



A woman's face is visible through a semi-transparent red and orange gradient overlay. Her features are partially obscured by the color, with her eyes, nose, and mouth being the most prominent elements.

NYKAA

Your Feedback. Our Project.

INDEX

- Introduction to e-commerce
- About Nykaa
- Literature review
- Project details
- Objectives
- Hypothesis
- Methodology
- Analysis
- Results & Conclusion
- Annexure
- Limitation & Further Scope
- Summary
- References
- Acknowledgement

INTRODUCTION TO E-COMMERCE

The term electronic commerce refers to a business model where companies and individuals buy and sell goods and services over the Internet using desktop computers, laptops, smartphones etc.

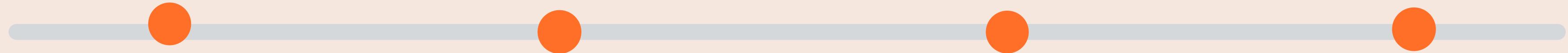
- Almost anything can be purchased through ecommerce today
- It can be a substitute for brick-and-mortar stores, though some businesses choose to maintain both

Top E-commerce companies in India - Amazon, Flipkart, Myntra etc.

Nykaa company is one of the top 15 e-commerce companies in India, standing at the 8th position.

NYKAA GROWTH STORY

Timeline



April 2012

Launched as an ecommerce portal curating a range of beauty and wellness products

2015

Company expanded from online-only to an omnichannel model and launched its collection of in-house beauty products via Nykaa Cosmetics

October 2020

launched Nykaa Man, India's first multi-brand ecommerce store for men's grooming.

December 2020

Nykaa Fashion launched its first store in Delhi, making the fashion business omnichannel

VISION

Bring inspiration and joy to people, everywhere, everyday.

MISSION

To create a world where our consumers have access to a finely curated, authentic assortment of products and services that delight and elevate the human spirit.



November 2, 2021 - Nykaa witnessed a blockbuster IPO with its shares being 81.78 times, raising ₹5,352 crore.

As per IANS-CVoter Issues, \$6.5 billion is the net worth of Falguni Nayar, founder of Nykaa. She owned a 53.5% stake in the company and became India's wealthiest self-made female billionaire.

Nykaa won as ‘The Most Innovative eCommerce Company’ at The eTales Awards 2015.

Nykaa also hosts beauty and fashion content via Nykaa TV, its YouTube channel, containing informational videos about beauty, cosmetics, and styling

Nykaa has adopted an inventory model. This ensures that all the bulk orders are received at once instead of one item at a time, which is the case with Flipkart or Amazon.

In 2020, Nykaa became the first Indian unicorn startup headed by a woman

HIGHLIGHTS

INTERNATIONAL OPERATIONS

Nykaa's fashion e-commerce store has launched international shipping to 13 countries such as USA, UK, Canada etc as the multi-brand business works to expand its distribution network.



LITERATURE REVIEW

Research papers related to our objective were studied and are stated below

01

Umesh Maiya and Mavy Miranda (2013) conducted a study on buying behavior towards cosmetics. The twin factors which everybody focused on were the Quality and Price

02

A. Goswami (2013) studied the customer satisfaction towards online shopping with special reference to teenagers of Jorhat Town. This research indicated that price factor should be given more importance

03

Nikhashemi (2013) found that customer perceived quality, perceived security and online payment process have significant positive effect on increasing customer satisfaction towards internet shopping

PARTICULARS OF THE DATA

- **DATA COLLECTION** - Primary data collected using an online questionnaire
- **VARIABLES INCLUDED** - Socio-demographic like Age, Gender, Occupation etc.
 - User of Nykaa? - Yes/No
 - If No, what is the reason?
 - If Yes, what is the level of satisfaction with various elements of Nykaa shopping app like Website Design, Product Variety, Payment method etc. ?
- **SAMPLE DESCRIPTION** - 92 respondents including both users and non-users of Nykaa
 - Female & Male respondents' age ranging from 18-30

Data on Excel

Name	Age	Gender	Are you a user of Nykaa?	Current Location	Occupation	Annual family income
Shivanshi	20	Female	Yes	Karnataka	Student	5-10 lakhs
Jinal	30	Female	Yes	Tamil Nadu	Employed	More than 15 lakhs
Jaishvin Kaur	21	Female	Yes	Karnataka	Student	5-10 lakhs
Mehak Agrawal	21	Female	Yes	Karnataka	Student	10-15 lakhs
Shruti	21	Female	Yes	Karnataka	Student	5-10 lakhs
Disha Zatakia	26	Female	Yes	Maharashtra	Employed	Less than 5 lakhs
Archi Jalan	20	Female	Yes	Jharkhand	Student	More than 15 lakhs
Shreya	20	Female	Yes	Karnataka	Student	More than 15 lakhs
Khushi Agarwal	21	Female	Yes	West Bengal	Student	5-10 lakhs
Apoorva Gupta	20	Female	Yes	Manipur	Student	Less than 5 lakhs
Shaily	21	Female	Yes	West Bengal	Student	5-10 lakhs
Samhita Rambhatla	20	Female	Yes	Karnataka	Student	More than 15 lakhs
Ruju	20	Female	Yes	Gujarat	Student	5-10 lakhs
Sadhika	19	Female	Yes	Karnataka	Student	5-10 lakhs
Bhumika	20	Female	Yes	Karnataka	Student	5-10 lakhs
Komal	21	Female	Yes	Maharashtra	Student	10-15 lakhs

Questionnaire details

Where, D1 = Website Design

D2 = Product Variety

D3 = Information Quality

D4 = Product Quality

D5 = Payment Method

D6 = Discounts & Offer

D7 = Security

D8 = Delivery Service

D9 = E-service Quality

CS = Overall Customer satisfaction

Response for each question was obtained in the form of a 5 point Likert scale

i.e., 1 = Strongly Disagree

2 = Disagree

3 = Neutral

4 = Agree

5 = Strongly Agree

Questionnaire details

- Each domain consisted of 3 questions
- Therefore, every respondent held a score which varied between 3 to 15 for each domain in the questionnaire. 3 - implying dissatisfaction and 15 - implying satisfaction towards that domain

OBJECTIVES

- 01 To study factors influencing customer satisfaction towards Nykaa
- 02 To study the relationship between website characteristics and customer satisfaction
- 03 To know the influencing factors among socio-demographic variables on customer satisfaction

HYPOTHESIS

Principal Component Analysis

H_{01} : Correlation coefficients are not statistically significant

Vs

H_{11} : Correlation coefficients are statistically significant

Bartlett's Test

H_{02} : The model is not statistically significant.

Vs

H_{12} : The model is statistically significant.

HYPOTHESIS

Multiple Linear Regression

H_{03} : The coefficient of independent variables is zero, i.e. $\beta=0$

Vs

H_{13} : The coefficient of independent variables is not zero, i.e. $\beta \neq 0$

METHODOLOGY

- **DESCRIPTIVE STATISTICS (USERS)** - Crosstabs
 - Clustered Bar chart
 - Scatterplot
- **DIMENSION REDUCTION (DOMAINS)** - Principal Component Analysis
- **CORRELATION (DOMAINS & SATISFACTION)** - Multiple Linear Regression
- **CORRELATION (SOCIO-DEMOGRAPHIC)** - Classification Decision Trees
 - (using Classification and Regression Trees model)
- **DESCRIPTIVE STATISTICS (NON-USERS)** - Crosstabs
 - Clustered Bar charts

ANALYSIS

DESCRIPTIVE STATISTICS (USERS) :

- Association between "Time spent on Nykaa's website" and "Occupation" -

Average daily hours spent on Nykaa

Occupation

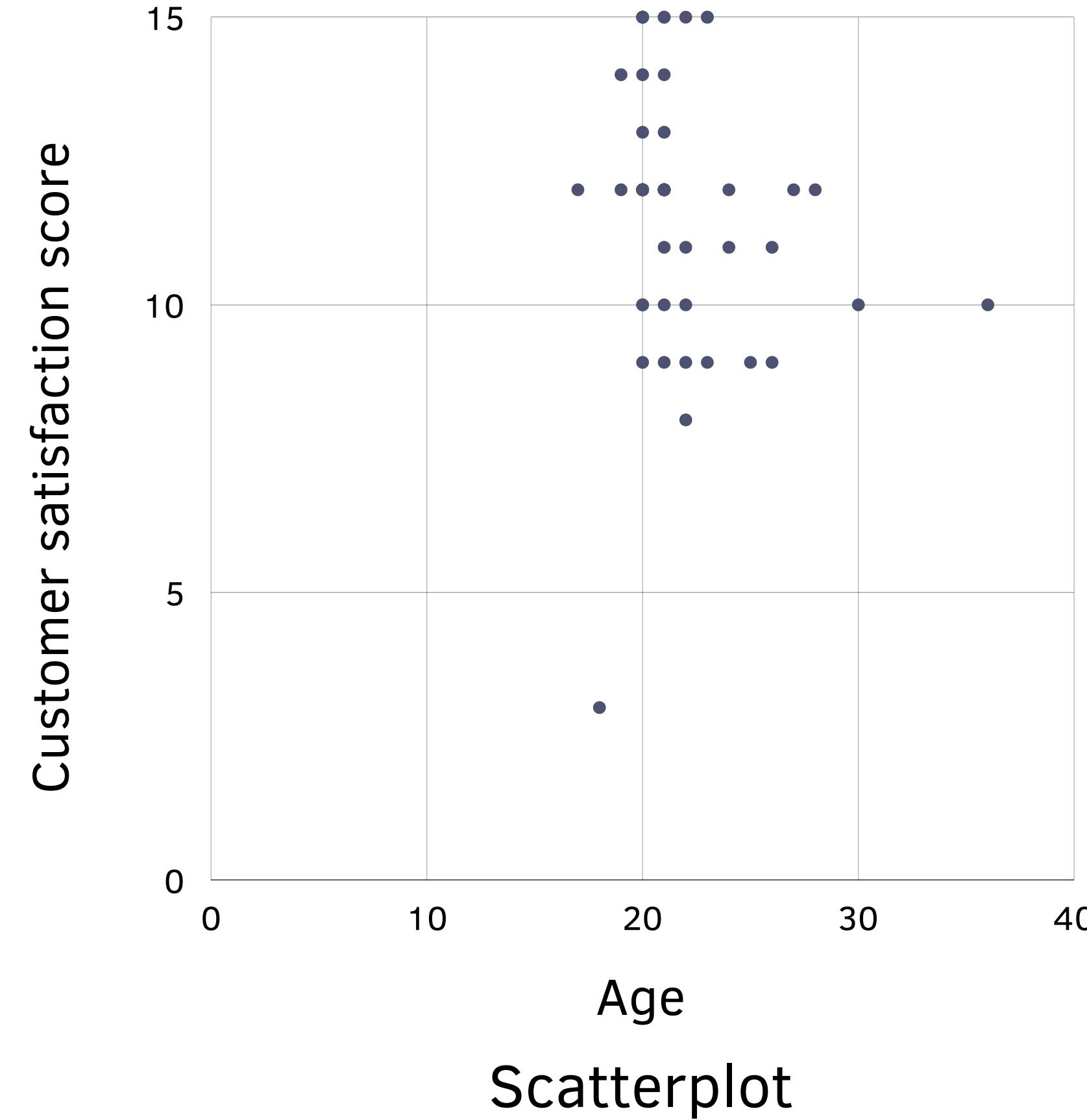
	Less than 1 hour	1-2 hours	More than 2 hours	Total
Student	36 (92.31%)	2 (5.13%)	1 (2.56%)	39
Employed	8 (80%)	1 (10%)	1 (10%)	10
Self-employed	2 (50%)	2 (50%)	0 (0%)	4
Total	46	5	2	53

Crosstabulation

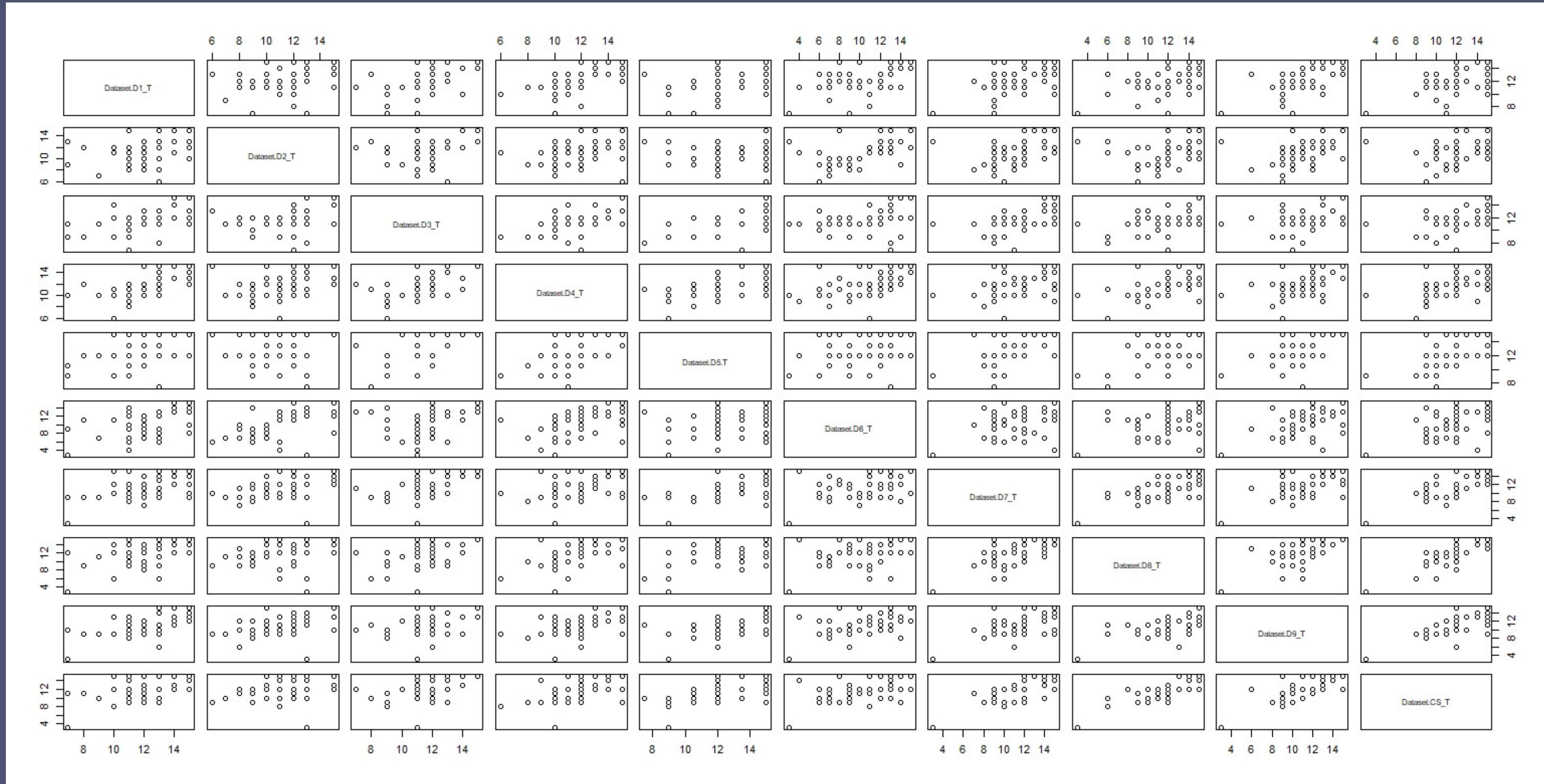


Clustered bar chart

- Correlation between "Age" and "Satisfaction score" -



- Pair plot of (D1, D2,..., D9) and Customer satisfaction -



DIMENSION REDUCTION - PCA :

- Reading data into R -

```
library(psych)
```

```
x <- read.csv("C:/Users/HP/Downloads/Mastersheet.csv")
```

```
y <- x[,12:20]
```

- Eigen values of 9 factors -

```
[1] 4.6550078 1.1471199 0.7602726 0.6451495 0.5063151 0.4384801 0.4078643
```

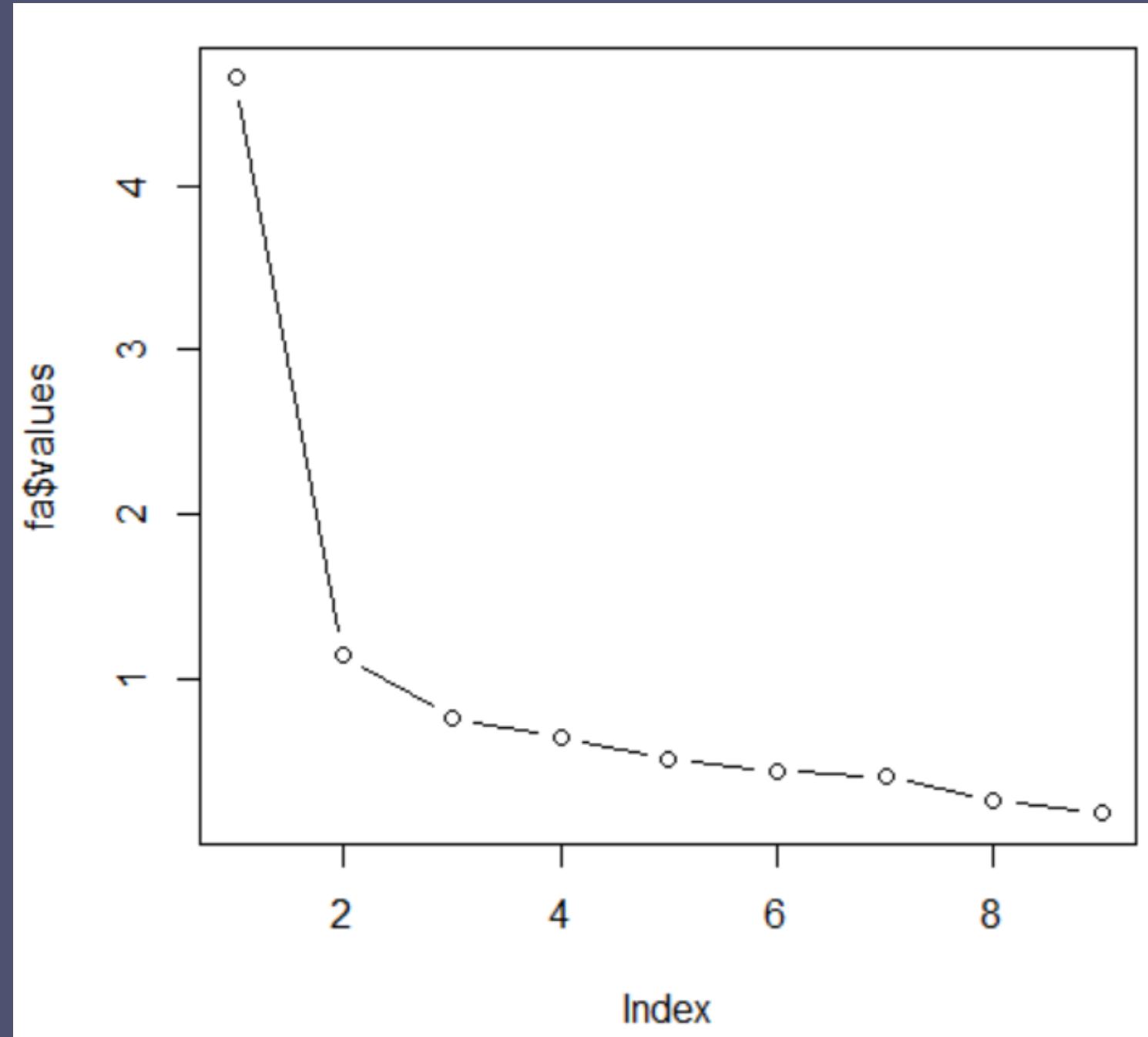
```
[8] 0.2600518 0.1797389
```

- KMO and Bartlett's Test

Bartlett's Test of Sphericity Approx. Chi-Square 216.853

df 36 Sig. .0001

- Factor Loadings of 2 components -
- Scree plot -



	PC1	PC2
SS loadings	4.655	1.147
Proportion Var	0.517	0.127
Cumulative Var	0.517	0.645

- Factor Loadings of 2 rotated components -

	RC1	RC2
Website design	0.701	0.296
Product variety	0.154	0.851
Information quality	0.730	0.122
Product quality	0.725	0.284
Payment method	0.515	-0.029
Discounts & Offers	0.204	0.838
Security	0.804	0.338
Delivery service	0.786	0.214
E-service quality	0.770	0.437

	RC1	RC2
SS loadings	3.841	1.961
Proportion Var	0.427	0.218
Cumulative Var	0.427	0.645

- Communalities -

	RC1	RC2	h ²	u ²	com
D1_T	0.70	0.30	0.58	0.42	1.3
D2_T	0.15	0.85	0.75	0.25	1.1
D3_T	0.73	0.12	0.55	0.45	1.1
D4_T	0.72	0.28	0.61	0.39	1.3
D5_T	0.51	-0.03	0.66	0.34	1.0
D6_T	0.20	0.84	0.74	0.26	1.1
D7_T	0.80	0.34	0.61	0.39	1.4
D8_T	0.79	0.21	0.66	0.34	1.1
D9_T	0.77	0.44	0.64	0.36	1.7

	RC1	RC2
SS loadings		3.84
Proportion Var		0.43
Cumulative Var		0.43
Proportion Explained	0.66	0.34
Cumulative Proportion	0.66	1.00

Therefore, from principal component analysis we obtain 5 domains D2, D6, D7, D8 and D9.

Where,

D2 = Product Variety

D6 = Discounts & Offers

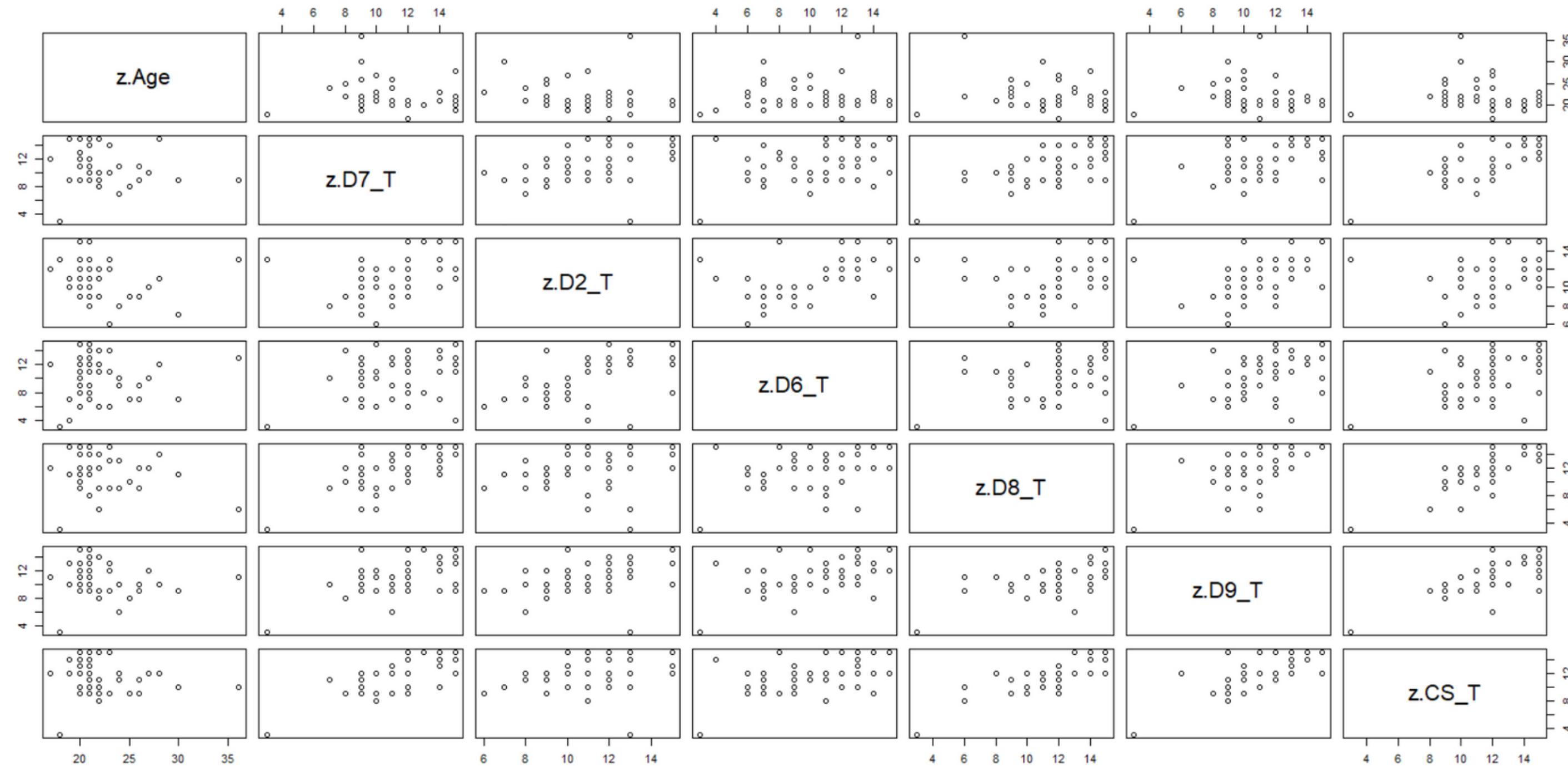
D7 = Security

D8 = Delivery Service

D9 = E - Service Quality

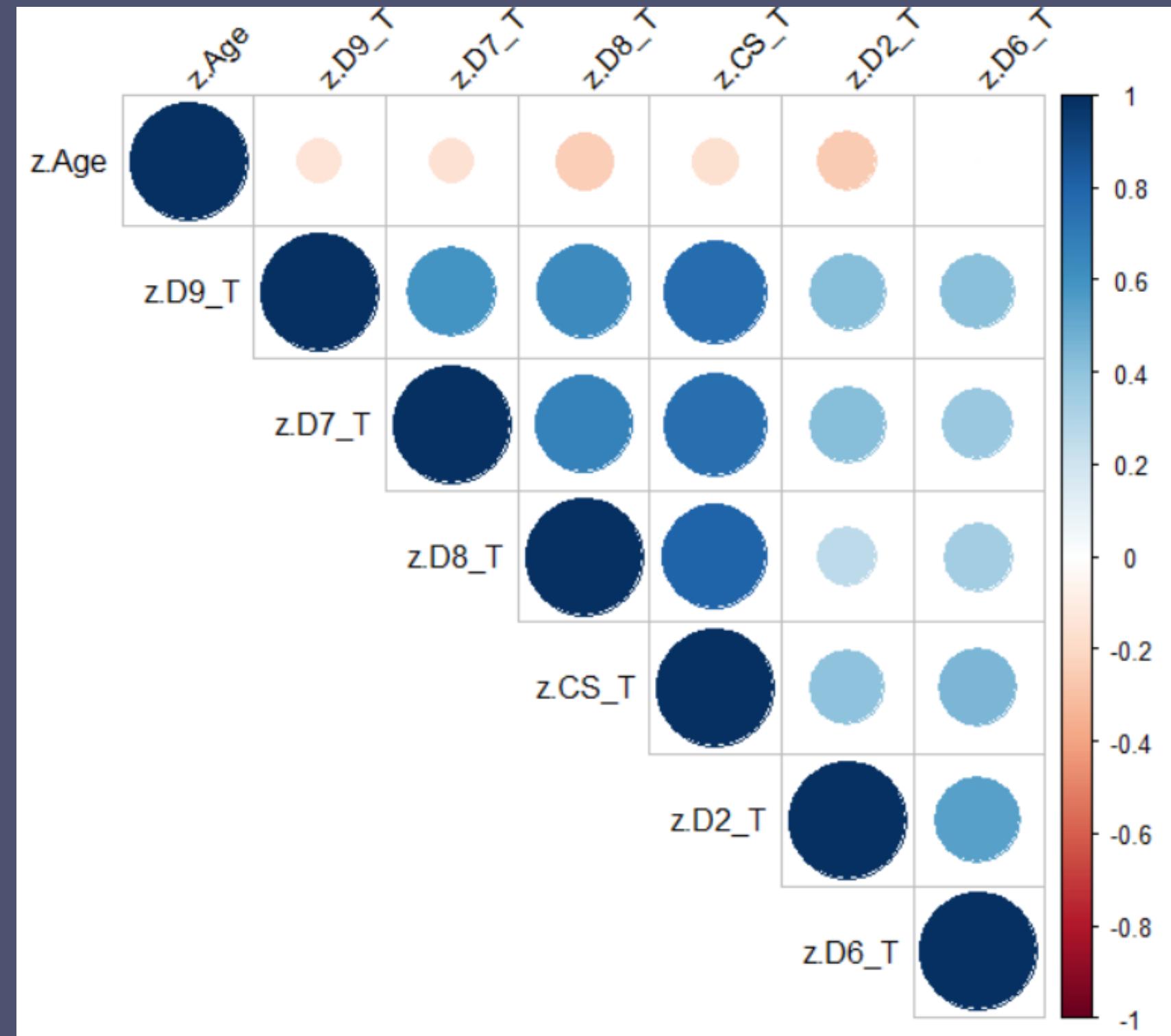
CORRELATION BETWEEN (Age, D2, D6, D7, D8, D9) & CS - MLR:

- Assumptions (Linearity) -



Pair plot

- Linearity -



Corrplot

- Multicollinearity -

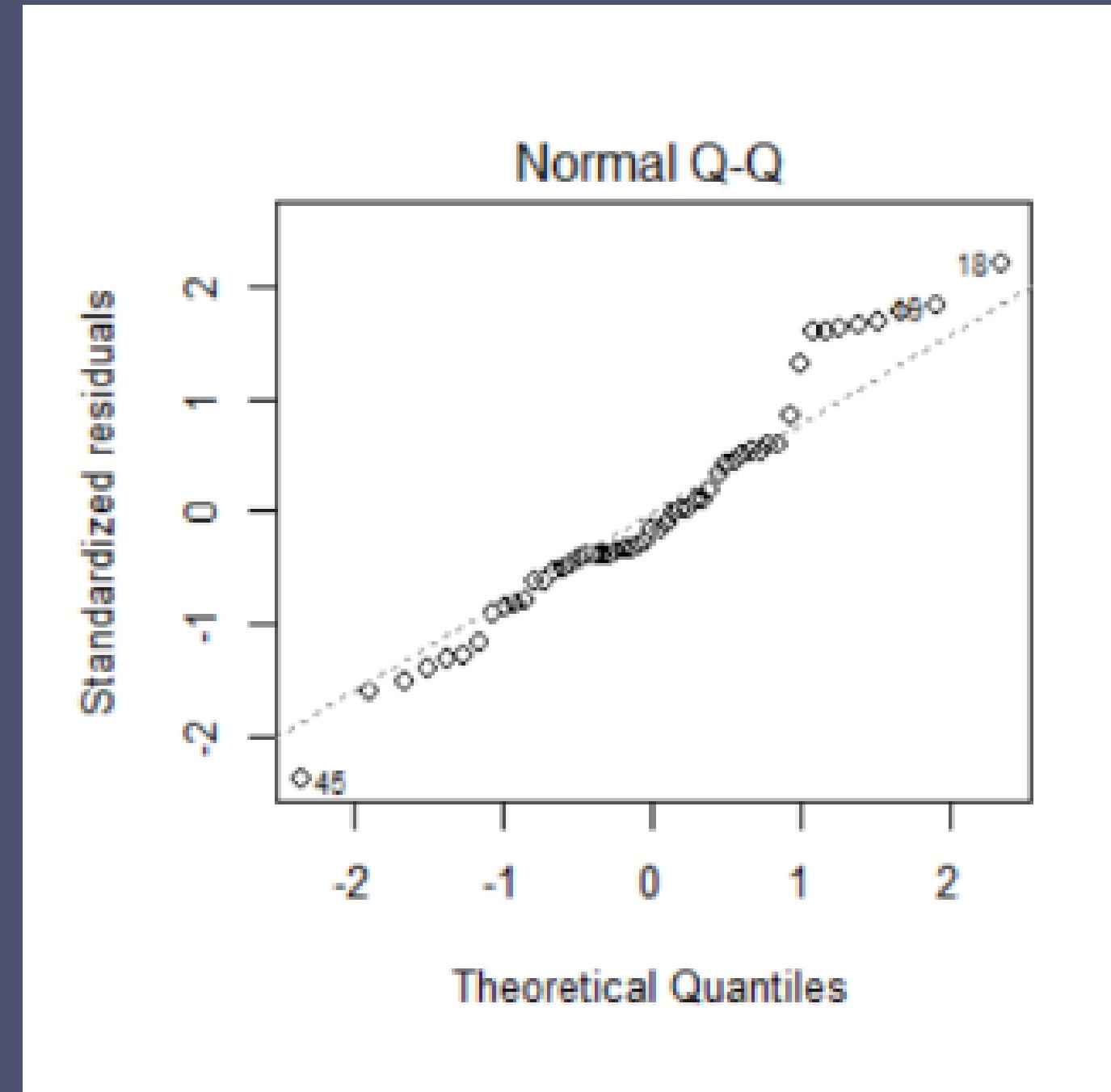
	z.Age	z.D7_T	z.D2_T	z.D6_T	z.D8_T	z.D9_T	z.CS_T
z.Age	1.00	-0.15	-0.26	0.0036	-0.25	-0.14	-0.16
z.D7_T	-0.15	1.00	0.43	0.37	0.68	0.59	0.76
z.D2_T	-0.26	0.43	1.00	0.54	0.26	0.43	0.40
z.D6_T	0.0036	0.37	0.54	1.00	0.35	0.41	0.45
z.D8_T	-0.25	0.68	0.26	0.35	1.00	0.63	0.80
z.D9_T	-0.14	0.59	0.43	0.41	0.63	1.00	0.77
z.CS_T	-0.16	0.76	0.40	0.45	0.80	0.77	1.00

Correlation Matrix

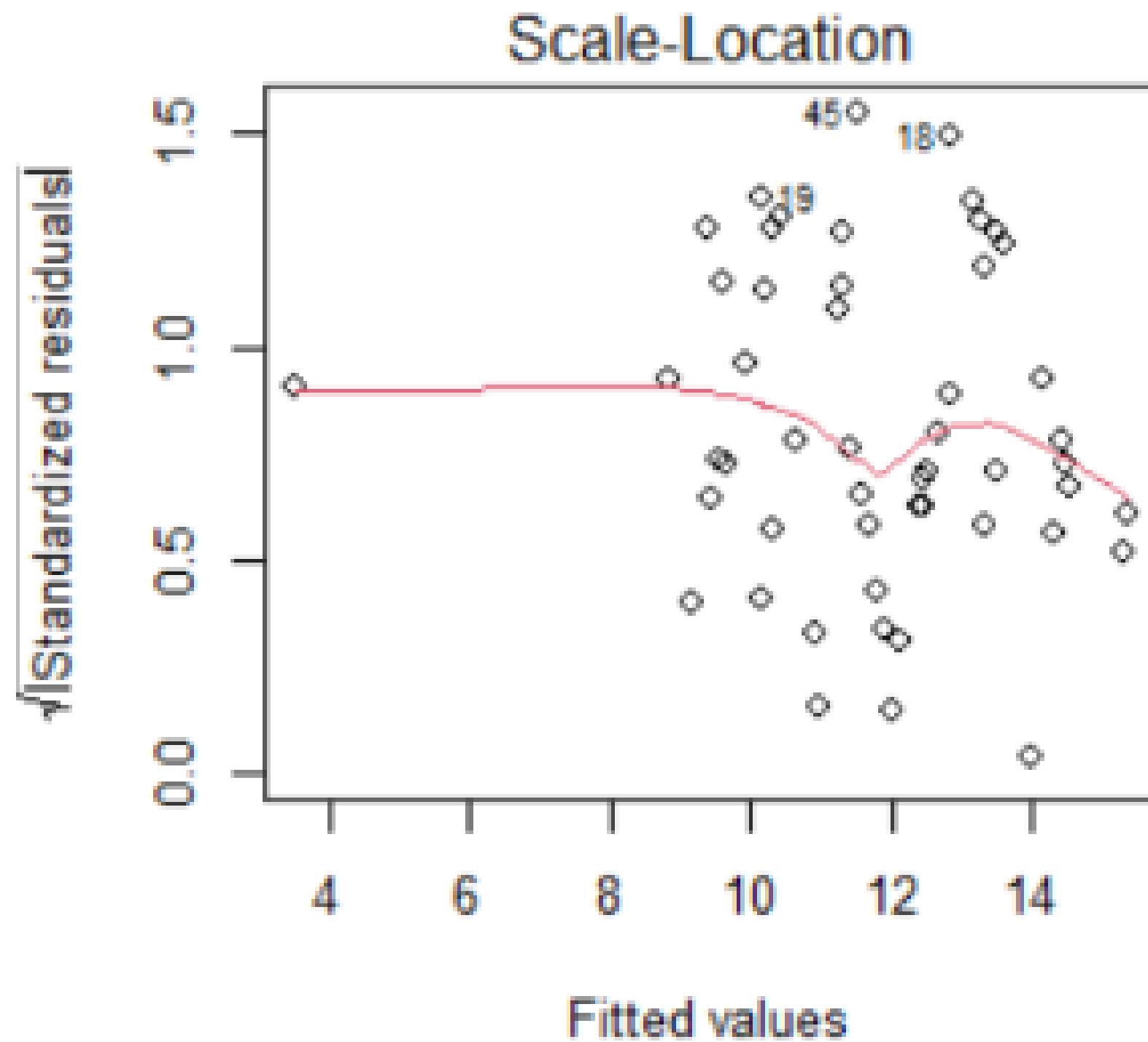
- Normality of residuals -

```
data: model.diag.metrics$resid  
W = 0.95784, p-value = 0.059
```

Shapiro-Wilks test
(Normality of Residuals)



- Heteroscedasticity -



- Multiple Linear Regression -

Residuals:

	Min	1Q	Median	3Q	Max
	-2.5011	-0.5303	-0.1636	0.5282	2.1843

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.143024	1.736096	0.082	0.934700
Age	0.018536	0.052025	0.356	0.723248
Security	0.254496	0.091475	2.782	0.007805 **
Product variety	0.008814	0.097615	0.090	0.928448
Discounts & offers	0.061221	0.066942	0.915	0.365201
Delivery service	0.341917	0.089567	3.817	0.000401 ***
E-service quality	0.324276	0.090570	3.580	0.000823 ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

- R squared and p value -

Residual standard error: 1.077 on 46 degrees of freedom

Multiple R-squared: 0.8035, Adjusted R-squared: 0.7779

F-statistic: 31.35 on 6 and 46 DF, p-value: 1.107e-14

Therefore, from multiple linear regression we obtain 3 domains which significantly influence customer satisfaction which are Security, Delivery Service and E - Service Quality

CLASSIFICATION DECISION TREE

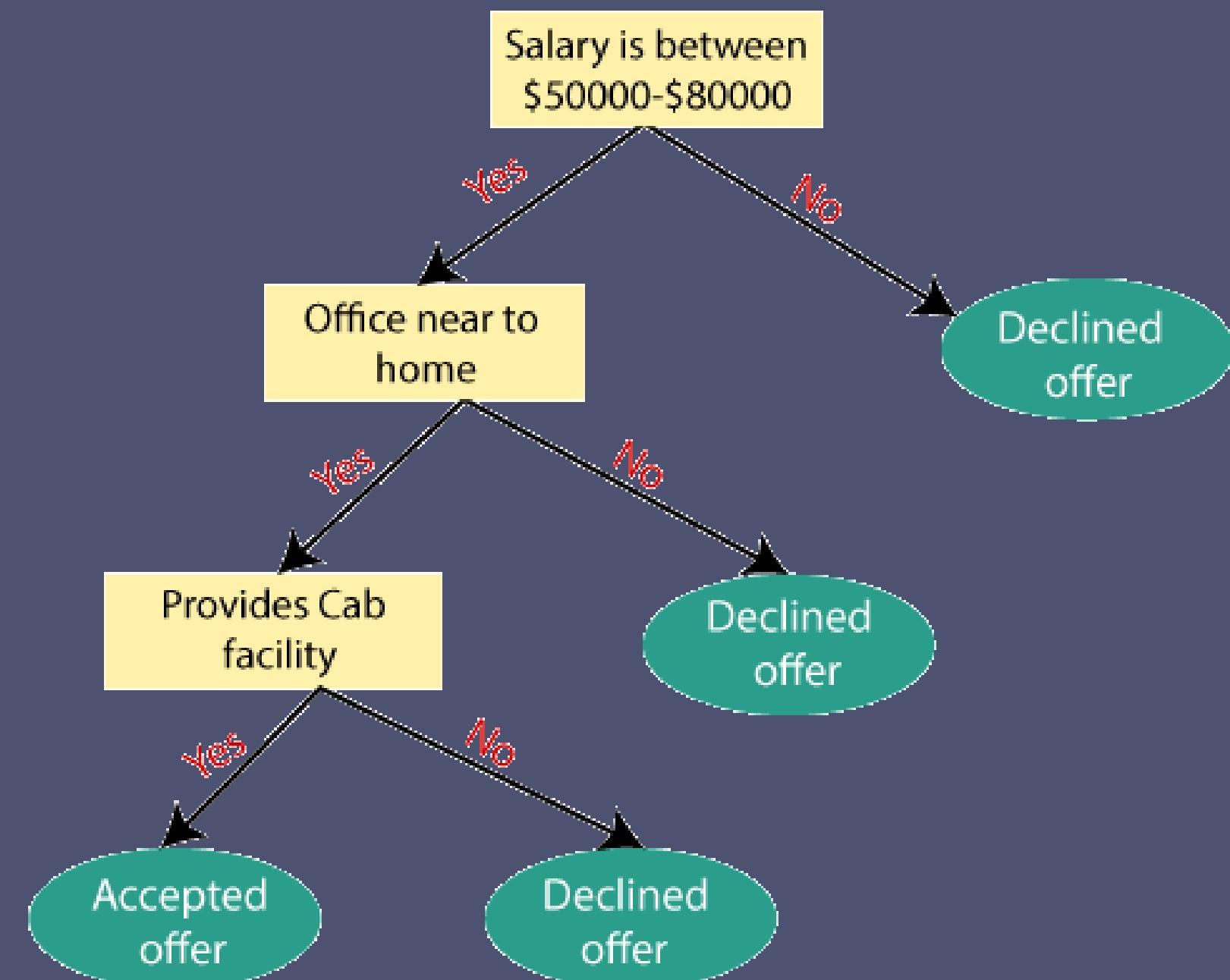
DTs are a **non-parametric supervised learning method** used for classification and regression. It creates a model that predicts the value of a target variable by learning simple decision rules inferred from the data features.

Classification and Regression Trees or CART is a term introduced by Leo Breiman to refer to Decision Tree algorithms used for classification or regression predictive modeling problems.

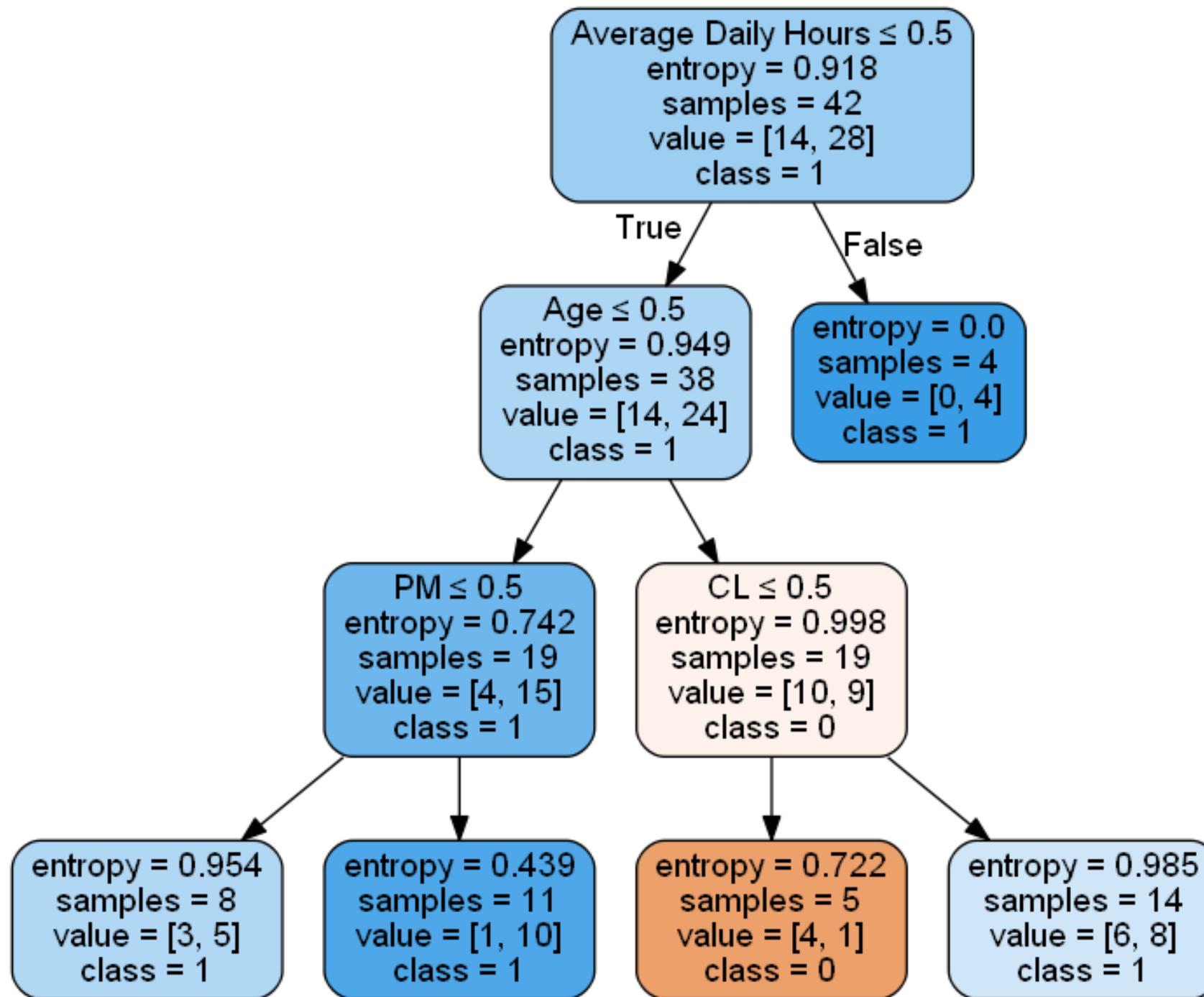
The representation for the CART model is a **binary tree**. Each root node represents a single input variable (x) and a split point on that variable (assuming the variable is numeric). The leaf nodes of the tree contain an output variable (y) which is used to make a prediction.

Given below is an example of a binary decision tree -

Suppose, there is a candidate who has a job offer and wants to decide whether he should accept the offer or not.



CLASSIFICATION DECISION TREE - CART METHOD :



- Average daily hours spent on Nykaa :
 $< 1 \text{ hr} = 0 ; \geq 1 \text{ hr} = 1$
- Age :
 $< 21 = 0 ; \geq 21 = 1$
- Payment method :
COD = 0 ; Online payment = 1
- Current location :
North/East = 0 ; South/West = 1
- Customer satisfaction :
Dissatisfied = 0 ; Satisfied = 1

DESCRIPTIVE STATISTICS (NON - USERS) :

- Association of "Nykaa Usage" & "Gender" -

		Gender		
		Female	Male	Total
Nykaa Usage	User	45 (84.91%)	8 (15.09%)	53
	Non-user	17 (43.59%)	22 (56.41%)	39
	Total	62	30	92

Crosstabulation

- Representation of "Other shopping websites used" & "Gender" -



Bar chart

- Representation of "Reasons behind not using Nykaa" & "Gender" -

Reasons	Males	Females	Total
Not interested in beauty products	7 (50%)	7 (50%)	14
Never had the need to buy	10 (71.43%)	4 (28.57%)	14
Prefer other online shopping platforms	2 (50%)	2 (50%)	4
Would rather go to the store and buy than shopping online	1 (25%)	3 (75%)	4
Not aware of the products available	2 (66.67%)	1 (33.33%)	3
Total	22	17	39

RESULTS

Principle Component Analysis

Here, we reject the null hypothesis at 5% level of significance,
i.e., The correlation coefficients are statistically significant

Bartlett's Test

Here, we reject the null hypothesis at 5% level of significance,
i.e., The model is statistically significant

Multiple Linear Regression

Here, we reject the null hypothesis at 5% level of significance,
i.e., The coefficient of the independent variables D7, D8 and D9 is not equal to 0

CONCLUSION

- 1) Out of the 9 domains considered, Product variety, Discounts & offers, Security, Delivery service and E-service quality are the domains which have the highest influence on customer satisfaction
- 2) Security, Delivery service and E-service quality are the three online shopping features of Nykaa that influence customer satisfaction significantly
- 3) From the decision tree we obtain that people spending more than 1 hr on Nykaa are satisfied with the online shopping app. As per our sample size, people residing in the North East region are not satisfied whereas people from South West are satisfied with Nykaa

ANNEXURE

Pair plot of (D1, D2,.., D9) and Customer satisfaction -

- R code :

```
Data=x[,12:21]  
plot(Data)
```

Dimension Reduction - PCA -

- R Code :

```
library(psych)
x <- read.csv("C:/Users/HP/Downloads/Mastersheet.csv")
y <- x[,12:20]

fa<- principal(y,nfactors = 9,rotate = "none",covar = FALSE)
fa$values
plot(fa$values, type = "b")
fa1<- principal(y,nfactors = 2,rotate = "none",covar = FALSE)
fa1$loadings
fa2 <- principal(y,nfactors = 2,rotate = "varimax",covar = FALSE)
fa2$loadings
fa2
```

Multiple Linear Regression -

- R Code :

```
z<-Nykaa_users[,c(2,12:21)]  
z  
mlr<-lm(CS_T~Age+D7_T+D2_T+D6_T+D8_T+D9_T,data=z)  
summary(mlr)  
par(mfrow = c(2, 2))  
plot(mlr)  
library(broom)  
model.diag.metrics <- augment(mlr)  
head(model.diag.metrics)  
shapiro.test(model.diag.metrics$resid)
```

Classification Decision Tree -

- Python Code :

```
import pandas as pd
from sklearn.tree import DecisionTreeClassifier # Import Decision Tree Classifier
from sklearn.model_selection import train_test_split # Import train_test_split function
from sklearn import metrics #Import scikit-learn metrics module for accuracy calculation
from sklearn.preprocessing import LabelEncoder
data = pd.read_csv("finaltreedata.csv")
data.head()
def object_cols(data):
    return list(data.select_dtypes(include='object').columns)
obj_col = object_cols(data)
obj_col
le = LabelEncoder()
```

```
for col in obj_col:  
    data[col] = le.fit_transform(data[col])  
feature_cols = ['Age', 'CL', 'Occupation', 'Annual family income','Average Daily Hours','PM']  
x = data[feature_cols] # Features  
y = data.CS_Levels # Target variable  
X_train, X_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=1)  
# Create Decision Tree classifier object  
clf = DecisionTreeClassifier()  
# Train Decision Tree Classifier  
clf = clf.fit(X_train,y_train)  
#Predict the response for test dataset  
y_pred = clf.predict(X_test)  
print("Accuracy:",metrics.accuracy_score(y_test, y_pred))  
dot_data = StringIO()
```

```
export_graphviz(clf, out_file=dot_data,
    filled=True, rounded=True,
    special_characters=True, feature_names = feature_cols,class_names=['0','1'])
graph = pydotplus.graph_from_dot_data(dot_data.getvalue())
graph.write_png('decisiontree.png')
Image(graph.create_png())
# Create Decision Tree classifier object
clf = DecisionTreeClassifier(criterion="entropy", max_depth=3)
# Train Decision Tree Classifier
clf = clf.fit(X_train,y_train)
#Predict the response for test dataset
y_pred = clf.predict(X_test)
# Model Accuracy, how often is the classifier correct?
print("Accuracy:",metrics.accuracy_score(y_test, y_pred))
dot_data = StringIO()
export_graphviz(clf, out_file=dot_data,
    filled=True, rounded=True,
    special_characters=True, feature_names = feature_cols,class_names=['0','1'])
graph = pydotplus.graph_from_dot_data(dot_data.getvalue())
graph.write_png('decisiontree.png')
Image(graph.create_png())
```

LIMITATION

- 1) Number of people participated in our survey is 92 respondents. Taking time constraint as an important factor in mind
- 2) Since we used a 5 point Likert Scale, obtaining the middle value (3 - Neutral) as a response usually does not help in any kind of analysis
- 3) Our domain knowledge for e-commerce platforms could be limited. That is, there might be other factors that influence a customer's satisfaction towards the platform

SCOPE

- 1) Nykaa can market its products to men through advertising, which appear to be not aware of the products offered on Nykaa
- 2) This study can also be expanded to the offline shopping experience of Nykaa's customers
- 3) The study can be extended further to conduct a comparative analysis between Nykaa and its competitors Myntra and Sephora.

SUMMARY

We collected data from users and non-users of Nykaa using a main data gathering strategy in this study.

To gain a thorough grasp of the variables impacting Nykaa's customer satisfaction, we used descriptive statistics, principal component analysis, multiple linear regression, and classification decision trees.

We were also able to analyze the behavior of non-users of Nykaa by using exploratory data analysis.

REFERENCES

ABOUT NYKAA

Retrieved from -

<https://www.nykaa.com/>

<https://startuptalky.com/nykaa-success-story/>

QUESTIONNAIRE STRUCTURE

Retrieved from -

<http://www.isihome.ir/freearticle/ISIHome.ir-22032.pdf>

http://mdu.ac.in/UpFiles/UpPdfFiles/2021/Jun/2_06-17-2021_11-22-44_Chapter-5.pdf

STATISTICAL MODELS

Retrieved from -

<http://www.isihome.ir/freearticle/ISIHome.ir-22032.pdf>

ACKNOWLEDGEMENT

We would like to take this opportunity to thank our Mentor **Dr. Santosh C D** for his humble guidance and time to time involvement throughout without which it would not be possible to compile the project.

We would also like to thank **Prof. Sukhesh Kumar** and **Dr. Roudra Chakraborty** for their suggestions and critical assessment which helped us improve on our short comings in our project and presentation.

We are highly grateful to all the people who participated in the survey and generously provided us the data for our project.

THANK YOU !

Happy Cart. Happy You.

Mentored by - Dr. Santosh C D

Prepared by : Mahima Doshi - 75251219004, Mohita Ahuja - 75251219002
Rishi R - 75251219014, Suhasi Gohil - 75251219005