Electric Vehicle Market Segmentation

Submitted By

Mohammad Sameer

1. OBJECTIVES

The objective of this report is to analyze market segmentation within India's electric vehicle (EV) sector to assist a start-up in making strategic decisions regarding:

1. The type of electric vehicle to manufacture.
2. The target customer segment to focus on.

Due to limited data availability, this study will primarily concentrate on selecting the appropriate type of electric vehicle. Identifying detailed demographic data for potential customers is beyond the scope of this research, as it would require highly specialized datasets. However, to gain insights into the ideal customer segment, the study will employ the 'benefits sought' segmentation approach, which highlights consumer preferences and expectations.

2. DATA COLLECTION

Due to lack of availability of data, we will be performing segmentation for the state of Telangana only. The original datasets can be found on:

https://data.telangana.gov.in/search/?theme=Automobile

From this portal we will be using:

1. Regional Transportation authority vehicle online sales data. (For the month of Oct-Nov 2023)
2. Regional Transportation authority vehicle online registration data. (For the month of Oct-Nov 2023)

Both of these datasets consisted of sales and registration data of all kinds of vehicles for the month of October and November in the state of Telangana. Particularly, the datasets consisted of following variables of interest:

1. Sales data: Model description, fuel type, vehicle class, make year, seat capacity, insurance validity, second vehicle or not, category, maker name.
2. Registration data: Body type, cc, cylinder, fuel, hp.

Since many vehicles in both of these datasets overlap, we will be combining both of these datasets to construct a final dataset with the following variables:

1. Company
2. Model
3. Fuel
4. Vehicle Class
5. Insurance Validity (in years)
6. Vehicle Age (in months)
7. Seats
8. Second Vehicle (Yes/No)
9. Category
10. Body Type
11. CC
12. Cylinders
13. HP

3. EXPLORATORY DATA ANALYSIS

The dataset has 8691 rows and 14 columns.

	S.No	Company	Model	Fuel	Vehicle Class	Insurance Validity (in years)	Vehicle age (in months)	Seats	Second Vehicle	Category	Body Type	CC	Cylinders	HP
0	1	Maruti Suzuki India Ltd.,	MARUTI - CIAZ SMART HYBRID SIGMA 1.5L 5MT BSVI...	Petrol Electric	Motor Cab	3	6	5	N	Transport	Sedan	1462	4	103.25
1	2	Maruti Suzuki India Ltd.,	MARUTI - CIAZ SMART HYBRID SIGMA 1.5L 5MT BSVI...	Petrol Electric	Motor Cab	3	6	5	N	Transport	Sedan	1462	4	103.25
2	3	Maruti Suzuki India Ltd.,	MARUTI - CIAZ SMART HYBRID SIGMA 1.5L 5MT BSVI...	Petrol Electric	Motor Cab	3	6	5	N	Transport	Sedan	1462	4	103.25
3	4	Tvs Motor Company Ltd	TVS IQUBE ELECTRIC S BOV	Battery	Motor Cycle	5	6	2	N	Non Transport	Solo	0	0	9.39
4	5	Tvs Motor Company Ltd	TVS IQUBE ELECTRIC SMARTXONNECT BOV	Battery	Motor Cycle	5	4	2	N	Non Transport	Solo	0	0	9.39

The categorical variables in the dataset are:

```
['Company',
 'Model',
 'Fuel',
 'Vehicle Class',
 'Second Vehicle',
 'Category',
 'Body Type']
```

Numerical variables are as follows:

```
['S.No',
 'Insurance Validity (in years)',
 'Vehicle age (in months)',
 'Seats',
 'CC',
 'Cylinders',
 'HP']
```

We will be analysing the data using bar charts and count plots. The graph underneath shows t -he count of companies, that registered sales in the month of October and November in the st- -ate. From the count plot we can see that the top players in the market are:

1. TVS Motors
2. Ola
3. Bajaj Auto
4. Tata
5. Hero
6. Quantum Energy
7. Pur Energy
8. Greaves Electric Mobility
9. MG Motors

Apart from them there are numerous other companies which occupy a small share of the market.
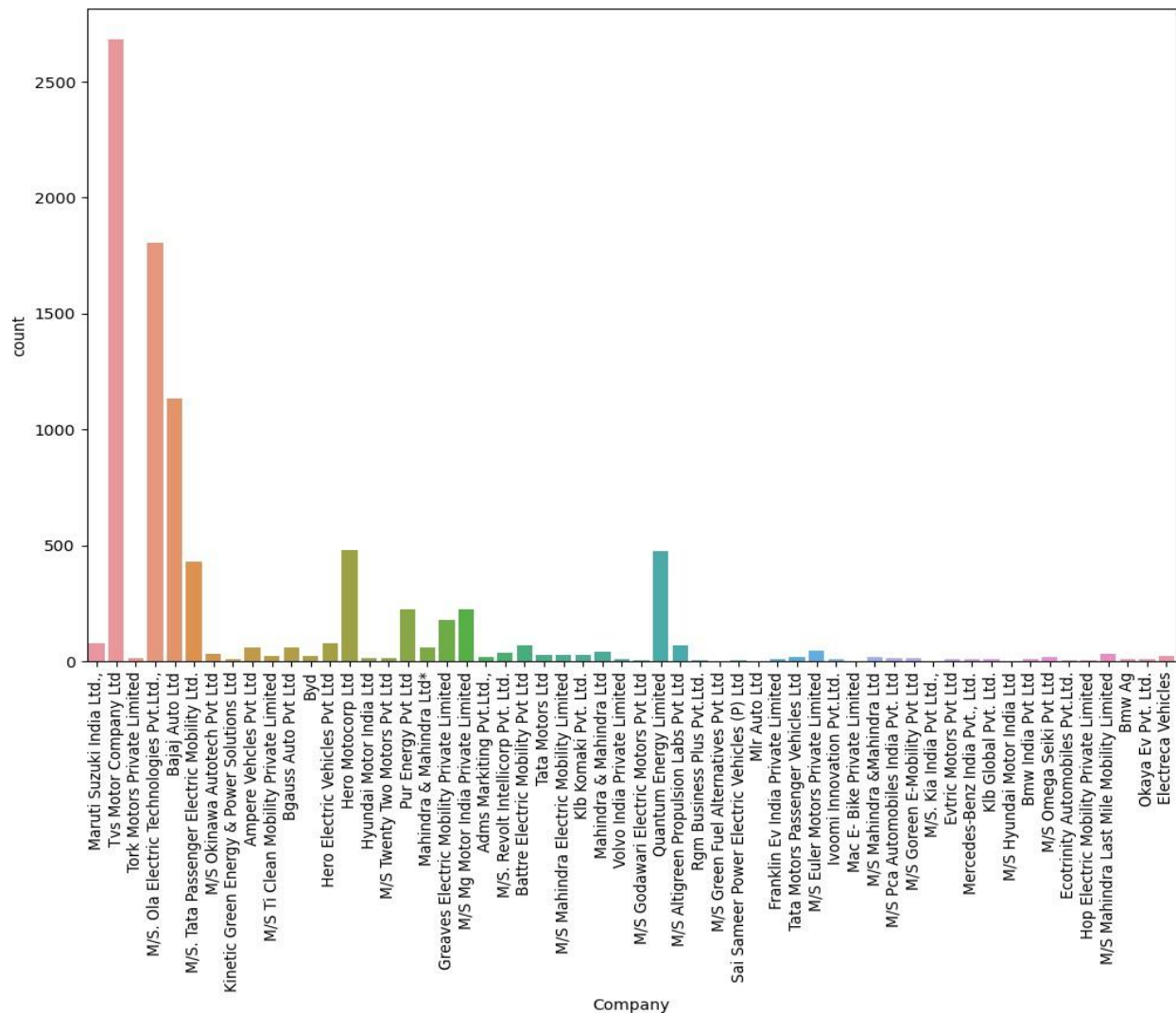
The top selling models in the market are:

1. TVS IQUBE ELECTRIC S BOV
2. OLA S1 AIR BOV
3. CHETAK 2413 PREMIUM BOV
4. TVS IQUBE ELECTRIC SMARTXONNECT BOV
5. OLA S1 PRO (GEN2) BOV
6. VIDA V1 PRO BOV
7. TVS IQUBE ELECTRIC BOV
8. BZINESS PRO BOV
9. MAGNUS EX BOV
10. BZINESS LITE BOV

The market is dominated heavily by battery operated vehicles. There are a very few vehicles that are hybrid in nature (i.e. Petrol Electric).
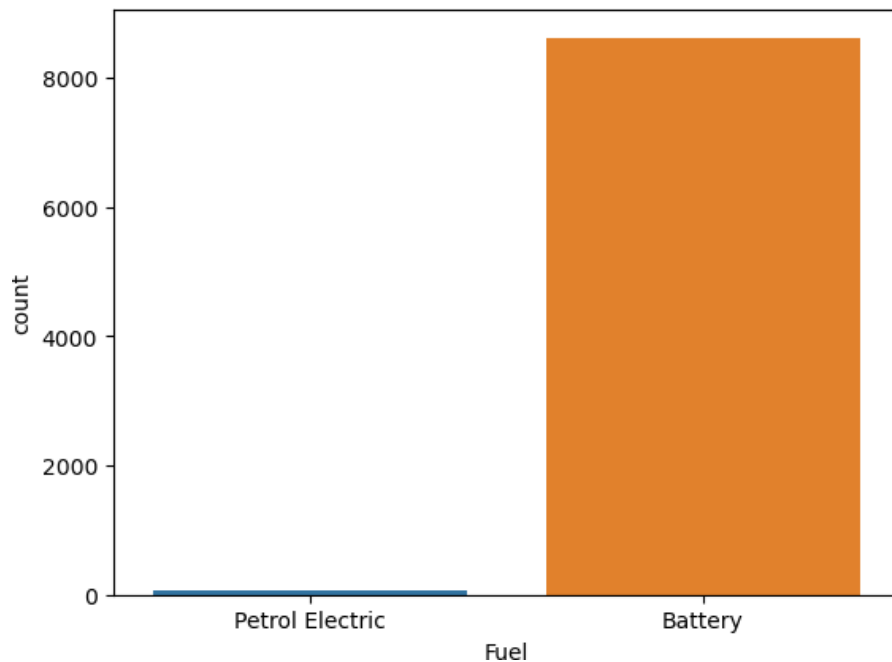

NOTE: Another significant fuel type that is popular in the EV market is the Diesel Electric type, but due to unavailability of vehicle specifications for such fuel type, we did not include it in our dataset.

The next plot shows vehicle class of the different EVs currently in use. The most popular vehicle class is Motor Cycle.
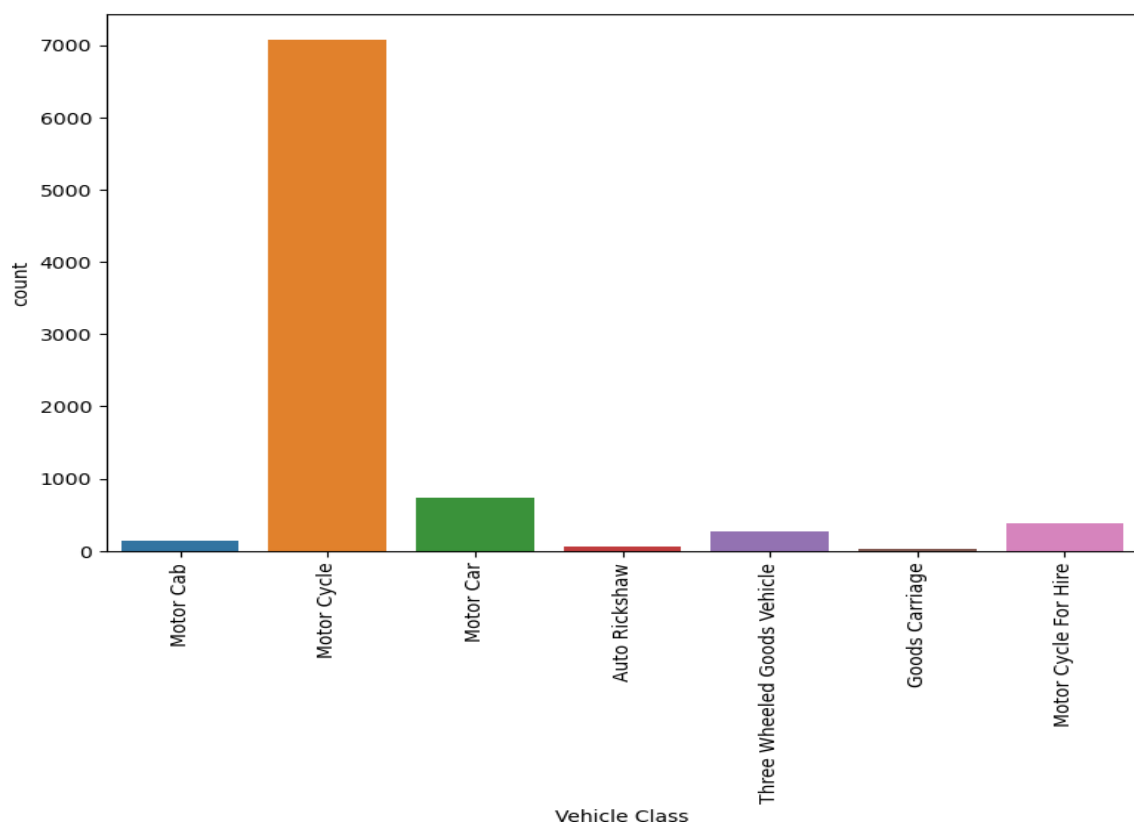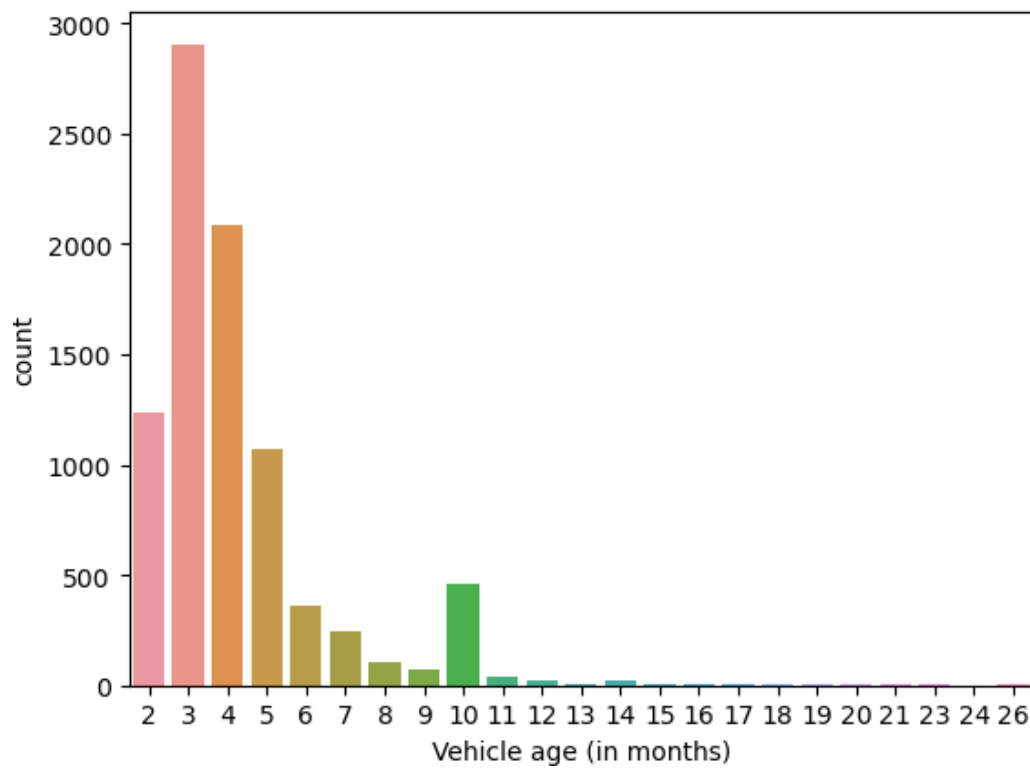
This pattern is supported by the seat capacity of the on-road vehicles. Most vehicles have a seat capacity of two.
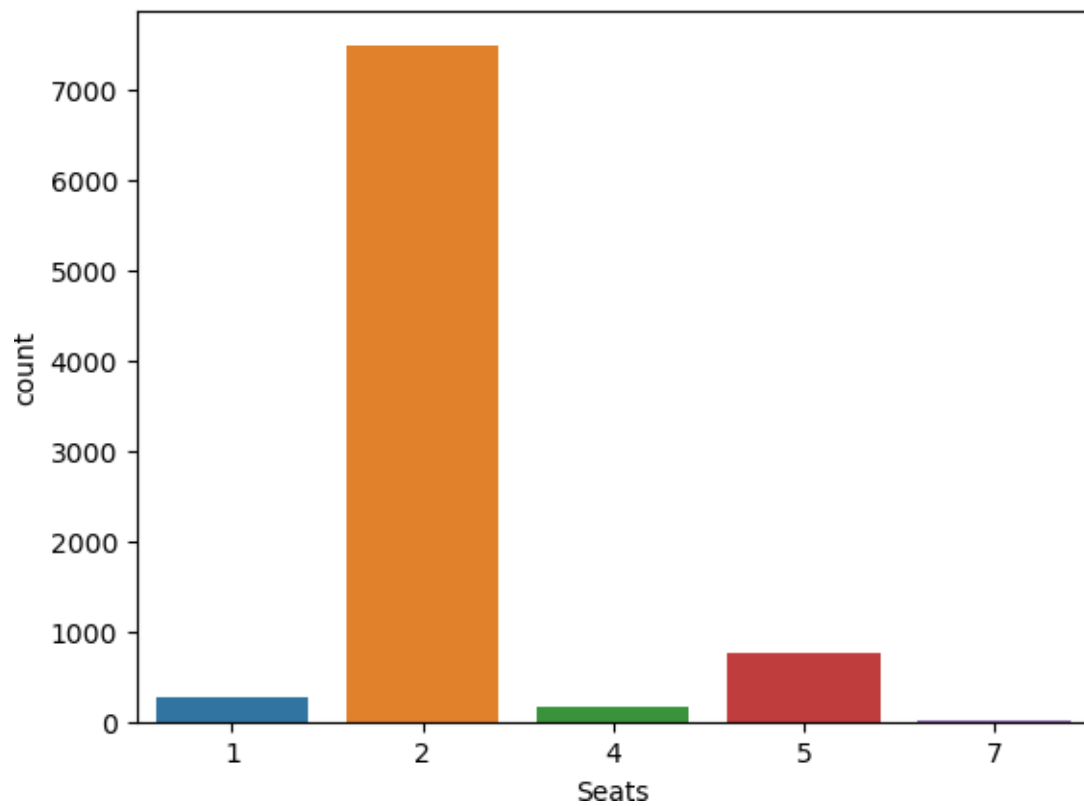
Insurance Validity of 5 years is also the most sought-after insurance term duration among EV purchasers. Any other length of insurance validities not preferred much.
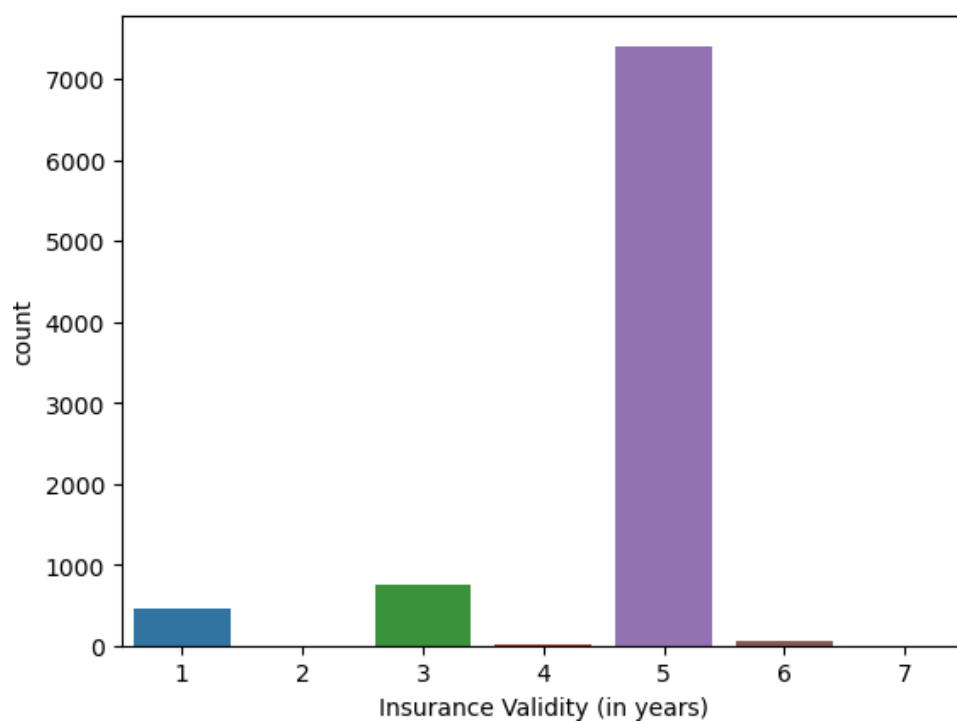
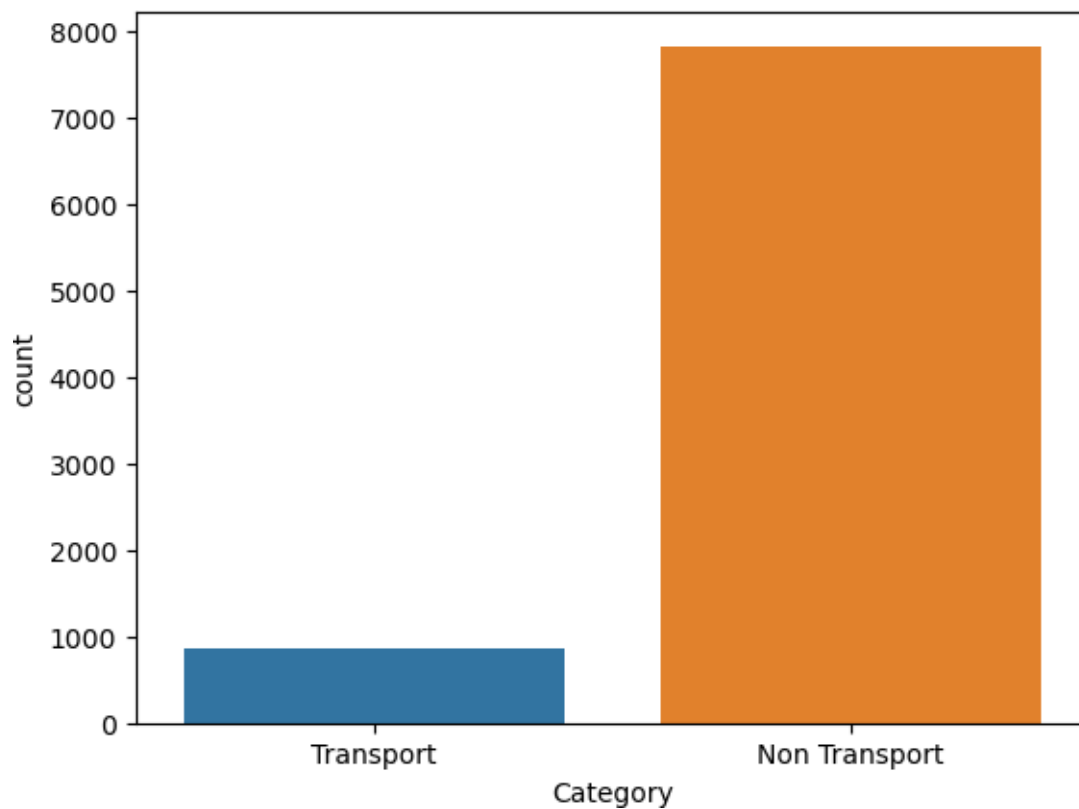EV Purchasers are also very less likely to purchase older vehicles:

Vehicles older than a year, have very few numbers of willing customers. This pattern is reflected in the following plot as well. There are a very few buyers of second hand EVs.
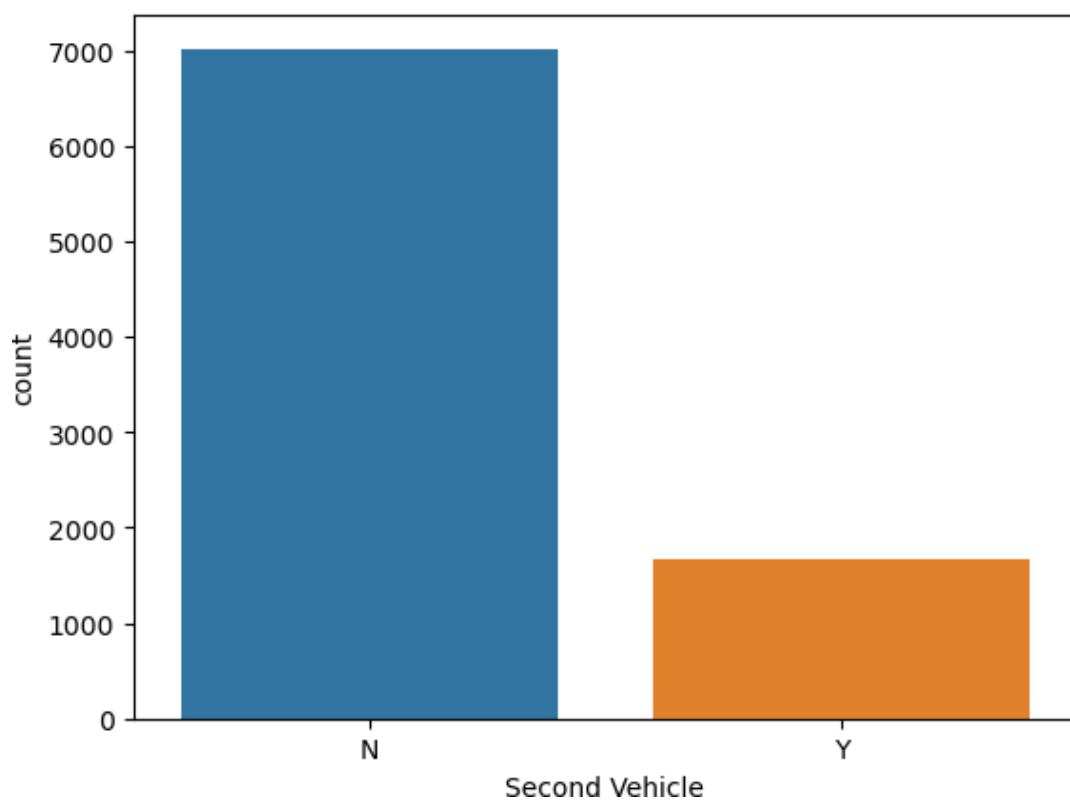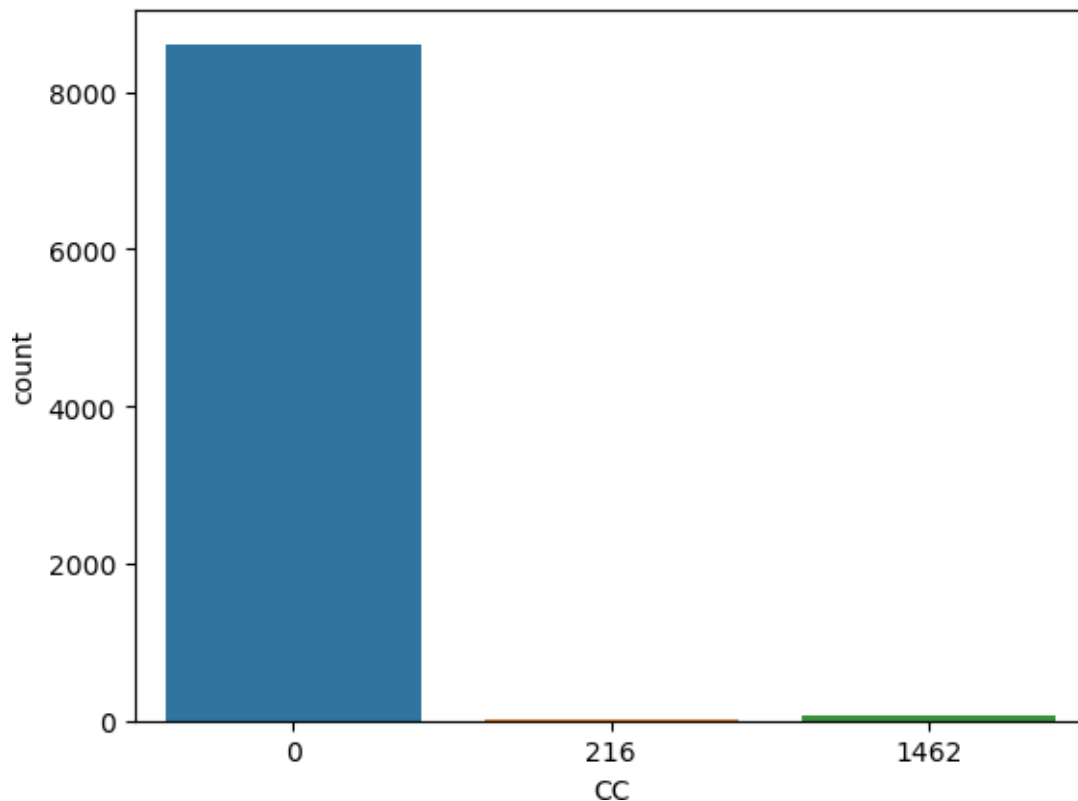
The major part of the EVs fall under the non-transport category:

The most popular body type is solo, this is because the most popular type of EVs in the market are motorcycles.
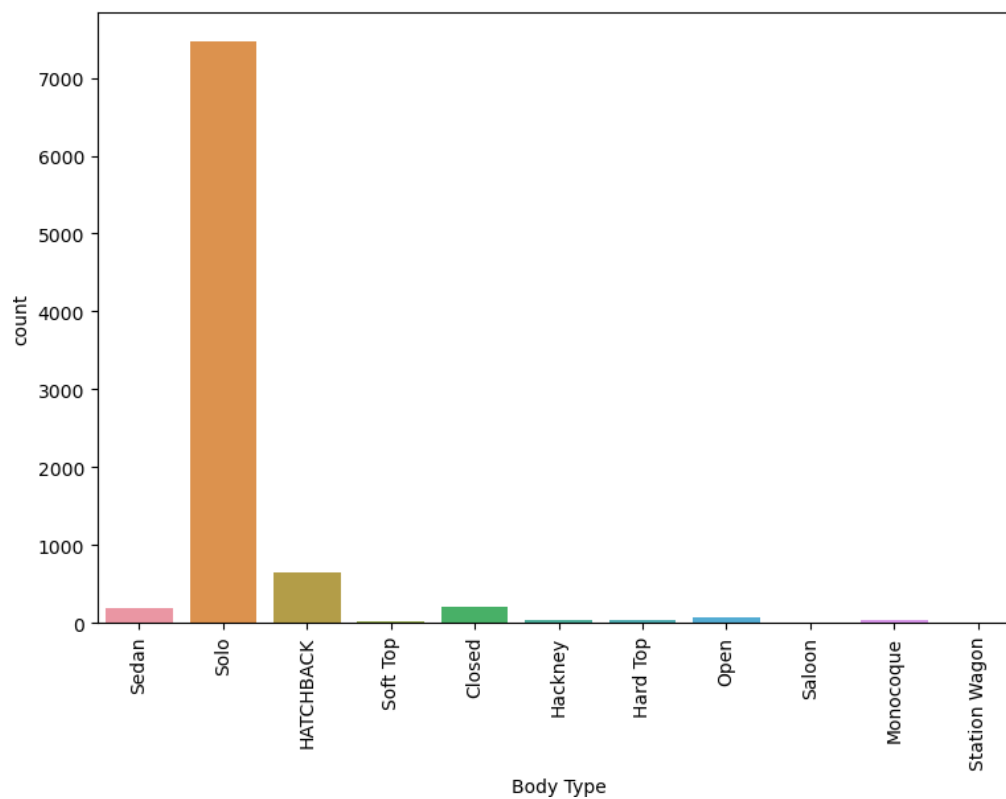
As we saw earlier, there are a very few hybrid vehicles in the market. Therefore the 'CC' of most of the vehicles is zero.

4. DATA PREPROCESSING

Variables such as 'Serial Number', 'Company' and 'model' are not required for the segmentation algorithm and therefore we will be dropping them. Apart from them, we need to use one hot encoding for the following categorical variables:

1. Fuel
2. Second Vehicle
3. Vehicle Class
4. Category
5. Body Type

The dimensions of our data therefore become 8691 rows and 25 columns. It is generally advisable to standardize our data before proceeding for the segmentation. So, we will be using the standard scaler to standardize our data. We are now ready to extract our segments.

5. SEGMENT EXTRACTION

For our case, we will be using an unsupervised machine learning technique called K-Means Clustering to extract our segments. But the algorithm requires the user to specify the number of clusters (k) that he/she wants to extract. There are various methods of determining the

ideal number of clusters. We will be using the elbow method here. For a specified range of clusters between (2 and 10), we have the following results.

Distortion Score Elbow for KMeans Clustering
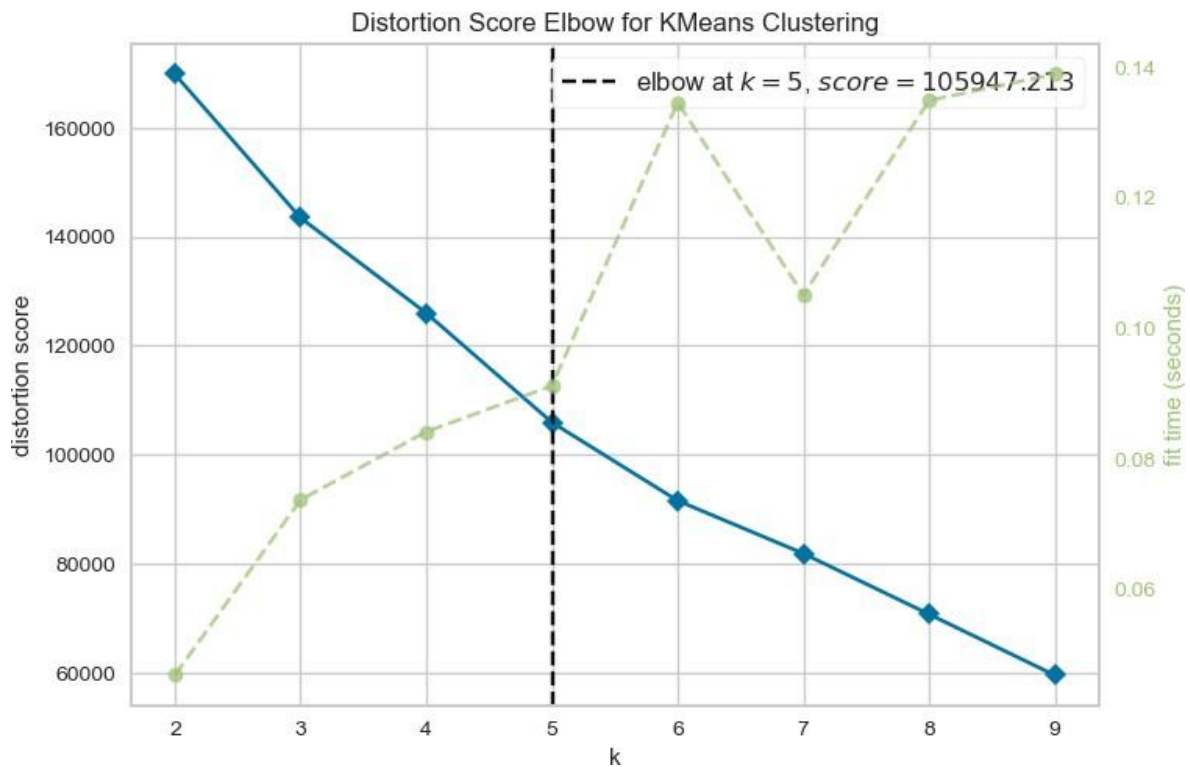
So, the elbow occurs at k = 5, therefore the ideal number of clusters would be 5. We can now initiate the algorithm. The next step is to visualize these clusters on a graph. Since, there are around 25 variables, it is not possible to represent it in 2D space. So, we will be using Principal component analysis to show the clusters on a 2D space.

Visualizing Clusters in Two Dimensions Using PCA
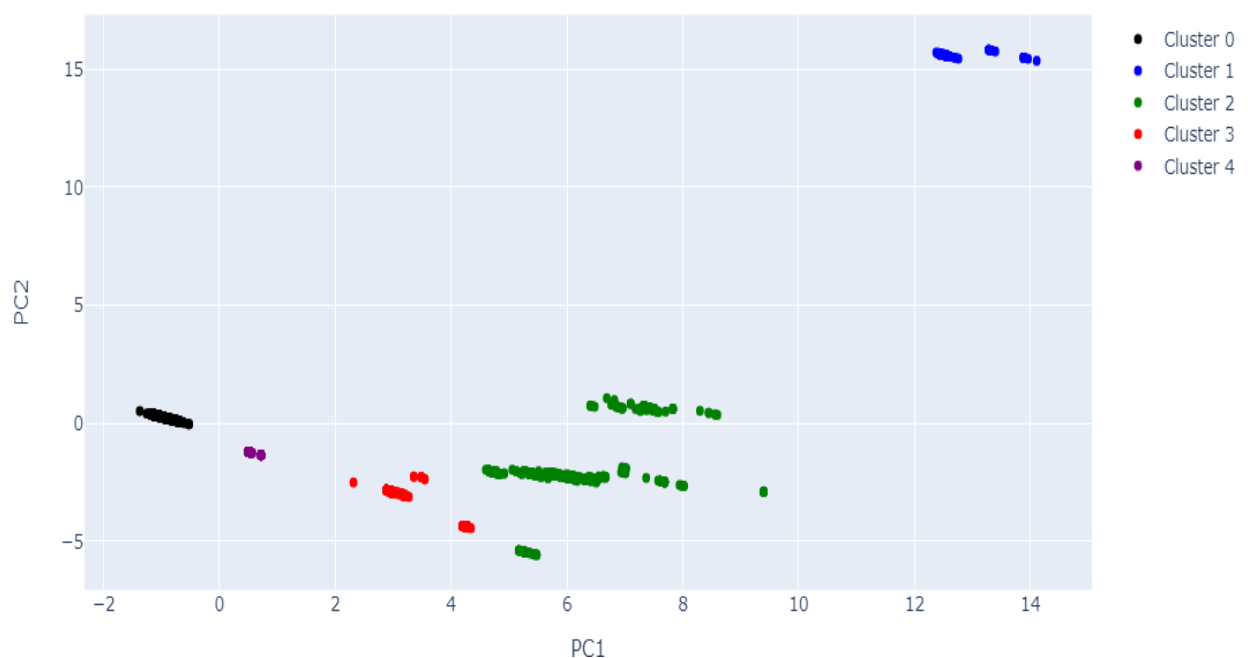
Since we have cluster labels for each data we can attach those labels back onto our original unprocessed data, to analyze each segment. We therefore have the following five segments with the following dimensions:

1. Segment 1 (7078 rows and 14 columns)
2. Segment 2 (76 rows and 14 columns)
3. Segment 3 (857 rows and 14 columns)
4. Segment 4 (294 rows and 14 columns)
5. Segment 5 (386 rows and 14 columns)

6. ANALYSIS OF EACH SEGMENT

	SEGMENT1	SEGMENT2	SEGMENT3	SEGMENT4	SEGMENT5
Share in total market size	81%	0.87%	9.8%	3.38%	4.44%
Top Players with their respective market shares	TVS (37.8%) Ola (25.48 %) Bajaj (15.59%) Hero Motocorp(6.78%) Pur Energy (3%) Greaves Electric2.543%) Quantum Energy (1.24%)	Maruti Suzuki (100%)	Tata (50.18%) Mg Motor (26.37%) Mahindra & Mahindra (4.67%) Ti Clean Mobility (2.80%) Byd (2.57%) Mahindra Electric Mobility (2.33%) Tata Passenger Vehicles (1.98%) Pca Automobile (1.63%) Hyundai (1.40%) Mahindra &Mahindra (1.28%) Mercedes-Benz (1.28%)	Altigreen (23.13%) Mahindra & Mahindra (18.71%) Euler Motors (15.65%) Mahindra Last Mile Mobility (10.54%) Bajaj (10.20%) Tata (9.8%) Omega Seiki (6.8%) Mahindra Electric Mobility (2.38%) Mahindra &Mahindra (2.041%) Mlr Auto (0.68%)	Quantum Energy (100%)
Fuel Type	Battery	Petrol Electric	Battery	Battery	Battery
Vehicle Classes present	Motorcycle	Motor Cab, Motor Car	Motor Car, Auto rickshaw, Motor Cab	Three Wheeled goods vehicle, Goods carriage	Motor Cycle for hire
Insurance Validity	5 years (99.18%) 6 years (0.08%) 7 years (0.014%)	3 years (88.15%) 1 year (11.84%)	3 years (80.05%) 1 years (18.90%) 4 years (0.93%) 2 years (0.12%)	1 years (99.66%) 3 years (0.34%)	5 years (100%)
Vehicle age (in months)	3(37.67%) 4(22.76%) 2 (17.067%) 5 (11.66%) 6 (3.60%) 10 (3.31%)	6 (23.68%) 4 (22.37%) 7 (19.74%) 5 (15.79%) 8 (6.58%) 3 (5.26%)	4 (24.74%) 3 (21.94%) 7 (15.17%) 5 (10.85%) 6 (8.17%) 8 (7.35%)	4 (46.26%) 3 (15.98%) 5 (14.29%) 10 (10.88%) 6 (4.76%) 9 (2.38%)	10 (45.59%) 4 (28.49%) 5 (25.38%) 6 (0.52%)

	7 (1.37%) 8 (0.48%)	10 (2.63%) 9 (1.32%) 11(1.32%) 12 (1.32%)	9 (4.78%) 2 (3.38%) 10 (1.75%) 14 (0.58%) 13 (0.35%) 15 (0.35%) 12 (0.23%) 23 (0.12%) 11 (0.12%) 18 (0.12%)	8 (2.38%) 7 (1.36%) 11 (1.36%) 12 (0.34%)	
Seat Capacity	2	5	5 (79.69%) 4 (19.60%) 7 (0.70%)	1 (90.14%) 2 (9.86%)	2
Second hand vehicle?	NO (78.64%) YES (21.36%)	NO (86.84%) YES (13.16%)	NO (82.15%) YES (17.85%)	NO (100%)	NO (100%)
Category	Non-Transport 100%	Non-Transport 75% Transport 25%	Non-Transport (80.047%) Transport (19.95%)	Transport 100%	Transport 100%
Body type	Solo	Sedan	HATCHBACK (74.67%) Sedan (12.37%) Hard Top (5.13%) Hackney (3.62%) Soft Top (2.80%) Saloon (0.82) Station Wagon (0.58%)	Closed (67.69%) Open (22.12%) Monocoque (10.2%)	Solo
CC	0	1462	0	0 (97.62%) 216 (2.38%)	0
HP	Majorly between 6HP to 10HP	Entirely between 103.2 HP and 103.4 HP	Majorly between 0HP and 200HP	Majorly between 10 HP and 20 HP	Majorly between 1.6 HP and 2 HP

Since the segmentation was not performed for analyzing customer perception or behavior regarding an existing product, and rather for deciding which vehicle to produce and then which consumers to target; we will first have to decide that first and then analyze customer preference. To do that we can divide the segment descriptors into two categories

1. Those which would allow us to decide which vehicles to produce
2. Those which help us understand customer preferences for that chosen segment.

Category 1 (For choosing vehicle)	Category 2 (For understanding consumer preferences)
1. Fuel 2. Vehicle Class 3. Body type 4. Category	1. HP 2. CC 3. Second hand vehicle 4. Vehicle age 5. Insurance Validity

Now the ideal market segment which our hypothetical company would choose should have the following characteristics:

1. Large Enough
2. No monopoly
3. Profitable
4. Financially feasible to enter.

Using these criteria, we can evaluate each segment:

1. Segment 1:
   a. **Largest Segment**
   b. **Competitive**
   c. **Profitable – Because this segment produces motorcycles which are the most common type of EV in India.**
2. Segment 2:
   a. **Smallest Segment**
   b. **Monopoly of Maruti**
   c. **Potential of profit**
3. Segment 3:
   a. **Second largest segment**
   b. **Competitive**
   c. **Profitability higher than segment 2 because of larger size and more vehicle classes.**
4. Segment 4:
   a. **Second smallest segment**
   b. **Competitive**
   c. **Low Profitability because Three Wheeled goods vehicle, Goods carriage have a very few numbers of buyers.**
5. Segment 5:
   a. **Third largest (or third smallest) segment**
   b. **Monopoly**
   c. **Profitability ambiguous because motorcycles for hire are not very popular.**

So, we can safely remove segments 2, 4 and 5 from our study. We will focus only on Segment 1 and 3. Let's compare those characteristics where these segments actually differ:

	SEGMENT1	SEGMENT3
Share in total market size	81%	9.8%
Vehicle Classes present	Motorcycle	Motor Car, Auto rickshaw, Motor Cab
Insurance Validity	5 years (99.18%) 6 years (0.08%) 7 years (0.014%)	3 years (80.05%) 1 years (18.90%) 4 years (0.93%) 2 years (0.12%)
Vehicle age (in months)	3(37.67%) 4(22.76%) 2 (17.067%) 5 (11.66%) 6 (3.60%) 10 (3.31%) 7 (1.37%)	4 (24.74%) 3 (21.94%) 7 (15.17%) 5 (10.85%) 6 (8.17%) 8 (7.35%) 9 (4.78%)

	8 (0.48%)	2 (3.38%) 10 (1.75%) 14 (0.58%) 13 (0.35%) 15 (0.35%) 12 (0.23%) 23 (0.12%) 11 (0.12%) 18 (0.12%)
Seat Capacity	2	5 (79.69%) 4 (19.60%) 7 (0.70%)
Category	Non-Transport 100%	Non-Transport (80.047%) Transport (19.95%)
Body type	Solo	HATCHBACK (74.67%) Sedan (12.37%) Hard Top (5.13%) Hackney (3.62%) Soft Top (2.80%) Saloon (0.82) Station Wagon (0.58%)
HP	Majorly between 6HP to 10HP	Majorly between 0HP and 200HP

Now both of these segments have their pros and cons:

1. Segment 1 is larger while segment 3 is smaller
2. Segment 1 produces Motorcycles which are most popular EVs in India.
3. It is difficult to sell vehicles older than a year in Segment 1, while its relatively easier in segment 3.
4. Segment 1 consumers also prefer longer insurance validity than segment 3, which can be difficult for the company.
5. The other characteristics such as body type and HP are natural consequences of the type of vehicle that one chooses so that is not a determining criterion.

So, depending upon what priorities and financial capabilities the company has, it can choose either of the two segments and customize their vehicles accordingly.

7. Conclusion

Although the available data was extensive, it lacked the specificity required for conducting a market segmentation study that could yield actionable insights in a real-world business context. The dataset was missing several critical variables necessary for a comprehensive analysis, including:

1. **Customer Income Levels:** Understanding the income distribution of potential customers is crucial for determining affordability and aligning product pricing with market demand. Without this data, it becomes challenging to identify which income segments are most likely to adopt electric vehicles (EVs).

2. **Vehicle Pricing Information:** Detailed price points across various electric vehicle models are essential to assess market competitiveness and consumer price sensitivity. This data would enable pricing strategies tailored to different market segments.

3. **Consumer Usage Patterns:** Insights into how consumers intend to use electric vehicles—such as daily commuting, long-distance travel, or commercial purposes—are vital for product development and positioning. This information informs design choices, battery range, and feature prioritization.

4. **Detailed Vehicle Specifications:** Technical features like battery capacity, charging speed, safety features, and smart technologies significantly influence consumer preferences. A deeper understanding of these specifications allows for better product differentiation in a competitive market.

Access to customized datasets containing these variables is limited and typically available through premium data platforms like Statista, which require a subscription. Utilizing data from such sources could significantly enhance the depth and accuracy of the market segmentation analysis.

In terms of analytical methods, while K-Means Clustering is a widely used technique for segmenting markets, relying solely on it may overlook complex market structures. Incorporating additional clustering methods could provide a more nuanced understanding of the market:

- **Hierarchical Clustering:** This method builds a tree-like structure of clusters, allowing for the identification of nested groupings and relationships between market segments. It is especially useful for discovering subtle patterns within data.

- **Density-Based Clustering (DBSCAN):** This technique identifies clusters of varying shapes and sizes based on data density, effectively handling noise and outliers. It is well-suited for uncovering non-linear relationships and irregular market patterns.

By integrating multiple clustering approaches, the analysis can deliver more robust and actionable market segmentation, providing deeper insights into consumer behavior and preferences.