Stability of Electric Vehicle in Indian Market

IBM-306 Marketing Research Project

Utkarsh Aashu Mishra	
Tejas Patil	17117093
Tarun Kumar	17117089
Lavkush Verma	17117087
Prakhar Mishra	17117045
	17117058

Abstract

The Indian Automobile industry is currently the 5th largest in the world and is set to take over as the 3rd largest automobile industry by 2030. Currently, there has been a lot of initiatives taken by the government in order to curb pollution caused due to automobiles. Electric Vehicle market can be seen as a new entry into the sector. The study is dedicated to the analysis of the stability and sustainability of this new market in a country like India. Various factors are considered by the concerned students of the group, importance of which were analyzed after a successful data collection through surveys and field visits. A simple exploratory probabilistic research design was constructed and analysis was done on the basis of demographics, knowledge on EVs, their opinion towards the factors and their expected satisfaction. Interesting results were observed with the factor and cluster analysis of the data obtained. Moreover, a linear regression analysis was also performed in order to estimate the willingness of the customers to buy an Electric Vehicle on the basis of their responses on the opinions. Finally, an Artificial Neural Network model was implemented with the available data to classify effectively the customers into their willingness to accept or buy an EV. Apart from information on the EV market, opinions were also collected and analysis was performed about their thoughts of accepting more public vehicles to curb out pollution.

Index

1. Introduction

The Indian government has set up an ambitious target of having only electric vehicles being sold in the country. The Indian auto industry is expected to see an 8-12 percent increase in its hiring during FY19. In order to promote the sale of electric vehicles in the Indian market, the government launched **FAME scheme (Faster Adoption and Manufacturing of Hybrid and Electric vehicles)** in India, as a part of the National Electric Mobility Mission Plan 2020, under which the government would provide certain incentives to lower the purchasing cost of electric vehicles The scheme has 4 focus areas i.e. Technology Development, Demand Creation, Pilot Projects and Charging Infrastructure. Overall, the government is expected to spend around Rs. 14,000 Crores for this scheme, which includes incentives to customers for purchasing electric vehicles, incentives to the manufacturers for research and development besides developing the charging infrastructure.

India's planned EV journey, on the other hand, has many dead ends. Instead of developing EVs, automakers in India have been busy recalibrating vehicles to meet the Bharat Standard-VI (BS-VI) emission norms. By 2020, India will switch to BS-VI standards to cut vehicular emissions. These are in line with European guidelines. Manufacturers who have already pumped in money to meet this deadline have little incentive to rev up on EVs, which needs billions more in investment and thousands of engineering hours, despite no assured demand. There is the question of creating a market for EVs. Even those who plan to make such vehicles are reluctant participants because of the high costs and lack of profit. India, though, is a value-conscious market and the bulk of the vehicles on its roads are entry-level ones where the price is everything.

In a country that is a power deficit, the current EV plan doesn't seem to be the right option. **India's EV policy aims to cut emissions, even though more than 70% of electricity is generated from conventional sources such as coal, fossil fuels, oil, petroleum, and natural gas.** With thermal power as the major source of power production, more will be needed to seamlessly charge the batteries for electric vehicles, thereby reducing the credibility of India's emission-reduction goals.

GOI is hoping consumers will overcome cost hurdles through subsidies it offers. That may be so. But it's not possible to race past inadequate infrastructure. **For EVs to run smoothly, India will need assured excess power supply that can be fed to charging stations throughout the country.**

In the meantime, if the government is serious about reducing emissions and encouraging cleaner vehicles, it needs to vastly **improve public transport and shared mobility so that fewer commuters opt for private vehicles**. It must curb the ownership of vehicles per household and make it prohibitively expensive to own more than one vehicle. It should also **incentivize remote working where possible, cutting back on the need to travel.**

2. Research Process

Various factors were identified which were needed to be addressed through this study. Price plays a vital role in the acceptance of electric vehicles by consumers in a country like India. Moreover, inadequate infrastructure for charging locations as well as charging time makes Electric Vehicles non-preferable for long-distance journeys as compared to Combustion engine vehicles. But, electric vehicles address the most significant problem of increasing pollution.

GOI has made many incentives as discussed in the previous section but Electric Vehicle is not the only way to reduce such pollution. Shifting more to public vehicles will serve as a good argument for this concern. With all these significant characteristics of the introduction of Electric Vehicles in the Indian Market, a deductive research model will be developed in the course of the study. The concerned null and alternative hypotheses developed are as follows:

Null Hypothesis, H_0:

1. **There will be no increase in the demand with the introduction and increasing Electric Vehicle supply.**
2. **Also people will not try to shift more to public vehicles to curb pollution.**

Alternate Hypothesis, H_a:

1. **There will be a high positive response on the demand in such a new market of Electric Vehicles with a good satisfaction to people's need as well as pollution.**
2. **Also, people will try to shift to more public vehicles in order to curb pollution, unless a proper infrastructure for smooth EV running is established.**

3. Survey Conducted

This research is supported by a **customer preference survey** which is an inclusive survey, where despite its small sample size, has incorporated **people's thoughts from different backgrounds**. The **word of mouth of vehicle owners** has been taken separately to better analyze the customer preference and situation. Special attention has been paid for this inclusiveness of the survey.

This research has focused on key factors, including **policies, charging infrastructure and vehicle attributes**. Moreover, the model has considered those factors from the consumer preference view. **Secondary data** has been obtained from governmental and non-governmental organizations. Primary data has been obtained from **field visits and questionnaire surveys** to estimate different parameters of our model.

This research is also focused on the **manufacturers of EVs and power supply companies**, the key factors for an automotive company to manufacture EVs are the **government incentives to production cost, serviceability, marketing** and for power supplying companies **are demand, change in power distribution system, infrastructure development**.

But we were unable to collect data from Companies due to difficulty in contacts with executives and time constraints.

3A. Consumer Survey

This survey was dedicated to the study of consumer behavior in facilitating the growth of Electric Vehicles in India. The survey was conducted in the form of a questionnaire as well as in-depth interviews with groups of known people.

Main Focus Group: The active working population as well as the younger generation who are going to join the working population in the next 2-5 years and their demographics were identified. Owning a vehicle and average time spending on travel were a few of the parameters used for clustering the type of responses.

Basic Information: Awareness of Electric Vehicles among the respondents as a scaled measure was introduced such that a suitable weight can be given to their

responses. An example can be, an individual with a good knowledge of EVs will be able to give more significant responses than an individual with less knowledge of EVs. For such cases, it was also asked if they believe that in India there exists sufficient technology to manufacture an affordable EV. The actual awareness as a combination of these two factors can be verified.

Information on Various Factors: Importance was given to the analysis of various factors that may encourage them to buy an EV or may restrict them to do so. Finally, their plans regarding buying an EV was collected. In order to take the public opinion on another vital component of our study, they were also asked whether an increase in the use of public vehicles rather than private ones will help in the reduction of the level of pollution.

4. Sampling Technique

The main objective of the study revolves around the **market sustainability of Electric Vehicles in India**. This market stability is controlled by a lot of factors:

1. The interest of automobile producing companies in introducing EVs,
2. The efficiency of Electricity supplying companies in setting up appropriate infrastructure for charging purposes, and finally
3. The consumer behavior in acceptance or rejection of such a new market.

With all these factors in mind, the target population comprises of respondents with **potential consumer capacities** i.e. the active working population as well as the younger generation who are going to join the working population in the next 2-5 years.

The sampling frame was created on the requirements of the research. The list included the **desired age groups** as well as their **potential to be a suitable consumer**. Moreover, as contacts with the company executives was a difficult task due to time constraint we couldn't get the responses from them.

A **probabilistic sampling technique** is used for the consumers and a **non-probabilistic sampling technique** is for the executives of the automobile and electricity-supplying companies.

For the consumers, **Simple Random Sampling** is used as it is a trustworthy method of obtaining information and for the automotive and electricity supplying

companies, **Judgmental or Purposive Sampling** is used the respondents are selected solely on the basis of requirements for the research.

Sample size was selected on the basis of confidence level where,
Confidence level corresponds to a **Z-score** as follows:

90% confidence – Z Score = 1.645
95% confidence – Z Score = 1.96
99% confidence – Z Score = 2.576

For our study, we had chosen a **95% confidence level, .5 standard-deviation, and a margin of error (confidence interval) of +/- 5%.**

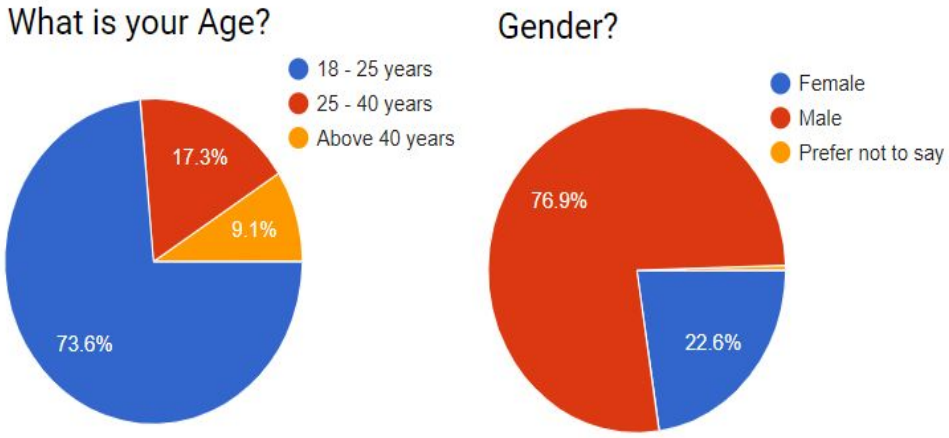
Thus, from formula

$$n = \frac{Z^2\, s\,(1-s)}{\alpha^2} = \frac{(1.96)^2\, 0.5\,(1-0.5)}{(0.05)^2} = 384.16 \approx 385$$

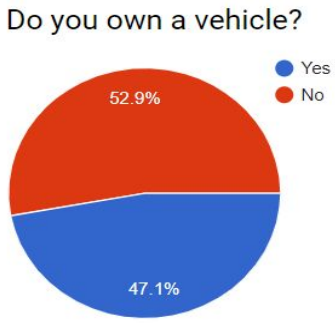
On the basis of the above criteria total **385 responses** are required for the survey. We were able to obtain **416 responses from the consumer survey.**

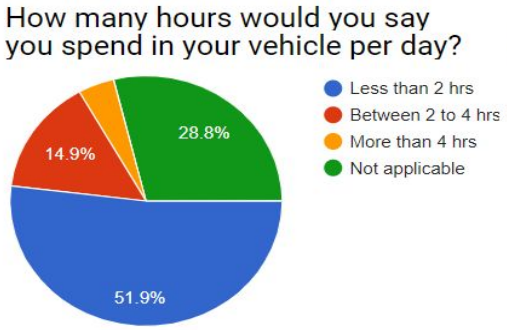
5. Response Collection and Analysis

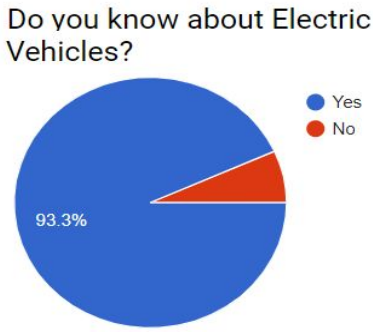
5 A. Responses based on Demographics:



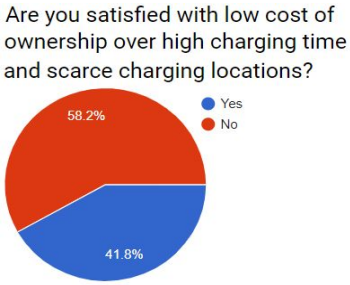
5 B. Responses based on Vehicles and Electric Vehicles:

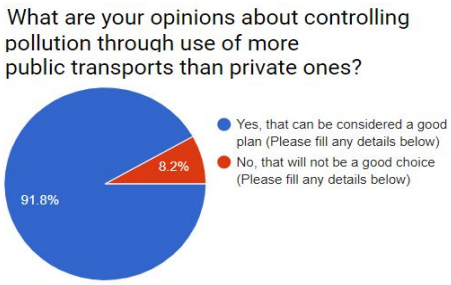


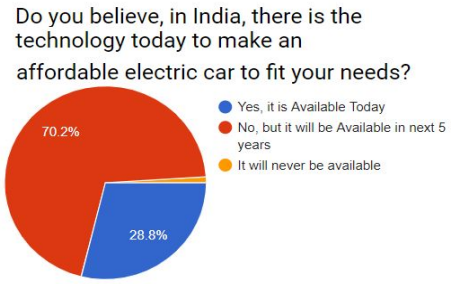


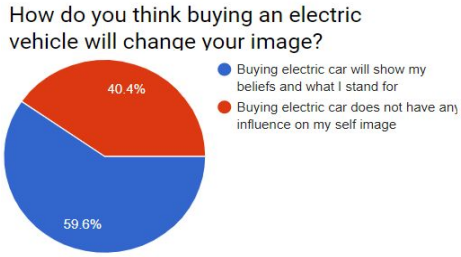


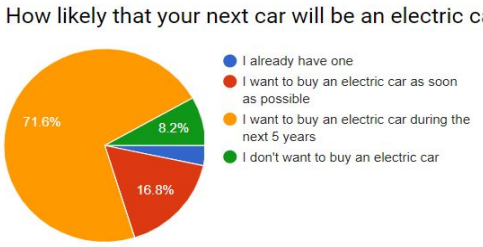
5 C. Responses based on Opinions:











6. Analysis

6 A. Factor Analysis:

Various factors that were collected through the prepared surveys were analyzed and appropriate test statistics were obtained. The **IBM SPSS Statistics** software package was used and the results were visualized.

The analysis for the **positive factors** motivating market growth is as follows.

Descriptive Statistics

	Mean	Std. Deviation	Analysis N
Positive Environment	4.0144	1.14693	416
Low Noise	3.1731	1.69445	416
Cheap Operation	2.9471	1.81084	416
Popularity	1.7596	1.31714	416
Image	3.3846	1.96503	416

Correlation Matrix

	Positive Environment	Low Noise	Cheap Operation	Popularity	Image
Positive Environment	1.000	.073	-.046	-.113	.053
Low Noise	.073	1.000	.223	.049	.038
Cheap Operation	-.046	.223	1.000	.039	-.051
Popularity	-.113	.049	.039	1.000	.073
Image	.053	.038	-.051	.073	1.000

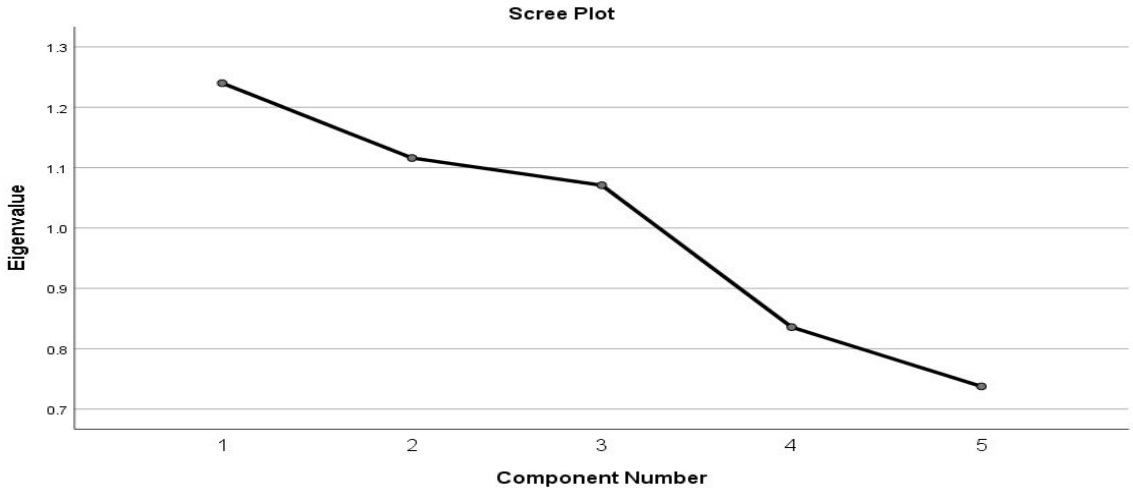
KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.476
Bartlett's Test of Sphericity	Approx. Chi-Square	37.395
	df	10
	Sig.	.000

The corresponding **correlation** between the factors can be observed from the above table. Also, the **chi-square value** of **37.395** and appropriate **sampling adequacy** of **0.476** was obtained. Next, the significance of each factor based on **eigenvalue method** is shown below.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.240	24.798	24.798	1.240	24.798	24.798	1.232	24.650	24.650
2	1.116	22.320	47.118	1.116	22.320	47.118	1.121	22.425	47.074
3	1.071	21.416	68.534	1.071	21.416	68.534	1.073	21.460	68.534
4	.836	16.715	85.249						
5	.738	14.751	100.000						



Rotated Component Matrix[a]

	Component		
	1	2	3
Low_Noise	.785		
Cheap_Op	.768		
Positivr_Env		.796	
Popularity		-.669	.438
Image			.865

The factor analysis of the positive factors generates several tables and graphs for interpreting the results. The first table is of Descriptive Statistics show the mean and standard deviation for the data collected for each factor and also shows the number of respondents in the sample. The KMO test tells us whether or not enough items are predicted by each factor The Bartlett test should be significant this means that the variables are correlated highly enough to provide a reasonable basis for factor analysis and in our case, it is less than 0.05 which is good. Extraction communalities are estimates of the variance in each variable accounted for by the factors or components. A correlation matrix determines the correlation among the various factors considered for the study, it shows that there is either a weak negative or weak positive correlation among the factors influencing the sales of EV. The determinant of the correlation matrix should be greater than 0.001 and for our study, it is 0.913. The Total Variance Explained table shows how the variance is divided among all the factors. The first three factors have eigenvalues (a measure of explained variance) greater than 1.0, which is a common criterion for a factor to be useful. When the eigenvalue is less than 1.0 the factor explains less

information than a single item would have explained. Rotation makes it so that, as much as possible, different items are explained or predicted by different underlying factors, and each factor explains more than one item. The analysis has sorted the five factors into three overlapping groups of items the table shows low noise and cheap operation had high factor loading on component 1 and positive environmental effect and popularity had the highest loading on component 2 and image had the highest loading on factor 3.

Now, analysis of the **negative factors** slowing market growth are as follows.

Descriptive Statistics

	Mean	Std. Deviation	Analysis N
Recharge Time	3.3029	1.71727	416
Price	3.4856	1.64567	416
Recharge Locations	3.7596	1.40564	416
Range	2.5625	1.56342	416
Maintenance	2.2837	1.78651	416

Correlation Matrix

	Recharge Time	Price	Recharge Locations	Range	Maintenance
Recharge Time	1.000	-.057	.176	.145	.090
Price	-.057	1.000	-.039	.130	.115
Recharge Locations	.176	-.039	1.000	.279	.010
Range	.145	.130	.279	1.000	.169
Maintenance	.090	.115	.010	.169	1.000

KMO and Bartlett's Test

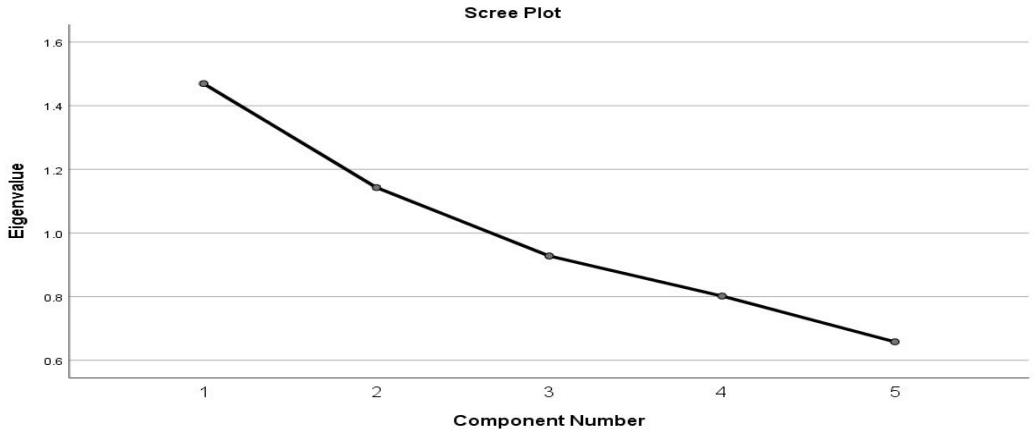
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.542
Bartlett's Test of Sphericity	**Approx. Chi-Square**	80.798
	df	10
	Sig.	.000

Again, the corresponding **correlation** between the factors can be observed from the above table. Also, the **chi-square value** of **80.798** and appropriate

sampling adequacy of **0.542** was obtained. Next, the significance of each factor based on **eigenvalue method** is shown below.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.469	29.389	29.389	1.469	29.389	29.389	1.388	27.752	27.752
2	1.143	22.858	52.248	1.143	22.858	52.248	1.225	24.496	52.248
3	.928	18.555	70.803						
4	.802	16.037	86.840						
5	.658	13.160	100.000						



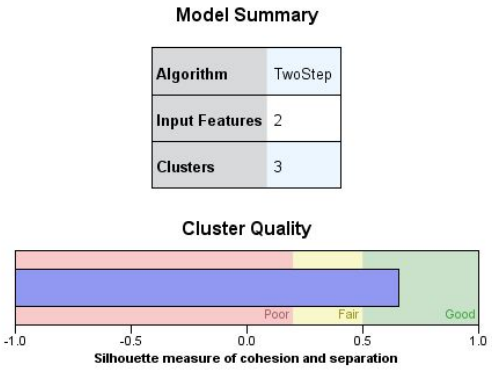
Rotated Component Matrix[a]

	Component	
	1	2
Recharge_Locs	.752	
Recharge_Time	.643	
Range	.590	.477
Price		.760
Maintenance		.645

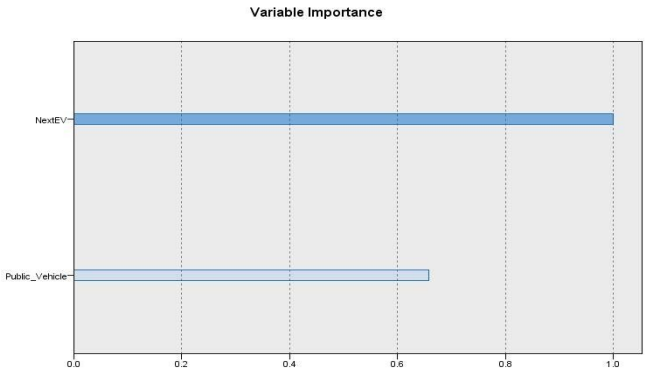
The factor analysis of the negative factors which would prevent the consumers from buying and EV gives several table and graphs for interpreting the results. The Bartlett test should be significant and in our case it is less than 0.05 which is good. A correlation matrix determines the correlation among the various factors considered for the study, it shows that there is either a weak negative or weak positive correlation among the factors influencing the sales of EV. The determinant of the correlation matrix should be greater than 0.001 and for our study, it is 0.822. The Total Variance Explained table shows how the variance is divided among all the factors. The first two factors have greater than 1.0, which is a common criterion for a factor to be useful. Rotation makes it so that, as much as possible, different items are explained or predicted by different underlying factors, and each factor explains more than one item. The analysis has sorted the five factors into two overlapping groups of items the table shows range, recharge time and recharge location are grouped together and had the highest factor loading on component one and price and maintenance had highest factor loading on component 2.

6B. Cluster Analysis

The **TwoStep Cluster Analysis** procedure is an exploratory tool designed to reveal natural groupings (or clusters) within a dataset that would otherwise not be apparent. The algorithm employed by this procedure has several desirable features that differentiate it from traditional clustering techniques like **Handling of categorical and continuous variables** and **Automatic selection of the number of clusters.**

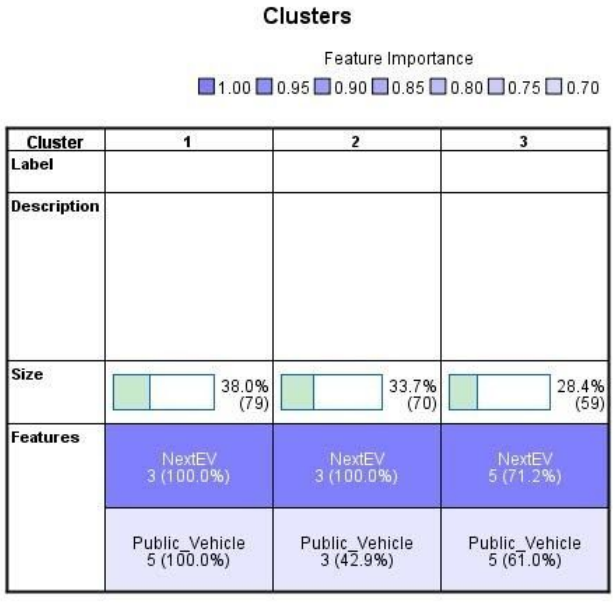


Here we have selected the following categorical variables: **nextEV (willingness to buy an EV)** and **Public vehicle (Opinion about shifting more to public vehicle)** and no continuous variable.



The cluster analysis of the data was performed on determining groups of data with respect to their interest in buying Electric Vehicles as well as their opinion towards the use of more public vehicles.

The **Silhouette method** was used to interpret and validate the consistency within the clusters of data. A good category Silhouette score was obtained. The data can now be evaluated for various evaluating variables as Age, Gender, their day to day driving times.



From the above table the clusters are:

- **38% of respondants** can **buy EV in next 3 years** and **supports public vehicle with rating 5** on a scale from 1 to 5, where 1 being lowest support and 5 being full support.
- **33.7% of respondents** can **buy EV in the next 3 years** and **supports public vehicle with rating 3**.

- **28.4% of respondents can buy EV in the next 5 years** and **supports public vehicle with rating 5.**

6C. Regression Analysis

After the factor analysis of the collected data, a **multiple linear regression model** was constructed with the help of **SPSS Statistics software**. The time duration in which the consumer is willing to buy the car was taken as the dependent variable while all the other demographic, factor and opinion based variables were considered as independent variables.

The results were obtained as the coefficients of all the independent variables and the predicted estimate was visualized. Also, the **ANOVA analysis** results were visualized.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.412[a]	.170	.130	.969	.170	4.268	19	396	.000

a. Predictors: (Constant), Public_v, Cheap_op, Recharge_time, Popularity, Satisfaction, Climate, Age, Price, Image, Pos_Env, Hours, Gender, Tech, Maintain, EV, Recharge_Loc, Vehicle, Low_Noise, Range

b. Dependent Variable: NextEV

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	76.139	19	4.007	4.268	.000[b]
	Residual	371.823	396	.939		
	Total	447.962	415			

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	1.466	.482		3.041	.003	.518	2.414					
	Age	.006	.008	.041	.819	.413	-.009	.022	.052	.041	.038	.838	1.194
	Gender	.052	.061	.042	.856	.392	-.067	.171	.016	.043	.039	.879	1.138
	Vehicle	.078	.052	.075	1.491	.137	-.025	.181	.115	.075	.068	.822	1.216
	Hours	.033	.049	.032	.665	.507	-.064	.130	.024	.033	.030	.912	1.096
	Climate	.084	.209	.019	.403	.687	-.326	.494	.028	.020	.018	.911	1.097
	EV	-.072	.103	-.035	-.699	.485	-.274	.130	-.086	-.035	-.032	.851	1.175
	Tech	.107	.039	.137	2.737	.006	.030	.183	.198	.136	.125	.843	1.187
	Pos_Env	.198	.043	.218	4.555	.000	.113	.283	.248	.223	.209	.911	1.098
	Low_Noise	-.064	.032	-.104	-2.003	.046	-.127	-.001	-.078	-.100	-.092	.776	1.289
	Cheap_op	-.026	.029	-.045	-.904	.367	-.082	.030	-.116	-.045	-.041	.845	1.184
	Popularity	.003	.039	.003	.068	.946	-.073	.079	-.021	.003	.003	.872	1.147
	Image	.105	.025	.199	4.137	.000	.055	.155	.210	.204	.189	.909	1.100
	Recharge_time	.042	.030	.069	1.369	.172	-.018	.102	.073	.069	.063	.825	1.212
	Price	-.005	.030	-.008	-.163	.871	-.064	.054	-.057	-.008	-.007	.924	1.082
	Recharge_Loc	.033	.038	.044	.850	.396	-.043	.108	-.026	.043	.039	.775	1.291
	Range	.001	.035	.001	.014	.989	-.069	.070	-.066	.001	.001	.750	1.333
	Maintain	-.059	.029	-.102	-2.064	.040	-.115	-.003	-.032	-.103	-.095	.865	1.156
	Satisfaction	.053	.034	.077	1.544	.123	-.014	.120	.130	.077	.071	.843	1.186
	Public_v	.010	.048	.011	.216	.829	-.084	.105	.055	.011	.010	.858	1.166

7. Artificial Neural Network Model

Finally, to go for a more advanced kind of predictive analysis, the motive was planned to predict the time in which the customer wants to buy an Electric Vehicle. All the other variables were considered as independent variables.

Hence an **Artificial Neural Network** was trained with the input features and the output was obtained in the range of 1 to 5. The data was checked and the model was verified not to be overfitting. A **70-30 training-validation division** was used to train the model.

Now, for any other customer who has filled the responses of our questionnaire, the algorithm can efficiently predict the approximate time duration in which he is willing to buy an Electric Vehicle.

The **ANN model specifications** are as follows:

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense_1 (Dense)              (None, 128)               2560
_________________________________________________________________
dense_2 (Dense)              (None, 256)               33024
_________________________________________________________________
dropout_1 (Dropout)          (None, 256)               0
_________________________________________________________________
dense_3 (Dense)              (None, 64)                16448
_________________________________________________________________
dropout_2 (Dropout)          (None, 64)                0
_________________________________________________________________
dense_4 (Dense)              (None, 1)                 65
=================================================================
Total params: 52,097
Trainable params: 52,097
Non-trainable params: 0
```

8. Inferences

The result of the **factor analysis** says that there are **three positive components** which has **eigenvalue greater than one** and there is **not much difference** between the importance of each factor. Also, there are **two negative components** which has **eigenvalue greater than one** and both the component **explains up to half of the total variance in the data**. **Positive Environmental Effect, Effect on their Image, and Low noise caused** are identified as the major positive factors and **Less recharge locations available and high price of EVs** are identified as the major negative factors.

On the basis of **cluster analysis**, the major factors chosen were their **opinion towards shifting more to public vehicles** and their **willingness to buy an Electric Vehicle**. There was a **lot of willingness in the younger generation** to move towards the electric vehicle market, although a good amount of response also pointed towards **shifting to more public vehicles unless a proper infrastructure for the Electric Vehicles is set up**.

Although, the **regression analysis** was not effective because of **non-linearity** in the data, the **Artificial Neural Network** is expected to behave more accurately to **predict the willingness of people to buy an Electric Vehicle** on the basis of their responses based on demographics and opinion about the other governing factors.

Future Scope

The impact of the work will increase with the proper analysis of responses collected from **company executives of Automotive and Electricity Supplying companies**. Although, the study comprises of approaching the research domain with carious angles of analysis methods, the **accuracy and validity can be increased with more response collection and reach**. With all this incorporated in the future works, a better study can be presented.