

Presented by:



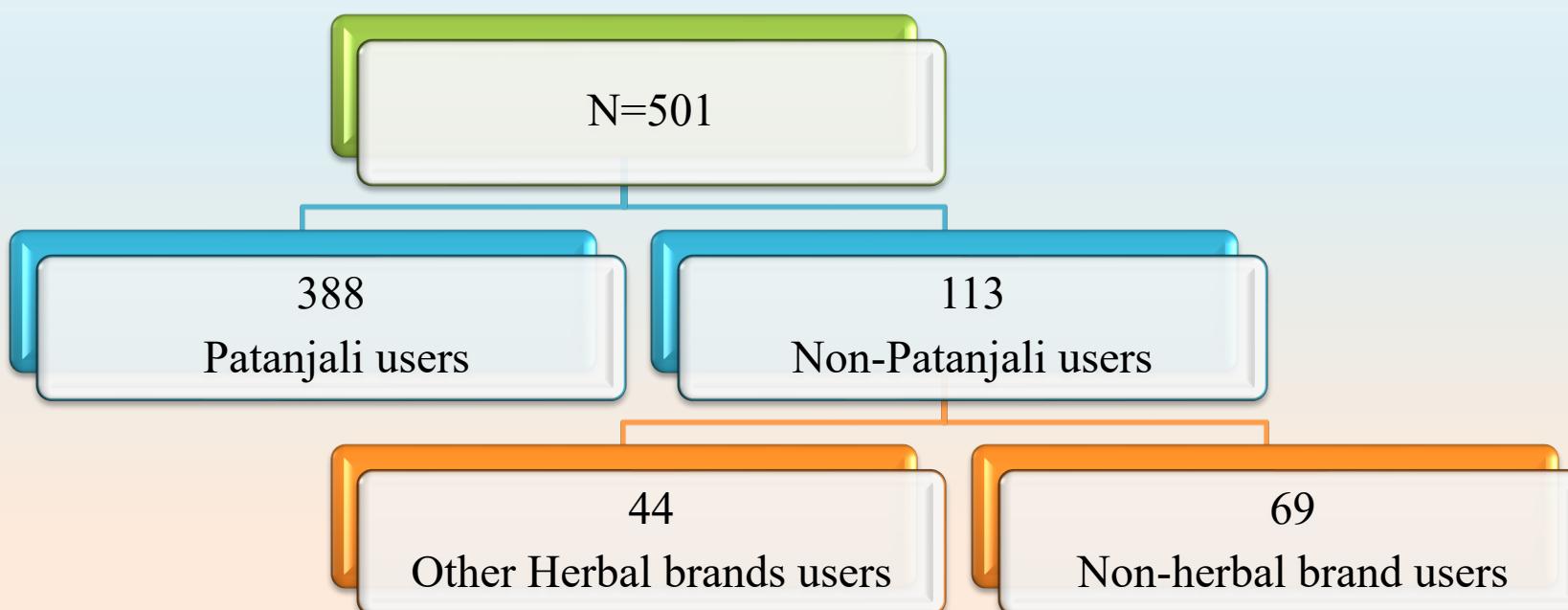
HETVI GALA	2019205
MANSI MAKWANA	2019208
VAIBHAVI OZA	2019210
AKSHAY BAROT	2019218

INTRODUCTION

- Patanjali Ayurved Kendra Ltd was founded by Swami Ramdev and Acharya Balakrishna on September 27th, 2007. It started its operation in Katmandu, Nepal to provide holistic, natural and effective ayurvedic treatment.
- “Patanjali“ is represented by Guru Swami and the company is set up with a motive to provide high quality Indian products at fair price which are organic and natural.

RESEARCH METHODOLOGY & SAMPLING TECHNIQUE

- For collecting Primary data, a survey approach was used only for Mumbai city.
- A Pilot survey of sample size 50 personally & made the necessary changes.
- The data was also encoded according to the techniques used.
- Convenient sampling was used as the sampling method.



SWOT ANALYSIS

OBJECTIVE 1

To study the internal and external environments of a company, through the identification and analysis of the strengths and weaknesses of the organization, and the opportunities and threats to which it is exposed.

- SWOT Analysis is a simple but useful framework for analyzing your organization's strengths, weaknesses, opportunities, and threats.



Strength

Baba Ramdev

Strong Patriotism

Comparatively Low Prices

Ayurveda & Herbal

Weakness

Over dependency on Ramdev

Lack of experienced management graduates

Low number of manufacturing units

Product Dependence

Penetration pricing is not long term

Opportunities

Growing Organic sector

Rural Expansion

Going global

Tie-ups

Diversification

Threats

Price war

Increasing Competition

Poor reap

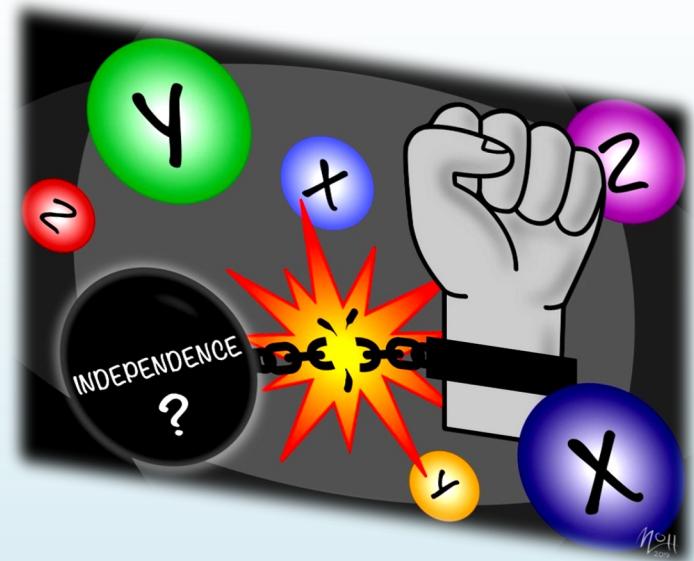
Negative word of mouth

CHI SQUARE TEST OF INDEPENDENCE

OBJECTIVE 2

To test the association between gender and their health-conscious behaviour towards the Patanjali products.

- Chi Square test of independence measures whether there is a relationship between two categorical variables.
- The Chi Square statistic is a non-parametric tool designed to analyze group differences when the dependent variable is measured at nominal level.



SUB-OBJECTIVE FOR FURTHER ANALYSIS

1. To test the association between gender and their consideration over doctor's recommendation while buying the Patanjali products.
2. To test the association between gender and their experience over any side effects while using the Patanjali products.

HYPOTHESIS

HO1: There is no significant association between the gender and their consideration over doctor's recommendation.

H11: Not HO1

HO2: There is no significant association gender and their experience over any side effects.

H12: Not HO2

CASE 1

Gender * Doctor's recommendation Cross tabulation

		Doctor's recommendation		Total
		Yes	No	
Gender	MALE	Count	90	99
		Expected Count	97.5	91.5
		% within Gender	47.6%	52.4%
	FEMALE	Count	105	84
		Expected Count	97.5	91.5
		% within Gender	55.6%	44.4%
Total		Count	195	183
		Expected Count	195.0	183.0
		% within Gender	51.6%	48.4%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2.383 ^a	1	.123		
Continuity Correction ^b	2.076	1	.150		
Likelihood Ratio	2.386	1	.122		
Fisher's Exact Test				.150	.075
Linear-by-Linear Association	2.377	1	.123		
N of Valid Cases	378				

p-value(0.123)>0.05
Thus, we fail to reject H₀₁.
Hence, we conclude that there is no association between gender & doctor's recommendation while buying the products, i.e., they are independent of each other.

Gender * side effects Cross tabulation

		Side effects		Total
		Yes	No	
Gender	MALE	Count	41	148
		Expected Count	42.5	146.5
		% within Gender	21.7%	78.3%
	FEMALE	Count	44	145
		Expected Count	42.5	146.5
		% within Gender	23.3%	76.7%
Total		Count	85	293
		Expected Count	85.0	293.0
		% within Gender	22.5%	77.5%

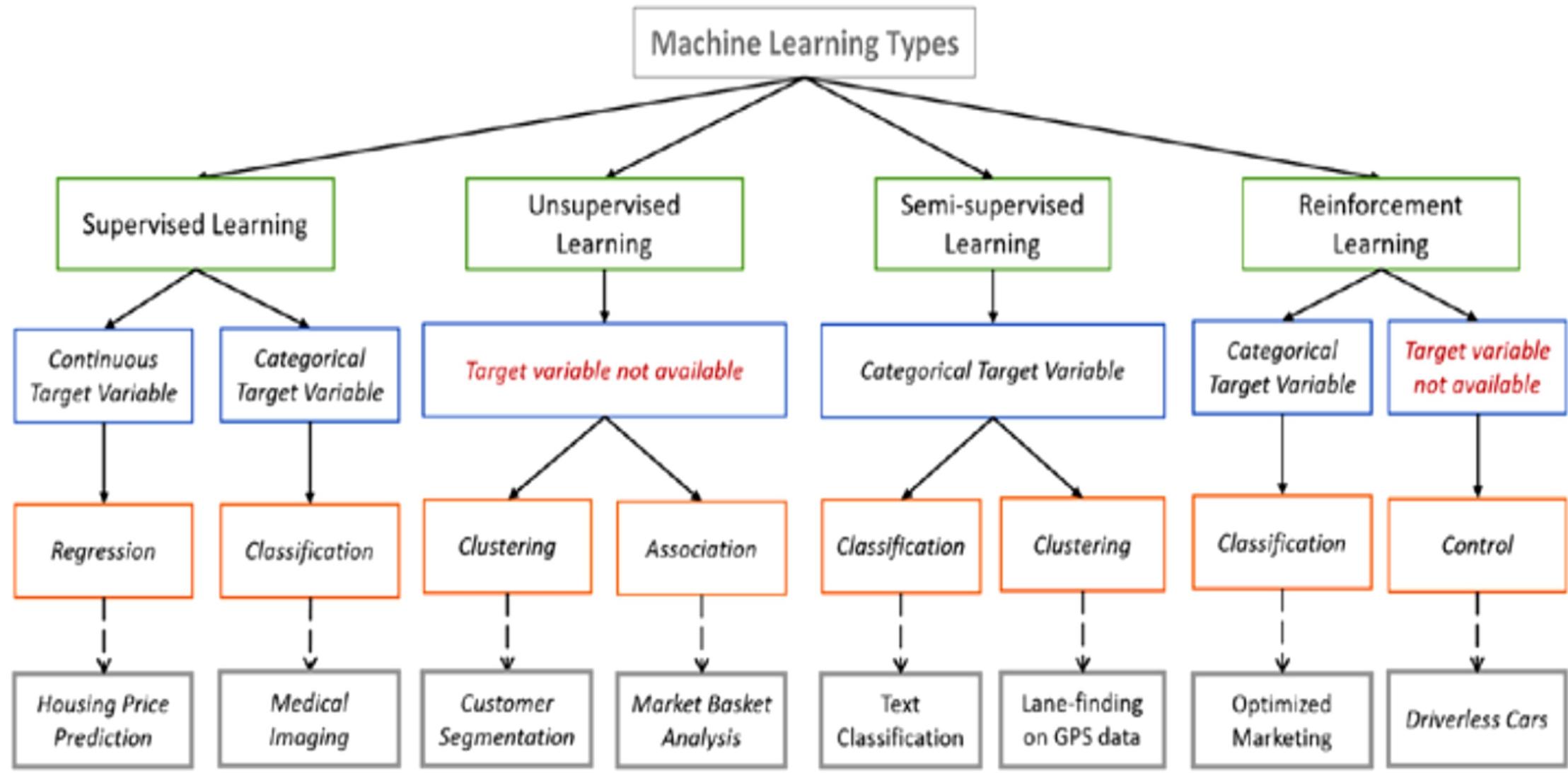
Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.137 ^a	1	.712		
Continuity Correction ^b	.061	1	.805		
Likelihood Ratio	.137	1	.712		
Fisher's Exact Test				.805	.403
Linear-by-Linear Association	.136	1	.712		
N of Valid Cases	378				

p-value(0.712)>0.05

Thus, we fail to reject H₀₂.

Hence, we conclude that there is no association between gender & their experience over any side effects while using the products, i.e., they are independent of each other.



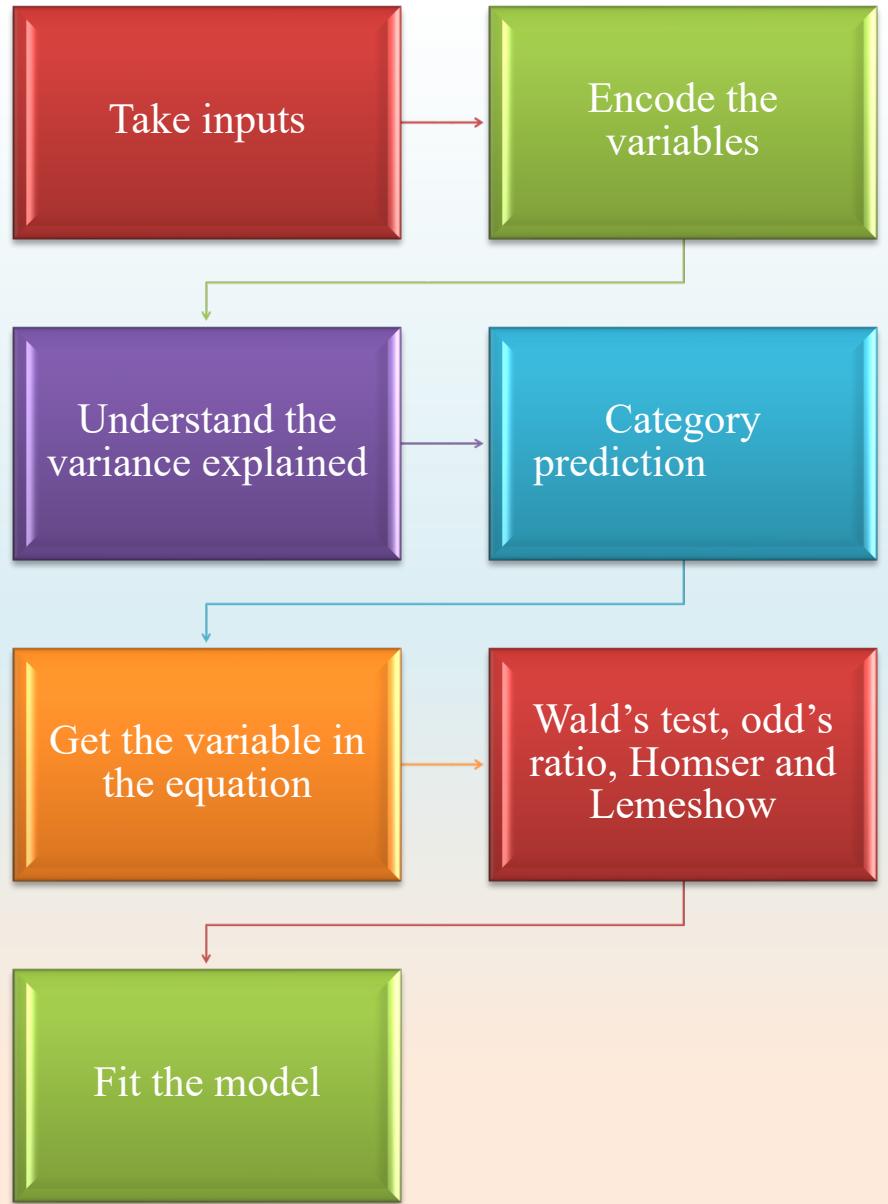
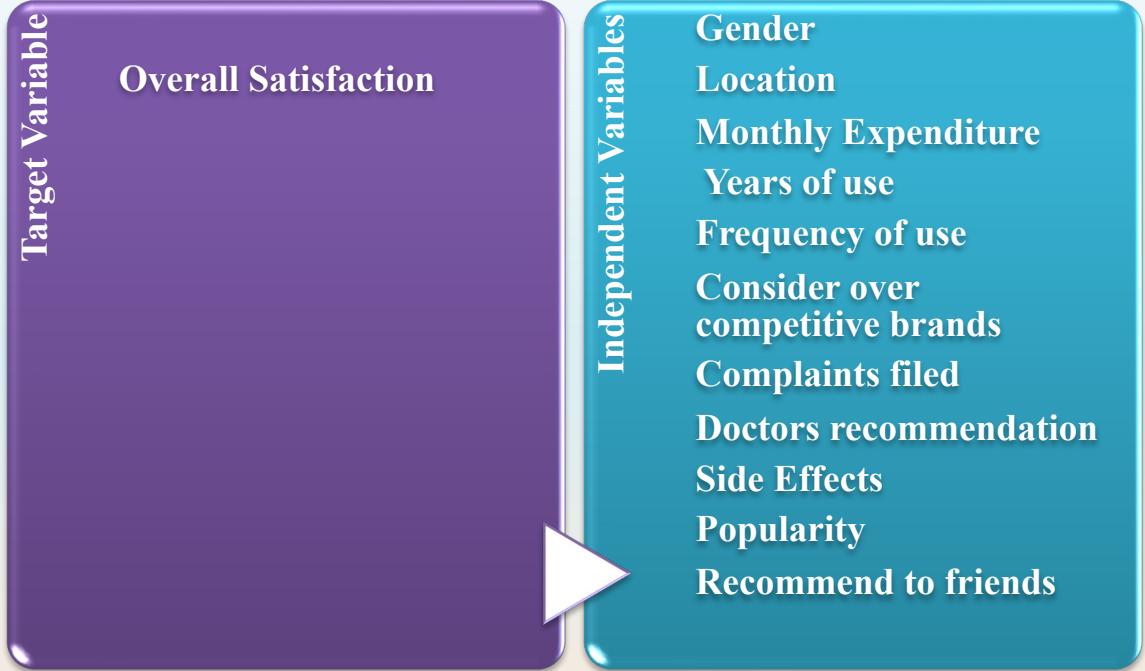
BINARY LOGISTIC REGRESSION

OBJECTIVE 3

To study the influence of various factors on overall satisfaction for the Patanjali users.

- Logistic regression is an extension of simple linear regression, where the dependent variable is dichotomous or binary in nature.
- Logistic regression is the statistical technique used to predict the relationship between predictors (our independent variables) and a predicted variable (the dependent variable) where the dependent variable is binary .





Multicollinearity

Model	Coefficients	
	Tolerance	VIF
1	gender	.962
	location	.964
	Years of use	.927
	Frequency of use	.772
	Monthly Expenditure	.866
	Consider over competitive brands	.851
	Complaints filed	.860
	Doctors Recommendation	.922
	Side Effects	.758
	Popularity	.937
	Recommend to friends	.686

Since, all the VIF's are less than 10, we can conclude that there is no multicollinearity between the independent variables.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	140.593*	.393	.680

The explained variation in the dependent variable based on our model ranges from 39.3% to 68%.

Classification Table

	Observed	Predicted			Percentage Correct	
		Overall, are you satisfied with the Patanjali products?				
		No	Yes			
Step 1	Overall, are you satisfied with the Patanjali products? No	33	27	55.0		
	Yes	14	314	95.7		
	Overall Percentage			89.4		

a. The cut value is .500

Sensitivity= 95.7%

Specificity= 55.0%

Positive predicted value= $100 * (314 / (314 + 27)) = 92.08\%$

Negative predicted value= $100 * (33 / (33 + 14)) = 70.21\%$

	B	S.E.	Wald	df	Sig.	Exp(B)	% C.I. for EXP(B)	
							Lower	Upper
	new_gender(1)	-1.395	.501	7.756	1	.005	.248	.093 .662
	new_location			3.004	2	.223		
	new_location(1)	19.565	7057.631	.000	1	.998	314108944.867	.000 .
	new_location(2)	.856	.494	3.004	1	.083	2.354	.894 6.196
	new_term			13.118	3	.004		
	new_term(1)	-1.404	1.139	1.521	1	.217	.246	.026 2.288
	new_term(2)	-1.668	.563	8.764	1	.003	.189	.063 .569
	new_term(3)	-3.262	1.028	10.075	1	.002	.038	.005 .287
	new_use			7.771	3	.051		
	new_use(1)	1.069	.626	2.914	1	.088	2.913	.853 9.944
	new_use(2)	19.803	3465.637	.000	1	.995	398314346.688	.000 .
Step 1*	new_use(3)	1.821	.653	7.769	1	.005	6.176	1.717 22.219
	new_monthlyspend			.294	3	.961		
	new_monthlyspend(1)	-.859	1.636	.276	1	.600	.424	.017 10.452
	new_monthlyspend(2)	1.096	9263.596	.000	1	1.000	2.993	.000 .
	new_monthlyspend(3)	-.161	.769	.044	1	.834	.851	.189 3.841
	competitive_brand(1)	1.779	.533	11.133	1	.001	5.922	2.083 16.835
	complaints(1)	-.467	.518	.811	1	.368	.627	.227 1.731
	doctor_rec(1)	-.201	.503	.160	1	.689	.818	.305 2.193
	side_effects(1)	-.213	.522	.167	1	.683	.808	.291 2.246
	popularity(1)	-1.676	.794	4.462	1	.035	.187	.039 .886
	recommendation(1)	3.185	.597	28.499	1	.000	24.166	7.505 77.809
	Constant	1.258	1.074	1.371	1	.242	3.519	

The statistical significance of the test is found in the "Sig." column. From these results, you can see that variables with p-value>0.05, did not add significantly to the model.

WALD'S TEST

Hypothesis:

$$H_0 : \beta_i = 0 ; i=1,2,3,4,5$$

$$H_1 : \beta_i \neq 0 ; i=1,2,3,4,5$$

Test Statistic: Under H_0 , the following test statistic follows the standard normal distribution.

$$W = \frac{\beta}{S.E(\beta ii)} \sim (0,1)$$

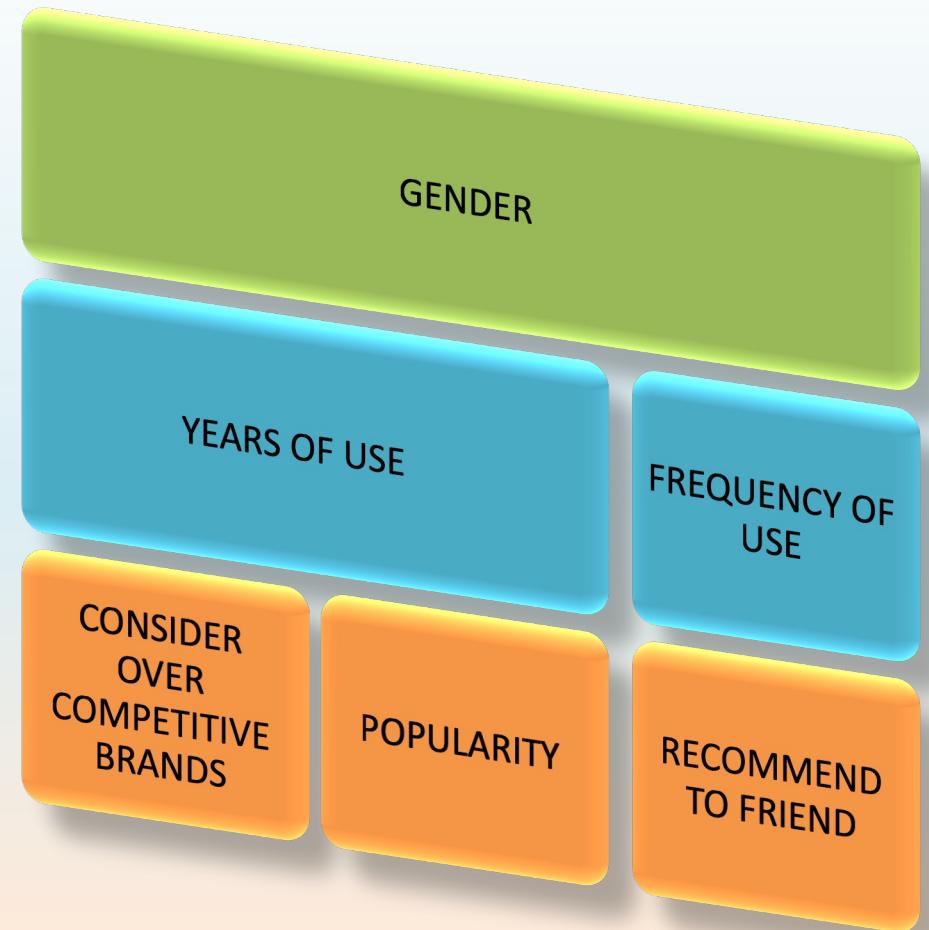
DECISION CRITERIA: Reject H_0 when p-value < 0.05

Variable	p-value
Gender	0.005
Location	0.223
Years of use(new_term)	0.004
Frequency of use(new_use)	0.051
Monthly Expenditure	0.951
Consider over Competitive brands	0.001
Complaints filed	0.368
Doctors recommendation	0.689
Side Effects	0.683
Popularity	0.035
Recommend to a friend	0

Hence, our influential variables are Gender, Years of use, Frequency of use, Consider over Competitive brands, Popularity and Recommendation to a friend.

ODD'S RATIO

Variables	ODD'S RATIO
new_gender(1)	0.248
new_term(1)	0.246
new_term(2)	0.189
new_term(3)	0.038
new_use(1)	2.193
new_use(2)	398314346.688
new_use(3)	6.167
Competitive_brand(1)	5.922
Popularity(1)	0.187
Recommendation(1)	24.166



HOSMER & LEMESHOW (TEST FOR GOODNESS OF FIT)

H_0 : Fitted model is a good fit

H_1 : Fitted model is not a good fit

DECISION CRITERIA: Reject H_0 when p-value < 0.05

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	9.448	8	.306

Since p-value > 0.05, we fail to reject H_0 and conclude that our model is a good fit.

CONFUSION MATRIX

Actual	Predicted	
	No	Yes
No	33	27
Yes	14	314

$$\text{Accuracy} = (\text{TN} + \text{TP}) / (\text{TN} + \text{TP} + \text{FN} + \text{FP})$$

Accuracy: 89.4%

FITTED MODEL

$$g(x) = 1.258 - 1.395 * \text{new_gender}(1) - 1.668 * \text{new term}(2) - 3.262 * \text{new term}(3) + 1.821 * \text{new_use}(3) + 1.779 * \text{competitive_brand}(1) - 1.676 * \text{popularity}(1) + 3.185 * \text{recommendation}(1)$$

new_gender(1)	Male
new_term(2)	Less than a year
new_term(3)	More than 5 years
new_use(3)	Sometimes
competitive_brand(1)	Yes
popularity(1)	Yes
Recommendation(1)	Yes

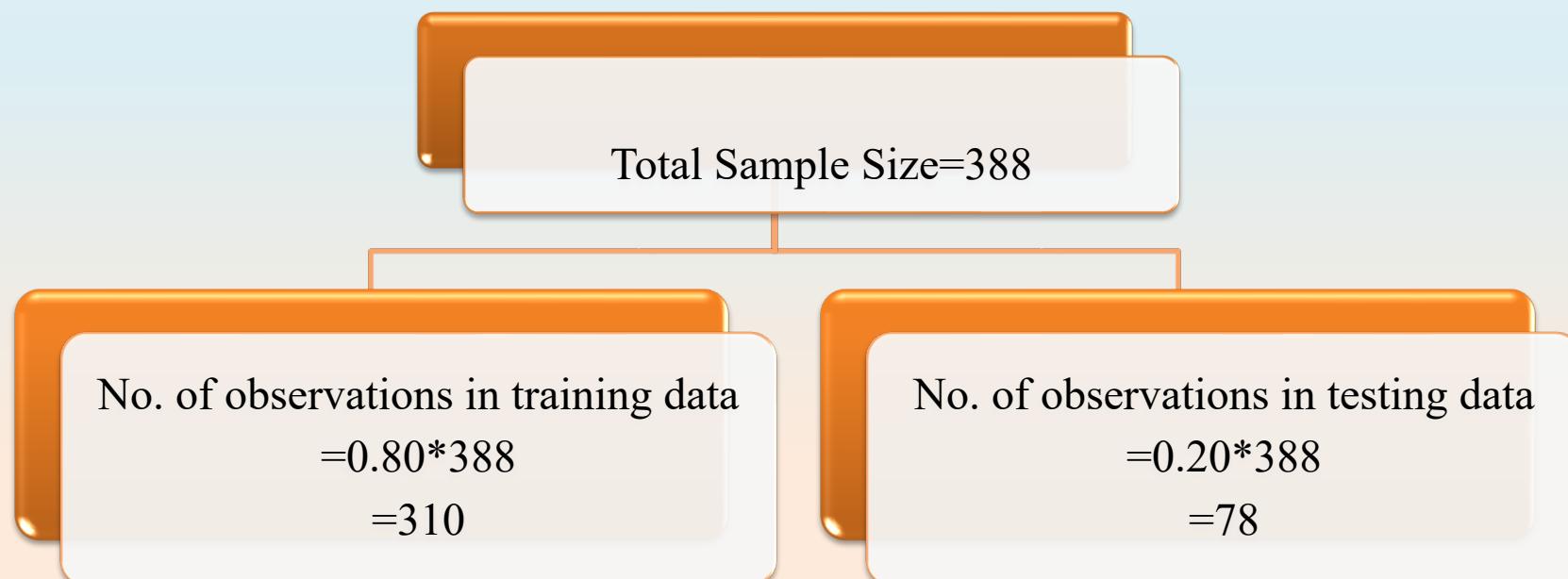
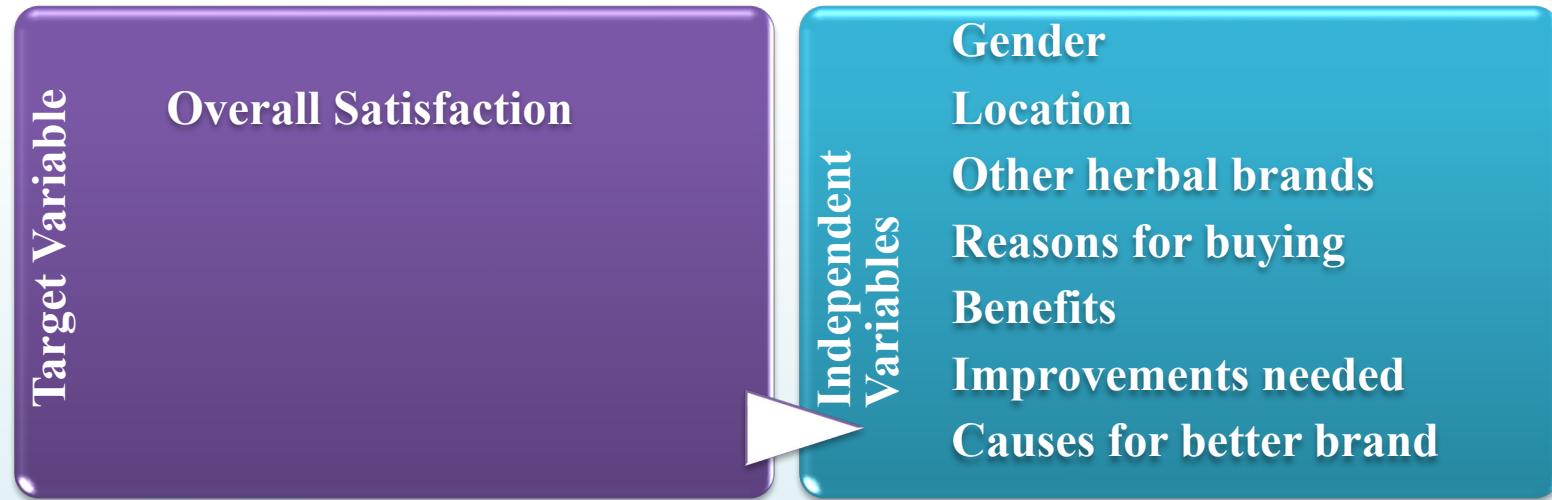
Hence, Gender(male), Years of use(less than a year & more than 5 years), frequency of use(sometimes), consider over competitive brand(yes), popularity(yes) and recommendation to a friend(yes) are the factors that significantly influence the overall satisfaction of Patanjali users.

ARTIFICIAL NEURAL NETWORK & K-NEAREST NEIGHBORS

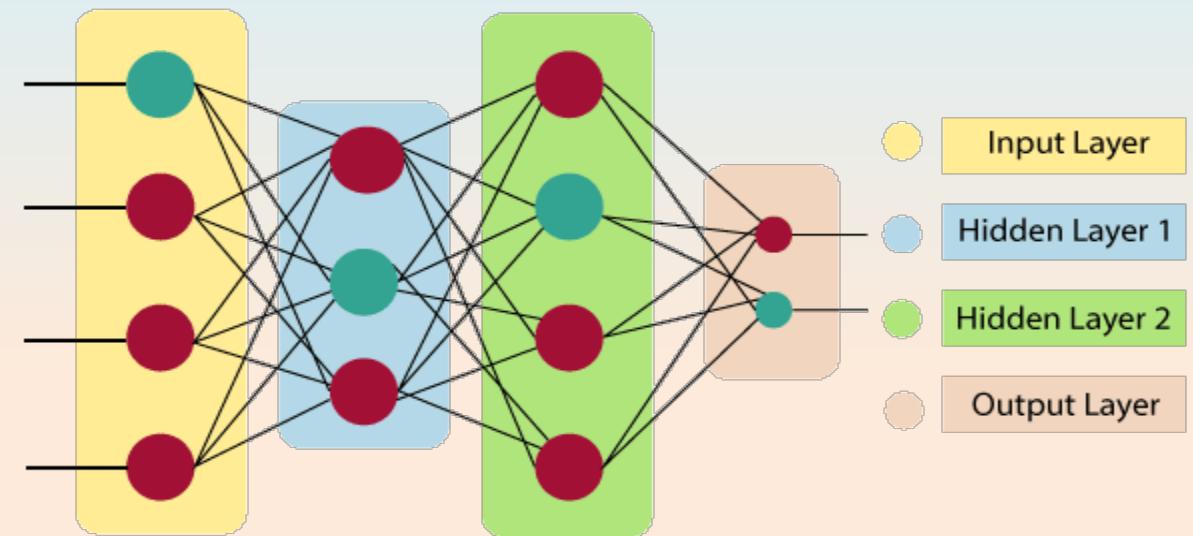
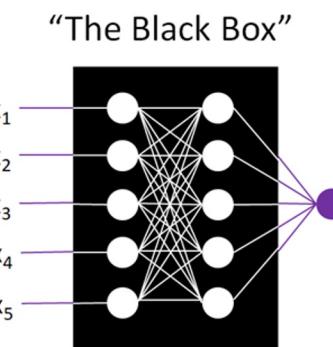
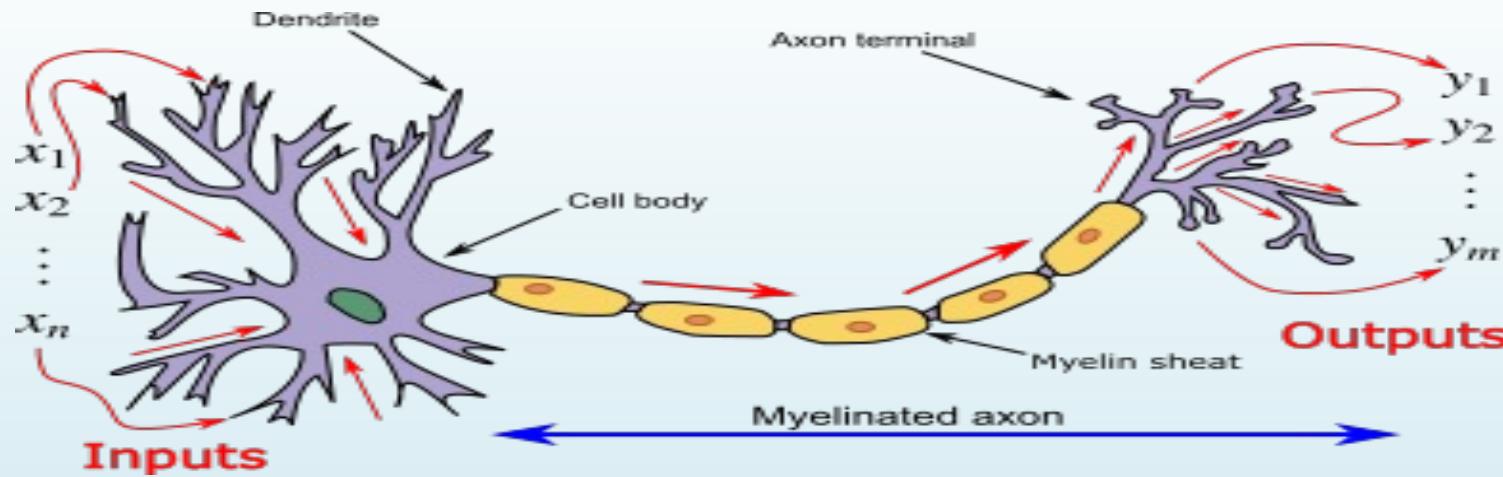
OBJECTIVE 4

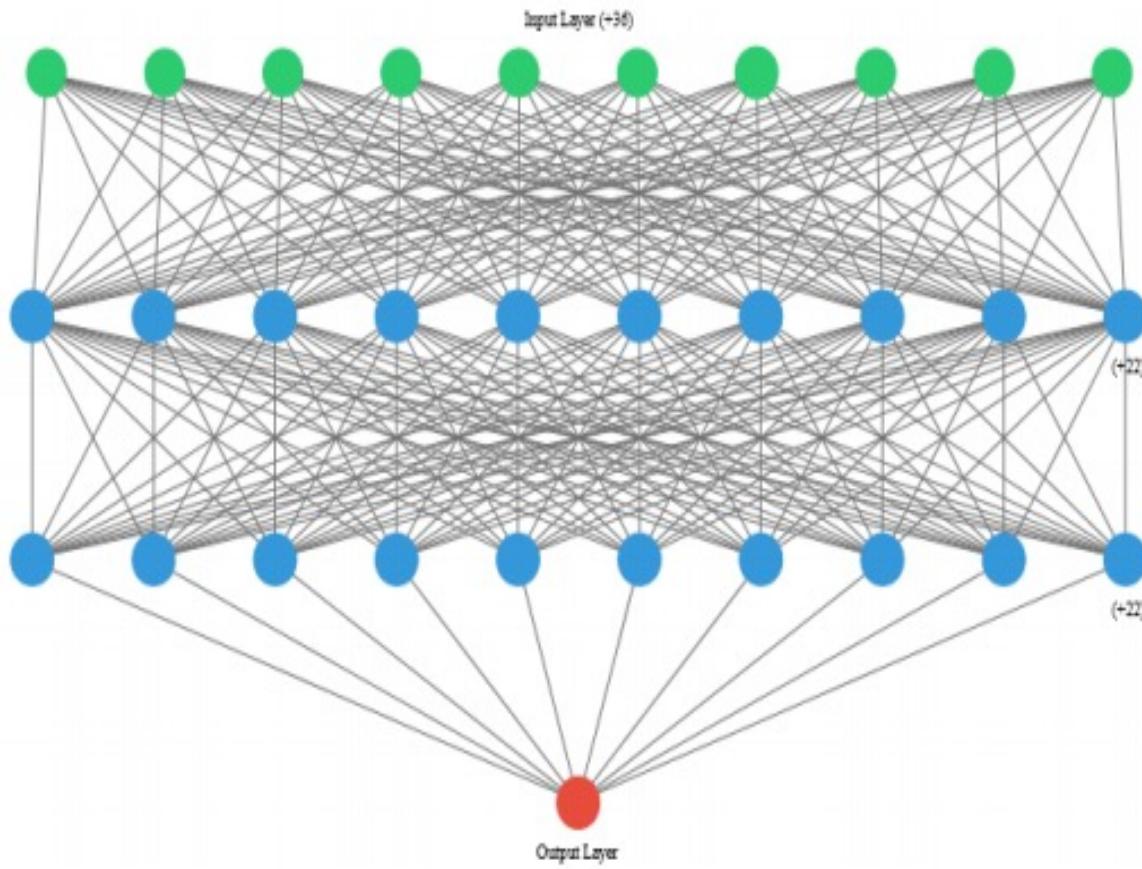
To predict if the Non-Patanjali users will be overall satisfied after using Patanjali products.

- An artificial neural network (ANN) is the component of artificial intelligence that is meant to simulate the functioning of a human brain.
- The k-nearest neighbors (KNN) algorithm is a simple, supervised machine learning algorithm that can be used to solve both classification and regression problems.
- Here, two techniques (KNN & ANN) are used to build our model & check the accuracy score. The model with the best accuracy score is used for further analysis to predict the overall satisfaction of non- Patanjali users.

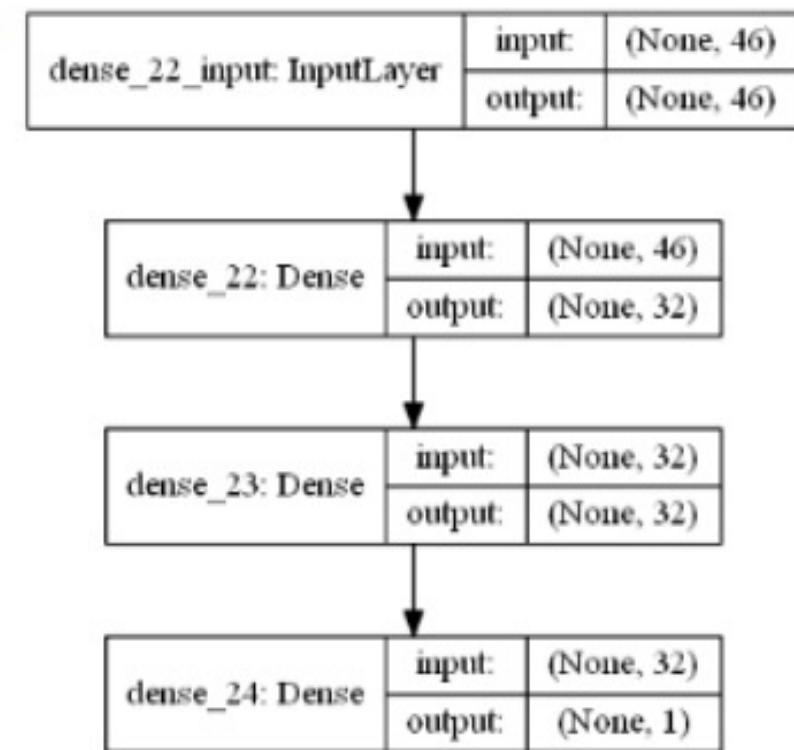


ARTIFICIAL NEURAL NETWORK

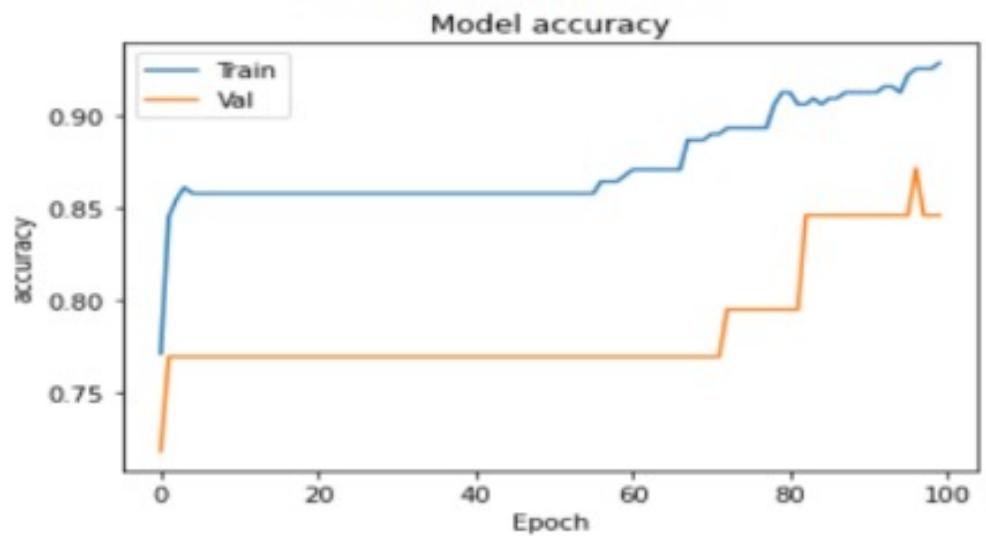




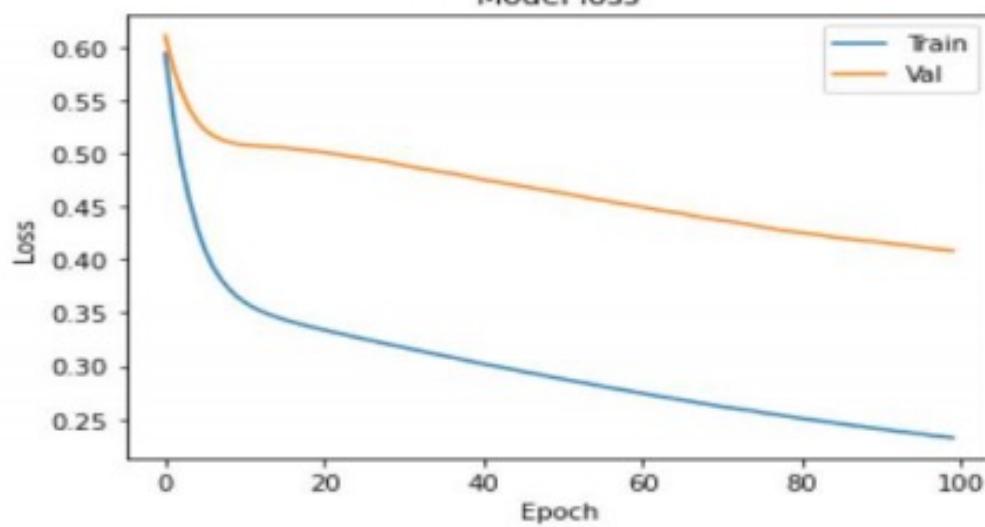
Out[133]:



MODEL ACCURACY PLOT



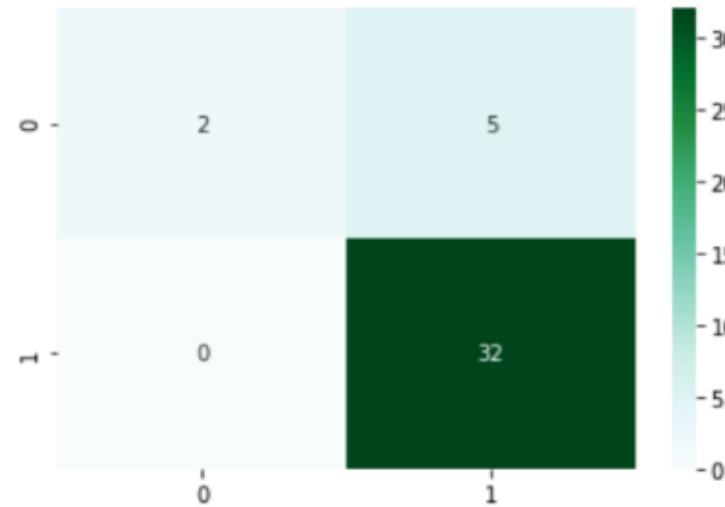
MODEL LOSS PLOT



CONFUSION MATRIX

		Predicted	
		Yes	No
Actual	Yes	TP=2	FP=5
	No	FN=0	TN=32

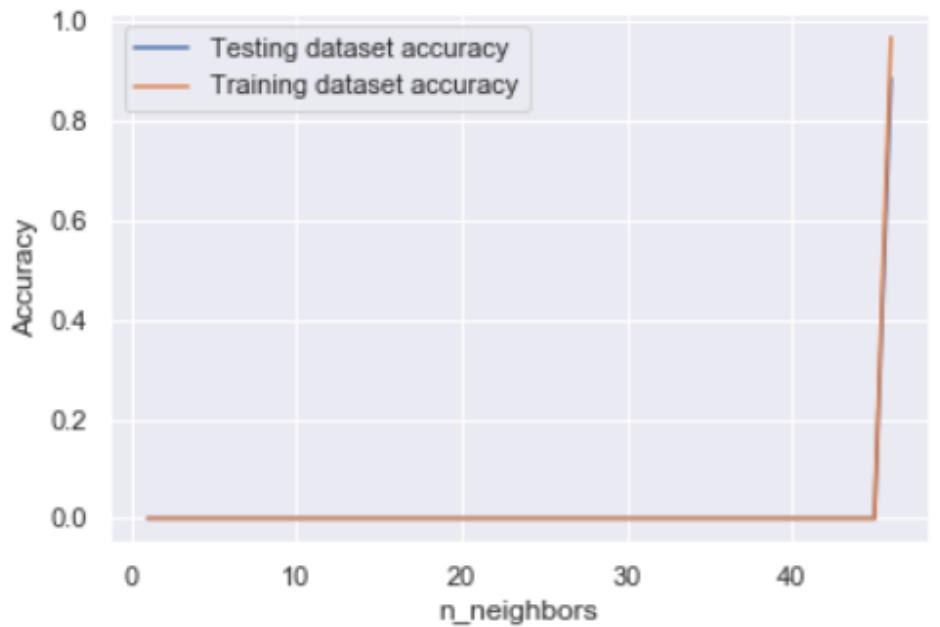
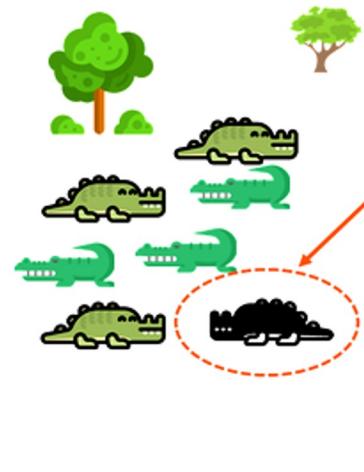
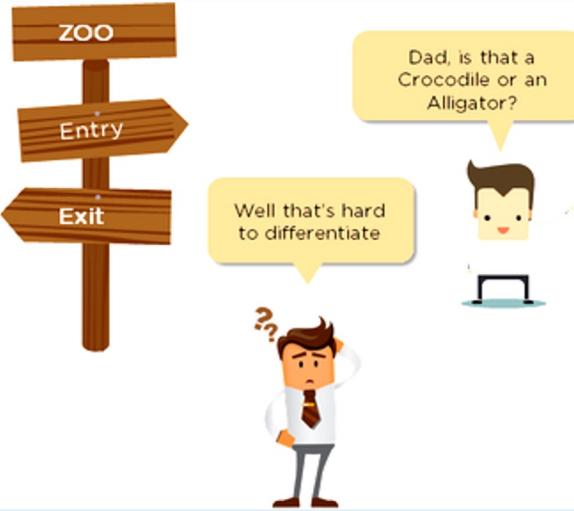
CONFUSION MATRIX PLOT



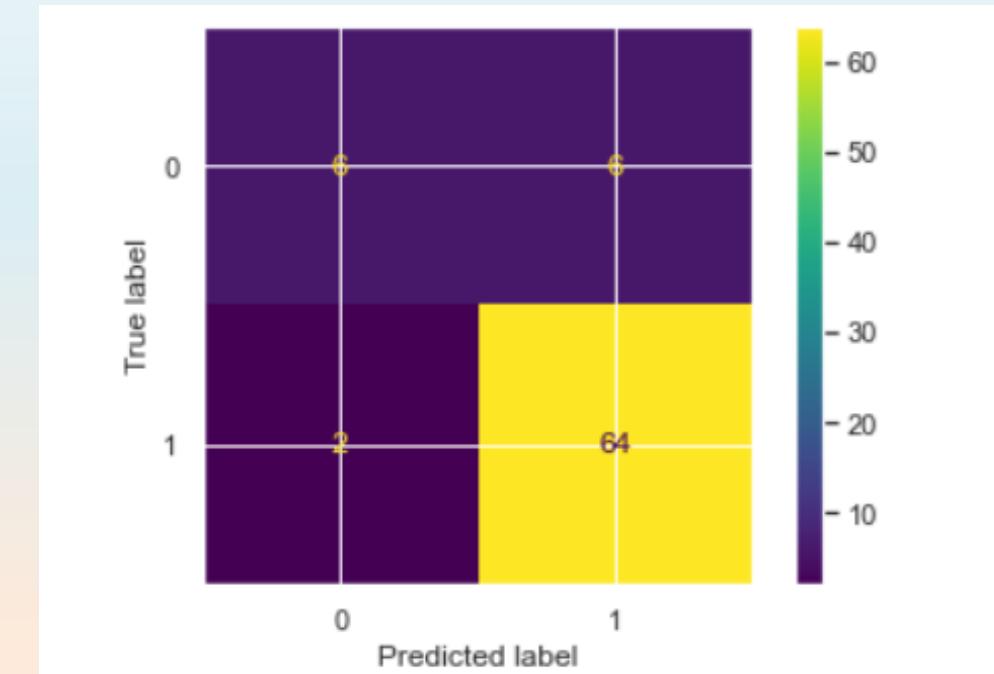
$$\text{Accuracy} = (\text{TN} + \text{TP}) / (\text{TN} + \text{TP} + \text{FN} + \text{FP})$$

Accuracy score = 87.17%

K-NEAREST NEIGHBORS



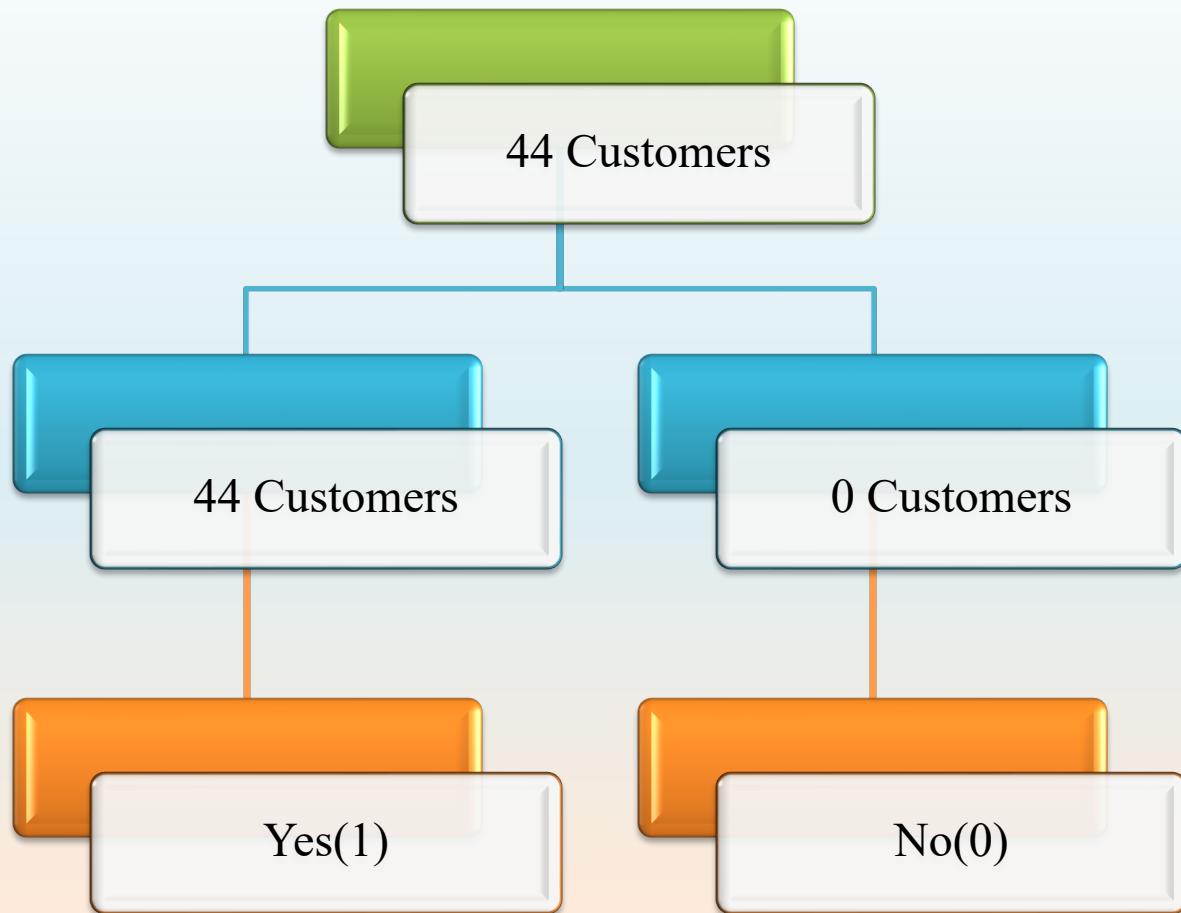
Actual	Predicted	
	Yes	No
Yes	TP=6	FP=6
No	FN=2	TN=64



$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{TP} + \text{FN} + \text{FP}}$$

Accuracy score = 89.74%

PREDICTIONS USING K-NN



DECISION TREES

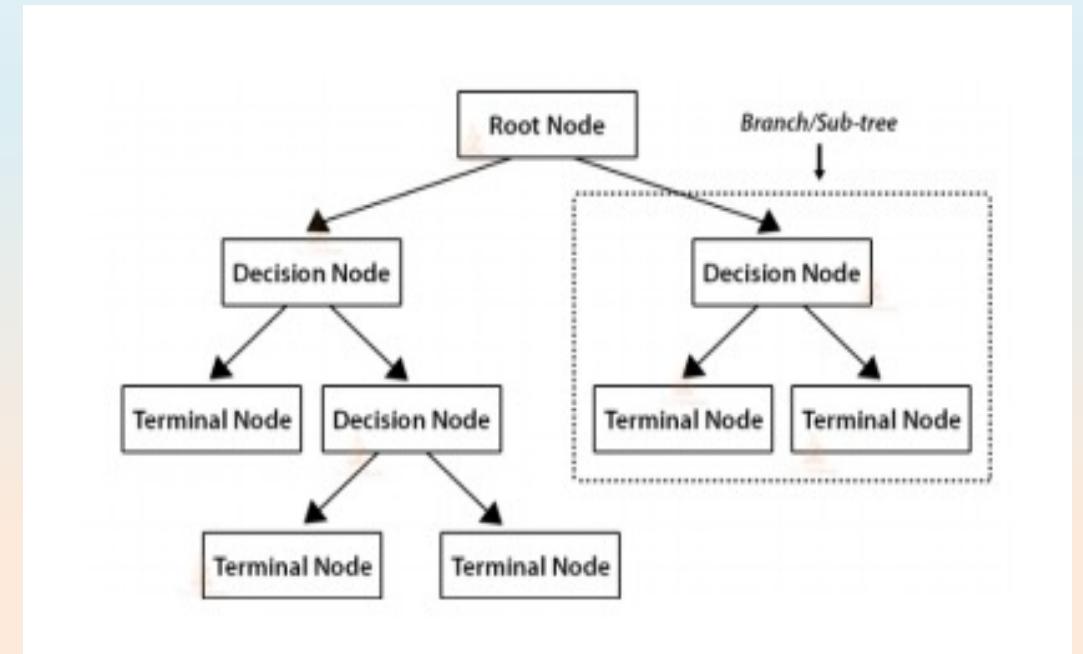
OBJECTIVE 5

To predict how often Customer use Patanjali products based on Socio-Demographic factor, Sources of Awareness and their Buying and Spending Patterns.

- Decision tree analysis is a powerful decision-making tool which initiates a structured nonparametric approach for problem-solving.

TERMINOLOGIES

- Root Node
- Decision Node
- Terminal Node
- Branch
- Splitting
- Pruning

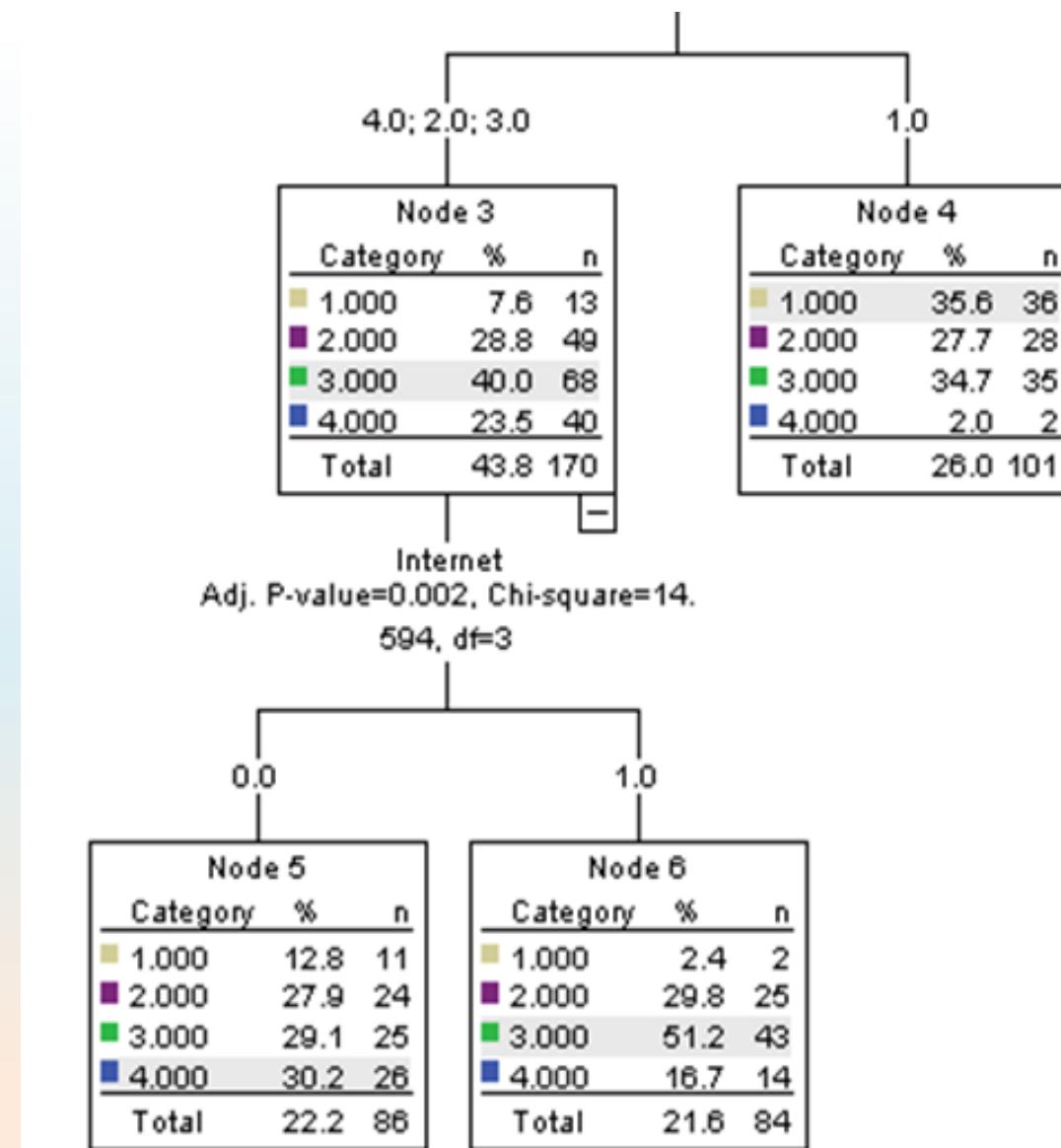
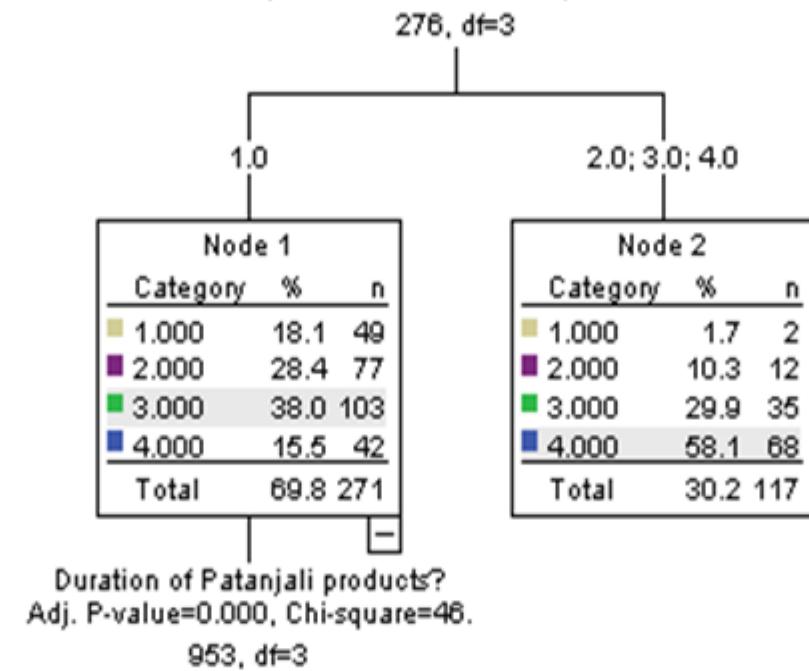
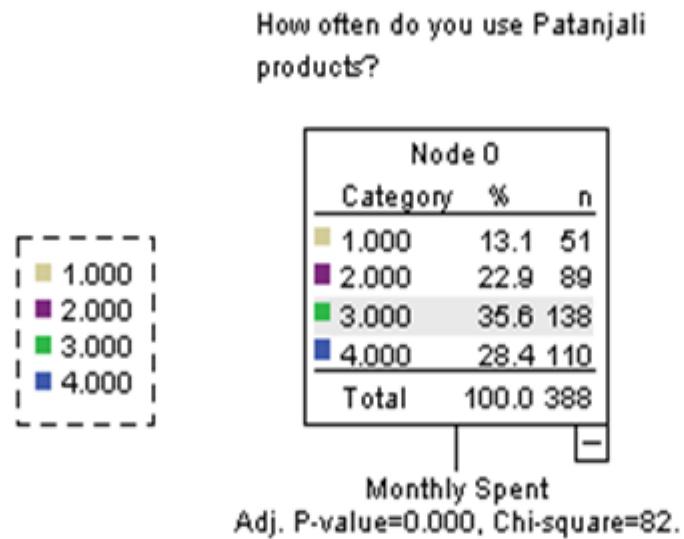


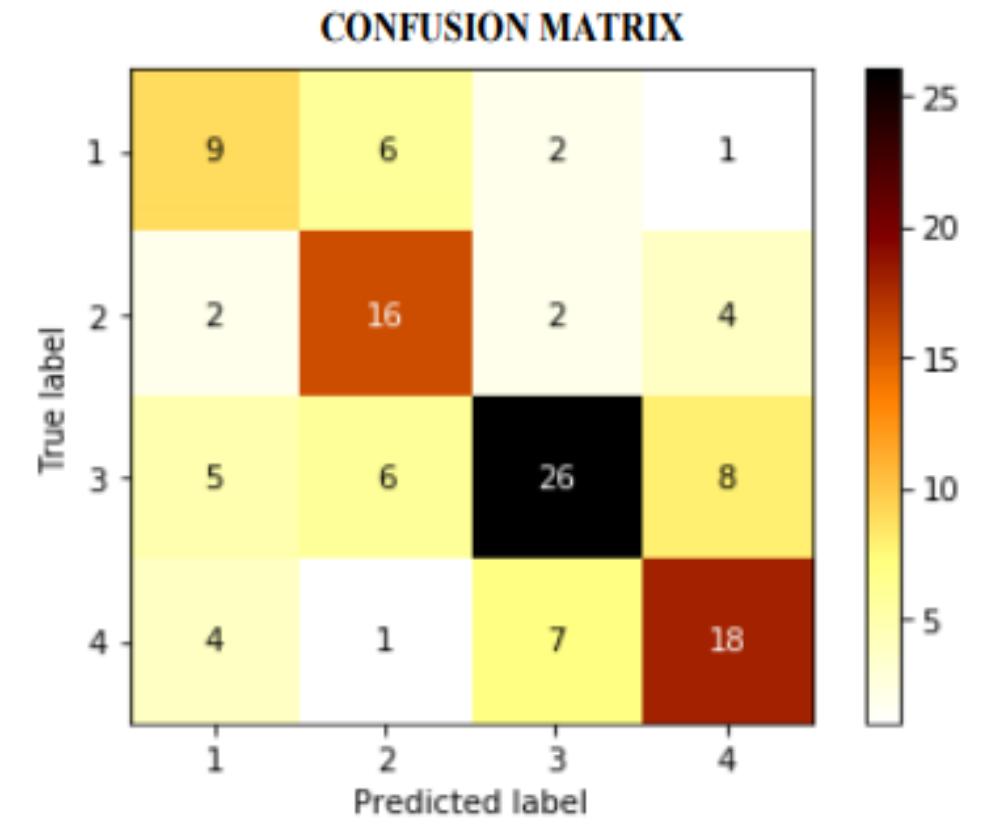
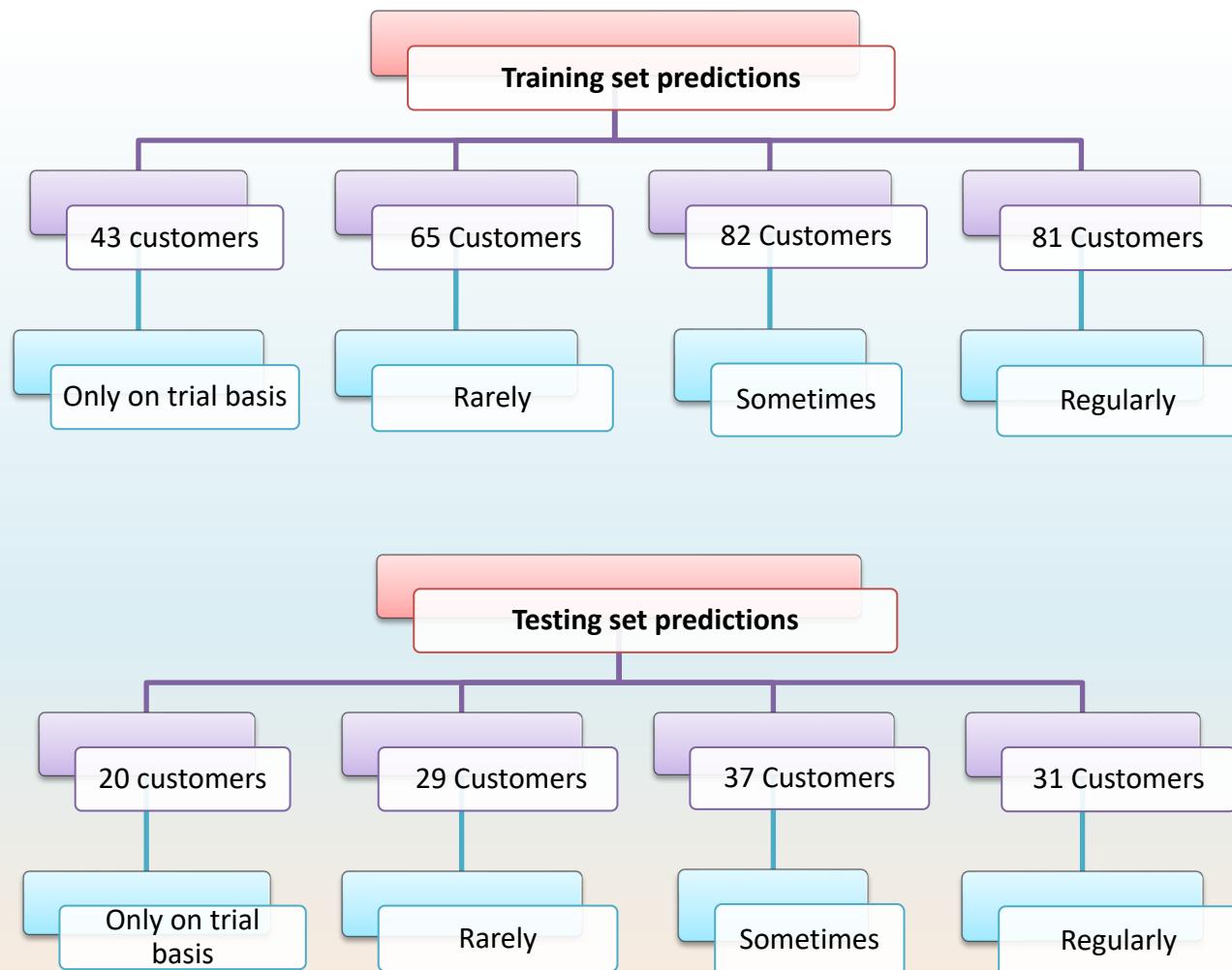
VARIABLE IDENTIFICATION



DATA SPLITTING







Sometimes

Regularly

Rarely

Only on trial basis

Accuracy score = 58.97%

MARKET BASKET ANALYSIS USING APRIORI ALGORITHM

OBJECTIVE 6

To identify next product purchase that might interest a customer.

- The process of discovering frequent item sets in large transactional database is called Market Basket Analysis.



APRIORI ALGORITHM

Systematically identify item set's 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 set's and keep only those with a confidence greater than a pre-specified threshold.

- **Association Rule:** $X \rightarrow Y$
X: Rule antecedent
Y: Rule consequent
- Association rule analysis is a technique to uncover how items are associated to each other. There are three common ways to measure association.
- **Support:** This says how popular an item set is, it is number of times an item appears in total number of transaction in other word we say frequency of item.
- **Confidence:** This says how likely item Y is purchased when item X is purchased, expressed as $\{X \rightarrow Y\}$.
- **Lift:** It is ratio of expected confidence to observed confidence. it is described as confidence of Y when item X was already known(x/y) to the confidence of Y when X item is unknown.

$$\text{Support} = \frac{\text{frq}(X, Y)}{N}$$
$$\text{Rule: } X \Rightarrow Y \longrightarrow \text{Confidence} = \frac{\text{frq}(X, Y)}{\text{frq}(X)}$$
$$\text{Lift} = \frac{\text{Support}}{\text{Supp}(X) \times \text{Supp}(Y)}$$

Steps



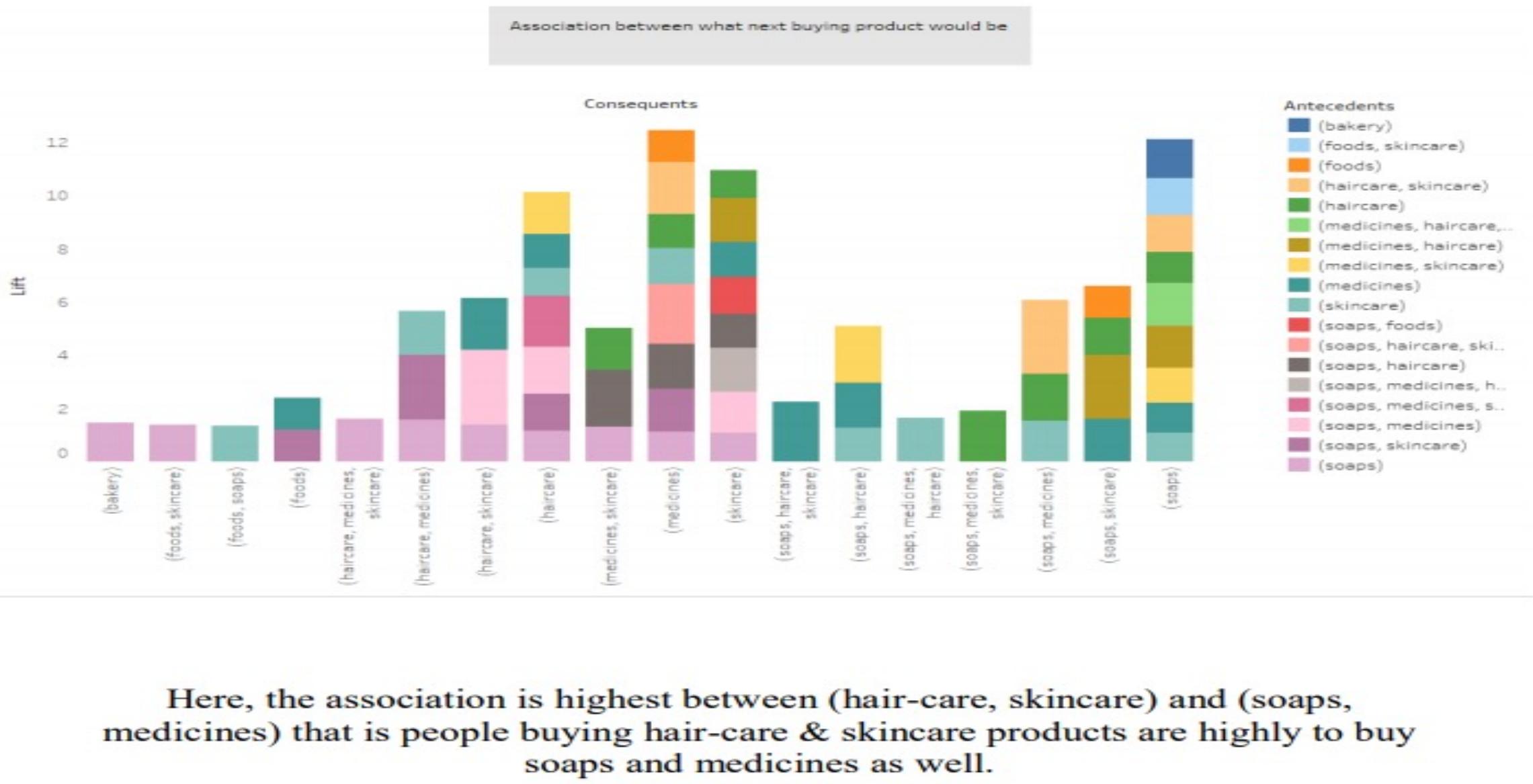
Determine the support of item sets in the transactional database, and select the minimum support and confidence.

Take all supports in the transaction with higher support value than the minimum or selected support value.

Find all the rules of these subsets that have higher confidence value than the threshold or minimum confidence.

Sort the rules as decreasing order of lift.

Products											
	Skin care	Soaps	Medicine	Haircare	Dental care	Spices	Dairy	Bakery	Food & Beverage	Others	
50	(soaps, medicines)		(haircare, skincare)		0.131443	0.211340	0.077320	0.588235	2.783357	0.049540	1.915317
9		(haircare)		(soaps)	0.355670	0.494845	0.206186	0.579710	1.171498	0.030184	1.201920
7		(medicines)		(soaps)	0.234536	0.494845	0.131443	0.560440	1.132555	0.015384	1.149227
36		(medicines, skincare)		(haircare)	0.175258	0.355670	0.097938	0.558824	1.571185	0.035604	1.460481
1		(skincare)		(soaps)	0.561856	0.494845	0.301546	0.536697	1.084576	0.023515	1.090334
48	(soaps, haircare, skincare)			(medicines)	0.146907	0.234536	0.077320	0.526316	2.244072	0.042865	1.615979
23		(soaps, skincare)		(haircare)	0.301546	0.355670	0.146907	0.487179	1.369751	0.039656	1.256443
20		(medicines)		(soaps, skincare)	0.234536	0.301546	0.113402	0.483516	1.603456	0.042679	1.352325
35		(haircare, skincare)		(medicines)	0.211340	0.234536	0.097938	0.463415	1.975878	0.048371	1.426546
13		(medicines)		(haircare)	0.234536	0.355670	0.105670	0.450549	1.266762	0.022253	1.172680



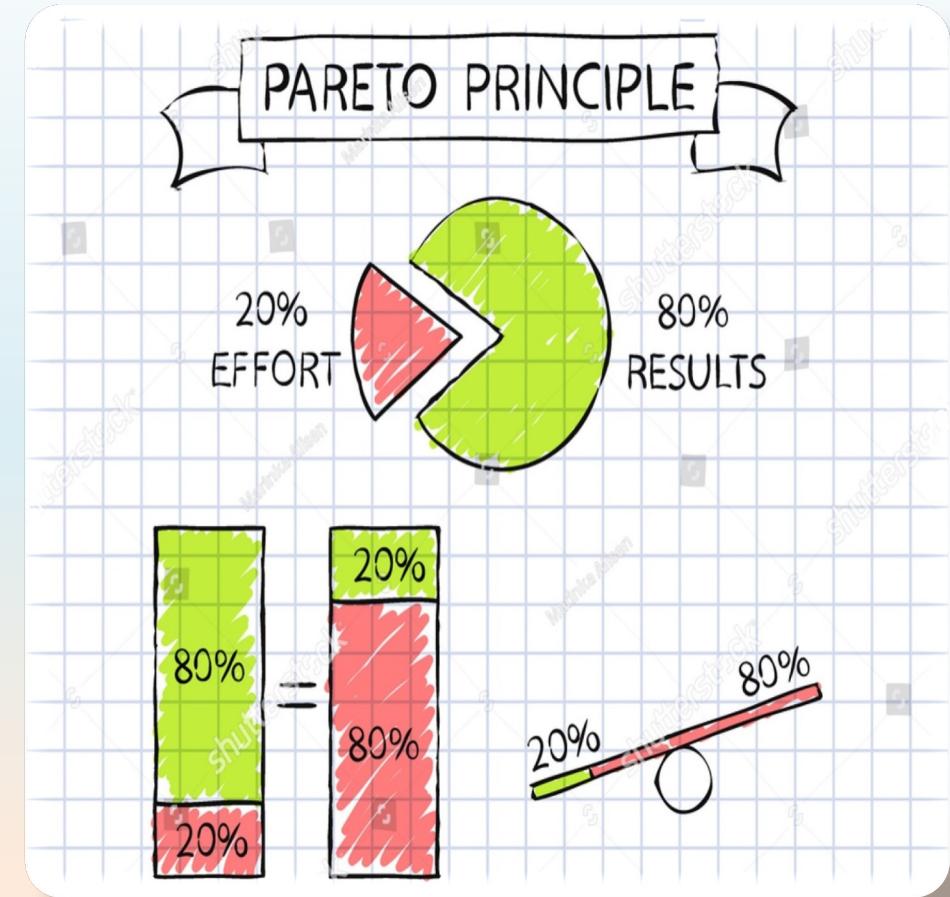
Here, the association is highest between (hair-care, skincare) and (soaps, medicines) that is people buying hair-care & skincare products are highly likely to buy soaps and medicines as well.

PARETO ANALYSIS

OBJECTIVE 7

To study the factors that affects Patanjali from being a better brand.

- Pareto Analysis is a statistical technique in decision-making.
- It uses the Pareto Principle (also known as the 80/20 rule) the idea that by doing 20% of the work you can generate 80% of the benefit of doing the entire job.
- A Pareto chart is a basic quality tool that helps you identify the most frequent defects, complaints, or any other factor you can count and categorize.



1. To identify factor that affects preferences of people while buying Patanjali products.

Variables	Frequency	Cumulative %
Health-wise	239	0.302532
Quality-wise	208	0.565823
Price wise	183	0.797468
Brand wise	68	0.883544
Quantity wise	67	0.968354
Demand wise	25	1

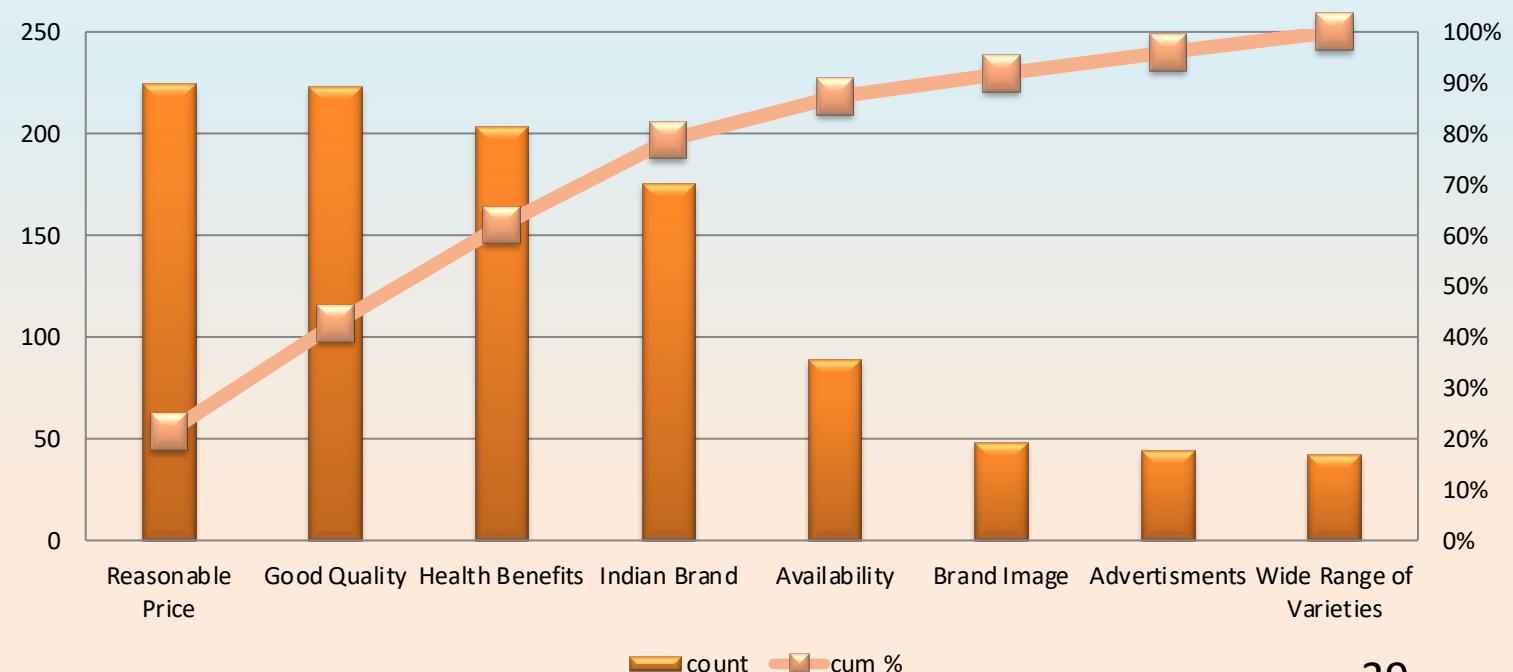
Health-wise, Quality-wise & Price-wise are the major factors that affects the preferences of people while buying Patanjali products.



2. To identify reasons for not buying Patanjali Products.

Reasons for buying Patanjali products	Frequency	Cumulative %
Reasonable Price	224	0.21374
Good Quality	223	0.426527
Health Benefits	203	0.620229
Indian Brand	175	0.787214
Availability	89	0.872137
Brand Image	48	0.917939
Advertisements	44	0.959924
Wide Range of Varieties	42	1

Lack of availability, lack of brand image, lack of advertisements & lack of wide ranges of varieties are the reasons for not buying the Patanjali product.



3. To identify in which category Patanjali needs improvement so as to be a better brand.

Categories for improvement	Frequency	Cumulative %
Online Services	145	0.20195
Ayurvedic Standards	144	0.402507
Customer Satisfaction	124	0.575209
Marketing Strategy	122	0.745125
Availability	85	0.86351
Quality	71	0.962396
Annual Sales	27	1

So as to improvise as a brand, Patanjali needs to improve the online services, ayurvedic standards, customer satisfaction and Marketing strategy.

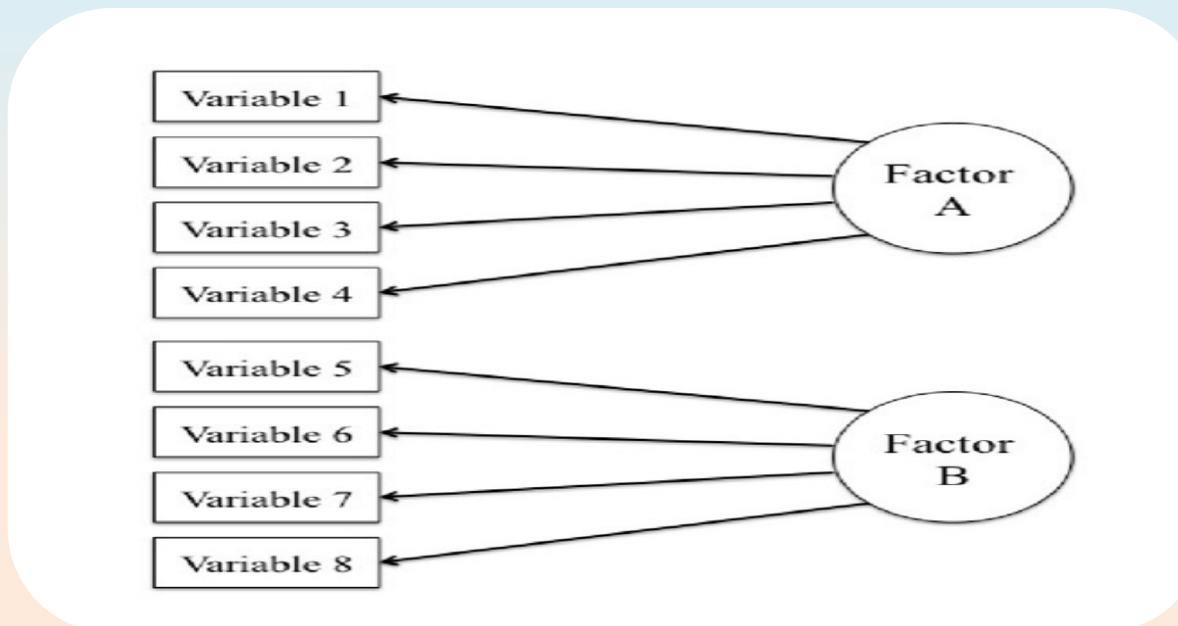
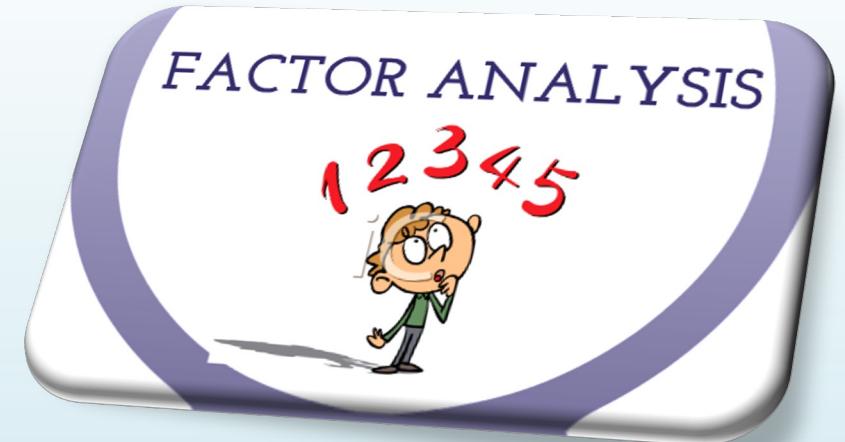


FACTOR ANALYSIS

OBJECTIVE 8

To determine the factors that influences the people for buying Patanjali products.

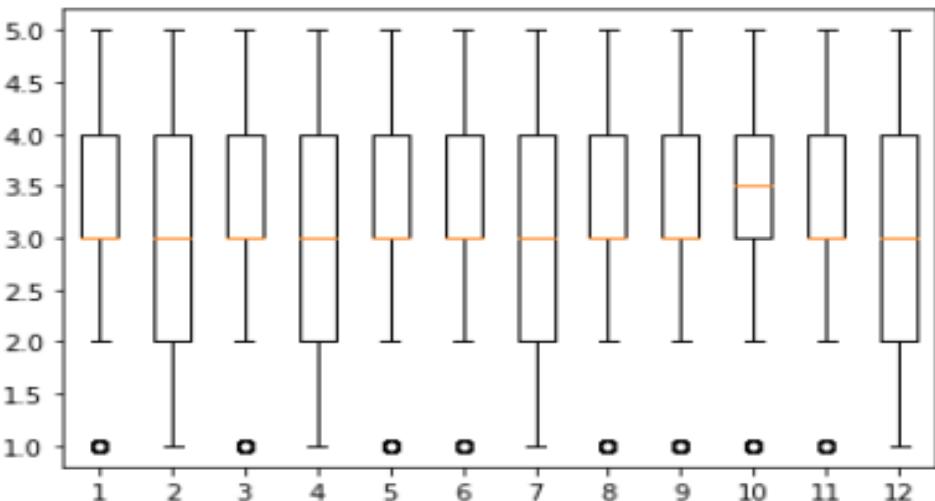
- Factor analysis is a statistical method used to describe variability among observed, correlated variables in terms of a potentially lower number of unobserved variables called factors.



Variables considered

- V1: Popularity
- V2: Advertisements
- V3: Quality
- V4: Availability
- V5: Purity
- V6: Organic
- V7: Appealing to youngsters
- V8: Increase the promotion
- V9: Brand Ambassadors
- V10: Low price
- V11: Beneficial
- V12: Customer service

OUTLIERS



No outliers present

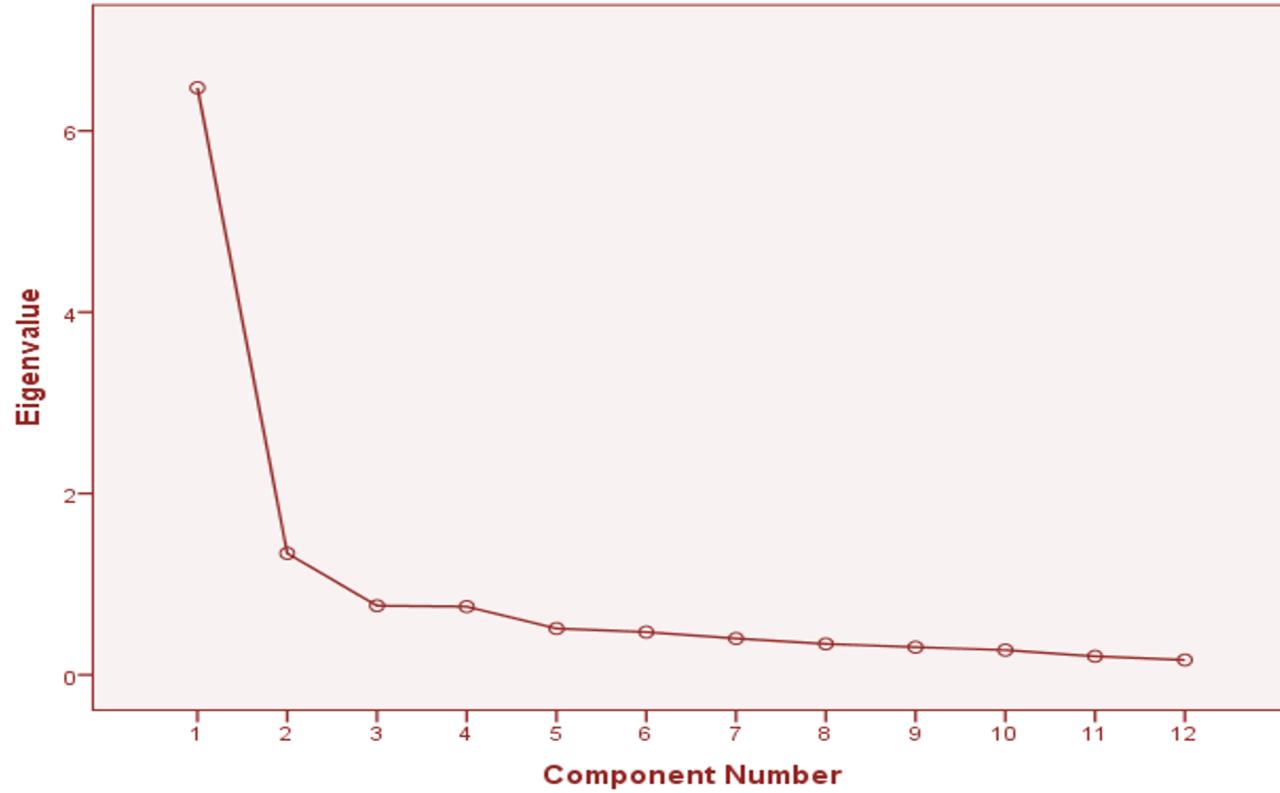
ASSUMPTIONS CHECKING

MULTICOLLINEARITY

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.904
Bartlett's Test of Sphericity	Approx. Chi-Square	2939.04
	df	2
	Sig.	.000

KMO Criterion: $0.904 > 0.5$
Bartlett's Test of Sphericity: p-value(0.000) < 0.05

Scree Plot



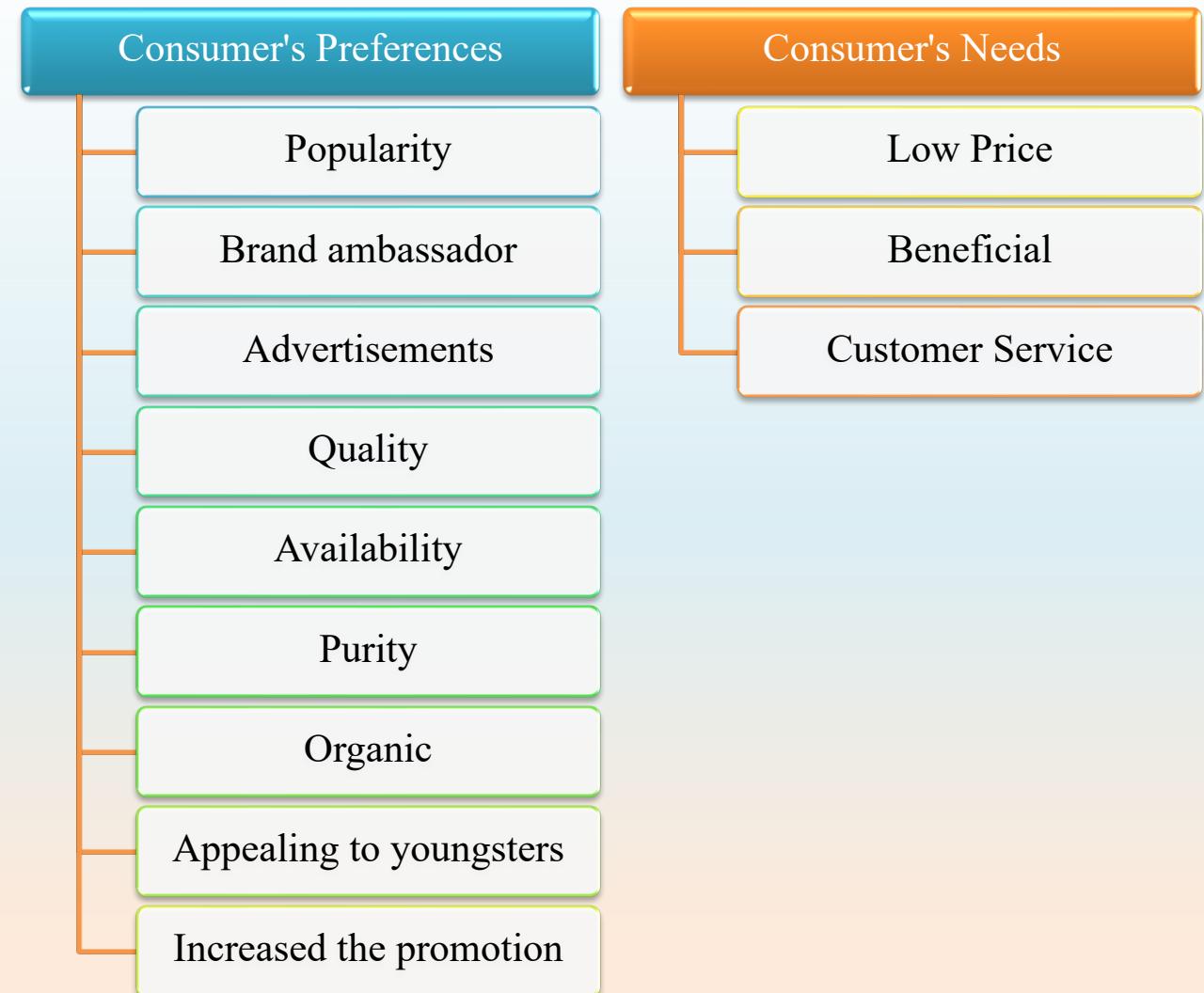
The point where the slope of the curve is clearly leveling off (the “elbow”) indicates the number of factors that should be generated by the analysis which is after the 2th component.

Rotated Component Matrix^a

	Component	
	1	2
Popularity	.768	.259
Advertisements	.736	.143
Quality	.797	.211
Availability	.725	.258
Purity	.815	.252
Organic	.821	.327
Appealing to Youngsters	.747	.132
Increase the promotion	.615	.423
Brand ambassadors	.525	.433
Low Price	.080	.853
Beneficial	.322	.819
Customer Service	.315	.783

Varimax Method

	R1	R2
V1	0.768	
V2	0.736	
V3	0.797	
V4	0.725	
V5	0.815	
V6	0.821	
V7	0.747	
V8	0.615	
V9	0.525	
V10		0.853
V11		0.819
V12		0.783

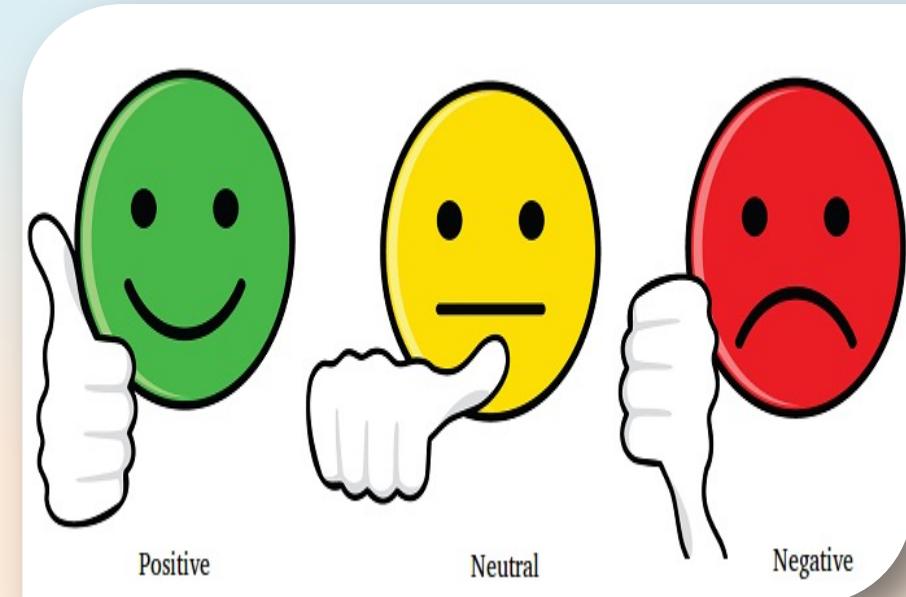


SENTIMENTAL ANALYSIS

OBJECTIVE 9

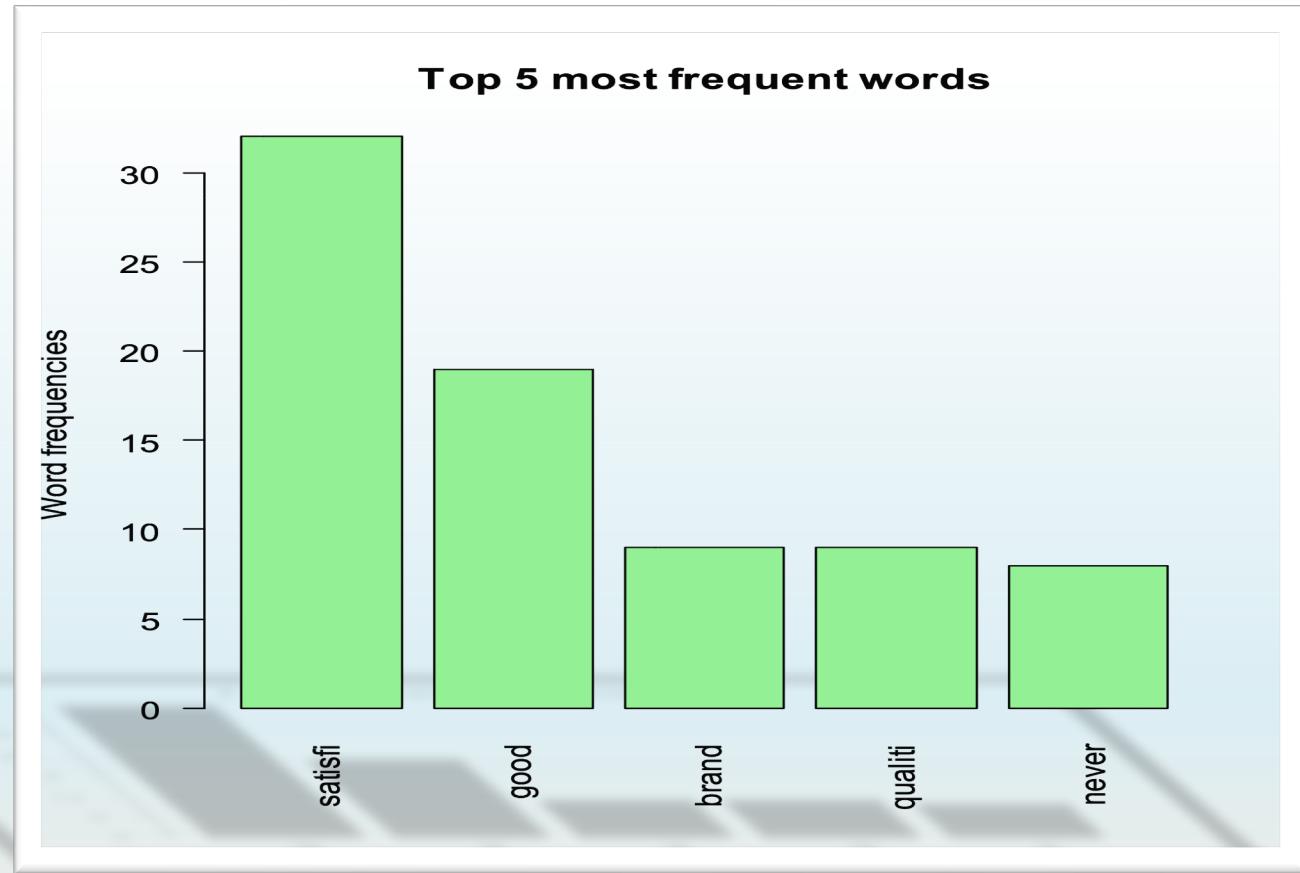
To analyze the sentimental behaviour of the customers towards the brand Patanjali.

- Sentiment analysis (or opinion mining) is a natural language processing technique used to determine whether data is positive, negative or neutral.



5 MOST FREQUENT WORDS

	Word	Freq
satisfied	Satisfy	32
good	good	19
brand	brand	9
quality	quality	9
never	never	8



WORD ASSOCIATION

\$work							
togeth							
0.4							
\$good							
integr	synergi						
0.28	0.28						
\$health							
declin	happen	noth	real	sentiment	suppli	wors	
0.29	0.29	0.29	0.29	0.29	0.29	0.29	
\$several							
bad							
0.26							
\$great							
journey	satisfact	march	goal	pursu	toward	hard	
0.52	0.52	0.36	0.35	0.28	0.26	0.26	
\$feel							
across	board	harsh	system	somewhat			
0.33	0.32	0.32	0.32	0.29			
\$improv							
room	perfect	propel	thik	attitud			
0.41	0.35	0.35	0.35	0.32			

SENTIMENT SCORES

SYUZHET METHOD

Min	1 st Qu.	Median	Mean	3 rd Qu.	Max
-1.45	0.9	1.6	1.88	2.9	9

BING METHOD

Min	1 st Qu.	Median	Mean	3 rd Qu.	Max
-3	1	-3	2.07	3	9

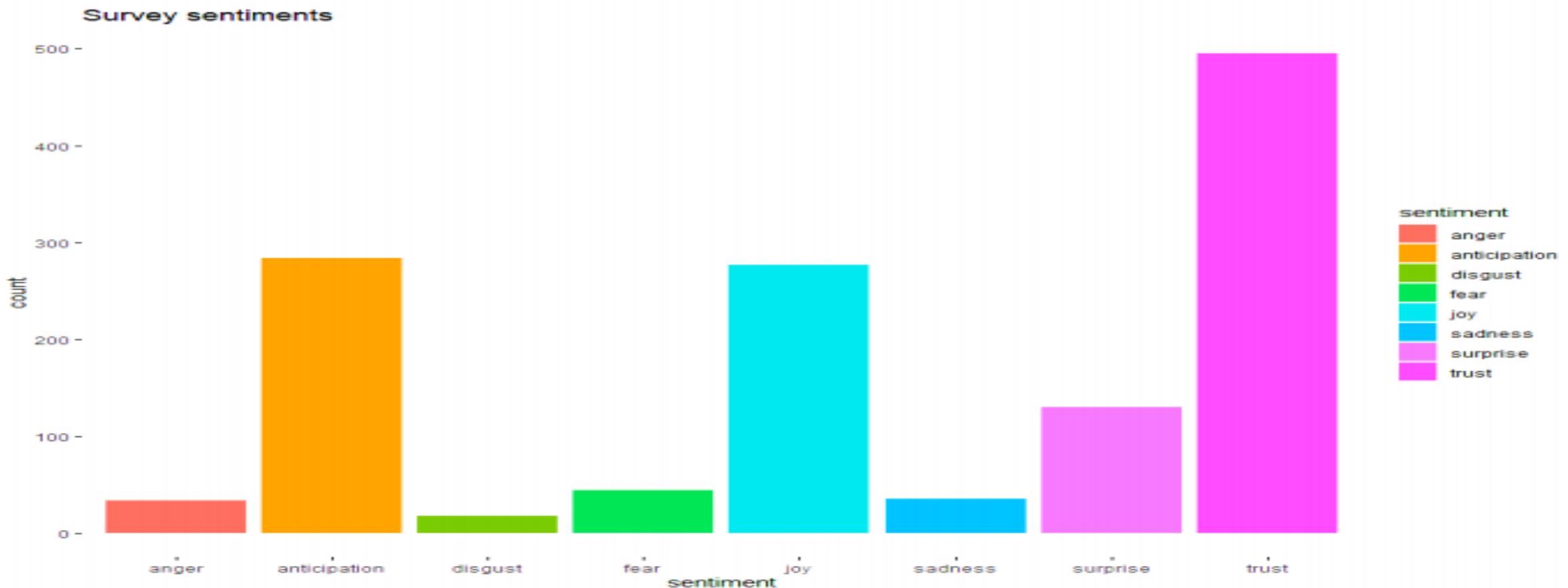
AFFIN METHOD

Min	1 st Qu.	Median	Mean	3 rd Qu.	Max
-6	2	4	4.36	7	18

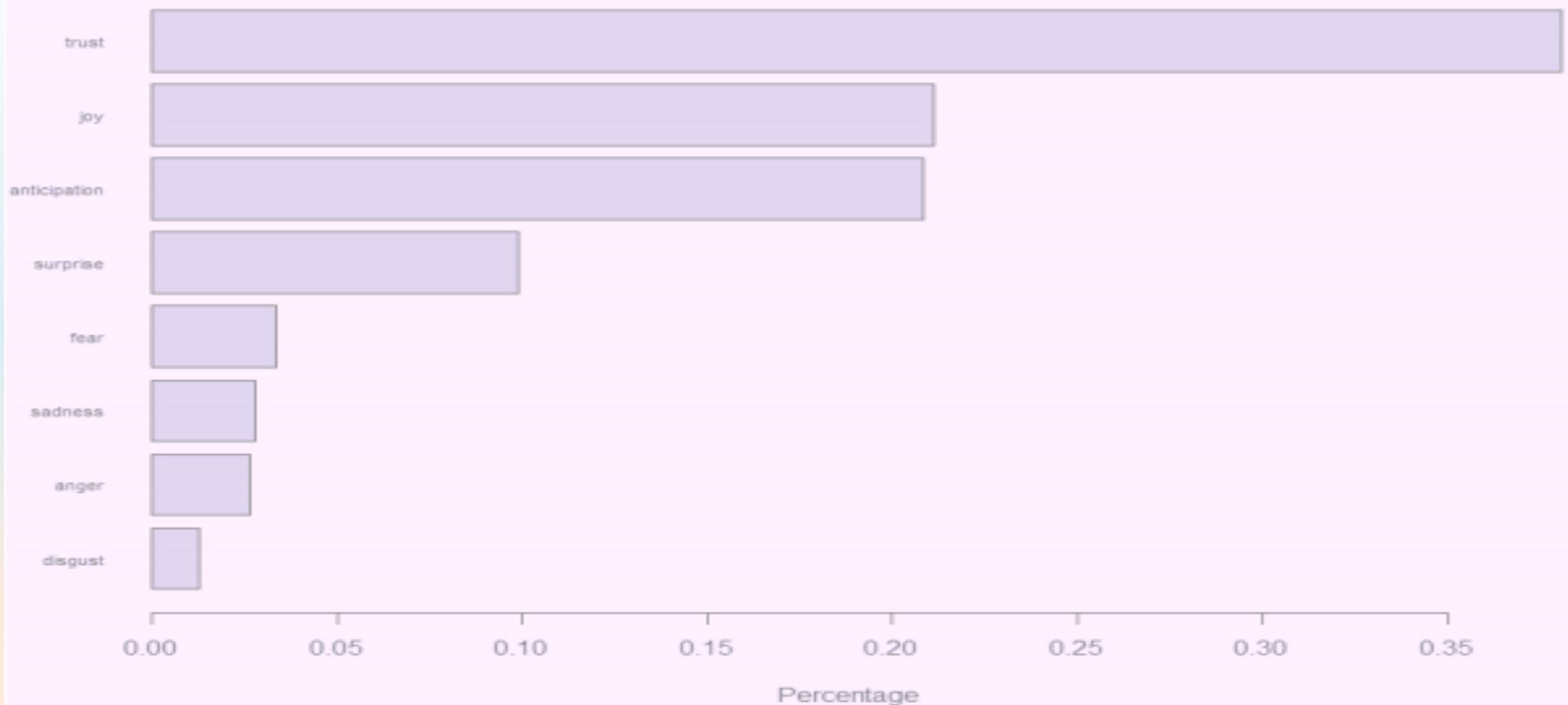
EMOTION CLASSIFICATION

Reviews	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust	Negative	Positive
1	0	1	0	0	1	0	0	2	1	2
2	0	3	0	1	0	0	0	1	1	5
3	0	1	0	0	1	1	0	0	0	4
4	0	3	0	0	2	0	1	1	2	2
5	0	2	0	0	2	0	0	3	1	3
6	0	0	0	0	0	0	0	4	0	1
7	0	2	0	0	2	0	0	4	0	6
8	0	4	0	0	4	0	0	4	1	5
9	0	3	0	0	3	0	0	3	0	5
10	1	1	0	1	0	0	0	1	0	3

Bar Plot showing the count of words in the text, associated with each emotion



Emotions in Text



Conclusions

SWOT ANALYSIS

- ✓ Patanjali Ayurveda Ltd have a strong impact on popularity due to the reputation of Baba Ramdev, patriotism, low prices of products & purity.
- ✓ The company needs to improve in the fields of manufacturing units, over dependency on Baba Ramdev, no longer penetration pricing, product dependence & lack of experienced graduates.
- ✓ The company have great opportunities in growth of organic sector, can be expanded in rural markets, it can be expanded globally, huge tie-ups and to be diversified.
- ✓ Price war, increasing competition, negative word- of mouth & poor reap can affect the company's growth & reputation.

CHI-SQUARE TEST OF INDEPENDENCE

- ✓ Patanjali product users(male or female) have no association with the health consciousness.
- ✓ That is, there is no association between gender & doctor's recommendation while buying the products.
- ✓ Also, there is no association between gender & side effects while using the Patanjali products.

FACTOR ANALYSIS

- ✓ **Consumer's Preferences:** Popularity, Purity, Brand ambassadors, Quality, Organic, Increased promotion, Appealing to youngsters, Advertisements, Availability
- ✓ **Consumer's Needs:** Low price, Beneficial, Customer Service

SENTIMENTAL ANALYSIS

- ✓ The analysis shows that Patanjali users have a very high emotion of "Trust" on their products & a very low emotion of "Disgust" for the products.
- ✓ Hence, this is a good sign of team health.

LOGISTIC REGRESSION

- ✓ Patanjali have the highest influence on years of use(less than a year & more than 5 years), frequency of use(sometimes), consideration over other brands, popularity & recommendation to their friends for the overall satisfaction of their customers.

ARTIFICIAL NEURAL NETWORKS & K-NEAREST NEIGHBORS

✓ These techniques were conducted to predict if non-Patanjali users will be overall satisfied after using the Patanjali products.

✓ After checking the accuracy for both the techniques, K-NN having the highest accuracy, we conclude that all the non-Patanjali users will be overall satisfied after using the Patanjali products.

DECISION TREES

✓ After making predictions for training & testing sets, majority of the people will use Patanjali products sometimes, followed by people using them regularly and rarely.

✓ Less number of people will use them only on trial basis.

MARKET BASKET ANALYSIS USING APRIORI ALGORITHM

- ✓ The customers would prefer buying soaps & medicines after buying haircare & skincare products on their next purchase, as this technique shows the maximum association between these products.

PARETO ANALYSIS

- ✓ Health-wise, Quality-wise & Price-wise are the major factors that affects the preferences of people while buying the products.
- ✓ Lack of availability, lack of brand image, lack of advertisements & lack of wide range of varieties are the reasons for not buying the Patanjali products.
- ✓ So as to improvise as a brand, Patanjali needs to improve the online services, Ayurvedic standards, customer satisfaction and Marketing strategy.

References

- <http://www.analyticsvidhya.com/>
- <https://stackoverflow.com/>
- <https://www.patanjaliayurved.net/>
- <https://bstrategyhub.com/patanjali-swot-analysis-2019-swot-analysis-of-patanjali/>
- <https://github.com/>
- <https://libguides.library.kent.edu/spss/chisquare#:~:text=The%20Chi%2DSquare%20Test%20of%20Independence%20determines%20whether%20there%20is,Chi%2DSquare%20Test%20of%20Association.>
- <https://www.kdnuggets.com/2015/05/top-10-data-mining-algorithms-explained.html>
- <https://pythonprogramminglanguage.com/top-data-mining-algorithms/>
- <https://www.geeksforgeeks.org/machine-learning/>

Thank You