

Analyzing Retail Marketing in Theni district using Data Mining Approaches

Jasmine B^{#1}, Shantha Mary Joshitta R^{*2}, Durga Devi R^{@3}

^{#1,2}Asst. Professors, Dept. of Computer Science, JayarajAnnapackiam college for women (Autonomous), Periyakulam, TheniDt, Tamilnadu, India.

^{@3}PG Student, JayarajAnnapackiamcollege for women (Autonomous), Periyakulam, TheniDt, Tamilnadu, India.

¹Jeromia2585@gmail.com²rjoshitta@gmail.com³durgadevi97dd@gmail.com

Abstract-Retailing is the link that connects the manufacturer and the consumer. It analyses a set of business behaviours that adds value to the products and the services sold to the end consumer for their personal, family or household use. This study uses data mining algorithms for discovering patterns to make better strategies and decisions in retail stores. In this paper a data mining case study is conducted in Theni District, Tamilnadu, India to analysing the factors attempts, how data mining can be applied in retail industry to improve market campaign. Data were collected from the public using a set of questioner and its analyzed using R tool.

Keywords: Association Rules, Apriori Algorithm, Retail Sector, Data mining Algorithms, Theni District.

I. INTRODUCTION

Retailer is the self or organization who delivers products to final consumer in the channel of distribution. Goods and services are created for utilization and use by people, it is vendor who assumes the role of taking the goods to its final destiny of utilization. Indian Retail industry is one of the pillars of Indian economy and accounts for 14 to 15 percent of its Gross Domestic Product (GDP). The Indian retail market is estimated to be US\$ 500 billion and one of the top five retail markets in the world by economic value. India is one of the fastest growing retail markets in the world, with 1.2 billion people. India's retailing industry was essentially owner manned small shops. In 2014, larger format convenience stores and supermarkets accounted for about 4 percent of the industry, and these were present only in large metropolitan centers. India's retail and logistics industry employs about 40 million Indians (3.3% of Indian population). Any organization selling to final consumers whether it is a producer, trader or retailer is doing retailing. It does not matter how the goods or services are sold (by person, mail, telephone, vending machine, or internet or where they are sold-in a store, on the street, or in the consumer's home). A Retailer thus, provides value creating functions like assortment of products and services to the consumers, breaking bulk, holding inventory and provides services to customers, manufacturers and wholesalers.

II. BACKGROUND OF THE STUDY

Retail is the final stage of any economic activity. By virtue of this fact, retail occupies an insignificant place in the world financial system. A retailer or retail store is any business enterprise whose sale volume comes primarily from retailing. These are the final business entities in a

distribution channel that links manufacturers to customers. Manufacturers typically make products and sell them to retailers or wholesalers. Wholesalers resell these products to the retailers and finally, retailers resell these products to the ultimate consumers.

Retailing thus, may be understood as the final step in the distribution of commodities, for consumption by the end consumers. Put simply, any firm that sells products to the final consumer is performing the function of retailing. It thus consists of all activities involved in the marketing of goods and services directly to the consumers, for their personal, family or household use. In an age where customer is the king and marketers are focusing on customer delight, retail may be redefined as the first point of customer contact. The distribution of final products begins with the producer and ends at the last consumer, Between two of them there is a middleman – the retailer.

III. REVIEW OF THE RELATED STUDY

JyotiVohra et al., [11] discussed on the use of data mining to gather the information from retail market and used different techniques of data mining for retail industry. These techniques show that association rule mining plays an important role in retail market. The study showed the conventional techniques such as the frequent - buyer program or the market basket analysis can only provide a profile of customer's purchasing affinities. It's described only the combinations, which are in customers shopping carts, but cannot tell us why. The reason why some products are frequently bought together, like detergent and fabric softener, is apparent, while other combinations, like bottled juice and cold remedies or beer and diapers, are not so easily explained. Apriori algorithm performs best as compare to other traditional techniques.

Diptimayee Mishra et al., [12] analyzed the different trends of data mining and its relative applications from past to present and discussed how effectively. The retail industry was also realizing that data mining gave them a competitive advantage. A majority of the banks in emergent countries (particularly in the public sector) are not usually known to exploit their information.

Jayadatta et al., [8] has presented an overview and major challenges affected unorganized retail sector in India. The study focused on the prosperous sector of Indian economy. The study also showed that there was ample growth for

unorganized retail sector though some challenges exist in the form of inefficient supply chain management, implementation of new technologies and e-commerce, e marketing as well as getting approval from the government. Major importance of unorganized retail sector in India and major findings of research survey conducted helped unorganized retailers to devise and frame strategies to face the challenges and opportunities in this sector.

Chaturvedi Molly[13] focused on the present structure of Indian Retail Sector, the major sub-sectors in controlled and conventional retail and changes in the relative share of various sub-sectors over last few years and saturation of organized retail in various segments. The analysis also covered the opportunities and emerging challenges before Indian retail sector in view of recent policy changes by Government of India. With India's large 'young' inhabitants and high domestic utilization, the macro trends for the sector look positive. Organized retailing offers huge potential for future growth of retailing in India. This paper provided information about the growth of retailing in India, and also paying attention on the challenges faced by organized retail sector in India, and also emphasized on major players of retailers in India and customer services provided by the retailers. This paper was also deals with various retail formats and the opportunities for the growth of retail industry in India and also provided some suggestions to defeat the challenges

Badma et al., focused on the market space and growth for organised retail market and the unique challenges and it was found that competition from unorganised retailers, changing consumer preference, entrants of global retail giants etc.,. Finally they analyzed to overcome the challenge the organised retailers need to improve infrastructure, strengthening supply chain, real estate availability. It concluded a successful organized retailer was one who understands and fulfilled the customers.

IV. SIGNIFICANCE OF THE STUDY

The present study is an important challenge to scrutinize the role of retail market in Theni district. The study will prove helpful to examine lacunae in the existing market practice, problems and shortcomings in the field of transportation development etc. promote, it can help to explore some untapped areas where potential for development exists and minor adjustment of the flaws in the present system of cost-effective policy. The study will also prove helpful in enhancing the GDP, generating employment avenues and creation the country self-reliant. The study may also be equally important for all those interested in undertaking similar studies in the background of retail markets because certain methods and approaches evolved and employed should be helpful in carrying out further studies of parallel nature. The present study is likely to provide valuable information to the investors and commodities about the success of the retail sector in Theni district. The constraints faced and suggestions made by the study will be of immense help for

policy makers and development programme initiators to plan the future programmes more effectively.

V. OBJECTIVES OF THE STUDY

Despite certain attempts made to study different aspects of retailing industry in Theni district, there remain certain very relevant gaps. In order to fill the gaps, the present study was undertaken. With this background the present study, the following objectives have been formulated.

- ❖ To discuss the environment of retailing in Theni district, Tamilnadu.
- ❖ To study the present scenario of retailing in Theni district, Tamilnadu.
- ❖ To analyze the positive and negative impacts of the reforms to be undertaken.
- ❖ To evaluate the change in the customer's requirements in retail markets of Theni district

For accomplishing the objectives of the study, both secondary and primary data have been utilized. To estimate the overall position of retail market in Theni, secondary data have been collected from various published sources and websites. Interpretation of the data is more about qualitative terms than on quantitative terms. Primary data is the first hand information and the analysis of the present study largely based on this data which has been collected through direct interview method.

Consistent with the objectives of the study, different techniques have been used for the analysis of data. The data were presented through tables, diagrams and figures and simple statistical techniques of analysis, percentages have been applied. Interpretation of data is based on rigorous exercises aiming at the achievement of the study objectives and findings of the existing studies.

VI. METHODOLOGY

The present study is an investigative research conducted among the people in Theni district. In order to follow the aim and objectives outlined in the above section, a content analysis of information gained from the research process was conducted to establish the underlying trends in location to find common sectors. The methodology of the research work is depicted in Fig.1.

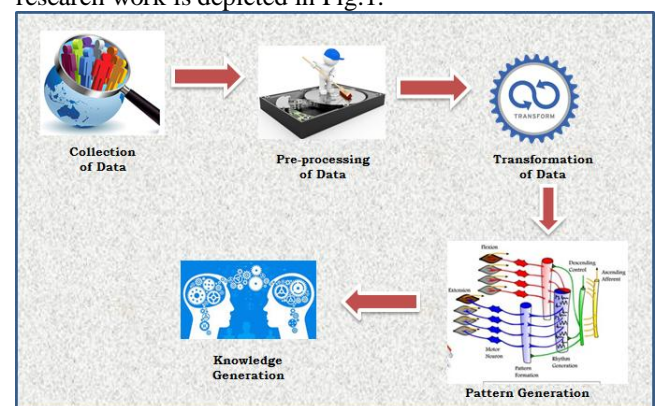


Fig. 1. Methodology of the Proposed System

The first stage is involved in gathering the secondary information from people. A summary of interpretations is presented below. Third stage, an analysis was carried out by making specific assumptions in a hypothetical situation. In the last stage, on the basis of the results and interpretations, specific postulates were framed on each postulate hypothesis and tested through quantitative research using data mining algorithm.

Data mining Technique: Data mining is the core process of knowledge discovery in database. It is the process of extraction of useful patterns from the large database. To analyze the large amount of collected information, the area of Knowledge Discovery in Database (KDD) provides techniques which extract interesting patterns in a reasonable amount of time. data mining is the application of efficient algorithms to detect the desired patterns contained within the given data. Data mining is the extraction of hidden descriptive or predictive information from large databases.

Association Rule: An association rule is one of the forms $x \Rightarrow y$. and each rule has two basic needs: support and confidence. Things that occur often together can be associated to each other. These together occurring things form a frequent itemset. Conclusions based on the frequent item sets make association rules.

Apriori Algorithm: Apriori algorithm is a fundamental algorithm mining association rule. It contains two processes: • Detect all frequent itemsets by scanning db. • Form strong association rules in the frequent itemsets. Process one needs to scan DB several times, which consumes a lot of time and space. As a result, what needs to be improved is the mining competency of frequent group of things in DB.

Apriori algorithm is a significant algorithm for mining frequent itemsets for Boolean association rules. Apriori algorithm is formed by Agrawal and Srikantin 1994. It is the most fundamental and important algorithm for mining frequent itemsets. Apriori is used to detect all frequent itemsets in a provided database db. The keynote of Apriori algorithm is to form multiple passes over the database. It employs an repetitive approach called as a breadth-first search (level-wise search). The flow diagram of the Apriori algorithm is presented in Fig.2.

- ❖ **Frequent Itemsets:** The itemsets which has minimum help (denoted by I_i for i^{th} -itemsets), Apriori property: any subgroup of frequent things must be frequent.
- ❖ **Join Operation:** to detect I_k , a group of candidate k -group of things is developed by adding I_{k-1} with itself.

$$\text{Support}(X \Rightarrow Y) = \text{Support}(XUY) = P(XUY).$$

- ❖ **Confidence:** The confidence is defined as a conditional probability

$$\text{Confidence}(X \Rightarrow Y) = \text{Support}(XUY) / \text{Support}(X) = P(Y/X).$$

- ❖ **Lift:** is the ratio of the probability that L and R occur together to the multiple of the two individual probabilities for L and R, i.e.

$$\text{lift} = Pr(L,R) / Pr(L).Pr(R).$$

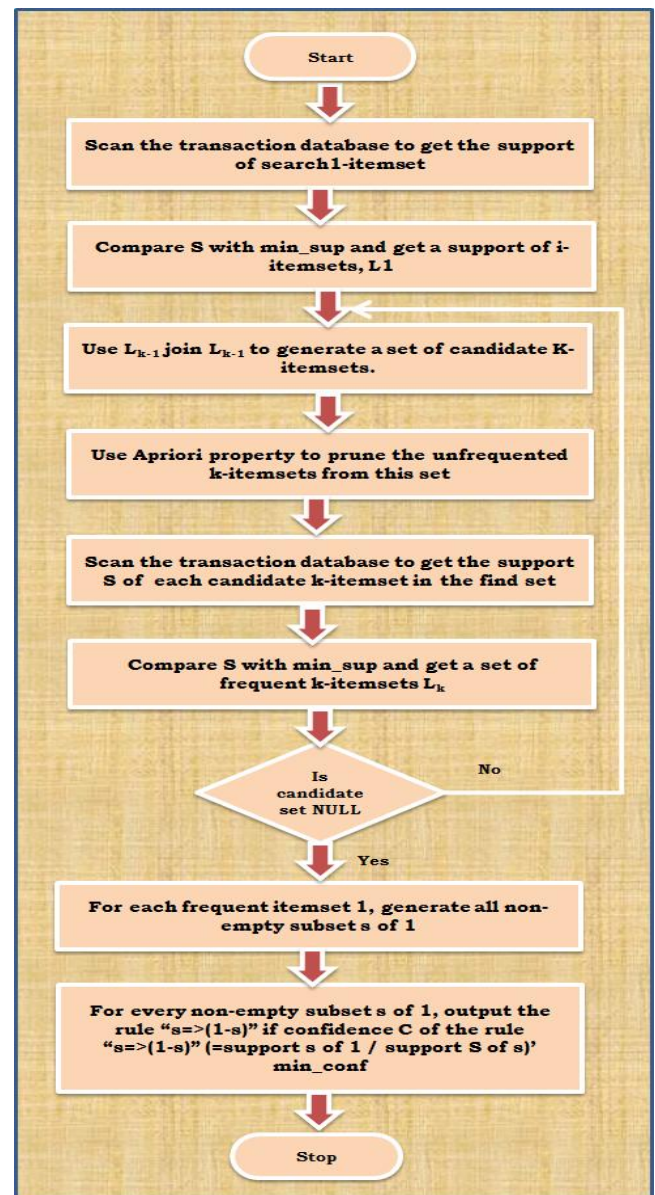


Fig. 2. Flow diagram of the Algorithm

- ❖ **Conviction:** is similar to lift, but it measures the effect of the right-hand-side not being true. It also inverts the ratio. So, a conviction is measured as:

$$\text{conviction} = Pr(L).Pr(not R) / Pr(L,R)$$

VII. EXPERIMENT ANALYSIS

Retail Sector One can also apply data mining for the following purposes.

- A. **Acquiring and Retaining Customer:** It is more costly to reach new customers than to get existing one. So by knowing existing customers' purchase behavior, direct marketer can predict customers need and interest in buying particular product. Using this type of prediction retailer can retain existing customers by providing discounts or offer, attract customers and acquire customers.

B. *Market Basket Analysis:* Market basket analysis is a technique in understanding what items are likely to be purchased together according to association rule. It provides valuable indications about customers, shopping patterns by showing associations among various items. This type of item association is useful for shelf design, deciding the location and promotion of items by means of combination. So that customers can easily locate item and this analysis helps in product cross-selling.

C. *Customer Segmentation and Target Marketing:* Segmentation is to divide the market into several parts by certain characters. Data mining can be used in grouping or clustering customers based on the behavior. This type of information is useful to define similar customers in a cluster, holding on good customers and identify likely responders for target marketing. Many industries have already applied data mining technology and gaining advantages over its competitors.

D. *Analysis of data*

```

➤ itemsetList = LIST(items(fsets), decode = FALSE);
➤ allConfidence = quality(fsets)$support /
  sapply(itemsetList, function(x)+
    max(singleSupport[as.character(x)]));
➤ quality(fsets) = cbind(quality(fsets),
  allConfidence);
➤ summary(fsets);
  
```

set of 10 itemsets

most frequent items:

item185 item200 item211 item317 item548 (Other)

1 1 1 1 1 5

Min. 1st Qu. Median Mean 3rd Qu. Max.

1 1 1 1 1 1

summary of quality measures:

supportallConfidence

Min.:0.05100 Min. :1

1st Qu.:0.05575 1st Qu.:1

Median :0.06150 Median :1

Mean :0.06830 Mean :1

3rd Qu.:0.07125 3rd Qu.:1

Max. :0.12200 Max. :1

includes transaction ID lists: FALSE

mining info:

datantransactions support

trans 1000 0.05

```

➤ itemFrequencyPlot(tr,support=0.1)
  
```

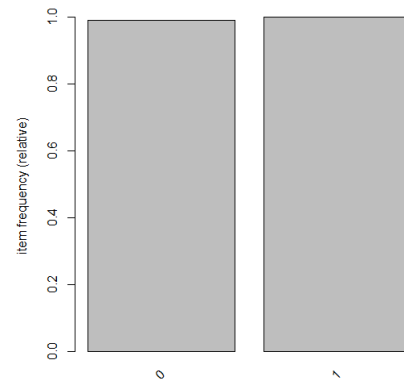


Fig.3 Support and confidence analysis

```

➤ interestMeasure(rules,c("support","chiSquare","confidence",
"conviction","cosine","coverage","leverage","lift",
"oddsRatio"),tr)
  
```

	Support	chi Squared	confidence	conviction
	cosine	coverage	leverage	lift
1	0.99	NA	0.99	1 0.9949874
	1.00	0	1	
2	1.00	NA	1.00	NA 1.0000000
	1.00	0	1	
3	0.99	NA	1.00	NA 0.9949874
	0.99	0	1	
4	0.99	NA	0.99	1 0.9949874
	1.00	0	1	

```

➤ print(rf)
  
```

function (n, df1, df2, ncp)

{

if (missing(ncp))

.External(C_rf, n, df1, df2)

else (rchisq(n, df1, ncp = ncp)/df1)/(rchisq(n, df2)/df2)

}

```

➤ bytecode: 0x00000000234bf320>
  
```

```

➤ environment: namespace:stats>
  
```

```

➤ plot(df1, main = "")
  
```

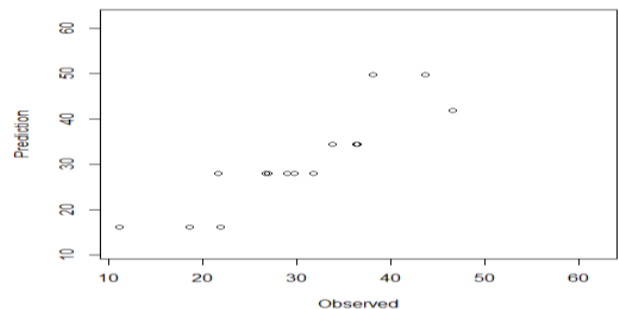


Fig 2 Scatter plot representation of frequent data

- detach(mtcars)
- set.seed(482)
- y <- rnorm(100)
- boxplot(y)

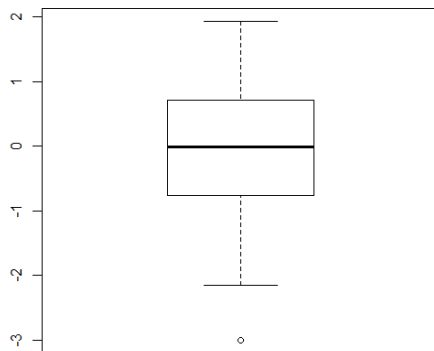


Fig 4 Boxplot representation of data

- set.seed(492)
- y <- rnorm(2000)
- x1 <- sample(letters[1:2], 2000, T)
- x2 <- sample(letters[1:2], 2000, T)
- lab_y <- sample(letters[1:4], 2000, T)
- ipcp(mtcars[c("mpg", "wt", "qsec", "disp", "hp")])

ID:6 Name: "Parallel coord. plot (default)"

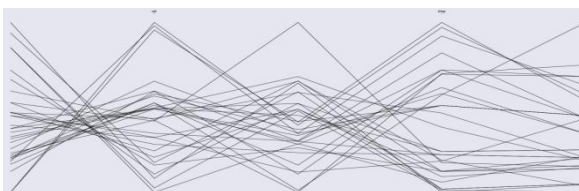


Fig 5 Parallel coordination of data

VIII. FINDING AND INTERPRETATIONS

The gathering of relevant and up-to-date information is a key business process. Information consists of organized facts and figures that have meaning within the context that the information is intended to be interpreted by people. Information is thus a valuable business commodity, and frequently businesses pay money for up-to-date and relevant information. The Retailing store support into the frequent itemset to be found by the following command in R tool.

- [1]. Gender based Shopping: According to censuses, 2015, Theni had population of 1,245,899 of which male and female were 625,683 and 620,216 respectively. It is evident to find out that Theni women show more interest and spend lot of time in shopping. They prefer to go to departmental stores where they get all the items at one place. Many of them express that they will buy house hold items, grocery items, and dressings not only for them but for the family as well. They prefer to shop at the festival times because of the chances of getting special offers for the purchases.

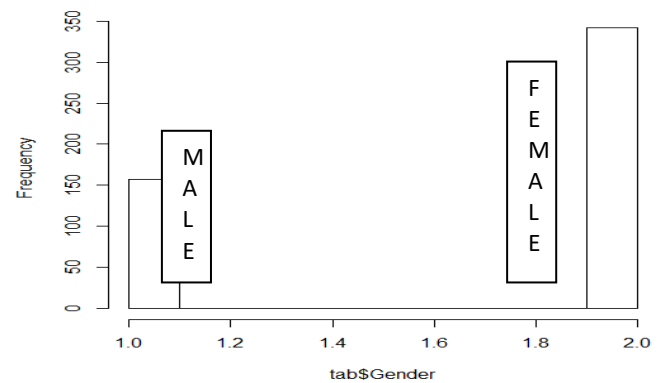


Fig 6 gender based shopping Analysis

- [2]. Degree based shopping: Censuses, 2015 shows, Male Literates of Theni district are 479,403 where as that of Female Literates 390,977. 85.03 per cent in Male Literacy is obtained where are Female Literacy is only 69.46 per cent. In Theni district people who have completed shows more interest in shopping. Many of them those who are not completing schooling feel that they could not make a good shopping. Moreover these kinds of people feel comfortable with nearby petty shops and give less preference to departmental stores. If some chances of going to departmental stores happened, they accompany any of their family members, relatives or friends who have completed any UG degree.

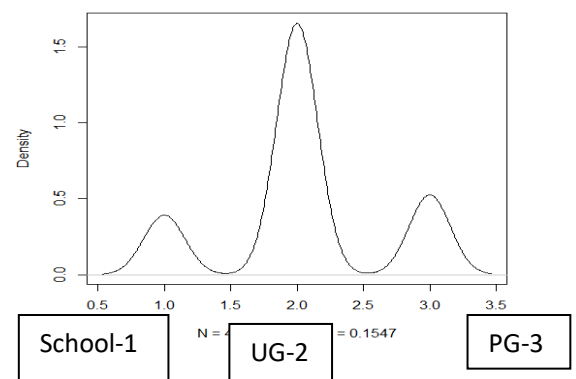


Fig 7 Degree based shopping analysis

- [3]. Community based Shopping: Though the Theni district has high number of scheduled cast people, the backward community people show more likeness in shopping. There may be many reason, one among them is the economic feasibility of this community people. Many of them are government employees and having their own business. It is another reason why the backward community people shopping more. The most backward community people come the second in shopping in departmental stores.

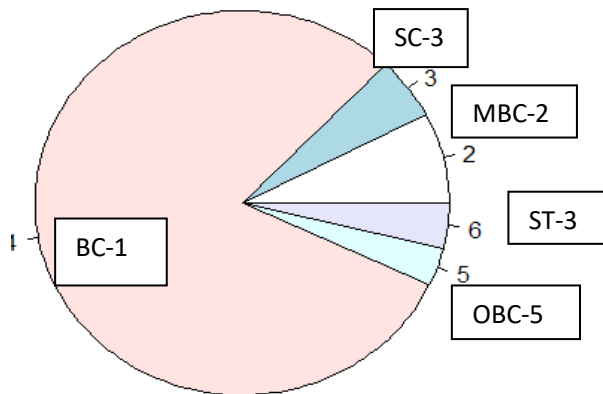


Fig 8 Community based shopping Analysis

Religion based Shopping:Theni district consists of 1011456 Hindus, 48066Muslims and 33830 according to the censuses 2015. So they come first on the religion based shopping.

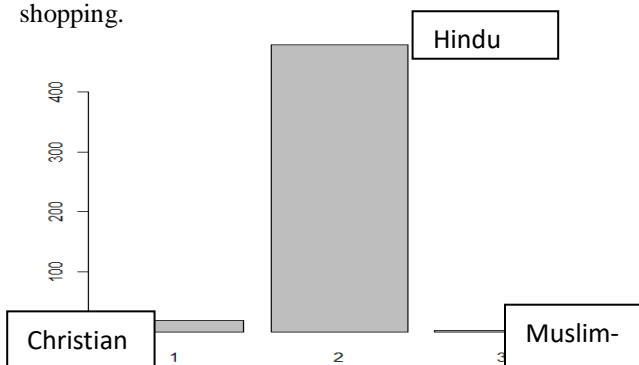


Fig. 9 Religion based shopping analysis

The multiple attributes are collected from the research tool and comparison of these variables is performed with the following commands.

Recommendation

Information collected by the researcher is analyzed by using R tool of Data mining. The retail market has a prosperous growth in Theni district as the quality of things, the customer services are concerned. As for us the geographical location of the next nearest city is concerned, people from Theni district has to travel minimum two hours to reach Madurai for all kinds of shopping. Moreover the rural villagers feel that they can shop in Theni rather than going to Madurai.

IX. LIMITATION OF THE STUDY

The study is based on different literatures and analysis of organized retail market. Retailing can be categorized as of different sectors like groceries, clothing and textiles, consumer durables, footwear, furniture and furnishing, catering services, jewellery and watches, books, music and gifts, mobile handsets and others.

Organized retailing is on continuous increase of its market share from the past. Lack of trained manpower, tax laws and government regulations and consumer awareness

delayed the growth of organized retailing. Until then the industry was dominated by the unorganized sector, it was a seller's market with a limited number of brands and little choice available to customers.

X. CONCLUSIONS

In the present study, the effects of store image variables, value observation, trust sensitivity and fulfillment on store loyalty were examined. The investigation was conducted in retail stores. However, this investigation is suggested to be carried out in different cities and regions where geographical and demographical variable are considered to affect the variables that determine the store loyalty. Therefore in the studies on the departmental stores, the variables of store loyalty can show differences. In our study, some effective variables about store loyalty or patronage behavior such as, social expectations, the social responsibility of the store, the cultural structure of customers were not included.

REFERENCES

- [1]. Akdogan, M. S., K. Gullu, and S. Babayigit, 2005, "A Study of Consumers' Perceptions Retail store," Erciyes University Journal of the Institute Social Sciences, 19:37-70.
- [2]. Sawmong, S. and O. Omar, 2004, "The Store of the UK's Retail Consumer," Journal of Academy of Business, 5:503-515.
- [3]. Andreassen, T. W., 1995, "(Dis) Satisfaction with Public Services: the Case of Public Transportation," The Journal of Services Marketing, 9:30-41.
- [4]. Arnold, S. J., J. Handelman, and D. J. Tigert, 1996, "Organizational Legitimacy and Store Patronage," Journal of Business Research, 35:229-239.
- [5]. Assael, H., 1992, Consumer Behavior and Marketing Action, Pws-Kent Publishing Company, Boston.
- [6]. PrafullaSudame and BrijeshSivathanuArthPrabhand, A Journal of Economics and Management Vol.2 Issue 8 August2013, ISSN2278-0629.
- [7]. Kalpana Singh (2012), "Retail Sector in India: Present Scenario, Emerging Opportunities and Challenges" IOSR Journal of Business and Management (IOSR-JBM) e-ISSN: 2278-487X, p-ISSN: 2319-7668. Volume 16, Issue 4. Ver. I (Apr. 2014), PP 72-81.
- [8]. SunitaSikri&DiptiWadhwa (2012) "Growth and Challenges of Retail Industry in India: An analysis" Asia Pacific Journal of Marketing and Management Review Vol.1 Issue 1, September 2012, ISSN 2319-2836.
- [9]. Gopal.V.V&Suryanarayana.A (2011), "Growth Drivers and Challenges for Organised Retailing in India" 2010 International Conference on Business and Economics Research vol.1 (2011) © (2011) IACSIT Press, Kuala Lumpur, Malaysia.
- [10]. VidushiHanda and Navneet Grover, International Journal of Multidisciplinary Research Vol.2 Issue 5 May 2012, ISSN 2231 5780.
- [11]. Er. JyotiVohra,Er. Jyoti "Data Mining Approach for Retail Knowledge Discovery". , Volume 6, Issue 3, March 2016 ISSN: 2277 128X
- [12]. Diptimayee Mishra, "Implications of Data Mining in Retail Sector of India: From the Recent Growth Perspective, IJCST Vol. 5, Issue Spl - 2, Jan - March 2014
- [13]. Chaturvedi Molly "A Study on the Opportunities and Challenges of Retail Sector in India", VOL.8 (3) 2017.
- [14]. D. Badma and Shanthi, "Challenges and Winning Strategies for organised Retailers", Volume 4, Issue 1, January 2016 pg. 49-58
- [15]. B. Jasmine and V. Selvakani, "Factors influencing the suicidal Ideation among Adelecents in Theni District", JASC, Volume V, Issue XII, December/2018.