IOSR Journal of Business and Management (IOSR-JBM)
e-ISSN: 2278-487X, p-ISSN: 2319-7668
PP 06-21
www.iosrjournals.org

Innovation Regarding Purchase Pattern Of Personal Care Products In India Using Market Basket Analysis

Ms. Debadrita Panda [1], Mr. Dipto Halder[2]
[1] *MBA(IISWBM),Assistant Professor, Brainware University.*
[2]*MBA(IISWBM),Executive Trainee, m-junction services Ltd.*

Abstract:
Objective*: Companies nowadays have a vast collection of data but not up to the mark in information extracted from that data. Big data is a valuable resource and although the concept is still new, companies in a variety of industries are relying on data mining for making strategic decisions. Facts that otherwise may go unnoticed can be now revealed by the techniques that sift through stored information.*
Market basket analysis is a very useful technique for finding out co-occurring items in consumer shopping baskets. Such information can be used as a basis for decisions about marketing activity such as promotional support, inventory control and cross-sale campaigns.
The main objective is to see how different products in a personal care store interrelate and how can these relations be exploited for marketing activities. Mining association rules from transactional data can provide us with valuable information regarding co-occurrences and co-purchases of products. Such information can be used as a basis for decisions several marketing activities such as promotional support, inventory control and cross-sale campaigns.
Approach*: Primary data of nearly 2000 people are collected using a structured questionnaire. Next it has been converted to a transactional dataset which has been used for association rule mining.*
Methods*: Market Basket Analysis using R has been implemented over here.*
Findings*: By setting parameters like support, confidence and lift, all the transactions have been analyzed and associations are noticed. Only sets having greater value compared to threshold support are taken for final association rule mining.*
Conclusion*: If the store space is utilized in a way keeping in mind the best possible combination of the products which are frequently bought by the customers, it will result in maximum returns & hence, a profitable business.*
Keywords - *Market Basket Analysis, Personal Care Products, Apriori Algorithm.*

I. Introduction:

Market Basket Analysis is a common and useful method of data analysis for marketing and retailing. The purpose of market basket analysis is to determine what products customers purchase together. The name comes from the idea of customers throwing all their purchases into a shopping cart (a "market basket") during grocery shopping. Knowing what products people purchase as a group can be very helpful to a retailer or to any other company. A store could use this information to place products frequently sold together into the same area, while a catalogue or World Wide Web merchant could use it to determine the layout of their catalogue and order form (e.g.- The "Frequently purchased together" tab popping every time we order purchase something online). Direct marketers could use the basket analysis results to determine what new products to offer to their prior customers.

Sometimes, items which can be sold together are obvious – e.g. every fast-food restaurant asks their customers "Would you like fries with that?" whenever they go through a drive-through window. However, sometimes the fact that certain items would sell well together is not so obvious. A well-known example is that a supermarket performing a basket analysis discovered that diapers and beer sell well together on Thursdays. Though the result does make sense – young couples stocking up on supplies for themselves and for their children before the weekend starts – it's not the sort of thing that someone would normally think of right away. The strength of market basket analysis is that by using computer data mining tools, it's not necessary for a person to think of what products consumers would logically buy together – instead, the customers' sales data is allowed to speak for itself. This is a good example of data-driven marketing.

Once it is known that customers who buy one product are likely to buy another, it becomes easier for the company to market the products together, or to make the buyers of one product the target prospects for another product. If customers who purchase diapers are likely to purchase beer, then a beer display just outside the diaper aisle will increase that probability. Likewise, if it's known that customers who buy a sweater and casual pants from a certain mail-order catalogue have a propensity towards buying a jacket from the same catalogue, sales of jackets can be increased if telephone representatives describe and offer the jacket to anyone who calls in to order the sweater and pants. Still better, the catalogue company can even provide an additional 5% discount on a package containing the sweater, pants, and jacket simultaneously and promote the complete package. The dollar amount of sales is guaranteed to go up this way. By targeting customers who are already known to be likely buyers, the effectiveness of marketing is significantly increased – regardless of if the marketing takes the form of in-store displays, catalogue layout design, or direct offers to customers. This is the purpose of market basket analysis – to improve the effectiveness of marketing and sales tactics using customer data already available to the company.

II. Literature Review

Over the past two decades a lot of attention has been devoted to the subject of data mining. While retailers are involved in this topic because of the absolute utility of market basket data, market analysts are interested because of the research and technical challenges they face while analyzing the data.

Increasing amount of data is being generated every second and this allows experts to search for meaningful associations among customer purchases. Customers make purchase decisions in several product categories on a single shopping trip. Interdependencies among products have faced increased attention recently as retailers are trying to improve their businesses by applying quantitative analyses to their data.

It is very important for retailers to get to know what their customers are buying. Some products have higher affinity to be sold together and hence the retailer can benefit from this affinity if special offers and promotions are developed for these products. It is also important to the retailer to cut off products from the assortment which are not generating profits. Deleting loss-making, declining and weak brands may help companies boost their profits and redistribute costs towards aspects of the more profitable brands. (Kumar, 2009) This is yet another reason why data mining is seen as a powerful tool for many businesses to regularly check if they are selling too many brands, identify weak ones and possibly merge them with healthy brands. Data mining techniques are highly valued for the useful information they provide so that the retailer can serve customers better and generate higher profits.

Chris Anderson in his book ‘The long tail: Why the future of business is selling less of more’ explains a concept of the ‘98% rule’, which is quite contrasting to the well-known 80/20 rule. In other words, 2% of the items a retailer sells are frequent, while 98% of the items have very low frequencies, which create a long tail distribution. This is why the presence of this ‘98% rule’ in the retail business created the need for data mining software and made quantitative analysis a must for retailers.(Anderson, 2006)

1. Find products with affinity to be sold together: A lot of research has been done in marketing to show that there are demand interdependencies among certain related products within a single store. Retailers tend to exploit this tendency by adjusting price promotions in a profit-maximising way. They can also exploit these product associations by incorporating them into promotional strategies.

Analyzing purchases in multiple categories allows retailers to benefit from promotion and other marketing activities. Incorporation of product interdependencies into a pricing strategy is an effective way of boosting profits.

For example, Mulhern and Leone(1991) studied the impact of price promotions on cake mix and cake frosting. Their main objective was to evaluate the overall profitability of implicit price bundling. Reducing the price of cake mix increase purchases of both cake mix and frosting and the overall profit improves. The study shows how promotions have positive impact on the sales of a complementary product.

Finding associations between product purchases is an effective way to adjust price promotions better and make better predictions on the effect of price bundling. Also, it is important to keep product complementarities in mind when making promotions. Complementary products often sell well together but this does not mean that they are a pair and a price increase in one of the set will not affect sales of the other one. Complementarity gives managers control over their customers’ buying behaviour, but co-occurrence of specific product categories in a single shopping basket is less controllable. Market basket analysis reveals all the underlying patterns of buying behaviour that cannot be simply observed. (Puneet Manchanda, Asim Ansari and Sunil Gupta, 1999). Analysing shopping baskets also shows multi-category dependencies across products which allows retailers to bundle new products that have not been discovered yet as a set.

2. Improve in-store settings and optimise product placement:

Gaining insight on product interdependencies can help retailers optimise store layout. It is an important aspect of retailing business because in-store settings may help increase sales if done right. It also influences buying behaviour, store traffic and the whole shopping atmosphere. If market basket analysis reveals that certain products are often purchased together, it is of great interest for the retailer to put these two items or categories of products close to each other to

facilitate the customer. Another option is to place them as far as possible from each other so that customers are exposed to much more products while trying to find the other product. However, the latter option may have negative consequences due to the fact the customers tend to get annoyed if they cannot find fast what they are looking for and need to waste time strolling around the whole store.

Optimisation of in-store settings may help improve shopping experience by reducing congestion and saving time for customers. With the right space planning the store benefits from increased cross product sales and impulse purchases. Moreover, store layout and atmosphere has a very strong impact on customer perceptions. A study made by Bill Merrilees and Dale Miller in 2001 showed that store layout and atmosphere has a positive effect on customer loyalty. In-store settings such as light, music, layout, appealing stock displays and easy to find goods are seen as determinants of pleasant and enjoyable shopping experience.

Various dimensions of store layout have positive effect on customers' purchase intentions and loyalty. This is why it is so crucial to extract knowledge from data so one can adjust store settings in order to improve customers' shopping experience.

3. Improve layout of the catalogue of e-commerce site:

Visual displays of products apply also to the catalogue of the firm online site. E-commerce website interface plays significant part of customers' perceptions. A key success factor for profitable e-commerce site is the layout. In order to be able to determine an optimised layout for website it is important to know the interdependencies among different products. A lot of research has been done in finding an optimal location, colouring and design for catalogues of e-commerce sites. The last step of successfully implementing a website strategy is to know how to place different products in order to maximise cross-sales. For instance, if we know which products have affinity to be sold together, we have to make sure that they are side by side on the same page on the website. It is also possible to provide discount in the form of shipping benefits for a group of products that have higher probabilities of selling together.

4. Control inventory based on product demand:

For the recent years, with more powerful analytical software it is possible to predict almost everything. It is now feasible to predict product demand based on data from past purchases, for example. For this objective it is important to know which products are related in terms of cross-sales.

Being able to find the probability of purchase for each product or a certain set of products is essential for controlling inventory. It has been observed that greater volume of products in the inventory can lead to higher levels of demand. (David R. Bell and Yasemin Boztu˘g, 2007). Many researchers have tried to give explanation for this phenomenon. Recent studies have found the impact of promotion on stockpilling and increased demand.

(Assunc¸ ˜ao, J. L., & Meyer, R. J., 1993) analyse the nature of the relationship which exists between price, promotion, sales and consumption. The authors' main finding is that price promotions encourage stockpilling, while on the other hand stockpilling rationally leads to increase in consumption.

However, the consumption time depends on the type of product that is associated with stockpiling. Foods and drinks are considered to be consumed faster than non-food goods. In this case, most of the beauty products cannot be stockpiled for long time due to extended consumption time. A face cream for example, can be used for 5-6 months before it is over. While a shampoo or toothpaste usually last not more than a month. Here comes the challenge of how many people are there in a single household. If the case is about a whole family stockpiling would be appropriate because families tend to shop more rare but in larger quantities. That is why it is harder to predict consumption time of products in a beauty store, but after examining which ones sell best, it will be very beneficial for the retailer so that he is always prepared with profit generating products available in stock.

III. Objectives For This Study-

The main point of interest for retailers is to understand dependencies among purchases. Consumers buy various combinations of products on a single shopping trip, but choice scenarios do not seem to be random to market analysts. These multicategory decisions result in the formation of consumers' "shopping baskets" which comprise the collection of categories that consumers purchase on a specific shopping trip.'

Overview of the Indian Personal Care Product Industry-

India's personal care industry is composed of hair care, bath products, skin care and cosmetics, and oral care. The sector is driven by rising income, rapid urbanization, and celebrity promotions. This industry accounts for 22% of the country's fast-moving consumer goods (FMCG), which is the term for Consumer Packaged Goods in India. Foreign direct investment in this sector totalled $691 million in 2014.

Hair care is a main category of this industry. A study by Nielsen, a market research firm, determined that shampoo is the most popular FMCG product in India. The $818 million shampoo segment is dominated by Hindustan Unilever Ltd., owned by U.K.-based Unilever. Its most popular brands are Sunsilk, Clear, and Clinic Plus. Hair oil is another important product, valued at $1.3 billion annually. India-based Marico's Parachute and Dabur are leaders in the production of branded coconut hair oil.

Estimated at $1 billion, the soap and bath category is significant. Soap is a prevalent product found in more than 90% of Indian households. The most common brands include Godrej's Cinthol, Reckitt Benckiser's Dettol, Wipro's Santoor, and Unilever's Lux, Dove, Hamam, and Lifebuoy. For men, shaving cream and razors are important personal care items. Procter & Gamble'sGillette is the most popular shaving cream and razor brand in India.

Within the cosmetics category, India's most prevalent products are skin creams, lotions, whitening creams, and makeup. Hindustan Unilever has three brands that are popular among Indian women—Fair & Lovely, Lakmé, and Ponds. Fair & Lovely was the world's first skin lightening cream and is the company's leading skin care brand. Colgate Palmolive's Charmis moisturizer is also prominent. The majority of the demand for cosmetics comes from working men and women. L'Oreal Paris develops both skin care and cosmetic products for India. New York-based Revlon expanded further to smaller cities in India, generating $40 million in revenues in 2014. The organic skin care category grows at over 20 percent annually and is expected to total $157 million in 2020, according to Azafran Innovacion, an organic skincare group. Large Indian organic skin care companies include Himalaya Herbals and Biotique. Both specialize in Ayurveda-based products. The oral care category is the smallest category; less than half of Indian consumers utilize western-style products such as toothpaste. Colgate Palmolive dominates more than half of this industry and was named India's most trusted brand four years in a row by a brand equity survey. Hindustan Unilever is another significant player with toothpaste brands Pepsodent and Close Up.

Market Size- The personal care market in India was estimated to be about Rs. 141 billion in 2006-07 (Rs 128 Billion in 2005-06) recording a value growth of about 10%. This includes two important daily hygiene product categories - personal wash (soap) and dental care (toothpaste and powders), which cannot really be classified as cosmetic products. Excluding these products, the size of the Indian cosmetics market is Rs. 64 billion (Rs. 57.5 billion in 2005).The major segments, by value, are skin care and shampoos, followed by men's grooming products (which includes also shaving accessories). A key market characteristic is the state of increasing competition and aggressive pricing. Major Segments in Personal Care Market-

Product Category	Rs. Billion	% Value
Shampoos	16.5	11.7%
Skin Care	21	14.9%
Hair Dyes & Colours	7	5.0%
Colour Cosmetics	3.5	2.5%
Oral Care	22.5	15.9%
Men's Grooming	12.75	9.0%
Personal Wash	55	38.9%
Deodorants & Perfumes	3.1	2.2%
Total	141.4	100%

Table: 1:India: Personal Care Market Segments, by Value Share-

(Source: Ace Global Private Limited)

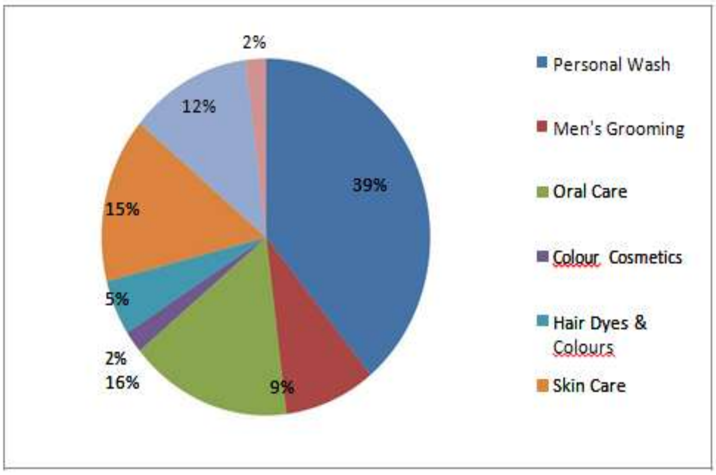


Fig:1: Personal Care Market Segments, by Value Share-(Source: *Ace Global Private Limited*)

How does Market Basket Analysis work?

To carry out an MBA, we need a data set of transactions. Each transaction represents a group of items or products that have been bought together and often referred to as an "item set". For example, one item set might be: {pencil, paper, staples, rubber} in which case all of these items have been bought in a single transaction. The transactions are analysed to identify rules of association. For example, one rule could be: {pencil, paper} => {rubber}. This means that if a customer has a transaction that contains a pencil and paper, then they are likely to be interested in also buying a rubber.

Before acting on a rule, a retailer needs to know whether there is sufficient evidence to suggest that it will result in a beneficial outcome. We therefore measure the strength of a rule by calculating the following three metrics:

Support: the percentage of transactions that contain all of the items in an item set (e.g., pencil, paper and rubber). The higher the support the more frequently the item set occurs. Rules with a high support are preferred since they are likely to be applicable to a large number of future transactions.

Confidence: the probability that a transaction that contains the items on the left hand side of the rule (in our example, pencil and paper) also contains the item on the right hand side (a rubber). The higher the confidence, the greater the likelihood that the item on the right hand side will be purchased or, in other words, the greater the return rate you can expect for a given rule.

Lift: the probability of all of the items in a rule occurring together (otherwise known as the support) divided by the product of the probabilities of the items on the left and right hand side occurring as if there was no association between them. For example, if pencil, paper and rubber occurred together in 2.5% of all transactions, pencil and paper in 10% of transactions and rubber in 8% of transactions, then the lift would be: 0.025/(0.1*0.08) = 3.125. A lift of more than 1 suggests that the presence of pencil and paper increases the probability that a rubber will also occur in the transaction. Overall, lift summarizes the strength of association between the products on the left and right hand side of the rule; the larger the lift the greater the link between the two products.

To perform a Market Basket Analysis and identify potential rules, a data mining algorithm called the *'Apriori algorithm*' is commonly used, which works in two steps:

- Systematically identify item sets that occur frequently in the data set with a support greater than a pre-specified threshold.
- Calculate the confidence of all possible rules given the frequent item sets and keep only those with a confidence greater than a pre-specified threshold.

BRAND RATINGS FOR MALE-

Fragrance & Body Care-

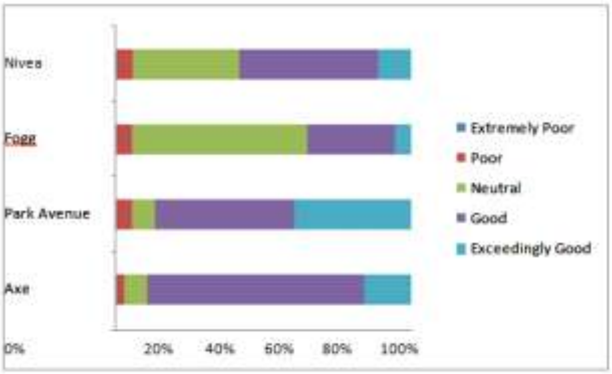


Fig:2

Skin Care-

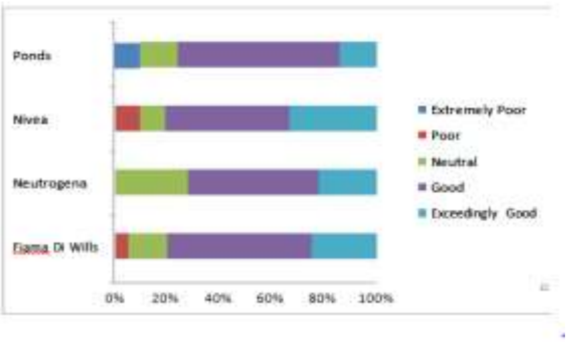


Fig:3

Face Care-

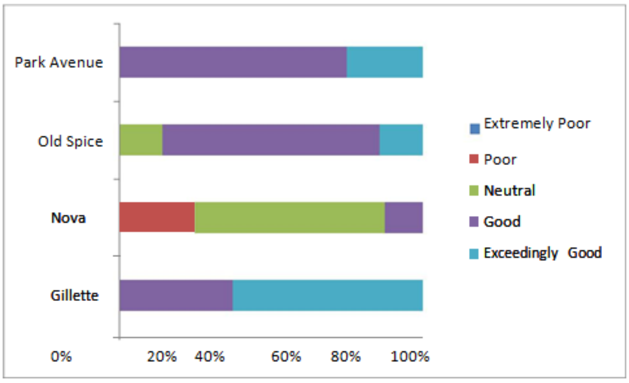


Fig:4

Oral Care-

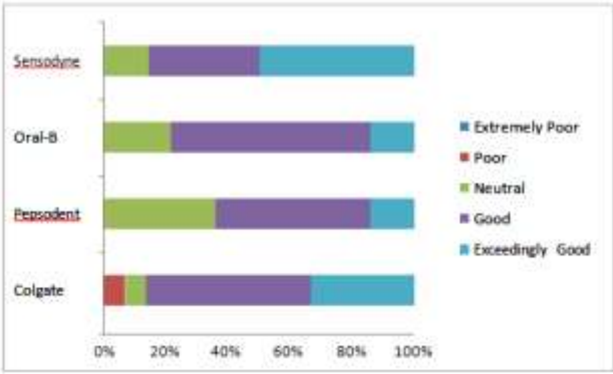


Fig:5

BRAND RATINGS FOR FEMALE-

Eye Care-

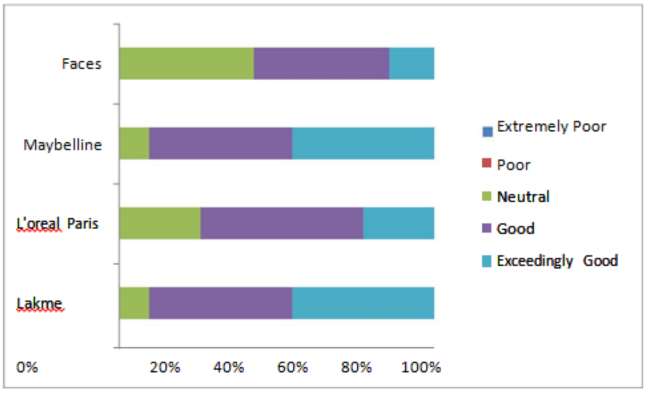


Fig:6

Lip Care-

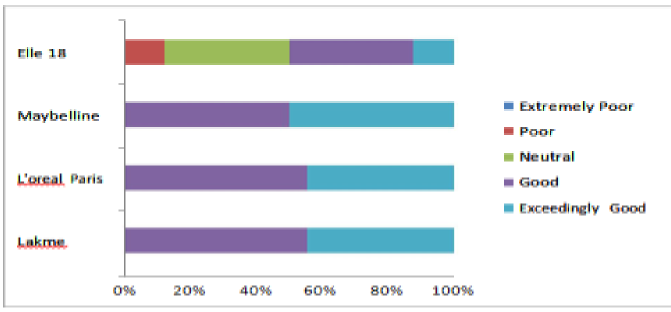


Fig:7

Face Care-

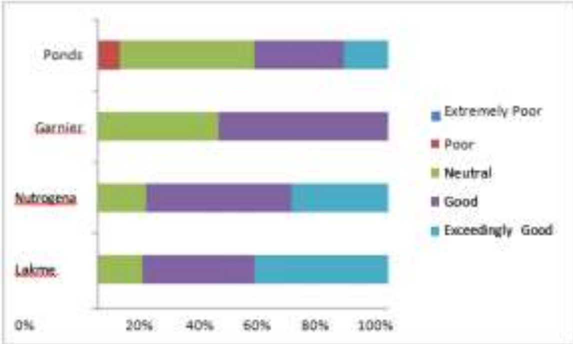


Fig:8

Hair Care-

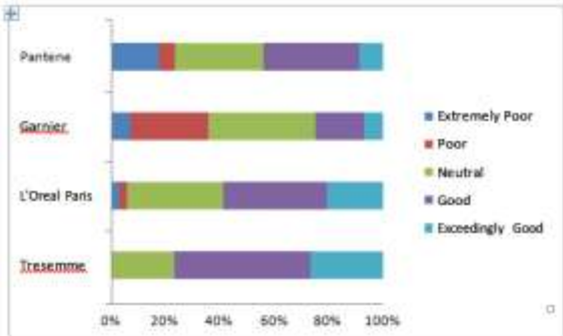


Fig:9

MARKET BASKET ANALYSIS FOR MALES-

Frequency of Occurrence of Items-

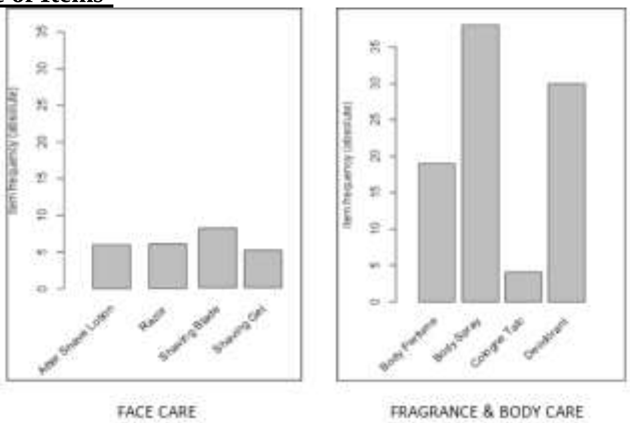


Fig:10

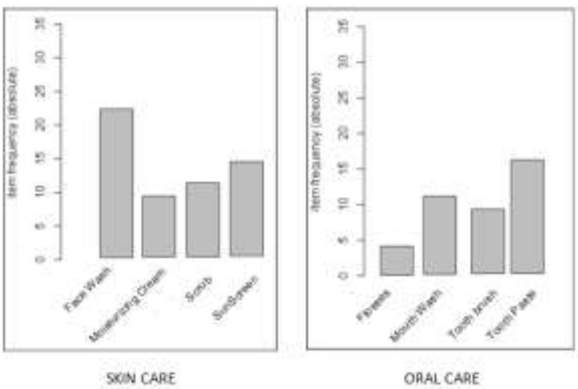
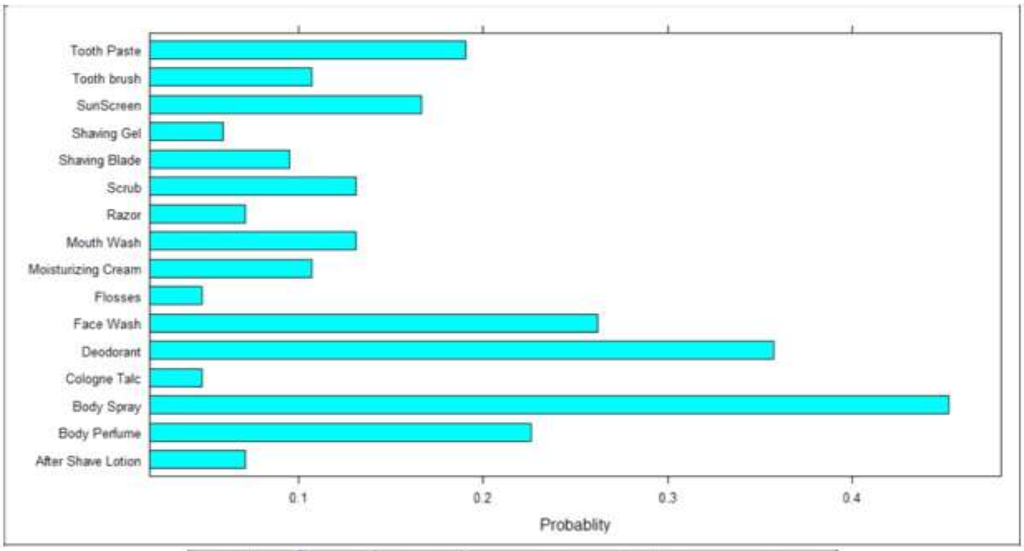


Fig:11

A bar plot of the support of all the items included-




```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.04762 0.07143 0.11900 0.15770 0.19940 0.45240
```

```
Apriori

Parameter specification:
 confidence minval smax arem  aval originalSupport support minlen maxlen target   ext
        0.5    0.1    1 none FALSE            TRUE    0.05      1     10  rules FALSE

Algorithmic control:
 filter tree heap memopt load sort verbose
    0.1 TRUE TRUE  FALSE TRUE    2    TRUE

Absolute minimum support count: 4

set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[16 item(s), 84 transaction(s)] done [0.00s].
sorting and recoding items ... [14 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 done [0.00s].
writing ... [33 rule(s)] done [0.00s].
creating S4 object  ... done [0.00s].
```

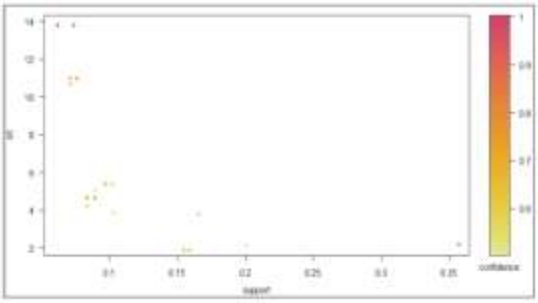
Conditional Probabilities of the occurrence of all items-

Tooth Paste (0.1904)	Flosses	0.0476
	Mouth Wash	0.1308
	Tooth Brush	0.1071
Body Spray (0.4523)	Body Perfume	0.0220
	Deodrant	0.1500
	Cologne Talc	0.0400

Shaving Blade (0.0954)	Shaving Gel	0.0595
	After Shave Lotion	0.0715
	Razor	0.0715
Face Wash (0.2619)	Scrub	0.1300
	Sunscreen	0.1600
	Moisturizing Cream	0.0100

Table:2

A scatter plot of the confidence, support and lift metrics-



Graph-based visualization of the rules in terms of lift-

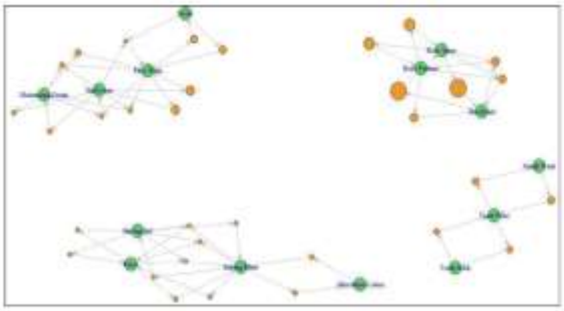


Fig: 13

MARKET BASKET ANALYSIS FOR FEMALES-

Frequency Of Occurrence Of Items-

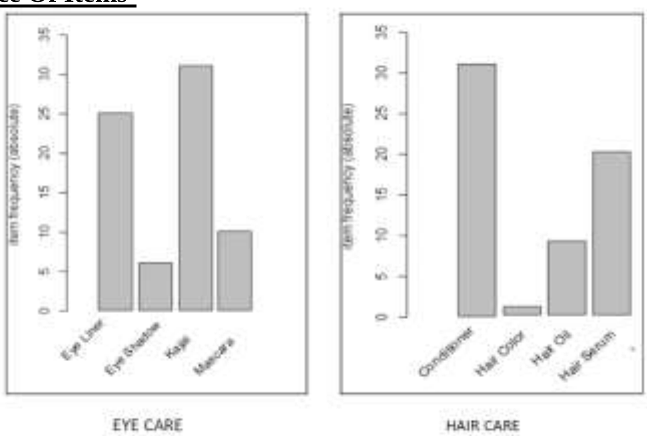
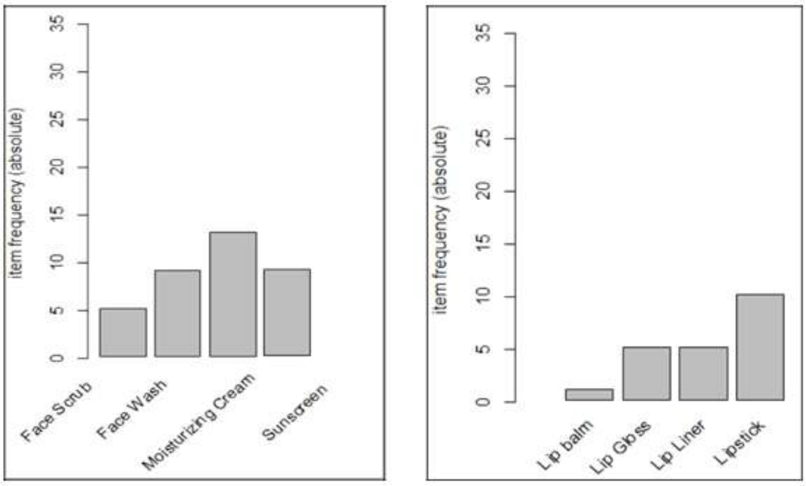


Fig:14



FACE CARE LIP CARE

Fig:15

Conditional Probabilities of the occurrence of all items-

Shampoo (0.3932)	Hair Colour	0.0112
	Hair Oil	0.1011
	Hair Serum	0.2247
Moisturizing Cream (0.1460)	Face Wash	0.1011
	Face Scrub	0.0562
	Sunscreen	0.1011
Lipstick (0.1123)	Lip Liner	0.0562
	Lip Gloss	0.0562
	Lip Balm	0.0112
Kajal (0.3485)	Mascara	0.1123
	Eye Liner	0.1008
	Eye Shadow	0.0674

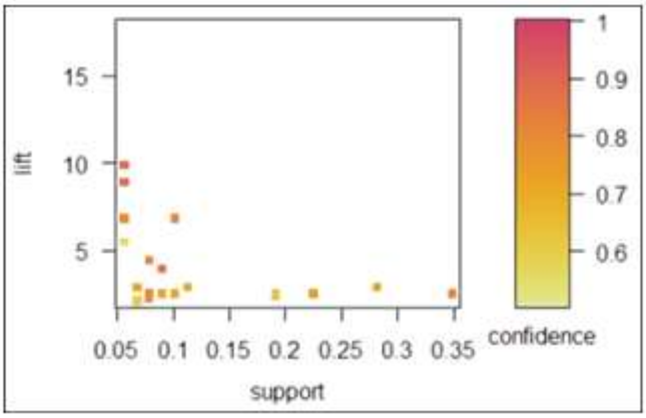
Table: 3

Market Basket Analysis-

19	{Hair Oil}	=>	{Hair Serum}	0.08988764	0.8888889	3.955556
47	{Hair Oil,Shampoo}	=>	{Hair Serum}	0.08988764	0.8888889	3.955556
29	{Shampoo}	=>	{Conditioner}	0.34831461	0.8857143	2.542857
44	{Hair Oil,Hair Serum}	=>	{Conditioner}	0.07865169	0.8750000	2.512097
58	{Hair Oil,Hair Serum,Shampoo}	=>	{Conditioner}	0.07865169	0.8750000	2.512097
24	{Hair Serum}	=>	{Conditioner}	0.19101124	0.8500000	2.440323
51	{Hair Serum,Shampoo}	=>	{Conditioner}	0.19101124	0.8500000	2.440323
23	{Kajal}	=>	{Eye Liner}	0.28089888	0.8064516	2.870968
20	{Hair Oil}	=>	{Conditioner}	0.07865169	0.7777778	2.232975
49	{Hair Oil,Shampoo}	=>	{Conditioner}	0.07865169	0.7777778	2.232975
14	{Moisturizing Cream}	=>	{Sunscreen}	0.10112360	0.6923077	6.846154
16	{Moisturizing Cream}	=>	{Face Wash}	0.10112360	0.6923077	6.846154
17	{Mascara}	=>	{Eye Liner}	0.06741573	0.6000000	2.136000
43	{Kajal,Mascara}	=>	{Eye Liner}	0.06741573	0.6000000	2.136000
27	{Shampoo}	=>	{Hair Serum}	0.22471910	0.5714286	2.542857
7	{Sunscreen}	=>	{Face Scrub}	0.05617978	0.5555556	9.888889
9	{Face Wash}	=>	{Face Scrub}	0.05617978	0.5555556	9.888889
11	{Sunscreen}	=>	{Face Wash}	0.05617978	0.5555556	5.493827
12	{Face Wash}	=>	{Sunscreen}	0.05617978	0.5555556	5.493827
35	{Moisturizing Cream,Sunscreen}	=>	{Face Scrub}	0.05617978	0.5555556	9.888889
38	{Face Wash,Moisturizing Cream}	=>	{Face Scrub}	0.05617978	0.5555556	9.888889
40	{Moisturizing Cream,Sunscreen}	=>	{Face Wash}	0.05617978	0.5555556	5.493827
41	{Face Wash,Moisturizing Cream}	=>	{Sunscreen}	0.05617978	0.5555556	5.493827
25	{Conditioner}	=>	{Hair Serum}	0.19101124	0.5483871	2.440323
52	{Conditioner,Shampoo}	=>	{Hair Serum}	0.19101124	0.5483871	2.440323
2	{Lipstick}	=>	{Lip Liner}	0.05617978	0.5000000	8.900000
4	{Lipstick}	=>	{Lip Gloss}	0.05617978	0.5000000	8.900000

Table: 4

A scatter plot of the confidence, support and lift metrics-



INFERENCE

Male Personal Care Products-

ORAL CARE-

{Tooth Paste} => {Mouth Wash} has a confidence value 68.75% indicating if someone buys tooth paste, they are 68.75% likely to buy mouth wash too.
The support value of 0.1309 indicates that 13.09% of the transactions in the data involve tooth paste purchases.
{Tooth Paste} => {Tooth Brush} has a confidence value 56.25% indicating if someone buys tooth paste, they are 56.25% likely to buy tooth brush too.
The support value of 0.1071 indicates that 10.71% of the transactions in the data involve tooth paste purchases.
Hence the support indicates goodness of the choice of rule and confidence indicates the correctness of the rule.
Lift is greater than 1 for both the transactions indicating that the presence of tooth paste has positively increased the probability of mouth wash (& tooth brush, in the latter case) occurring on this transaction.

FRAGRANCE & BODY CARE-

{Body Spray} => {Deodrant} has a confidence value 78.94% indicating if someone buys Body Spray, they are 78.94% likely to buy Deodrant too.
The support value of 0.3571 indicates that 35.71% of the transactions in the data involve Body Spray purchases.
{Body Perfume} => {Deodrant} has a confidence value 68.42% indicating if someone buys Body Perfume, they are 68.42% likely to buy Deodrant too.
The support value of 0.1547 indicates that 15.47% of the transactions in the data involve Body Perfume purchases.
{Body Perfume, Body Spray} => {Deodrant} has a confidence value 68.42% indicating if someone buys Body Perfume & Body Spray, they are 68.42% likely to buy Deodrant too.
The support value of 0.1547 indicates that 15.47% of the transactions in the data involve both Body Perfume & Body Spray purchases.
{Body Spray} => {Body Perfume} has a confidence value 50% indicating if someone buys Body Spray, they are 50% likely to buy Body Perfume too.
The support value of 0.2261 indicates that 22.61% of the transactions in the data involve Body Spray purchases.
Hence the support indicates goodness of the choice of rule and confidence indicates the correctness of the rule.
Lift is greater than 1 for all the transactions indicating that the presence of the product on the L.H.S has positively increased the probability of the occurrence the product on the R.H.S, on all transactions.

FACE CARE-

{Razor} => {Shaving Gel} has a confidence value 83.33% indicating if someone buys Razor, they are 83.33% likely to buy Shaving Gel too.
The support value of 0.0595 indicates that 5.95% of the transactions in the data involve Razor purchases.
{Razor, Shaving Blade} => {Shaving Gel} has a confidence value 83.33% indicating if someone buys Razor & Shaving Blade, they are 83.33% likely to buy Shaving Gel too.
The support value of 0.0595 indicates that 5.95% of the transactions in the data involve both Razor & Shaving Blade purchases.
{Shaving Blade} => {After Shave Lotion} has a confidence value 75% indicating if someone buys Shaving Blade, they are 75% likely to buy After Shave Lotion too.
The support value of 0.0714 indicates that 7.14% of the transactions in the data involve Shaving Blade purchases.

{Shaving Blade} => {Razor} has a confidence value 75% indicating if someone buys Shaving Blade, they are 75% likely to buy Razor too.
The support value of 0.0714 indicates that 7.14% of the transactions in the data involve Shaving Blade purchases.
{Shaving Blade} => {Shaving Gel} has a confidence value 62.5% indicating if someone buys Shaving Blade, they are 62.5% likely to buy Shaving Gel too.
The support value of 0.0595 indicates that 5.95% of the transactions in the data involve Shaving Blade purchases.
Hence the support indicates goodness of the choice of rule and confidence indicates the correctness of the rule.Lift is greater than 1 for all the transactions indicating that the presence of the product on the L.H.S has positively increased the probability of the occurrence the product on the R.H.S, on all transactions.

SKIN CARE-

{Moisturizing Cream} => {Sunscreen} has a confidence value 77.78% indicating if someone buys Moisturizing Cream, they are 77.78% likely to buy Sunscreen too.
The support value of 0.0833 indicates that 8.33% of the transactions in the data involve Moisturizing Cream purchases.
{Face Wash, Moisturizing Cream} => {Sunscreen} has a confidence value 77.78% indicating if someone buys Face Wash & Moisturizing Cream, they are 77.78% likely to buy Sunscreen too.
The support value of 0.0833 indicates that 8.33% of the transactions in the data involve both Face Wash & Moisturizing Cream purchases.
{Face Wash} => {Sunscreen} has a confidence value 63.63% indicating if someone buys Face Wash, they are 75% likely to buy Sunscreen too.
The support value of 0.1666 indicates that 16.67% of the transactions in the data involve Face Wash purchases.
{Sunscreen} => {Moisturizing Cream} has a confidence value 50% indicating if someone buys Sunscreen, they are 50% likely to buy Moisturizing Cream too.
The support value of 0.0833 indicates that 8.33% of the transactions in the data involve Sunscreen purchases.
{Face Wash} => {Scrub} has a confidence value 50% indicating if someone buys Face Wash, they are 50% likely to buy Scrub too.
The support value of 0.1309 indicates that 13.09% of the transactions in the data involve Face Wash purchases.
{Face Wash, Sunscreen} => {Moisturizing Cream} has a confidence value 50% indicating if someone buys Face Wash & Sunscreen, they are 50% likely to buy Moisturizing Cream too.
The support value of 0.0833 indicates that 8.33% of the transactions in the data involve Face Wash purchases.
Hence the support indicates goodness of the choice of rule and confidence indicates the correctness of the rule.
Lift is greater than 1 for all the transactions indicating that the presence of the product on the L.H.S has positively increased the probability of the occurrence the product on the R.H.S, on all transactions.

<u>Female Personal Care Products</u>-
EYE CARE-

{Kajal} => {Eye Liner} has a confidence value 80.64% indicating if someone buys Kajal, they are 80.64% likely to buy Eye Liner too.
The support value of 0.2808 indicates that 28.08% of the transactions in the data involve Kajal purchases.
{Mascara} => {Eye Liner} has a confidence value 60% indicating if someone buys Mascara, they are 60% likely to buy Eye Liner too.
The support value of 0.0674 indicates that 6.74% of the transactions in the data involve Mascara purchases.
{Kajal, Mascara} => {Eye Liner} has a confidence value 60% indicating if someone buys Kajal & Mascara, they are 60% likely to buy Eye Liner too.
The support value of 0.0674 indicates that 6.74% of the transactions in the data involve both Kajal & Mascara purchases.
Hence the support indicates goodness of the choice of rule and confidence indicates the correctness of the rule.
Lift is greater than 1 for all the transactions indicating that the presence of the product on the L.H.S has positively increased the probability of the occurrence the product on the R.H.S, on all transactions.

LIP CARE-

{Lipstick} => {Lip Liner} has a confidence value 50% indicating if someone buys Lipstick, they are 50% likely to buy Lip Liner too.
The support value of 0.0561 indicates that 5.61% of the transactions in the data involve Lipstick purchases.

{Lipstick} => {Lip Gloss} has a confidence value 50% indicating if someone buys Lipstick, they are 50% likely to buy Lip Gloss too.
The support value of 0.0561 indicates that 5.61% of the transactions in the data involve Lipstick purchases.
Hence the support indicates goodness of the choice of rule and confidence indicates the correctness of the rule.
Lift is greater than 1 for all the transactions indicating that the presence of the product on the L.H.S has positively increased the probability of the occurrence the product on the R.H.S, on all transactions.

HAIR CARE-

{Hair Oil} => {Hair Serum} has a confidence value 88.89% indicating if someone buys Hair Oil, they are 88.89% likely to buy Hair Serum too.
The support value of 0.0898 indicates that 8.98% of the transactions in the data involve Hair Oil purchases.
{Hair Oil, Shampoo} => {Hair Serum} has a confidence value 88.89% indicating if someone buys Hair Oil & Shampoo, they are 88.89% likely to buy Hair Serum too.
The support value of 0.0898 indicates that 8.98% of the transactions in the data involve both Hair Oil & Shampoo purchases.
{Shampoo} => {Conditioner} has a confidence value 88.57% indicating if someone buys Shampoo, they are 88.57% likely to buy Conditioner too.
The support value of 0.3483 indicates that 34.83% of the transactions in the data involve Shampoo purchases.
{Hair Oil, Hair Serum} => {Conditioner} has a confidence value 87.5% indicating if someone buys Hair Oil & Hair Serum, they are 87.5% likely to buy Conditioner too.
The support value of 0.0786 indicates that 7.86% of the transactions in the data involve Hair Oil & Hair Serum purchases.
{Hair Oil, Hair Serum, Shampoo} => {Conditioner} has a confidence value 87.5% indicating if someone buys Hair Oil, Hair Serum & Shampoo, they are 87.5% likely to buy Conditioner as well.
The support value of 0.0786 indicates that 7.86% of the transactions in the data involve Hair Oil, Hair Serum & Shampoo purchases.
{Hair Serum} => {Conditioner} has a confidence value 85% indicating if someone buys Hair Serum, they are 85% likely to buy Conditioner too.
The support value of 0.1910 indicates that 19.10% of the transactions in the data involve Hair Serum purchases.
{Hair Serum, Shampoo} => {Conditioner} has a confidence value 85% indicating if someone buys Hair Serum & Shampoo, they are 85% likely to buy Conditioner too.
The support value of 0.1910 indicates that 19.10% of the transactions in the data involve both Hair Serum & Shampoo purchases.
{Hair Oil} => {Conditioner} has a confidence value 77.78% indicating if someone buys Hair Oil, they are 77.78% likely to buy Conditioner too.
The support value of 0.0786 indicates that 7.86% of the transactions in the data involve Hair Oil purchases.
{Hair Oil, Shampoo} => {Conditioner} has a confidence value 77.78% indicating if someone buys Hair Oil & Shampoo, they are 77.78% likely to buy Conditioner too.
The support value of 0.0786 indicates that 7.86% of the transactions in the data involve both Hair Oil & Shampoo purchases.
{Shampoo} => {Hair Serum} has a confidence value 57.14% indicating if someone buys Shampoo, they are 57.14% likely to buy Hair Serum too.
The support value of 0.2247 indicates that 22.47% of the transactions in the data involve Shampoo purchases.
{Conditioner} => {Hair Serum} has a confidence value 54.83% indicating if someone buys Conditioner, they are 54.83% likely to buy Hair Serum too.
The support value of 0.1910 indicates that 19.10% of the transactions in the data involve Conditioner purchases.
{Conditioner, Shampoo} => {Hair Serum} has a confidence value 54.83% indicating if someone buys Conditioner & Shampoo, they are 54.83% likely to buy Hair Serum too.
The support value of 0.1910 indicates that 19.10% of the transactions in the data involve both Conditioner & Shampoo purchases.
Hence the support indicates goodness of the choice of rule and confidence indicates the correctness of the rule.
Lift is greater than 1 for all the transactions indicating that the presence of the product on the L.H.S has positively increased the probability of the occurrence the product on the R.H.S, on all transactions.

FACE CARE-

{Moisturizing Cream} => {Sunscreen} has a confidence value 69.23% indicating if someone buys Moisturizing Cream, they are 69.23% likely to buy Sunscreen too.

The support value of 0.1011 indicates that 10.11% of the transactions in the data involve Moisturizing Cream purchases.
{Moisturizing Cream} => {Face Wash} has a confidence value 69.23% indicating if someone buys Moisturizing Cream, they are 69.23% likely to buy Face Wash too. The support value of 0.1011 indicates that 10.11% of the transactions in the data involve Moisturizing Cream purchases.
{Sunscreen} => {Face Scrub} has a confidence value 55.56% indicating if someone buys Sunscreen, they are 55.56% likely to buy Face Scrub too.
The support value of 0.0561 indicates that 5.61% of the transactions in the data involve Sunscreen purchases.
{Face Wash} => {Face Scrub} has a confidence value 55.56% indicating if someone buys Face Wash, they are 55.56% likely to buy Face Scrub too.
The support value of 0.0561 indicates that 5.61% of the transactions in the data involve Face Wash purchases.
{Sunscreen} => {Face Wash} has a confidence value 55.56% indicating if someone buys Sunscreen, they are 55.56% likely to buy Face Wash too.
The support value of 0.0561 indicates that 5.61% of the transactions in the data involve Sunscreen purchases.
{Face Wash} => {Sunscreen} has a confidence value 55.56% indicating if someone buys Face Wash, they are 55.56% likely to buy Sunscreen too.
The support value of 0.0561 indicates that 5.61% of the transactions in the data involve Face Wash purchases.
{Moisturizing Cream, Sunscreen} => {Face Scrub} has a confidence value 55.56% indicating if someone buys Moisturizing Cream & Sunscreen, they are 55.56% likely to buy Face Scrub too.
The support value of 0.0561 indicates that 5.61% of the transactions in the data involve both Moisturizing Cream & Sunscreen purchases.
{Face Wash, Moisturizing Cream} => {Face Scrub} has a confidence value 55.56% indicating if someone buys Face Wash & Moisturizing Cream, they are 55.56% likely to buy Face Scrub too.
The support value of 0.0561 indicates that 5.61% of the transactions in the data involve both Face Wash & Moisturizing Cream purchases.
{Moisturizing Cream, Sunscreen} => {Face Wash} has a confidence value 55.56% indicating if someone buys Moisturizing Cream & Sunscreen, they are 55.56% likely to buy Face Wash too.
The support value of 0.0561 indicates that 5.61% of the transactions in the data involve both Moisturizing Cream & Sunscreen purchases.
{Face Wash, Moisturizing Cream} => {Sunscreen} has a confidence value 55.56% indicating if someone buys Face Wash & Moisturizing Cream, they are 55.56% likely to buy Sunscreen too.
The support value of 0.0561 indicates that 5.61% of the transactions in the data involve both Face Wash & Moisturizing Cream purchases.
Hence the support indicates goodness of the choice of rule and confidence indicates the correctness of the rule.
Lift is greater than 1 for all the transactions indicating that the presence of the product on the L.H.S has positively increased the probability of the occurrence the product on the R.H.S, on all transactions.

IV. Conclusion

MALES-

ORAL CARE-

{Tooth Paste} => {Mouth Wash} has the maximum confidence value of 68.75% indicating these two products should be placed close to each other to facilitate purchase by customers.
Majority of the respondents voted Sensodyne to be the best brand, closely followed by Colgate.
Oral B & Pepsodent are also considered to be good brands.

FRAGRANCE & BODY CARE-

{Body Spray} => {Deodrant} has the maximum confidence value of 78.94% indicating people purchasing body sprays will have a high propensity for purchasing deodorants as well. Thus they need to be conveniently placed to catch attention of buyers.
Park Avenue is perceived to be the best quality brand among majority of the respondents, followed by Axe.

FACE CARE-

{Razor} => {Shaving Gel} & {Razor, Shaving Blade} => {Shaving Gel} both have a confidence value of 83.33%. Thus in either case, when people will buy razor (with or without shaving blade), they are likely to buy shaving gel along with that. Thus, placing all three of them together will significantly increase sales.
Gillette is the perceived to be the best quality brand. 25% of the respondents rated Nova to be a poor quality brand.

SKIN CARE-

{Moisturizing Cream} => {Sunscreen} & {Face Wash, Moisturizing Cream} => {Sunscreen} has a maximum confidence value of 77.78%. Thus people buying moisturizing cream alone or with face wash will be inclined to buy sunscreen as well. Keeping them in close vicinity will facilitate purchase.

All the brands- Ponds, Nivea, Neutrogena & Fiama Di Wills are almost at par in terms of goodness of quality. However 10% of the respondents have rated Ponds to be "extremely poor" & Nivea to be "poor".

FEMALES-

EYE CARE-

{Kajal} => {Eye Liner} has the maximum confidence value of 80.64%. Female buyers looking for Kajal will be likely to want for an eye liner as well along with that. Suggesting the same will positively influence their purchase decision.

Maybelline & Lakme are perceived to be exceedingly good by majority of the respondents.

LIP CARE-

{Lipstick} => {Lip Liner} has the maximum confidence value of 50%. Thus, lipstick & lip liner are suitable products which are likely to be bought together. Retailers must use this information to their advantage.

Lakme, L'oreal Paris & Maybelline are perceived to be the exceedingly good brands whereas Elle 18 is more of an average brand.

HAIR CARE-

This category shows high confidence values which indicate that hair care products when placed suitably together are likely to be sold together & increase sales.

{Hair Oil} => {Hair Serum} & {Hair Oil, Shampoo} => {Hair Serum} has the maximum confidence value of 88.89%, closely followed by {Shampoo} => {Conditioner} having a confidence value 88.57%.

Coming next, both {Hair Oil, Hair Serum} => {Conditioner} & {Hair Oil, Hair Serum, Shampoo} => {Conditioner} have a high confidence value of 87.5%.

Lastly {Hair Serum} => {Conditioner} & {Hair Serum, Shampoo} => {Conditioner} have confidence value 85%, which is also quite high.

Thus, when products like hair oil, shampoo, conditioner, hair serum are placed in close proximity, they will highly facilitate purchase.

Tresemme is perceived to be the best quality brand for haircare closely followed by L'oreal Paris. Garnier is rated as a poor quality brand by majority.

FACE CARE-

{Moisturizing Cream} => {Sunscreen} & {Moisturizing Cream} => {Face Wash} have the maximum confidence value of 69.23% indicating that a person buying Moisturizing cream is likely to buy Sunscreen or Face Wash along with it. Thus shelf space should be utilized keeping this in mind.

Lakme emerges as the best quality brand for Face care followed by Neutrogena.

FINAL CONCLUSION-

MALES-

The top 5 transactions with the highest confidence values are-

{Razor} => {Shaving Gel} & {Razor, Shaving Blade} => {Shaving Gel} both having the maximum confidence value of 83.33%.

{Body Spray} => {Deodrant} having a high confidence value of 78.94%.

{Moisturizing Cream} => {Sunscreen} & {Face Wash, Moisturizing Cream} => {Sunscreen} having a high confidence value of 77.78%.

FEMALES-The top 5 transactions with the highest confidence values are-

{Hair Oil} => {Hair Serum} & {Hair Oil, Shampoo} => {Hair Serum} having the maximum confidence value of 88.89%.

{Shampoo} => {Conditioner} having a high confidence value of 88.57%.

{Hair Oil, Hair Serum} => {Conditioner} & {Hair Oil, Hair Serum, Shampoo} => {Conditioner} having a high confidence value of 87.5%.

Thus, if the store space is utilized in a way keeping in mind the best possible combination of the products which are frequently bought by the customers, it will result in maximum returns & hence, a profitable business.

References

[1]. Agrawal, R., Imieliński, T., & Swami, A. 1993. Mining association rules between sets of items in large databases.

[2]. In P. Buneman, & S. Jajodia (Eds.), *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*: 207-216. New York: Association for Computing Machinery.

[3]. Aguinis, H., Beaty, J. C., Boik, R. J., & Pierce, C. A. 2005. Effect size and power in assessing moderating effects of categorical variables using multiple regression: A 30-year review. *Journal of Applied Psychology*, 90: 94-107.

[4]. Aguinis, H., Boyd, B. K., Pierce, C. A., & Short, J. C. 2011. Walking new avenues in management research methods and theories: Bridging micro and macro domains. *Journal of Management*, 37: 395-403.

[5]. Aguinis, H., Dalton, D. R., Bosco, F. A., Pierce, C. A., & Dalton, C. M. 2011. Meta-analytic choices and judgment calls: Implications for theory building and testing, obtained effect sizes, and scholarly impact. *Journal of Management*, 37: 5-38.

[6]. Alvarez, S. A. 2003. *Chi-squared computation for association rules: Preliminary results*. Technical report BC-CS- 2003-01. Chestnut Hill, MA: Boston College.

[7]. Aumann, Y., & Lindell, Y. 2003. A statistical theory for quantitative association rules. *Journal of Intelligent Information Systems*, 20: 255-283.
[8]. Berry, M. J. A., & Linoff, G. S. 2004. *Data mining techniques for marketing, sales, and customer relationship management* (2nd ed.). Indianapolis, IN: Wiley.
[9]. Bowen, H. P. 2012. Testing moderating hypotheses in limited dependent variable and other nonlinear models: Secondary versus total interactions. *Journal of Management*, 38: 860-889.
[10]. Boyd, B. K., Gove, S., & Hitt, M. A. 2005. Construct measurement in strategic management research: Illusion or reality? *Strategic Management Journal*, 26: 239-257.