



Project Report On Customer Retention Analysis

(E-retail factors for customer activation and
retention: A case study from Indian
e-commerce *customers*)

Submitted By:
B. SHEEBARANI

ACKNOWLEDGMENT

I would like to express my sincere thanks of gratitude to my mentors from Data Trained Academy and Flip Robo Company for letting me work on this project. Their suggestions and directions have helped me in the completion of this project successfully.

References :

1. <https://www.researchgate.net/publication>
2. <https://blog.accessdevelopment.com/the-top-ten-benefits-of-customers-retention>
3. <https://www.researchgate.net/publication/301706084> Is Customer Retention Beneficial for Customers A Conceptual Background
4. <https://www.bigcommerce.com/blog/build-an-ecommerce-website/#how-to-build-an-ecommerce-website-step-by-step>

TABLE OF CONTENTS

1. Introduction

- Business Problem Framing
- Conceptual Background of the Domain Problem

2. Analytical Problem Framing

- *Mathematical/Analytical Modeling 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)*
- Visualization

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 customers' lifetime of purchases.

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 them 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:

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

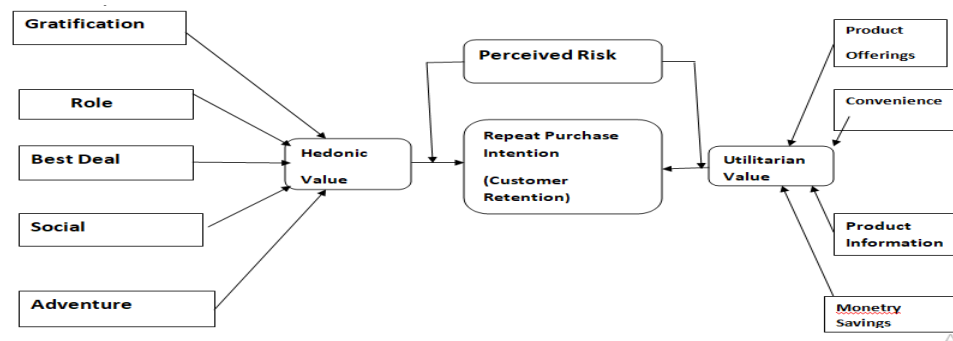
(loyalty) positively. The data is collected from the Indian online shoppers. Results indicate the e-retail success factors, which are very much critical for customer satisfaction.

Utilitarian Value:

Utilitarian value is an objective value which provides some functional benefits to the consumers and helps consumers to accomplish practical tasks.

Hedonistic Value:

Hedonistic value is subjective (Psychological) value which provides an experiential satisfaction. In other words, the immediate psychological gratification that comes from experiencing some activity or from consumption of a product.



Conceptual Background of the Domain Problem:

The problem statement examined how customers form expectations on technology based self-service quality and suggested five main attributed 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.

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.

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.

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

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.

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 Modeling 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 maps.

Data Source 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 Flip Robo which is in the format xlsx.

- There are two excel sheets one is detailed datasheet and other one is encoded datasheet. 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 (integer data type).
- While describing the data I found skewness and outliers present in the columns. Since all the columns are categorical so I haven't removed outliers and skewness

Data Pre-processing Done:

At first, I have imported the necessary libraries and dataset & Checked the dimension of the dataset, date 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')
```

- I have checked the 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 them using appropriate values.
- Visualized each feature using seaborn and matplotlib libraries by plotting count plot, pie plot, distribution plot, box plot and factor plot.
- 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 ().

Libraries:

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

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import warnings
import os
import scipy as stats
from sklearn.preprocessing import LabelEncoder
%matplotlib inline
warnings.filterwarnings('ignore')
```

MODEL/S DEVELOPMENT AND EVALUATION

Identification of possible problem-solving approaches (methods) :

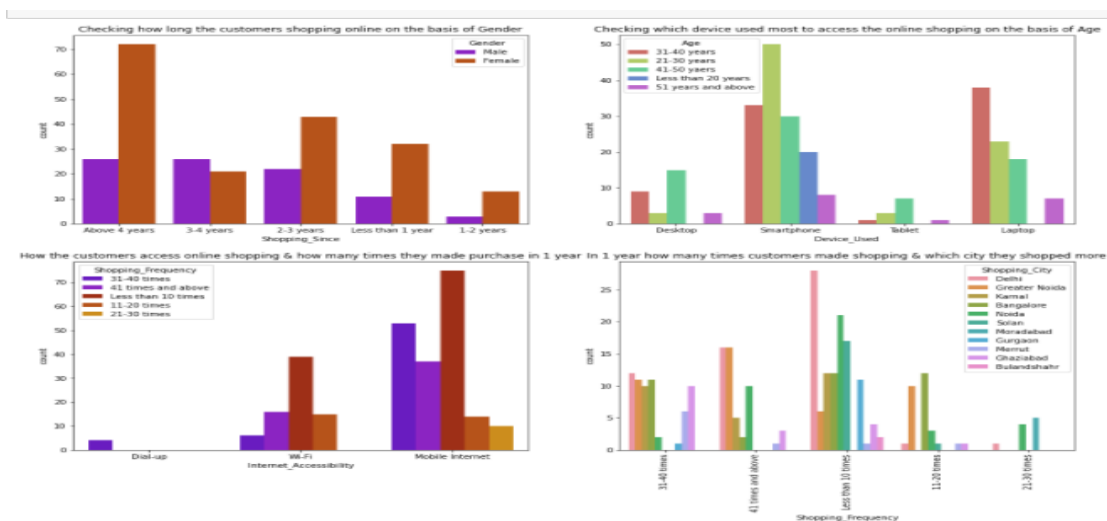
- I have used “.replace()” function to replace unwanted entries in the columns.
- Used “Label Encoder” method to encode the categorical features into numerical.
- Described the statistical details of the features using “.describe()” method.
- Used “Pearson’s method” to check the correlation between the features.
- Performed both univariate and bivariate analysis using seaborn and matplotlib.

Visualizations :

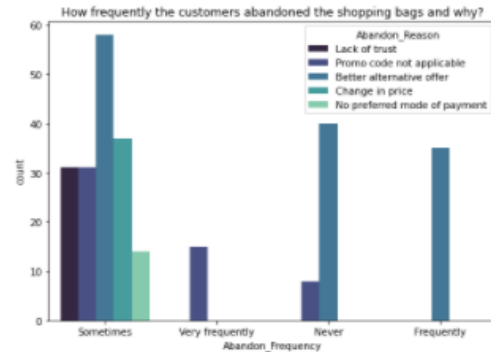
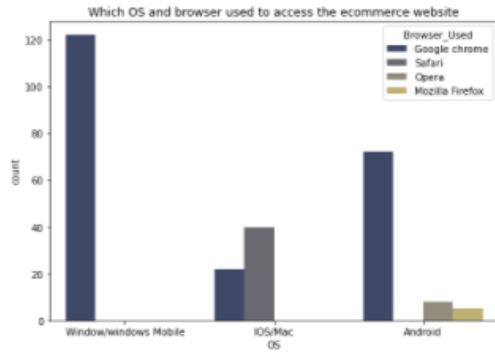
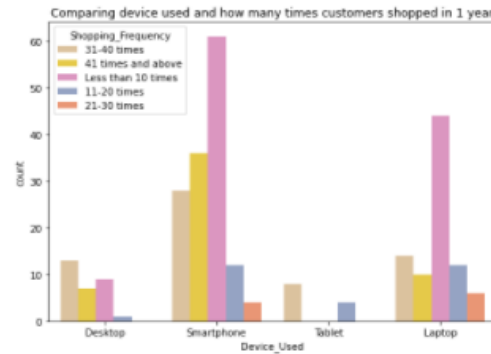
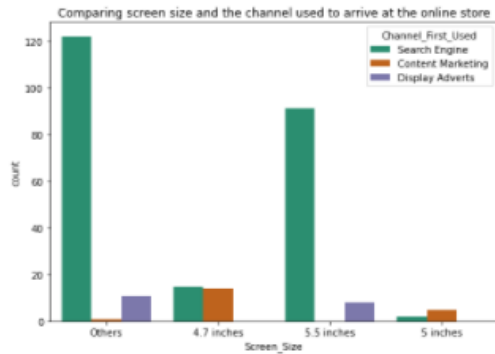
I have performed both univariate and bivariate analysis to visualize the data. In univariate analysis I have used pie plots, count plots and distribution plot and in bivariate analysis I have used count plot, factor plot and boxplots. Here I will be showing only bivariate analysis plots to analyse the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

Observation From the 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 year. 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.

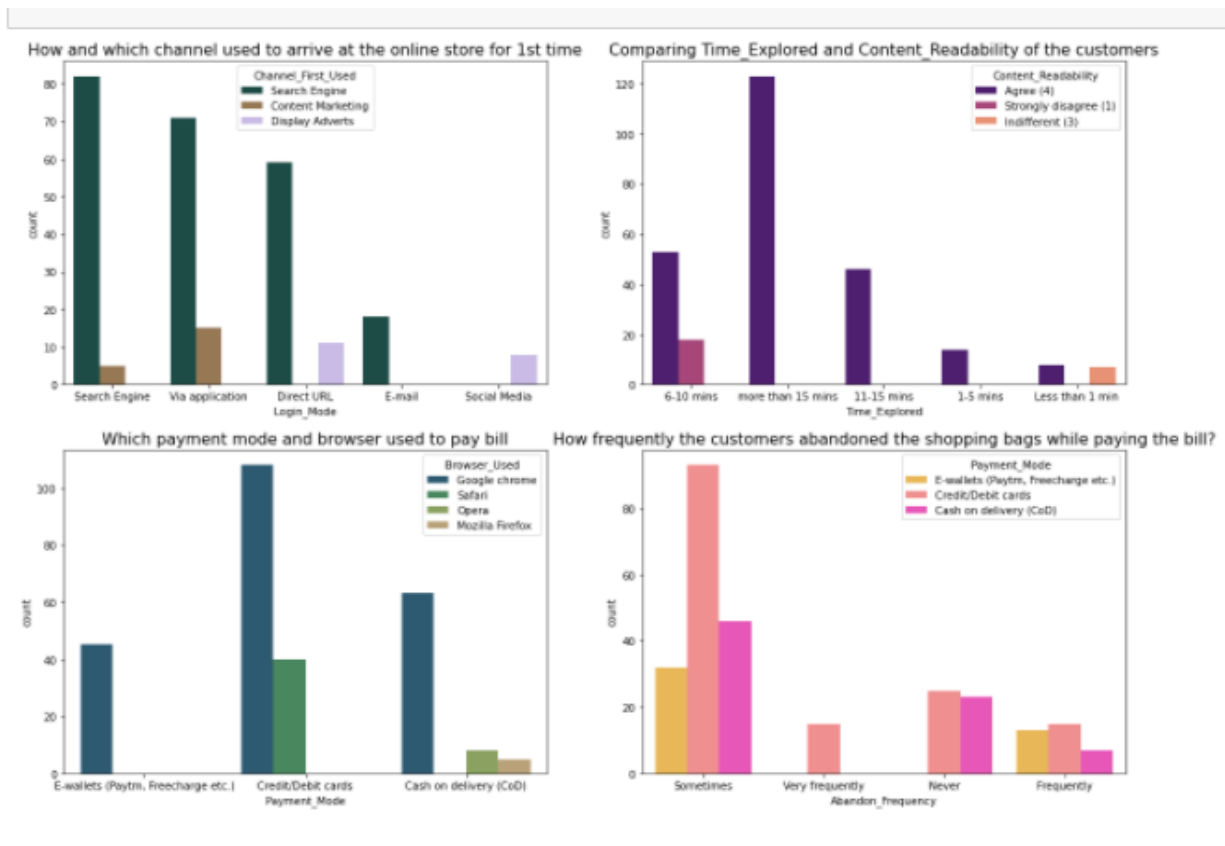


- ❖ Many customers whose age between 31-40 years and 21-30 years used Smartphone's followed by Laptops to access the online shopping websites.
- ❖ Most of the customers access the shopping websites more than 31-40 times in 1 year through Mobile Internet to shop the products also most of the customers who used mobile internet to access the online shopping website made online purchase less than 10 times in a year

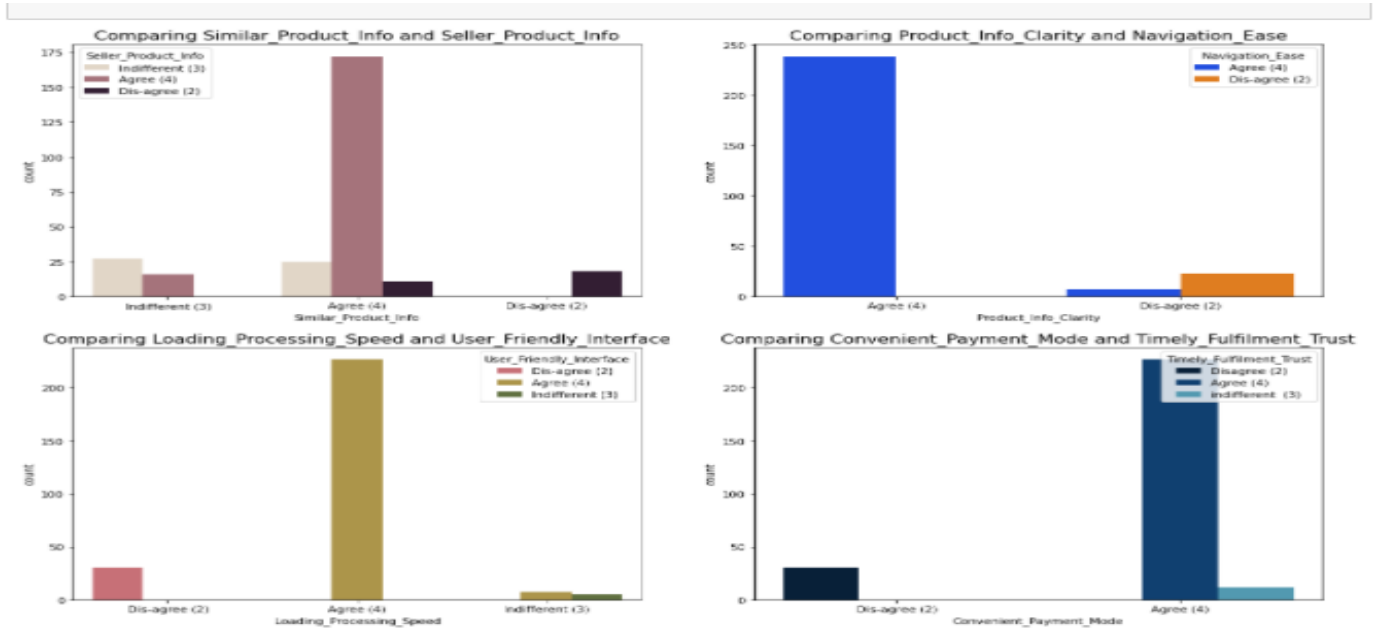


- The customers having their mobile screen size say 6 inches have followed search engine channel to arrive at their favorite online store for the first time.
- Most of the customers used Smartphone's 31-40 times in a year to access the ecommerce websites to shop the products.
- Many customers having windows operating system in their device ran Google chrome to access the ecommerce shopping websites and some of the customers having IOS/Mac operation system used Google chrome as well as Safari to reach the online shopping store.

OBSERVATION OF THE BELOW PLOT:



- 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 reshopping the product.
- 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 if 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.



Observation From the above Plots:

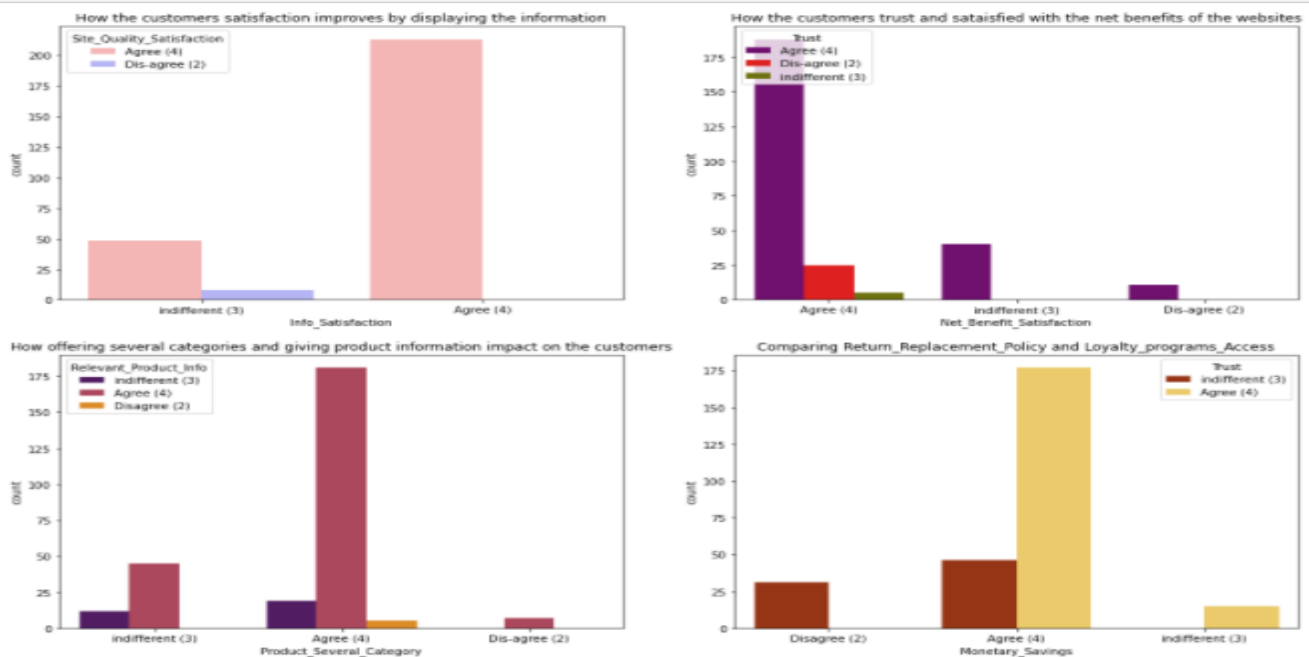
Most of the customers agreed that the information on similar product to the one highlighted is important for product comparison and also Complete information on listed seller and product being offered is important for purchase decision in order to buy a product, the ecommerce website must give the complete 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. 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 fulfill 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 case on delivery .So if the retailers provides 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. 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 will enhance the companies 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 one channel is not available then customers can easily reach out to other channel to fulfill their benefits.

Most of the customers believed that they enjoy online shopping also shopping 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.



Observation 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 .
- 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 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.
- 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. Similarly, the remaining plots are predicted as same before.

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 frustrated 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 favourite websites for the customers and Flip Kart was the second favourite webapp 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, Also, some of the dimensions like ease of navigation, loading and speed, late delivery etc did not have impact on the ecommerce websites.
- 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 endeavour 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 consumer' 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 :

- Customers 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.