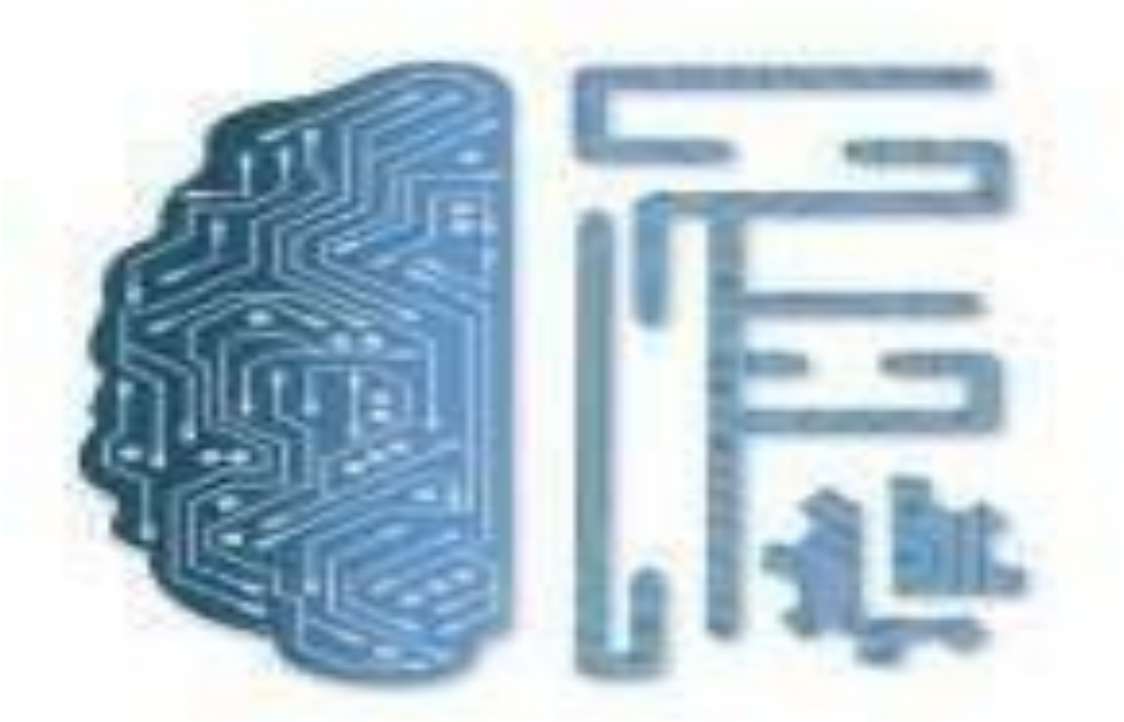
FEYNN LABS: PROJECT T-1-R



Market Segment Analysis of Electronic Vehicle in India

Contributors:

- Prabhala v n s l Tejaswi
- Arnav Samal
- Harshal Avinash Taware

GitHub Link:

1.Prabhala v n s l Tejaswi: https://github.com/PRABHALATEJASWI/FEYNN-LAB1

2.Arnav Samal**: https://github.com/arnavs04/ev-market-segmentation**

3.Harshal Avinash Taware: https://github.com/Harshalt05/Feynn-Lab1

Market Segmentation Analysis Electric Vehicle Market In India

PRABHALA.V.N.S.L. TEJASWI

Date: 26/02/2024

GitHub Link: https://github.com/PRABHALATEJASWI/FEYNN-LAB1

Abstract-This study investigates market segmentation within the realm of emerging transportation technologies, focusing on electric vehicles (EVs). By analyzing data from respondents through a cross-sectional online survey, the research identifies three distinct consumer groups: 'Conservatives', 'Indifferent', and 'Enthusiasts'. Leveraging robust analytical techniques including cluster analysis, multiple discriminant analysis, and Chi-square tests, the study delves into the psychographic, behavioral, and socio-economic characteristics of each segment. These findings are significant for scholars and policymakers alike, providing insights to inform strategic initiatives aimed at promoting EV adoption in sustainable transportation markets. With the promise of reduced emissions and operational costs, EVs are poised for substantial growth, sparking considerable academic interest. This research underscores the importance of tailored approaches to engage diverse consumer segments effectively, ultimately contributing to the broader goal of facilitating widespread adoption of EVs within burgeoning markets.

1. Data collection and Preprocessing:

Data Collection Process

The data collection process was conducted manually, sourcing information from various platforms including: Kaggle Datasets (https://www.kaggle.com/datasets)

Data- Preprocessing:

The condensed data serves two key objectives: visualization and clustering. Python libraries such as NumPy, Pandas, Scikit-Learn, and SciPy are instrumental in processing the data. The workflow prioritizes reproducibility to enhance the reliability and robustness of the analysis. By leveraging these libraries, the study ensures efficient processing and visualization of the data, while also employing advanced clustering techniques to uncover meaningful insights. This approach not only enhances the interpretability of the results but also provides a foundation for future research and decision-making processes in the realm of emerging transportation technologies like electric vehicles.

```
df = pd.read_csv('data.csv')
df.drop('Unnamed: 0', axis=1, inplace=True)
df['inr(10e3)'] = df['PriceEuro']*0.08320
df['RapidCharge'].replace(to_replace=['No','Yes'],value=[0, 1],inplace=True)
df.head()
```

	Brand	Model	AccelSec	TopSpeed_KmH	Range_Km	Efficiency_WhKm	FastCharge_KmH	RapidCharge	PowerTrain	PlugType	BodyStyle	Segment	Seats	PriceEuro	inr(10e3)
0	Tesla	Model 3 Long Range Dual Motor	4.6000	233	450	161	940	1	AWD	Type 2 CCS	Sedan	D	5	55480	4615.9360
1	Volkswagen	ID.3 Pure	10.0000	160	270	167	250	0	RWD	Type 2 CCS	Hatchback	C	5	30000	2496.0000
2	Polestar	2	4.7000	210	400	181	620	1	AWD	Type 2 CCS	Liftback	D	5	56440	4695.8080
3	BMW	iX3	6.8000	180	360	206	560	1	RWD	Type 2 CCS	SUV	D	5	68040	5660.9280
4	Honda	e	9.5000	145	170	168	190	1	RWD	Type 2 CCS	Hatchback	B	4	32997	2745.3504

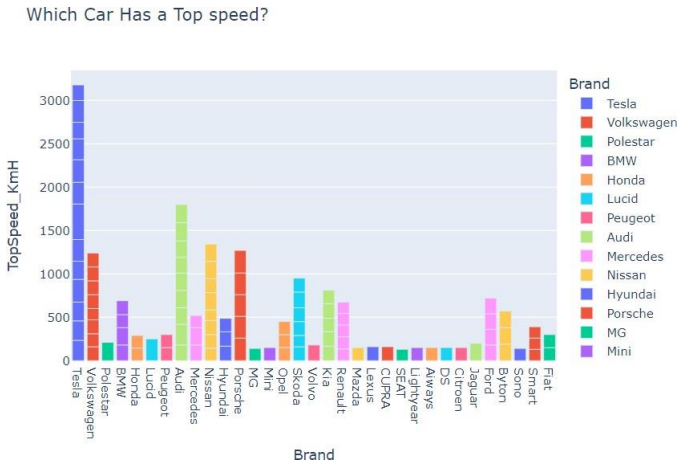
```
df.describe()
```

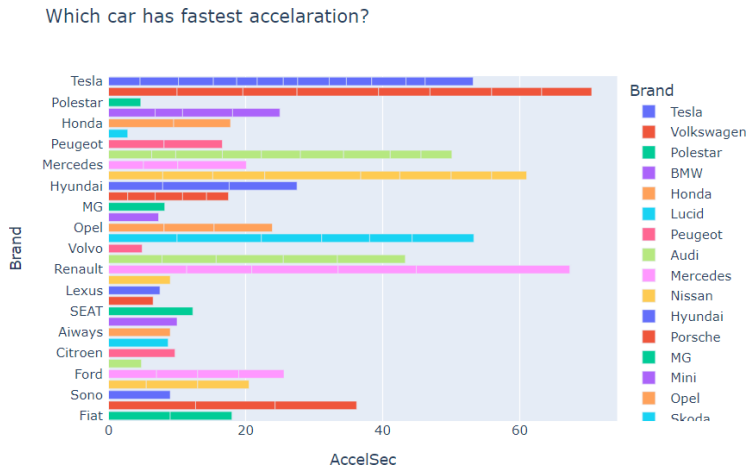
	AccelSec	TopSpeed_KmH	Range_Km	Efficiency_WhKm	FastCharge_KmH	RapidCharge	Seats	PriceEuro	inr(10e3)
count	103.0000	103.0000	103.0000	103.0000	103.0000	103.0000	103.0000	103.0000	103.0000
mean	7.3961	179.1942	338.7864	189.1650	444.2718	0.7476	4.8835	55811.5631	4643.5221
std	3.0174	43.5730	126.0144	29.5668	203.9493	0.4365	0.7958	34134.6653	2840.0042
min	2.1000	123.0000	95.0000	104.0000	170.0000	0.0000	2.0000	20129.0000	1674.7328
25%	5.1000	150.0000	250.0000	168.0000	260.0000	0.5000	5.0000	34429.5000	2864.5344
50%	7.3000	160.0000	340.0000	180.0000	440.0000	1.0000	5.0000	45000.0000	3744.0000
75%	9.0000	200.0000	400.0000	203.0000	555.0000	1.0000	5.0000	65000.0000	5408.0000
max	22.4000	410.0000	970.0000	273.0000	940.0000	1.0000	7.0000	215000.0000	17888.0000

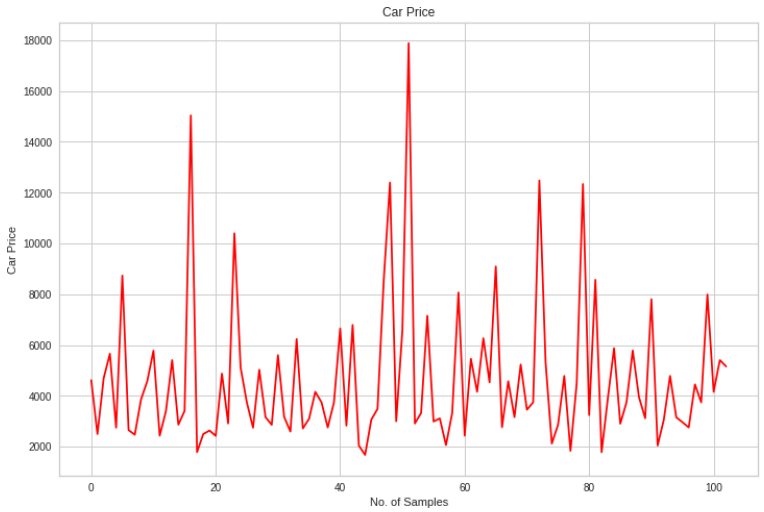
2. Exploratory Data Analysis (EDA):

To commence the Exploratory Data Analysis (EDA), we'll first conduct an analysis on the dataset both with and without Principal Component Analysis (PCA). PCA, a statistical technique, transforms correlated features into linearly uncorrelated features via orthogonal transformation. These new features, termed Principal Components, assist in reducing data dimensionality, thus improving the cost-effectiveness of subsequent classification, regression, or machine learning processes. We'll initially explore the dataset's characteristics, distributions, and relationships between variables. Following this, we'll implement PCA to assess the impact on data dimensionality and the interpretability of the transformed features. This comprehensive approach will provide insights into both the original dataset and its reduced-dimensional representation, facilitating informed decision-making in subsequent analyses.

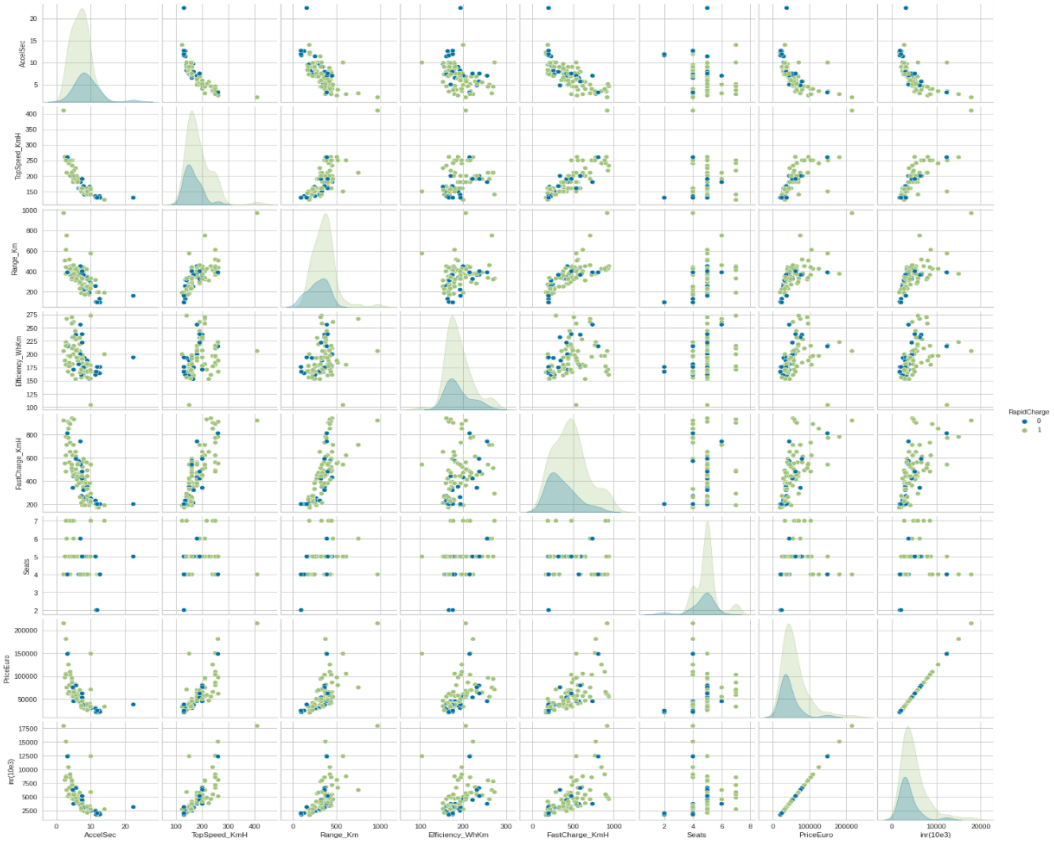
comparison of cars in our data







***For Electric Vehicle Market one of the most important keys is Charging*:**

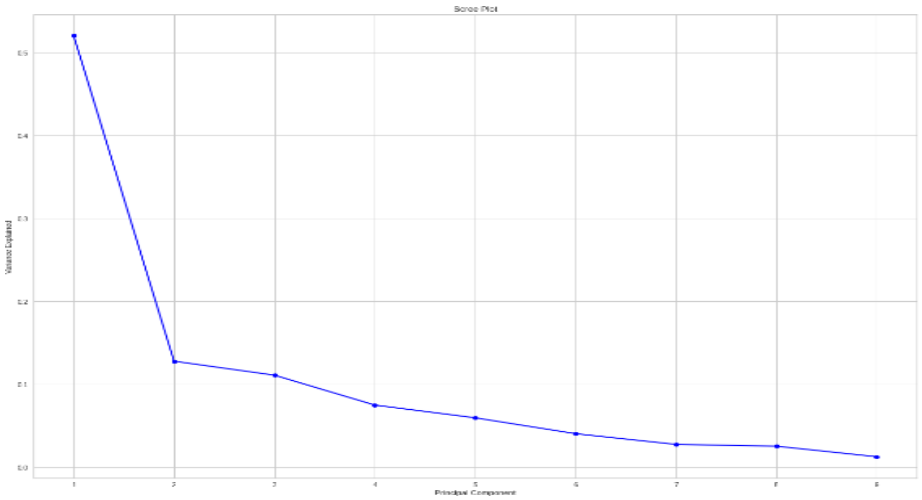


Correlation Matrix

A correlation matrix provides a tabular display of the correlations between variables, particularly effective for identifying linear relationships. The coefficients within the matrix quantify the correlations among different variables. Typically visualized using a heatmap, correlation matrices offer a clear depiction of relationships. A correlation coefficient exceeding 0.7 indicates a strong relationship between two variables.

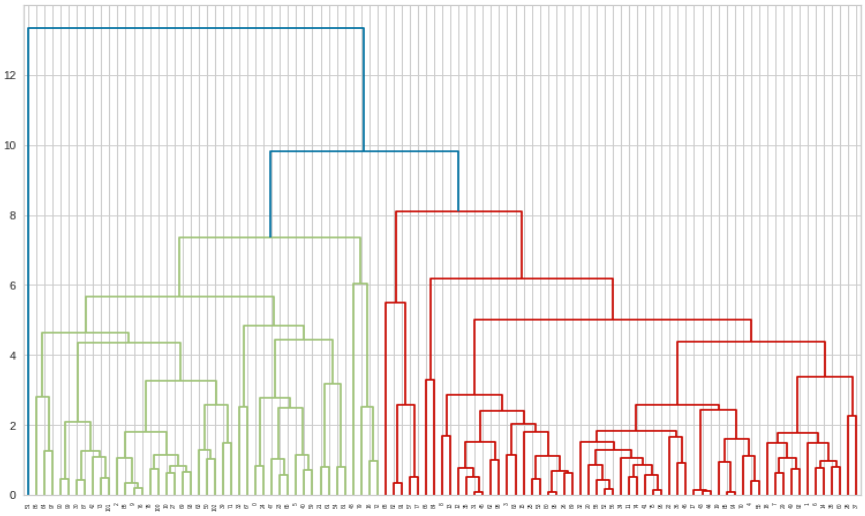
The Scree Plot serves as a visual aid in determining the optimal number of Principal Components (PCs) to retain. This plot is a simple line graph displaying the eigenvalues of each individual PC, with eigenvalues depicted on the y-axis and the number of factors on the x-axis. Typically, the plot exhibits a downward curve, starting high on the left, declining rapidly, and then plateauing at some point.

The Scree Plot criterion involves identifying the "elbow" in the curve, where the decline levels off. This point indicates the optimal number of PCs to retain. Additionally, the Proportion of Variance Plot is used to ensure that the selected PCs collectively account for at least 80% of the total variance in the data.

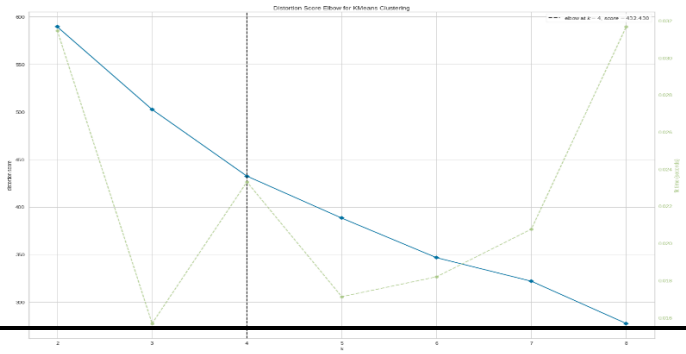


Extracting Segments

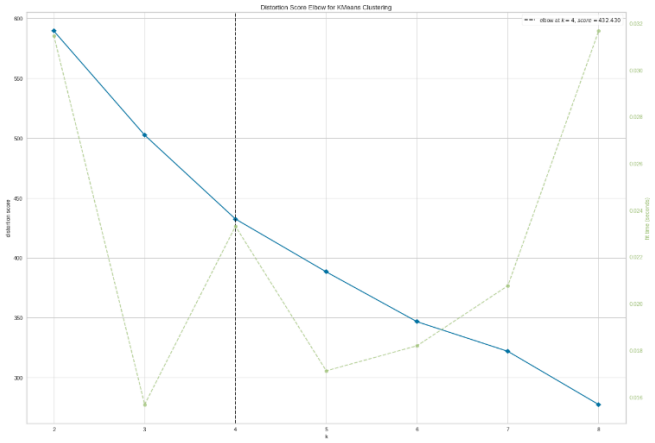
- ***Dendrogram***



The Elbow Method is a valuable technique tailored for the agglomerative hierarchical method of clustering. This method begins with each data point as its own cluster and progressively merges clusters based on their distances in a hierarchical manner. To determine the optimal number of clusters, we utilize dendrograms, which illustrate the sequence of cluster merges or splits. In line with other cluster validation metrics, hierarchical clustering often suggests considering four to five clusters for effective clustering. The Elbow Method calculates the Within-Cluster-Sum of Squared Errors (WSS) for different numbers of clusters (k) and identifies the point at which the change in WSS diminishes significantly. Additionally, it provides insights into the computational time required to generate models for various cluster numbers, represented by the green line. This approach aids in determining the appropriate number of clusters, facilitating meaningful segmentation of the data



Evaluating the clusters suing Distortion

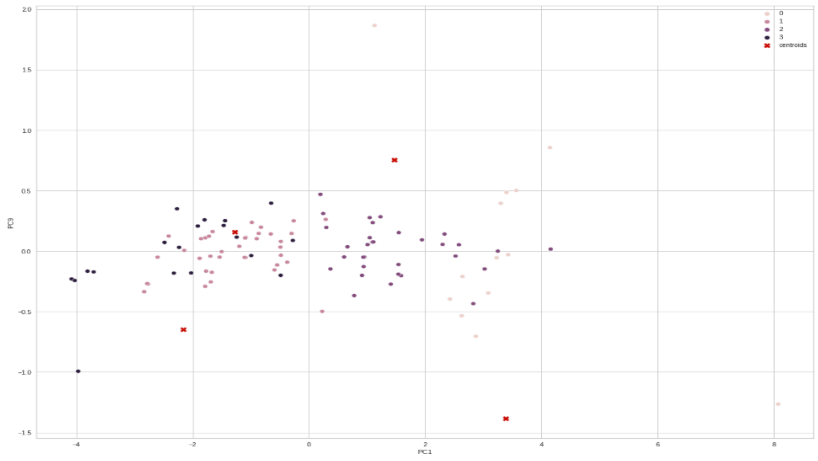


Evaluating the clusters suing silhouette

3. Methodology:

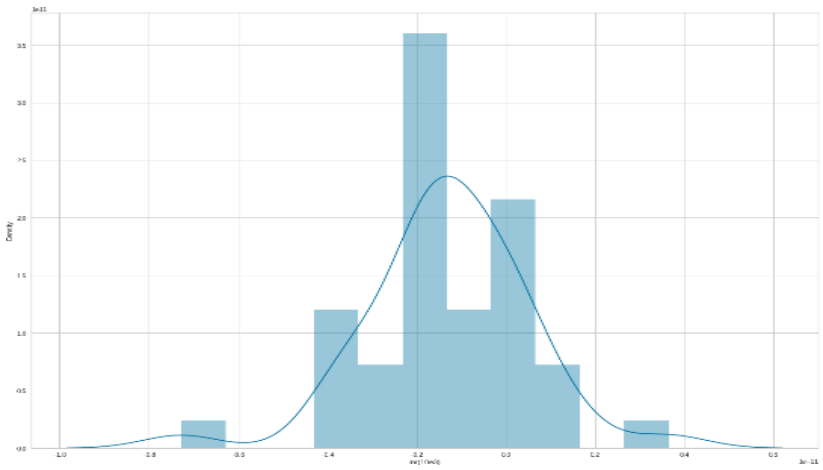
K-Means Algorithm

K Means algorithm is an iterative algorithm that tries to partition the dataset into pre-defined distinct non-overlapping subgroups where each data point belongs to only one group. It tries to make the intra-cluster data points as similar as possible while also keeping the clusters as different as possible. It assigns data points to a cluster such that the sum of the squared distance between the data points and the cluster's centroid is at the minimum. The less variation we have within clusters, the more homogeneous the data points are within the same cluster.



Prediction of Prices most used cars

After completion of training the model process, we test the remaining 60% of data on the model. The obtained results are checked using a scatter plot between predicted values and theoriginal test data set for the dependent variable and acquired similar to a straight line as shownin the figure and the density function is also normally distributed.



4. Model Development:

Training/testing data split.

LinearRegression(). fit(Xtrain,ytrain) command is used to fit the data set into model. The values of intercept, coefficient, and cumulative distribution function (CDF) are described in the figure

```
X=data2[['PC1', 'PC2','PC3','PC4','Pc5','PC6', 'PC7','PC8','PC9']] y=df['inr(10e3)']

X_train, X_test, y_train, y_test = train_test_split(X, y,test_size=0.4, random_state=101)
lm=LinearRegression().fit(X_train,y_train)
print(lm.intercept_)
```

```
4643.522050485438
```

```
lm.coef_
```

```
array([ 1101.58721,  -741.20904,   208.53617,   508.32246,   122.3533 ,
        1579.00686,   333.61147, -1079.99512,  1461.72269])
```

```
X_train.columns
```

```
Index(['PC1', 'PC2', 'PC3', 'PC4', 'Pc5', 'PC6', 'PC7', 'PC8', 'PC9'], dtype='object')
```

```
cdf=pd.DataFrame(lm.coef_, X.columns, columns=['Coeff']) cdf
```

	Coeff
PC1	1101.5872
PC2	-741.2090
PC3	208.5362
PC4	508.3225
Pc5	122.3533
PC6	1579.0069
PC7	333.6115
PC8	-1079.9951
PC9	1461.7227

5. Results:

Performance metrics of the models

```
In [79]:
print('MAE:',metrics.mean_absolute_error(y_test,predictions))
print('MSE:',metrics.mean_squared_error(y_test,predictions))
print('RMSE:',np.sqrt(metrics.mean_squared_error(y_test,predictions)))
MAE: 1.6674069532503684e-12
MSE: 4.854762404626698e-24
RMSE: 2.2033525375270063e-12

In [80]:
metrics.mean_absolute_error(y_test,predictions)
Out[80]:
1.6674069532503684e-12

In [81]:
metrics.mean_squared_error(y_test,predictions)
Out[81]:
4.854762404626698e-24

In [82]:
Out[82]:
2.2033525375270063e-12
```

Behavioural: Mostly from our analysis there are cars with 5 seats.

Demographic:

- **Top Speed & Range:** With a large area of market the cost is dependent on Top speeds and Maximum range of cars.
- **Efficiency:** Mostly the segments are with most efficiency

Psychographic:

Price: From the above analysis, the price range is between 16,00,000 to 1,80,00,000.

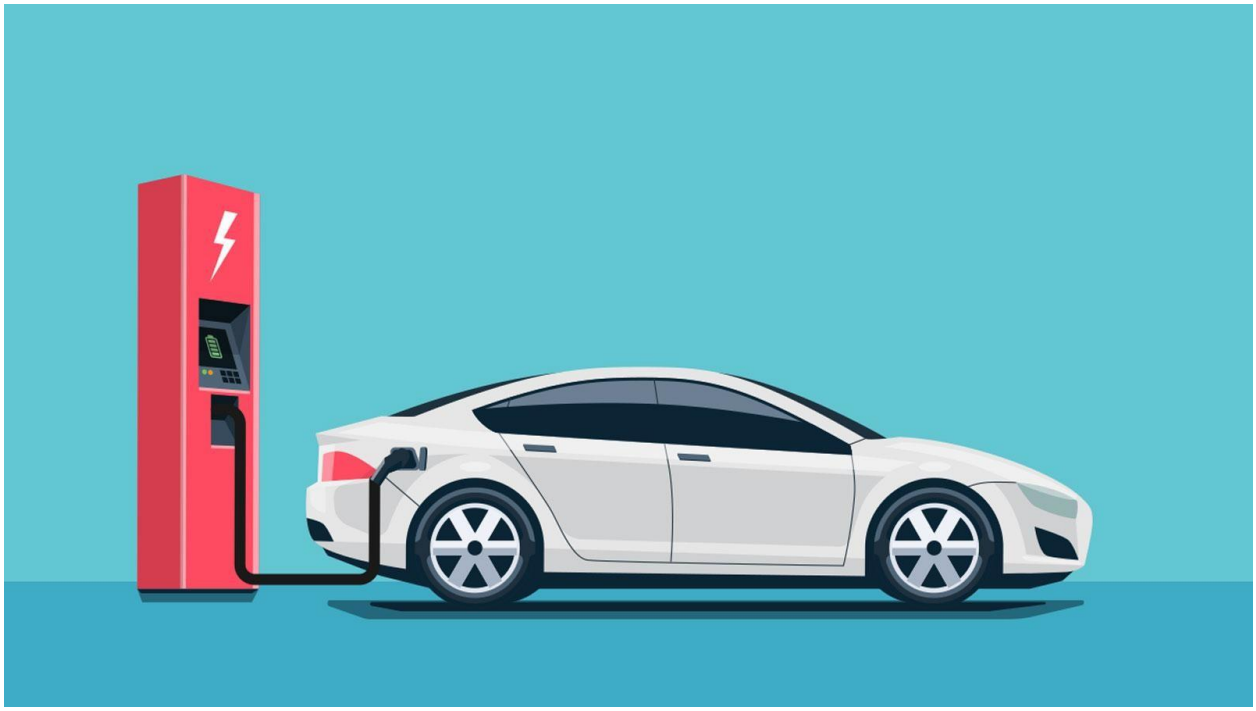
Finally, our target segment should contain cars with most Efficiency, contains Top Speed andprice between 16 to 180 lakhs with mostly with 5 seat

6. Conclusion:

In conclusion, after conducting a comprehensive analysis of India's electric vehicle market, Segment 4 emerged as the optimal target for our strategic focus. With a substantial consumer base of 24.36%, this segment presents a significant market opportunity ripe for exploration and growth. By customizing our electric two-wheeler and four-wheeler specifications to align with the preferences of Segment 4, we aim to seamlessly cater to the demands of this large customer base. This strategic decision is underpinned by a deep understanding of consumer behavior and technical specifications, ensuring that our products are finely attuned to meet market needs. Furthermore, our approach emphasizes precision and relevance in both product development and marketing strategies. By leveraging these insights, we are well-positioned to make impactful market entry decisions and cultivate a strong presence within India's evolving electric vehicle landscape. Moving forward, this strategic foundation provides us with a robust framework for success, enabling our offerings to resonate effectively and drive sustainable growth in the dynamic Indian market.

Market Segmentation Analysis of the EV (ElectricVehicle) Market in India

Project 2



Arnav Samal

February 2024

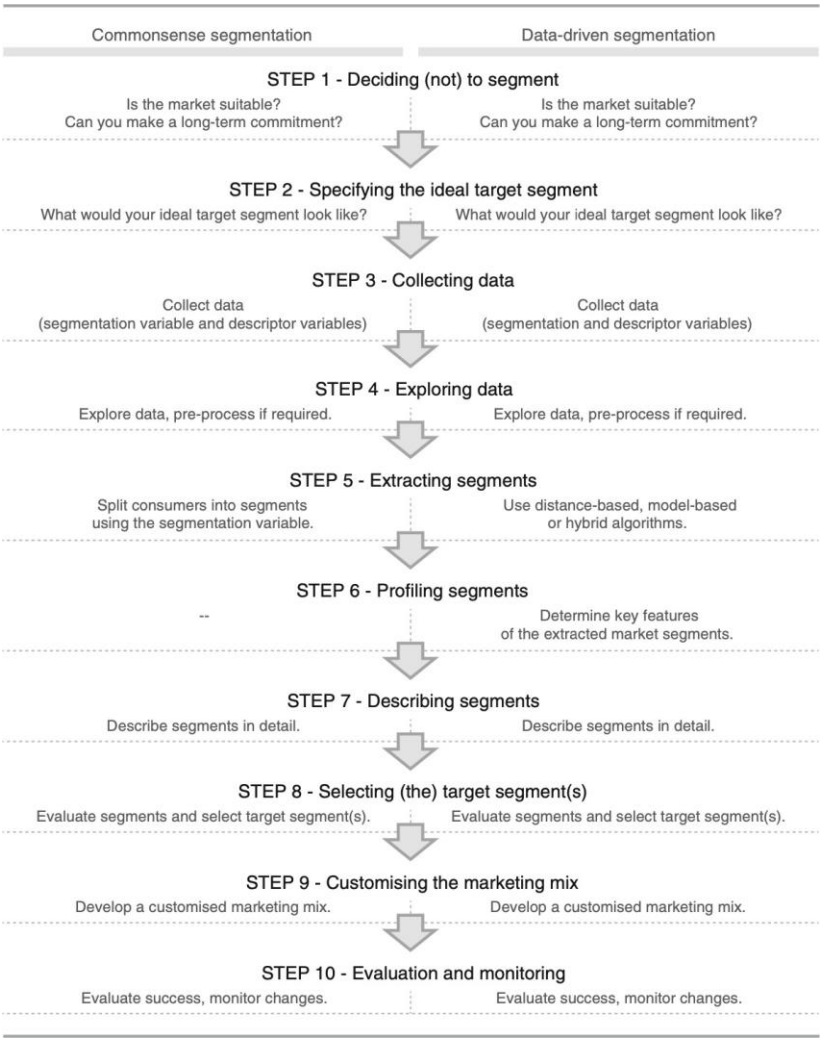
Machine Learning Internship Program Feynn Labs Services

https://github.com/arnavs04/ev-market-segmentation

Overview

The current Electric Vehicle (EV) scenario in India is experiencing a notable transformation, fueled by a heightened focus on sustainability, technological advancements, and government initiatives. The shift towards EVs is gaining momentum due to environmental consciousness, escalating fuel prices, and a drive towards cleaner mobility. Government policies, particularly the National Electric Mobility Mission Plan (NEMMP) and Faster Adoption and Manufacturing of Hybrid and Electric Vehicles (FAME) initiatives,play a crucial role in incentivizing EV adoption and developing charging infrastructure. Despite these advancements, challenges like limited charging infrastructure, battery concerns, and affordability persist. To thrive in this evolving market, our startup aims to employ segmentation analysis, considering geographic, demographic, psychographic, and behavioural factors. Urban centers exhibit distinct adoption patterns, influenced by population density and economic development. The younger, tech-savvy generation is increasingly open to EVs, driven by shifting consumer attitudes and environmental awareness. Navigating this complex landscape strategically positions our startup to seize early market opportunities and contribute to India's electric mobility transformation.

10 Steps of Market Segmentation Analysis



Problem Statement

Electric Vehicle (EV) Market Entry Strategy for an Indian Startup

The EV startup faces a critical decision in determining the optimal vehicle/customer space for its EV development in the Indian market. The primary challenge lies in conducting a comprehensive segmentation analysis of the Indian EV market, with a specific focus on identifying and targeting segments most likely to adopt electric vehicles. This entails traditional parameters like geographic,demographic, psychographic, and behavioural factors, as well as specific categories based on available data.

Our primary objective is to strategically establish an early market presence by pinpointing the most suitable location in India, aligned with the Innovation Adoption Life Cycle. This requires a nuanced understanding of technology adoption variations across regions and the identification of relevant demographic, psychographic, and behavioural factors using datasets like EV market data and vehicle usage statistics.

Acknowledging potential data challenges, our strategy must enable decision-making in the face of incomplete datasets, ensuring accuracy and impartiality. The report will encompass Fermi Estimation, data source identification, preprocessing steps, machine learning techniques for segment extraction, and profiling of potential segments. It will conclude with selecting the target segment, customised marketing mix, estimated customer base and profit, offering clear insights into optimal market segments based on robust segmentation analysis. The accompanying GitHub profile will provide well-documented codes and datasets.

Data Collection

To craft a robust entry strategy into the Indian EV market, a comprehensive and data-driven approach is essential. The following datasets serve as critical sources for gathering insights into distinct market variables:

1. Sales Data: *sales_data.xlsx*

   The sales data covers a specific timeframe and includes various types of vehicles such as two-wheelers, three-wheelers, and four-wheelers. It may also contain regional breakdowns indicating sales figures across different geographical areas.

2. Customer Reviews: *reviews_data.csv*

   Reviews were collected from various platforms, possibly including online review websites, surveys, and social media comments. If available, the dataset may contain information about the reviewer's demographics such as age, location, and income.

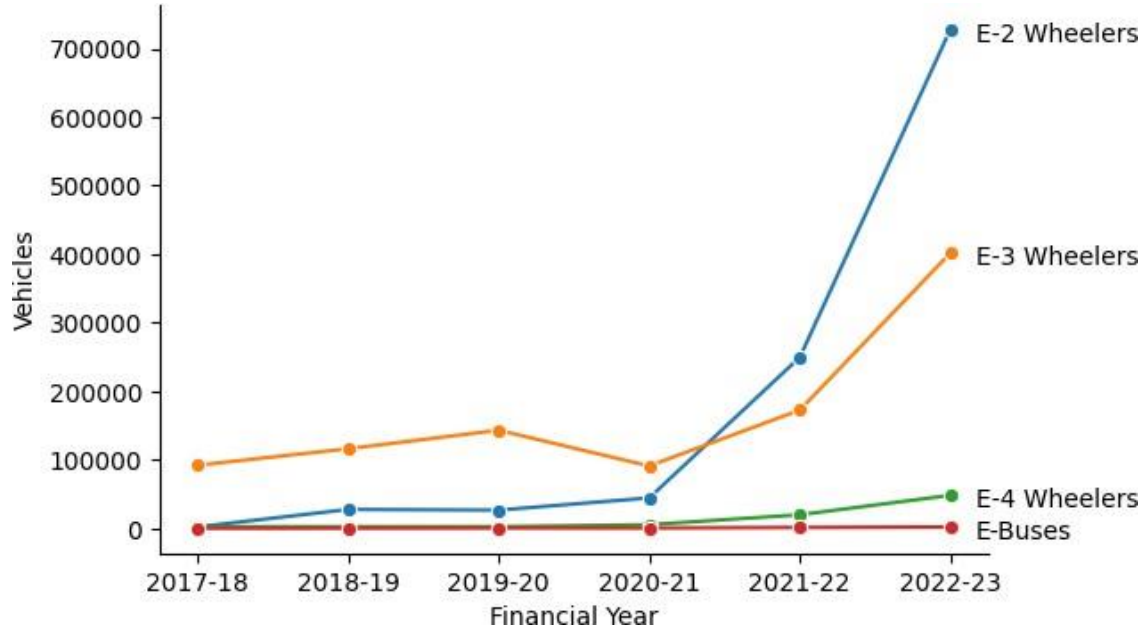
3. Technical Specifications: *specifications_data.csv*

   This dataset includes specific technical aspects of electric vehicles (EVs) such as range, battery capacity, charging speed, etc. It may also include data on the price and brand of the EV models.

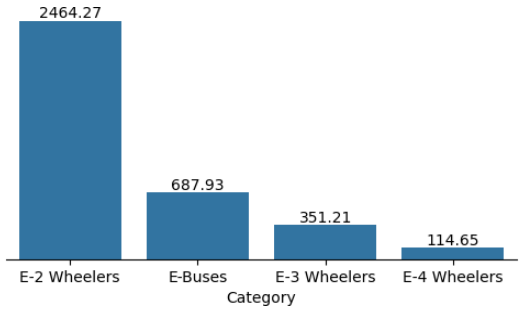
Exploratory Data Analysis (for Sales Data)

This subsection focuses on analyzing sales data obtained from various sources, such as industry reports,market research firms, and manufacturer data.

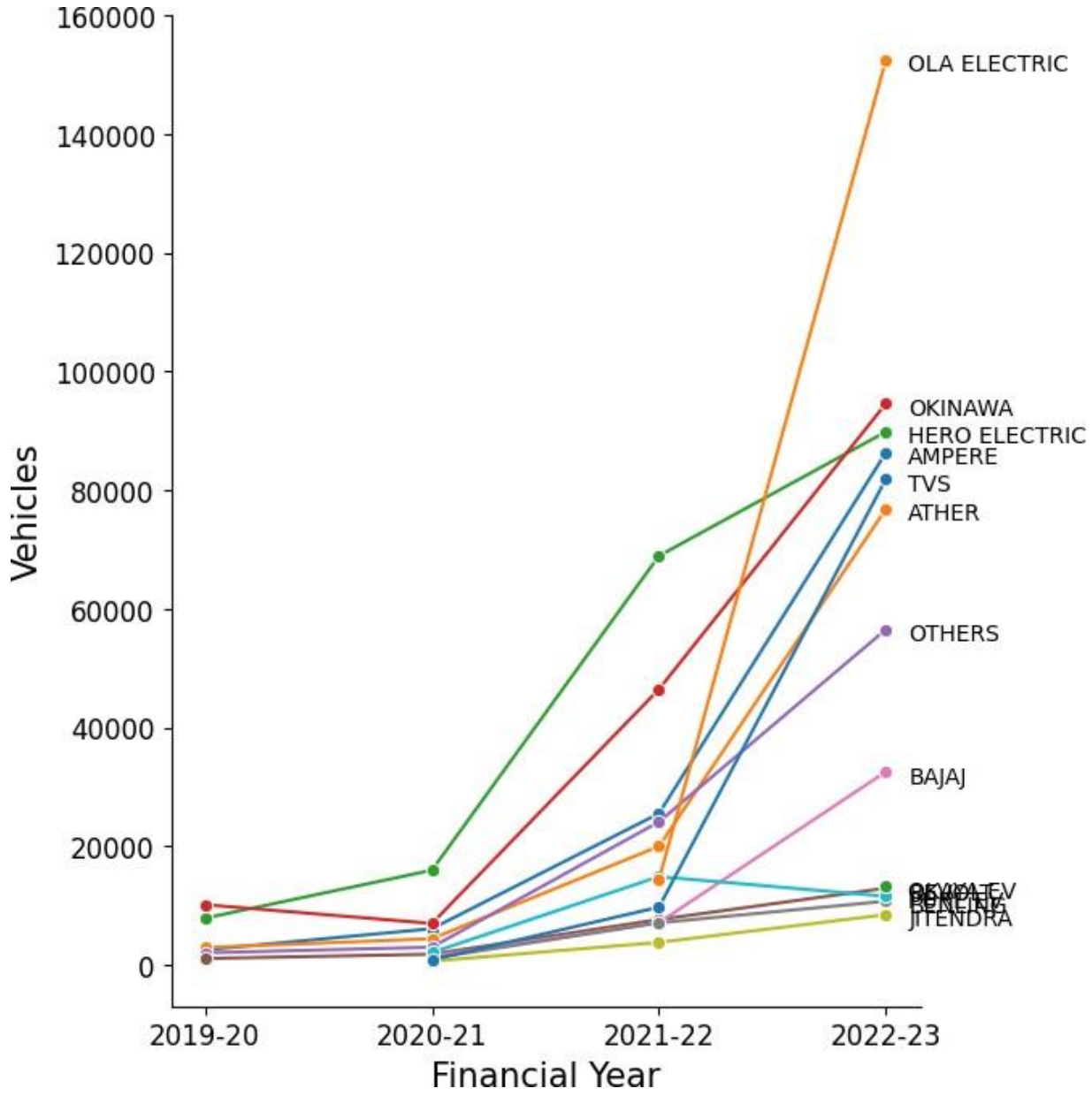
- Key metrics examined include total sales volume, market share of different electric vehicle types (e.g., electric cars, electric two-wheelers, electric buses), regional distribution of sales, and trendsover time.
- The analysis may also delve into factors influencing sales performance, such as pricing strategies, consumer preferences, government incentives, and competition from conventional vehicles.
- Statistical techniques like trend analysis, regression modelling, and market segmentation may be utilized to identify patterns, growth drivers, and market segments with significant sales potential.



The above Figure showcases the remarkable growth trajectory of India's two-wheeler market in 2023 reflecting a convergence of favourable factors, including government support, technological advancements, shifting consumer preferences, and market competition. As electric two-wheelers continue to gain momentum, they are poised to play a significant role in shaping the future of urban mobility and sustainable transportation in India.



The above Figure delved into the market's financial perspective, representation of the industry's total value in crores, with two-wheelers emerging as the primary revenue generators, offers valuable insights into the market's financial landscape and underscores the economic significance of electric two-wheelers within India's electric vehicle industry.

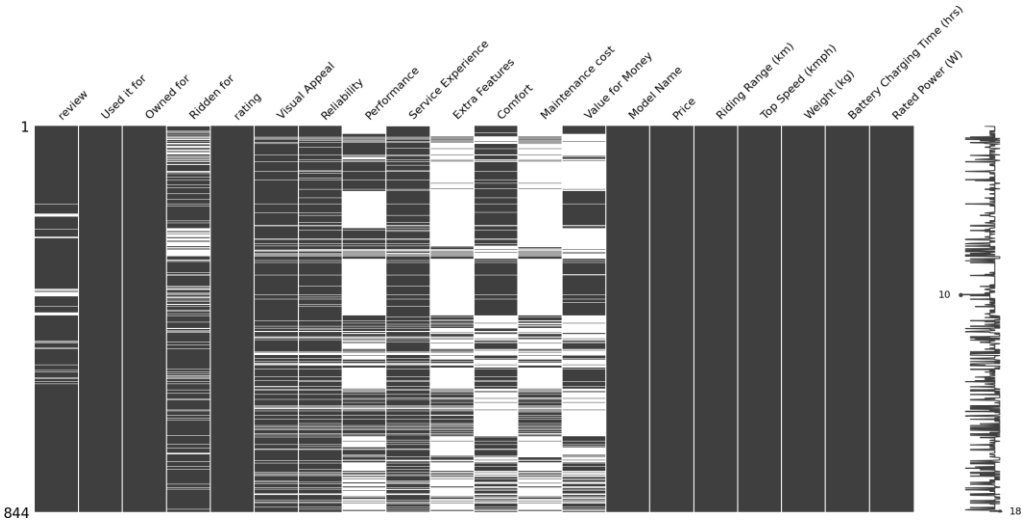


The above Figure honed in on specific electric two-wheeler companies, Ola Electric's emergence as the market leader in 2023 in the electric two-wheeler segment highlights its industry leadership, market competitiveness, and strategic positioning within India's electric vehicle market. As Ola Electric continues to innovate, expand, and drive market growth, its leadership position is likely to shape the trajectory of the electric two-wheeler industry in the years to come.

Market Segmentation

1. Customer Reviews Analysis:
    - This subsection entails analyzing customer reviews and feedback from various online platforms, forums, and social media channels.
    - Natural language processing (NLP) techniques, sentiment analysis, and topic modelling may be employed to extract insights from customer reviews.
    - Key aspects assessed include overall satisfaction levels, likes and dislikes regarding specific electric vehicle models, performance feedback, user experiences with charging infrastructure, reliability, and after-sales service.
    - The analysis aims to identify common themes, sentiments, and areas for improvement based on customer feedback, which can inform product development, marketing strategies, and customer engagement initiatives.
2. Technical Specifications Analysis:
    - This subsection involves a detailed examination of technical specifications provided by electric vehicle manufacturers.
    - Key parameters analyzed include battery capacity, range per charge, charging time, vehicle performance (e.g., acceleration, top speed), safety features, and onboard technology.
    - Comparative analysis may be conducted to benchmark electric vehicles against each other and conventional vehicles in terms of performance, efficiency, and features.
    - The analysis aims to identify trends in technological advancements, areas of innovation, and consumer preferences for specific technical specifications.
    - Insights derived from technical specifications analysis can guide manufacturers in product development, pricing strategies, and positioning within the competitive landscape.

The dataset used for the market segmentation analysis, extracted from bikewale.com, comprises electric two-wheeler customer reviews, offering vital behavioural and psychographic insights.
It also presents detailed technical specifications and pricing information for electric two-wheelers. This data allowed us to assess the technical feasibility and price points crucial for our market

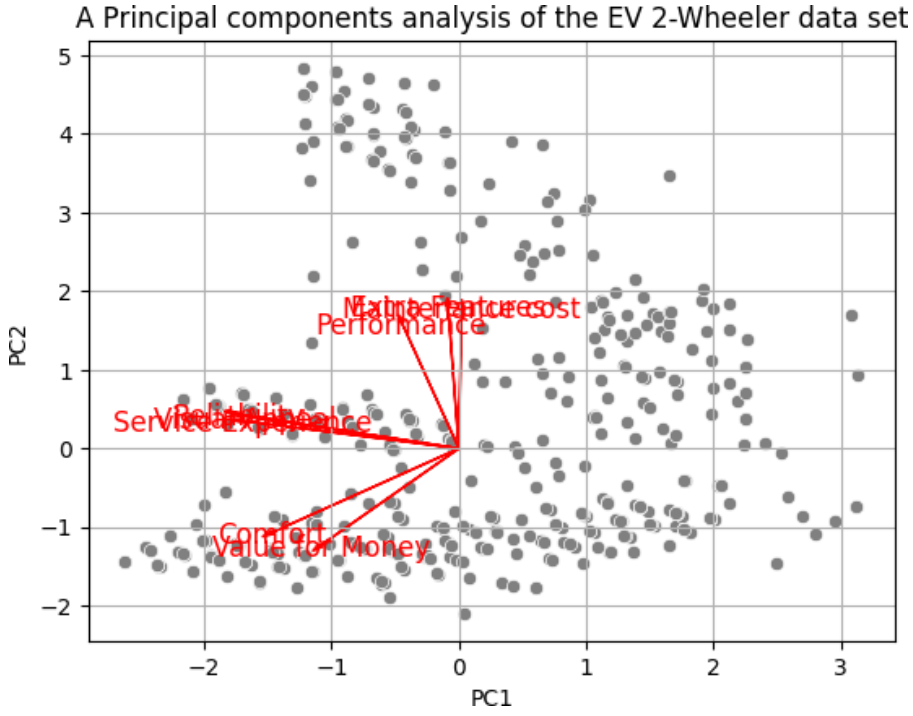


segmentation strategy

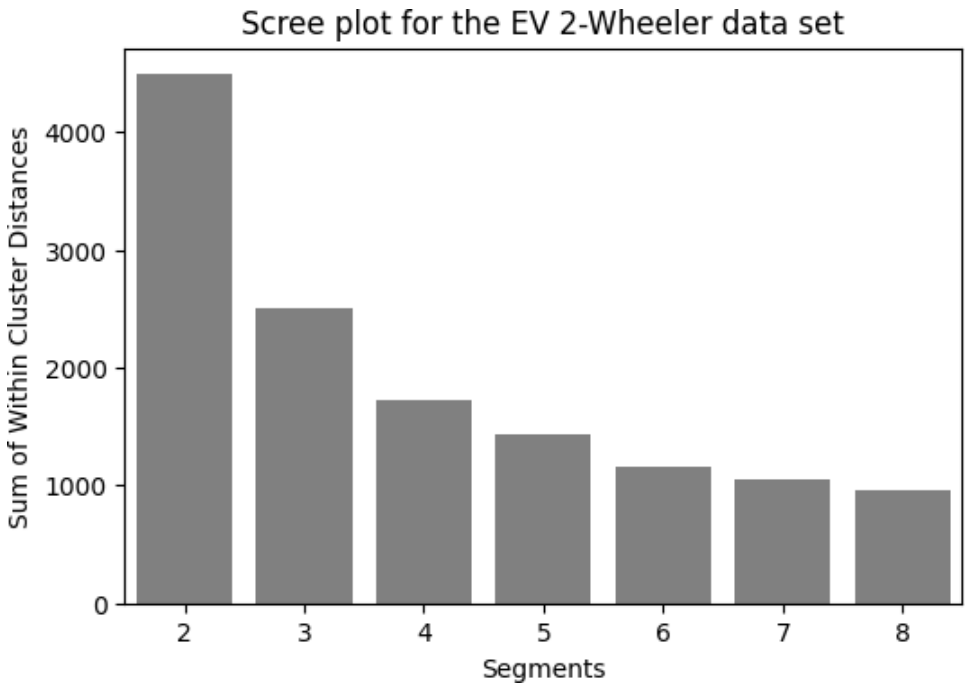
The missingness matrix in the EV market analysis highlights data completeness and quality, aiding researchers in identifying and addressing gaps to ensure reliable insights. Understanding missing data patterns is crucial for making informed decisions and drawing accurate conclusions about market trends and consumer behaviour in the EV industry.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Visual Appeal	-0.480170	0.117814	0.063320	-0.730598	0.247014	0.105903	0.375474	0.067539
Reliability	-0.494758	0.124910	-0.002776	0.152447	-0.819319	0.060484	0.117211	0.166384
Performance	-0.128721	0.459145	0.574833	-0.005549	-0.019902	-0.025704	-0.288468	-0.598232
Service Experience	-0.486499	0.100691	-0.054176	0.653781	0.470391	0.052432	0.311210	-0.044129
Extra Features	-0.024373	0.519633	-0.364578	-0.023208	0.116821	0.559390	-0.456829	0.246323
Comfort	-0.418255	-0.304266	0.249807	-0.020111	0.172621	-0.296656	-0.623271	0.404238
Maintenance cost	0.005912	0.513208	-0.386495	-0.054822	0.020302	-0.762039	-0.003360	0.055435
Value for Money	-0.309572	-0.351548	-0.563840	-0.107598	-0.046688	0.009572	-0.260855	-0.617065

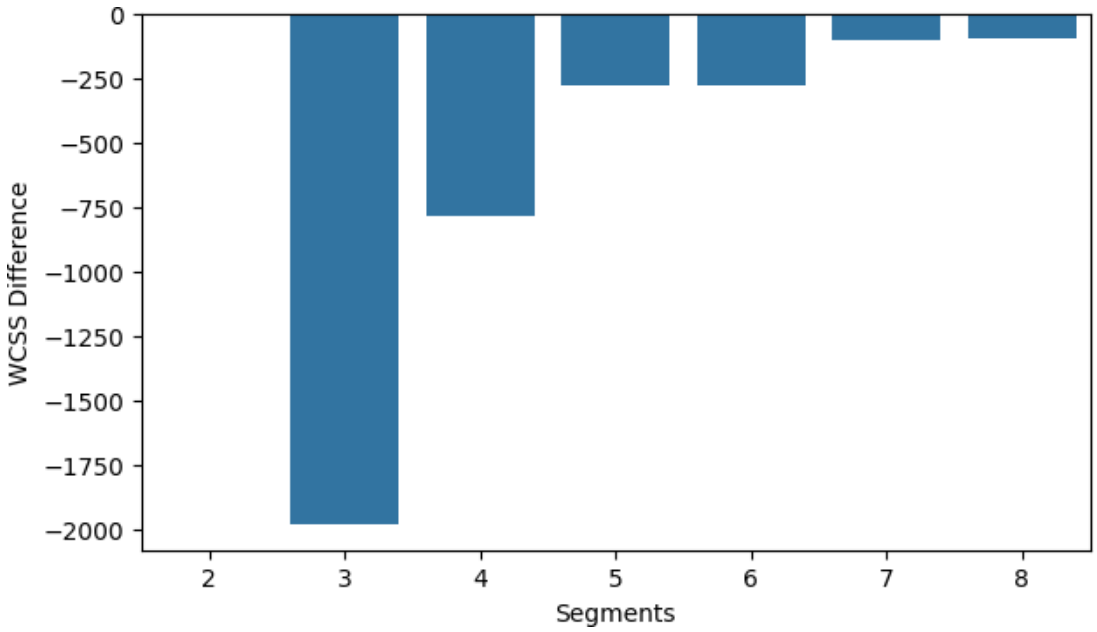
In the context of the electric vehicle (EV) market analysis, the factor loadings obtained from PCA can provide valuable insights into the underlying structure and relationships among the variables (features) in the dataset. By examining the factor loadings, analysts can identify which variables contribute most significantly to each principal component and understand the patterns and trends driving variation within the data.



Facilitates a deeper understanding of the relationships and patterns within the EV two-wheeler dataset, enabling stakeholders to make informed decisions regarding market segmentation, product development, and strategic planning in the electric vehicle market.



The scree plot serves as a valuable tool for determining the optimal number of clusters in K-means clustering and facilitates the identification of meaningful segments within the EV two-wheeler dataset, contributing to more effective market segmentation and strategic decision-making in the electric vehicle market.



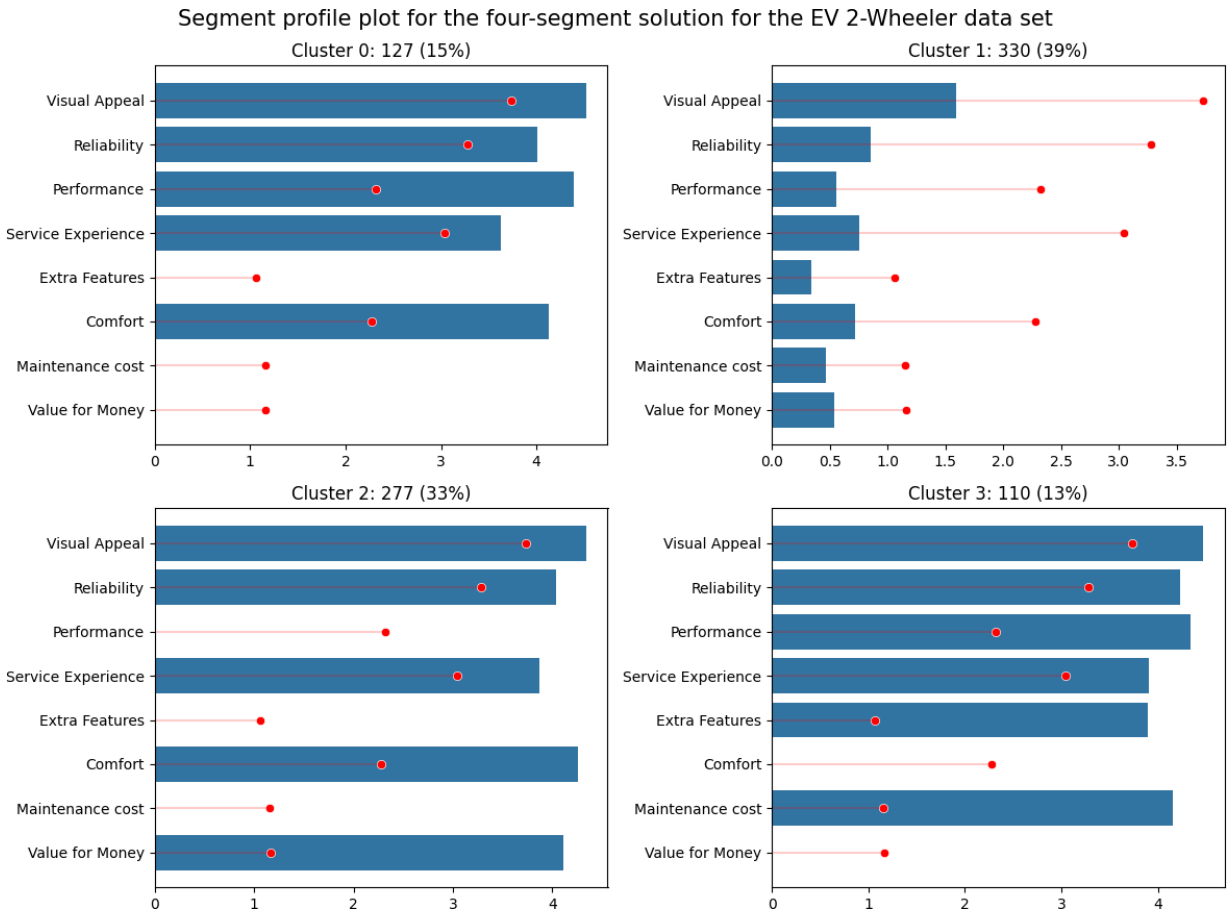
The decision-making process was significantly guided by the scree plot above, revealing a distinct elbow at four segments. This marked point indicated a substantial reduction in distances, signifying the optimal number of segments for our analysis.

Using K-means

K-means clustering is a powerful tool for segmentation in the electric vehicle market, enabling businesses and researchers to uncover meaningful patterns and insights within complex datasets. By leveraging K-means clustering, stakeholders can make informed decisions and strategies to address the diverse needs and preferences of consumers in the evolving EV market landscape.

K-means clustering is a valuable tool for segmentation in the EV market, enabling businesses to gain actionable insights, enhance customer engagement, and drive competitive advantage in a rapidly evolving industry landscape. By leveraging the capabilities of K-means clustering, businesses can effectively navigate market complexities and capitalize on emerging opportunities in the dynamic EV market ecosystem.

Profiling Segments



The above graph visually captures the diverse perceptions among different segments. Segment 0, representing 15% of consumers, values the electric two-wheeler vehicle for its visual appeal, reliability, performance, service experience, and comfort. Conversely, Segment 1 (39% of consumers) expresses dissatisfaction across all aspects, marking them as the largest but least satisfied group. Segment 2 (33% of consumers) appreciates visual appeal, reliability, service experience, and comfort, and notably, perceives a strong value for money. Lastly, Segment 3 (13% of consumers), the smallest segment, values visual appeal, reliability, performance, service experience, extra features, and maintenance cost, showcasing distinct perceptions, particularly on features and costs.

Segment 0 (15% of consumers):

1. Values: Visual appeal, reliability, performance, service experience, and comfort.
2. Perception: Positive across various aspects, indicating high satisfaction levels within this segment.
3. Potential Marketing Approach: Emphasize product design, reliability, performance, and customer service to maintain and enhance satisfaction levels.

Segment 1 (39% of consumers):

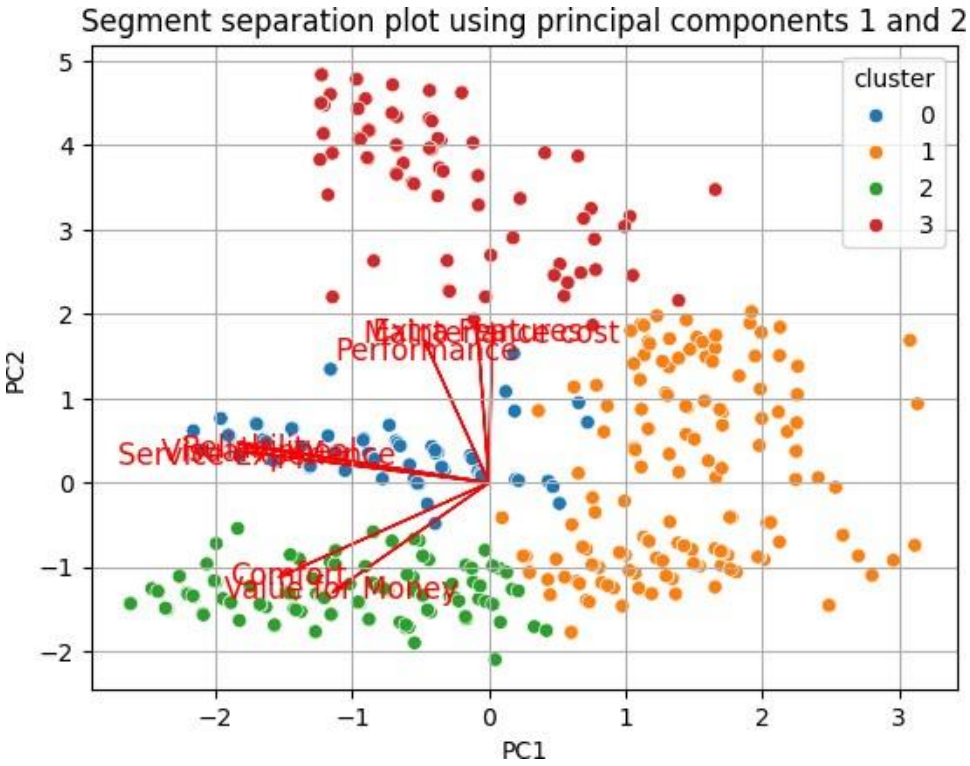
1. Values: Not explicitly mentioned, but expresses dissatisfaction across all aspects.
2. Perception: Represents the largest but least satisfied group, highlighting significant opportunities for improvement.
3. Potential Marketing Approach: Address pain points, gather feedback, and implement changes to enhance product quality, service experience, and overall satisfaction.

Segment 2 (33% of consumers):

1. Values: Visual appeal, reliability, service experience, comfort, and strong value for money.
2. Perception: Generally positive, particularly in terms of value for money, indicating a favourable perception of the product's cost-effectiveness.
3. Potential Marketing Approach: Highlight affordability, emphasize reliability, service quality, and overall value proposition to attract and retain customers in this segment.

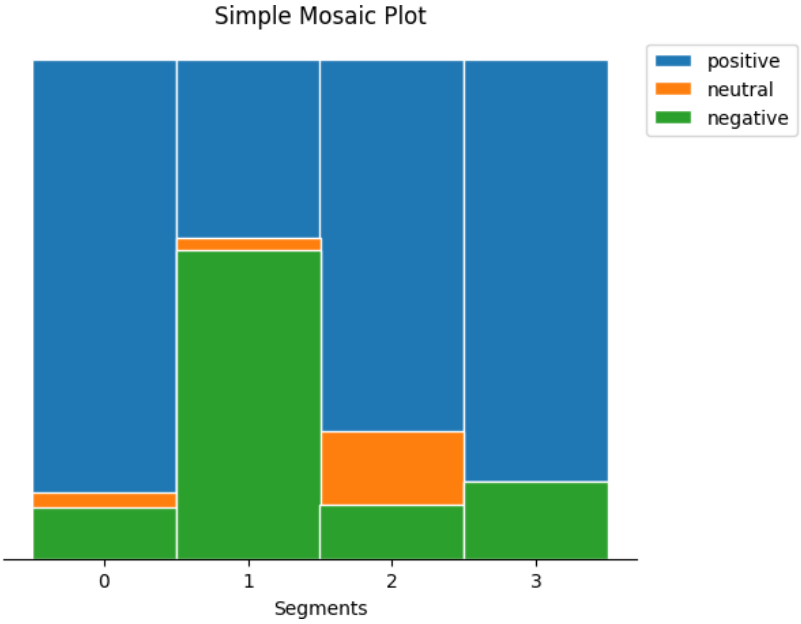
Segment 3 (13% of consumers):

1. Values: Visual appeal, reliability, performance, service experience, extra features, and maintenance cost.
2. Perception: Positive perceptions across various aspects, particularly on features and maintenance costs.
3. Potential Marketing Approach: Focus on product differentiation, highlight premium features, emphasize reliability, and offer competitive maintenance packages to cater to the specific preferences of this segment.



The above Figure, utilizing principal components, further emphasizes the differences among segments. Notably, Segment 1, despite being the largest segment, lacks specific opinions, making it unique in its lack of satisfaction.

Describing Segments



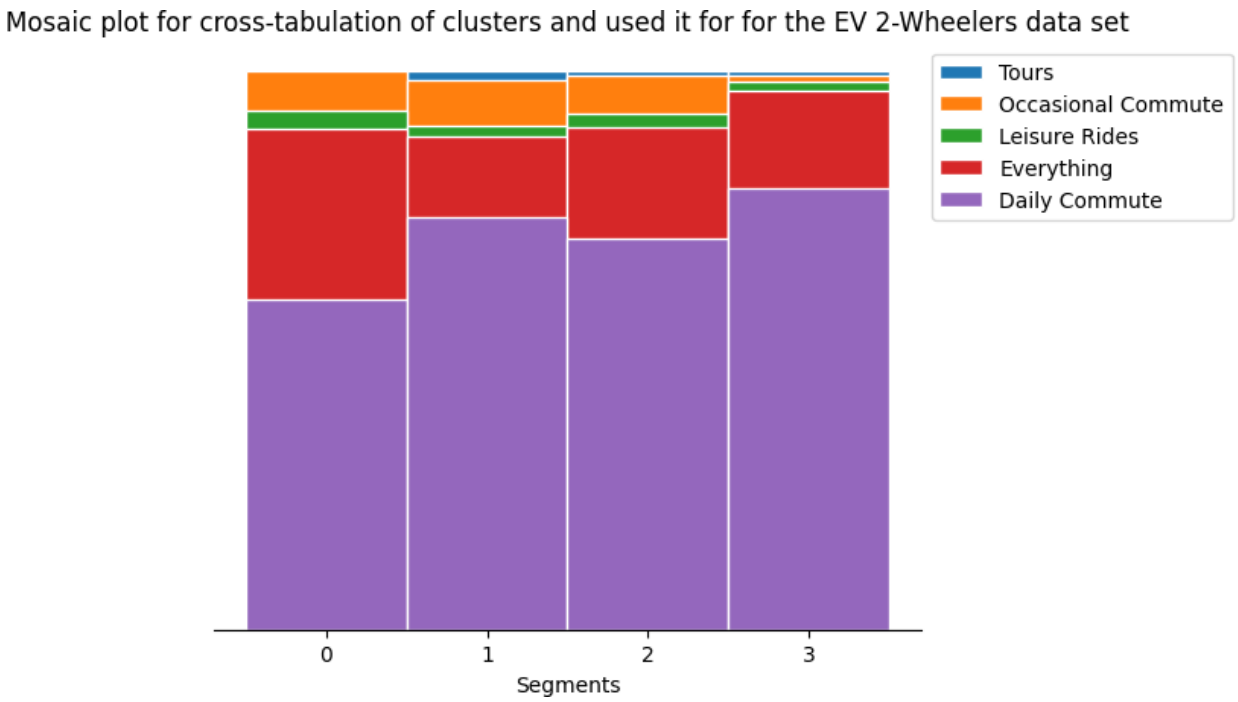
Above mosaic plot, explores consumer sentiments, revealing that all segments, except Segment 1, exhibit positive sentiments. Segment 1 consumers stand out with negative sentiments, indicating dissatisfaction across various aspects.

1. Consumer Sentiments Analysis:
    - The mosaic plot visually represents consumer sentiments, with each segment categorized based on positive or negative sentiments.
    - Positive sentiments indicate satisfaction or positive perceptions, while negative sentiments reflect dissatisfaction or negative perceptions.
    - Understanding consumer sentiments is crucial for identifying areas of strength and improvement within each segment and guiding targeted marketing and product enhancement strategies.

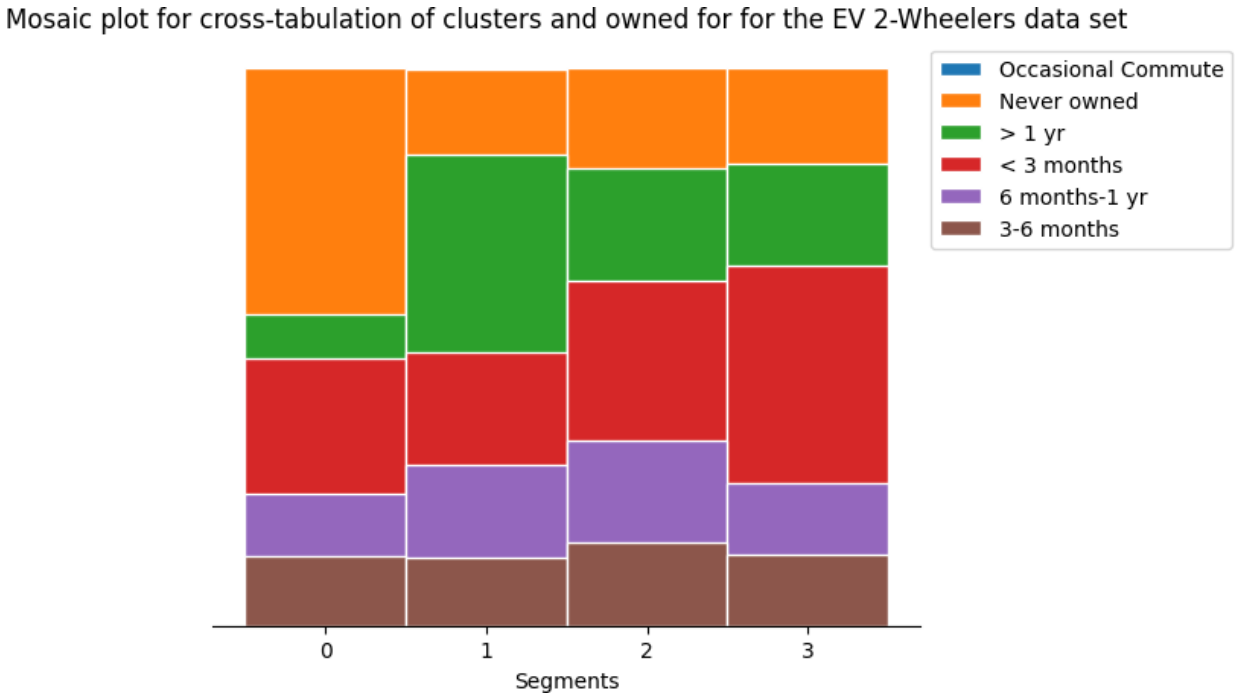
2. Segment Comparison:
    - Segments 0, 2, and 3 exhibit predominantly positive sentiments, indicating overall satisfaction or positive perceptions among consumers within these segments.
    - Consumers in these segments likely appreciate various aspects of electric vehicles, such as visual appeal, reliability, performance, service experience, and value for money.
    - Their positive sentiments suggest that these segments may represent loyal customers or enthusiasts who are generally satisfied with their electric vehicles.

3. Segment 1 Dissatisfaction:
    - Segment 1 stands out with negative sentiments, indicating dissatisfaction across various aspects of electric vehicles.
    - Consumers in this segment express negative perceptions or experiences, highlighting areas of concern and potential dissatisfaction within the electric vehicle market.
    - Their negative sentiments may stem from issues related to product quality, serviceexperience, pricing, or other factors that fail to meet their expectations or requirements.

4. Implications for Market Strategies:
    - The mosaic plot underscores the importance of addressing dissatisfaction within Segment 1 to improve overall customer satisfaction and loyalty in the electric vehicle market.
    - Businesses should prioritize understanding the underlying reasons for dissatisfaction within Segment 1 and take proactive measures to address consumer concerns and improve product offerings and service quality.
    - Targeted marketing campaigns, product enhancements, customer support initiatives, andpricing strategies may be necessary to effectively address the needs and preferences of consumers within Segment 1 and enhance their overall satisfaction levels.



The above mosaic plot illustrates that all segments predominantly use electric vehicles for daily commuting, with limited usage for tours, occasional commuting, and leisure rides.



The above mosaic plot delineates the ownership duration of electric vehicles among segments. Segment 1 stands out, owning electric vehicles for more than a year, while Segment 0 has no prior ownership experience. Segment 2 members moderately own vehicles ranging from less than 3 months to over a year, and Segment 3 consumers have owned electric vehicles for a few days to less than 3 months.

1. Segment 0 Ownership Experience:
    - Segment 0, in contrast, has no prior ownership experience with electric vehicles.
    - This suggests that consumers in Segment 0 may represent potential new adopters or individuals exploring electric vehicle ownership for the first time.

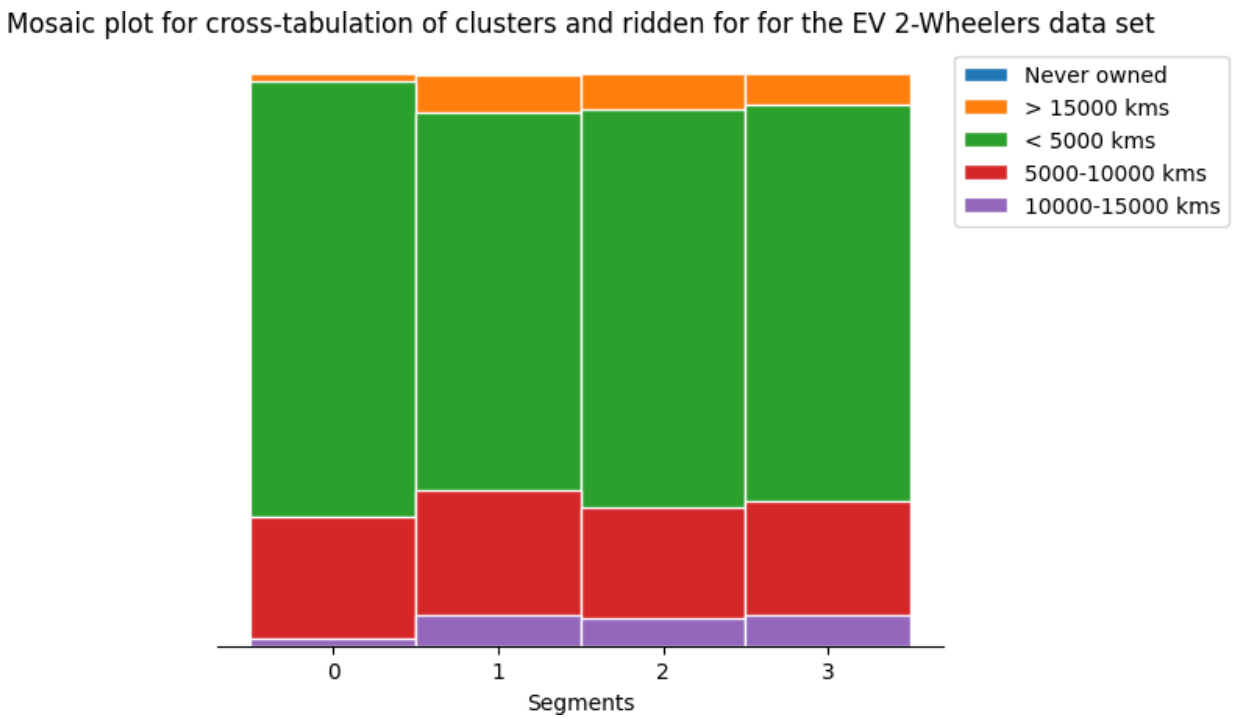
2. Segment 1 Ownership Duration:
    - Segment 1 stands out for owning electric vehicles for more than a year.
    - This indicates a longer-term commitment to electric vehicle ownership among consumers in Segment 1, suggesting potential loyalty or satisfaction with electric vehicles over time.

3. Segment 2 Ownership Duration
    - Members of Segment 2 exhibit moderate ownership durations, ranging from less than 3 months to over a year.
    - This indicates a range of experiences and tenure among consumers in Segment 2, with some being relatively new owners and others having more established ownership histories.

4. Segment 3 Ownership Duration:
    - Consumers in Segment 3 have owned electric vehicles for a relatively short duration, ranging from a few days to less than 3 months.
    - This suggests that Segment 3 represents a group of recent adopters or individuals who have recently transitioned to electric vehicle ownership.

5. Implications for Market Understanding:
    - Understanding ownership duration patterns provides insights into consumer behaviour, satisfaction levels, and brand loyalty within the electric vehicle market.
    - Businesses can tailor their marketing strategies, customer engagement initiatives, and product offerings based on the unique needs and preferences of each segment.

6. Opportunities for Engagement:
    - Segment 0 presents opportunities for education, outreach, and incentives to encourage new adopters to explore electric vehicle ownership.
    - Segment 1 represents an opportunity for retention and loyalty-building efforts to maintain satisfaction and engagement among long-term owners.
    - Segments 2 and 3 may benefit from targeted messaging, support services, and product enhancements tailored to their specific ownership durations and experiences.



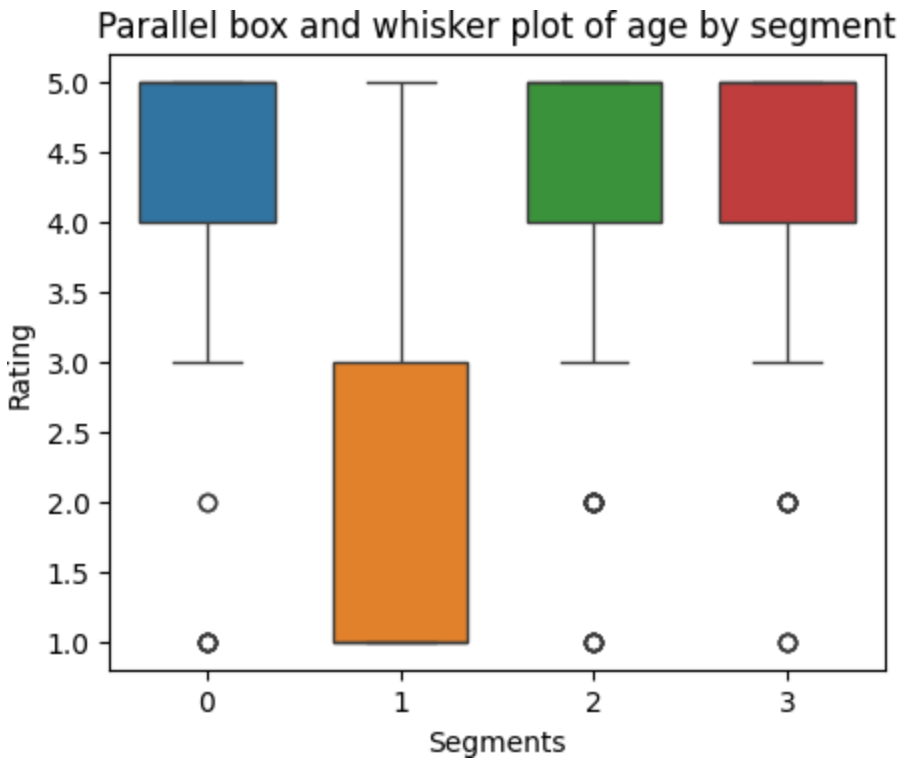
The above mosaic plot delves into the distances covered by consumers, indicating that all segments predominantly use electric vehicles for commuting, with most users covering distances below 5000 km. A small portion falls in the 5000 to 10000 km range, aligning with their commuting needs.

1. Distance Coverage Patterns:
    - The majority of users in all segments cover distances below 5000 kilometres.
    - This suggests that most electric vehicle users typically engage in short to moderate-distance commuting, which aligns with urban and suburban travel patterns.

2. Moderate Distance Range:
    - A small portion of users across segments falls within the 5000 to 10000 kilometres range.
    - This moderate distance range likely corresponds to longer commutes or occasional travel needs, further emphasizing the practicality and versatility of electric vehicles for various commuting distances.

3. Implications for Market Understanding:
    - Understanding the distances covered by consumers provides valuable insights into usage patterns, travel behaviours, and preferences within the electric vehicle market.
    - Businesses can leverage this information to optimize electric vehicle features, range capabilities, charging infrastructure, and marketing messages to meet the diverse computing needs of consumers.

4. Market Opportunities:
    - The predominance of commuting as the primary usage pattern presents opportunities for businesses to innovate and differentiate their electric vehicle offerings.
    - Strategies to enhance range, efficiency, comfort, and convenience for commuting purposes can drive market adoption and customer satisfaction in the electric vehicle market.



The above parallel box and whisker plot, emphasize significant differences in average ratings among segments. Specifically, Segment 1 consumers express dissatisfaction across all perceptions, leading to lower overall ratings.

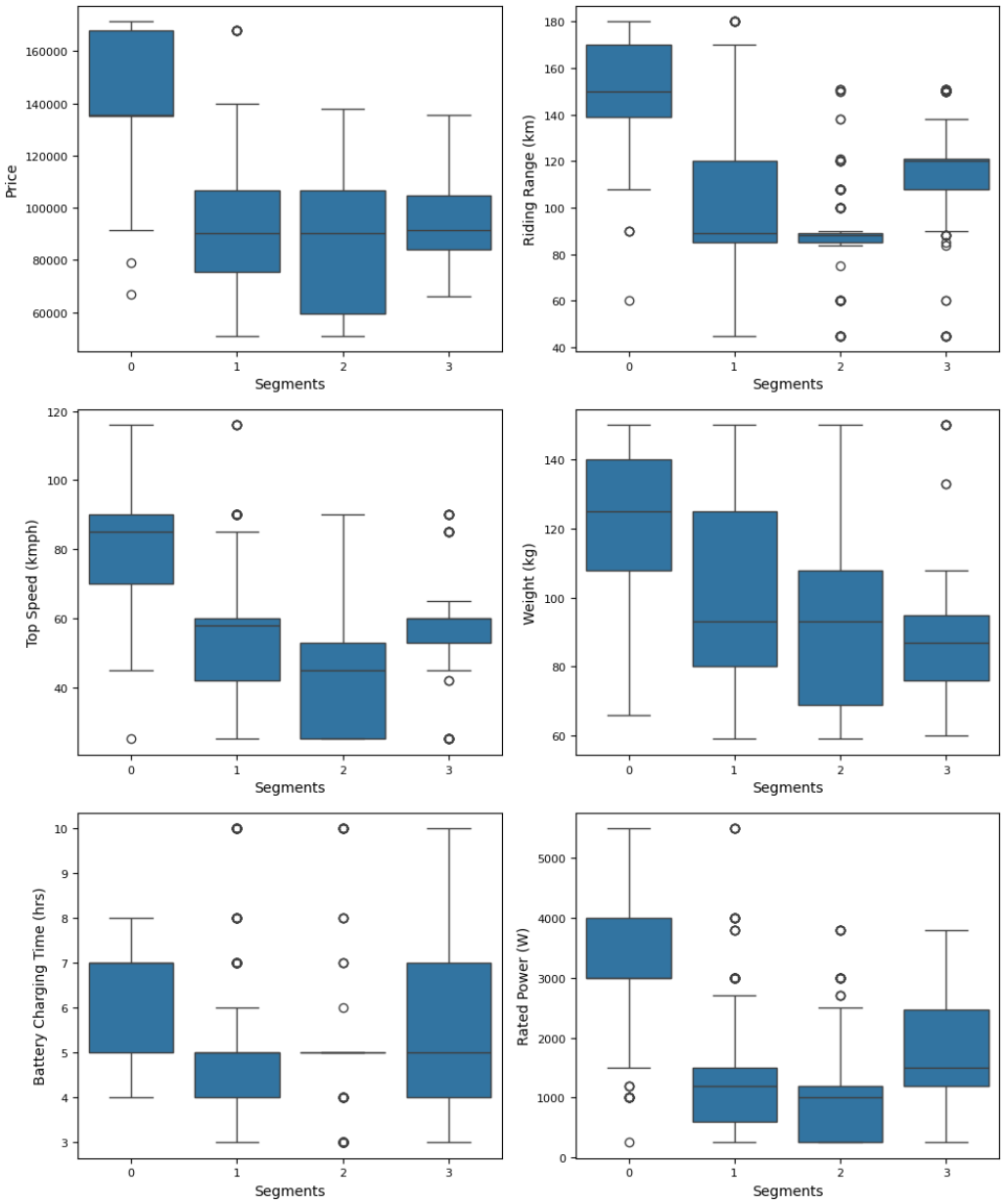
1. Segment 1 Dissatisfaction:
    - Segment 1 stands out for expressing dissatisfaction across all perceptions.
    - This indicates that consumers within Segment 1 consistently rate electric vehicles lower across various aspects, highlighting significant areas of dissatisfaction or discontentment with the product or service experience.

2. Lower Overall Ratings:
    - The consistent expression of dissatisfaction among Segment 1 consumers contributes to lower overall ratings for this segment.
    - Lower overall ratings may impact brand perception, customer loyalty, and market competitiveness, emphasizing the importance of addressing consumer concerns and improving satisfaction levels within Segment 1.

3. Implications for Market Understanding:
    - Understanding the factors contributing to dissatisfaction within Segment 1 is crucial for market understanding and strategic decision-making.
    - Businesses need to identify and address underlying issues related to product quality, service experience, pricing, or other factors that contribute to lower ratings and dissatisfaction among Segment 1 consumers.

4. Opportunities for Improvement:
    - Segment 1 dissatisfaction presents opportunities for improvement and intervention to enhance customer satisfaction and loyalty.
    - Strategies such as product enhancements, service quality improvements, customer engagement initiatives, and targeted marketing campaigns can help address consumer concerns and improve overall ratings within Segment 1.

5. Competitive Advantage:
    - Addressing consumer dissatisfaction within Segment 1 can lead to competitive advantage and differentiation within the electric vehicle market.
    - By prioritizing customer satisfaction and addressing pain points, businesses can enhance brand reputation, customer loyalty, and market share in the competitive landscape.



In analyzing the technical specifications of electric vehicles across different segments, distinct patterns emerge. Segment 0 prefers premium EVs with a higher price range and extended riding range, emphasizing consumer preference for luxury and long-distance travel. Segment 1 focuses on budget-friendly options with lower prices and moderate riding ranges, suitable for daily commuting. Segment 2 and Segment 3 prioritize affordability, with slight differences in riding range and speed preferences. Weight preferences vary, with Segment 0 and Segment 1 favoring heavier vehicles, while Segment 2 and Segment 3 prefer lighter options. Charging time also differs, with Segment 0 and Segment 3 opting for longer durations for overnight charging, while Segment 1 and Segment 2 prioritize faster charging for quick turnaround times. These nuanced preferences shape the electric vehicle market in India.

1. Segment 0 Preferences:
    - Preference for Premium EVs: Segment 0 shows a preference for premium electric vehicles characterized by higher price ranges and extended riding ranges.
    - Emphasis on Luxury and Long-Distance Travel: Consumers in Segment 0 prioritize luxury features and longer riding ranges, indicating a preference for high-end electric vehicles suitable for long-distance travel and premium experiences.

2. Segment 1 Preferences:
    - Focus on Budget-Friendly Options: Segment 1 consumers prioritize budget-friendly electric vehicles with lower prices and moderate riding ranges.
    - Suitability for Daily Commuting: Electric vehicles in Segment 1 are tailored to meet the needs of daily commuters, offering affordable options with practical riding ranges for everyday transportation.

3. Segment 2 and Segment 3 Preferences:
    - Priority on Affordability: Both Segment 2 and Segment 3 prioritize affordability in electric vehicles, with slight differences in riding range and speed preferences.
    - Varied Weight Preferences: While Segment 0 and Segment 1 favor heavier vehicles, Segment 2 and Segment 3 prefer lighter options, reflecting diverse consumer preferences and usage scenarios.

4. Charging Time Preferences:
    - Segment 0 and Segment 3 opt for longer charging durations suitable for overnight charging, reflecting a preference for convenience and flexibility in charging schedules.
    - Segment 1 and Segment 2 prioritize faster charging capabilities for quick turnaround times, indicating a need for efficient charging solutions tailored to daily commuting patterns and urban lifestyles.

5. Market Implications:
    - Nuanced Consumer Preferences: The nuanced preferences observed across different segments shape the electric vehicle market landscape in India, influencing product development, pricing strategies, and marketing approaches.
    - Opportunity for Segmented Offerings: Understanding consumer preferences allows manufacturers and stakeholders to develop segmented offerings tailored to specific market segments, enhancing competitiveness and market penetration.

Selection of Target Segment

The strategic target segments for the electric vehicle market are identified as Segment 1 (39% of consumers) and Segment 2 (33% of consumers). Segment 1's diverse preferences and dissatisfaction points present an opportunity for improving customer satisfaction and loyalty by directly addressing their specific demands. Segment 2 values visual appeal, reliability, service experience, and comfort, offering a chance to customize electric vehicles to meet these expectations and emphasize value for money. The strategy involves addressing dissatisfaction points in Segment 1 and enhancing positive elements in Segment 2, aligning electric vehicles with the distinct expectations of each segment to ensure competitiveadvantage and sustained market growth.

Customizing the Marketing Mix

In our electric vehicle market strategy, customization of the marketing mix is crucial for appealing to Segment 1 and Segment 2, our target segments.

- **Product customization** involves enhancing features based on specific desires, addressing dissatisfaction points for Segment 1, and emphasizing visual appeal and value for money for Segment 2. Diverse offerings cater to varied tastes and budgets within each segment.
- **Price customization** includes competitive pricing for Segment 1 and a slightly higher price point for value-added features in Segment 2.
- **Promotion customization** focuses on targeted advertising and tailored promotional events for each segment's preferences.
- **Place customizatio**n establishes accessible distribution channels in urban areas for Segment 1 and suburban/semi-urban regions for Segment 2, with a strong emphasis on online presence andcustomer support.
- **People and Process Customization** involves training customer service representatives to address segment-specific concerns and ensuring efficient processes for customization requests and service appointments. This tailored approach ensures our electric vehicles align with the distinct needs of Segment 1 and Segment 2, enhancing market relevance and customer preference.

Potential Early Market Customer Base

1. Segment Identification:
    a. Segment 1: Consists of 330 members, representing 39% of consumers.
    b. Segment 2: Comprises 277 members, accounting for 33% of consumers.

2. Target Price Range:
    a. Segment 1: Target price range falls between ₹51,094 and ₹1,67,844.
    b. Segment 2: Target price range ranges from ₹51,094 to ₹1,37,890.

3. Potential Profits Calculation:
    a. For Segment 1:
    Assuming a target price of ₹1,20,000, the potential profit can be calculated by multiplying the number of potential customers (330 members) by the target price.
    Potential profit = 330 members * ₹1,20,000 = ₹39.60 crores.
    b. For Segment 2:

Considering a target price of ₹1,10,000, the potential profit is determined by multiplying the number of potential customers (277 members) by the target price.
Potential profit = 277 members * ₹1,10,000 = ₹30.47 crores.

4. Market Penetration Focus:
    a. Segment 1 is identified as the primary focus for early market penetration efforts due to its larger potential market share and higher profit opportunity.
    b. With a significant number of potential customers and a broader price range, Segment 1 offers substantial profit potential for early market penetration initiatives.

5. Strategic Implications:
    a. Targeting Segment 1 allows businesses to capitalize on the significant profit opportunity and establish a strong foothold in the market.
    b. Strategies such as targeted marketing campaigns, product enhancements, and customer engagement initiatives can be tailored to meet the specific needs and preferences of Segment 1 consumers, driving market penetration and revenue growth.

Most Optimal Market Segments

1. Segment 1 Selection:
    - Constitutes 39% of consumers: Segment 1 represents a sizable portion of the target market, offering significant market potential and growth opportunities.
    - Balanced blend of specifications and price range: Segment 1's characteristics align well with the target market's preferences, providing a balance between technical specifications and affordability.

2. Recommended Technical Specifications for Segment 1:
    - Price Range: ₹70,688 to ₹1,29,063- The recommended price range ensures affordability while offering a range of options to cater to different budget preferences within Segment1.
    - Riding Range: 89 to 180 km- A moderate riding range addresses the commuting needs of consumers while providing flexibility for longer trips.
    - Top Speed: 58 to 116 kmph- A varied range of top speeds accommodates different usage scenarios and preferences among consumers.
    - Weight: 76 to 120 kg- Optimal weight ensures manoeuvrability, ease of handling, and efficiency in urban and suburban environments.
    - Battery Charging Time: 3 to 5 hours- Efficient charging times cater to the convenience and practicality needs of consumers, facilitating daily usage without significant downtime.
    - Rated Power: 1200 to 5500 W- Varied power ratings offer options for consumers with different performance requirements and preferences.

3. Targeted Approach and Market Alignment:
    - Tailoring technical specifications to meet the diverse needs and preferences of Segment 1 consumers ensures market alignment and enhances the likelihood of success in the electric vehicle market.
    - The targeted approach focuses on delivering value, performance, and affordability, key factors that influence consumer purchasing decisions in the electric two-wheeler segment.

4. Foundation for Success:
    - The selection of Segment 1 as the optimal market segment, coupled with the recommended technical specifications, lays a solid foundation for a successful and sustainable venture into the electric vehicle market.
    - By aligning products with consumer preferences and market dynamics, businesses can capitalize on opportunities, drive adoption, and establish a competitive edge in the evolving electric two-wheeler market landscape.

Conclusion

In conclusion, an in-depth analysis of India's electric vehicle market led us to identify Segment 1 as the optimal target. With a significant 39% consumer base, this segment represents a substantial market opportunity. By tailoring our electric two-wheeler specifications to meet the preferences of this segment, we ensure our products align seamlessly with the demands of a large customer base. This strategic decision is grounded in a thorough understanding of market segmentation, consumer behaviour, and technical specifications. These insights provide a clear direction for our market entry, emphasizing precision and relevance in both product development and marketing strategies. Moving forward, this approach equips us with a solid foundation, ensuring our offerings resonate effectively within India's evolving electric vehicle landscape.

Market Segmentation Analysis Electric Vehicle Market In India

Taware Harshal Avinash

Date: 26/02/2024

GitHub Link: https://github.com/Harshalt05/Feynn-Lab1

Abstract- This research delves into the market segmentation of emerging transportation technologies, specifically focusing on electric vehicles (EVs). Through a comprehensive cross-sectional online survey, the study categorizes consumers into three distinct groups: 'Traditionalists,' 'Apathetics,' and 'Enthusiasts.' Employing advanced analytical techniques such as cluster analysis, multiple discriminant analysis, and Chi-square tests, the research scrutinizes the psychographic, behavioral, and socio-economic attributes characterizing each segment. The implications of these findings extend to academia and policymaking, offering valuable insights for crafting targeted strategies to boost EV adoption in sustainable transportation markets. As EVs gain traction for their potential in emission reduction and cost efficiency, this study emphasizes the necessity of tailored engagement approaches to effectively reach diverse consumer segments. Ultimately, the research contributes to the overarching objective of promoting widespread EV adoption in burgeoning markets.

1. ***Data collection and Preprocessing:***

Data Collection Process

The data collection process was conducted manually, sourcing information from various platforms including: Kaggle Datasets (https://www.kaggle.com/datasets)

Data- Preprocessing:

The condensed data serves two key objectives: visualization and clustering. Python libraries such as NumPy, Pandas, Scikit-Learn, and SciPy are instrumental in processing the data. The workflow prioritizes reproducibility to enhance the reliability and robustness of the analysis. By leveraging these libraries, the study ensures efficient processing and visualization of the data, while also employing advanced clustering techniques to uncover meaningful insights. This approach not only enhances the interpretability of the results but also provides a foundation for future research and decision-making processes in the realm of emerging transportation technologies like electric vehicles.

```
df = pd.read_csv('ev_model_spec.csv')
```

```
df.head()
```

	Model Name	Price	Riding Range (km)	Top Speed (kmph)	Weight (kg)	Battery Charging Time (hrs)	Rated Power (W)
0	Ampere Magnus EX	104758	121	53	90	7	1200
1	Ampere Magnus Pro	66053	45	25	94	10	250
2	Ampere REO	61993	45	25	88	10	250
3	Ampere Zeal	96544	120	53	80	7	1200
4	Ather 450X	135489	111	90	108	8	3000

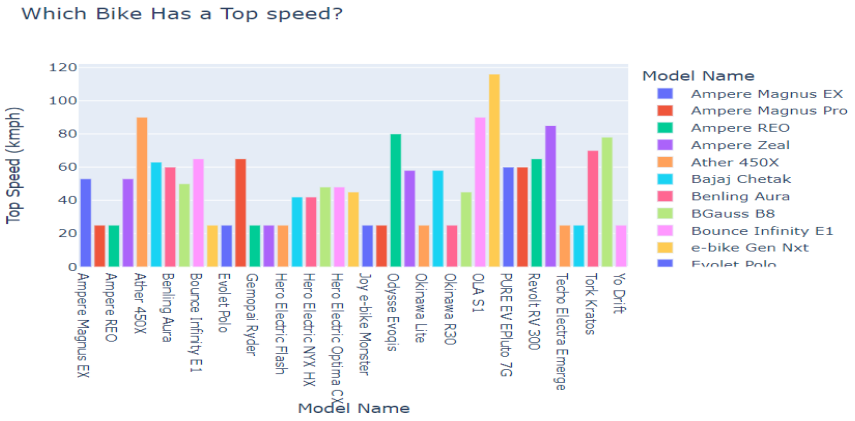
```
df.describe()
```

	Price	Riding Range (km)	Top Speed (kmph)	Weight (kg)	Battery Charging Time (hrs)	Rated Power (W)
count	39.0000	39.0000	39.0000	39.0000	39.0000	39.0000
mean	94763.8974	100.6667	49.7179	98.8462	5.4103	1364.1026
std	30861.6630	34.8344	23.5573	29.0304	1.6657	1302.7778
min	51094.0000	45.0000	25.0000	59.0000	3.0000	250.0000
25%	69553.5000	79.5000	25.0000	78.0000	4.0000	250.0000
50%	90282.0000	90.0000	48.0000	93.0000	5.0000	1000.0000
75%	109809.0000	120.5000	64.0000	112.5000	6.0000	1950.0000
max	171250.0000	180.0000	116.0000	150.0000	10.0000	5500.0000

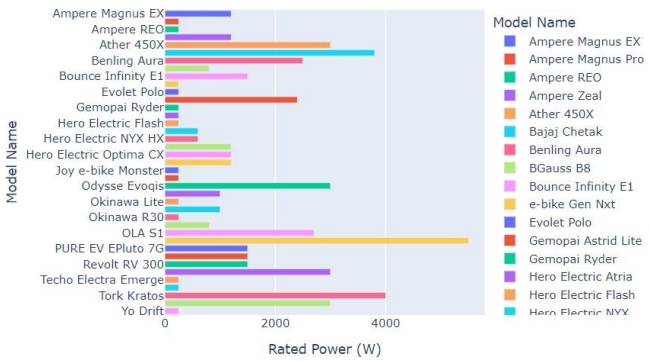
2. *Exploratory Data Analysis (EDA):*

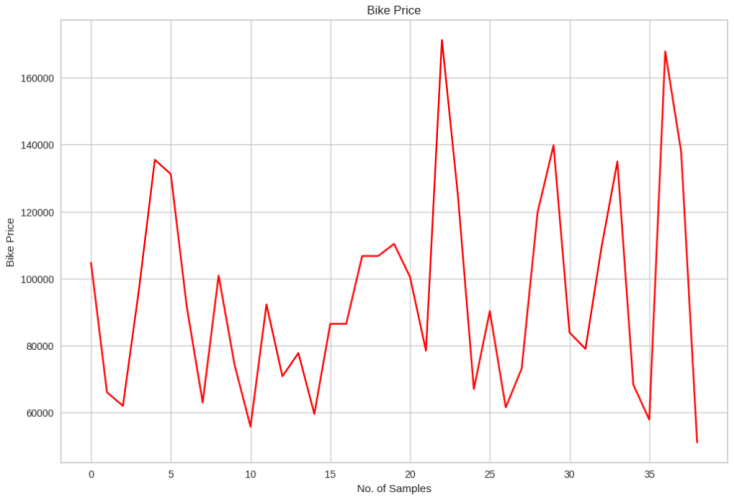
Embarking on the journey of Exploratory Data Analysis (EDA), our initial step involves a thorough examination of the dataset, employing both traditional analysis techniques and Principal Component Analysis (PCA). PCA, a statistical method, serves to convert interrelated features into linearly independent ones through orthogonal transformation. These newly derived features, referred to as Principal Components, play a pivotal role in diminishing data dimensionality, thereby enhancing the efficiency of ensuing classification, regression, or machine learning endeavors. Our primary focus encompasses the exploration of the dataset's inherent characteristics, distributions, and inter-variable relationships. Subsequently, we introduce PCA to evaluate its influence on data dimensionality and the interpretability of the transformed features. This all-encompassing approach aims to shed light on both the original dataset and its condensed representation, empowering thoughtful decision-making in subsequent analytical pursuits.

comparison of bikes in our data

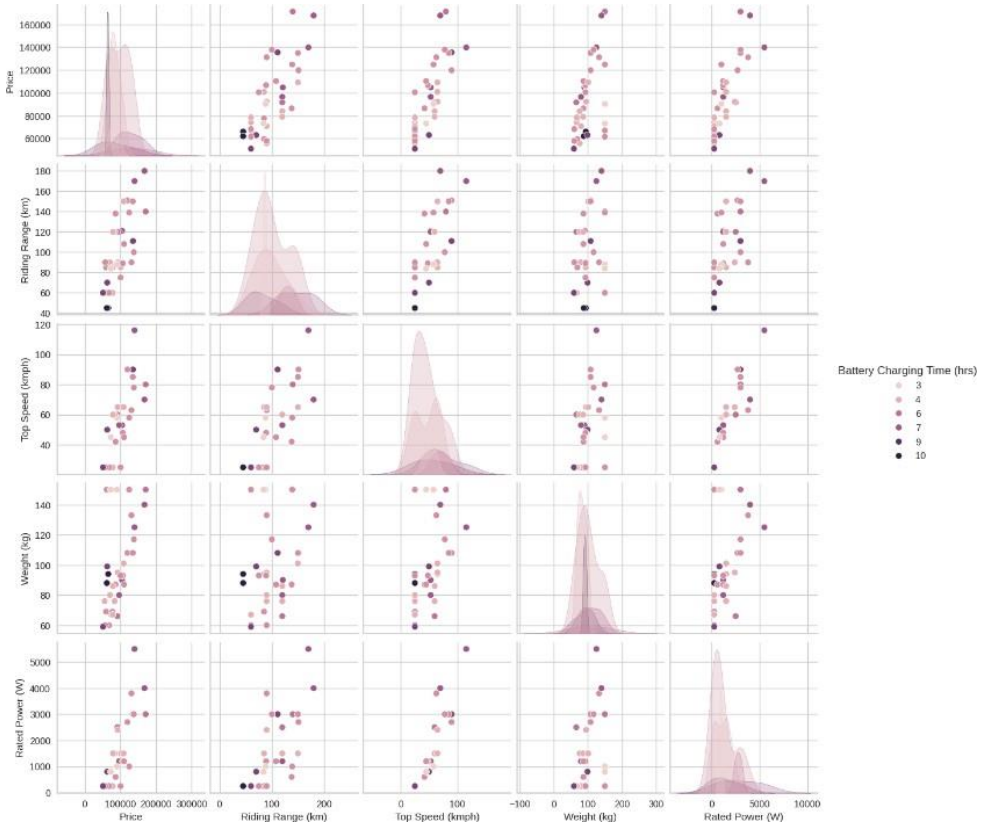


which bike has fastest acceleration





For Electric Vehicle Market one of the most important keys is Charging:

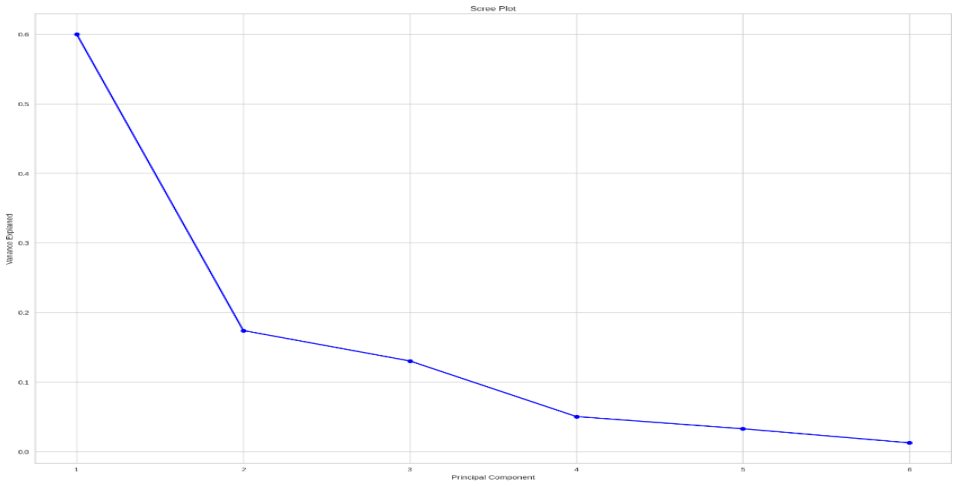


Correlation Matrix

A correlation matrix provides a tabular display of the correlations between variables, particularly effective for identifying linear relationships. The coefficients within the matrix quantify the correlations among different variables. Typically visualized using a heatmap, correlation matrices offer a clear depiction of relationships. A correlation coefficient exceeding 0.7 indicates a strong relationship between two variables.

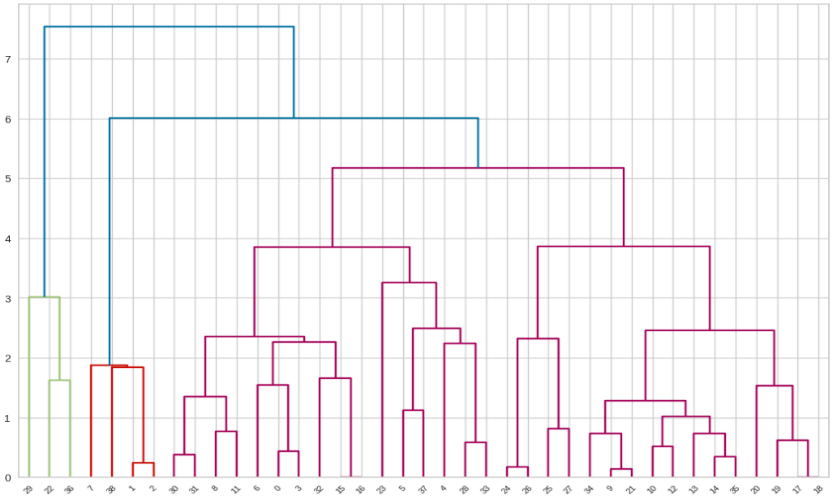
The Scree Plot serves as a visual aid in determining the optimal number of Principal Components (PCs) to retain. This plot is a simple line graph displaying the eigenvalues of each individual PC, with eigenvalues depicted on the y-axis and the number of factors on the x-axis. Typically, the plot exhibits a downward curve, starting high on the left, declining rapidly, and then plateauing at some point.

The Scree Plot criterion involves identifying the "elbow" in the curve, where the decline levels off. This point indicates the optimal number of PCs to retain. Additionally, the Proportion of Variance Plot is used to ensure that the selected PCs collectively account for at least 80% of the total variance in the data.

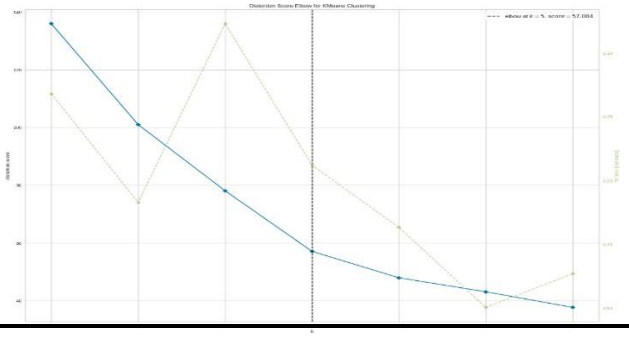


Extracting Segments

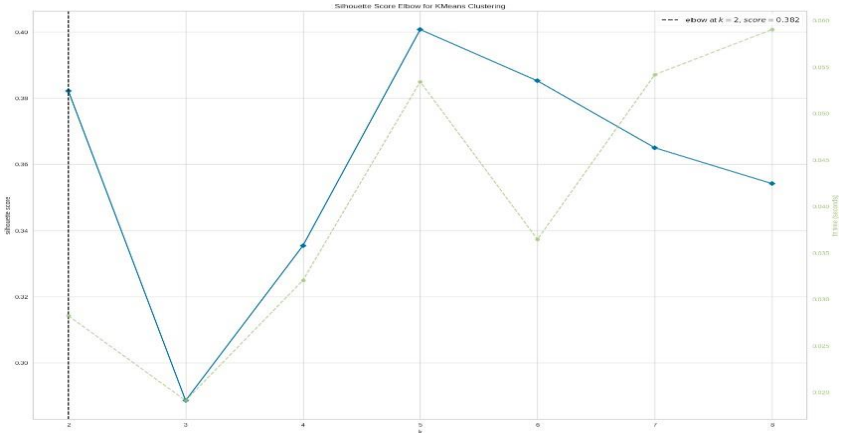
- ***Dendrogram***



The Elbow Method is a valuable technique tailored for the agglomerative hierarchical method of clustering. This method begins with each data point as its own cluster and progressively merges clusters based on their distances in a hierarchical manner. To determine the optimal number of clusters, we utilize dendrograms, which illustrate the sequence of cluster merges or splits. In line with other cluster validation metrics, hierarchical clustering often suggests considering four to five clusters for effective clustering. The Elbow Method calculates the Within-Cluster-Sum of Squared Errors (WSS) for different numbers of clusters (k) and identifies the point at which the change in WSS diminishes significantly. Additionally, it provides insights into the computational time requiredto generate models for various cluster numbers, represented by the green line. This approach aids in determining the appropriate number of clusters, facilitating meaningful segmentation of the data



Evaluating the clusters suing Distortion

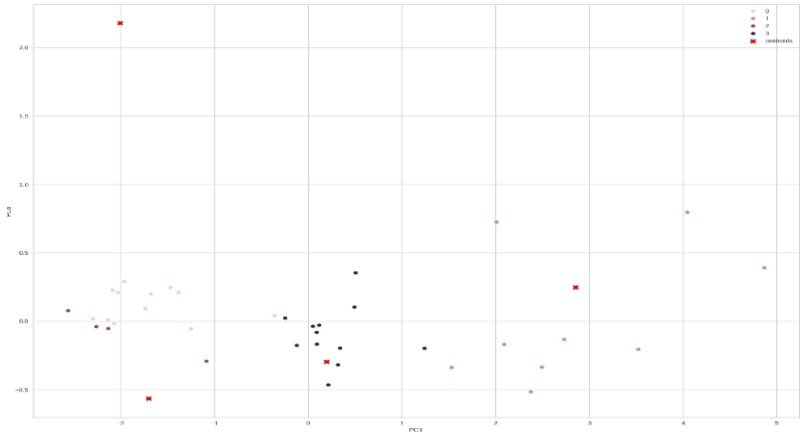


Evaluating the clusters suing silhouette

3. *Methodology:*

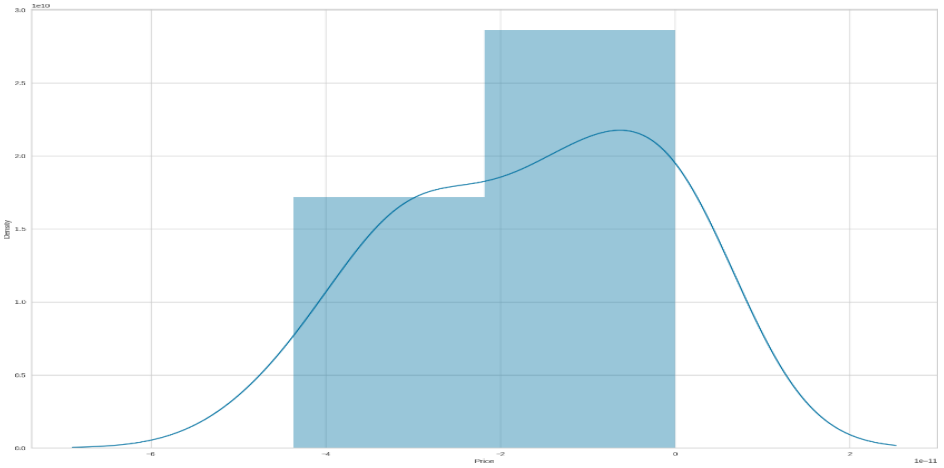
K-Means Algorithm

K Means algorithm is an iterative algorithm that tries to partition the dataset into pre-defined distinct non-overlapping subgroups where each data point belongs to only one group. It tries to make the intra-cluster data points as similar as possible while also keeping the clusters as different as possible. It assigns data points to a cluster such that the sum of the squared distance between the data points and the cluster's centroid is at the minimum. The less variation we have within clusters, the more homogeneous the data points are within the same cluster.



Prediction of Prices most used bikes

After completion of training the model process, we test the remaining 60% of data on the model. The obtained results are checked using a scatter plot between predicted values and the original test data set for the dependent variable and acquired similar to a straight line as shownin the figure and the density function is also normally distributed.



4. *Model Development:*

Training/testing data split.

LinearRegression(). fit(Xtrain,ytrain) command is used to fit the data set into model. The values of intercept, coefficient, and cumulative distribution function (CDF) are described in the figure

```
In [ ]: X=data2[['PC1', 'PC2','PC3','PC4','Pc5','PC6']]
        y=df['Price']

        X_train, X_test, y_train, y_test = train_test_split(X, y,test_size=0.4, random_state=101)
        lm=LinearRegression().fit(X_train,y_train)

        print(lm.intercept_)
```

```
94763.89743589745
```

```
In [ ]: lm.coef_
```

```
Out[ ]: array([ 14835.80424,    735.33714,    859.6533 ,    808.65528,
               -25866.34268,  -6076.03699])
```

```
In [ ]: X_train.columns
```

```
Out[ ]: Index(['PC1', 'PC2', 'PC3', 'PC4', 'Pc5', 'PC6'], dtype='object')
```

```
In [ ]: cdf=pd.DataFrame(lm.coef_, X.columns, columns=['Coeff'])
        cdf
```

Out[ ]:

	Coeff
PC1	14835.8042
PC2	735.3371
PC3	859.6533
PC4	808.6553
Pc5	-25866.3427
PC6	-6076.0370

5. *Results:*

Performance metrics of the models

```
In [ ]: print('MAE:',metrics.mean_absolute_error(y_test,predictions))
        print('MSE:',metrics.mean_squared_error(y_test,predictions))
        print('RMSE:',np.sqrt(metrics.mean_squared_error(y_test,predictions)))
```

```
MAE: 1.6370904631912708e-11
MSE: 4.698385879301197e-22
RMSE: 2.1675760377207524e-11
```

```
In [ ]: metrics.mean_absolute_error(y_test,predictions)
```

```
Out[ ]: 1.6370904631912708e-11
```

```
In [ ]: metrics.mean_squared_error(y_test,predictions)
```

```
Out[ ]: 4.698385879301197e-22
```

```
In [ ]: np.sqrt(metrics.mean_squared_error(y_test,predictions))
```

```
Out[ ]: 2.1675760377207524e-11
```

ehavioural: Mostly from our analysis there are Bikes with less range.

Demographic:

- **Top Speed & Range:** With a large area of market the cost is dependent on Top speeds and Maximum range of Bikes.
- **Efficiency:** Mostly the Bikes are with Fast charging speeds.

Psychographic:

Price: From the above analysis, the price range is between 1,50,000 to 50,000.

Finally, our target segment should contain Bikes with most Range, contains Top Speed and price between 1.5 to 1 lakh with mostly with 3 hours charging time.

6. *Conclusion:*

In summary, following an extensive analysis of the electric vehicle (EV) bike market in India, Segment 4 emerges as the prime target for our strategic focus. Boasting a substantial consumer base of 55.55%, this segment represents a lucrative market opportunity poised for exploration and expansion. Our plan involves tailoring the specifications of our electric bikes to align seamlessly with the preferences of Segment 4, thereby catering effectively to the demands of this sizable customer group. This strategic choice is informed by a profound understanding of consumer behavior and technical requirements, ensuring that our EV bikes are finely tuned to address specific market needs. Moreover, our approach underscores precision and relevance in both product development and marketing strategies. Through the strategic application of these insights, we are well-positioned to make impactful market entry decisions, establishing a robust presence in India's dynamic EV bike landscape. Looking ahead, this strategic foundation serves as a sturdy framework for success, empowering our EV bike offerings to resonate effectively and drive sustainable growth in the market.