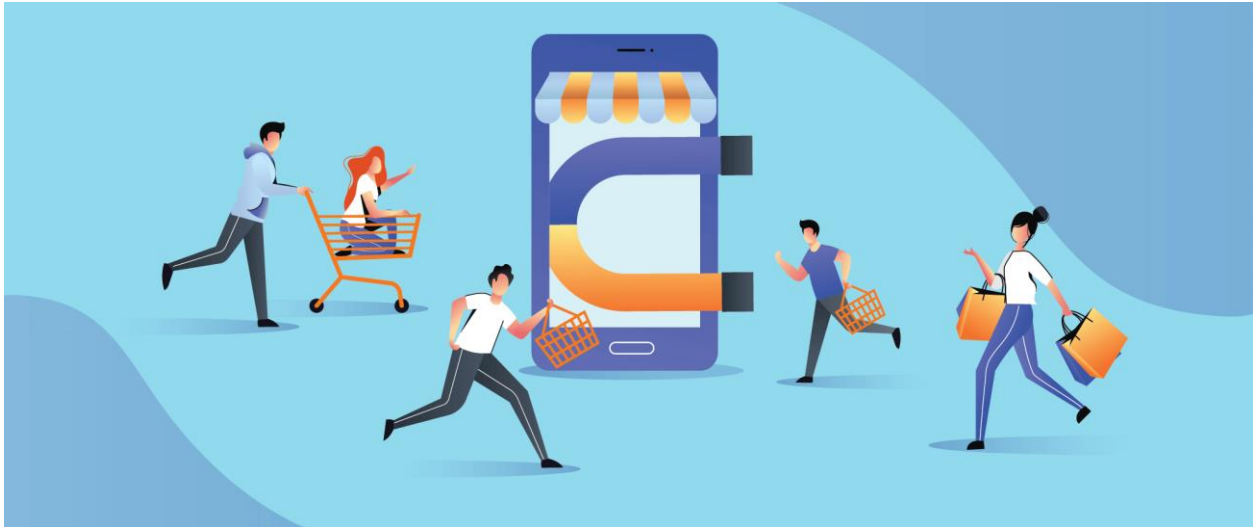




PROJECT REPORT



Data Analysis On Customer Retention in E-Commerce

Submission by
Himaja. I

ACKNOWLEDGMENT

I would like to express my sincere gratitude to my mentors from DataTrained and FlipRobo for helping me work on this project. Their timely and valuable inputs have helped me in the completion of this project successfully.

Finally, I would like to thank my family who have always been a great support to me and have been very helpful in various stages of project completion. Without their support it would not be possible to pursue my PGcourse and internship.

References:

- https://www.researchgate.net/publication/285927801_Customer_Retention
- <https://www.bigcommerce.com/blog/build-an-ecommerce-website/#how-to-build-an-ecommerce-website-step-by-step>
- https://www.researchgate.net/publication/301706084_Is_Customer_Retention_Beneficial_for_Customers_A_Conceptual_Background
- <https://blog.accessdevelopment.com/the-top-ten-benefits-of-customer-retention>

TABLE OF CONTENTS:

1. Introduction

- Business Problem Framing
- Conceptual Background of the Domain Problem
- Review of literature
- Motivation for the Problem Undertaken

2. Analytical Problem Framing

- Mathematical/ Analytical Modelling of the Problem
- Data Sources and their formats
- Data Pre-processing Done
- Hardware & Software Requirements & Tools Used

3. Model/s Development and Evaluation

- Identification of possible problem-solving approaches (methods)
- Visualizations

4. Conclusions

- Key Findings and Conclusions of the Study
- Learning outcomes of the study in respect of data science
- Limitations of this work and Scope for Future Work

INTRODUCTION

With the rapid global growth in electronic commerce (e-commerce), businesses are attempting to gain a competitive advantage by using e-commerce to interact with customers. Retaining customers in highly competitive business environments is critical for any company's survival because a lost customer represents more than the loss of the next sale. The company might lose all future sales and profits from that customer's lifetime of purchases.

Environment is critical for any company's survival

Nowadays, online shopping is a fast-growing phenomenon. Growing numbers of consumers shop online to purchase goods and services, gather product information or even browse for enjoyment. Online shopping environments are therefore playing an increasing role in the overall relationship between marketers and their consumers. That is, consumer purchases are mainly based on the cyberspace appearance such as pictures, images, quality information and video clips of the product, not on the actual experience. Due to this internet penetration, customers can just sit at their home, place their orders, pay via credit card, and wait until the goods are delivered to their home and they can easily shop their needy products via e-commerce rather than in physical stores.

Business Problem Framing:

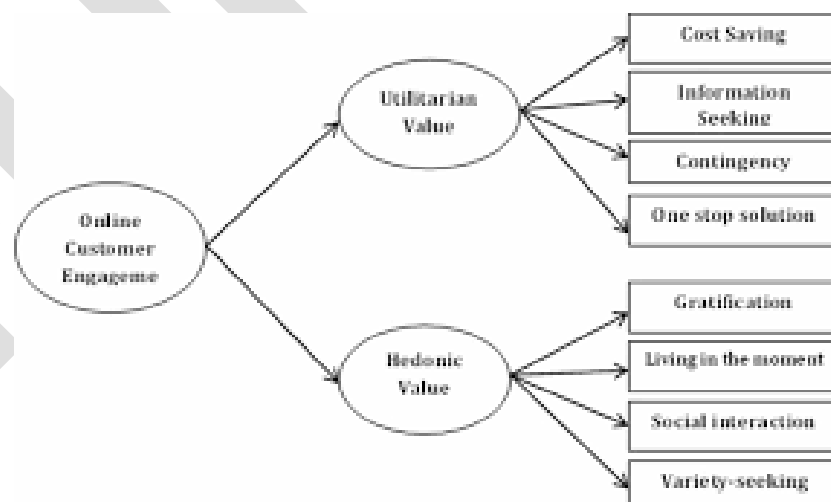
The online purchasing rate is increasing day by day. Customers' satisfaction is most important for any business development. The first question that comes into our mind is "What individual factors motivate a buyer to purchase online as well as restrain to purchase online?". We will discuss about this in the data analysis part. The main aim of this project is to explore consumer perceptions of internet-based e-retailing.

Customer satisfaction has emerged as one of the most important factors that guarantee the success of an 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.

Utilitarian value: Utilitarian means relating to practical purposes, or relating to the idea that usefulness is more important than beauty. When individuals are deciding what to do for themselves alone, they consider only their own utility. Such values are called as 'Utilitarian values'.

Hedonistic value: Hedonic value is defined as that value a customer receives based on the subject experience of fun and playfulness. The degree to which a product/service arouses emotions and creates pleasant experiences.



Conceptual Background of the Domain Problem:

The problem statement examined how customers form expectations on technology based self-service quality and suggested five main attributes of ecommerce store quality, that are service quality, system quality, information quality, trust and net benefit. The result of the study shows that, ease of use and enjoyment were also significant determinants of service quality, while speed of delivery and reliability had positive impact on service quality.

Many businesses focus on customer loyalty programs to remain competitive in today's marketplaces and explore innovative ways to keep their existing customers engaged. Customer retention is a simple concept. Yet, it often requires effort from every department to improve the customer experience and build lasting trust. So, what is customer retention and why it is so important. Let's know about it.

What is Customer Retention?

The customer retention is the process of engaging existing customers to continue buying products or services from their business. The goal of customer retention is retaining as many as customer as possible in the company.

One of the best ways to understand customer retention is to explore how to calculate customer retention rate for a given period.

How to Measure Customer Retention?

The simplest way to measure a customer retention rate is to look at the number of new customers acquired, number of previous customers, and total number of customers at the end of any chosen period.

Retention rate = $((\text{No of Total Customers} - \text{No of New Customers}) / \text{No of Previous Customers}) / 100$

The calculation provides a percentage that indicates how many customers have been retained over the period.

Why is Customer Retention Important?

- Customer retention is an express route to a business' financial success. It allows the companies to build long term, meaningful relationships with customers.
- It empowers customers to share feedback with the company team.
- It helps you understand how loyal and satisfied your customers are, how strong your customer service is, and your products are really worth their money or not.
- It enhances the brand reputation and understands future needs of the customers.

Review of literature:

This literature review is the integrated form of relationship marketing, customer loyalty, trust, reliability and privacy policy concept because all of these concepts are inter linked with each other.

The customer retention is possible only when the organization fulfil customers' needs and their expectations. A buyer goes through five stages while making a decision to purchase like problem recognition, information search, and evaluation of alternatives, purchase decision, and purchase behavior. These stages are best explainable when a buyer goes for buying costly items on online stores, such as a diamond jewelry, home appliances etc. However, in day-to-day purchase, consumers may not go through all these stages, since some commodities do not need information, and based on buyer's previous experience they will visit a store for the purchase.

In addition, companies are now using the Internet to build closer relationships with consumers and marketing partners. Besides competing in traditional marketplaces, they now have access to exciting new market spaces. Moreover, through the internet, consumers can learn about design, order, and pay for products and services, without ever leaving home and then, through the wonders of express delivery, they can receive their purchases in short time. From virtual reality displays that test new products to online virtual store that sell them, the technology explosion is affecting every aspect of marketing.

The internet has now become a truly global phenomenon.

Consumer's satisfaction and purchase decision on online shopping depends on few more issues. These are, e-store image, delivery and customers services, service quality and purchase behavior personalization, motivations for online purchase, trust, reliability, privacy, transaction and cost, incentive programs, web-site design, online interactivity, merchandise motivation, assurance, convenience, responsiveness etc. Following from the above consideration, this project will find out what individual factors motivate a buyer to purchase online and restrain to purchase online.

Motivation for the Problem Undertaken:

The main objective behind this problem is to help the ecommerce websites to find the e-retail factors for customers' activation and retention. The online shopping was considered by almost everyone to be convenient. The store is always open and it is easy, less stress and time saving to shop from home. Participants also seemed to agree that internet shopping gives access to a large range of stores and products from the entire world and offers best price to their consumers. So, in this project we will be analyzing how these online stores attracts the customers and whether the customers satisfied by their products or not.

ANALYTICAL PROBLEM **FRAMING**

Mathematical/ Analytical Modelling of the Problem:

The main thing that I found in problem statement is, the data given for us is unsupervised data. The problem statement contains both utilitarian value and hedonic value. I have performed both univariate and bivariate analysis to analyze these values using different plots like pie plot, count plot, distribution plot, factor plot etc. These plots give better pattern for analyzing the data. In this project I have done various mathematical and statistical analysis such as describing the statistical summary of the columns in which I found that the count is same for all the columns which means no null values present. Since the dataset contains object data type, I used label encoding method to convert the object data into numerical data. Checked for correlation between the features and visualized it using heat map.

Data Sources and their formats:

- The data is collected from the Indian online shoppers. Results indicate the e-retail success factors, which are very much critical for customer satisfaction. The dataset is provided by FlipRobo which is in the format xlsx.
- I have used detailed datasheet for the processing.
- The dataset contains 269 rows and 71 columns which is comprised of categorical columns. All the features contain object data type except the feature Pin code (unique values in integer data type).
- While describing the data I found skewness and outliers present in the columns. Since all the columns are categorical they need not be removed.

Data Pre-processing:

- Firstly, I have uploaded the dataset in the jupyter notebook.
- Then I have imported the necessary libraries and dataset.
- Checked the dimension of the dataset, data types of the features. Checked the information of the data frame using info() and checked the number of unique values present in the dataset.
- Since the columns had no proper names, so I renamed them with new name as below.

```
# Checking the new column names after renaming them
df.columns

Index(['Gender', 'Age', 'Shopping_City', 'Pincode', 'Shopping_Since',
      'Shopping_Frequency', 'Internet_Accessibility', 'Device_Used',
      'Screen_Size', 'OS', 'Browser_Used', 'Channel_First_Used', 'Login_Mode',
      'Time_Explored', 'Payment_Mode', 'Abandon_Frequency', 'Abandon_Reason',
      'Content_Readability', 'Similar_Product_Info', 'Seller_Product_Info',
      'Product_Info_Clarity', 'Navigation_Ease', 'Loading_Processing_Speed',
      'User_Friendly_Interface', 'Convenient_Payment_Mode',
      'Timely_Fulfilment_Trust', 'Customer_Support_Response',
      'Customer_Privacy_Guarantee', 'Various_Channel_Responses', 'Benefits',
      'Enjoy', 'Convenience', 'Return_Replacement_Policy',
      'Loyalty_Programs_Access', 'Info_Satisfaction',
      'Site_Quality_Satisfaction', 'Net_Benefit_Satisfaction', 'Trust',
      'Product_Several_Category', 'Relevant_Product_Info', 'Monetary_Savings',
      'Patronizing_Convenience', 'Adventure_Sense', 'Social_Status',
      'Gratification', 'Role_Fulfilment', 'Money_Worthy', 'Shopped_From',
      'Easy_Web_App', 'Visually_Appealing_WebApp', 'Product_Variety',
      'Complete_Product_Info', 'Fast_WebApp', 'Reliable_WebApp',
      'Quick_Purchase', 'Payment_Options_Availability', 'Fast_Delivery',
      'Customer_Privacy_Info', 'Financial_Security_Info',
      'Perceived_Trustworthiness', 'Multichannel_Assistance',
      'Long_Login_Time', 'Long_Display_Time', 'Late_Price_Declare',
      'Long>Loading_Time', 'Limited_Payment_Mode', 'Late_Delivery',
      'WebApp_Design_Change', 'Page_Disruption', 'WebApp_Efficiency',
      'Recommendation'],
      dtype='object')
```

These are the new column names in the dataset.

- I have checked for null values and found no null values present in the dataset and visualized it using heat map.
- Checked the value count of each column. By checking value count function, I found some of the columns contains duplicate and unwanted entries, so I have replaced those using appropriate values.
- Visualized each feature using seaborn and matplotlib libraries by plotting count plot, pie plot, distribution plot, box plot and factor plot.
- Performed label encoding.
- Checked the skewness and outliers. Since all the columns are categorical, so I haven't removed skewness and outliers.
- Checked the correlation between the features and visualized it using heat map.
- Described the data using describe().

Hardware & Software Requirements & Tools Used:

Hardware required:

- Processor: core i5 or above
- RAM: 8 GB or above
- ROM/SSD: 250 GB or above

Software required:

- Anaconda 3- language used
- Python 3Microsoft Excel

Libraries:

The important libraries that I have used for this project are below.

```
#importing required Libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings("ignore")
```

import numpy as np:

It is defined as a Python package used for performing the various numerical computations and processing of the multidimensional and single dimensional array elements. The calculations using Numpy arrays are faster than the normal Python array.

import pandas as pd:

Pandas is a Python library that is used for faster data analysis, data cleaning and data pre-processing. The data-frame term is coming from Pandas only.

import matplotlib.pyplot as plt and import seaborn as sns:

Matplotlib and Seaborn acts as the backbone of data visualization through Python.

Matplotlib: It is a Python library used for plotting graphs with the help of other libraries like Numpy and Pandas. It is a powerful tool for visualizing data in Python. It is used for creating statistical interferences and plotting 2D graphs of arrays.

Seaborn: It is also a Python library used for plotting graphs with the help of Matplotlib, Pandas, and Numpy. It is built on the roof of Matplotlib and is considered as a superset of the Matplotlib library. It helps in visualizing univariate and bivariate data.

From sklearn.preprocessing import LabelEncoder:

There are several encoding techniques like Label Encoder, OneHotEncoder, and Ordinal Encoder. In this project I have used Label Encoder technique to convert categorical data or object type data into numerical data.

MODEL/S DEVELOPMENT AND EVALUATION

Identification of possible problem-solving approaches (methods):

I have used “.replace()” function to replace unwanted entries in the columns.

Used “Label Encoder” method to encode the categorical features into numerical.

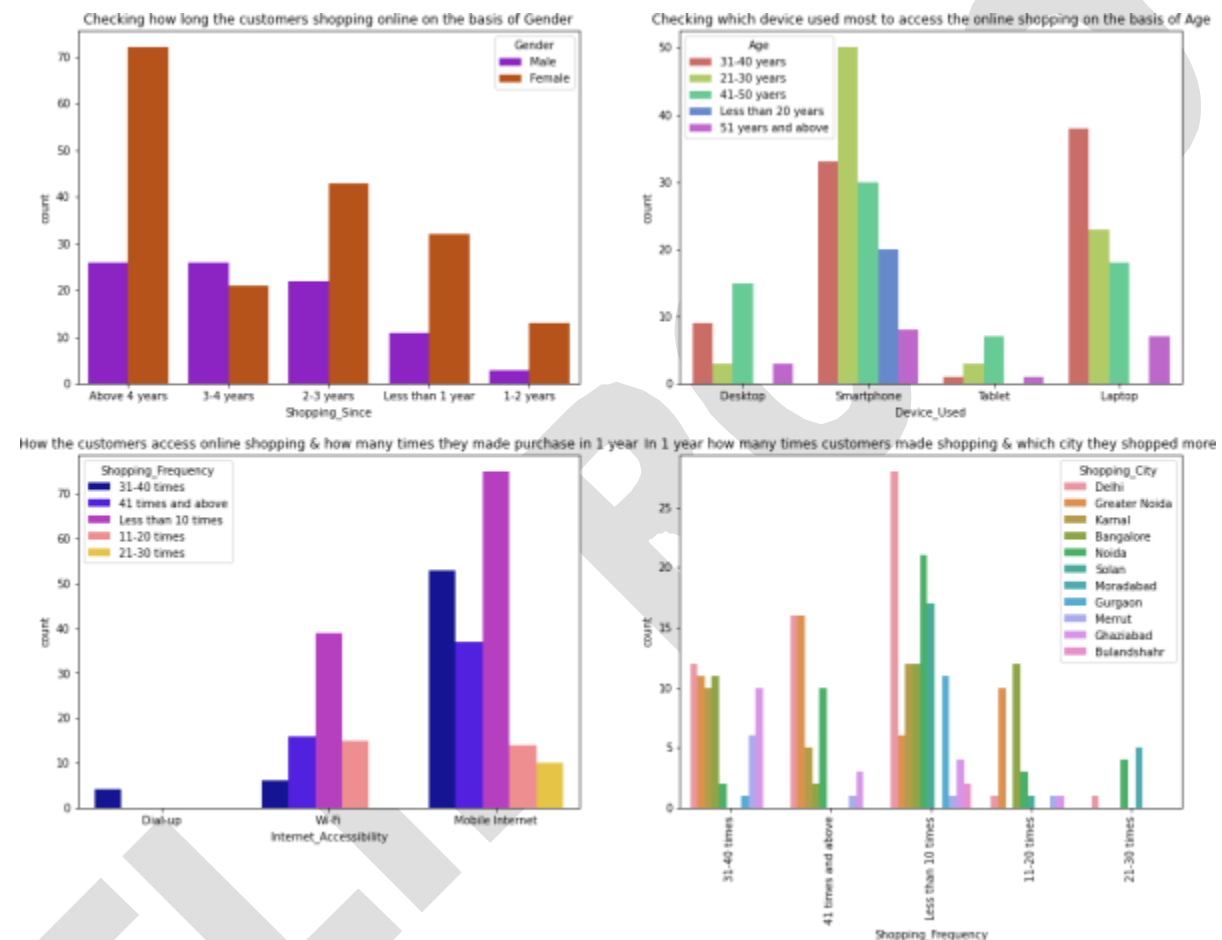
Described the statistical details of the features using “.describe()” method.

Used “Pearson’s method” to check the correlation between the features.

Performed both univariate and bivariate analysis using seaborn and matplotlib.

Visualizations:

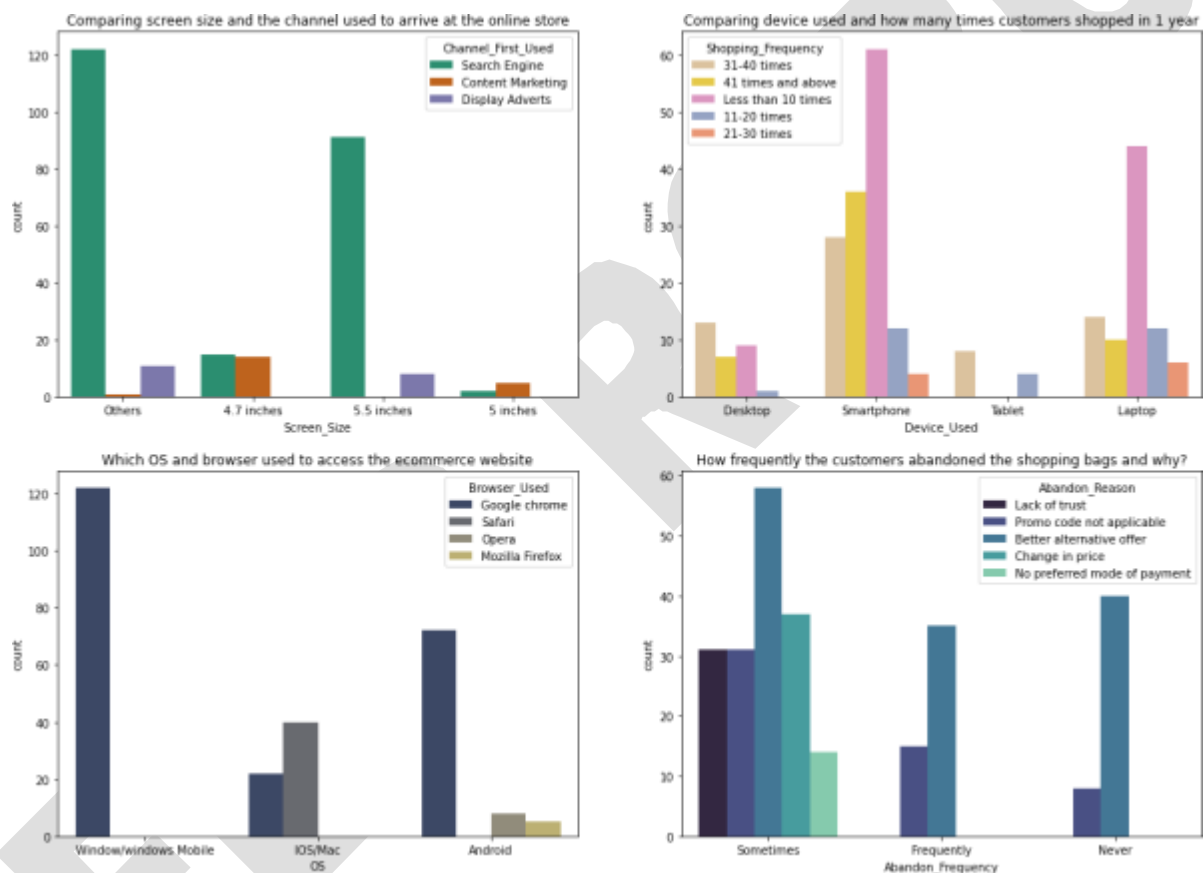
I have performed both univariate and bivariate analysis to visualize the data. In univariate analysis I have used pie plots, count plots and distribution plot and in bivariate analysis I have used count plot, factor plot and boxplots. Here I will be showing only bivariate analysis plots to analyze the data.



Observations from the above graphs:

- Most of the female customers shopped online from more than 4 years and the count is also high for the females who shopped from 2-3 years. And only few male customers shop online more than 4 years. Which means the female customers are more enthusiastic to buy products from the online shopping websites.
- Many customers whose age between 31-40 years and 21-30 years used Smartphones followed by Laptops to access the online shopping websites.

- Most of the customers access the shopping websites more than 31-40 times in 1 year through Mobile Internet to shop the products also most of the customers who used mobile internet to access the online shopping website made online purchase less than 10 times in a year. And only few of the customers used WIFI network to access the shopping store.
- Most of the customers used ecommerce websites less than 10times in a year from the city Delhi to shop the products.



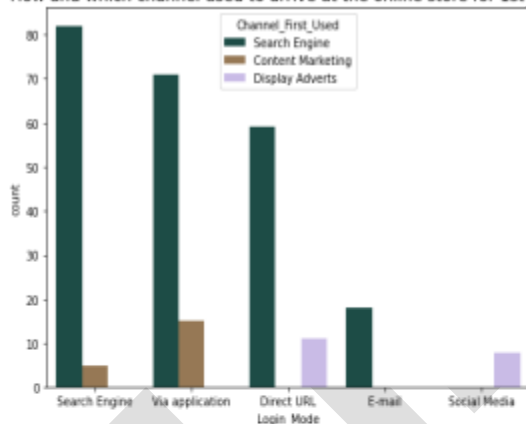
Observations from the above plots:

- The customers having their mobile screen size say 6 inches (others) have followed search engine channel to arrive at their favorite online store for the first time. Also, the customers who have their screen size 5.5 inches also used search engine channel to access the online shopping store.

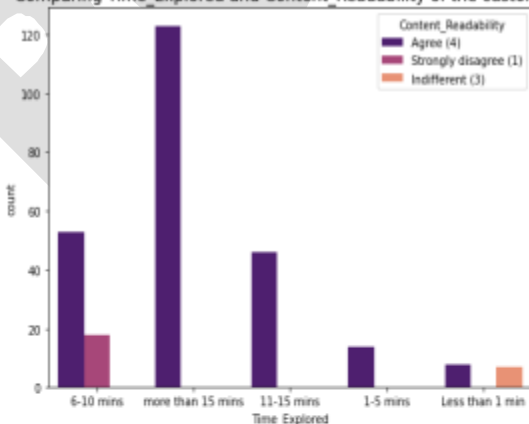
- Most of the customers used Smartphones 31-40 times in a year to access the ecommerce websites to shop the products.
- Many customers having windows operating system in their device ran Google chrome to access the ecommerce shopping websites and some of the customers having IOS/Mac operating system used Google chrome as well as Safari to reach the online shopping store.
- Due to Lack of trust on the ecommerce websites, sometimes most of the customers abandoned the websites and some of the customers abandoned the shopping website due to the promocode not applicable. which means, if the product is having the special price or some catalogue price rule is applicable on it. Then coupon code should not be applicable on the products.

So, it is important for the ecommerce companies to create discount price, offers, and coupon codes to retain the customers.

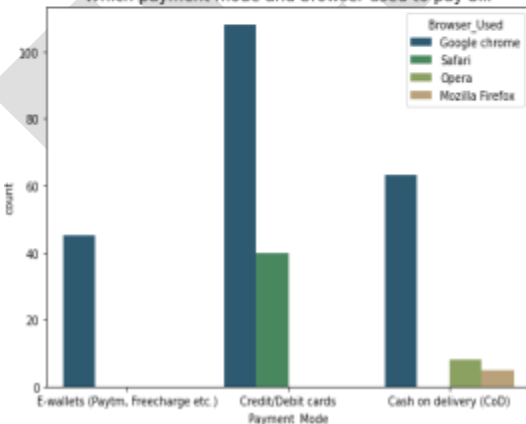
How and which channel used to arrive at the online store for 1st time



Comparing Time Explored and Content Readability of the customers



Which payment mode and browser used to pay bill

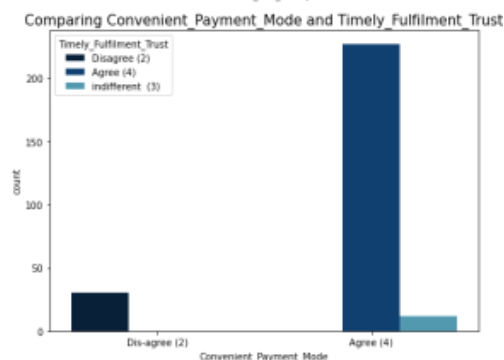
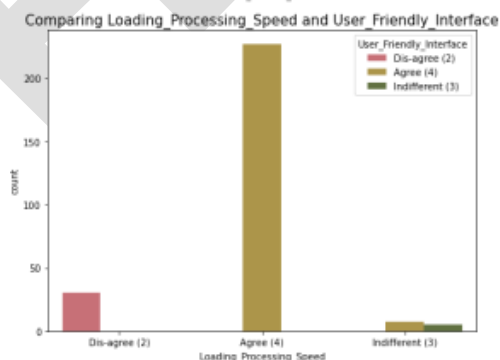
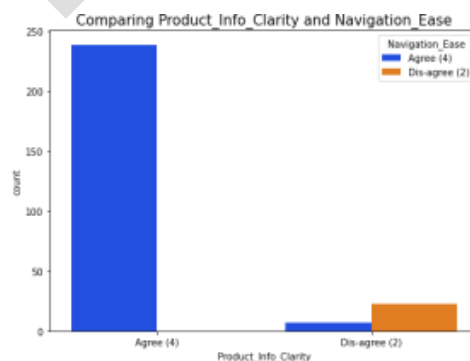
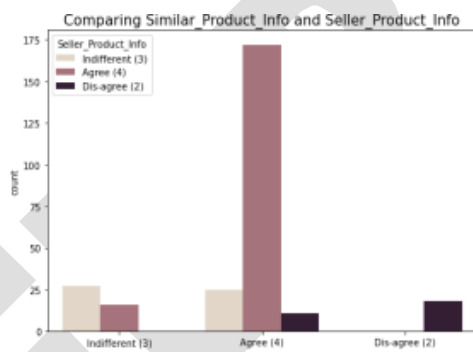


How frequently the customers abandoned the shopping bags while paying the bill?



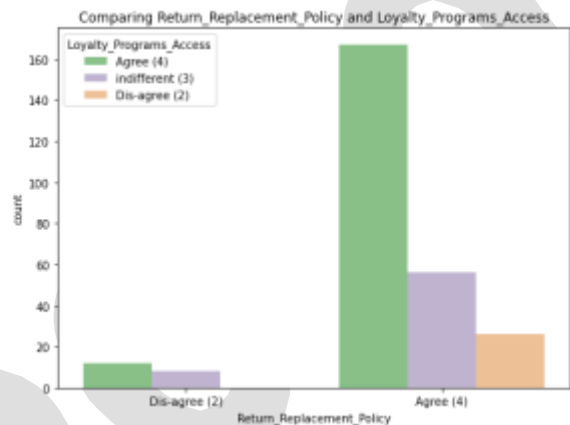
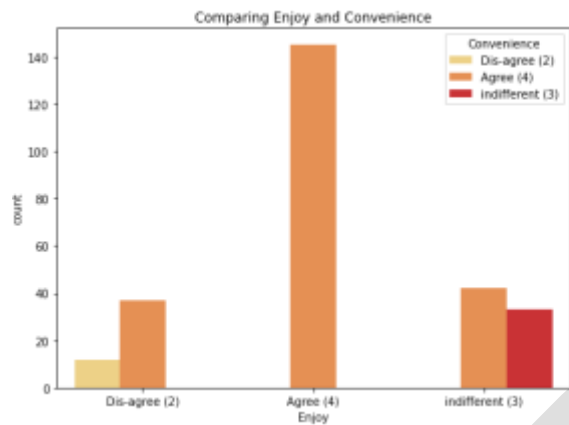
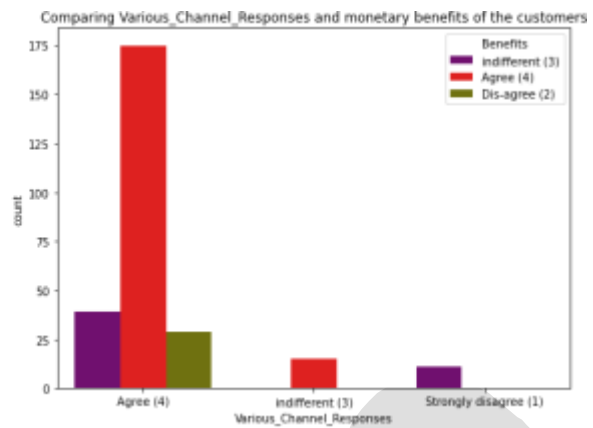
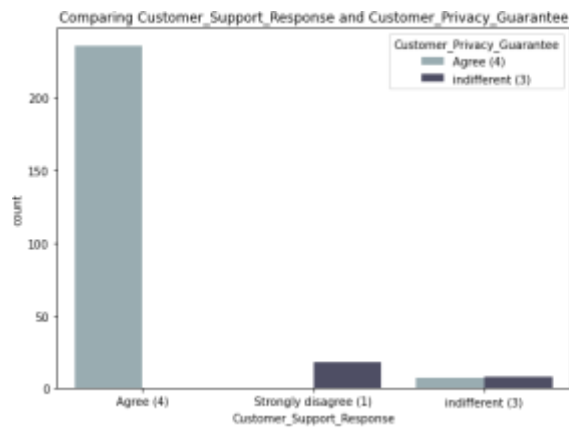
Observations from the above plots:

- Search engine is the most used channel by the customers to arrive their favorite store for the first time and after visit the Website for the first time, most of them used the same channel to reach the online retail store to re-shopping the products.
- Most of the customers agreed that the content on the website is easy to read and understand also they explored more than 15 mins before making the purchase decision and some of the customers strongly disagreed that the content is not good and they explored 6-10 mins before making the purchase decision. So, ecommerce websites should enable some images and it should contain clear structure, so that the customers can easily read and understand the content of the product.
- Most of the customers used google chrome to reach the websites and they preferred to pay their product price using Credit/Debit cards and only few of the customers used Safari browser to reach the e-retail websites.
- Sometimes the customers used to abandon their selected items and wants to leave without making payment and most of them making the payment using E-wallets methods.



Observations from the above plots:

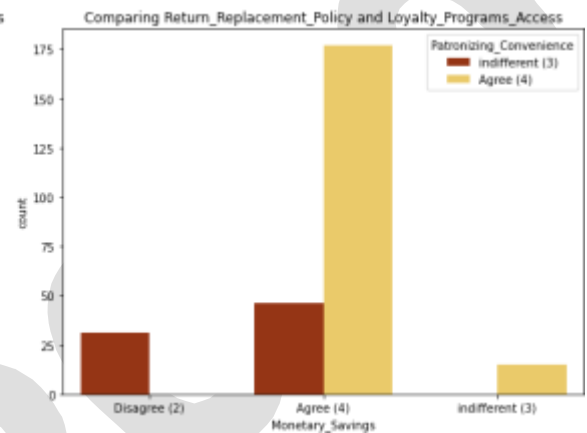
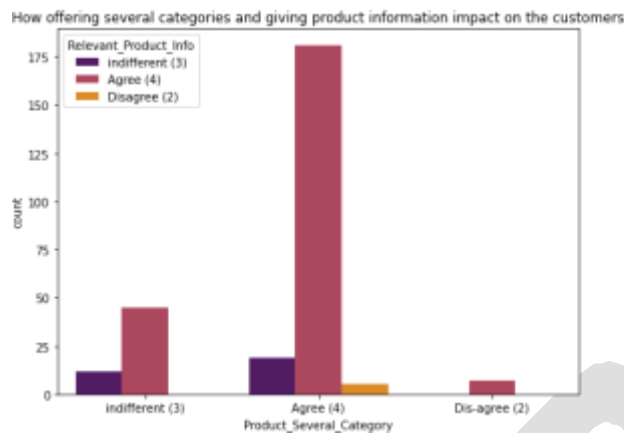
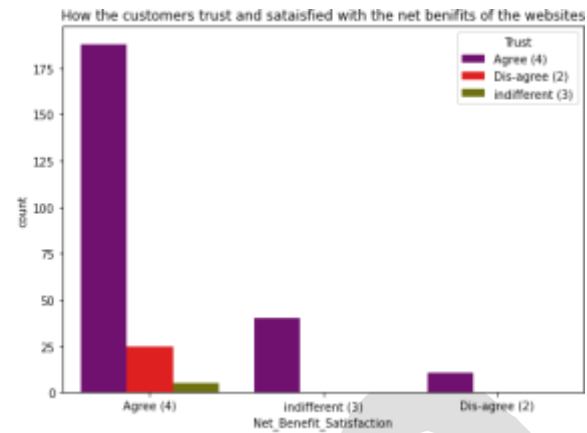
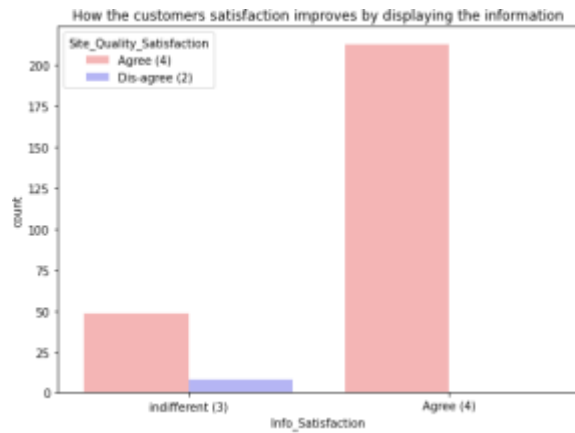
- ✚ Most of the customers agreed that the information on similar product to the one highlighted is important for product comparison and also Complete information on listed seller and product being offered is important for purchase decision. In order to buy a product, the ecommerce website must give the complete information about the product and seller information then only the customers can compare the product costs and its details in different websites and they tend to buy that particular product in a particular website.
- ✚ Around 90% of the customers agreed that they should be able to navigate the website easily and the products information in the website must be clearly stated their uses, lifetime, benefits etc. Then only more customers tend to buy those products and can shop easily.
- ✚ Most of the customers agreed with the user-friendly interface of the websites which can be easily loaded and processed also these websites' loading and processing capacity is very fast so that the customers like to shop in ecommerce websites. If these websites do not have this much of loading and processing speed then customers don't want to buy the products in this website and they tend to other websites or other options rather than this.
- ✚ Most of the customers agree to the trust that the online retail stores will fulfil its part of the transaction at the stipulated time also most of them very happy with the convenient payment modes given by the websites. In other words, the websites must provide all the possible ways of payment methods then only the customers shop frequently all the time the mode of the payment for customers may not possible sometimes they may choose cash on delivery. So, if the retailers provide all type of payment methods then the customers can easily make the payment also it enhances the sales of the ecommerce sites. And the transaction must also be given with some stipulated time otherwise the payments may be failed so they've to provide minimum amount of time which need to be fixed for all.



Observations from the above plots:

- Almost all the customers agreed that ecommerce websites have empathy towards them and these sites being able to guarantee the privacy of the customers. That is the online retailers must be able to resolve all the queries of the customers and they have to assure the customers keeping all their credential secured and should not share with others. If the websites give guarantee about the privacy, then the customers make shopping regularly which will enhance the company's sales.
- Most of the customers agreed that the online shopping gives monetary benefits and responsiveness, availability of several communication channels will help them more while shopping online which means if one channel is not available then customers can easily reach out to other channel to fulfil their benefits. So, it is important for the online e-tailer companies to provide various channels to communicate with the customers. The ecommerce websites should ask the feedback regarding their services, ratings of the products, reviews etc and also, they try to communicate with the customers in different social platform.

- ✚ Most of the customers believed that they enjoy online shopping also shopping online is convenient and flexible and some of the customers who disagreed with the enjoyment of the shopping, they are not convenient with the online shopping. Some customers shops online for their enjoyment purpose they are termed to be hedonistic, for them shopping online gives experiential satisfaction. They contribute much for the ecommerce companies by buying all the costly products randomly.
- ✚ Most of the customers agreed that return and replacement policy of the e-tailer is important for purchase decision also gaining access to loyalty programs is a benefit of shopping online. Many return policies have conditional agreements, such as time limits, that must be clearly defined and expressed at the time of purchase or else the customers won't get the chance to return their damaged or dissatisfied products due to this they may not access the same website if they want to shop again. It is evident from the fact that the customers actually not liking the products completely, they are just purchasing the products and returning them in case of any dissatisfaction. So, it is important for the online shopping websites to make easy return and replacement policy if they want to retain their customers. Also, by gaining access to loyalty programs, the customers get more and more rewards, increasing their engagement rate and thus bringing more profit to both company and customer.

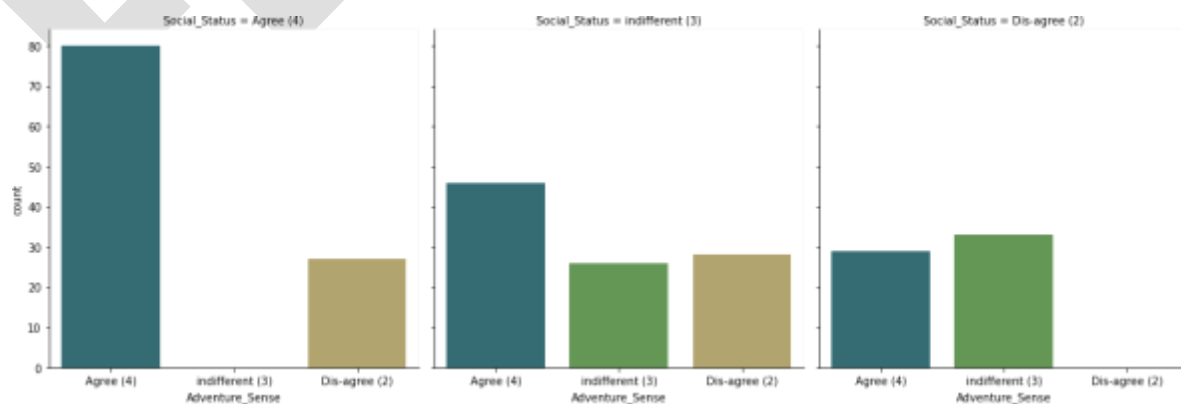


Observations from the above plots:

- Many customers agreed that displaying quality information on the website improves satisfaction of customers since they believe that displaying quality information have significant association with customer satisfaction and they are satisfied and happy while shopping on good quality websites. In order to obtain high levels of customer satisfaction, high service quality is needed, which often leads to favorable behavioral intentions also a website with good system quality, information quality, and electronic service quality is a key to success in e-commerce. So, the online retailers must display all the information about the product then only customers get an idea to buy the products regularly.
- Most of the customers agreed that net Benefit derived from shopping online can lead to users' satisfaction also they believe that user satisfaction cannot exist without trust. The e-tailer should provide crediting points (net benefits) so that the customers tend to buy frequently in order to gain points. Trust is

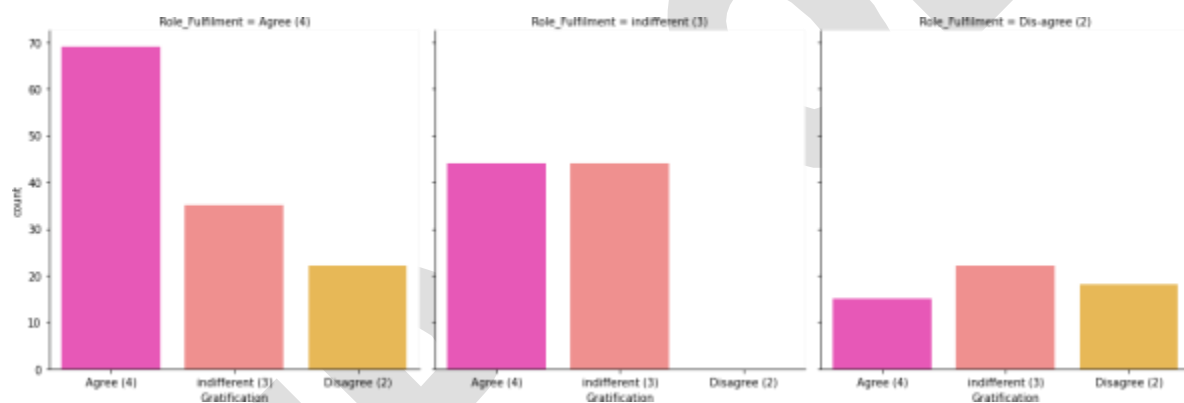
also a major factor for customers to decide whether to buy products from online stores or not also trust helps reduce uncertainty when the degree of familiarity between the customer and transaction security mechanism is insufficient. If customers have a high level of trust toward the website, it is more likely for them to have intention to purchase so it's important for the ecommerce website to make the customers get trust on them.

- The customers are more likely to purchase on the same websites if that website offers them a wide variety of products in several category and giving relevant information about the products. Having multiple product lines may allow to grow the ecommerce business and finding accurate and up-to-date information of the product must be stated clearly in the website so that the customers can buy the products without any confusion.
- In this digital and competitive world, everyone wants to save money, the ecommerce company need to know that the best way to sell online is to make the consumer feel that he is saving money doing so. And not just feel, online shopping should result in a lot of saving for the consumer. This saving would automatically get converted into trust and brand equity for the seller. To do this the online companies should offer the best deals and bargains to the consumer through social platforms. If the retailers give some discounted prices, then the customers can make money savings and they tend to purchase in the same websites regularly. Convenience is the important thing for ecommerce and most of the customers agreed with it.



Observations from the above plots:

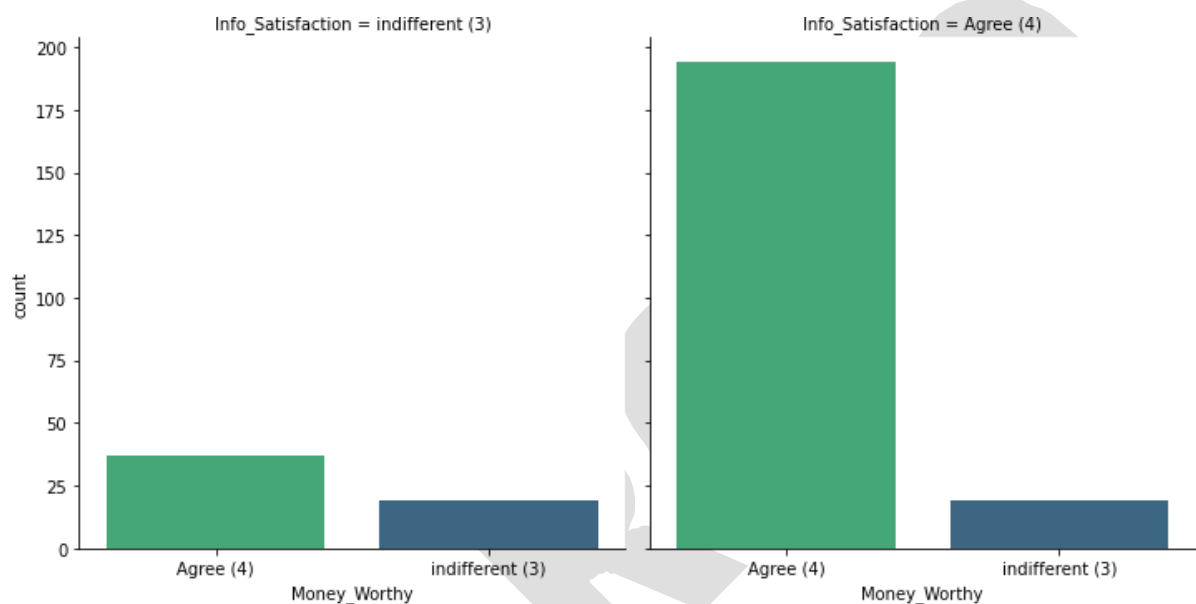
Most of the customers agreed that shopping on the website gives the sense of adventure. The adventures in the shopping websites gives positive activity to experience an amplified enjoyment to the customers while shopping on websites. They also believe that shopping on preferred e- tailer enhances the social status of the customers. Many customers think they are adventuring while shopping online as they search for low cost and high discount products to buy and prefer the same to the others. In this way they think that shopping in the website gives them the adventure. Shopping online won't affect anyone's status and the customers agreed that shopping on preferred e-tailer enhances their social status.



Observations from the above plots:

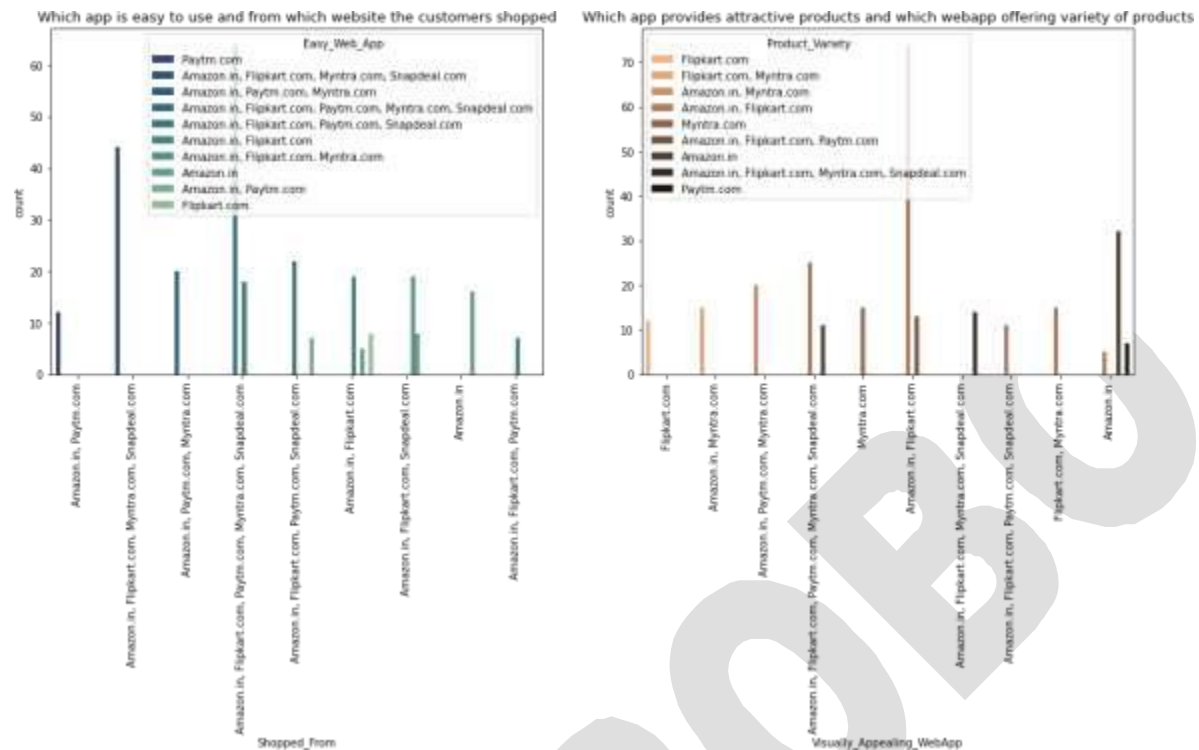
- Most of the customers agreed that they felt gratified while shopping on their favourite e-tailer. This is because the e-tailer companies can successfully make up for a mistake or a dissatisfied customer is to be equally expedient in addressing the customer's needs.
- Also, most of the customers agreed that shopping on online website helps them fulfil their certain roles. Fulfilment refers to activities that ensure customers receive what they ordered, including the time of delivery, order accuracy, and delivery condition, also the customers cannot see the product directly before they purchase it. Companies must ensure delivery timeliness, order accuracy, and delivery conditions to provide

superior service quality for customers. The companies must understand that the customer satisfaction is an indication of the customer's belief of the probability of a service leading to a positive feeling. If the companies give positive vibration to the customers by giving chance to fulfil their roles, then they shop more on that particular website.



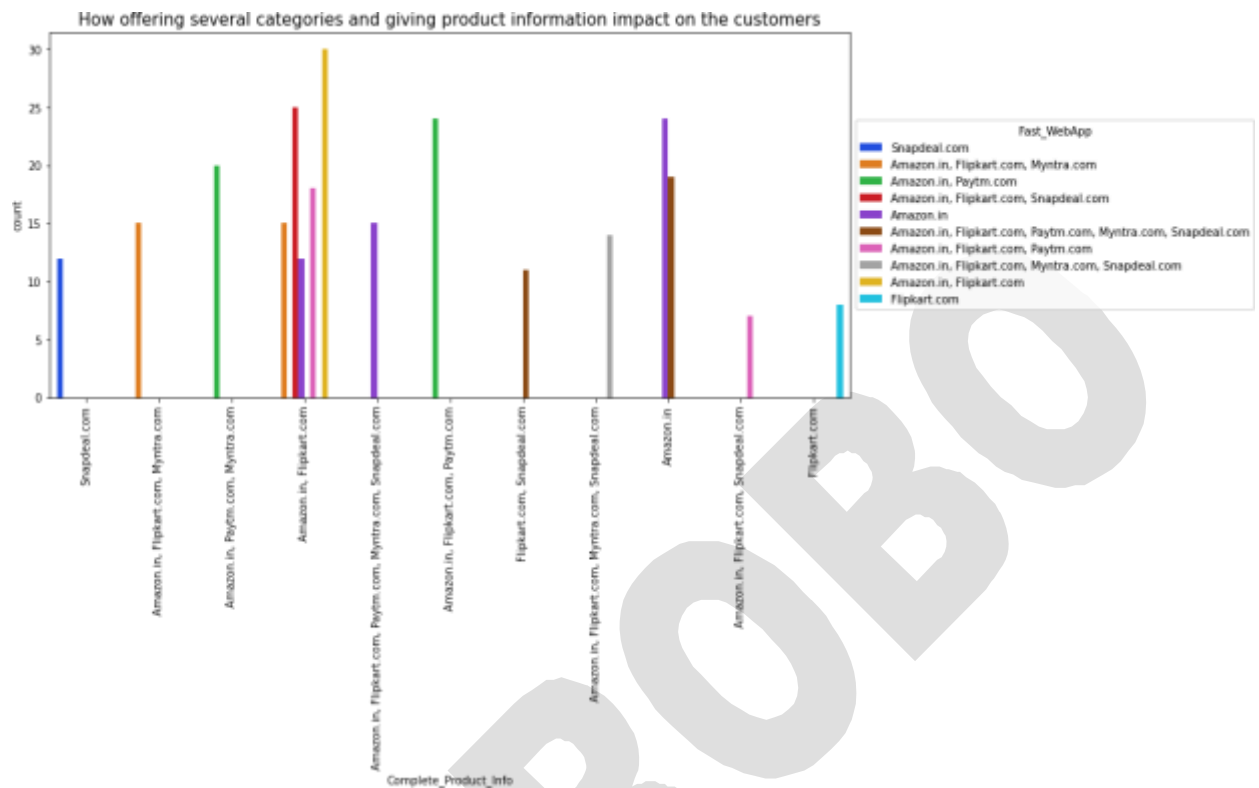
Observations from the above graph:

The customers should satisfy with their product that they shopped from the online store then only they agreed that they got value for the money they spent. The companies should display the quality information about the products so that the customers being able to purchase their product and thinks that it worth for money and this comes under utilitarian value.



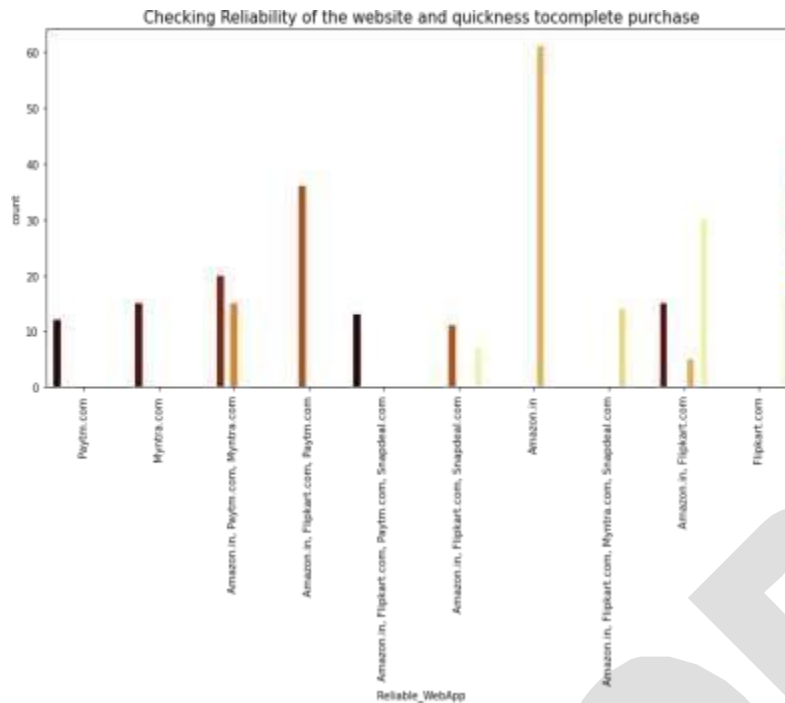
Observations from the above plots:

- There are many websites for selling the products among them Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com are easy to use and shop. Most of the customers used these websites more, this is because, these websites may provide less price products, good discounts and may have lots of varieties of similar products with different brand.
- Amazon and Flipkart have high visual appealing web-page layout compared to others that means these websites provides some colorful graphics on the homepage. The more people find the website attractive, there are higher chances that they will stay a little longer in that website, also these websites provide wild variety of products in an attractive manner which makes the customers to buy the product.



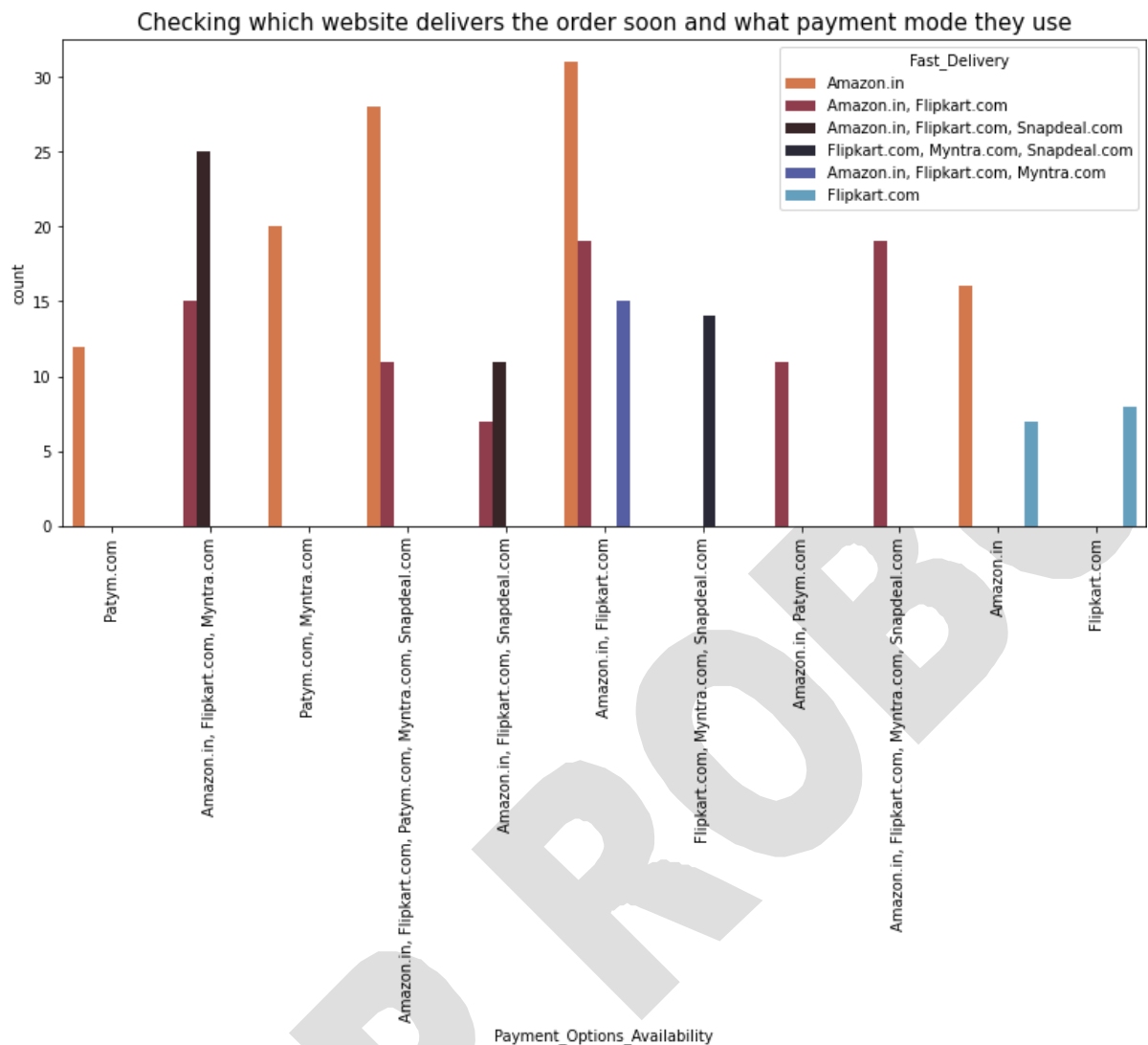
Observations from the above plot:

- The most common problem which is faced by the customers is server problem, in many websites these problems are common. Also, only few websites provide the relevant information about the products. These things may lead the customers to have bad impression on these websites.
- From the plot we can visualize that the amazon and flip kart websites gives complete and relevant information and these websites have no issue with the server and most of the customer liked the web speed of both amazon and flip kart.



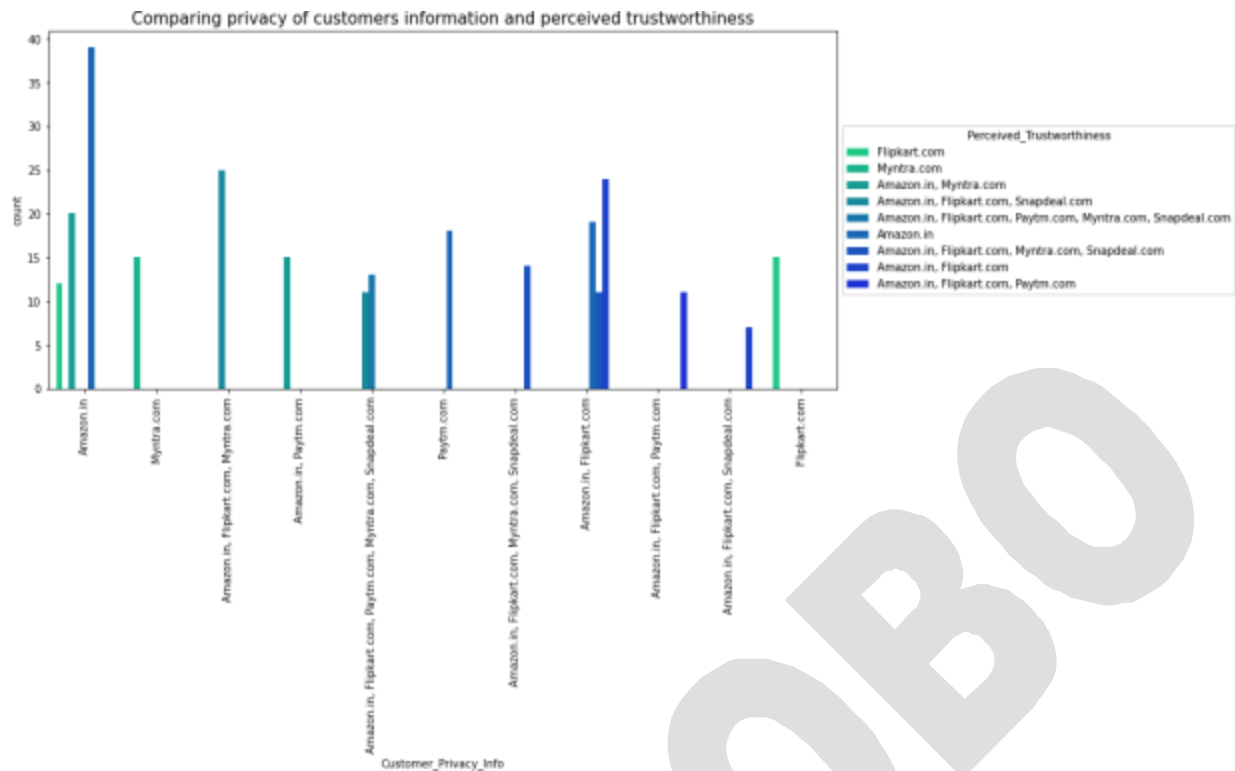
Observations from the above plot:

- The consumer determines the shop's reliability based on the information transmitted by the shop and certain sites offer customers the opportunity to purchase items that are used which means they are likely to be the most reliable. Some of the customers completes their purchase very quickly due to the discount, less price, free delivery charges etc provided by the ecommerce websites.
- From the plot we can notice amazon site is more reliable and most of the customers complete their purchase on amazon very quickly.



Observations:

- Having different types of payment methods will help the customers to pay the invoice easily using their choice of payment and if the websites have speedy delivery methods without delivery charge, then the customers like to buy the products in those websites.
- Here Amazon and Flipkart have several payment options and Amazon indeed has speedy order delivery compared to other websites.



Observations:

- Security/privacy refers to the security of credit card payments and privacy of shared information like name of the customer, address and phone number. Customers are always concerned whether the website would protect them against fraud after a transaction. So, website security and privacy are important to assess the service quality of online stores. The customers think that buying online means taking risk, in this case trust is more important thing for both merchant and customer.
- Most of the customers trusts amazon followed by flip kart in terms of keeping their privacy of data information secured and the customers who believes that amazon website keeps their financial information as secrete also trusts flip kart, Myntra, Snapdeal and Paytm in terms of keeping their financial information secured. Most of the customers believed that Amazon has perceived trustworthiness compared to others. Apart from this, customers believed that flip kart and Myntra also have perceived trustworthiness.

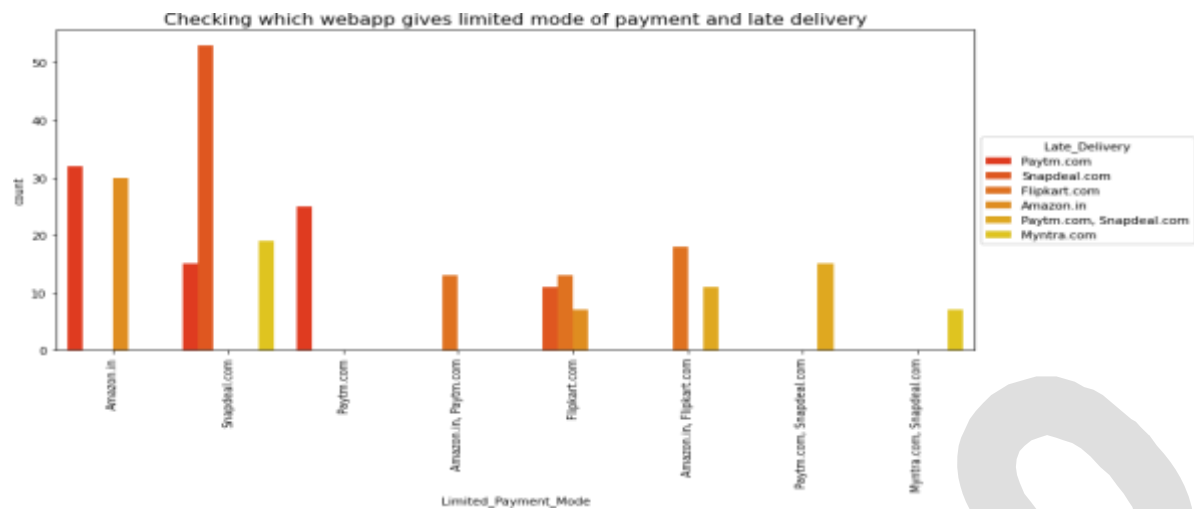
Observations from the above plot:

- The customers mostly choose amazon website for buying products as it gives promotions and sales periods in some days, on these days most of the customers attracted by the offers provided by the websites, wants to buy the products. So, amazon will take more time to allow the customers to get login into the site.
- When there is promotion or sales period, amazon and Myntra takes longer time to display the graphics and photos.

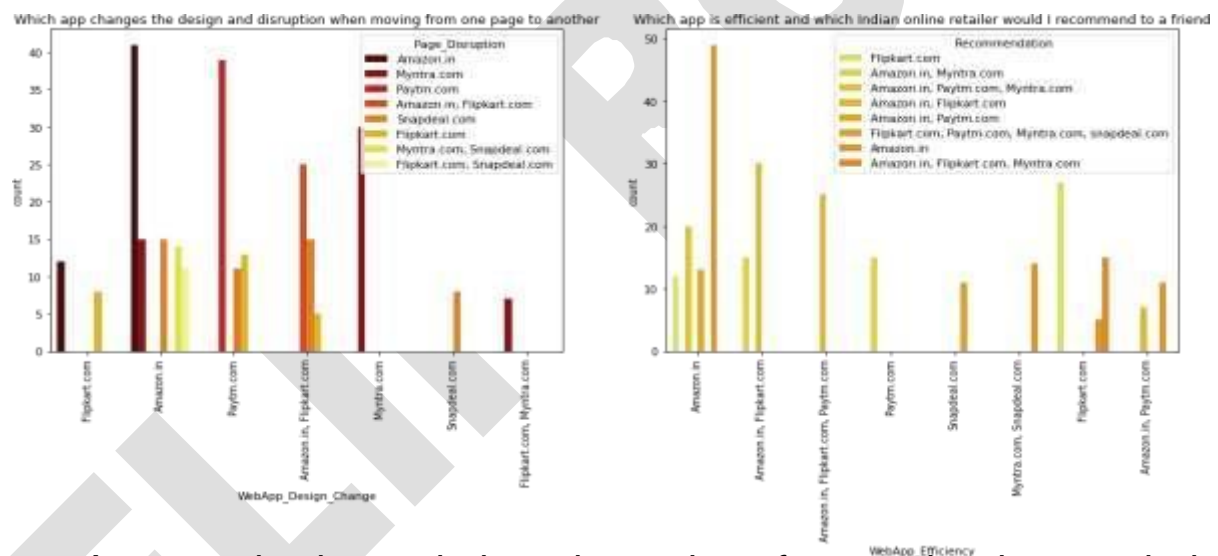


Observations from the above plot:

- When there is promotion and sales, Myntra takes time to load the page and it has late declaration of price in these days.
- Myntra declare the late price in order to clear the sales and they fix the price by comparing with other websites and they end up sales by providing benefits to the customers. In this time most of the customers tries to shop in this website so it takes long loading time.



Above count plot shows that Snapdeal has limited mode of payment on most of the products followed by Amazon. And Paytm takes more time to deliver the product. So, this website may not satisfy the customers due to late delivery.



- Amazon is the website where they frequently change their application designs in order to attract the customers and satisfies the customers' needs and they tend to make customers by updating everyday as per the trend. But the disadvantages of this website are when moving from one pageto other it slows down and sometimes it may shutdown.
- Amazon is the website which is more efficient as before and I suggest Amazon.com and Flipkart as a best Indian online retailer store for purchasing all types of products, as they provide enormous amounts of benefits.

Interpretation of the Results:

The results that were interpreted from the visualization are as follows:

From the survey we found that the customers agreed with certain things like ecommerce websites have empathy towards them, content must be easy to read and understand, similar products should be highlighted for product comparison, payment convenience, trustworthy and they felt gratified while shopping etc.

The respondents were asked to give ratings and feedback regarding certain statements that may be used to describe their ideal online store. We have found answers from customers like which websites among listed sites they are familiar with, they find easy to use, reliable, secured, about their delivery, change in web design, etc.

From the visualization we have found Amazon.in was the most used and favorite websites for the customers and Flip Kart was the second favorite web app while snapdeal.com and Paytm.com were the least used websites by the customers.

CONCLUSION

Key Findings and Conclusions of the Study:

Findings:

- ❖ In this project we have investigated ecommerce quality in online businesses and develop new knowledge to understand the most important dimensions of E-retail factor for customer activation and retention.
- ❖ This project aimed to enhance prior understanding of how ecommerce websites affected customer satisfaction, customer trust, and customer behavior, i.e., repurchase intention, customer loyalty, and site revisit.
- ❖ The dimensions like information about the products, convenient payment mode, Trust, Fulfilment, website design change, security/privacy and many others had a positive impact on the

ecommerce websites for customers.

- ❖ Thus, a company needs to pay attention to these dimensions more specifically and seek breakthroughs that can improve its performance and e-service quality.

Conclusion of the Study:

- ❖ The endeavor of this study is to identify the motivating factors towards online shopping and in which e-tailer the customers likely to shop more.
- ❖ In this project, I have done some feature engineering by replacing the unwanted entries by suitable values, found no null values, and renamed the columns by giving new names. Visualized the data using count plot, factor plot, pie plot and distribution plot, also encoded the object data into numerical using label encoding method. Checked the statistical summary of the dataset and checked for skewness, outliers and correlation between the features.
- ❖ From the analysis it was found that consumers purchasing decisions were dependent on various factors. All these motives motivate consumers to purchase products through online. According to consumers' opinions, "time saving" is the most important motivating factor for online shopping.
- ❖ Again "information availability", "open 24/7", "huge range of products/ brands", "reasonable prices", "various offers for online products", "easy ordering system", and "shopping fun" are other motivating factors for online shopping respectively. Also, "online payment system", "personal privacy or security issues", "delaying of delivery" and "lacks of personal customer service" are the main inhibitions of online shopping to the respondents.
- ❖ After visualizing the data, I found Amazon is the best online store where the customers trust on buying products and it has positive impact on the customers. Also, amazon and flip kart have increased customers' expectations. So, they are the best online retailer who makes the loyal customers and satisfies the customers.

- ❖ It was also observed that online shopping is not trustworthy and reliable to some consumers due to only online payment system and personal privacy. In addition, online security is a major concern for the consumer particularly in terms of fraud, privacy and hacking.

So, the organizers should make up their organization for better sales.

Learning outcomes of the study in respect of data science:

- ❖ Customer satisfaction and customer trust appeared as the outcomes of overall e-retail factor. The results of the analysis showed that e- retail factor had a positive impact on customer satisfaction. The majority of research done about e- retail factor states that customer satisfaction is the main determinant impacting on e-retail factor. It supports the idea that there is a significant relationship between e- retail factor and customer satisfaction. E-retail factor also had a positive impact on customer trust. The better the e-retail factor of a company, the higher the customer trust. Providing good service quality enhances customer satisfaction and customer trust.
- ❖ From the above analysis we found that the mains reasons or factors which attract consumers to do shopping online and then main reasons or obstacles which discourage consumers from shopping online. Therefore, from the analysis, it is found that most of the respondents use internet daily but most of the respondents do not use internet daily to buy products. Nearly half of the total respondents' opinions were that they would only use the internet to buy products when the need arises to do so.

Limitations of this work and Scope for Future Work:

Limitations:

- ❖ The limitation to the analysis is that there are more female customers who shop more on e-tailer websites compared to male customers, this is due to the minimum data. Since the data is very less so it's bit difficult to come to the conclusion on the retention rate.
- ❖ There was no dependent variable in the dataset. If the dataset contains target variable, then it gives good visualization and relation between the feature and target and the model prediction is also possible of the target variable present.

Recommendation to the online sellers (Scope for future work)

Based upon the findings and analysis, the following assumptions (recommendations to the online seller) are presented for the online sellers to make online shopping more popular, convenient, reliable and trustworthy.

- ❖ Transaction security and consumers data safety are principal concerns of online customers purchasing products or services online. Therefore, online vendors can assure their consumers' by offering personal information privacy, protection policy and guarantee for transaction security by improving their technological systems.
- ❖ Retailers should be careful about the annoying factors of online shopping such as being unable to access the website, long delays in completing online orders, inconsistencies in the items available online, mistakes in filling orders, and the hassle of returning goods.
- ❖ Online sellers can be more concerned about delivery times, delivery charge and product return policies. They can make it easier, quicker and reliable, so that consumers can enjoy the online shopping experience and they likes to shop in the particular websites regularly.
- ❖ The respondents thought that products' mixing up or they find different product at delivery time which is the main inhibition of online shopping, so that the sellers must be very cautious when it comes to delivery.
- ❖ Getting feedbacks from the customers is also on of the important

thing to improve the sales of the company. The e-tailer wants to keep the customer happy in order to build the successful business, but they easily fall into a trap of assuming that the customers will give feedback without being prompted. If the e-tailers are doing something wrong, most of the customers won't complain, they will just go elsewhere. So, it is important to ask customers how they really feel about their services.

FLIP PROBO