Market Basket Analysis and Product Affinity in Retail

* Dr. N. Malati
** Ruchika
*** Bhawna Warikoo

Market Basket Analysis (MBA) identifies links involving items or connecting categories of items that have a propensity to procure jointly (complements) or between

Abstract

Summary: *In the current scenario of dynamic consumer behavior, it has become important for the retailers to understand the purchasing patterns of their customers and develop marketing strategies that develops customer loyalty. The introduction of electronic point-in sale increased the use and application of transactional data in market basket analysis. The major aim of Market Basket Analysis is to identify relationships in the form of association rules between groups of products, items or categories enabling the retailers manage their merchandise well and attract the customers for frequent purchases.*

Purpose: *The purpose of the study is to determine the pattern and associations among the lifestyle products such as apparels, watches, wallets, bags etc to better understand the buying behavior or pattern of the customers.*
Research Design: Research design used for this study is correlational or Prospective Research Design. It attempts to explore relationships among the commodities to make customer purchase predictions with the data of the year 2013-2014. The 109102 transaction data was purchased from the ecommerce website and processed for analysis and pattern recognition of purchase of lifestyle products such as apparels, watches, wallets, bags etc. Apriori algorithm and C5.0 algorithm is used to analyze the data. The customer profiling on the basis of top five performing baskets in the four metropolitan cities Delhi, Mumbai, Chennai& Kolkata are taken as sample for studying the customer demographics.

Findings: *Market basket analysis helped the ecommerce website to specifically target their customer segments with propensity of making frequent purchases.*

Practical Implications: *Retailers may use MBA to make advertising and promotions more predictable by understanding how buyers respond to different offers and communications vehicles. Longitudinal use of MBA allows ecommerce websites to characterize the buying behavior of customers over time.*

Keywords: *Market basket analysis, customer, product affinity, purchase patterns.*

INTRODUCTION

Market Basket Analysis:

In the current scenario, analyzing shopping baskets has become quite appealing to retailers. Advanced technology made it possible for them to gather information on their customers and their buying habits. The introduction of electronic point-in sale increased the use and application of transactional data in market basket analysis. In retail business analyzing such information is highly useful for understanding buying behavior. Mining purchasing patterns allows retailers to adjust promotions, store settings and serve their customers in a better way. Identifying buying rules is crucial for every successful business. Transactional data is always used to better understand the buying patterns of the customers and adjust the promotion and advertising efforts accordingly.

Market Basket Analysis (MBA) identifies links involving items or connecting categories of items that have a propensity to procure jointly (complements) or between

* Faculty, DIAS,
** Faculty, DIAS
*** Student, DIAS

items customers rarely purchase together (substitutes). The main goal of MBA is to identify relationships that are association rules between groups of products, items or categories. Wal-Mart also found a statistically significant relationship between the purchase of beer and nappies. As a result of this finding, the supermarket chain is alleged to have the nappies next to the beer, resulting in increased sales of both.

1.1 Strengths and weaknesses of Market Basket Analysis

One of the main advantages of market basket analysis is that it is perfect for undirected data mining. The majority of data mining techniques are not used for undirected data mining, while market basket analysis can be easily applied to analyze big data and provide the user with an appropriate start. Another major advantage and strength of the analysis is its operational simplicity. However, as the number of items and transactions increases, the computations needed to generate association rules grow very quickly and even exponentially. A possible solution for this problem is to reduce the number of items. This can be easily done by generalizing the items. Also, these tools remain expensive, time - consuming to deploy, and still require significant IT and application expertise

2 Business Intelligence and Tableau.

Tableau Software is an American computer software company headquartered in Seattle, Washington. It produces a family of interactive data visualization products focused on business intelligence. Tableau is a dynamic data visualization tool which has been used in this study for preparing market basket analysis dashboards in order to produce dynamic insights from the transaction data set of customers.

Organizations are storing a variety of unstructured data from websites, infrastructure logs and sensors, and ecommerce velocity gives less and less time to interpret and act on information. It helps one to connect to one or more than 45 databases and formats supported by Tableau also publish web dashboards with live connections on corporate portal, SharePoint or wiki so that data is automatically refreshed. Mix-and-match multiple data sources from different database types in the same web dashboard.

3 Business Intelligence and SPSS Modeler

IBM SPSS Modeler is an extensive predictive analytics platform that is designed to bring predictive intelligence to decisions made by individuals, groups, systems and the enterprise. By providing a range of advanced algorithms and techniques that include text analytics, entity analytics, decision management and optimization, SPSS Modeler can help the organisations to take the right decisionsfrom the desktop or within operational system.

IBM SPSS MODELER COGNOS has been used in this research. It has helped in running the Apriori and C5.0 algorithm to give the analytical solution of the business problem.

Market Basket analysis promises to be the next step in the progression of ecommerce merchandising and promotion. Ecommerce websites has big databases which became challenging task in analyzing the data. The ecommerce merchants prefer to outsource the work to Research and Analytical companies to solve their data related problems and provide them various insights from their data and help them in gaining competitive edge in the industry. Annik Technology Services is one such company that provides Business Intelligence services. Annik customize its client data into meaningful insights by creating dynamic dashboards which represents the Big-Data into charts and graphs. The current study is conducted under the assistance of Annik Technology services.

Objective of the study:

The objective of this study is to understand product associations and buying pattern of the customers through Market Basket Analysis

LITERATURE REVIEW

1 Association Rule Mining

Association rule mining is to discover the associations and relations among items of sets of large data. It is a data mining technique which helps in finding affinity between items in from a large set of transaction data. The identified data forms the base for market basket analysis. Market basket analysis helps in analyzing buying habits of customer by focusing on the items they buy together and place them in their baskets. The discovered associations help retailers in tailoring their promotion strategies and strategically plan their inventor. An attempt is made to associate one product with another so whenever one product is purchased, the other is also purchased.

2 Market Basket Analysis

Market basket analysis is a technique that discovers relationships between pairs of products purchased together. It discovers interesting technique which can be used to find cross-sells opportunities between related products. The idea behind market basket analysis is simple. Simply examine the orders for products that have been purchased together. For example using market basket analysis we might uncover the fact that customers tend to buy hot dogs and buns together. Using this information the retailer might organize the store so that hot dogs and buns are next to each other. Market basket analysis started off with pairing products with each other which grew into analyising the cross category effect of prizing, the promotional effects and the bundling of products and offering discounts. If a supermarket knows that bread and butter tend to be purchased together, it can avoid offering price discounts on both at the same time. Almost all available literature have attempted to address the what, and how part of Market Basket Analysis. Results obtained reveal that retailers are using MBA to develop more profitable advertising and promotions target that can attract into the stores and increase the size and value of the basket of purchases among other things

Table 2.1 provides an overview of existing literature and methodology on market basket analysis.

2.1 TABLE OF RECORDS

Method and Selected References	Characteristics of the Analysis	Primary Task of the Analysis
Market basket analysis Beauty Products (Velislava Gancheva ,Bruno Jacobs 2013)	Exploratory	Mining association rules
Market Basket Analysis for a Supermarket (Loraine Charlet Annie M.C. and Ashok Kumar D, 2012)	Exploratory	Frequent Item set Mining
Bogdan Hoanca, Kenrick Mock,(2011)	Exploratory	Estimate Potential Revenue Increases for a Small University Bookstore
Larry Gordon, Partner The FactPoint Group. (2008)	Exploratory	Leading practices of Increasing market share by market basket analysis
Finite Mixture Model (Rick L Andrews , Imran S. Currim, 2002)	Exploratory	Identification of customer preference segments .
Multivariate Logistic Model (Gary J Russell, Ann Petersen, 2000)	Exploratory	Estimate and predict cross -category effects .
Inter-category Choice Dynamics (Pradeep K Chintagunta and SudeepHaldar, 1998)	Exploratory	Analysing purchase timing across categories .
Association Rules (Robert J. Hilderman, Colin L Carter, Howard J. Hamilton, and Nick Cercone), (Rakesh Agrawal, Sirkant Ramakrishnan, 1996)	Exploratory	Discovery of association rules
Pairwise Associations (Julander, 1992)	Exploratory	Represent relationships

RESEARCH METHODOLOGY

Research design used for this study is correlational or Prospective Research Design. It attempts to explore relationships among the commodities to make customer purchase predictions. The 109102 transaction data was purchased from the ecommerce website for analysis and pattern recognition of purchase of lifestyle products such as apparels, watches, wallets, bags etc.

With the rise of social media providing an additional platform for showing off or simply amplifying a stylized version of today's consumer, the lifestyle products have got a major role to play in any consumers life. It offers consumers a way to reinforce or supplement their identity by publicly associating themselves with it. So the current study has specially considered lifestyle products as consumer tend to purchase it more often with medium or high involvement.

Apriori algorithm and C5.0 algorithm are being used to analyze the data. The customer profiling on the basis of top five performing baskets in the four metropolitan cities Delhi, Mumbai, Chennai and Kolkata are taken as sample for studying the customer demographics. The analysis has been conducted on the data from the year 2013-14.

1. APRIORI ALGORITHM

R. Agarwa and R. Srikan(1994) raised the algorithm of Boolean association rule of mining frequent item sets. Using data mining techniques on transactional data leads to the generation of association rules and finds correlations between products in the records. The main concept of association rules is to examine all possible rules between items and turn them into 'if-then' statements.

Definition

Let I = {i1,i2,i3,. . . , im}is the set of all items available at the store.

By T = {t1,t2,t3, . . . , tn}defines the set of all transactions in the store.

Each transaction ti = {i2,i4,i9}contains a subset of items from the whole market basket dataset.

An itemset is every collection of zero or more items from the transaction database.

The number of items that occur in a transaction is called a transaction width.

Transaction tj contains an item set X if X is a subset of tj (X tj).

An association rule can be expressed in the form of X →Y, where X and Y are two disjoint itemsets (do not have any items in common).

X is an antecedent and Y is a consequent, in other words, X implies Y.

The main concept of association rules is to examine all possible rules between items and turn them into 'if-then' statements. In this case the 'if' part is X or the antecedent, while the 'then' part is Y or the consequent.

Antecedent → consequent [support, confidence]

The antecedent and consequent are often called rule body and rule head accordingly. The generated association rule relates the rule body with the rule head. There are several important criteria of an association rule: the frequency of occurrence, the importance of the relation and the reliability of the rule. Association Rule Mining (ARM) is about two basic parameters, support and confidence. They both measure the strength of an association rule. Support of an association rule is the percentage support of records that contain X U Y to the total number of records in the database. In this measure of strength, quantity is not taken into account. The support count increases by one for each time the item is encountered in a different transaction T from the database D.

Support can be derived from the following formula:

Support (XY) = support count xy/total number of transaction D

If the support of X and Y (a set of items) is 10%, it means that X and Y appear together in 10% of the transactions.

Confidence of an association rule is defined as the percentage of the number of transactions that contain XUY to the total number of records that contain X. In other words, confidence is a measure of the strength of association rules and is used to determine how frequently items from itemset Y appear in transactions that contain itemset X. Let's suppose a rule X →Y. Confidence demonstrates how likely it is to find Y in a transaction that contains X.

There is no association between X and Y if the importance is 0. If the importance score is positive, this means that the probability of Y increases when X is true. A negative importance score says the opposite: the probability of Y decreases when X is true.

It is also known as a Weight-of-Evidence (WOE). The importance is derived by the following

Formula:
I (X→Y) = log (P(X|Y) / P(Y| not X)

Generated rules can be grouped into rules that have direct and rules that have indirect relationships. If two rules, say R1 and R2, share at least one item (no matter if it is in the rule body or rule head), they belong to the same rule group and they are directly related. Indirectly related rules are such rules that do not contain the same item in both the rule body and rule head.

The association rules problem can be easily defined as it follows:

Given a threshold S (the minimum support) and a threshold c (the minimum confidence), the store intends to find all rules in the form of X → Y, where X and Y are sets of items, such that:

1. X and Y appear together in at least s% of the transactions.
2. Y occurs in at least c% of the transactions, in which X occurs. A given association rule is supported in the database, if it meets both the minimum support and minimum confidence criteria.

Formula:
Confidence(X/Y) = support xy/ support y
Lift measures the importance of a rule. The lift value is represented as the ratio of the confidence and the expected confidence of a rule. The lift can take over values between zero and infinity. In every association rule there is an antecedent and a consequent, also called rule body and rule head accordingly.

Rule body [item1] + [item2] → Rule head [item1 & item 2]

If the value of the lift is greater than 1, this means that both the rule body and the rule head appear more often together than expected. The occurrence of the rule body positively affects the occurrence of the rule head. The other way around, if the lift value is lower than 1, this means that both the rule body and rule head appear less often together than expected and the occurrence of the rule body negatively affects the occurrence of the rule head. However, if the lift value is near 1, the rule body and rule head appear together as often as expected. (Lift in an association rule)

Lift can be derived from the following formula:
L(X→Y) = c(X→Y) / P(Y) = P (X,Y) / (P(X)P(Y))

2. C5.0 AGORITHM

C5.0 is a data mining technique that gives more accurate and efficient result. The classification process under C5.0 generates fewer rules with low error rate in comparison to other techniques. In this research work proposed system use C5.0 classifier that performs feature selection and reduced error pruning techniques.

The new algorithm C5.0 represents that how the rule sets are generated with improved features.

This research work focuses on high accuracy and high prediction. The probability of event X occurring given that event Y has occurred (P(X|Y)) is proportional to the probability of event Y occurring given that event X has occurred multiplied by the probability of event X occurring ((P(Y|X)P(X))..

A rule-based classification technique used a collection of if ... then ... rules for classifying records. Rule-based classifier extracts a set of rules that show the relationships between the attributes of a dataset and the class label. Here the part of a rule is known as the antecedent of rule. The then part of the rule is consequent of rule. The Coverage of a rule is the number of instances that satisfy the previous rule. The Accuracy of a rule is the fragment of instances that satisfy both the previous and consistent rule, normalized by those satisfying the antecedent. Ideal rules should have both high coverage and high accuracy rates.

DATA ANALYSIS AND INTERPRETATION

The data analysis is divided into two parts. In the first part, the data is analyzed through aprori algorithm and in the second part the output of the first analysis is used as an input to an algorithm C5.0 which classified the data in accordance with the demographics variables of the buyers.
In the following table 1, the output of apriori algorithm is shown which has interpreted the 109102 transaction data purchased from the ecommerce website into antecedents and consequents with respect to support, confidence and lift.

Table 1 : Top 25 Baskets with Antecedents and Consequents

	Antecedent	Consequent	Support %	Confidence %	Lift
1	Winter Wear and Trousers Jeans	T-Shirts Shirts	9.00	82.78	2.04
2	Wallets Bags and Trousers Jeans	T-ShirtsShirts	7.57	81.89	2.01
3	Wallets Bags and Winter Wear	T-Shirts Shirts	7.33	81.30	2.00
4	Watches and Trousers Jeans	T-Shirts Shirts	7.75	76.15	1.87
5	Slippers Sandals	T-Shirts Shirts	9.48	76.15	1.73
6	Wallets Bags and Winter Wear	Trousers Jeans	7.33	65.04	2.72
7	Sports Shoes	T-Shirts Shirts	12.69	64.79	1.59
8	Sports Shoes and T-shirts Shirts	Trousers Jeans	8.22	63.77	2.67
9	Ethnic Wear	T-Shirts Shirts	12.63	63.68	1.57
10	Wallets Bags	T-Shirts Shirts	17.52	63.27	1.56
11	Wallets Bags and Trousers Jeans	Winter Wear	7.57	62.99	3.16
12	Formal Shoes	T-Shirts Shirts	11.98	62.19	1.53
13	Winter Wear	T-Shirts Shirts	19.96	61.79	1.52
14	Formal Shoes and T-Shirts Shirts	Trousers Jeans	7.45	61.60	2.58
15	Kurtas Kurtis Suits and Tops Tees Shirts	Trousers Jeans	9.00	60.93	1.50
16	Ethnic Wear and T-Shirts Shirts	T-Shirts Shirts	8.05	60.74	3.04
17	Winter Wear and T-Shirts Shirts	Trousers Jeans	12.34	60.39	2.53
18	Ethnic Wear and T-Shirts Shirts	Trousers Jeans	8.05	59.26	2.48
19	Winter Wear and T-Shirts Shirts	T-Shirts Shirts	20.38	58.48	1.44
20	Ethnic Wear and T-Shirts Shirts	T-Shirts Shirts	23.90	58.10	1.43
21	Watches	Wallets Bags	7.45	57.60	3.29
22	Trousers Jeans	Trousers Jeans	11.08	55.91	2.34
23	Winter Wear and Trousers Jeans and T-Shirts Shirts	Winter Wear	7.75	54.62	2.74
24	Wallets Bags and T-Shirts Shirts	Winter Wear	11.08	53.76	2.69
25	Watches and Trousers Jeans	Winter Wear	13.89	53.65	2.69

The output is again used as an input to an algorithm C5.0 which classified the data in accordance with the demographics variables of the buyers. The algorithm helped in determining the characteristics of the customers buying the products. The data from the above table has been interpreted into different insights. For explaining the relationship between the products and the various benefits derived from them, top five performing baskets are taken in descending order on the basis of their lifts. The various insights are explained in the subsequent sections.

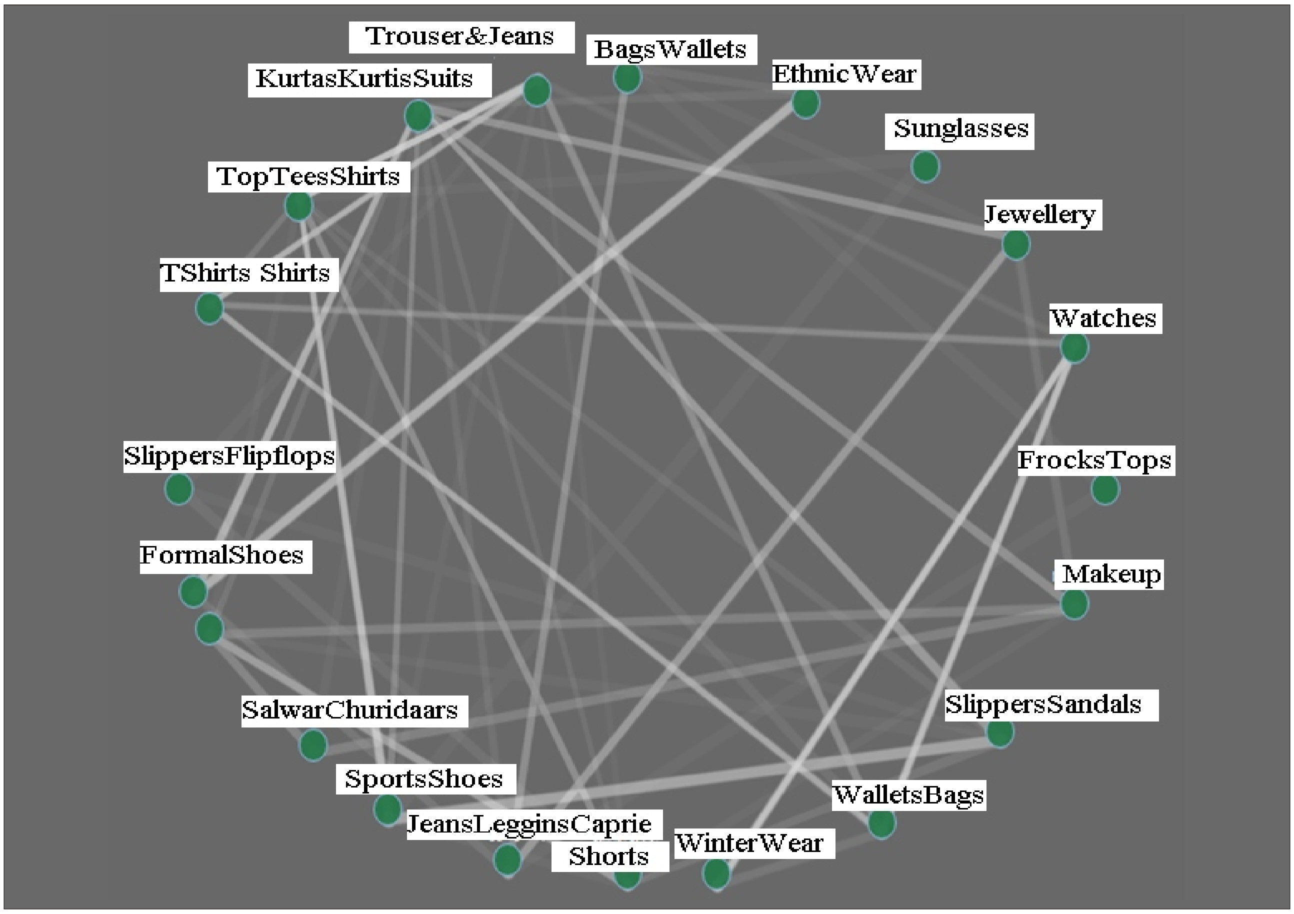
IMPORTANT CONTRIBUTIONS OF APIORI ALGORITHM AND C5.0

1 Associations between Products

The Apriori algorithm helps in providing the associations between the product categories and highlights the relationship between two or more item based on their lifts. This implies that products linking with each other have an association between them, which can be exploited for the purpose of product bundling and giving price discounts.

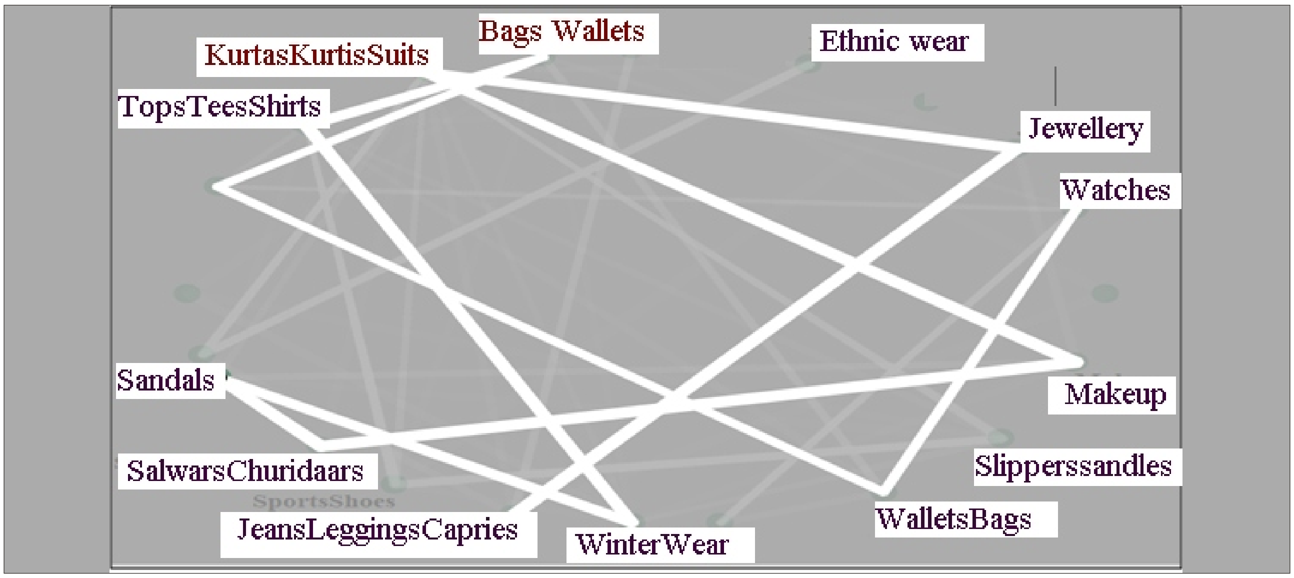
The network graph (snapshot 1.1) shows the pictorial representation of the tabular data highlighting the antecedents and consequents. It shows the product categories that are linked with each other on the basis of their lift. Higher the lift, thicker is the line adjoining them.

Snapshot 1.1 Selecting Product Category



The associations can be direct as well as indirect between the product subcategories. The main category of the product may not be related but product subcategories may have the association directly or indirectly (snapshot 1.2).

Snapshot 1.2 Selecting Product subcategory



This can also serve as the basis to judge the purchasing pattern of sub categories of product by the customers.

2 Customer Profiling

The customer profiling is a way to create a portrait of the customers to help the business to make design decisions concerning their service. The customers are divided into groups on the basis of similar goals and characteristics and each group is given a representative with a photo, a name, and a description. A small group of customer profiles or 'personas' are then used to make key design decisions with, e.g. “which of these products will help middle aged man's searching easier.” Further the clustering technique can be used to help the client in finding the different segments of the customer and their buying behavior on the basis of their demographic details. This helps the retailer in reducing its selling cost and time.

In this insight the client can view the demographics of its buyers, buying its five best performing baskets on the basis of state, city, gender, and marital status. The client can customize its segmenting, targeting and positioning strategies accordingly. This is one of the most important insights which can be extracted from the market basket analysis. Customer profiling helps in sending customized mails and recommendations in accordance with the buying behavior of the customer and the firm's sales strategies.

For instance in the current research (Snapshot 2), the sales of all the five baskets in the Tier-A-1 cities (Chennai, Mumbai, Kolkata and Delhi) it's found that Basket 1 has 25,493 buyers in the four metropolitan cities. It has the highest number of buyers in Chennai (12,256) followed by New Delhi (7,667), Mumbai (4,661) and Kolkata (909). Basket 2 has 23,876 buyers in the four metropolitan cities.It has the highest number of buyers in Chennai (11,752) followed by New Delhi (7,232), Mumbai (4,163) and Kolkata (729).Basket 3 has 22,583 buyers in the four metropolitan cities. It has the highest number of buyers in Chennai (11,274) followed by New Delhi (6,839), Mumbai (3804) and Kolkata (666). Basket 4 has 22,617 buyers in the four metropolitan cities. It has the highest number of buyers in Chennai (12,028) followed by New Delhi (5,451), Mumbai (4,389) and Kolkata (749).Basket 5109,107 buyers in the four metropolitan cities. It has the highest number of buyers in Chennai (9,322) followed by New Delhi (5,801), Mumbai (3,392) and Kolkata (592).It can be concluded that Chennai has the highest customer base of all the four baskets and Kolkata has the smallest. The total male buyers are 95,307 and female are 20,701 in the metropolitan. The married buyers are 47999 and unmarried 56,770 in the metropolitan.

From the above inferences it can be concluded that Basket 1 is the highest performing basket among all the baskets in these four states followed by Basket 2, Basket 4, Basket 3 and finally Basket 5. Also the male buyer has outnumbered the female buyers in these four states. The Single buyers are buying more from this online retail store when compared with the married buyers.

Snapshot 2: CUSTOMER PROFILING

BASKET 1 : T-shirts & shirts, Trousers & Jeans, Winter Wear

City	Marital Status	Gender	
New Delhi	Married	Female	901
		Male	2,908
	Single	Female	578
		Male	3,280
Mumbai	Married	Female	361
		Male	1,777
	Single	Female	225
		Male	2,098
Kolkata	Married	Female	108
		Male	358
	Single	Female	94
		Male	349
Chennai	Married	Female	1,329
		Male	5,170
	Single	Female	697
		Male	5,060

BASKET 2 : Wallets Bags and Trousers Jeans, Tshirts Shirts

City	Marital Status	Gender	
New Delhi	Married	Female	671
		Male	2,860
	Single	Female	413
		Male	3,288
Mumbai	Married	Female	303
		Male	1,648
	Single	Female	223
		Male	1,989
Kolkata	Married	Female	93
		Male	363
	Single	Female	35
		Male	238
Chennai	Married	Female	1,341
		Male	4,788
	Single	Female	631
		Male	4,992

BASKET 3 : Wallets Bags, Winter Wear & T-Shirts Shirts

City	Marital Status	Gender	
New Delhi	Married	Female	743
		Male	2,716
	Single	Female	309
		Male	3,071
Mumbai	Married	Female	230
		Male	1,648
	Single	Female	83
		Male	1,843
Kolkata	Married	Female	93
		Male	300
	Single	Female	35
		Male	238
Chennai	Married	Female	1,144
		Male	4,726
	Single	Female	667
		Male	4,737

BASKET 4 : Watches, Trousers Jeans and T-Shirts Shirts

City	Marital Status	Gender	
New Delhi	Married	Female	1,252
		Male	2,710
	Single	Female	869
		Male	3,059
Mumbai	Married	Female	483
		Male	1,637
	Single	Female	327
		Male	1,942
Kolkata	Married	Female	137
		Male	320
	Single	Female	93
		Male	292
Chennai	Married	Female	1,595
		Male	4,794
	Single	Female	1,090
		Male	4,549

BASKET 5 : Slippers Sandals& T-Shirts Shirts

City	Marital Status	Gender	
New Delhi	Married	Female	714
		Male	2,297
	Single	Female	371
		Male	2,419
Mumbai	Married	Female	344
		Male	1,140
	Single	Female	275
		Male	1,633
Kolkata	Married	Female	29
		Male	235
	Single	Female	35
		Male	293
Chennai	Married	Female	1,173
		Male	3,799
	Single	Female	607
		Male	3,743

3 Performance of Baskets

The performance of the baskets can be analyzed over days, week or months. From this insight a retailer can determine
a) The performing and non performing baskets
b) The cannibalism effects

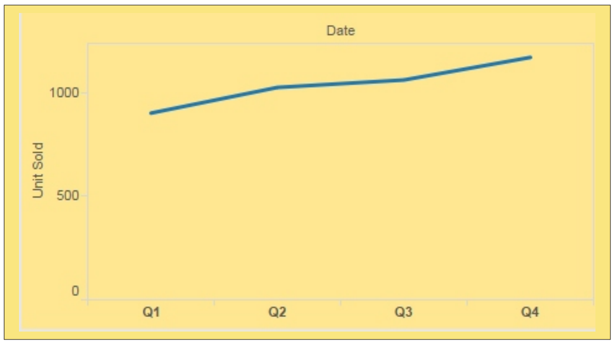
This makes the retailer in a better position to analyze the most and least profitable basket and accordingly take their marketing and advertising decisions.

In the current study, after analyzing every basket (Snapshot 3) it's concluded that the five baskets had similarity in their sales over the period with negligible differences in their performances. The baskets showed gradual increase over the quarters.

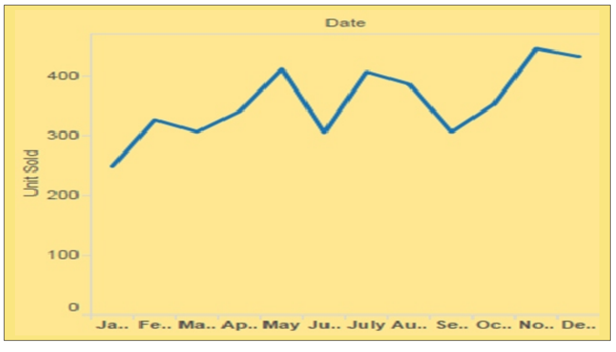
Snapshots 3: TOP 5 BASKET PERFORMANCE

BASKET 1-Tshirts &shirts, Trousers & Jeans, Winter Wear

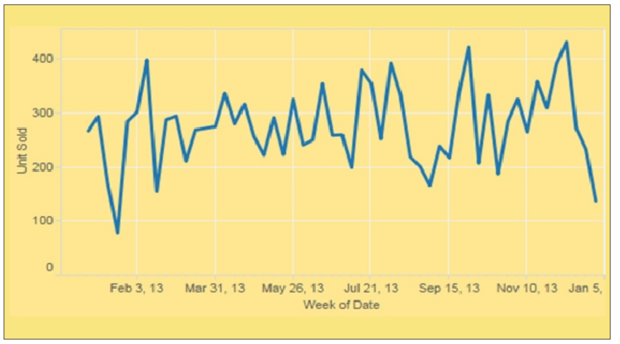
Performance per quarter (Year 2013-14)



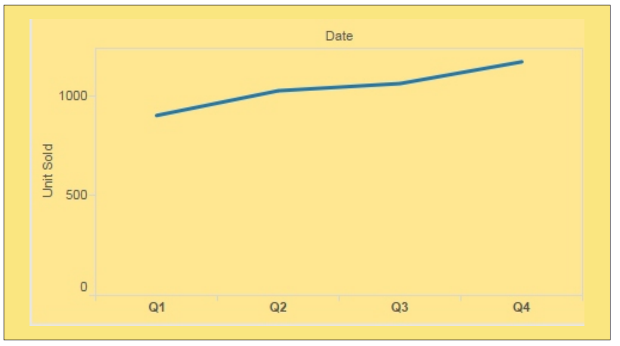
Performance per month (January 2013-December 2013)



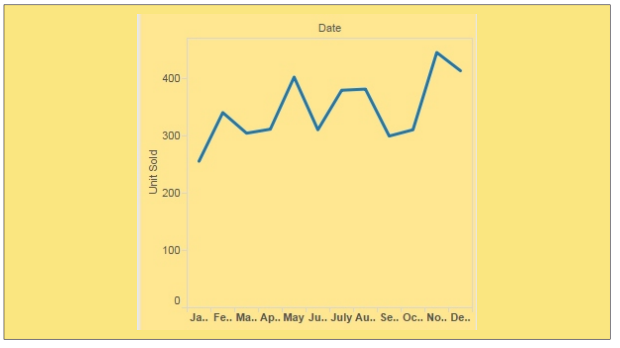
Performance per week (February 2013- January 2014)



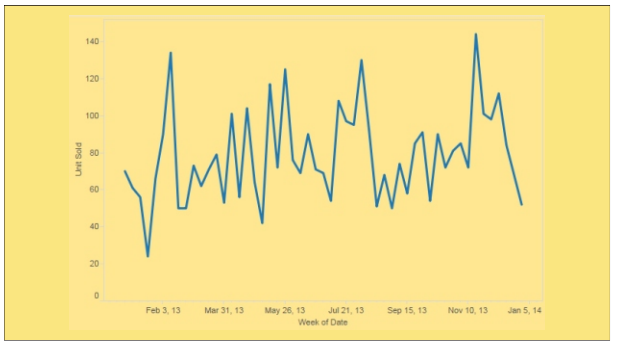
BASKET 2- Wallets Bags and Trousers Jeans, T-Shirt Performance per quarter (Year 2013-14)



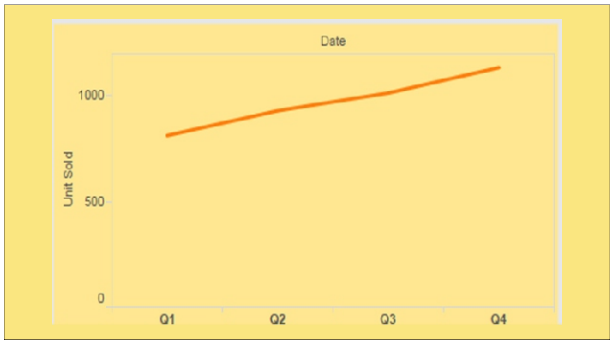
Performance per month (January 2013-December 2013)



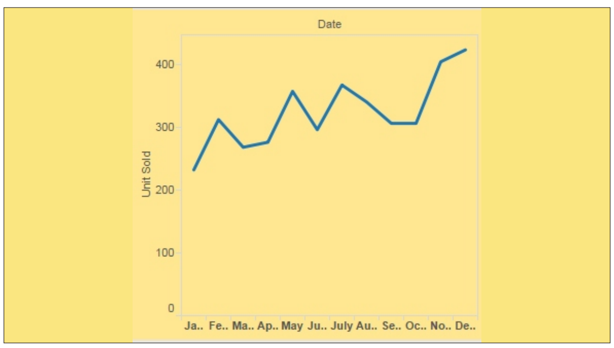
Performance per week (February 2013- January 2014)



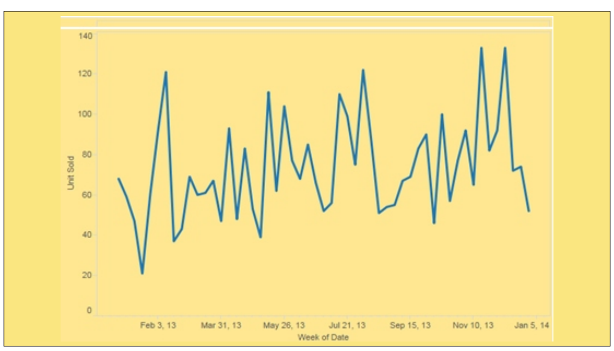
BASKET -3 Wallets Bags, Winter Wear& T-Shirts Shirts Performance per quarter (Year 2013-14)



Performance per month (January 2013-December 2013)

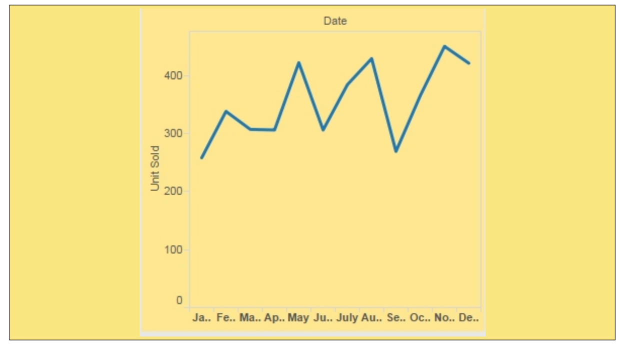


Performance per week (February 2013- January 2014)

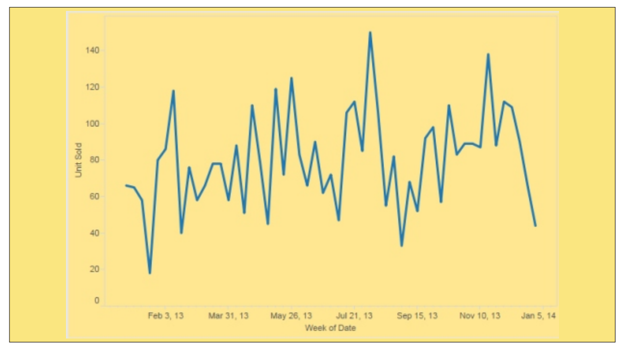


BASKET 4 Watches, Trousers Jeans and T-Shirts Shirts Performance per quarter (Year 2013-14)

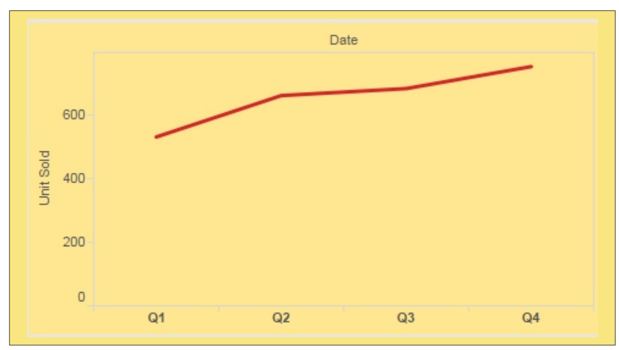
Performance per month (January 2013-December 2013)



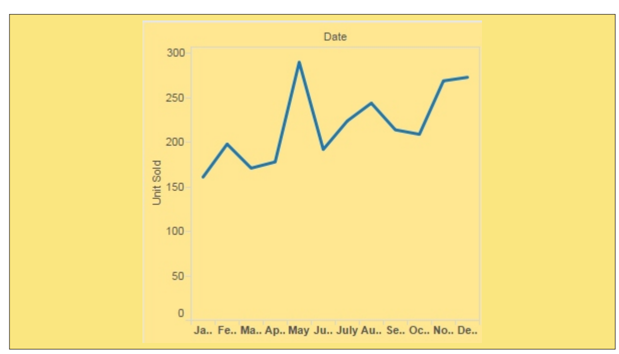
Performance per week (February 2013- January 2014)



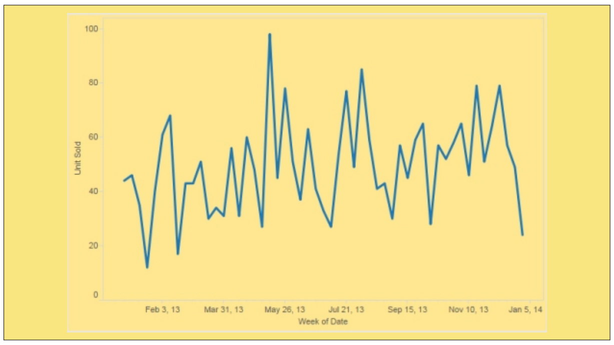
BASKET -5 Slippers Sandals& T-Shirts Shirts Performance per quarter (Year 2013-14)



Performance per month (January 2013-December 2013)



Performance per week (February 2013- January 2014)



4 Cross Sales

Cross-selling is the action or practice of selling an additional product or service to an existing customer. In practice, businesses define it as a practice of selling or suggesting related or complimentary products to a prospective customer at the time of making purchases. It is generated on the basis of affinity analysis performed after aprori algorithm and is further given to clients for easy referral and analysis. Client can flip through the antecedents and find out the five corresponding consequent for each product category or subcategory. Whenever any antecedent is selected it shows five of its consequent on the right side of the window. The client can note these relationships and manage their products and store in a way to guide the purchases of their customers and increase their own sales.

CONCLUSION

Progressive ecommerce websites see MBA as a strategic tool that will help them increase their success and provide them with the edge that they need. By using market basket analysis, leading ecommerce websites are increasing their competitiveness by focusing directly on the customer's buying habits, and then using that knowledge to quickly tailor their operations to the changing needs of their customers and trade areas. In traditional brick and mortar business, seller has an option of convincing the buyers which is not possible in case ecommerce websites where there are no face to face interactions of buyers and sellers. Also in traditional businesses the customers had less variety and knowledge of the products, but with the advent of technology specially ecommerce websites customers are highly aware and new websites are coming up at a rapid rate. This has led to the increase in need of business intelligence and analytical tools. Most e-retailers admit that they traditionally have been product focused and know surprisingly little about their customers. Their desire for more insight into customer buying behavior has driven many to invest in business analytics solutions. Additionally, leading ecommerce websites are beginning to equip buyers, merchandisers and planners, with powerful and convenient market basket analysis tools, improving success across the board.Market basket analysis can very well help the ecommerce website to specifically target their customer segments which can be converted into customers by understanding the individual's buying behavior over the period of purchase.

LIMITATIONS OF THE STUDY

Although market basket analysis is computationally simple and very efficient, there are several important limitations. The time constraint made it impossible to measure the impact of the study. Also, the limited knowledge of the ecommerce business acted as a constraint in applying price related experiments.

REFERENCES

1) B Hoanca, K Mock (2011), US Patent 7,986,816, Methods and Systems for Multiple Factor Authentication Using Gaze Tracking And Iris Scanning.
2) Charlet Annie and Ashok Kumar (2012), Market Basket Analysis for a Supermarket based on Frequent Itemset Mining, IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 5, No 3, September 2012.
3) Gancheva, V.K. (2013). Market basket analysis of beauty products. Economics. Retrieved from http://hdl.handle.net/2105/
4) Gary J Russell, Ann Petersen. (2000). Analysis of Cross-Category Dependence in Market Basket Selection. Journal of Retailing, Vol.76 (3), 367-392.
5) Julander, C.-R. (1992). Basket Analysis: A New Way of Analysing Scanner Data. International Journal of Retail and Distribution Management, Volume 20 (7), 10-18.
6) Larry Gordon, Partner The FactPoint Group. (2008), http://www.irgintl.com/pdf2/1.pdf
7) Leading Practices in Market Basket Analysis: How Top Retailers are Using Market Basket Analysis to Win Margin and Market Share." Www.irgintl.com/pdf2/1.pdf
8) M. Khattak, A. M. Khan, Sungyoung Lee and Young-Koo Lee. (2010). Analyzing Association Rule Mining and Clustering on Sales Day Data with XLMiner and Weka. International Journal of Database Theory and

Application, Vol. 3, No. Uclab.khu.ac.kr/resources/ publication/J_98.pdf

9) Pradeep K. Chintagunta, Inseong Song. (2006). Measuring Cross-Category Price Affects with Aggregate Store Data. Management Science, Vol. 52, No. 10, 15941609.
10) R. Agrawal, H. Mannila, R.Srikant, H.Toivonen, and A.I. Verkamo. Fast discovery of association rules. In U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, editors, Advances in Know ledge Discovery and Data Mining, pages 307{328, Menlo Park, CA, 1996. AAAI Press/MIT Press.
11) Rick L. Andrews, Imran S. Currim. (2002). Identifying segments with identical choice behaviors across product categories: An Inter category Logit Mixture model. International Journal of Research in Marketing, 19, 6579.
12) Robert (2014), The Relationship between (4Ps) & Market Basket Analysis. A Case Study Of Grocery Retail Shops In Gweru Zimbabwe, International Journal of Scientific & Technology Research, Volume 3, Issue 10.
13) Robert J. Hilderman, Colin L. Carter, Howard J. Hamilton, and Nick Cercone. Mining Market Basket Data Using Share Measures and Characterised Itemsets, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10 .1.1.21.9550&rep=rep1&type=pdf.
14) Velislava Gancheva, Bruno Jacobs (2013), Market basket analysis Beauty Products, Master of Science in Economics and Business, Erasmus University Rotterdam, Erasmus School of Economics.