

### **Psychographic Segmentation**

Segmentation based on psychographics involves grouping users by their attitudes, lifestyles, interests, or values. While this method of segmentation is more complex than demographic segmentation, it provides marketers with a more profound understanding of their audience.

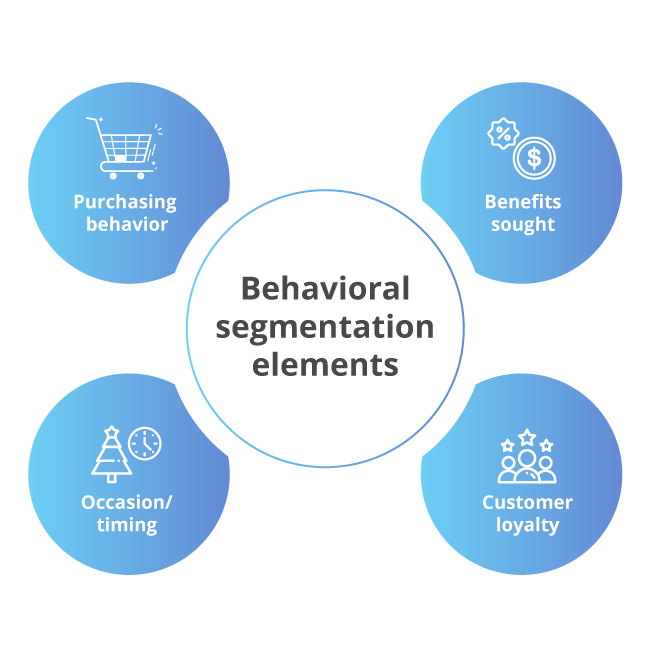
For example, if a consumer is highly concerned about the privacy features of her smartphone, she might opt to buy an iPhone.



### **Behavioral Segmentation**

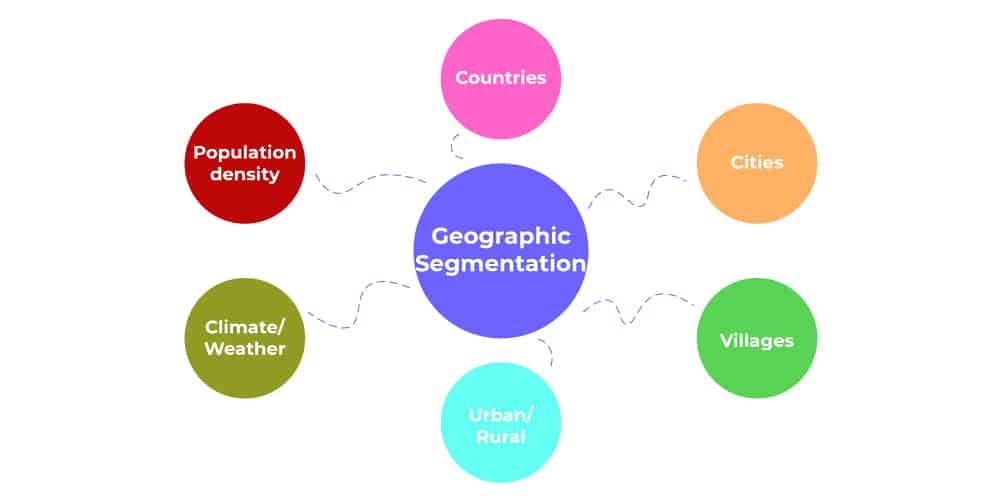
Behavioral segmentation makes segments based on the customer’s behaviour when interacting with a company or making purchase decisions. The segmentation is done by analyzing patterns of behaviour exhibited by customers. No meaning is removed, and no new sentences are added.

For example, buying patterns of a user, brand interaction, website interaction, engagement on an app or website, and loyalty.



### **Geographic Segmentation**

Geographic segmentation segments people based on geographical boundaries. Factors like climate, cultural preferences, and urban-suburban-rural divides are also used for the segmentation process.



# **EV Market Overview**

The electric vehicle (EV) market has experienced significant growth and transformation in recent years, driven by technological advancements, environmental concerns, government regulations, and shifting consumer preferences. The growth is primarily attributed to increasing government initiatives promoting clean energy, declining battery prices, improving charging infrastructure, and rising consumer awareness about environmental sustainability.

The electric vehicle market is highly competitive, with several established players and numerous startups entering the space. Major automotive manufacturers such as Tesla, Volkswagen, Nissan, BMW, and General Motors have been investing heavily in electric vehicle development and production. Additionally, technology companies like Rivian, Lucid Motors, and BYD have gained prominence in the EV market with innovative electric vehicle offerings. Collaboration and partnerships between traditional automakers and technology firms are becoming increasingly common to accelerate electric vehicle adoption and address challenges such as battery technology and manufacturing scale.

Advancements in battery technology, including solid-state batteries and improved energy density, are anticipated to further enhance the performance and affordability of electric vehicles. The rise of autonomous and connected electric vehicles is poised to revolutionize transportation and mobility services, offering new business models and opportunities for industry stakeholders.

# **Technique Used for Segmentation**

## **KMeans Algorithm**

K-Means Clustering is an Unsupervised Machine Learning algorithm, which groups the unlabeled dataset into different clusters. K means clustering is a method that groups data points into K clusters based on their distance from the center of the clusters. The first step of this method is to randomly assign the clusters' centroid in the space. Then, each data point is assigned to one of the clusters based on its distance from the centroid of the cluster. After assigning each point to one of the clusters, new cluster centroids are assigned. This process is run iteratively until it finds a good cluster. During the analysis, we assume that the number of clusters is given in advance, and we have to put points in one of the groups.

### **Algorithm:**

* Randomly initialize K cluster centroids in the feature space. These centroids can be either randomly selected data points or randomly generated points.
* For each data point, calculate the distance (e.g., Euclidean distance) to each centroid and assign the point to the nearest centroid.
* After all data points have been assigned to clusters, recalculate the centroids by taking the mean of all data points assigned to each cluster.
* Repeat the assignment and update steps until convergence criteria are met.

### **Choice of K**

* The selection of the number of clusters (K) is crucial and often requires domain knowledge or exploration of the dataset.
* Techniques such as the elbow method or silhouette analysis can be used to determine an optimal value of K.

# **Implementation**

## **Data Sources**

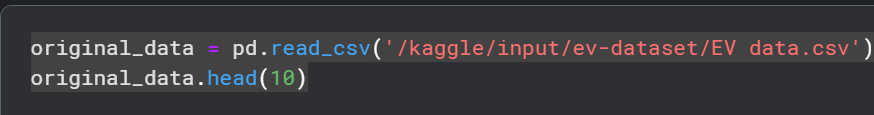
We used the gathered datasets which are related to the case.

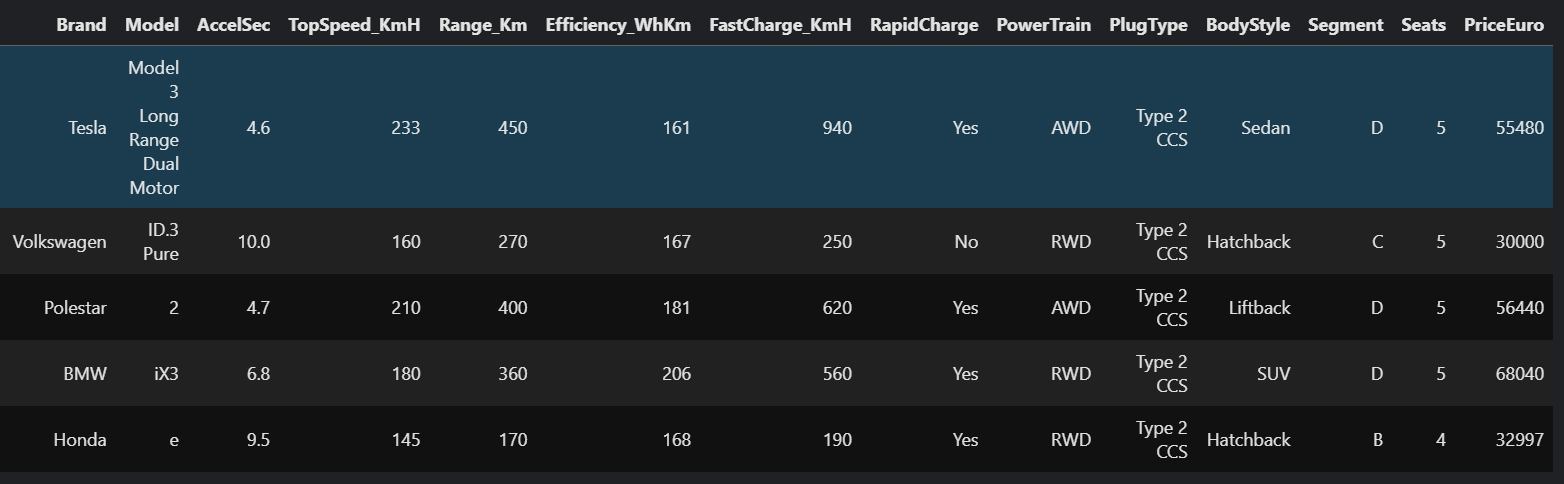
### **Libraries and Tools used:**

Pandas - Pandas is an open-source Python library widely used for data manipulation and analysis.

Numpy - NumPy is a library for numerical computing in Python

We used the dataset named “EV Data”.



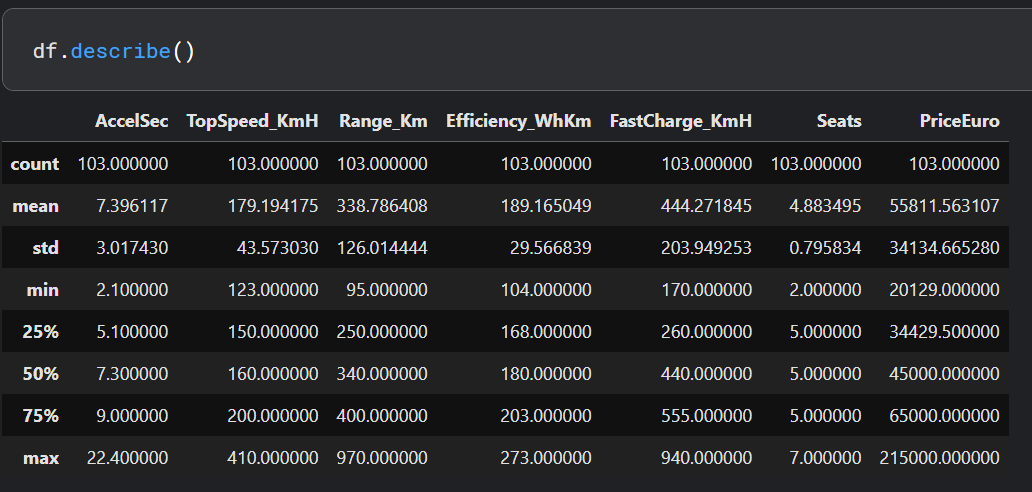


## **Data Preprocessing**

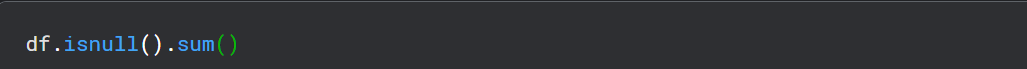
### **Libraries used:**

SKLearn - open-source machine learning library for Python.

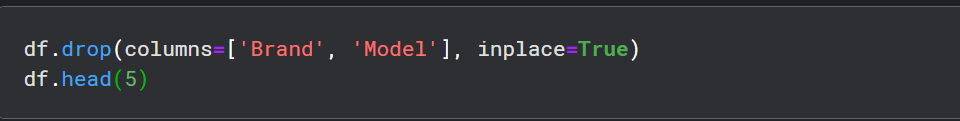
MatPlot library, Seaborn – libraries used for visualication.



Provides a summary of descriptive statistics for numerical columns in the DataFrame. It computes statistics such as count, mean, standard deviation and so on.

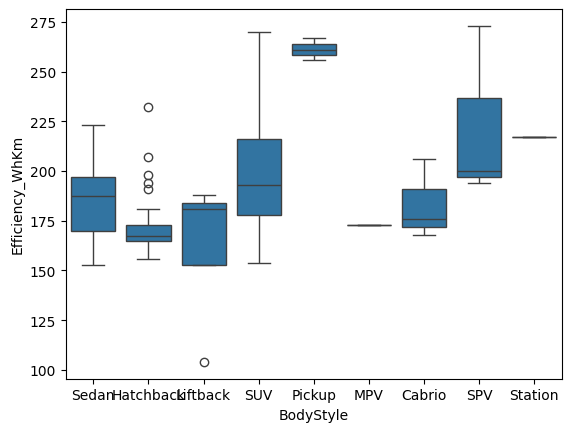


Checks for Null values.



Dropping features which are not relevant for segmentation,

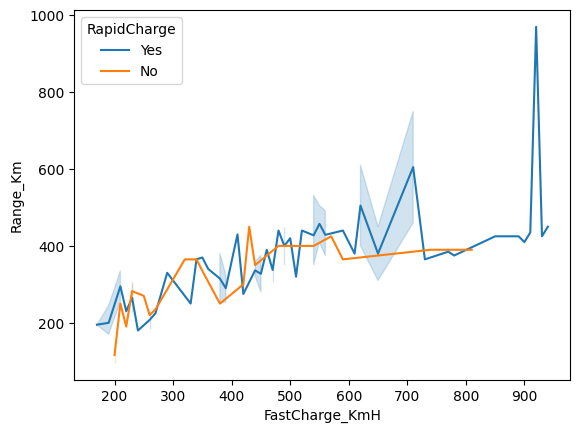
## **Analysis and Visualization**

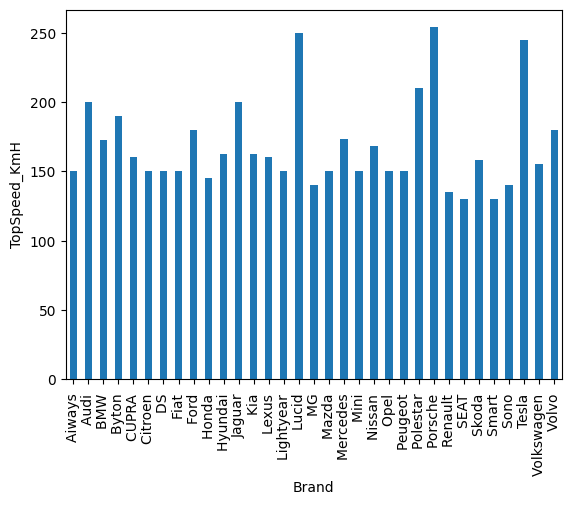


Box Plot to plot the Body Style to compare it with Efficieny\_WhHm.

The line plot provides insights into the relationship between the "FastCharge\_KmH", "Range\_Km", and "RapidCharge" variables in the data frame, allowing you to visually compare how the range varies with fast charging speed for different rapid charging capabilities.

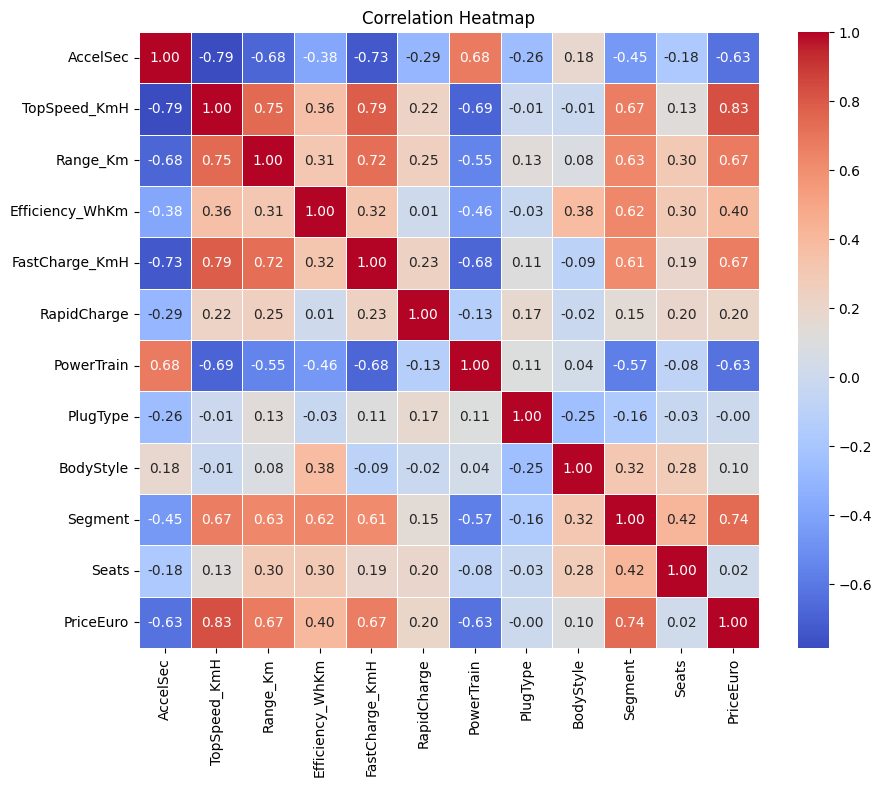
The third graph visually compares the average top speeds of vehicles across different brands, allowing you to quickly identify which brands tend to produce vehicles with higher or lower average top speeds.





The corr() method in Pandas DataFrame calculates the pairwise correlation between columns, returning a correlation matrix. This correlation matrix shows how each numerical column in the DataFrame is correlated with every other numerical column.

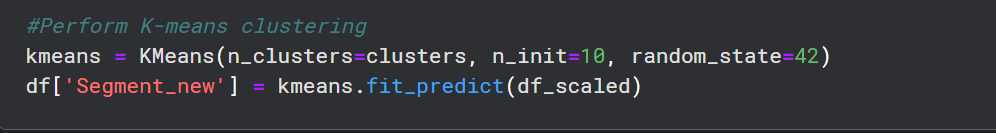
You can use the correlation matrix to identify patterns and relationships between variables in your dataset, which can be useful for feature selection, identifying multicollinearity, and understanding the data's underlying structure. Additionally, you can visualize the correlation matrix using heatmaps for easier interpretation.



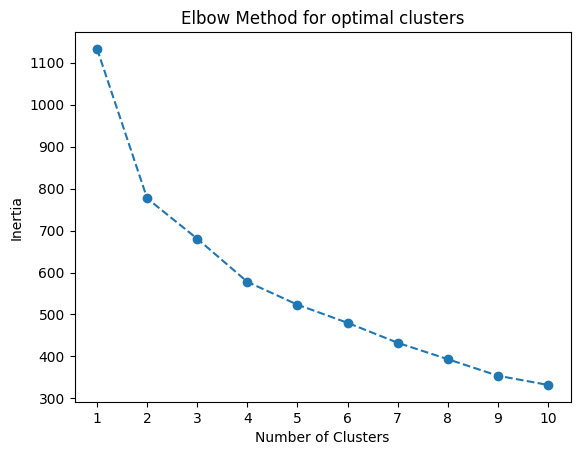
## **Segmentation**

### **KMeans Algorithm**

Using KMeans Algorithm for performing clustering on the dataset.

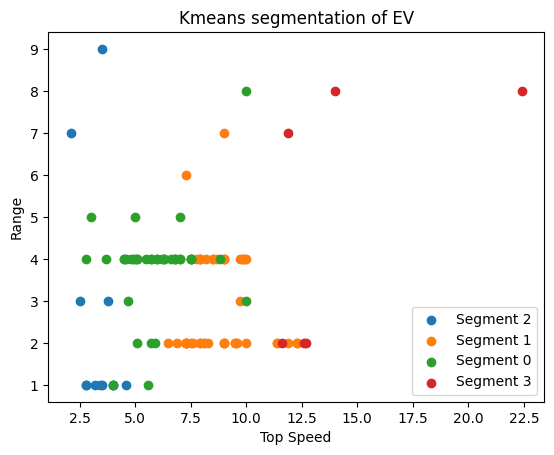
****

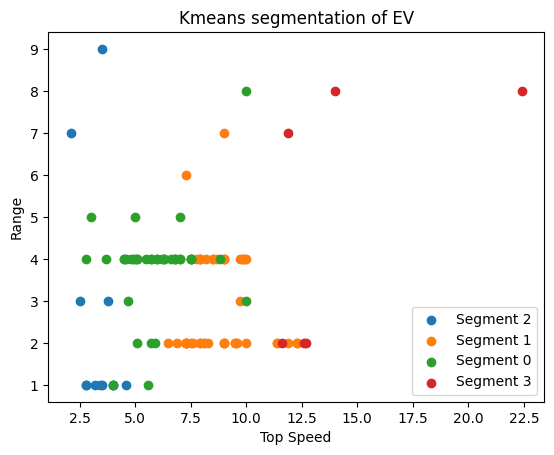
And we can use the Elbow method to find the optimum K value

.

As we can see after 4 clusters the inertia is sane to all the other clusters, so we assume that 4 is the optimal cluster.

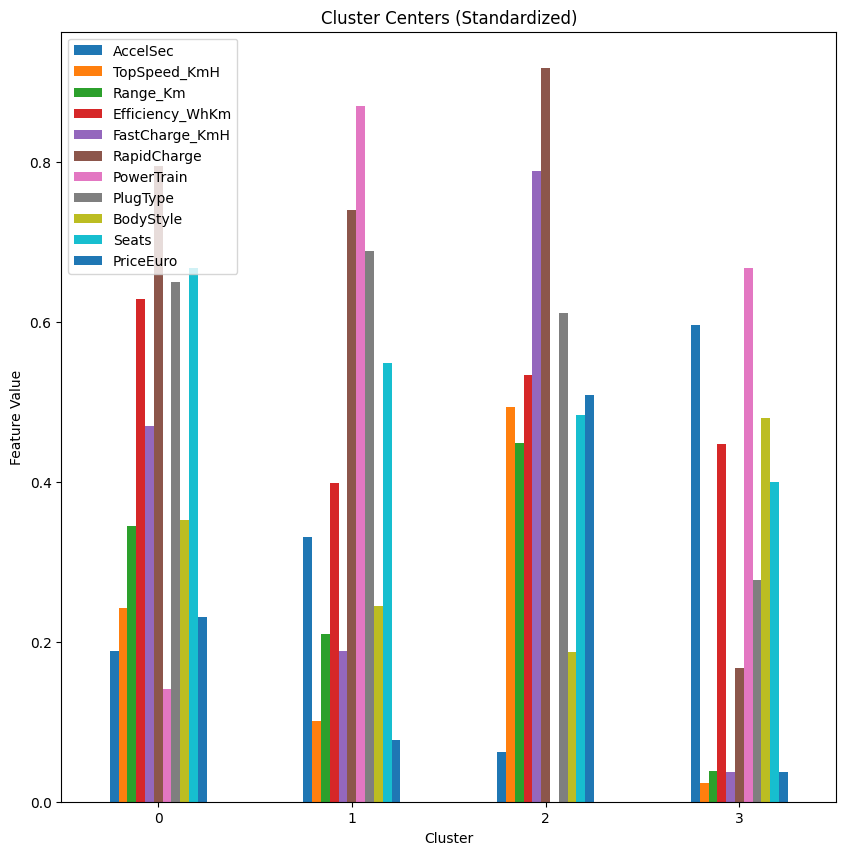
After segmentation, these graphs represent the clustering of the features.





## **Calculate Cluster Centers**

Calculating the cluster centre of the 4 clusters we determine the average or typical value of each feature for data points in that cluster. Features that exhibit significant differences between clusters, as these features are likely to be important for distinguishing between clusters.



The graph below showcases the strength or the distribution of data points across different clusters generated by the clustering algorithm.

