

Electric Vehicle Market Segmentation

*Navigating the Electric Vehicle Market through Behavior and
Vehicle Features Segmentation*

BY:

Sahil Barbade

Solo Team Member

Overview

- *Electric Vehicles (EVs) represent a significant shift in automotive technology and are poised to revolutionize the transportation industry. Unlike traditional internal combustion engine vehicles that rely on gasoline or diesel, EVs are powered by electric motors, which draw energy from rechargeable batteries. This fundamental difference offers several advantages, including reduced greenhouse gas emissions, decreased dependence on fossil fuels, and potentially lower operating costs.*
- *There are primarily three types of Electric Vehicles:*
 - *BEVs: Fully electric, zero-emission vehicles running solely on battery power, charged by plugging into a power source.*
 - *PHEVs: Combine electric motor and internal combustion engine, offering electric mode for short trips and switching to gas for longer ranges.*
 - *HEVs: Use both electric motor and internal combustion engine, but can't be plugged in; battery charged via braking and engine, offering improved fuel efficiency but not fully electric.*
- *The adoption of EVs has been steadily increasing worldwide due to factors such as advancements in battery technology, government incentives, environmental concerns, and consumer demand for sustainable transportation options. Major automotive manufacturers are investing heavily in EV development and production, leading to a growing variety of electric models available to consumers across different vehicle segments.*
- *Despite the growing popularity of EVs, challenges such as limited driving range, longer charging times compared to refueling gasoline vehicles, and the need for infrastructure development, including charging stations, remain. However, ongoing research and development efforts aim to address these challenges and accelerate the transition to electric mobility.*

- *Overall, Electric Vehicles represent a transformative force in the automotive industry, offering the potential to reduce greenhouse gas emissions, enhance energy security, and reshape the way people and goods are transported.*

Market Overview

The Electric Vehicle (EV) market is experiencing a remarkable surge globally, emerging as a transformative force within the automotive industry. With a growing emphasis on sustainability and eco-consciousness, EVs have garnered significant attention as viable alternatives to traditional Internal Combustion Engine (ICE) vehicles. This overview delves into the key trends, challenges, and opportunities shaping the EV landscape across different regions.

Current Scenario in India:

- **Low Penetration:** *Despite the growing interest, EV penetration in India remains relatively low, accounting for only about 1% of total vehicle sales. However, there is substantial potential for expansion.*
- **Ambitious Government Targets:** *The Indian government has set ambitious targets to bolster EV adoption, aiming for 30% of private cars, 70% of commercial vehicles, and 80% of two- and three-wheelers to be electric by 2030. This signals a strong commitment to transitioning towards cleaner mobility solutions.*
- **Challenges and Hurdles:** *Several challenges hinder widespread EV adoption in India, including high upfront vehicle costs compared to ICE vehicles, inadequate charging infrastructure, and technological limitations associated with battery technology.*
- **Leading Models and Market Segments:** *Top-selling EV models in India include the Tata Nexon EV, Tata Tigor EV, MG ZS EV, and Hyundai Kona Electric. Notably, the market is currently dominated by two and three-wheeler EVs, spearheaded by companies like Hero Electric, Okinawa, and Ampere.*
- **Government Incentives:** *To incentivize EV adoption, the Indian government offers various subsidies and benefits under schemes like FAME (Faster Adoption and Manufacturing of Electric Vehicles), along with GST reductions and income tax benefits.*

Global Perspective:

- **China's Dominance:** China stands out as the global leader in the EV market, commanding over 40% of all new energy vehicle sales. Its robust infrastructure and supportive policies have propelled remarkable growth in EV adoption.
- **European Market Dynamics:** Europe has emerged as another major EV market, with countries like Norway, the Netherlands, Germany, and the UK leading the charge. Several European nations have announced ambitious plans to phase out ICE vehicles in the coming years.
- **Rising Trends in the US:** The United States is witnessing a surge in EV adoption, buoyed by Tesla's success and the introduction of more EV options from established automakers like GM and Ford. California remains a prominent hub for EV sales and innovation.
- **Global Automaker Initiatives:** Major automotive manufacturers worldwide are actively embracing electrification to comply with stringent emission regulations and cater to evolving consumer preferences. This trend underscores the industry's shift towards sustainable mobility solutions.
- **Tesla's Influence:** Tesla continues to exert significant influence in the high-end luxury EV segment globally, with its Model S, X, and 3 models commanding considerable market share and setting benchmarks for innovation.

Conclusion:

While India is still in the early stages of EV adoption compared to more mature markets like China and Europe, the government's ambitious targets and supportive policies, coupled with declining battery costs, signal promising growth prospects. Globally, the EV market is witnessing unprecedented momentum, driven by technological advancements, regulatory pressures, and evolving consumer preferences. As the industry continues to evolve, collaborations between governments, automakers, and stakeholders will play a crucial role in accelerating the transition towards a sustainable mobility future.

Problem Breakdown (Fermi Distribution)

- ***Demographic Segmentation Analysis:***
 - *In-depth examination of consumer demographics such as age, income, education, and household composition.*
 - *Understanding how these demographic factors influence electric vehicle purchasing decisions and preferences for specific vehicle features.*
- ***Vehicle Features Segmentation Study:***
 - *Detailed analysis of consumer preferences for various vehicle features including range, battery capacity, performance, technology integration (e.g., autonomous driving, connectivity), and design aesthetics.*
 - *Identifying key features that different consumer segments prioritize and their willingness to pay premiums for them.*
- ***Market Opportunities Identification:***
 - *Identifying market opportunities based on demographic and vehicle features segmentation.*
 - *Assessing potential target segments and their specific preferences to tailor products and marketing strategies accordingly.*
- ***Customized Marketing Strategies Development:***
 - *Developing customized marketing strategies tailored to specific demographic segments.*
 - *Crafting messaging and advertising campaigns that resonate with the unique preferences and needs of different consumer segments.*
- ***Product Development Insights:***

- Utilizing demographic insights to inform product development and innovation efforts.
- Identifying opportunities to enhance existing features or introduce new ones that align with the preferences of target demographic segments.
- **Overcoming Barriers to Adoption:**
 - Exploring strategies to overcome barriers to electric vehicle adoption among certain demographic groups.
 - Addressing concerns such as affordability, range anxiety, or lack of awareness through targeted initiatives.
- **Market Expansion Strategies:**
 - Investigating opportunities for market expansion by targeting underrepresented demographic segments or emerging consumer trends.
 - Developing strategies to tap into new market segments and drive overall industry growth.

By focusing on these areas of demographic and vehicle features segmentation, the report aims to provide actionable insights to stakeholders in the electric vehicle industry, facilitating informed decision-making and strategic planning.

Methodology

In this report we worked majorly in demographic and vehicle feature segmentation i.e. first 2 parts of problem breakdown and for this we used following datasets:

1) Datasets

For Behavioral Segmentation:

- *Dataset:* https://github.com/BairagiSaurabh/Electric-Vehicle-Market-Segmentation/blob/main/Behavioural%20Segmentation/behavioural_dataset.csv
- *The dataset comprises demographic and financial attributes of individuals used for behavioral segmentation analysis over more than 100 samples. It includes information on age, profession, marital status, education, number of dependents, personal loan status, total salary, and loan amount. This data allows for understanding consumer behavior regarding personal loans, identifying target segments, and devising tailored marketing strategies.*

Behavioural Segmentation

```
behaviour_path = "/content/behavioural_dataset.csv"
data = pd.read_csv(behaviour_path)
data
```

	Age	Profession	Marrital Status	Education	No of Dependents	Personal loan	Total Salary	Price
0	27	Salaried	Single	Post Graduate	0	Yes	800000	800000
1	35	Salaried	Married	Post Graduate	2	Yes	2000000	1000000
2	45	Business	Married	Graduate	4	Yes	1800000	1200000
3	41	Business	Married	Post Graduate	3	No	2200000	1200000
4	31	Salaried	Married	Post Graduate	2	Yes	2600000	1600000
...
94	27	Business	Single	Graduate	0	No	2400000	1600000
95	50	Salaried	Married	Post Graduate	3	No	5100000	1600000
96	51	Business	Married	Graduate	2	Yes	2200000	1100000
97	51	Salaried	Married	Post Graduate	2	No	4000000	1500000
98	51	Salaried	Married	Post Graduate	2	Yes	2200000	1100000

For Vehicle Specification Segmentation:

- **Dataset:** <https://www.kaggle.com/datasets/sjsumanth/electric-data-analysis>
- *The dataset provides details on various electric vehicle (EV) models, including brand, model name, performance metrics, range, charging capabilities, powertrain, plug type, body style, segment, seating capacity, and price in euros. It offers insights into EV specifications, aiding consumers in comparing and selecting vehicles based on their preferences and needs.*

Vehicle Segmentation (Vehicle Features)

```

okk_path = "/content/ElectricCarData_Clean.csv"
okk = pd.read_csv(okk_path)
okk

```

	Brand	Model	AccelSec	TopSpeed_KmH	Range_Km	Efficiency_WhKm	FastCharge_KmH	RapidCharge	PowerTrain	PlugType	BodyStyle	Segment	Seats	PriceEuro
0	Tesla	Model 3 Long Range Dual Motor	4.6	233	450	161	940	Yes	AWD	Type 2 CCS	Sedan	D	5	55480
1	Volkswagen	ID.3 Pure	10.0	160	270	167	250	Yes	RWD	Type 2 CCS	Hatchback	C	5	30000
2	Polestar	2	4.7	210	400	181	620	Yes	AWD	Type 2 CCS	Liftback	D	5	56440
3	BMW	iX3	6.8	180	360	206	560	Yes	RWD	Type 2 CCS	SUV	D	5	68040
4	Honda	e	9.5	145	170	168	190	Yes	RWD	Type 2 CCS	Hatchback	B	4	32997
...
98	Nissan	Ariya 63kWh	7.5	160	330	191	440	Yes	FWD	Type 2 CCS	Hatchback	C	5	45000
99	Audi	e-tron S Sportback 55 quattro	4.5	210	335	258	540	Yes	AWD	Type 2 CCS	SUV	E	5	96050
100	Nissan	Ariya e-4ORCE 63kWh	5.9	200	325	194	440	Yes	AWD	Type 2 CCS	Hatchback	C	5	50000
101	Nissan	Ariya e-4ORCE 87kWh Performance	5.1	200	375	232	450	Yes	AWD	Type 2 CCS	Hatchback	C	5	65000
102	Byton	M-Byte 95 kWh 2WD	7.5	190	400	238	480	Yes	AWD	Type 2 CCS	SUV	E	5	62000

103 rows x 14 columns

2) Data Preprocessing

- **Handling Categorical Variables:**

- Categorical variables like 'Profession', 'Marital Status', 'Education', and 'Car_Loan' were encoded using **LabelEncoder** from sklearn and used **pd.dummies** to perform one-hot encoding of nominal variables. This step converted categorical data into numerical format, making it suitable for ML techniques like KNN clustering etc.


```
[ ] cluster_data[:, 0] = cluster_data[:, 0].astype(float)
cluster_data[:, 4] = cluster_data[:, 4].astype(float)
cluster_data[:, 6] = cluster_data[:, 6].astype(float)
cluster_data[:, 7] = cluster_data[:, 7].astype(float)

[ ] cols = ['Profession', 'Marrital Status', 'Education', 'Car_Loan']
data[cols] = data[cols].apply(LabelEncoder().fit_transform)
data
```

	Age	Profession	Marrital Status	Education	No of Dependents	Car_Loan	Total Salary	EV_Price
0	-1.498630	1	1	1	-1.642313	1	-1.406760	-0.904843
1	-0.211304	1	0	1	-0.136859	1	-0.258937	-0.445579
2	1.397855	0	0	0	1.368594	1	-0.450240	0.013685
3	0.754191	0	0	1	0.615867	0	-0.067633	0.013685
4	-0.854967	1	0	1	-0.136859	1	0.314975	0.932213
...
94	-1.498630	0	1	0	-1.642313	0	0.123671	0.932213
95	2.202434	1	0	1	0.615867	0	2.706274	0.932213
96	2.363350	0	0	0	-0.136859	1	-0.067633	-0.215947
97	2.363350	1	0	1	-0.136859	0	1.654102	0.702581
98	2.363350	1	0	1	-0.136859	1	-0.067633	-0.215947

99 rows x 8 columns

- **Standardization of Continuous Variables:**

- Continuous variables such as 'Age', 'No of Dependents', 'Total Salary', and 'EV_Price' were standardized using **StandardScaler** from **sklearn**. Standardization scaled the features such that they have a mean of 0 and a standard deviation of 1, which improved the performance of machine learning algorithms to great scale.

- **Checking for Missing Values:**

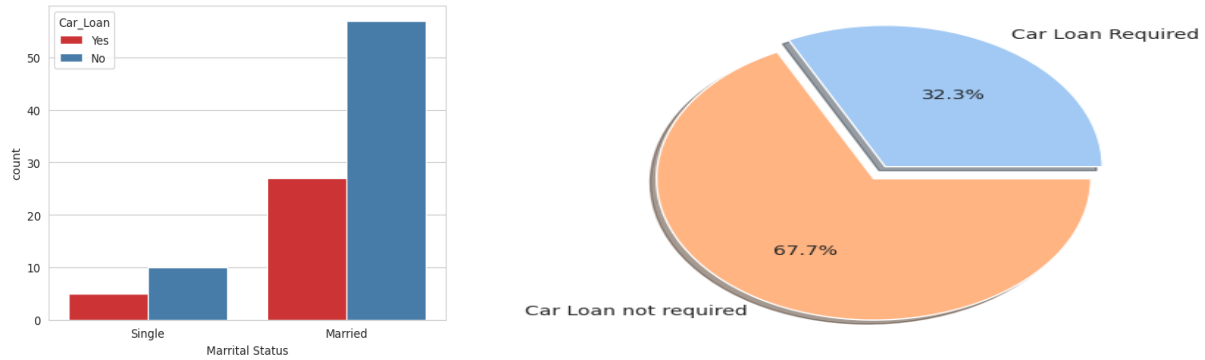
- Missing values were checked using **pd.isnull(data).sum()** to ensure that there are no missing values in the dataset. Dealing with missing values was crucial as they were adversely affecting the performance of machine learning models.

- **EDA Analysis:**

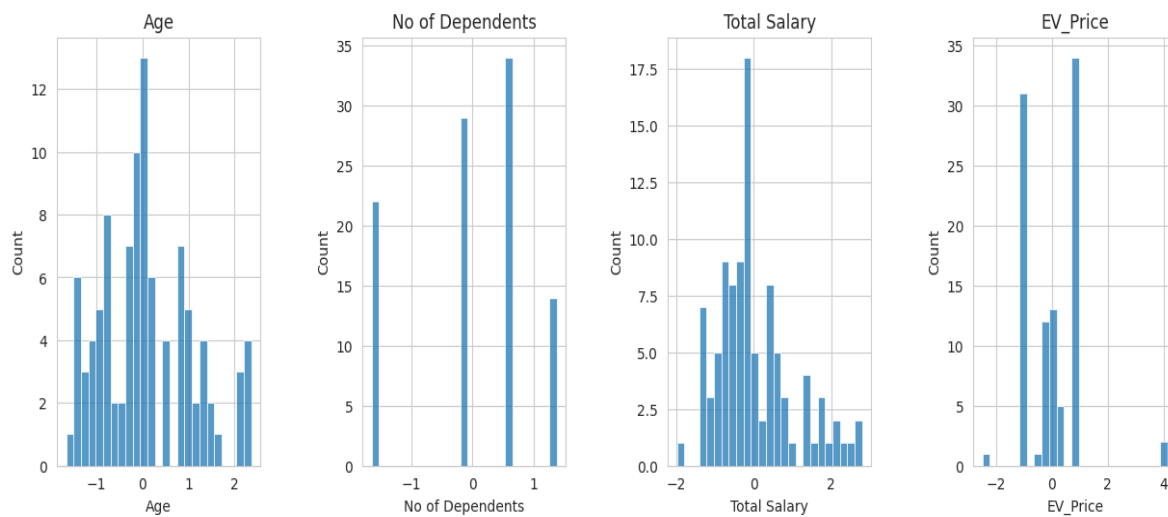
- Checked variation of continual variables and distribution of categorical variables i.e. whether they are normally distributed or skewed or have quite unbalanced distribution (SMOTE could be applied if present)

- Created a correlation map using heatmap to get intuitive understanding of how variables are dependent or varying w.r.t each other.

Behavioral Segmentation Features

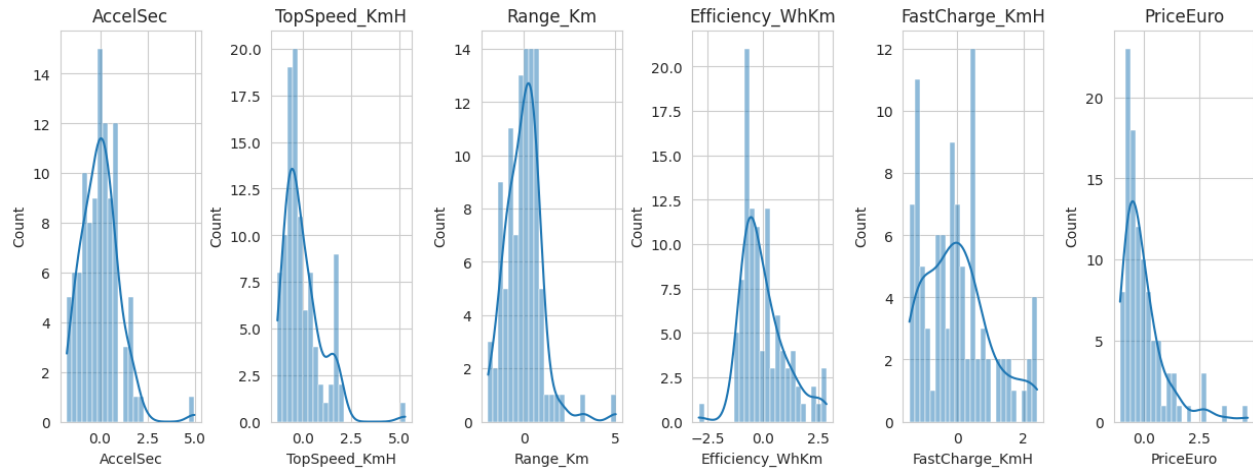


Categorical Variables Segmentation

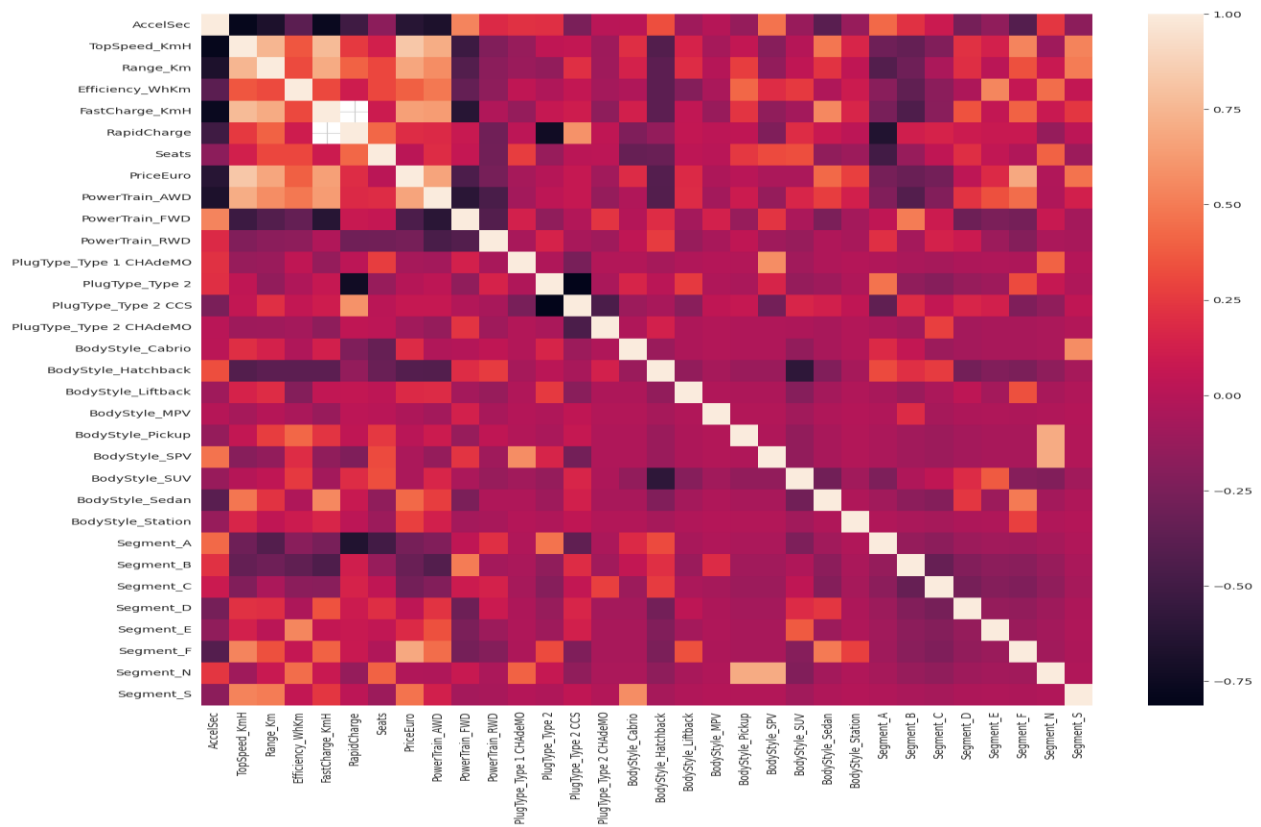


Continuous Variables Variation

2) Vehicular Segmentation



Continuous Variables Variation (Almost normal)

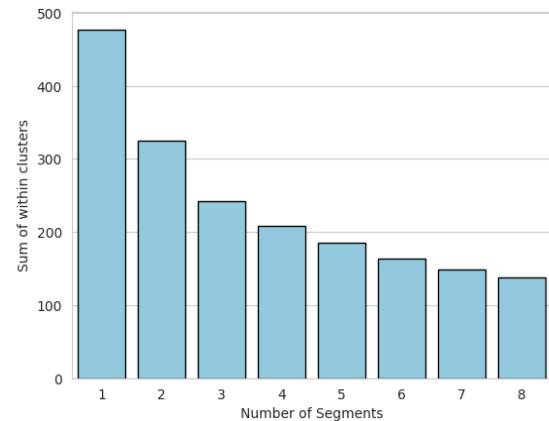
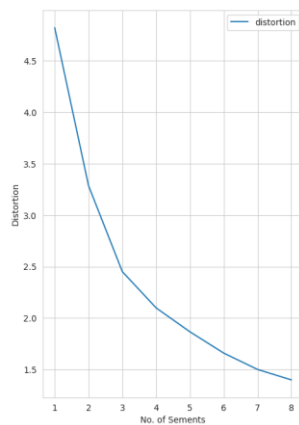
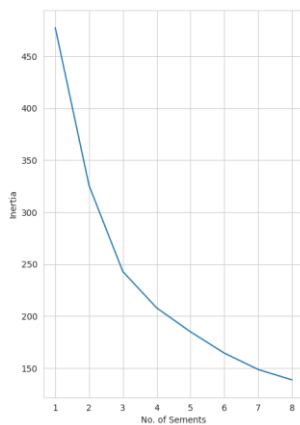


Correlation map for EV features Segmentation

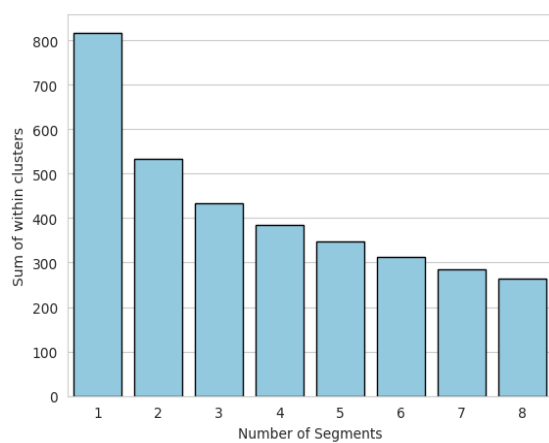
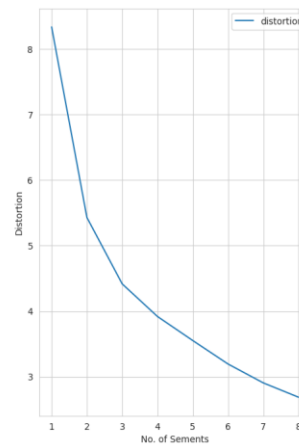
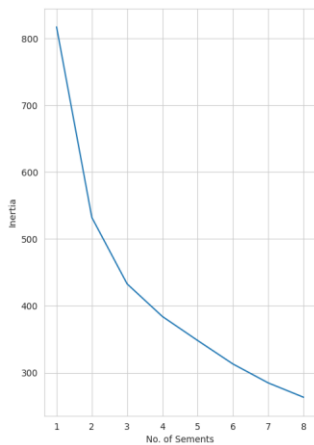
3) Segment Extraction (ML techniques used)

i) ML Techniques:

- *KNN – Clustering: K-means clustering was applied to the dataset with varying numbers of clusters, ranging from 1 to 9.*
- *The algorithm iteratively partitioned the data into 'k' clusters based on the similarity of data points to the cluster centroids*
- *Inertia and distortion values are calculated for each clustering solution ranging from 1 to 9 clusters. These values were plotted against the number of clusters. The plot allowed for the identification of the "elbow point," where the rate of decrease in inertia or distortion slows down significantly, indicating a potential optimal number of clusters.*
- *The "elbow" was observed at 3 clusters, suggesting that 3 clusters was the optimal choice for segmenting the data.*

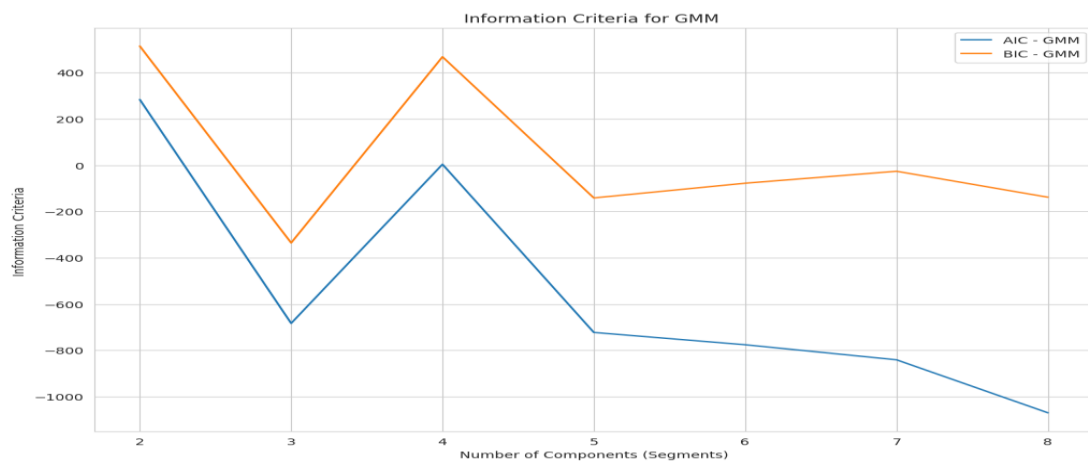


Behavioural Segmentation Clustering

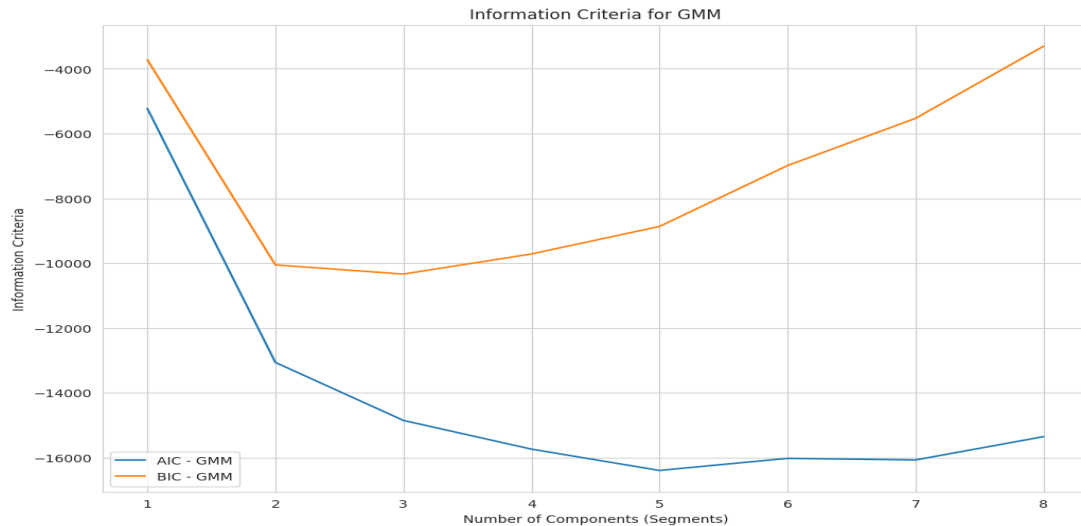


Vehicular Segmentation Clustering

- GMM was applied to the dataset with varying numbers of components (clusters), ranging from 1 to 9. The algorithm iteratively fitted the GMM to the data for each number of components and estimated the model parameters using an expectation-maximization (EM) algorithm.
- Information criteria, such as the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), were used to evaluate the goodness of fit of the GMM models. Lower values of AIC and BIC indicated better model fit, balancing the goodness of fit and the complexity of the model. AIC penalized the model complexity less severely, while BIC penalized it more rigorously.
- By examining the plot of AIC or BIC against the number of components, the "elbow point" was observed at 3 clusters in the plot of information criteria suggesting that 3 clusters may be the optimal choice for segmenting the data according to the GMM model.



- Behavioural Segmentation Gaussian Mixture

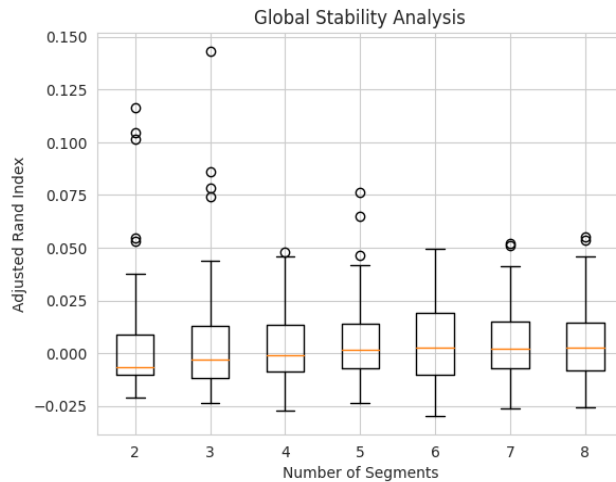


- Vehicular Segmentation Gaussian Mixture

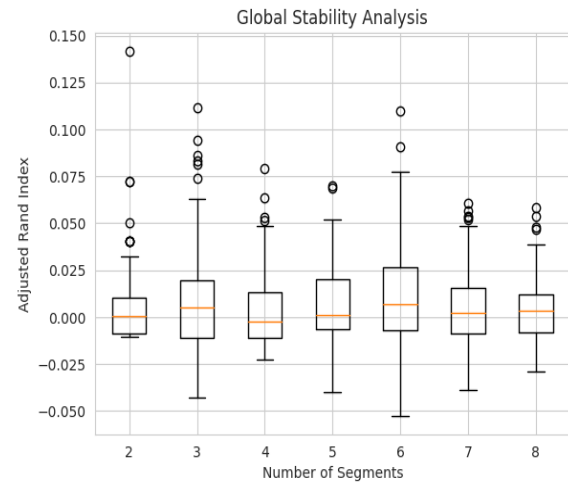
Profiling and analyzing potential segments

1) Global Stability Analysis

- Global stability analysis assessed the stability of clustering solutions across different random initializations or algorithms. It involved evaluating how consistent the clustering results were when the algorithm is applied multiple times to the same dataset with different starting conditions.
- One common metric used for global stability analysis was the adjusted Rand index (ARI). ARI measured the similarity between two clusterings, considering all pairs of samples and counting pairs that are assigned to the same or different clusters in the true and predicted clusterings.
- In the provided code, global stability analysis was performed by applying K-means clustering with varying numbers of clusters (from 1 to 9) multiple times ($n_{boot} = 100$ times) and calculating ARI for each run.
- The stability of clustering solutions decreased as the number of clusters increased but beyond 3 clusters this unstability pumped to considerable extent as shown in following box and whisker plot



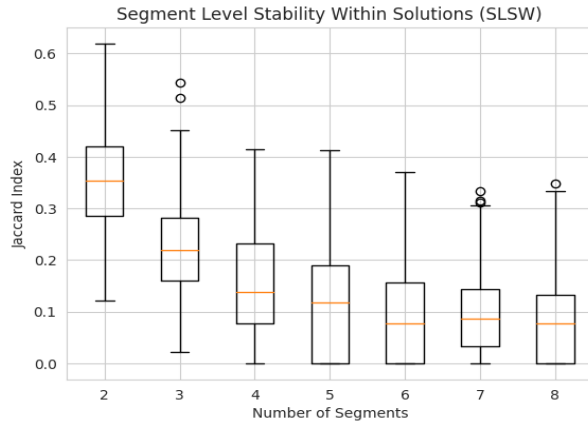
Behavioural Segmentation



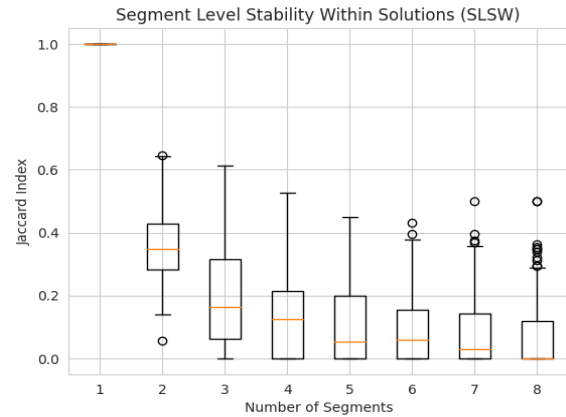
Vehicular Segmentation

2) **Stability Within Segments:**

- *Stability within segments assessed the stability of individual clusters or segments across different iterations or samples. It provided insights into the robustness of each cluster and helps in identifying stable and well-defined segments.*
- *In the provided code, stability within segments was evaluated using bootstrapping. For each bootstrap sample, K-means clustering was applied with a fixed number of clusters, and the resulting cluster labels were compared across different bootstraps to calculate stability scores.*
- *Jaccard index was used to measure the stability of individual clusters across bootstraps.*
- *Stability scores was visualized for each cluster, providing insights into the consistency of segment assignments like after 3 segments the stability scores dropped significantly.*



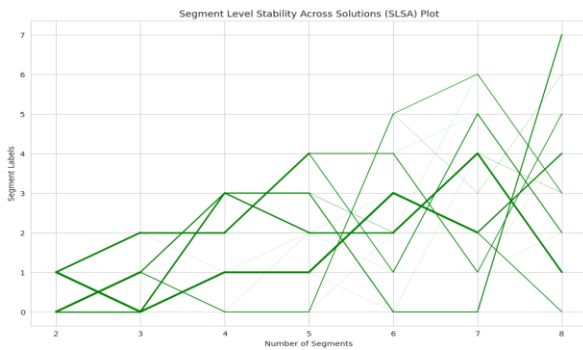
Behavioural Segmentation



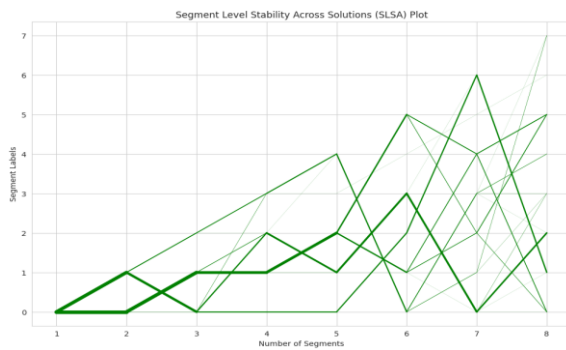
Vehicular Segmentation

3) **Stability Across Segments:**

- The SLSA plot compared the clustering results obtained from different numbers of clusters (1 to 8 in this case) to identify consistent and stable segments across solutions. It examined the stability of segments by tracking how cluster labels changed as the number of clusters varies.
- The plot helped in identifying segments that remain consistent across different clustering solutions. Stable segments were indicated by thick lines that maintain a consistent pattern like less division or partitions across the range of cluster solutions.
- Understanding the stability of these segments enhanced the interpretability and reliability of our clustering results, providing more confidence in the identified clusters and their characteristics.



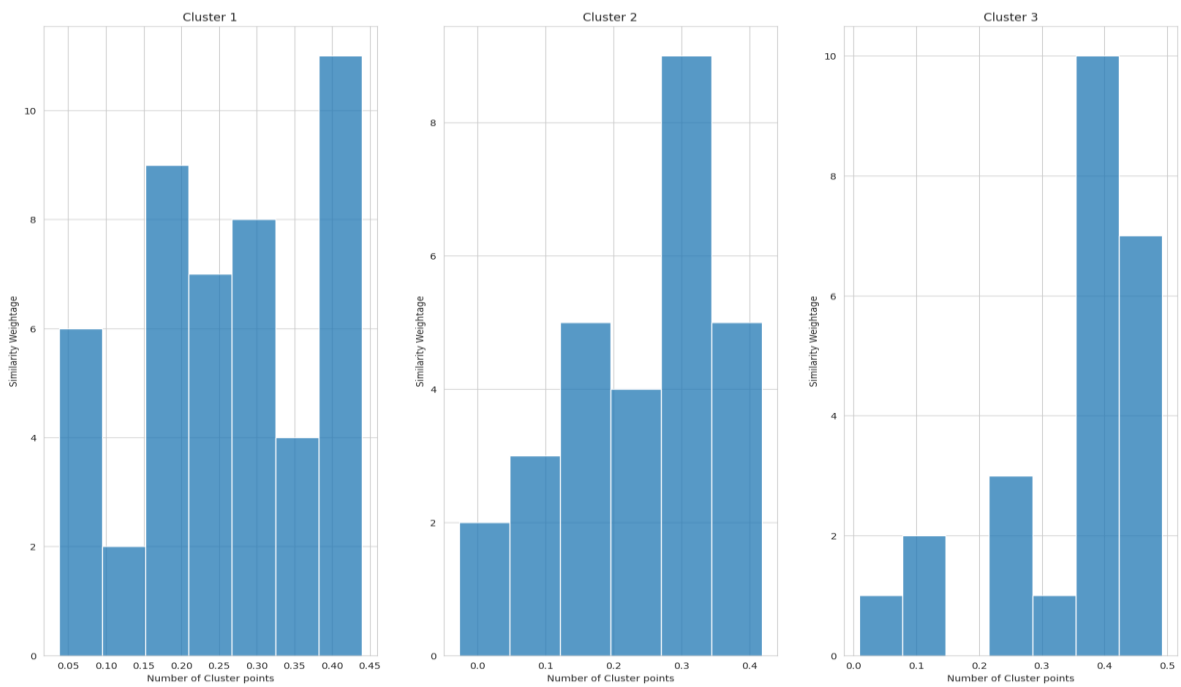
Behavioural Segmentation



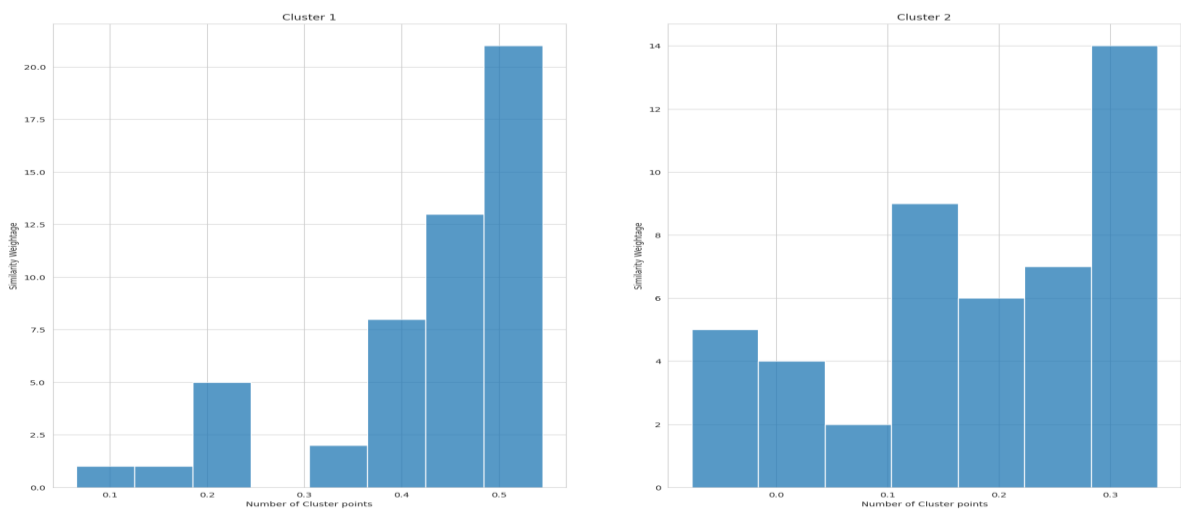
Vehicular Segmentation

4) Similarity weightage or cluster tightness using Silhouette Scores

The silhouette analysis provided valuable insights by assessing the quality and separation of clusters in the dataset. By computing silhouette scores for each sample, it helped evaluate how well-clustered the data points were within their respective clusters compared to neighboring clusters. This analysis offered a retrospective understanding of the effectiveness of the clustering algorithm and the appropriateness of the chosen number of clusters.



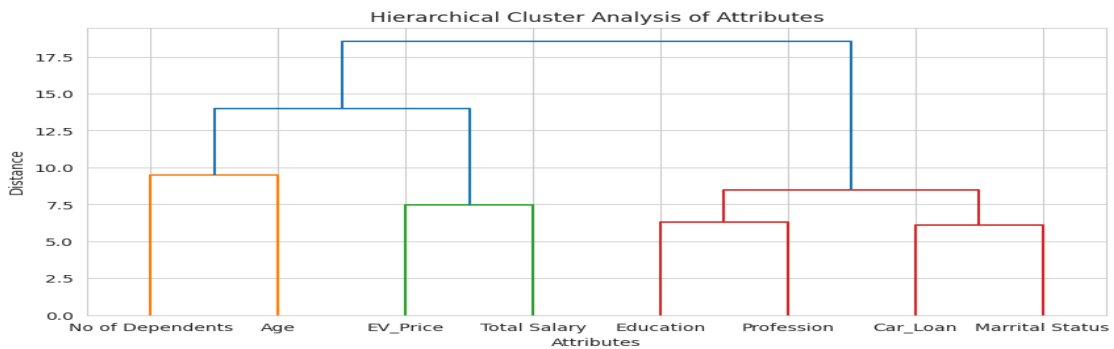
Behavioural Segmentation



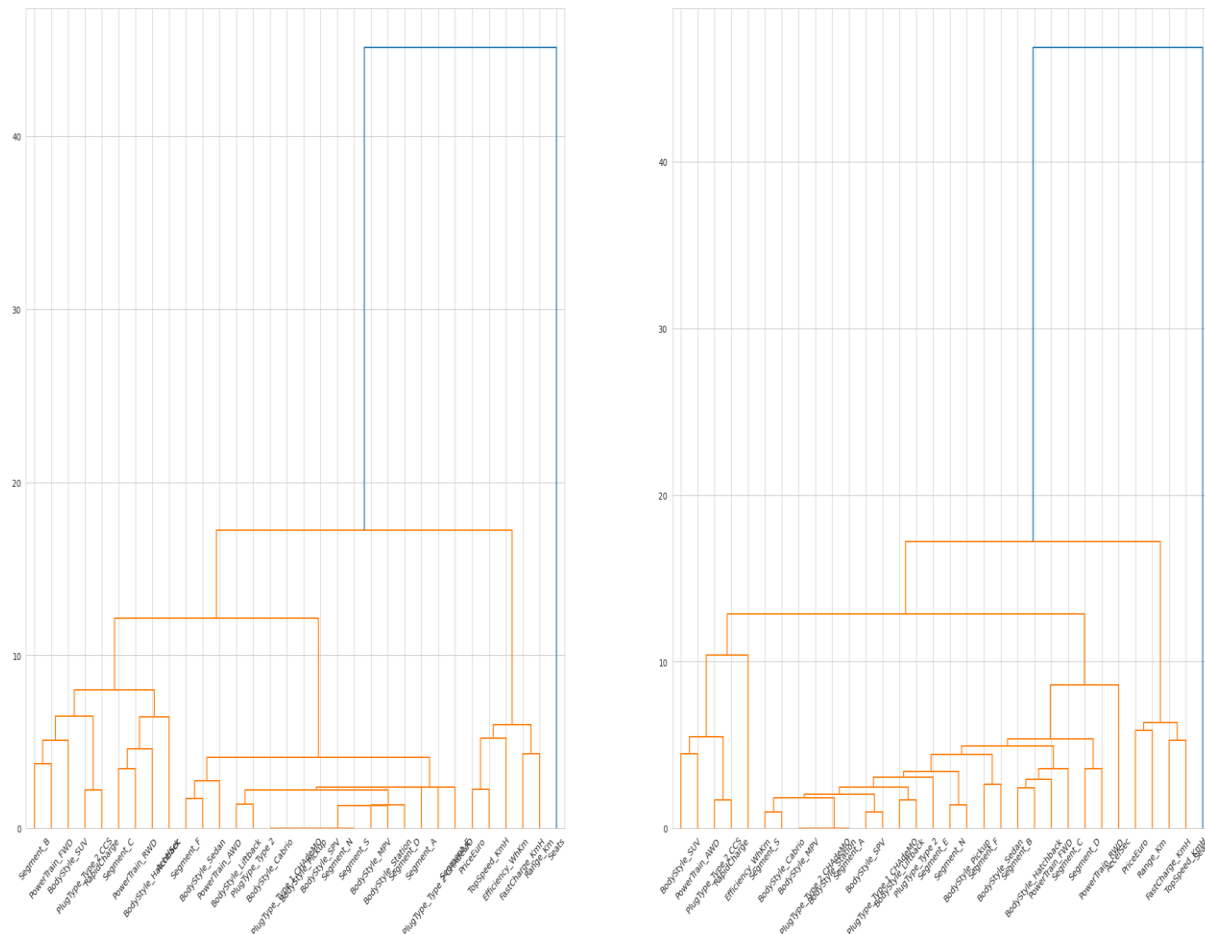
Segregation and Selection of target segment

1) Hierarchical Clustering

- Hierarchical clustering is a method of cluster analysis that builds a hierarchy of clusters. It starts by treating each data point as a separate cluster and then iteratively merges the closest clusters until all points belong to a single cluster. This results in a dendrogram, which visually represents the hierarchy of clusters and the relationships between them. This was done on an overall basis and within cluster basis.
- It was employed as it can help in segregating attributes or features into distinct groups based on their similarities or dissimilarities. This segregation aided in organizing and understanding the structure of the dataset.
- By examining the dendrogram, we can identify clusters of attributes that are closely related and may represent distinct segments or patterns within the data. This information could guide us in the selection of target segments for further analysis or targeted marketing strategies.



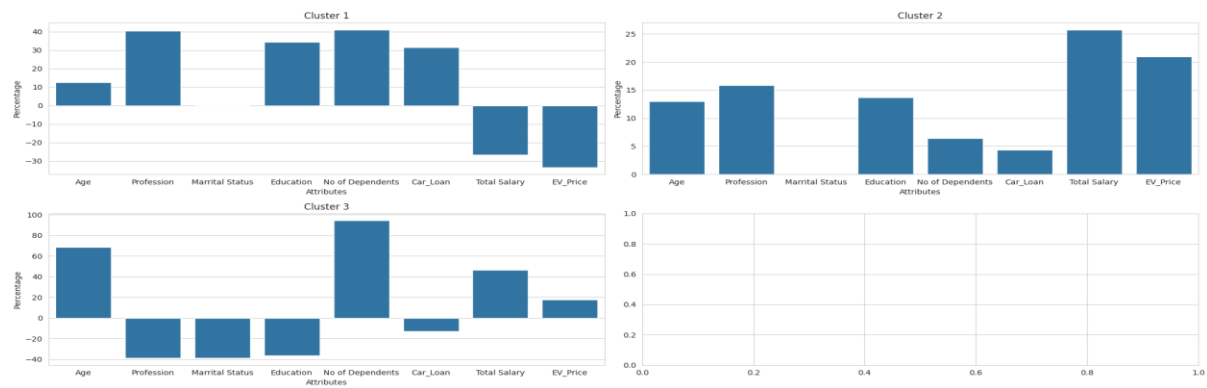
Behavioural Segmentation (Overall)



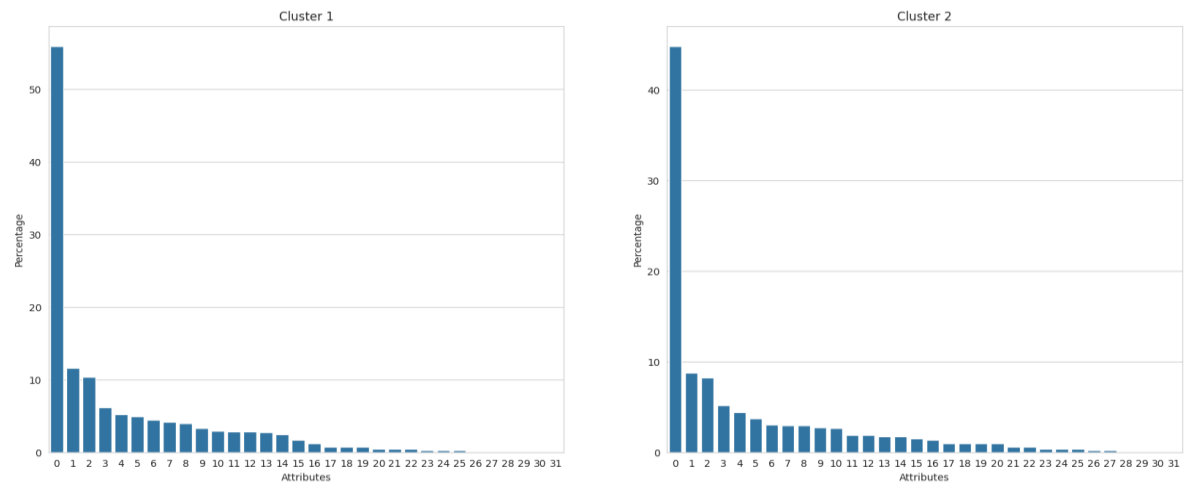
Vehicular Segmentation (Cluster Wise)

2) *Feature Weightage within selected Clusters.*

- *By analyzing the bar plots, we can identify which attributes have the highest percentage contribution within each cluster. Attributes with higher percentages were considered more important or influential in defining the characteristics of that cluster.*
- *Features with significant weightage differences across clusters were likely to be discriminative. Such features played a crucial role in distinguishing one cluster from another and capturing the underlying patterns or characteristics unique to each cluster.*
- *Focusing on features with higher weightage within clusters simplified the analysis by narrowing down the set of relevant attributes and then perform their selection.*

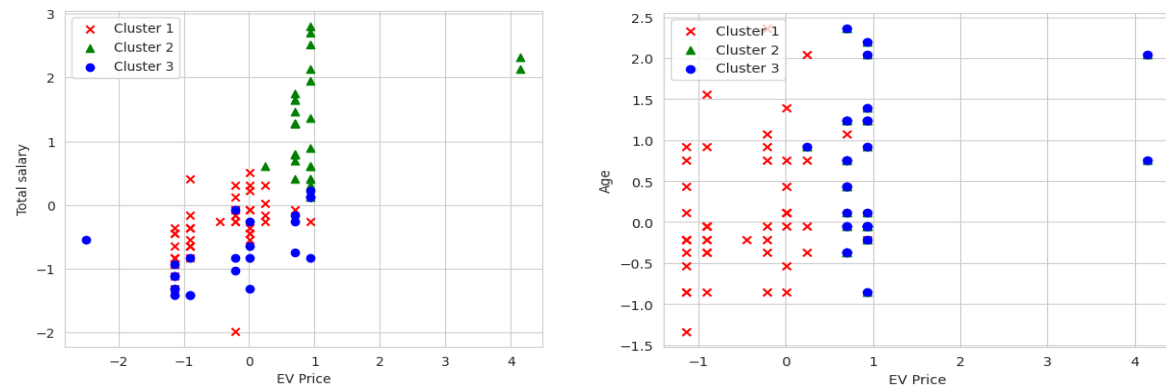


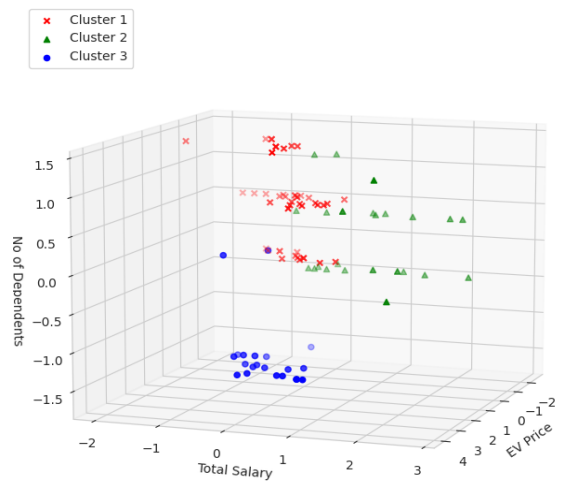
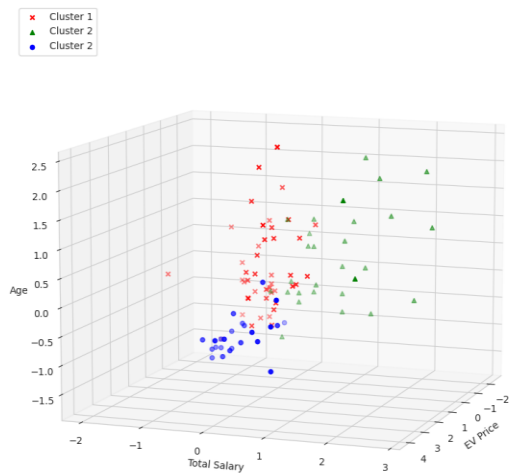
Behavioural Segmentation



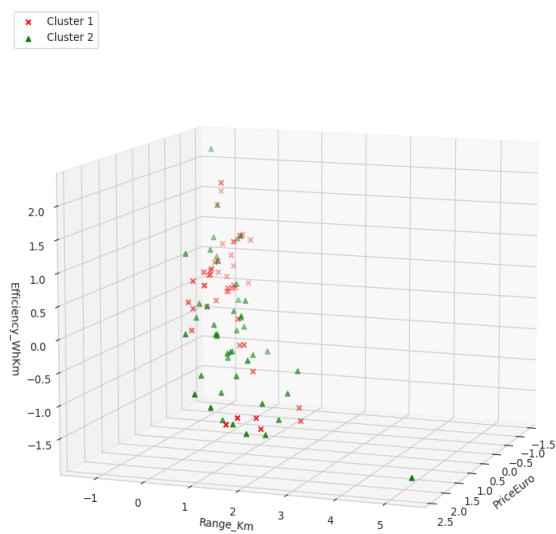
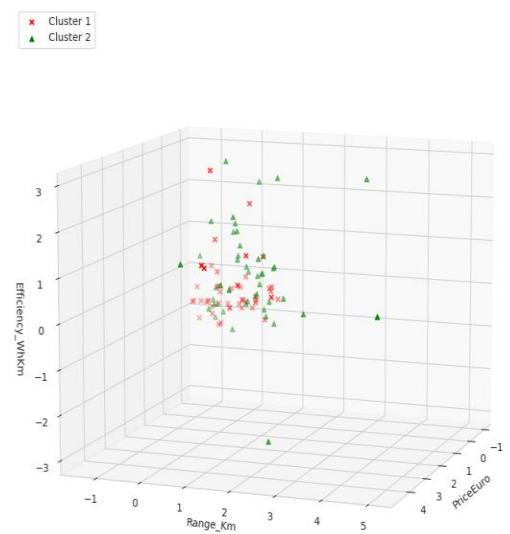
Vehicular Segmentation

3) Visualization of these targeted features





Behavioural Segmentation



Vehicular Segmentation

Results and Analysis

- *Based on demographic portfolios of various samples or individuals, we selected about 3 main segmentation areas, and for each of these segmented clusters, financial background, like salary and EV vehicle price, were dominated (in cluster 2), i.e. one USP for EV startup could play in the pricing domain and launching its product in different economic zones based on population average incomes.*
- *Similarly, the average family size and their age groups (based on cluster 3) can be other features on which potential segments can be created like in rural areas, mostly joint families and a wider age range of people resides so customized product can be designed for these areas while in tier-1 and tier-2 mostly nuclear families have less no. of dependent and their age group has narrow range so different product can be launched according to their needs.*
- *Third differentiator could be their profession or education, and their previous loan backgrounds like most working office people require compact, more average providing vehicles than other taxis taking driver which prefer spacious to take maximum fare and more fuel-efficient vehicle products. Similarly, their car-loans payments can categorize these segments based on their stringent nature of repaying in time.*
- *In terms of vehicle features, the startups should be focusing on their vehicle pricing which can segment their market by placing their product in premium or affordable market and then they must take distinctive features into account of those bifurcated market like how the average or efficiency of vehicle the market requires, or they prefer more speed or acceleration gaining in less time feature etc. like if startup want to launch sports EV vehicle then they should first target premium market and choose their pricing accordingly since these market prefer more range and more speed like fancy features which other affordable market domain don't prefer or priorities.*
- *Another could be their charging features like their powertrain, plug type, rapid charge options etc. since these are features which can give them an upper lift since it can enhance the appeal and competitiveness of electric vehicles. These features cater to different consumer preferences and usage scenarios like regular driving individuals or occasionally traveling to long range drivers etc.*

- *The range of body types of vehicles like sedan, SUV etc. the startup wants to launch can also categorize these segments since different body types have their distinct requirements or providing like their efficiency, speed or acceleration in unit time or charging features included and what segments the vehicle belong to like (segment D, C etc.) which are responsible for creating of quite distinct market category thus catering to different usage and requirement of customers.*

Business Perspective for EV Startup based on Performed Market Segmentation

A rough estimate analysis of potential market segments for an EV startup in India based on the information provided:

- *Potential customer base in the early market:*
 - *Urban metro cities (Tier 1): 200,000 customers*
 - *Other urban cities (Tier 2): 100,000 customers*
 - *Rural areas: 50,000 customers*

Total potential customer base: 350,000

- *Target price range:*
 - *Assuming the startup launches a compact sedan in the affordable segment (Rs. 10-15 lakhs), the target price range could be Rs. 12-14 lakhs.*

Potential profit in the early market:

- *Potential Customer Base: 350,000 Target Price Range: Rs. 12-14 lakhs
Average Selling Price: Rs. 13 lakhs Potential Profit = Potential Customer Base *
Average Selling Price = 350,000 * Rs. 13 lakhs = Rs. 4,550 Crores*

Most optimal market segments:

Based on the analysis of customer demographics, lifestyle patterns and vehicle preferences, the most optimal early market segments for an Indian EV startup would be:

- *Urban metro cities (Tier 1 young nuclear families prefer compact sedans/SUVs)*
- *Salaried professionals in other urban cities (Tier 2 - compact sedan segment)*
- *Joint families in rural areas (SUV/multipurpose vehicle catering to family needs)*

Targeting these segments allows maximizing sales potential while meeting distinct customer needs through customized offerings. It also allows testing products across urban and rural markets before mass scale up.

Implementation

- *Google Collab Notebook:*

<https://colab.research.google.com/drive/1xnZqZZjnFIrrkBpFfVAUICfyeccox-0d?usp=sharing>

- *GitHub Repository:*

https://github.com/SahilBarbade1203/Feynn_Labs/blob/main/EV_Vehicle_Market_Segmentation.ipynb