

Project Report

# Customer Retention

## TABLE OF CONTENT

Sr No	Content	Page Number
1	Introduction	3
2	Discussion Context/Background	4
3	EDA – Exploratory Data Analysis	5 - 22
4	Conclusion	23

# Introduction

**Customer retention** refers to the ability of a company or product to retain its customers over some specified period. High customer retention means customers of the product or business tend to return to, continue to buy or in some other way not defect to another product or business, or to non-use entirely. Selling organizations generally attempt to reduce customer defections. Customer retention starts with the first contact an organization has with a customer and continues throughout the entire lifetime of a relationship and successful retention efforts take this entire lifecycle into account. A company's ability to attract and retain new customers is related not only to its product or services, but also to the way it services its existing customers, the value the customers actually perceive as a result of utilizing the solutions, and the reputation it creates within and across the marketplace.

## **What is E commerce?**

E-commerce (electronic commerce) is the activity of electronically buying or selling of products on online services or over the Internet. E-commerce draws on technologies such as mobile commerce, electronic funds transfer, supply chain management, Internet marketing, online transaction processing, electronic data interchange (EDI), inventory management systems, and automated data collection systems. E-commerce is in turn driven by the technological advances of the semiconductor industry, and is the largest sector of the electronics industry.

## Context/Background of the Project

Customer satisfaction has emerged as one of the most important factors that guarantee the success of online store; it has been posited as a key stimulant of purchase, repurchase intentions and customer loyalty. A comprehensive review of the literature, theories and models have been carried out to propose the models for customer activation and customer retention. Five major factors that contributed to the success of an e-commerce store have been identified as: service quality, system quality, information quality, trust and net benefit. The research furthermore investigated the factors that influence the online customers repeat purchase intention. The combination of both utilitarian value and hedonistic values are needed to affect the repeat purchase intention (loyalty) positively. The data is collected from the Indian online shoppers. Results indicate the e-retail success factors, which are very much critical for customer satisfaction. I have followed below process.

1. Getting data from the source
2. Correction and cleaning of data.
3. Converting data to Data Frame
4. Used EDA to analyse the data
5. Data Visualization
6. Converting data to Graphs and tables

## Libraries and Tools Used for the Project

1. **Python (Jupyter Notebook)**
2. **Numpy**
3. **Pandas**
4. **Seaborn**
5. **Matplotlib**
6. **Warnings**

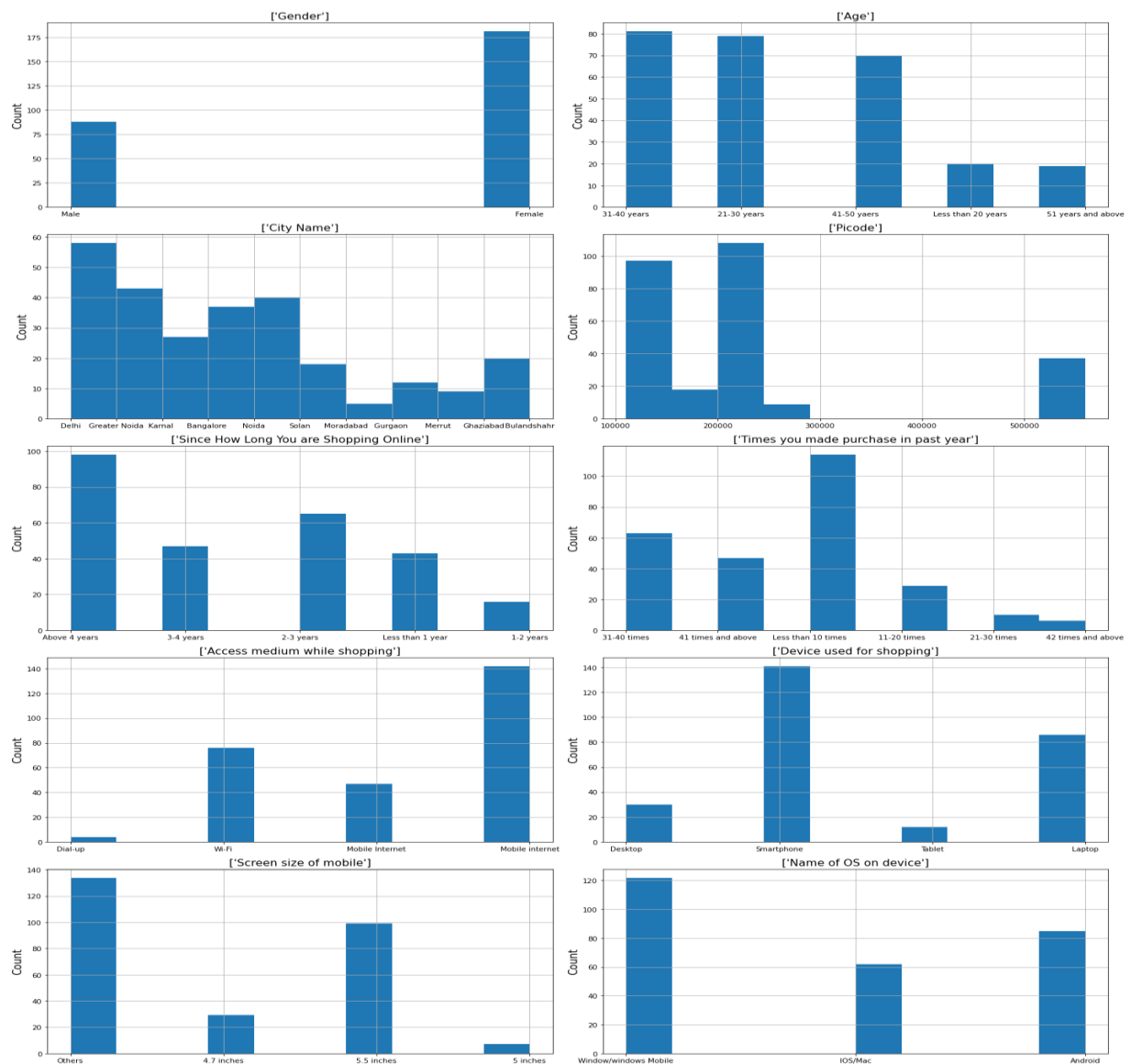
# Exploratory Data Analysis

Exploratory data analysis (EDA) is used by data scientists to analyze and investigate data sets and summarize their main characteristics, often employing data visualization methods. It helps determine how best to manipulate data sources to get the answers you need, making it easier for data scientists to discover patterns, spot anomalies, test a hypothesis, or check assumptions.

EDA is primarily used to see what data can reveal beyond the formal modeling or hypothesis testing task and provides a better understanding of data set variables and the relationships between them. It can also help determine if the statistical techniques you are considering for data analysis are appropriate. Originally developed by American mathematician John Tukey in the 1970s, EDA techniques continue to be a widely used method in the data discovery process today. I have done below analysis on this project.

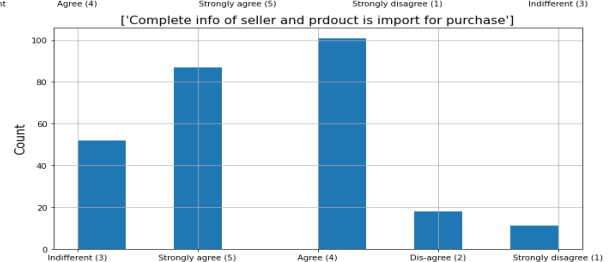
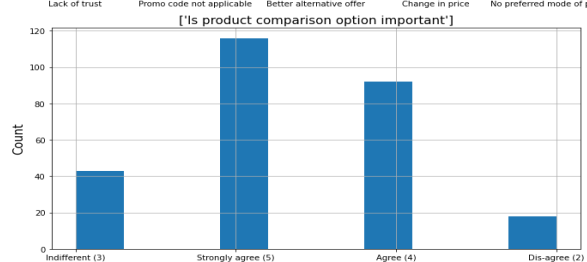
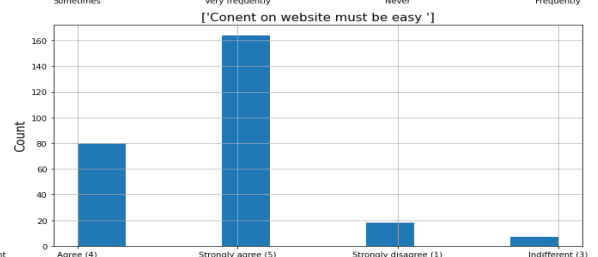
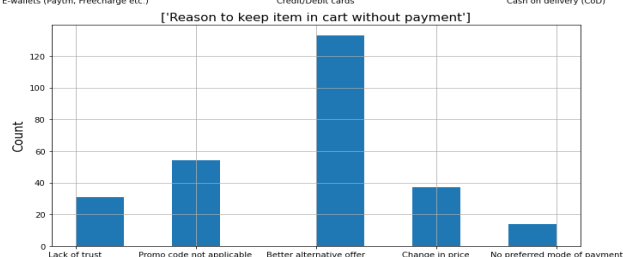
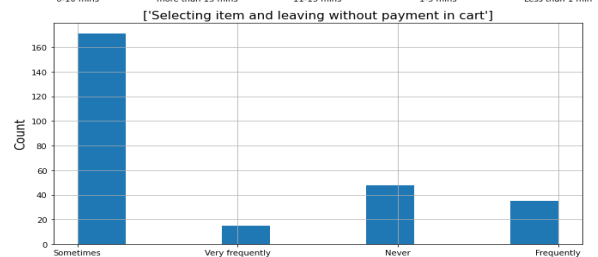
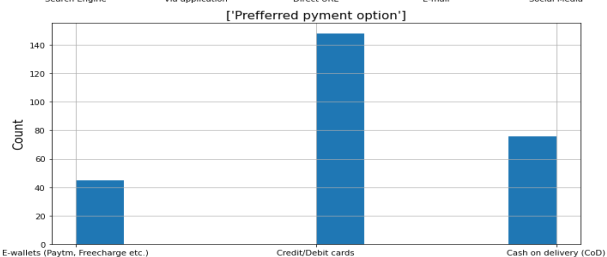
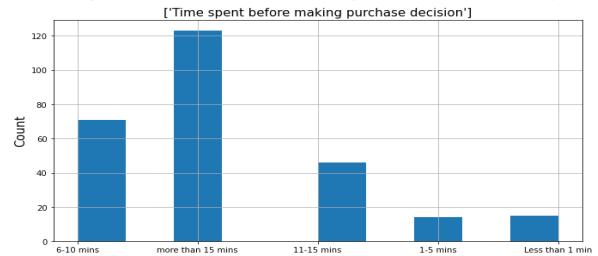
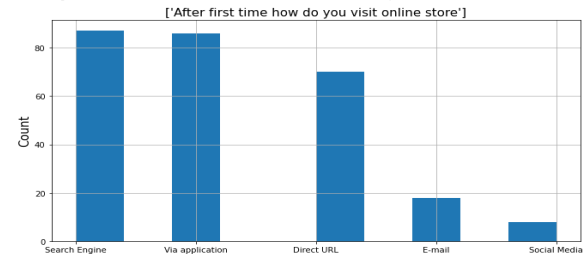
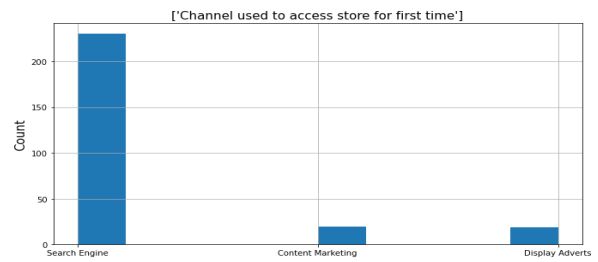
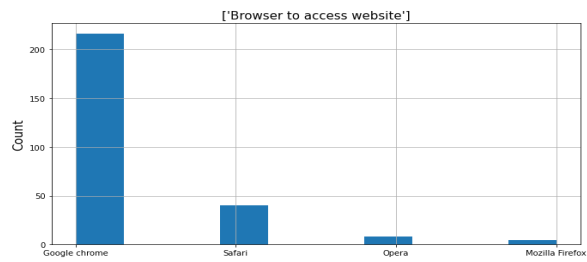
1. Column wise data analysis
2. Comparison with target column
3. Correlation of dataset set using heatmap
4. Correlation of all features with target column (bar)

## Column wise data analysis



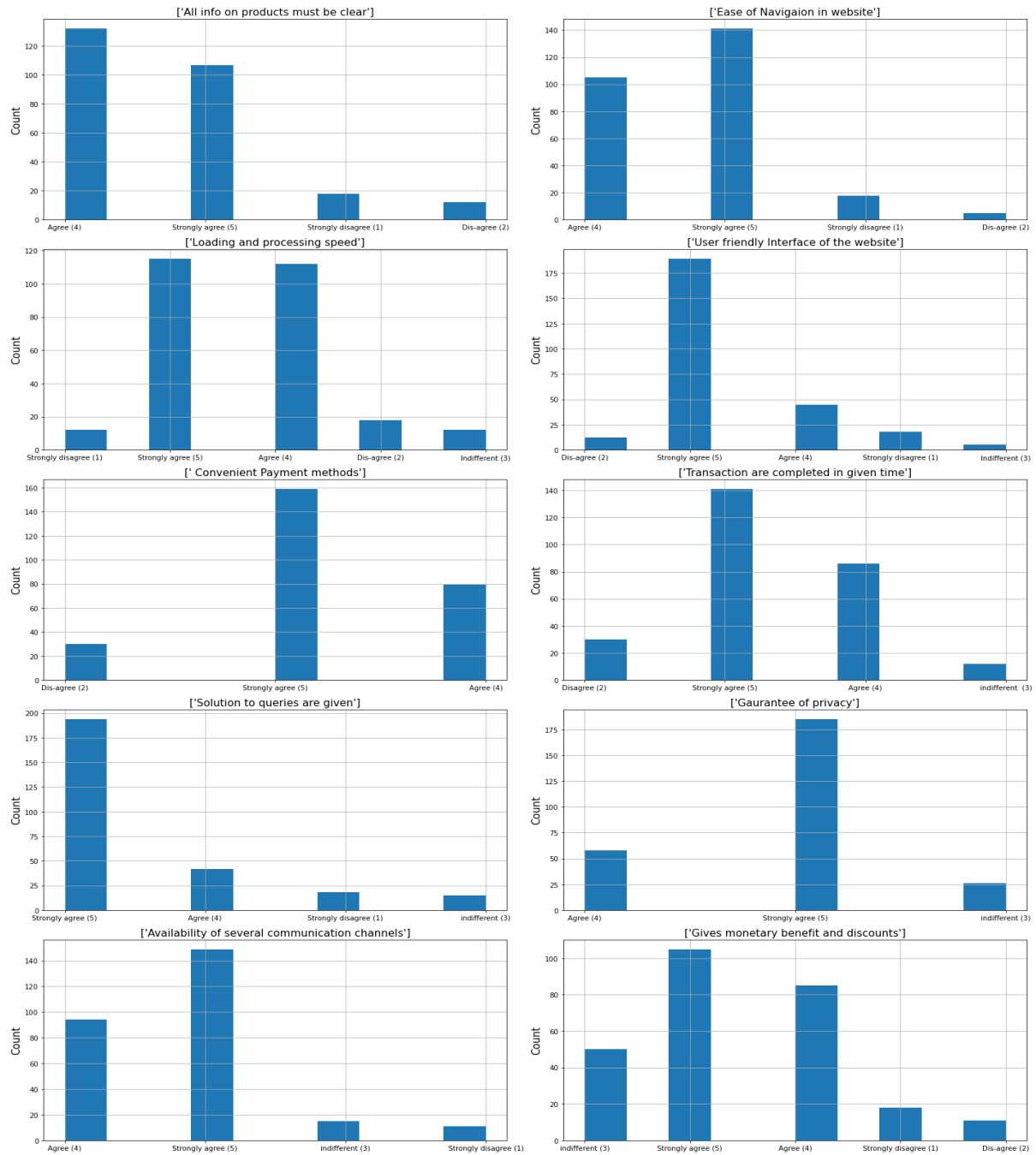
## Observations

1. Majority of the age group is between 21-30 years and 31-40 years, there is little difference in this age category. In which female customers are more than male customers
2. Observed that most of the online shopping is done in Delhi
3. Most of the people are shopping since last 4 years
4. Mainly shopping is done via mobile using mobile internet.



## Observation

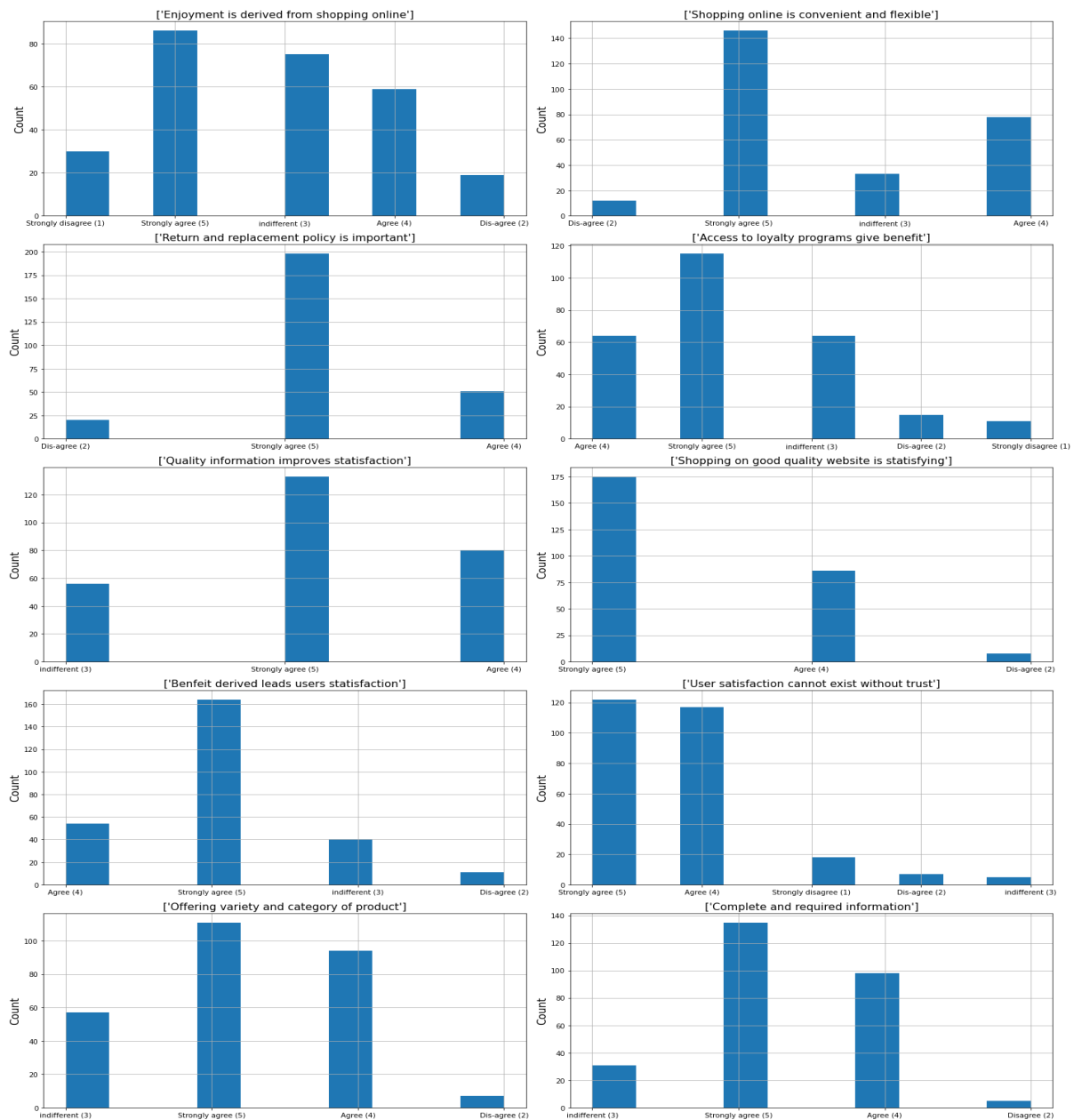
1. It's been observed that first time customers visit these sites via search engine and mainly from Google Chrome, other web browsers are no way near.
2. From next time most of the people get tuned to application, however same amount of people also prefers using search engine to visit the site.
3. On an average approx. more than 15 mins are spent by the customers on these sites.
4. Most payments are done by credit cards followed by COD and E-wallets respectively.



## Observation

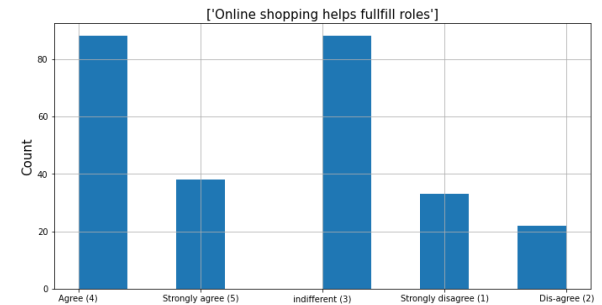
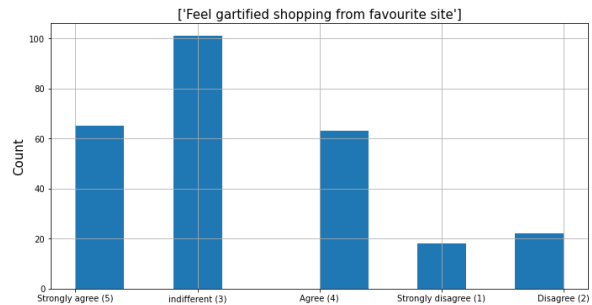
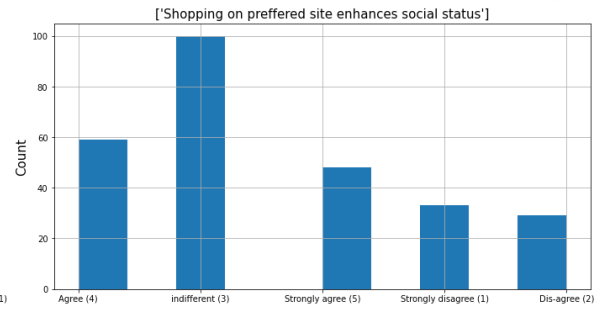
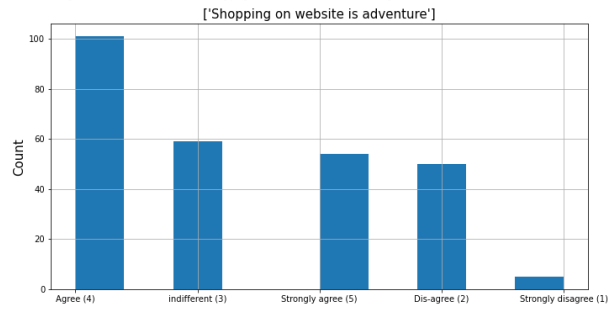
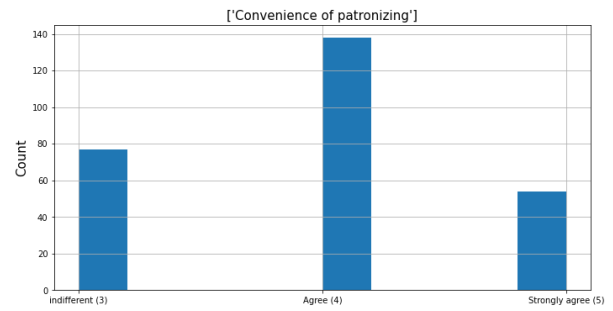
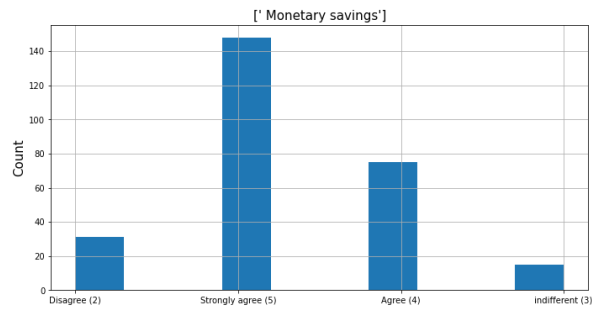
1. Mostly people have agreed that these websites give ease to navigate the product on website and are user friendly.
2. There is ease in make payments as payment mentioned are very convenient.
3. Also, solution to any query is given in specified time, customers also agree upon that their privacy is safe.
4. Due to available of several communication channels it becomes easy to communicate which helps in smoothing customer journey.





## Observations

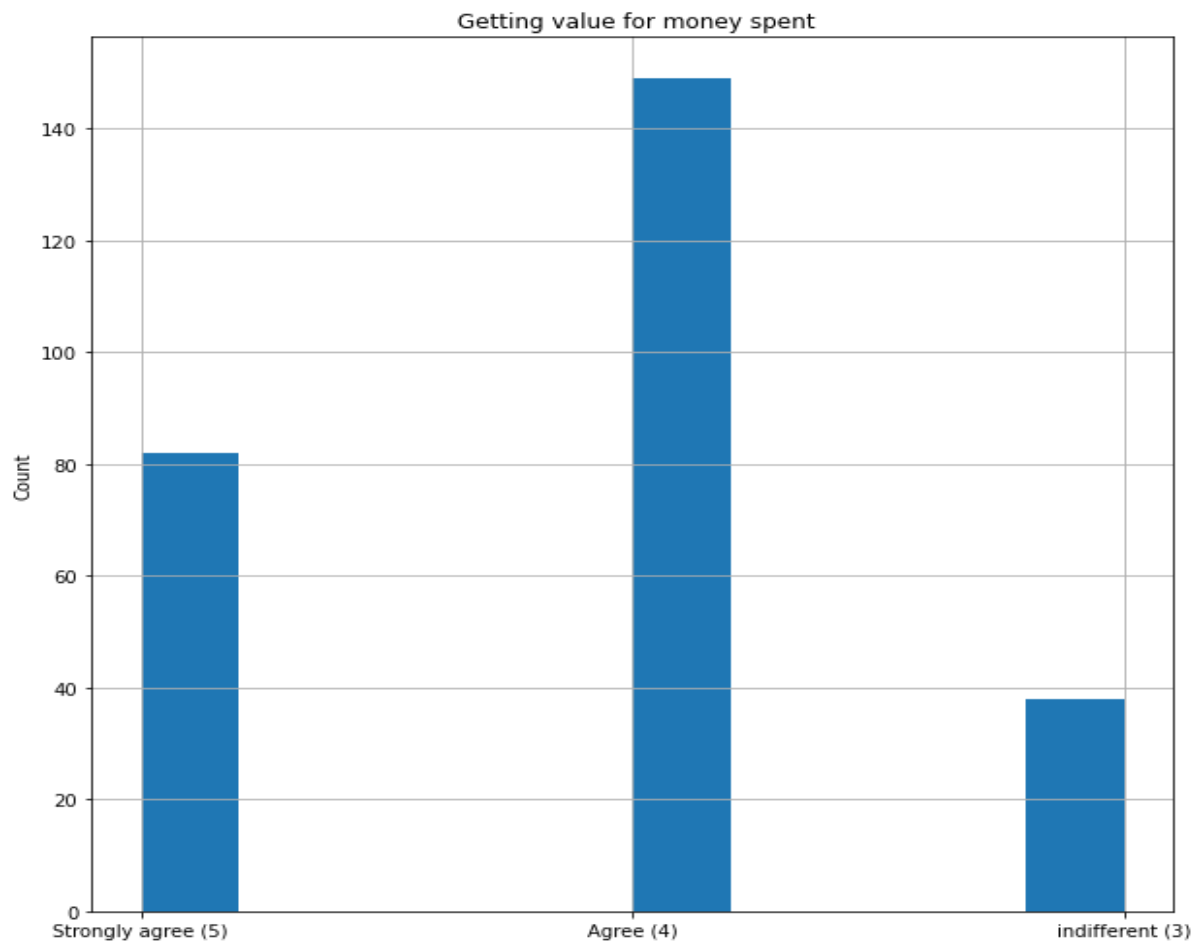
1. Lots of customers agree that online shopping is easy, flexible and convenient, where there is not much positive ratio in enjoyment of online shopping as it has mixed reviews from customers.
2. According to most of the customers, return and replacement policy of the company is very important along quality of information improves satisfaction of the customers as its gives convenience of shopping online.
3. Most of people also agree that quality, variety, complete information of product and trust are import for the customer to revisit these sites.



## Observation

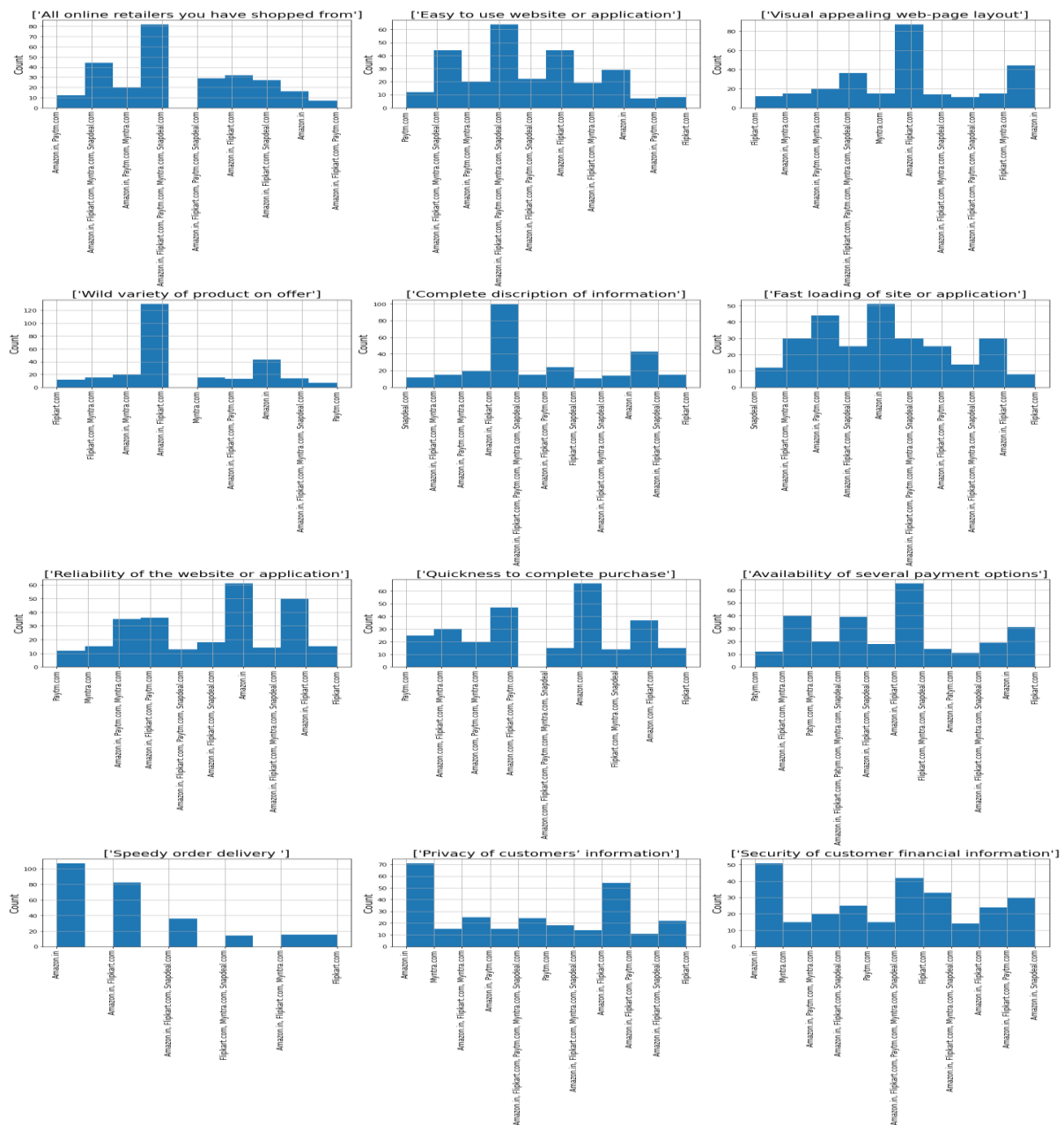
1. Most of the customers agrees that they have saved money while shopping online.
2. While most of the customer preferred to be patronized, but large amount of them is also not sure about this.
3. Also, there are mixed reviews of customer if they feel adventure in shopping.
4. Many people are not sure if online shopping affects their social status.
5. Most of the people are confused or not sure if they are gratified from favourite site.
6. According to most of the customers online shopping helps fulling roles.

## Value for money spent



Sr No	Getting value for money spent	Number of Customers
1	Agree (4)	149
2	Strongly agree (5)	82
3	indifferent (3)	38

## Comparison with target column

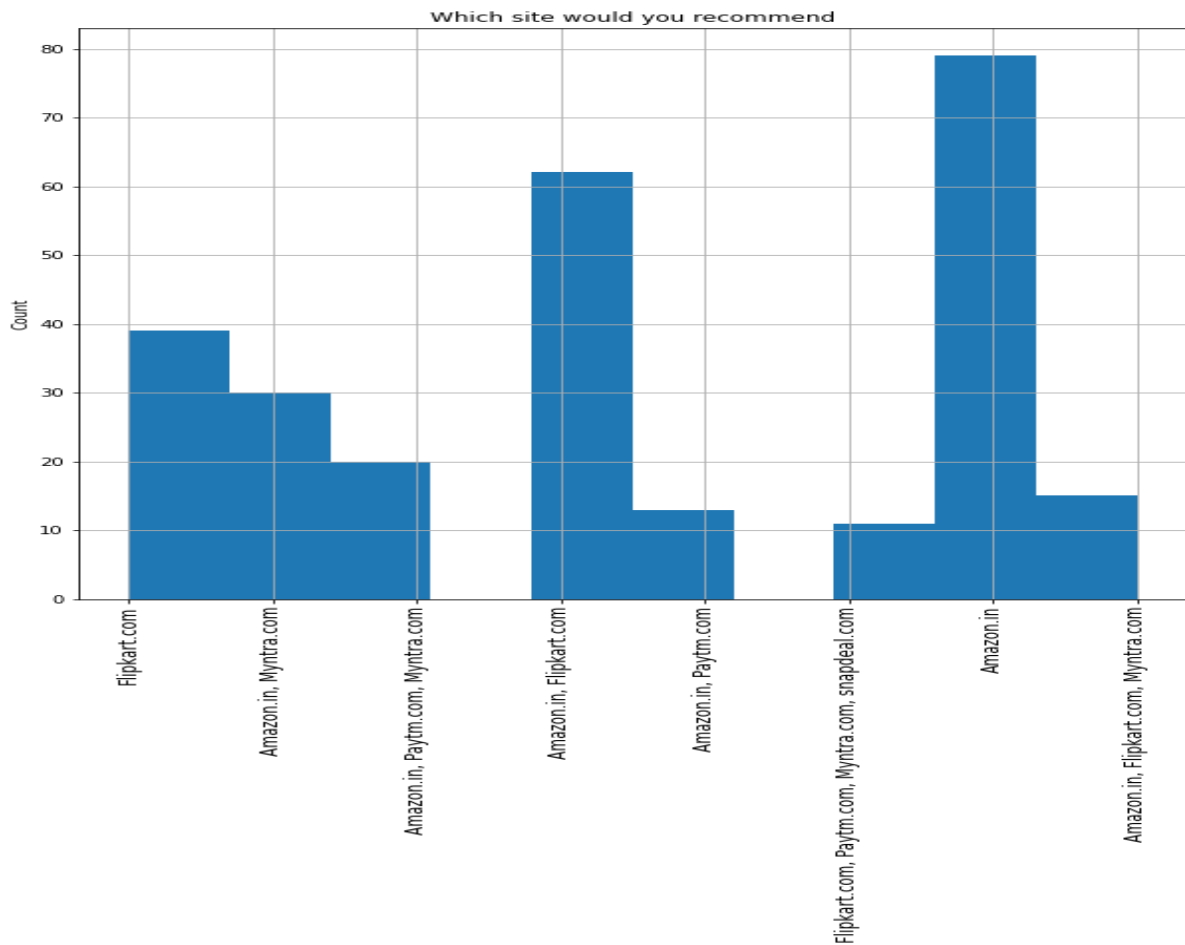


## Observations

1. Above data states that every has shopped from amazon n at least one time.
2. Most people also have voted for amazon for being easy to use site followed by flipkart. They both are also good in page layout or visual effects along with variety of products which they provide.
3. These two also provide proper description of products as well.
4. Many positive votes are in favour of amazon, it's on top number in providing delivery on time followed by flipkart and Snapdeal.
5. Customers have also voted highest to amazon for data privacy and keeping their financial information secured.

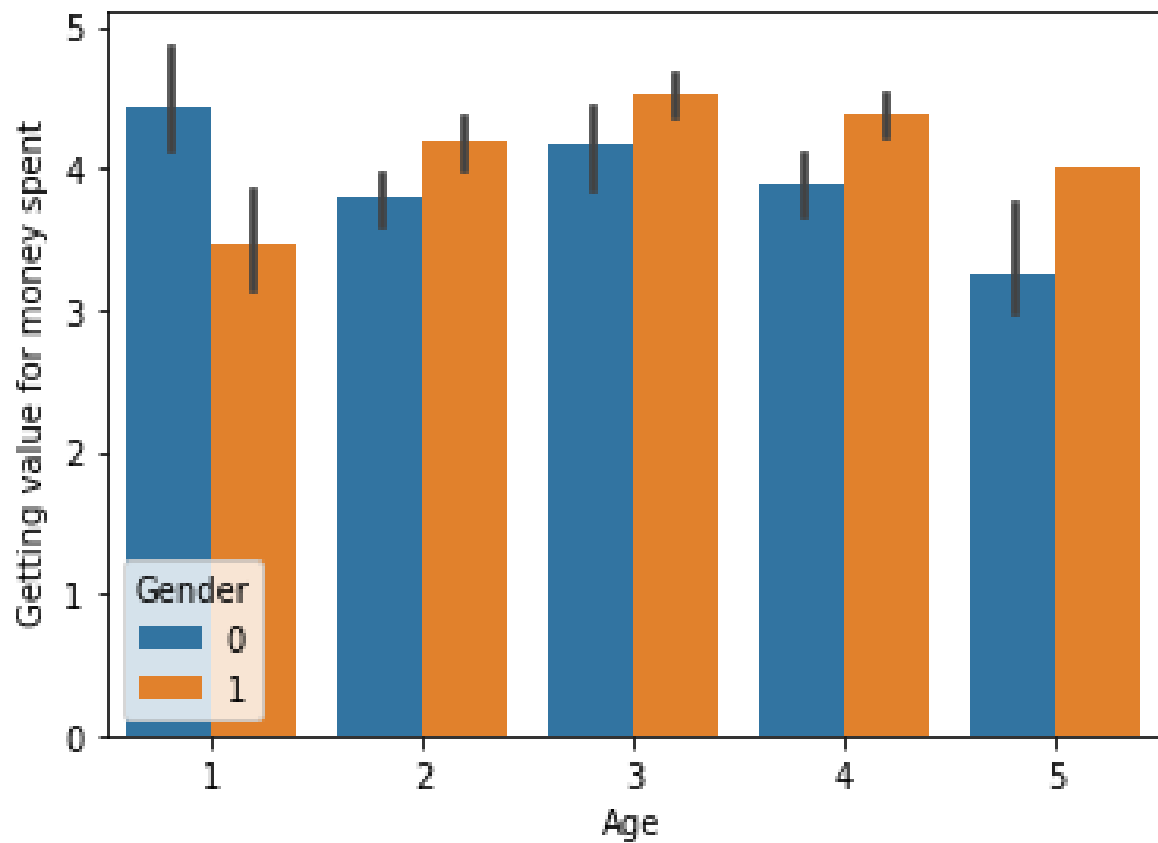


### Most recommended site



Sr No	Which site would you recommend	Number of Customers
1	Amazon.in	79
2	Amazon.in, Flipkart.com	62
3	Amazon.in, Flipkart.com, Myntra.com	15
4	Amazon.in, Myntra.com	30
5	Amazon.in, Paytm.com	13
6	Amazon.in, Paytm.com, Myntra.com	20
7	Flipkart.com	39
8	Flipkart.com, Paytm.com, Myntra.com, snapdeal.com	11

Most recommended site (Gender and age wise)



### Age

1 = Less than 20,

2 = 21-30 Years,

3 = 31-40 Years,

4 = 41 - 50 Years,

5 = 51 Years and above

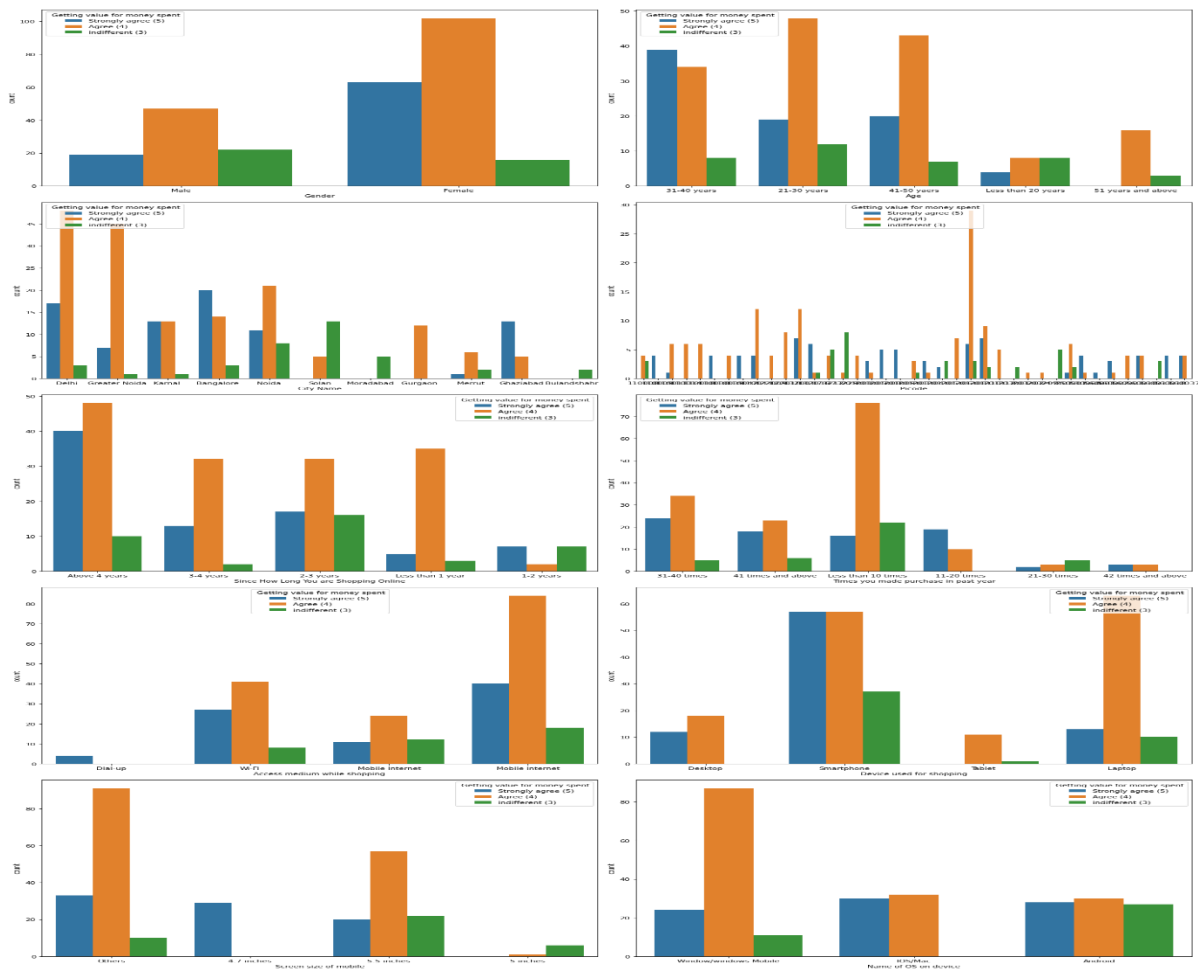
### Gender

0 = Male

1 = Female

**Females are more satisfied by the value for money spent in all other categories except where age is less than 20 years.**

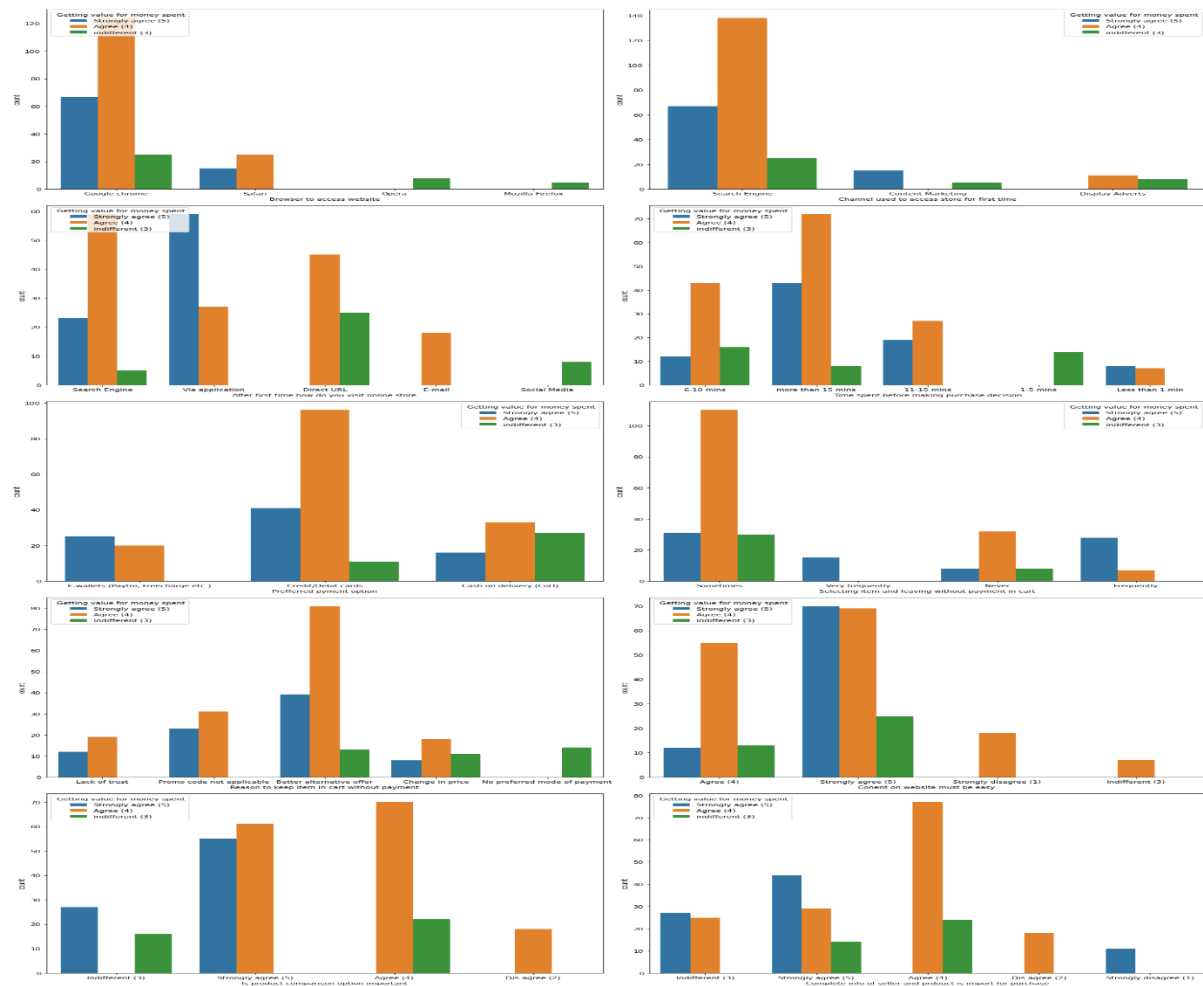
## Comparison with target column



## Observations

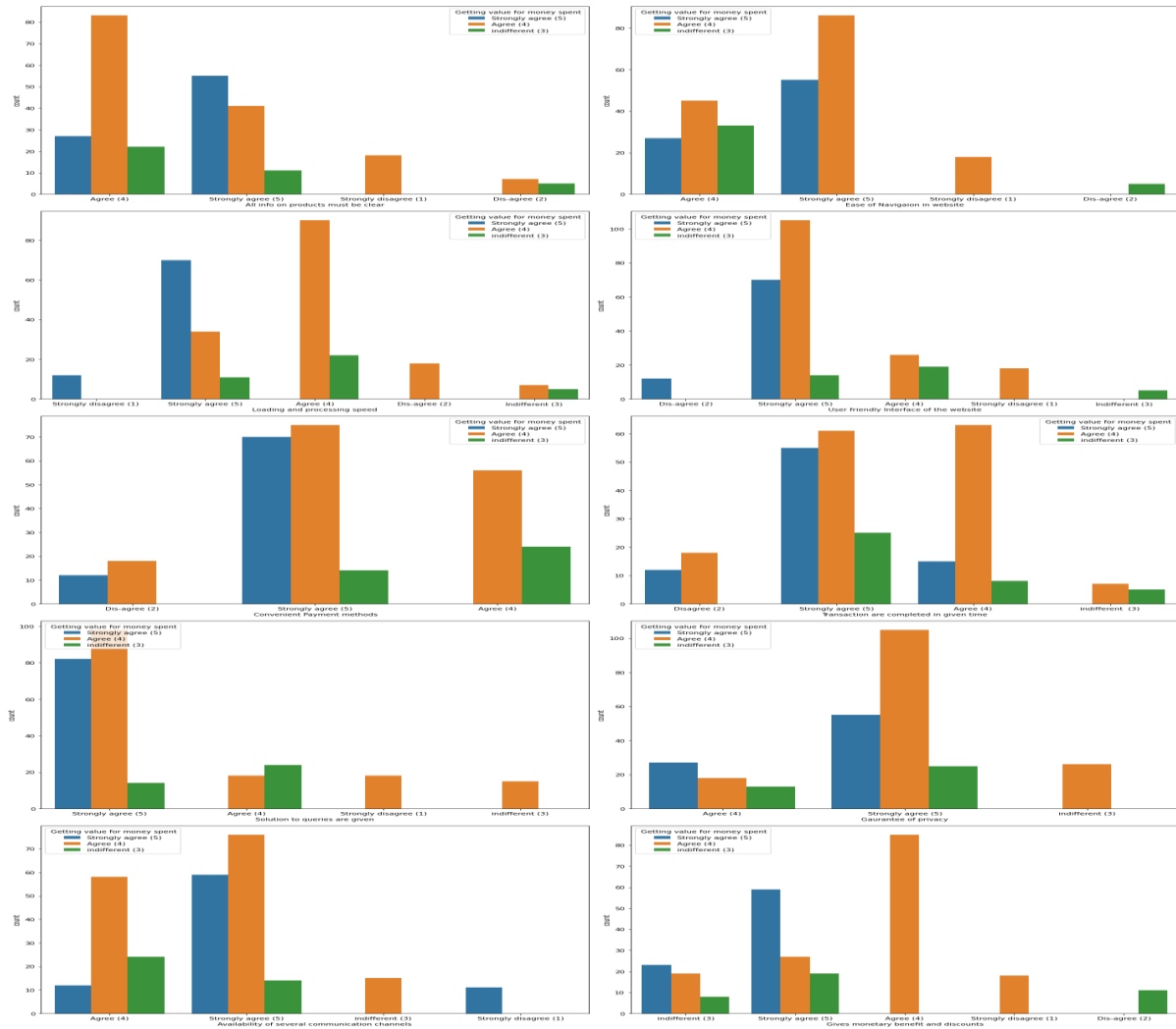
1. As per above graph there are more female customers than male and females are more, they have agreed that they are getting true value for the money they have sent.
2. Age category of 21-30 are more as compared to 31-40, however there is not much difference in their feed most of the people have agreed that they are getting value for the money spent.
3. Customers from Delhi and Greater Noida have agreed upon the same, but customers from Solan and Moradabad and Bulandshahr don't agree with this.
4. The ratio of customers which are shopping since long agree but customers who are shopping for 1-2 years have poor ratio.
5. Customer who shops for less than 10 times have poor ratio of happy customer, however as the number gets increasing the ratio of happy customers getting increasing.
6. Smart phone users are happier in this case as compared with ratio of other users.





## Observations

1. People using chrome have agreed that they are getting true value for money, however customers using opera and Mozilla fire fox seems to be dissatisfied as there is no customer agree to the value for money sent.
2. Customers who spent more than 15 minutes before shopping are satisfied that they are getting true value for money where as in another scenario the ratio gets depraved.
3. There are more agreeing credit card customers however the ration gets bad where people who have selected COD.

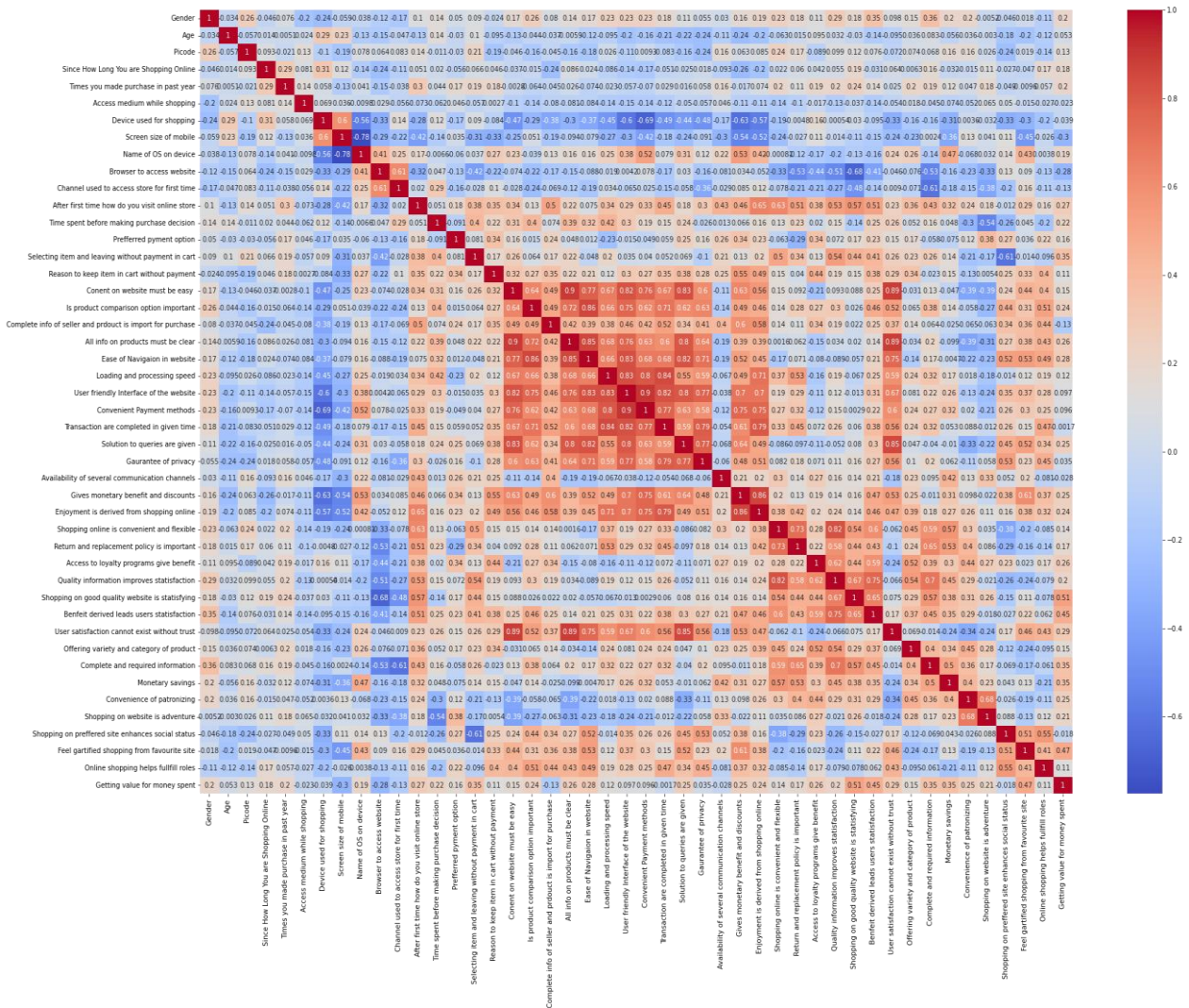


## Observations

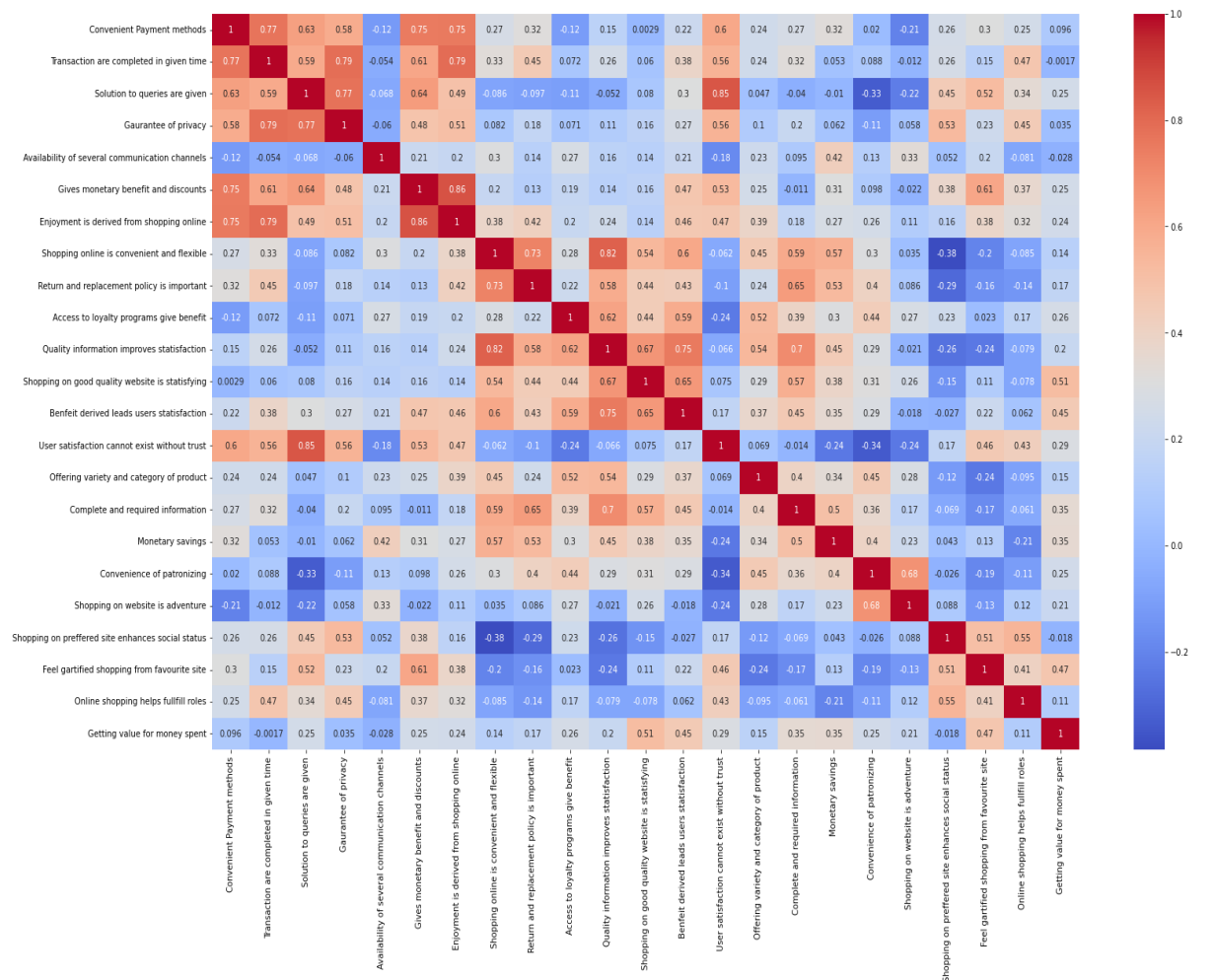
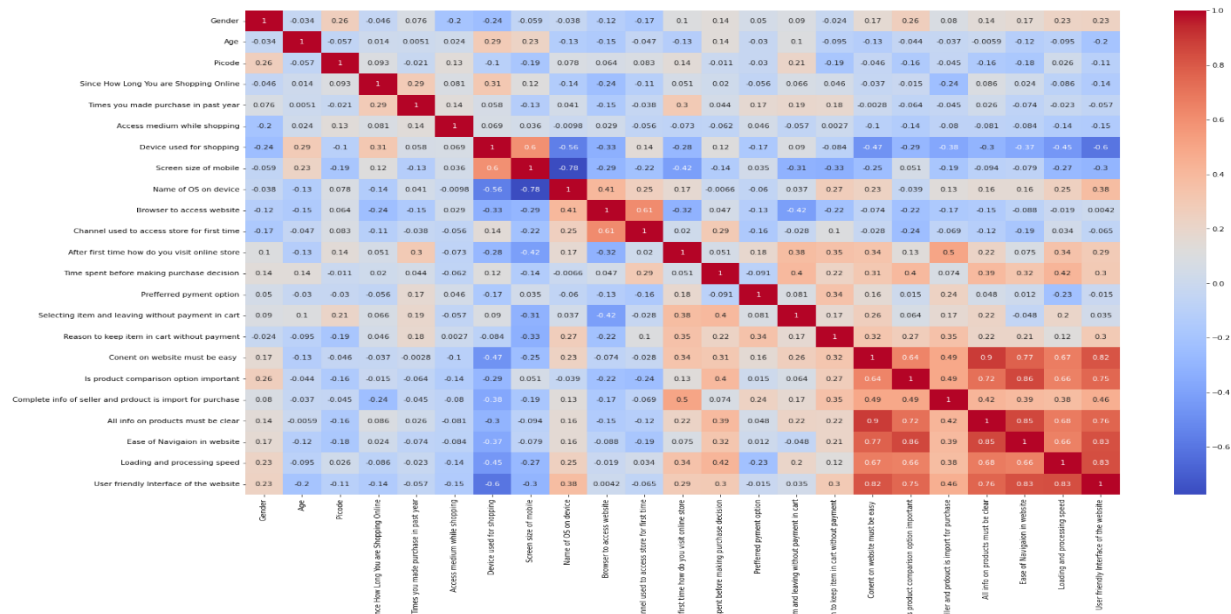
1. The customer who thinks websites has convenient method of payments are happy and agree that they are getting true value for the money they spent.
2. Where solution to the customers query is answered they seem to be happier and here the strongly agree with the value for money spent.
3. Same with the data privacy, it is also related to the value for money spent, it gives positive indication to the customers.
4. The person who gets more discounts seems to be happier and agreeing.
5. Availability of several communication medium to the customers also gives pleasure to the customer which helps in getting positive response.



## Correlation of dataset set using heatmap

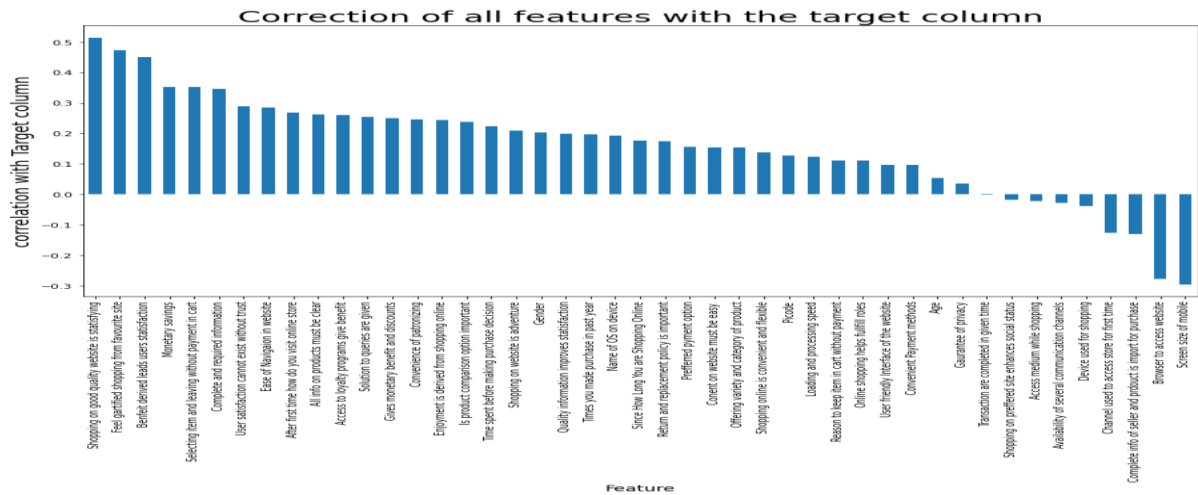


As pervious map was bit congested so divided it into 2 parts for more clear correlation. There is correlation of features with each other, which states where is relation between columns.



This is second part of our heatmap states correlation of columns with each other.

## Correlation of all features with targe column(bar)



## Final observation

All other columns are in positive relation to getting value for money spent however there are some exceptions which I have mentioned below. Below factor does not have any effect on value of money. Guarantee of privacy, Transaction are completed in given time, Shopping on preferred site enhances social status, Access medium while shopping, Availability of several communication channels, Device used for shopping, Channel used to access store for first time, Complete info of seller and product is import for purchase, Browser to access website and Screen size of mobile

## Conclusion

The ease of usage of internet and these websites along with quality, attractiveness of the web pages and websites along with other proper and arrange customer support makes a customer to be a repeated user of these sites. It is also observed that shopping on good quality of website have the most positive feedback for the customer along with this discount and multiple payment options also attracts the customer attention along with other factors.