Customer Retention Project

Submitted By:-

Mansi Nagpal

Problem Statement & Understanding

Problem Statement

A new project is allocated and the project name is Customer Retention. The attached zipped folder contains the below files:
• Customer_retention_case_study
• Customer_retention_dataset

Submission file needed to contain the below three files:
• 1. The jupyter notebook solution
• 2. The project report in pdf format
• 3. A PowerPoint presentation containing problem statement and understanding, EDA steps and visualizations, Steps and assumptions used to complete the project, analysis, and conclusion.

INTRODUCTION

This is a case study on Indian E-commerce customers to know which are the major factor for activation and retention of the customers, Customer satisfaction has emerged as one of the most important factors that guarantee the success of online store

Activation/Retention

For this project a detailed research has been conducted to increase the purchase ,activation and retention of customers.

Five major factors

Service quality, System quality, Information quality, trust and net benefits are the key factors.

Loyalty

The combination of both utilitarian and hedonistic values are need to affect the repeat purchase(loyalty)

Project Flow

Basic outlook over the project

This slide talk about the aim of the project,Informations to be drawn for better conclusions etc.

Processing of data

This slide talk about preprocessing of the data

Univariate Analysis

This slide talk about the analysis of independent variables

Bi variate Analysis

This slide talk about the relationship between two variables.

Multivariate Analysis

Here we are going to look the relationship between more than 2 variable

Conclusion

This final slide talk about the summary of the information drawn from the analysis

Basic outlook over the data

● ● ● ● ●

Understanding We are provided with a dataset which contains a more than 47 columns of factors that directly affects a online shopping store. It contains the data of a survey taken from different age group peoples from different areas of India. It shows what website they prefer more to shop and recommend to other peoples as well. 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.

Preprocessing of data

Categorical columns

Since all the columns are of categorical type we did the label encoding, There were no null values, In some cases column name was out of simplicity so we renamed it for future analysis purpose

Dictionary

We made separate dictionary for each columns to keep which label encoded value infers what.

Exploratory Data Analysis

● ● ● ● ●

Since the dimension of the dataset is significantly wide there was lump sum amount of data in different angles , We picked up relevant information through 3 types of following analysis.
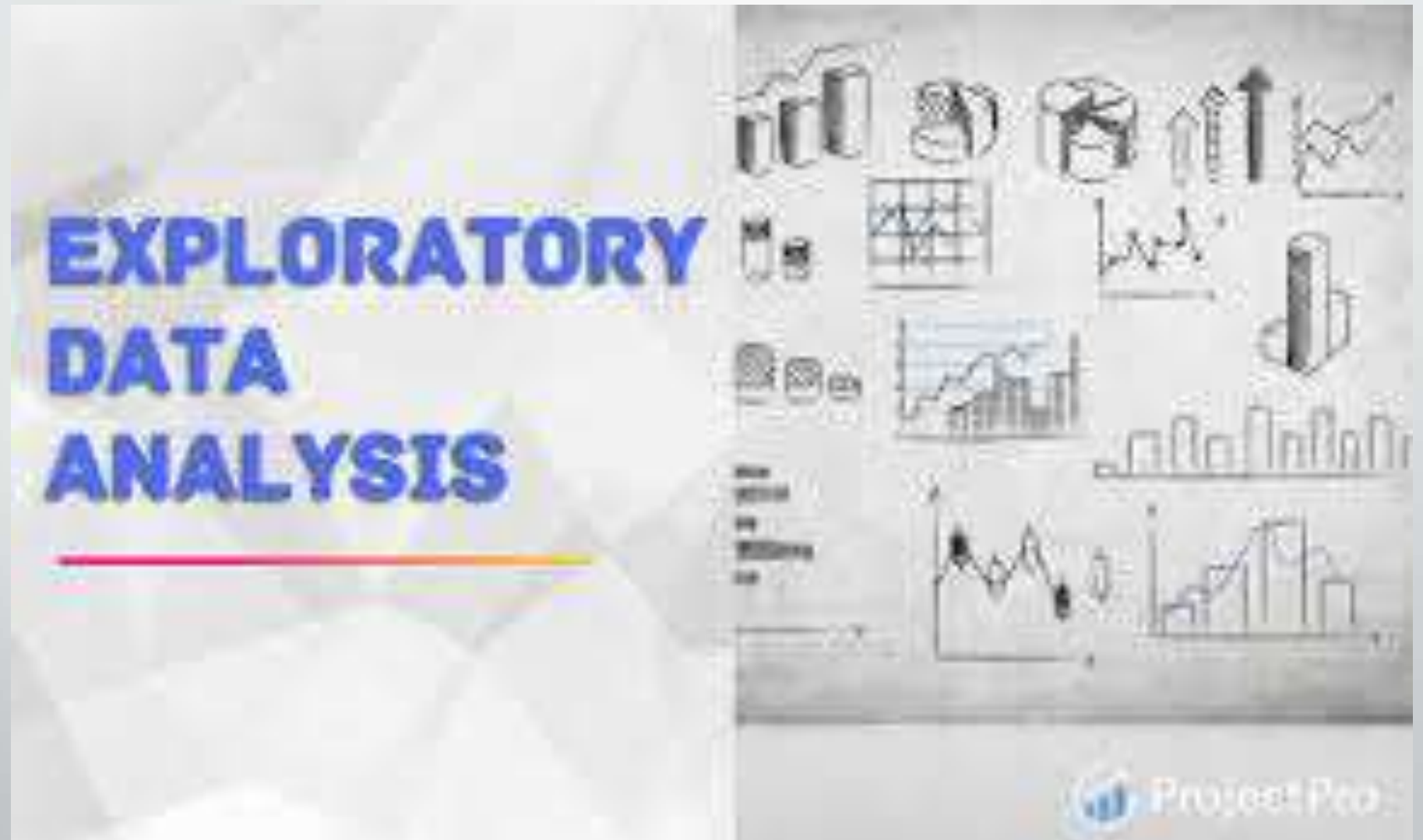
Univariate Data Analysis

In this case we analyze only one variable at a time, there is detailed information in the project report , Here we are going through important features only.

General information's of online retail shop users

❑ In this particular dataset around 67% of users are females and left 33% are male
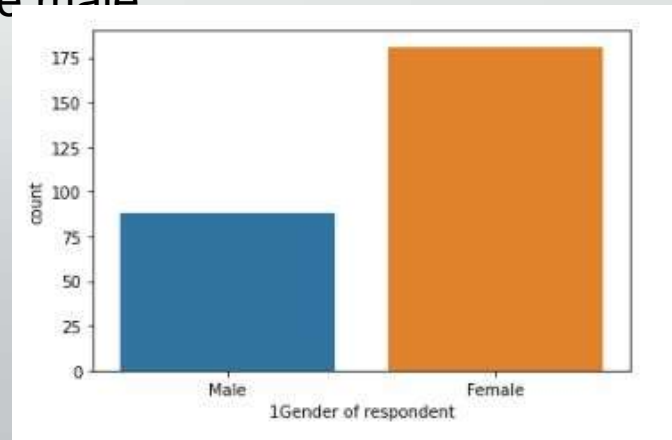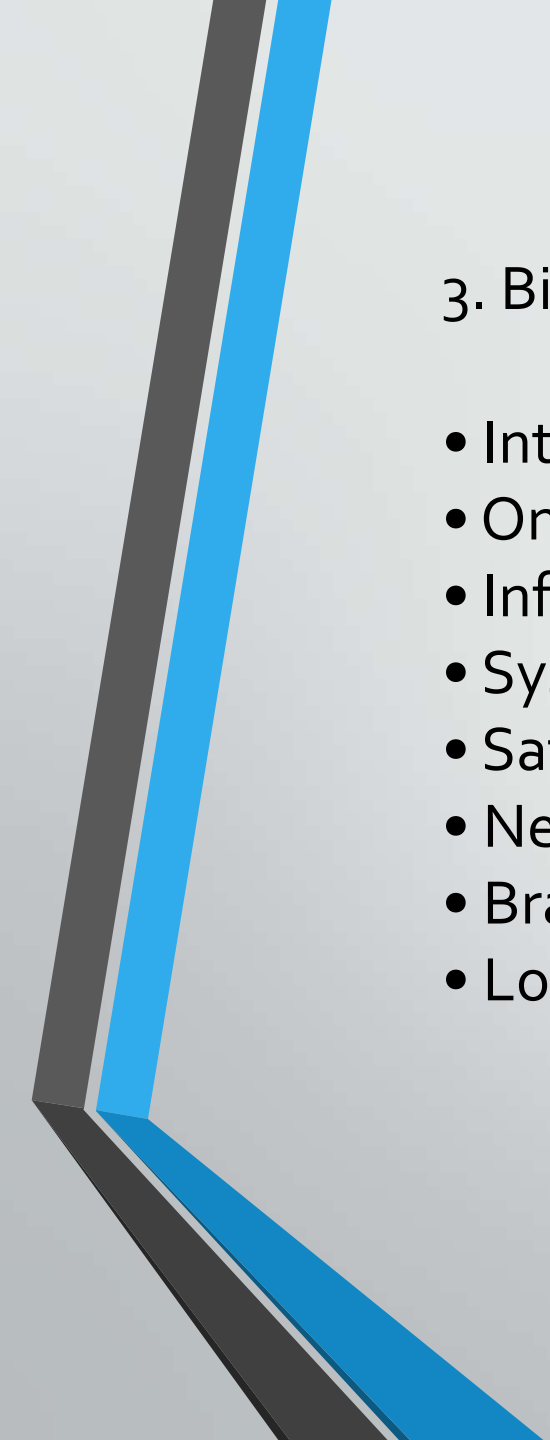
EDA Steps

1. Checking the attributes:

• Shape of the dataset
• Columns present in the dataset
 • Datatypes of each column in the dataset
• Pre-processing the columns name like removing the number and extra space from it
• Unique values in each column
• Checking the null values in the dataset

2. Univariate analysis
• Analyzed the personal details like age, city, gender, how many time they shop in the past year, since how long they are shopping.
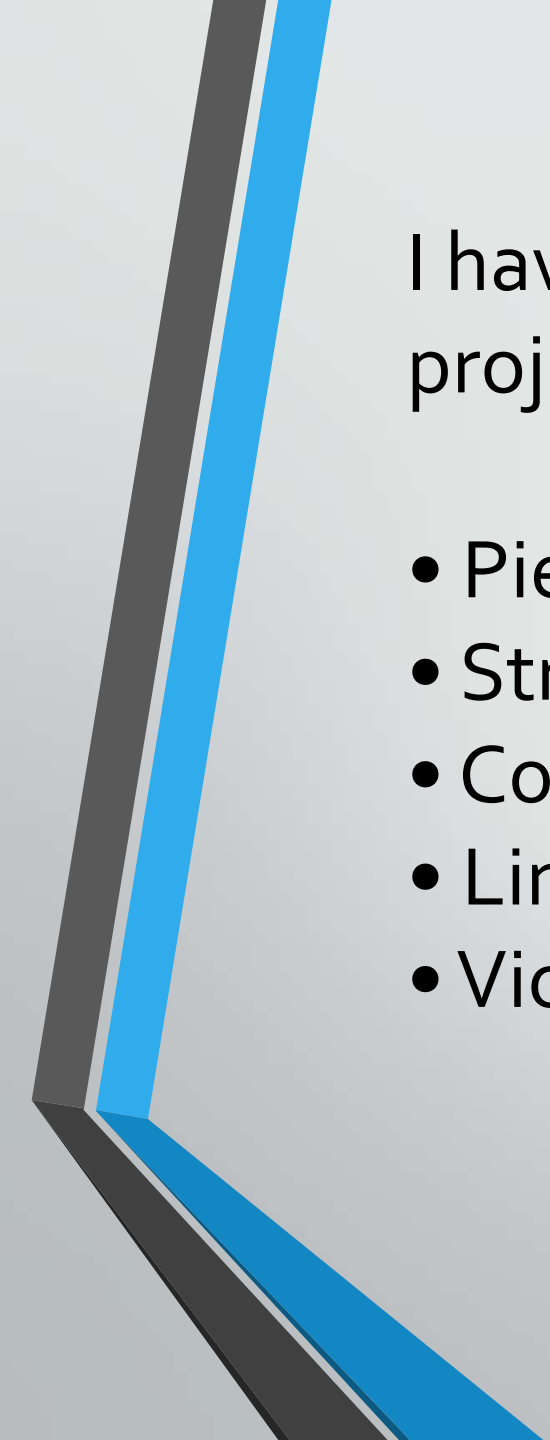
3. Bivariate Analysis:

- Intention of repeat purchase
- Online retailing
- Information
- System quality
- Satisfaction
- Net benefits
- Brand image
- Loyalty

Visualizations

I have used the following visualizations in my project to analyze the data in different ways:

- Pie chart
- Strip plot
- Count plot
- Line plot
- Violin plot

Let's see the analysis

1. Personal Info - It contains the personal details of a person like age, city,gender, how many times they shop in the past year, since how long they areshopping. We use the pie plot to see the different percent in the dataset and determine which category has the superiority. We found these conclusions after plotting:

• There is double the number of women than men who have taken this survey.
• Most of the people are in their 30's followed by 20's, teenagers and senior citizen are the least in number.
• Most of the people belong from Delhi, Noida and Bangalore, ambiguity can also be seen as Noida has two categories (Noida and Greater Noida) which need to be handled
• Most of the people shopping online have been shopping from a long time.
• Majority of people shop online 10 times a year, ambiguity can also be seen for range 42 times and above which needs to be handled

2. Intention of repeat purchase –

It means how much a person is intended to purchase the product again through online stores. We use the line plot and violin plot to analyse the different factors of repeat purchase and got the following
conclusions:

• Heavy shoppers who shop more than 41 times a year shop from all the online brands, some of the people who shop for 32-40 and less than 10 times a year seem to exclude myntra. People shop from Amazon and flipkart whatever be the case
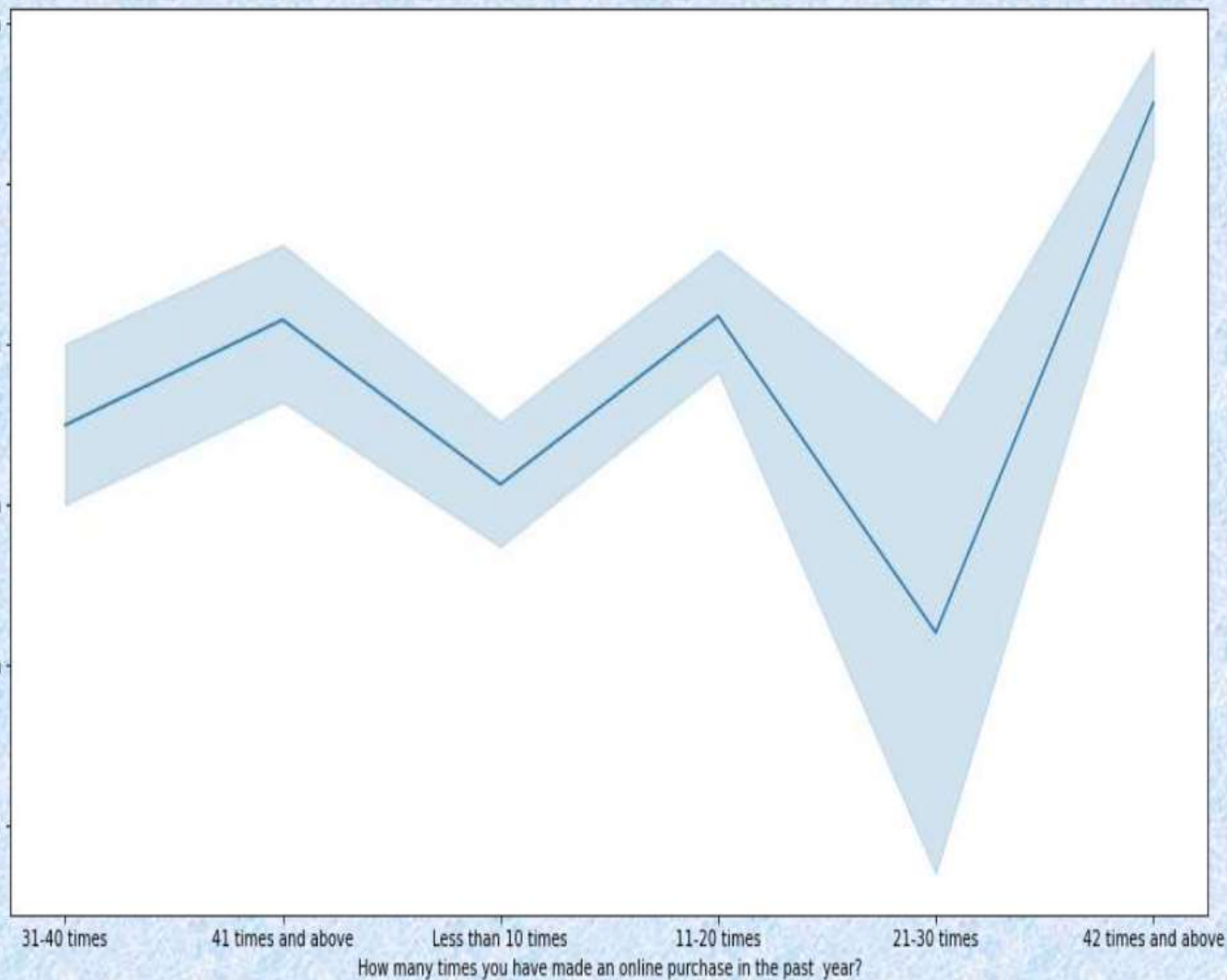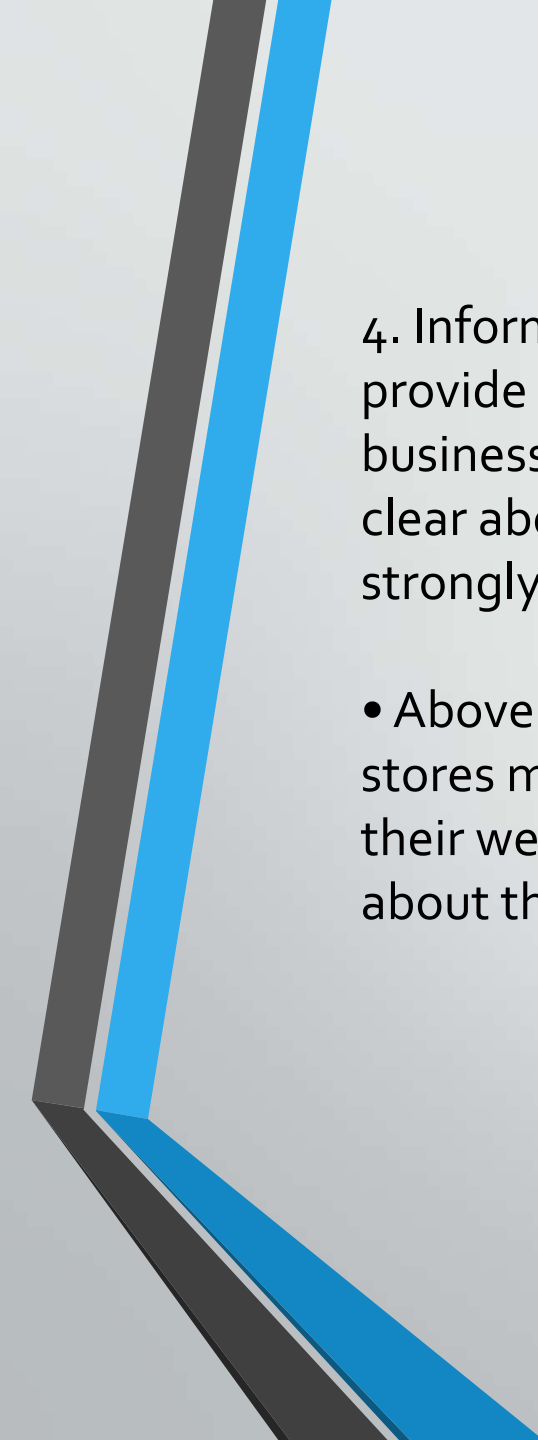
3. Online Retailing – We try to find out which age group shop the most from which city and how they reach to their favorite online store, how long do you think before a purchase decision. We use different plots to conclude on this like line plot, count plot and obtain the following results:
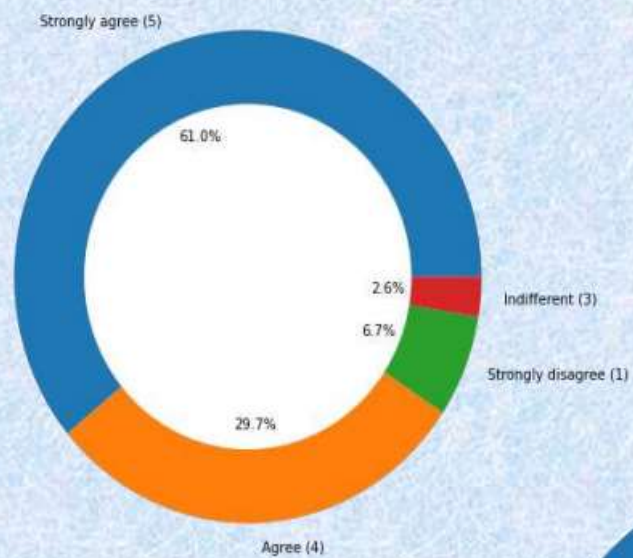
• Highest number of people have been shopping online for above 4 years except for the age group below 20 years and above 50 years. People who are shopping online for 1-2 years does not include teenagers and elder people.

• In lines, we can see that density of female customers is more than male. Men living in Bangalore and Ghaziabad shop have shopped online for less than 1 year. Highest number of men shopping online belong from Delhi and Noida, while men from Moradabad have been shopping online for the longest. Women from Meerut and Noida have shopped the longest.

• Men & Women both spends a lot of time to make a purchase decision on Amazon

• Also they spend a less than 1 min before making a purchase decision on Amazon, Flipkart, Paytm

• They spend a lot while surfing through many websites.

• Even though people who are shopping online for more than 3 years use search engine in large number which indicates that first time users always tried the search engine to shop on their favorite online store for the first time.

• Even though people who are shopping online for more than 3 years do not use the application rather use search engine and direct URL's in large number which indicates that online brands should update all their platforms rather than just application.
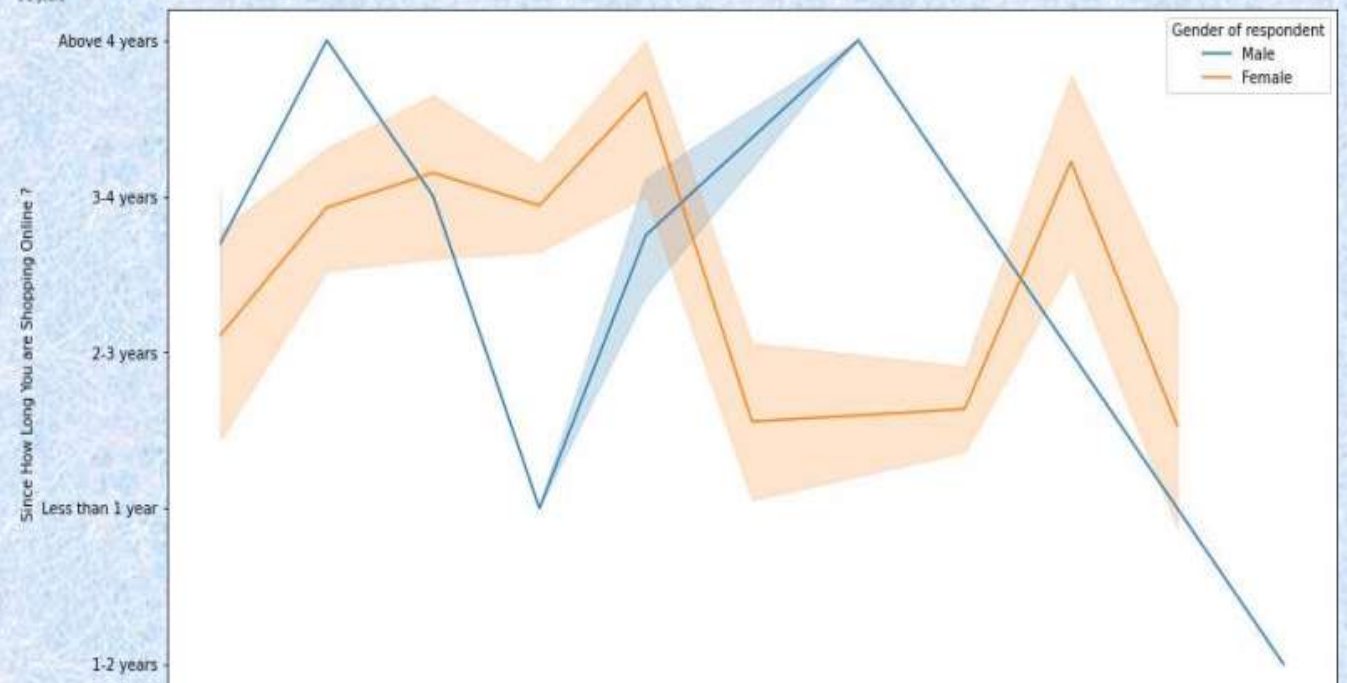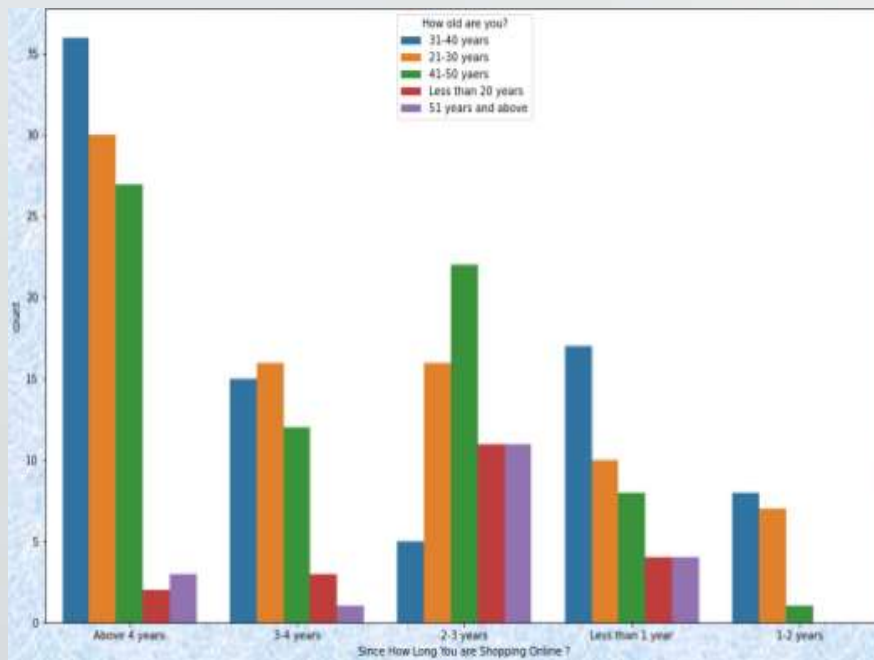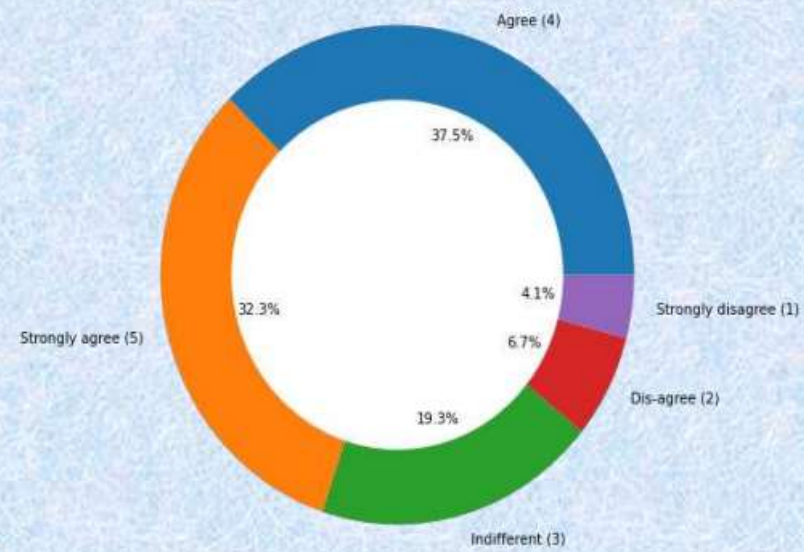
4. Information – Here information means the details regarding a product provide by the different online shopping websites. This play a important role in the business of e-commerce as detailed & correct information makes the buyer more clear about the product. We have used the pie plot to compare how many buyers strongly agree to this point and get to know:

• Above comparisons shows that both men or women strongly agree that online stores must mention a complete and relevant information about the product on their websites. Provision of complete information leaves a clear perception about the product.

Strongly agree (5) — 61.0%
Agree (4) — 29.7%
Indifferent (3) — 2.6%
Strongly disagree (1) — 6.7%
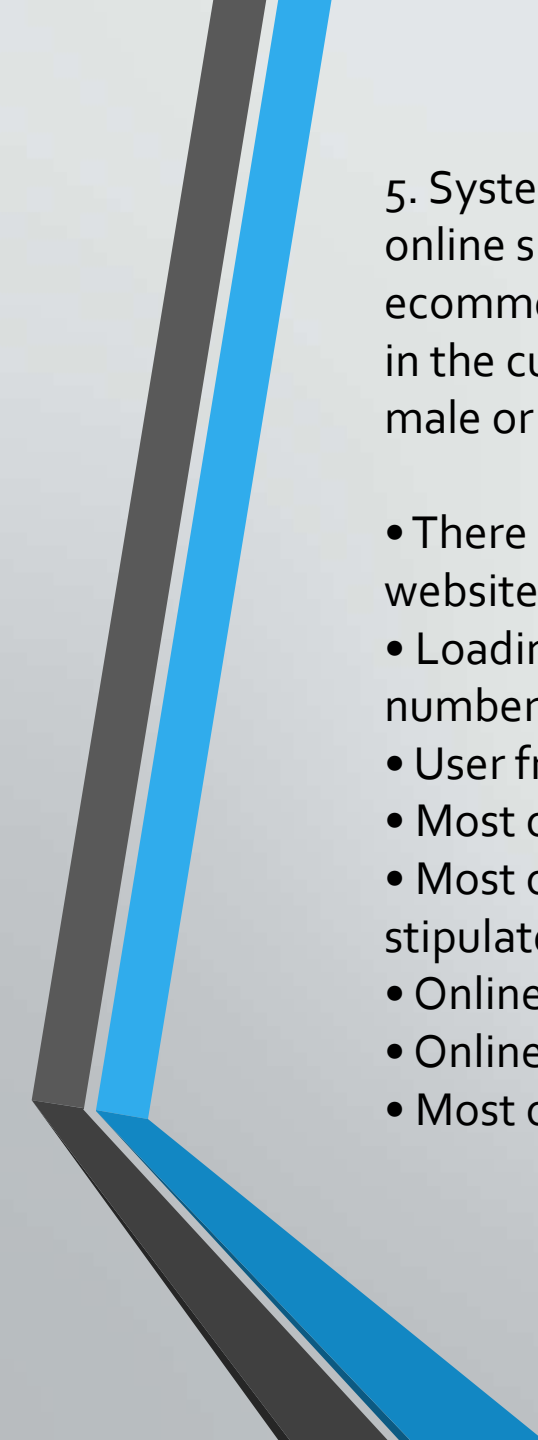
The content on the website must be easy to read and understand

Agree (4) — 37.5%
Strongly agree (5) — 32.3%
Strongly disagree (1) — 4.1%
Dis-agree (2) — 6.7%
Indifferent (3) — 19.3%

Complete information on listed seller and product being offered is important for purchase decision.

Strongly agree (5) — 50.2%
Agree (4) — 36.4%
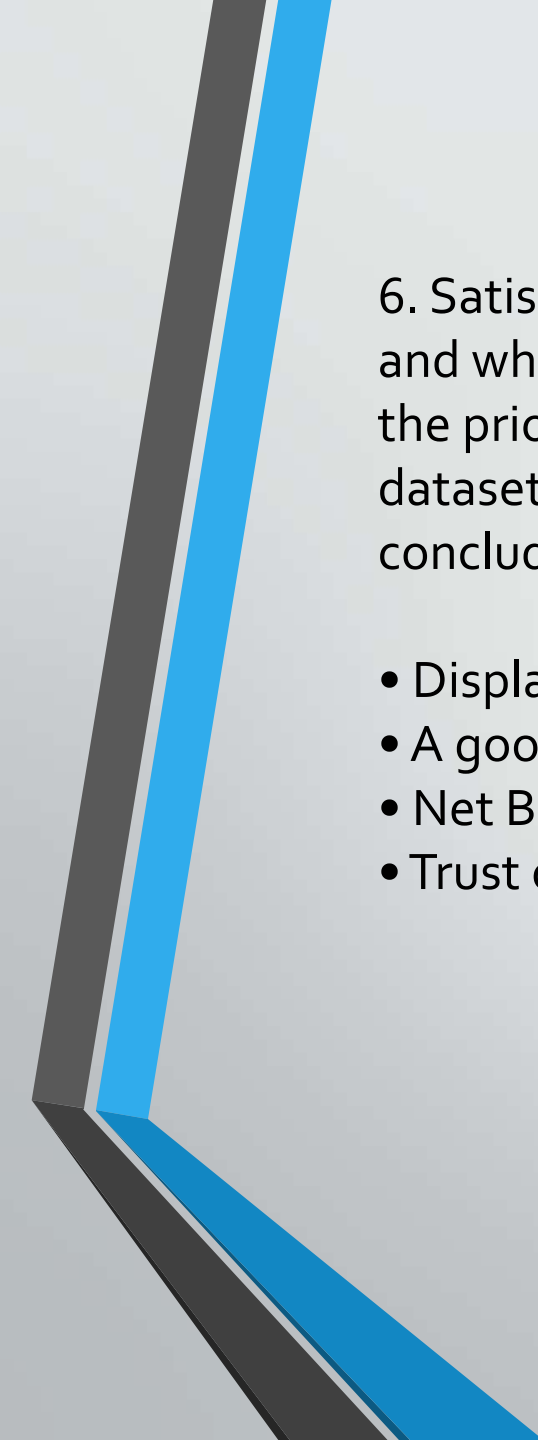Disagree (2) — 1.9%
indifferent (3) — 11.5%
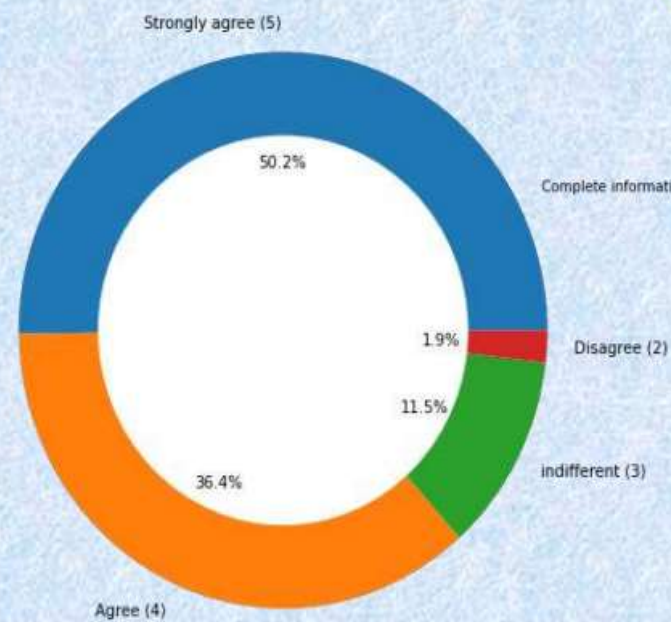
Provision of complete and relevant product information

5. System Quality – It refers to the website quality or the user interface qualities of an online shopping website. If the website is slow or not maintain the quality then the ecommerce business can face the loss of customers as well so it plays a significant role in the customers retention. We have used the stripplot to compare the thoughts foe male or female about the system quality of different shopping portals.
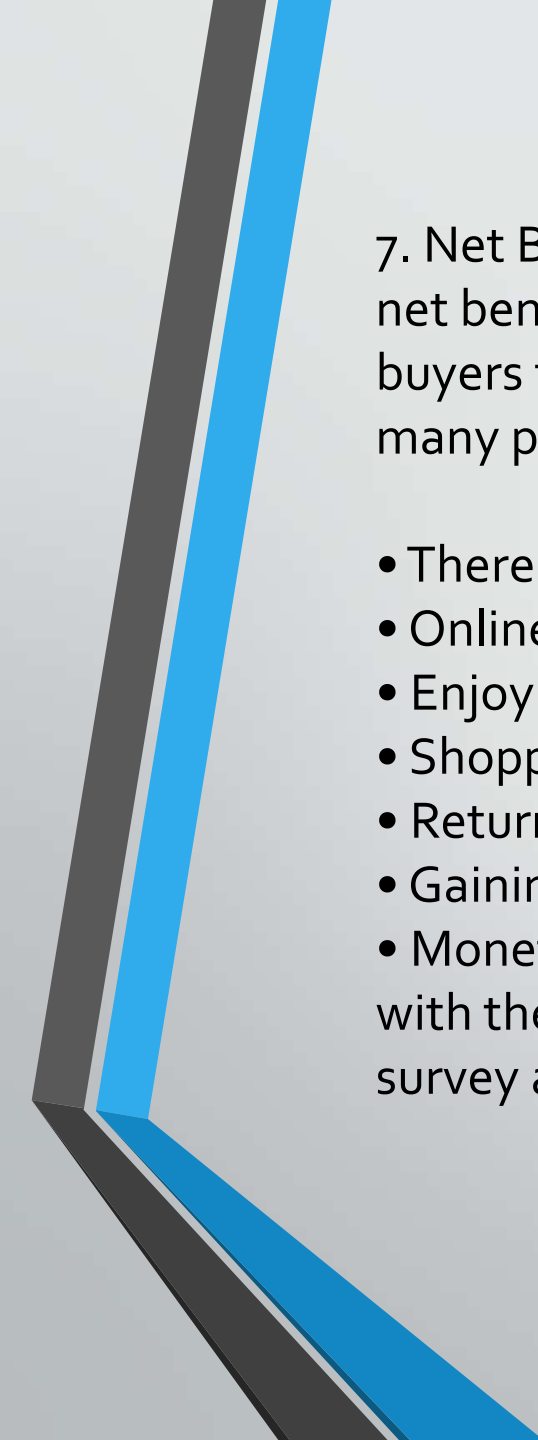
• There are peoples who are strongly disagree that there is an ease of navigation on the website in any of the online store
• Loading & processing speed of any online shopping website can create a difference in the number of customers.
• User friendly interface entact the customer
• Most of the websites doesn't have the convinient payment method
• Most of users trust that the online retail store will fulfill its part of the transaction at the stipulated time
• Online websites have a good backend customer care services to their customers
• Online websites are able to provide the privacy to their customers
• Most of the online stores have a good responsiveness
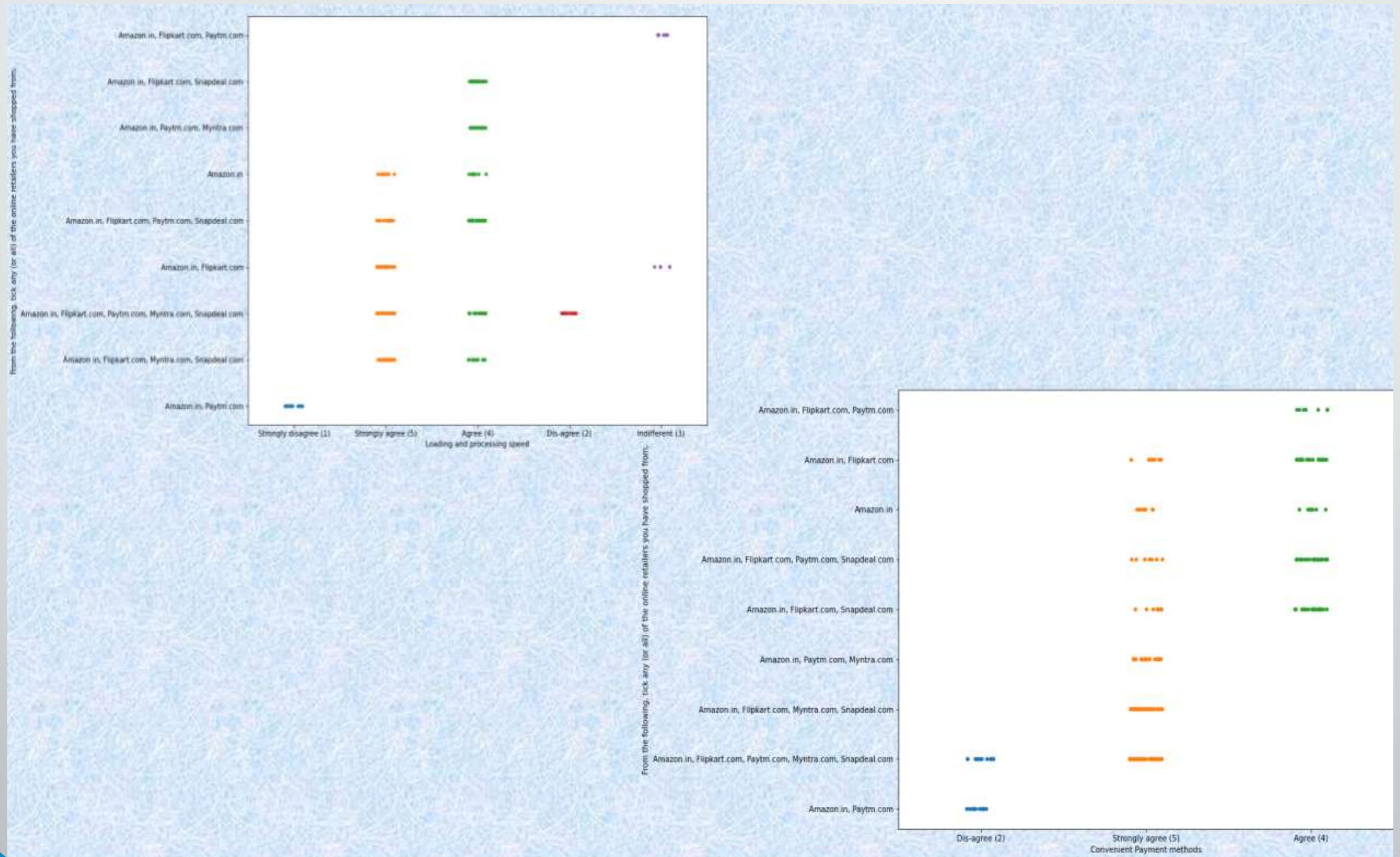
6. Satisfaction – How much a customer is satisfied by the online retailing store and which factors affect the satisfaction. Satisfaction of a customer should be the priority of every online shopping store. We have some factors in the dataset, so we used the pie plot to compare the thoughts of buyers and conclude that these four factors strongly agreed by the buyers:

- Displaying quality information
- A good website with better system quality
- Net Benefits from the online store
- Trust on the website by the users
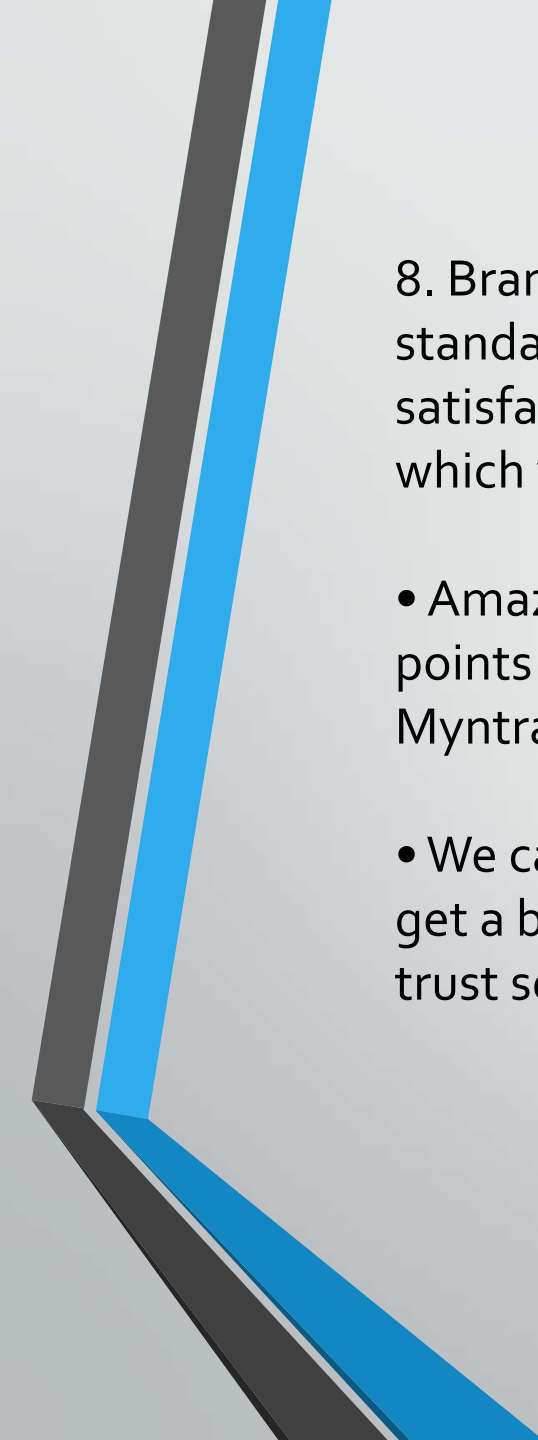
7. Net Benefits - A person buy online for the benefits and looking forward to his net benefits on the purchase. Shopping online provides great benefits to the buyers that's why they attract towards it. We used count plot to analyze how many people(he/she) agreed that benefits plays a important role.

- There are some benefits users expects after shopping online:
- Online shopping gives monetary benefit and discount
- Enjoyment is derived from shopping online
- Shopping online is convenient and flexible
- Return and replacement policy of the e-tailer is important for purchase decision
- Gaining access to loyalty programs is a benefit of shopping online
- Monetary savings

with the above comparison we can conclude that Mens participated actively in the survey and there are some of the things peoples are indifferent with others more
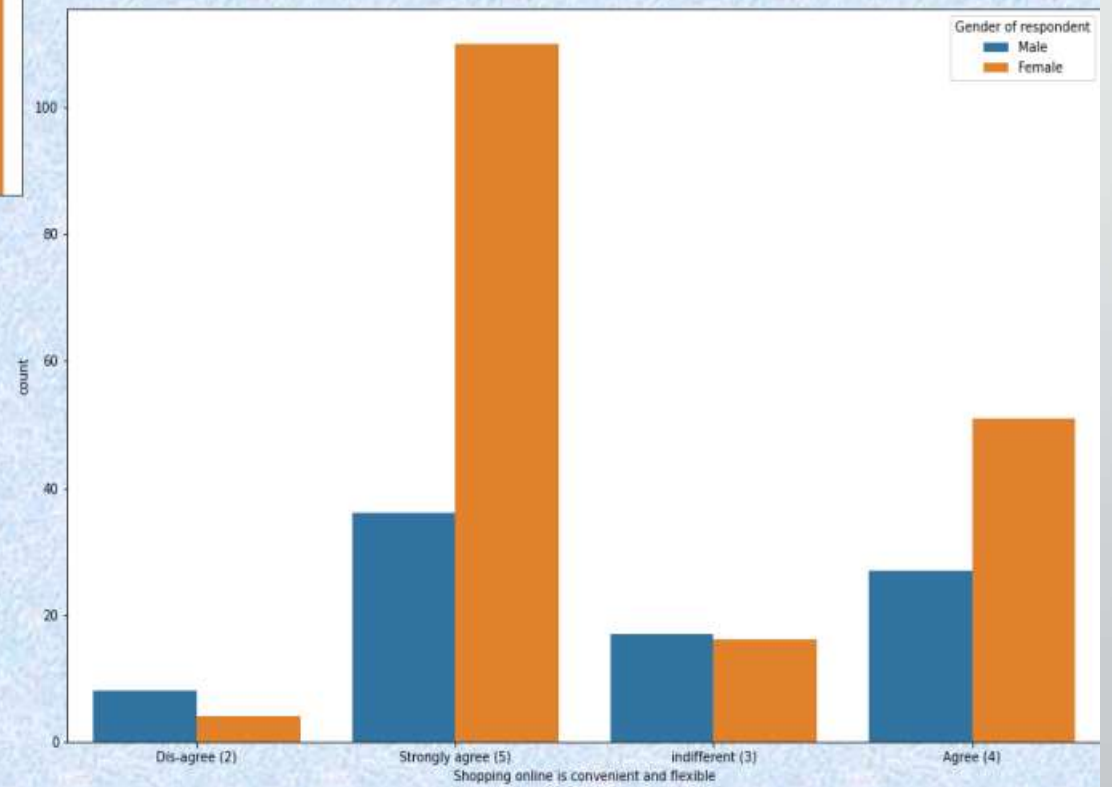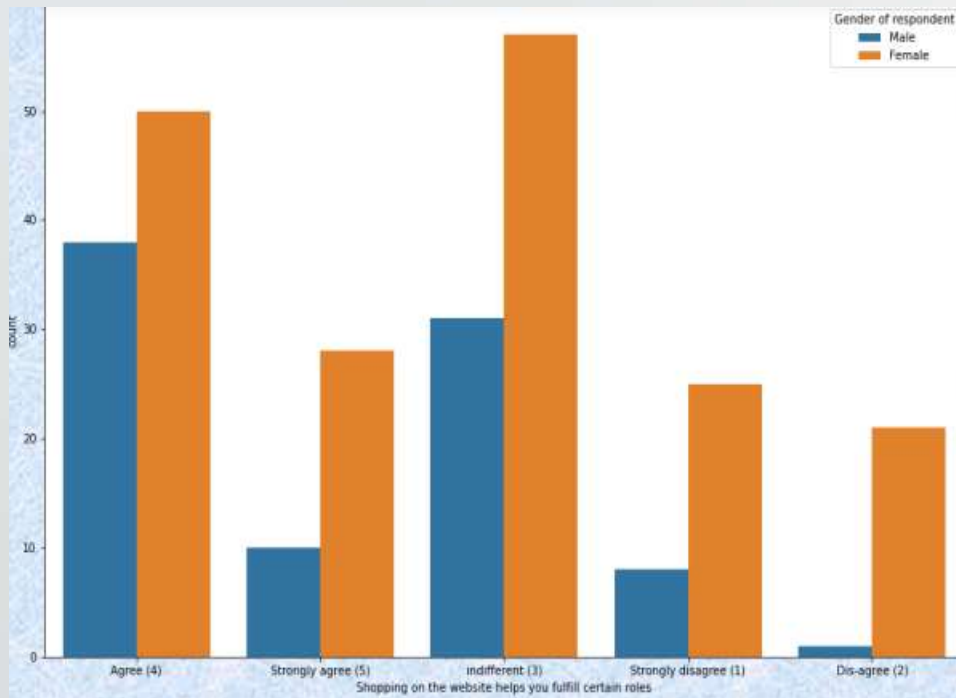
8. Brand Image – A shopping platform becomes a brand when it maintains its standards and able to provide the best quality products, support & services, satisfaction continuously to their customers. By the use of pie plot we try to find out which website performs well in the market and maintain its standards & quality.

• Amazon, Flipkart have been had the highest votes for having all the positive points and have maintained a very good brand image followed by Paytm and the Myntra.

• We can clearly see that most of the time people abandon the bag is because they get a better alternative offer or promo code not applicable. There is also lack of trust seen in amazon, flip kart and Paytm by some people.

Wild variety of product on offer

- Amazon.in, Flipkart.com — 48.3%
- Amazon.in — 16.0%
- Amazon.in, Myntra.com — 7.4%
- Flipkart.com, Myntra.com — 5.6%
- 5.6%
- Myntra.com — 5.2%
- Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com — 4.8%
- Flipkart.com — 4.5%
- Paytm.com — 2.6%
- Amazon.in, Flipkart.com, Paytm.com

Speedy order delivery

- Amazon.in — 39.8%
- Amazon.in, Flipkart.com — 30.5%
- Amazon.in, Flipkart.com, Snapdeal.com — 13.4%
- Amazon.in, Flipkart.com, Myntra.com — 5.6%
- Flipkart.com — 5.6%
- Flipkart.com, Myntra.com, Snapdeal.com — 5.2%

Perceived Trustworthiness

- Amazon.in — 28.3%
- Amazon.in, Myntra.com — 13.4%
- Amazon.in, Flipkart.com — 13.0%
- 11.5%
- Flipkart.com — 10.0%
- Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com — 9.3%
- Myntra.com — 5.6%
- Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com — 4.8%
- Amazon.in, Flipkart.com, Paytm.com — 4.1%

9. Loyalty - Loyalty is what a customer shows towards a website even after facing some issues and continue shopping on that website. Loyalty of a customer is what a E-commerce business earns.

• Customers seem to be more loyal to amazon, flipkart and paytm as even though many of them have given negative remarks about them still they would recommend these platforms to their friend
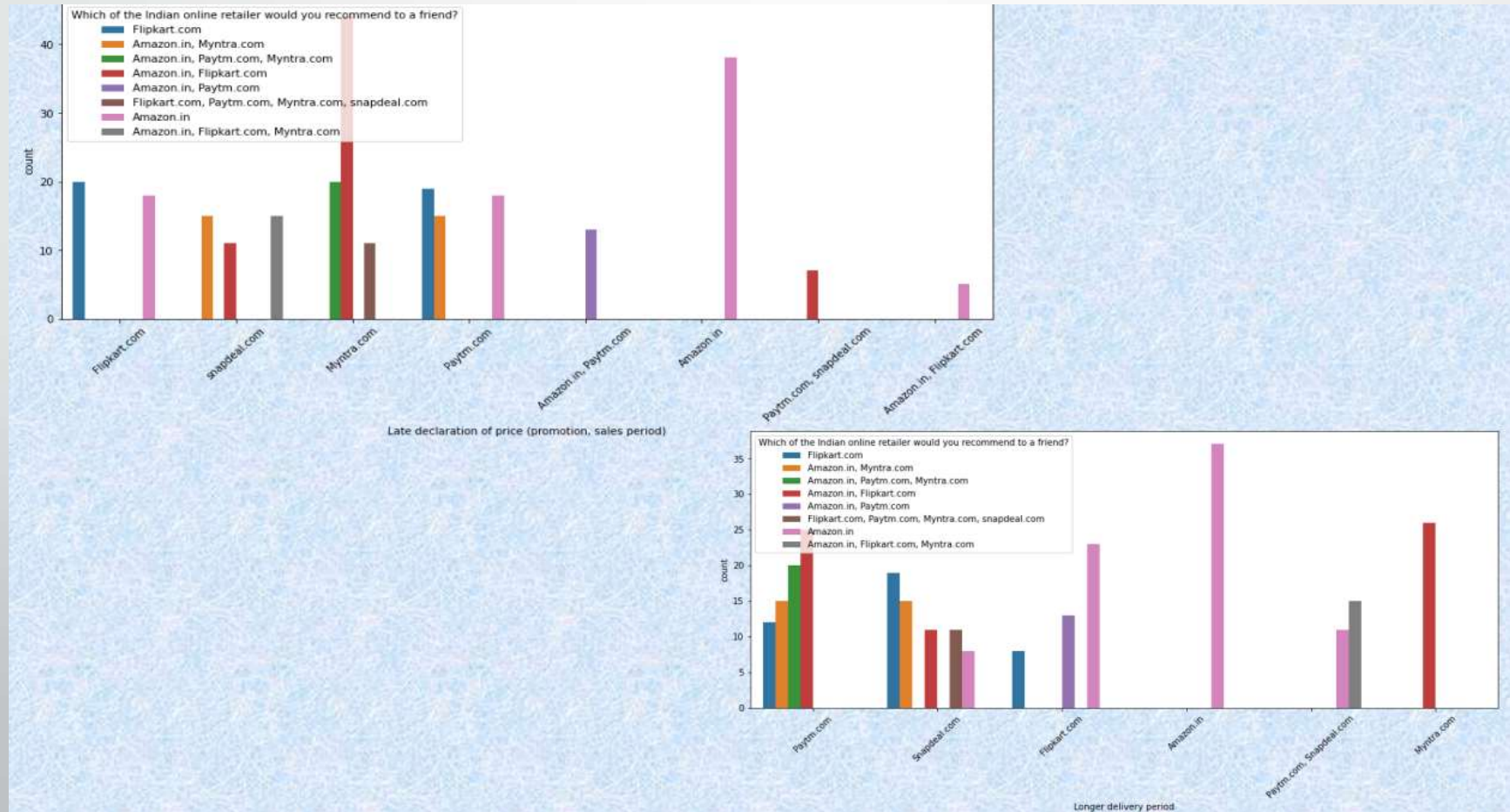
Conclusion

The results of this study suggest following outputs which might be useful for E-commerce websites to extend their business

• The cost of the product, the reliability of the E-commerce company and the return policies all play an equally important role in deciding the buying behavior of online customers. The cost is an important factor as it was the basic criteria used by online retailers to attract customers. The reliability of the E-commerce company is also important, as it is even required in offline retail. It is important because customers are paying online, so they need to be sure of security of the online transaction. The return policies are important because in online retail customer does not get to feel the product. Thus, he wants to be sure that it will be possible to return the product if he does not like it in real. Whereas, the logistics factor, which included Cash on delivery option, One day delivery and the quality of packaging plays a secondary role in this process though these are Must-be-quality. This is so because these all does not interfere with the real product and people believe that this is the basic value that E-commerce websites provide.

• All the websites were not equally preferred by online customers. Amazon was the most preferred followed by Flipkart. This can be explained easily by previous result that we got. These two companies are most trusted in the industry and hence, have a huge reliability. Also, the sellers listed on these websites are generally from Tier 1 cities as compared to Snapdeal and PayTM which have more sellers from tier 2 and 3 cities. Also, these websites have the most lenient return policies as compared to others and also the time required to process a return is low for these.

Thank you