

A Study on Factors that Influence the Choice of Motor Bikes in Chennai

Dr. J. Padmavathi*, J. Parthasarathi

Department of Computer Science & Application, SRM Institute of Science & Technology, Chennai, Tamil Nadu, India

ABSTRACT

Article Info

Volume 7, Issue 3

Page Number: 692-697

Publication Issue :

May-June-2021

Article History

Accepted : 12 June 2021

Published : 20 June 2021

The automobile industry is facing intense competition and there is an urge among the youngsters to own a popular bike. This study was undergone to analyze the factors that influence the choice of bike and to evaluate the customer behavior in decision making in buying bikes. The paper analyses the various brand that are looked for by different categories of users. It also performs a predictive analysis to customer behavior and their buying pattern based on age of the customer, style of the bike, cost, speed, technology etc.

Keywords - Data Clustering, Data Mining, Data cleaning, brand preference, choice of bike

I. INTRODUCTION

The automobile industry is facing a stiff completion in promoting their bikes and to achieve high targets, especially with mixed consumers. The expert committee and the data scientist observe the various trends, shopping patterns of all the people belonging to various age groups and categorize their preferences and needs. An analysis of the customer's behavior in terms of customer consumption patterns, customer preferences, customer motivation, customer buying process and shopping behavior is very much helpful to formulate a firm's marketing strategy. This results in predictive analysis which helps the firms to raise their level in the competitive market by making new strategies.

A. CLUSTERING IN DATA MINING

The ability to form meaningful groups of objects and insights are one of the most fundamental

modes of intelligence .Cluster analysis is a tool for exploring the structure of data. It is the process of grouping objects into group of clusters such that the objects belonging same cluster are similar and objects from other clusters are dissimilar. Objects can be described in terms of attributes, relationship with objects. Clustering is unsupervised learning technique. Cluster analysis is a more primitive technique. In this we don't make assumptions on concerning the number of groups or the group structure. Groupings are done based on the similarities and dissimilarities. The inputs are similarity measures/ the data from which similarities can be computed

B. LITERATURE REVIEW

Dr. I. Savarimuthu **K. Chandrakumar (2017) in his study states that it is very important to understand the meaning of perception and brand.

Both are associated when it comes to purchase decision of two wheelers

Dr. T. Thirupathi and M. Anthonisamy (2015) have analysed that buyers of two Wheelers , specially buyers of Bajaj in Salem pay more attention on mileage, color, design and look of the two wheelers, and availability of spare parts, it will be beneficial for the companies to concentrate more on these aspects of production.

Dr. Ronald Mani & Mr. Debasis Tripathy(2013) have stated that consumer making a purchase decision will be affected by Cultural Factors, Social Factors, Personal Factors and Psychological Factors. The marketer must be aware of these Factors in order to develop an appropriate marketing plan for its target market.

Reddy, Mallikarjuna,K. (2006), states that there are four major factors which influences on the buying behavior of consumer such as Cultural, Social, Personal and Psychological. It was found that as far as mileage is concerned, Hero Moto Corp was ruling the market. Yamaha and TVS were far below the expectations of the consumers based on mileage per liter of petrol. Also, Hero Moto Corp bike was more popular with the students, TVS was more popular with the employees, whereas businessmen opted for Yamaha. Regarding satisfaction of motor bikes 85 respondents have stated that Yamaha satisfaction was highest and the lowest is for TVS bikes.

Kanojia, A.K. (2011)^[6] States that top 4 motorcycle segment control 93% market share. Hero Moto Corp controls executive segment, Bajaj has premium segment and Honda has the dominant position in scooter segment. The author studied about various factors responsible for Hero Moto Corp being able to sustain leadership like highest network penetration,

highest customer satisfaction with one of the fastest complaint resolution.

Md Ashaduzzaman et al (2018) ^[10], concluded that Marketer should use the customer analysis report findings for new brand building, brand extension, market extensions of motorbikes in Bangladesh.

C. Objective of The Study

- To study the customer preferences.
- To analyze if brand conversion takes place.
- To analyze if producers meet the expectations of a customer.

Data clustering help to identify, learn, or predict the nature of new data items — especially how new data can be linked with making predictions. For example, in pattern recognition, analyzing patterns in the data can help us develop predictive analytics — in this case, predicting the nature of future data items that can fit well with established patterns.

II. METHODS AND MATERIAL

A. Data Mining

The goal of Data Mining is to make data more usable. It is based on mathematical and scientific methods to identify patterns or trends. On the other hand, data analysis helps in proving a hypothesis or taking business decisions. It business intelligence and analytics models for better results based on abstract^[7] Data mining is an emerging multidisciplinary field which facilitates discovering of previously unknown correlations, patterns and trends from large amounts of data stored in multiple data sources. It is a powerful new technology with great potential to help businesses make full use of the available data for competitive advantages. It also draws attention to useful applications, giving a small collection of real-life examples of data mining

implementations from the business to the scientific world. Data clustering reveals structure in the data by extracting natural groupings from a dataset. Therefore discovering clusters is an essential step toward formulating ideas and hypotheses about the structure of our data and deriving insights to better understand it. Similarity measure in clustering is used to analyze and group data. Similarity is hard to define. It is a measured based feature. To purify the similarities, the following steps has to be performed

- Generate features
- Clean features
- Normalize features[transform]
- Reduce features[reduce dimensionality]

Similarity can be measured using Feature projection-project the data into a feature space. The distance in feature space becomes the similarity.

Edit distance-measure similarity between two objects by transforming one object into another and measure how much effort it takes for this process.

B. Clustering

Clustering is the process of making a group of abstract objects into classes of similar objects.

A cluster of data objects can be treated as one group. While doing cluster analysis, we first partition the set of data into groups based on data similarity measure and then assign the labels to the groups. The main advantage of clustering over classification is that, it is adaptable to changes and helps single out useful features that distinguish different groups.

1) **Applications of Cluster Analysis:** Clustering analysis is broadly used in many applications such as market research, pattern recognition, data analysis, and image processing. Clustering can also help marketers discover distinct groups in their customer base. And they can characterize their customer groups based on the purchasing

patterns. In the field of biology, it can be used to derive plant and animal taxonomies, categorize genes with similar functionalities and gain insight into structures inherent to populations. Clustering also helps in classifying documents on the web for information discovery. Clustering is also used in outlier detection applications such as detection of credit card fraud. As a data mining function, cluster analysis serves as a tool to gain insight into the distribution of data to observe characteristics of each cluster.

2) Requirements of Clustering in Data Mining:-

The following points throw light on why clustering is required in data mining –

- i) Scalability – we need highly scalable clustering algorithms to deal with large databases.
- ii) Ability to deal with different kinds of attributes – Algorithms should be capable to be applied on any kind of data such as interval-based (numerical) data, categorical, and binary data.
- iii) Discovery of clusters with attribute shape– the clustering algorithm should be capable of detecting clusters of arbitrary shape. They should not be bounded to only distance measures that tend to find spherical cluster of small sizes.
- iv) High dimensionality – the clustering algorithm should not only be able to handle low-dimensional data but also the high dimensional space.
- v) Ability to deal with noisy data – Databases contain noisy, missing or erroneous data. Some algorithms are sensitive to such data and may lead to poor quality clusters.
- vi) Interpretability – the clustering results should be interpretable, comprehensible, and usable.

3) Clustering Methods

Clustering methods can be classified into the following categories –

- Partitioning Method
- Hierarchical Method
- Density-based Method
- Grid-Based Method
- Model-Based Method

4) Constraint-based Method :In this paper partitioning method is used to analyze data. Suppose we are given a database of 'n' objects and the partitioning method constructs 'k' partition of data. Each partition will represent a cluster and $k \leq n$. It means that it will classify the data into k groups, which satisfy the following requirements:-

1. Each group contains at least one object.
2. Each object must belong to exactly one group.

Thus, for a given number of partitions (say k), the partitioning method will create an initial partition and then uses the iterative relocation technique to improve the partitioning by moving objects from one group to other. This is k-means partitioning. K-means algorithm is used to analyze the data and clusters^[9].

K-Means Algorithm:

1. Decide on a value for k.
2. Initialize the k cluster centers (randomly)
3. Decide the class membership of n objects by assigning them to the nearest cluster centers.
4. Re-estimate the k cluster centers
5. If none of the n objects changed membership in the last iteration, exit else go to 4.

III. RESULTS AND DISCUSSION

Data clustering can also be a way to model data. It represents a larger body of data by clusters or cluster representatives. In addition, the analysis may seek simply to partition the data into groups of similar items — as when market segmentation partitions target-market data into groups such as customers who share the same interests, customers who have common needs etc. Identifying clusters of similar customers can help analyst to develop a marketing strategy that addresses the needs of specific clusters.

A. Data Collection

The study was targeted on college students who are eager and crazy in buying bikes with different features like speed, style, cost, pride to own, etc. The study was not focus on a particular brand, instead on the teenagers. A descriptive questionnaire was framed and data was collected from 200 students. 50 records were discarded in cleaning process due to inconsistency and irrelevancy. Outliers were not considered. Cleaned data set had 150 records which included 135 students and 15 working people who were pursuing their part-time studies.

B. Result Discussion

The data on analysis was clustered into 2 groups. The number of customers who owned a bike previously was 79 out of 135. In the sample test data 55 out of 71 were willing to buy a new bike and the survey recorded various brand counts that are tabulated as follows:

Brand	Count(Own ing a bike)	Count(Willing to buy)
Honda	20	11
Bajaj	9	9
Yamaha	17	8
Suzuki	3	-
Hero	6	-
TVS	5	2
Royal Enfield(RE)	16	15
Piaggio	2	3
KTM		
BMW		
Mahindra		
Total	79	

Figure1: FREQUENCY OF BUYING

In our study, it was observed that the factors that influence the purchase of a particular brand was based on mileage, style, speed, model, Pride in owning a bike (status), love for brand and technology such as ABS, emission norms, vehicle torque, etc. The study further analyzed customers who were willing to buy a bike. This category included new purchase and exchange of pre-owned bike to new ones. 40 customers were willing to go for new bikes. It was observed that 20% students show interest in buying Royal Enfield stating that enhancement in technology such as drum brakes replaced with disc braking system and ABS. 12% showed interest in Yamaha due to mileage and 8% showed interest in Honda. Safety riders opted gearless bikes.

A	B	C	D	E	F	G	H	I	J	K	L
age	speed	mileage	brand	price	comfort	Based on Ads	Availability	safety	sales service	after sales	buy
18	yes	no	yes	no	yes	yes	yes	yes	yes	no	yes
18	yes	no	no	no	no	no	no	yes	yes	yes	yes
18	yes	yes	yes	no	yes	yes	no	no	no	no	yes
18	yes	yes	no	no	yes	no	no	no	yes	yes	yes
19	yes	yes	yes	no	yes	yes	no	yes	yes	no	yes
19	yes	yes	yes	no	no	no	no	no	no	yes	yes
20	yes	yes	yes	no	yes	no	no	no	no	yes	yes
20	yes	no	no	yes	no	yes	no	yes	yes	no	yes
20	yes	no	yes	no	yes	yes	no	no	yes	yes	yes

Figure 2 : Attributes used dataset

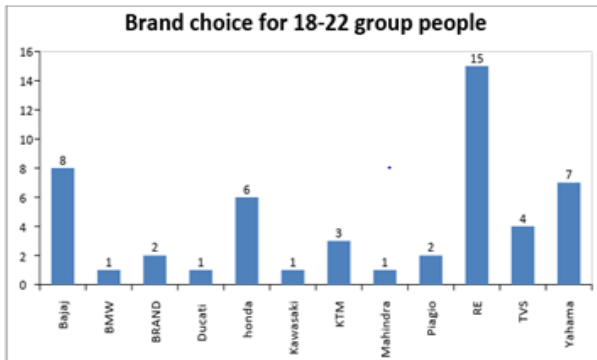


Fig: Graph indicating various brand preferences

```

Cluster centroids:
Attribute      Full Data      Clusters
              (150)          (88)          (62)
-----
Age           20.4467        20.125        20.9032
Occupation    Student        Student       Student
brand         RE             Honda         RE
mileage       yes            yes           high
cost          moderate      moderate
style         yes           yes           yes
speed         yes           yes           yes

Clustered Instances
0      88 ( 59%)
1      62 ( 41%)

Class attribute: comfort
Classes to Clusters:
0 1 <-- assigned to cluster
88 62 1 yes

Cluster 0 <-- yes
Cluster 1 <-- No class
    
```

Figure4: Clusters based on “buy”, “not buy”

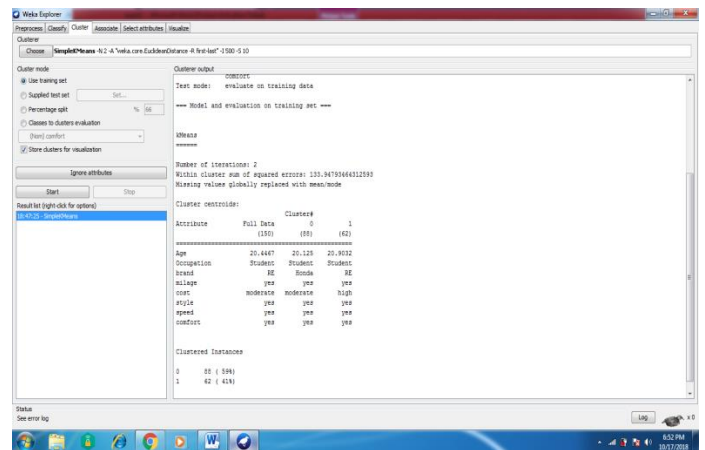


Figure 5: Output Screen-1

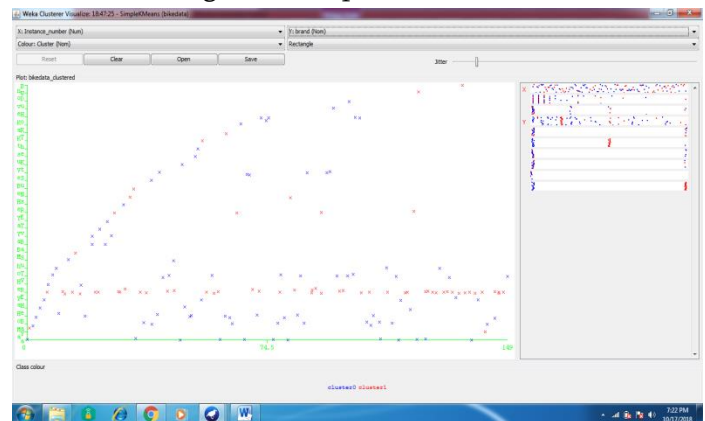


Figure 6 : Cluster visualization based on brand.

IV. CONCLUSION

The study shows that sales of bikes depend on various factors such as speed, style, cost, pride to own etc. Speedy and stylish bikes are more preferred by students. It was quite obvious that 96% of the customers in the age group 18 to 21 were ready to buy bikes based speed, brand, comfort, ads, sales service and after service. The research may further be extended to customer relationship management and brand conversion.

V. REFERENCES

- [1]. Dr. I. Savarimuthu **K. Chandrakumar, A Study On Customer Perception Towards Major Brands Of Two Wheelers In Villupuram Town And Its Impact On Buying Behaviour(2009), Journal Of Exclusive Management Science – September 2017 - Vol 6 Issue 09 – Issn 2277-5684 1 www.Jems.Net.In
- [2]. Dr. T. Thirupathi, M. Anthonisamy, Brand Preference Of Two Wheelers – (A Study With Special Reference To Bajaj Two Wheeler Users In Salem District), International Journal Of Advanced Scientific Research & Development Vol. 02, Iss. 03, Ver. I, Jul – Sep’ 2015, Pp. 57 – 67
- [3]. Dr. Ronald Mani &Mr. Debasis Tripathy, A Study On Consumer Buying Behavior Towards Two Wheeler Bikes In Context To Indian Market, International Journal Of Advanced Research In Management (Ijarm), Issn 0976 – 6324 (Print), Issn 0976 – 6332 (Online), Volume 4, Issue 1, January- April 2013 © Iaeme 65
- [4]. Dr. Ronald Mani &Mr. Debasis Tripathy, A Study On Consumer Buying Behavior Towardstwo Wheeler Bikes In Context To Indian Market, International Journal Of Advanced Researchin Management (Ijarm)Issn 0976 – 6324 (Print), Issn 0976 – 6332 (Online) ,Volume 4, Issue 1, January- April 2013, pp. 65-73
- [5]. Reddy, Mallikarjuna, K. (2006), “Consumers Behavior Towards Two wheeler Motor Bikes”, Osmania Journal of Management, Vol. II, No. 2, June.
- [6]. Kanojia, A. K. (2011), “Analyzing the state of competition in Indian two wheeler industry”, New Delhi, Competition Commission of India.
- [7]. N. Mlambo, “Data Mining: Techniques, Key Challenges and Approaches for Improvement”, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 6, Issue 3, March 2016
- [8]. Jiawei Han and Micheline Kamber, Data Mining: Concepts and Techniques, Morgan Kaufmann Publishers, 2003.
- [9]. Blog by Imad Dabbura Data Scientist , <https://towardsdatascience.com/k-means-clustering-algorithm-applications-evaluation-methods-and-drawbacks-aa03e644b48a>.
- [10]. Md Ashaduzzaman, Imranual Hoque, Farzana Mahub, Factors Influencing Bike Purchase Decisions: A Study of Consumers in Bangladesh, <https://www.researchgate.net/publication/327634116> , September, 2018.

Cite this article as :

Dr. J. Padmavathi, J. Parthasarathi, "A study on factors that influence the choice of Motor Bikes in Chennai.", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 7, Issue 3, pp.692-697, May-June-2021. Available at
doi : <https://doi.org/10.32628/CSEIT2173150>
Journal URL : <https://ijsrcseit.com/CSEIT2173150>