CUSTOMER RETENTION PROJECT

Submitted by:

Priya Rajput

ACKNOWLEDGMENT

I would like to thank Flip Robo Technologies for providing me with the opportunity to work on this project. It is indeed a pleasure for me to have worked on this project.

I am also grateful to Miss Sapna for her constant guidance and support.

TABLE OF CONTENTS

INTRODUCTION

Customer retention enables the company to increase profitability and revenue. Thus, the small increase in customer retention could have a positive impact on profitability. Customer retention indicates customer's intention to repurchase a service or a product from the service provider. It should be a continuous process to find a customer and retain them in a long-term relationship

Business Problem Framing

"We do not need many customers to buy once, we need one customer to buy many times". That is a very familiar slogan of companies providing products and services. Attracting new customers is important to businesses, but they are not necessarily the best customers. Because if they only buy once and do not come back, the business will not be able to make much profit. That's why businesses need to keep finding a way to hold them back, make them trust the brand and keep buying. That's why customer retention arises**.**

Conceptual Background of the Domain Problem

Customer retention is the collection of activities a business uses to increase the number of repeat customers and to increase the profitability of each existing customer. 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.

Customers also demand greater choice and would appear to place even greater emphasis on value for money. Flexible payment methods and sophisticated channels of distribution have also played their part. These factors combined with the growth of the car owning middle class enjoying economic stability have also fuelled its growth. To cope with consumer demands retailers are constantly reshaping their business strategies and striving to offer the right products and services both to meet existing needs and to anticipate and stimulate future demand.

Review of Literature

Customer retention survives when the companies can fulfil customer expectations and additionally maintain it in long-term relationships to ensure long-term buying decisions [1–2]. The topic of customer retention is argued in business economics commonly within the perspective of relationship marketing, which considers customer relationships as one of the primary concerns with the long-term objective of developing and maintaining them [3]. Many previous studies indicated that companies should always manage customer satisfaction to achieve the retention stage. According to [3] "satisfaction is an overall customer attitude

towards a service provider". Authors added that satisfaction is an emotional reaction regarding what customers expect and what they receive, including the fulfilment of needs and goals. Customer retention states a desired outcome in the future to satisfaction, so long-term of relationship is demonstrated by satisfaction. Although customer satisfaction does not guarantee repurchase, it still plays a vital role in ensuring customer retention. While many studies on customer retention had long focused on customer satisfaction, additional factors are stated as an influence in customer retention, such as trust and commitment. "The Commitment-Trust Theory of Relationship Marketing," which is the most influential Relationship Marketing, suggests that the centre of successful relationship marketing is the relationship of commitment and trust. They urged the importance of commitment and trust that leads to build a positive correlation between company and customers and encourage efficiency, productivity, and effectiveness. The degree of trust between service provider and customer is significantly influenced by the quality of the service, which results in an effective commitment to the provider, and enhancing commitment is important since it leads to an intention to invest further and reinforce the relationship with the provider.

Motivation for the Problem Undertaken

The combination of both utilitarian value and hedonistic values are needed to affect the repeat purchase intention positively. Objective behind the problem is to help ecommerce websites to find E-retail factors for customer activation and retention.

Analytical Problem Framing

Mathematical/ Analytical Modelling of the Problem

In this project we have performed various mathematical and statistical analysis. We checked description or statistical summary of the data using describe, info and unique value count.

- The dataset consist of 269 rows and 71 columns.
- All of the attributes were of 'object' type except the pin code (int) so we change this int64 to object
- Dataset did not contain any null values.

```
In [14]: #Loading the dataset
         df=pd.read_excel(r'C:\Users\Admin\Desktop\customer_retention_dataset.xlsx')
         df.head()
```

Out[14]:

	1Gender of respondent	2 How old are you?	3 Which city do you shop online from?	4 What is the Pin Code of where you shop online from?	5 Since How Long You are Shopping Online ?	6 How many times you have made an online purchase in the past 1 year?	7 How do you access the internet while shopping on-line?	8 Which device do you use to access the online shopping?	9 What is the screen size of your mobile device? \t\t\t\t\t\t	10 What is the operating system (OS) of your device? \t\t\t	...	Longer time to get logged in (promotion, sales period)	Longer time in displaying graphics and photos (promotion, sales period)	Late declaration of price (promotion, sales period)	
0	Male	31-40 years	Delhi	110009	Above 4 years	31-40 times	Dial-up	Desktop	Others	Window/windows Mobile	...	Amazon.in	Amazon.in	Flipkart.com	
1	Female	21-30 years	Delhi	110030	Above 4 years	41 times and above	Wi-Fi	Smartphone	4.7 inches	IOS/Mac	...	Amazon.in, Flipkart.com	Myntra.com	snapdeal.com	
2	Female	21-30 years	Greater Noida	201308	3-4 years	41 times and above	Mobile Internet	Smartphone	5.5 inches	Android	...	Myntra.com	Myntra.com	Myntra.com	
3	Male	21-30 years	Karnal	132001	3-4 years	Less than 10 times	Mobile Internet	Smartphone	5.5 inches	IOS/Mac	...	Snapdeal.com	Myntra.com, Snapdeal.com	Myntra.com	
4	Female	21-30 years	Bangalore	530068	2-3 years	11-20 times	Wi-Fi	Smartphone	4.7 inches	IOS/Mac	...	Flipkart.com, Paytm.com	Paytm.com	Paytm.com	

5 rows × 71 columns

```
In [15]: df.shape
Out[15]: (269, 71)
```

```
In [25]: #finding the null values
         df.isna().sum()
Out[25]: Gender                    0
         Age                       0
         City                      0
         PinCode                   0
         ShoppingSince             0
                                  ..
         LongerDeliveryPeriod      0
         WebApp_DesignChange       0
         PageDisruption            0
         EebApp_Efficiency         0
         Recommendation            0
         Length: 71, dtype: int64
```

```
In [26]: df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 269 entries, 0 to 268
Data columns (total 71 columns):
 #   Column                          Non-Null Count  Dtype
---  ------                          --------------  -----
 0   Gender                          269 non-null    object
 1   Age                             269 non-null    object
 2   City                            269 non-null    object
 3   PinCode                         269 non-null    int64
 4   ShoppingSince                   269 non-null    object
 5   ShoppingFrequencySince_1year    269 non-null    object
 6   Internet_Accessibility          269 non-null    object
 7   Device_Used                     269 non-null    object
 8   Screen_Size                     269 non-null    object
 9   OS                              269 non-null    object
 10  Browser_Used                    269 non-null    object
 11  Channel_FirstUsed               269 non-null    object
 12  Login_Mode                      269 non-null    object
 13  Deciding_Time_spent             269 non-null    object
 14  Payment_Mode                    269 non-null    object
 15  AbandonMaking_PaymentFrequency  269 non-null    object
```

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 Flip Robo which is in the format xlsx.

There are two excel sheets one is detailed datasheet and other one is coded datasheet. We 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.

Data Pre-processing Done

First, We have imported the necessary libraries and dataset.

```
In [1]: import numpy as np
        import pandas as pd
        import seaborn as sns
        import matplotlib.pyplot as plt
        import warnings
        warnings.filterwarnings('ignore')
        %matplotlib inline
```

Checked the dimension of the dataset, information of the data frame using info() and unique values present in the dataset.

```
In [15]: df.shape
```

```
In [29]: #Checking unique values of each
         df.nunique()

Out[29]: Gender                    2
         Age                       5
         City                     11
         PinCode                  39
         ShoppingSince             5
                                  ..
         LongerDeliveryPeriod      6
         WebApp_DesignChange       7
         PageDisruption            8
         EebApp_Efficiency         8
         Recommendation            8
         Length: 71, dtype: int64
```

```
In [36]: #Lets check the value count of e
         for i in df.columns:
                 print(i)
                 print(df[i].value_counts
                 print('*****************

         31-40 years          81
         21-30 years          79
         41-50 yaers          70
         Less than 20 years   20
         51 years and above   19
         Name: Age, dtype: int64
         ****************************
         City
```

```
In [26]: df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 269 entries, 0 to 268
Data columns (total 71 columns):
 #   Column                        Non-Null Count  Dtype
---  ------                        --------------  -----
 0   Gender                        269 non-null    object
 1   Age                           269 non-null    object
 2   City                          269 non-null    object
 3   PinCode                       269 non-null    int64
 4   ShoppingSince                 269 non-null    object
 5   ShoppingFrequencySince_1year  269 non-null    object
 6   Internet_Accessibility        269 non-null    object
 7   Device_Used                   269 non-null    object
 8   Screen_Size                   269 non-null    object
 9   OS                            269 non-null    object
 10  Browser_Used                  269 non-null    object
 11  Channel_FirstUsed             269 non-null    object
 12  Login_Mode                    269 non-null    object
 13  Deciding_Time_spent           269 non-null    object
 14  Payment_Mode                  269 non-null    object
 15  AbandonMaking_PaymentFrequency 269 non-null   object
 16  Abandon_Reason                269 non-null    object
 17  Content_Readability           269 non-null    object
 18  HighlightSame_ProductInfo     269 non-null    object
 19  Seller_Info                   269 non-null    object
 20  ProductInfo_Clarity           269 non-null    object
 21  NavigationEase                269 non-null    object
 22  Loading_Speed                 269 non-null    object
 23  UserFriendly_Interface        269 non-null    object
 24  Convenient_PaymentMode        269 non-null    object
 25  timelyFulfilmentTrust         269 non-null    object
 26  CustSupportResponse           269 non-null    object
 27  CustPrivacyGuarantee          269 non-null    object
```

Since the columns had no proper names, so we renamed them with new name as below.

```
In [22]: # Let's rename the new column names to the original datastes

columns = ['Gender','Age','City','PinCode','ShoppingSince','ShoppingFrequencySince_1year','Internet_Accessibility','Device_Used',
           'Screen_Size','OS', 'Browser_Used','Channel_FirstUsed','Login_Mode','Deciding_Time_spent','Payment_Mode',
           'AbandonMaking_PaymentFrequency','Abandon_Reason','Content_Readability','HighlightSame_ProductInfo',
           'Seller_Info','ProductInfo_Clarity','NavigationEase','Loading_Speed','UserFriendly_Interface',
           'Convenient_PaymentMode','timelyFulfilmentTrust','CustSupportResponse','CustPrivacyGuarantee',
           'VariousChannelResponses','OnlineShoppingBenefit','enjoyment','convenient','returnReplacementPolicy',
           'LoyaltyProgramsAccess','InfoSatisfaction','siteQualitySatisfaction','netBenefitSatisfaction','trust',
           'ProductSeveralCategory','relevantProductInfo','MonetarySavings','patronizingConvenience','AdventureSense',
           'EnhanceSocialStatus','gratification','roleFulfilment','moneyWorthy','shoppedFrom','EasyWebApp',
           'VisuallyAppealingWebApp','ProductVariety','CompleteProductInfo','fastWebApp','reliableWebApp','quickPurchase',
           'PaymentOptionsAvailability','SpeedyDelivery','custInfoPrivacy','financial_InfoSecurity','perceivedTrustworthiness',
           'MultiChannelAssistance','LongLoginTime','LongDisplayTime','LatePriceDeclare','LongLoadingTime','LimitedPaymentMode',
           'LongerDeliveryPeriod','WebApp_DesignChange','PageDisruption','EebApp_Efficiency','Recommendation']

In [23]: df.columns = columns
```

We have checked the null values and found no null values present in the dataset and visualized it using heat map.

By value_counts function we checked the count of each column.

```
[36]: #Lets check the value count of each column to see if there are any unexpected and unwanted entries present in the column
      for i in df.columns:
              print(i)
              print(df[i].value_counts())
              print('*****************************')
31-40 years           81
21-30 years           79
41-50 yaers           70
Less than 20 years    20
51 years and above    19
Name: Age, dtype: int64
*****************************
City
Delhi           58
Greater Noida   43
Noida           40
Bangalore       37
Karnal          27
Solan           18
```

Visualized each feature using seaborn and matplotlib libraries by plotting count plot,box plot and factor plot.

Described the data using describe().

```
In [43]: df.describe()
```

Out[43]:

	ShoppingSince	ShoppingFrequencySince_1year	Internet_Accessibility	Device_Used	Screen_Size	OS	...	LongLoginTime	LongDisplayTime	LatePr
)	269	269	269	269	269	269	...	269	269	
)	5	6	4	4	4	3	...	10	10	
)	Above 4 years	Less than 10 times	Mobile internet	Smartphone	Others	Window/windows Mobile	...	Amazon.in	Amazon.in, Flipkart.com	
)	98	114	142	141	134	122	...	57	60	

Data Inputs- Logic- Output Relationships

EDA was performed by creating valuable insights using various visualization libraries.

Hardware and Software Requirements and Tools Used

Hardware required:
Processor: core i3
RAM: 8 GB

Software required:
Anaconda 3- language used Python 3
Microsoft Excel

Model/s Development and Evaluation

Identification of possible problem-solving approaches (methods)

Firstly we imported all the required libraries and load the dataset. Then did a statistical analysis on the dataset .Then deeply analyzed the data by finding a relationship between each attribute and the Customer according to given details and tried to find the factors that will help in Customer Retention.

Visualizations

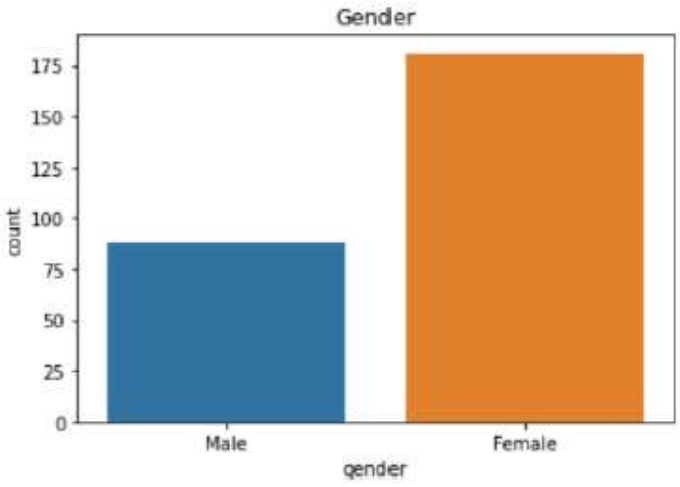
We have performed both univariate and bivariate analysis to visualize the data. In univariate analysis we have used count plots and in bivariate analysis we used countplot and factor plot. Here,we have used loop for Categorical column to save time in univariate analysis.

In [45]:
```
# Separating categorical columns
categorical_col=[]
for i in df.dtypes.index:
    if df.dtypes[i]=='object':
        categorical_col.append(i)
print(categorical_col)
```

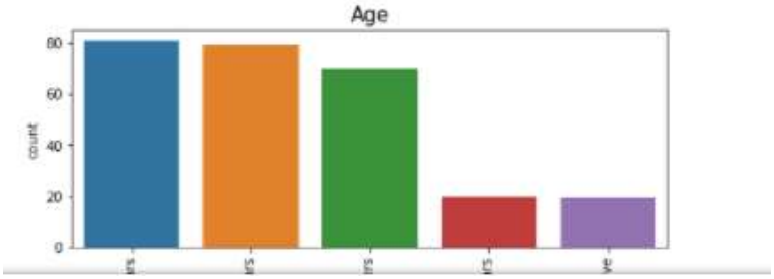
```
['Gender', 'Age', 'City', 'PinCode', 'ShoppingSince', 'ShoppingFrequencySince_1year', 'Internet_Accessibility', 'Device_Used',
'Screen_Size', 'OS', 'Browser_Used', 'Channel_FirstUsed', 'Login_Mode', 'Deciding_Time_spent', 'Payment_Mode', 'AbandonMaking_P
aymentFrequency', 'Abandon_Reason', 'Content_Readability', 'HighlightSame_ProductInfo', 'Seller_Info', 'ProductInfo_Clarity',
'NavigationEase', 'Loading_Speed', 'UserFriendly_Interface', 'Convenient_PaymentMode', 'timelyFulfilmentTrust', 'CustSupportRes
ponse', 'CustPrivacyGuarantee', 'VariousChannelResponses', 'OnlineShoppingBenefit', 'enjoyment', 'convenient', 'returnReplaceme
ntPolicy', 'LoyaltyProgramsAccess', 'InfoSatisfaction', 'siteQualitySatisfaction', 'netBenefitSatisfaction', 'trust', 'ProductS
everalCategory', 'relevantProductInfo', 'MonetarySavings', 'patronizingConvenience', 'AdventureSense', 'EnhanceSocialStatus',
'gratification', 'roleFulfilment', 'moneyWorthy', 'shoppedFrom', 'EasyWebApp', 'VisuallyAppealingWebApp', 'ProductVariety', 'Co
mpleteProductInfo', 'fastWebApp', 'reliableWebApp', 'quickPurchase', 'PaymentOptionsAvailability', 'SpeedyDelivery', 'custInfoP
rivacy', 'financial_InfoSecurity', 'perceivedTrustworthiness', 'MultiChannelAssistance', 'LongLoginTime', 'LongDisplayTime', 'L
atePriceDeclare', 'LongLoadingTime', 'LimitedPaymentMode', 'LongerDeliveryPeriod', 'WebApp_DesignChange', 'PageDisruption', 'Ee
bApp_Efficiency', 'Recommendation']
```

[104]:
```
#Count plot for all categorical columns
for i in df[categorical_col]:
    plt.figure(figsize = (10, 8))
    sns.countplot(df[i])
    print(round(df[i].value_counts()/269*100),2)
    plt.title(i,fontsize=15)
    plt.xticks(rotation=90)
    plt.show()
```

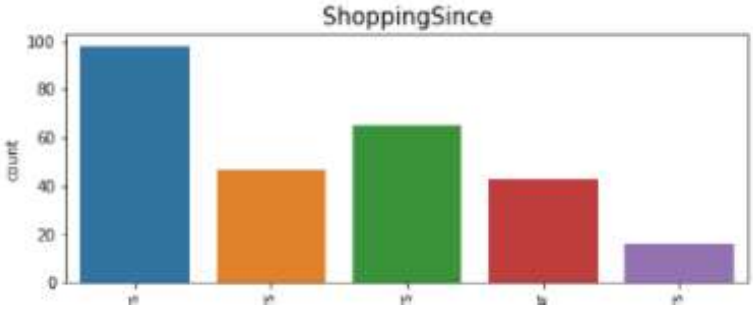
```
Female    67.0
Male      33.0
Name: gender, dtype: float64 2
```



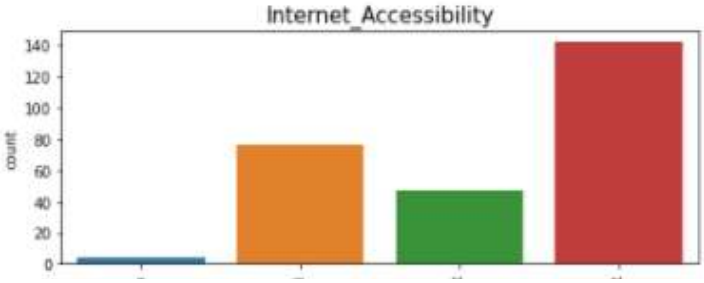
```
31-40 years           30.0
21-30 years           29.0
41-50 yaers           26.0
Less than 20 years     7.0
51 years and above     7.0
Name: Age, dtype: float64 2
```



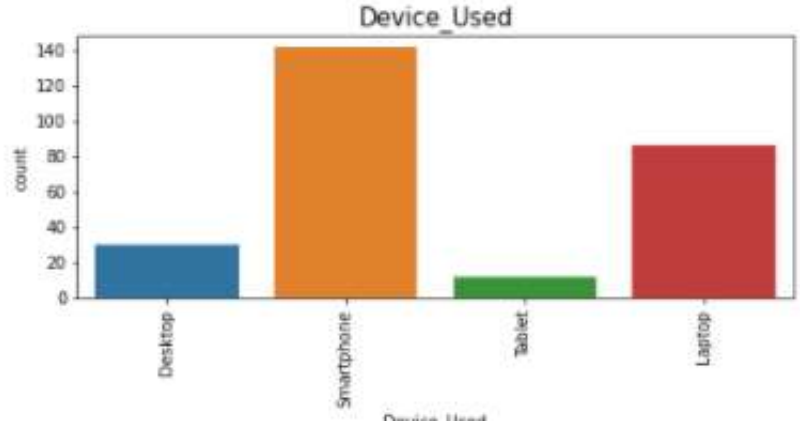

```
Above 4 years       36.0
2-3 years           24.0
3-4 years           17.0
Less than 1 year    16.0
1-2 years            6.0
Name: ShoppingSince, dtype: float64 2
```

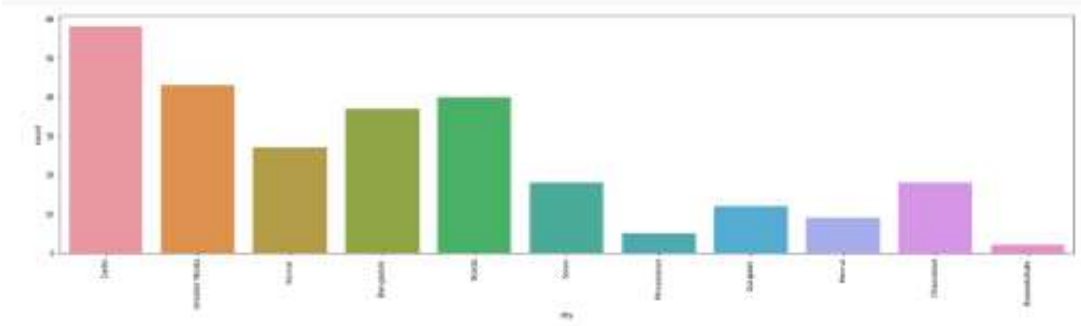



```
Mobile internet    53.0
Wi-Fi              28.0
Mobile Internet    17.0
Dial-up             1.0
Name: Internet_Accessibility, dtype: float64 2
```

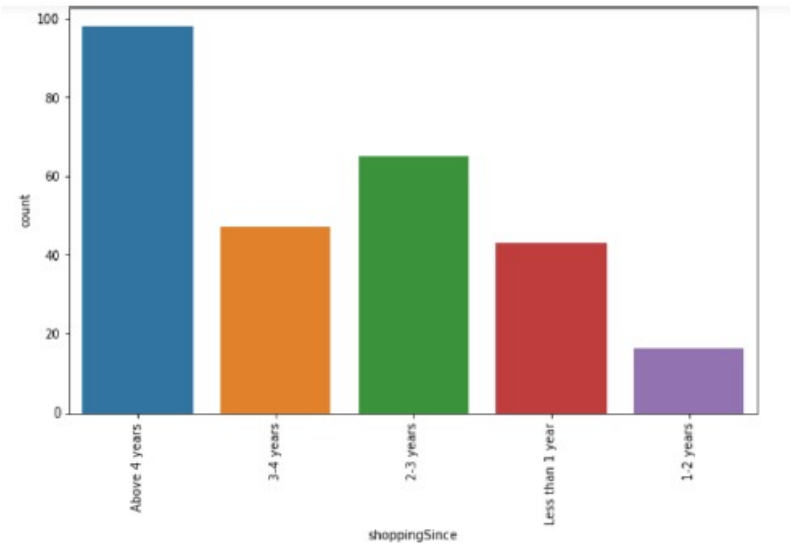


Name: Device_Used, dtype: float64 2

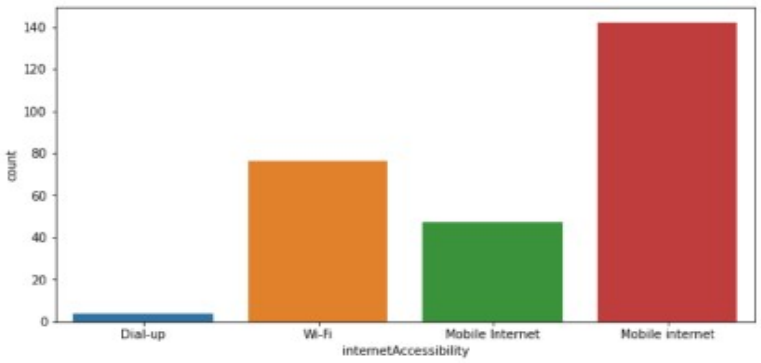




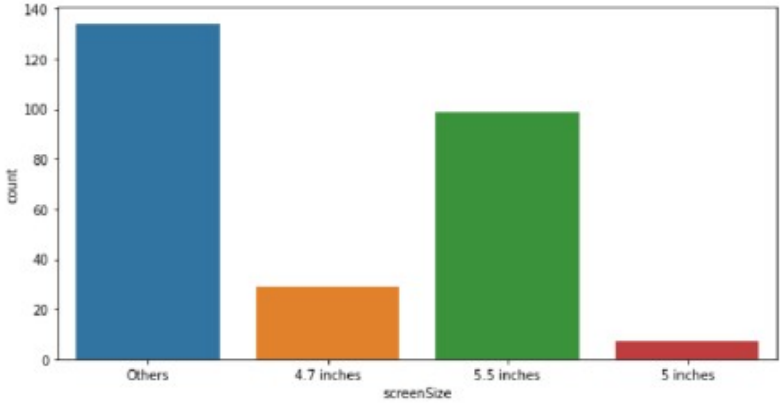
```
Delhi           22.0
Greater Noida   16.0
Noida           15.0
Bangalore       14.0
Karnal          10.0
Solan            7.0
Ghaziabad        7.0
Gurgaon          4.0
Marrut           3.0
Moradabad        2.0
Bulandshahr      1.0
Name: city, dtype: float64 )
```



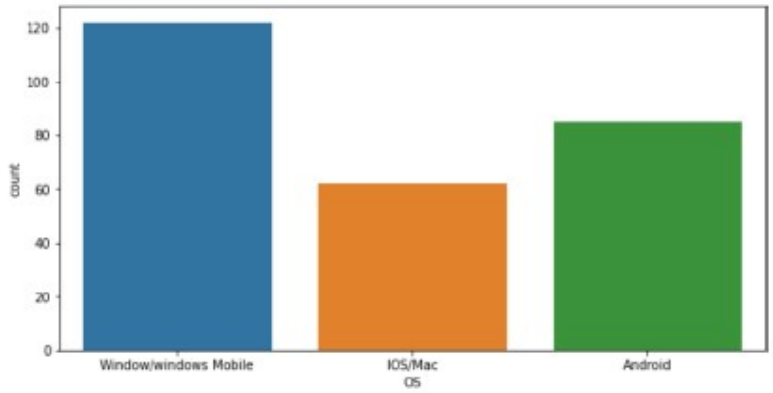

```
Above 4 years       36.0
2-3 years           24.0
3-4 years           17.0
Less than 1 year    16.0
1-2 years            6.0
```



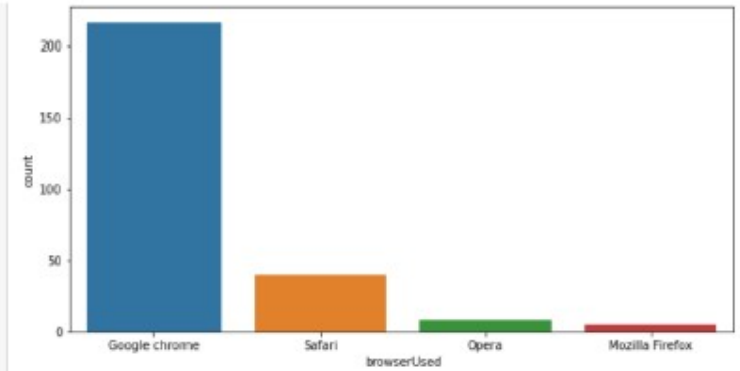

```
Mobile internet    53.0
Wi-Fi              28.0
Mobile Internet    17.0
Dial-up             1.0
Name: internetAccessibility, dtype: float64 2
```



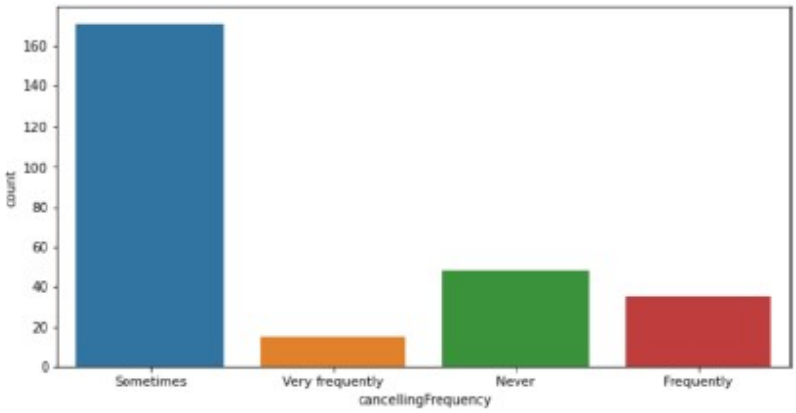

```
Others        50.0
5.5 inches    37.0
4.7 inches    11.0
5 inches       3.0
Name: screenSize, dtype: float64 2
```



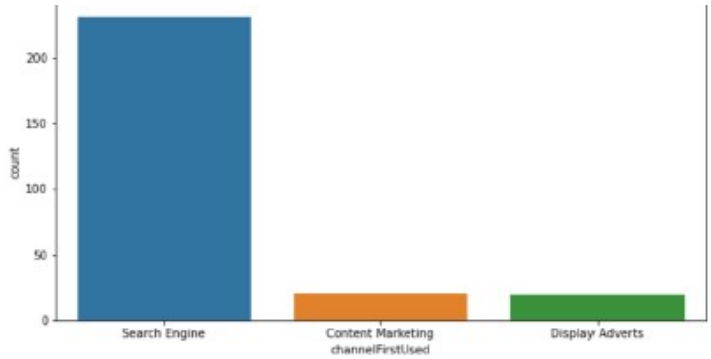

```
Window/windows Mobile    45.0
Android                  32.0
IOS/Mac                  23.0
Name: OS, dtype: float64 2
```



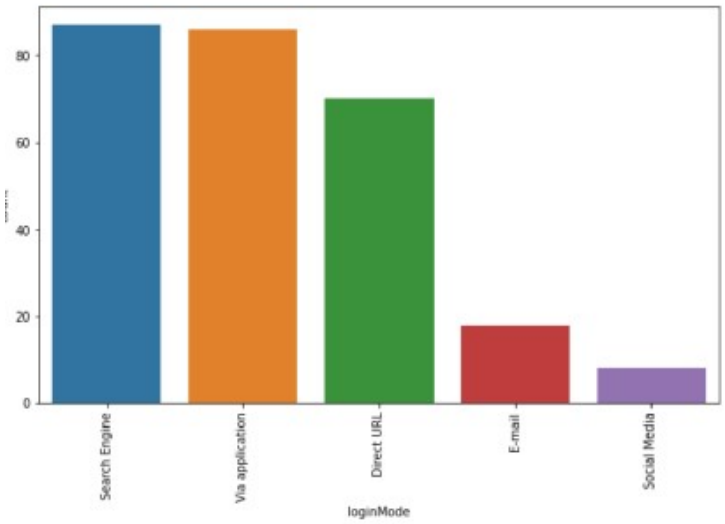

```
Google chrome      80.0
Safari             15.0
Opera               3.0
Mozilla Firefox     2.0
Name: browserUsed, dtype: float64 2
```



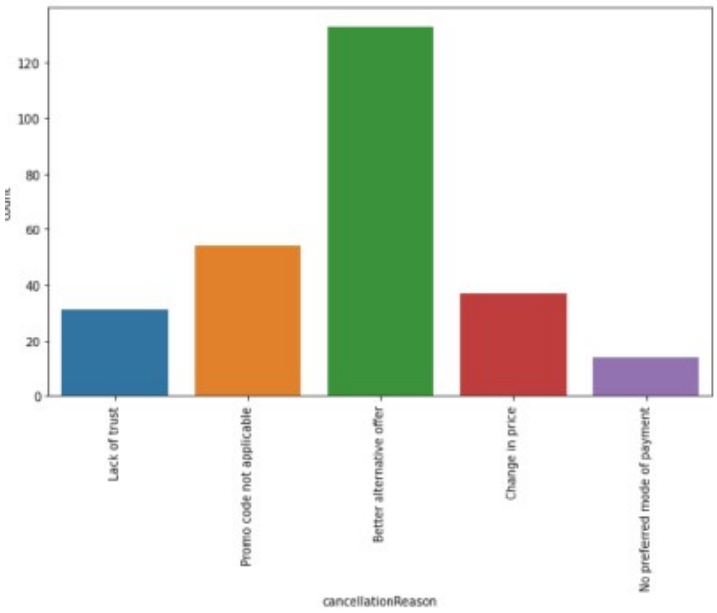

```
Sometimes          64.0
Never              18.0
Frequently         13.0
Very frequently     6.0
Name: cancellingFrequency, dtype: float64 2
```

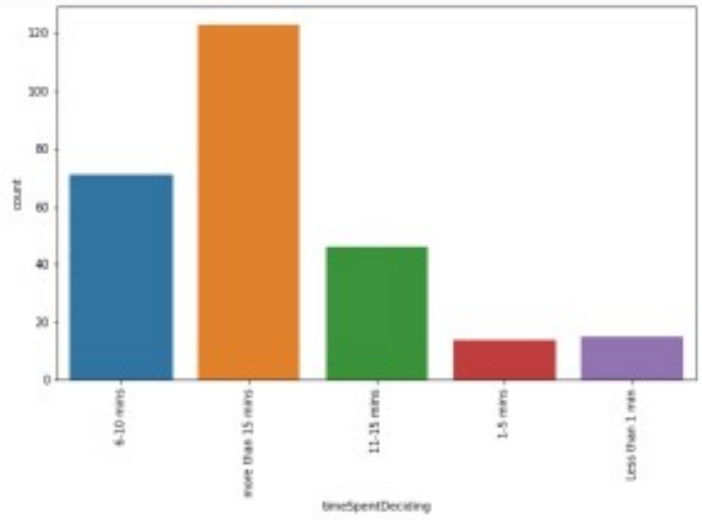



```
earch Engine        86.0
ontent Marketing     7.0
isplay Adverts       7.0
```

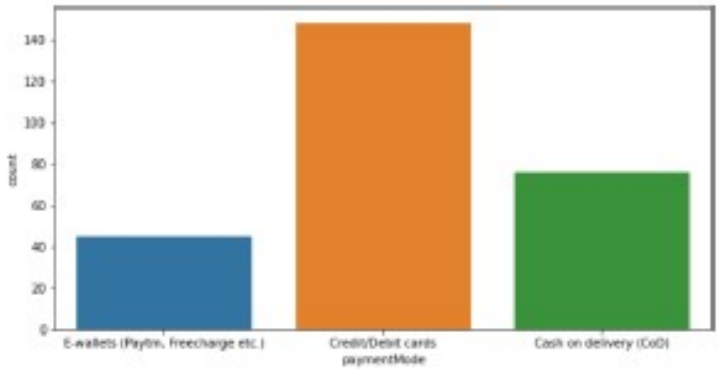



```
arch Engine      32.0
a application    32.0
```

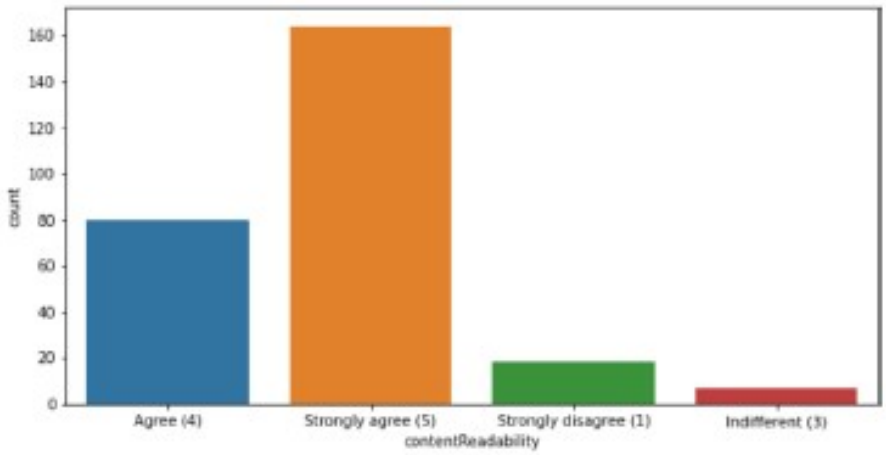




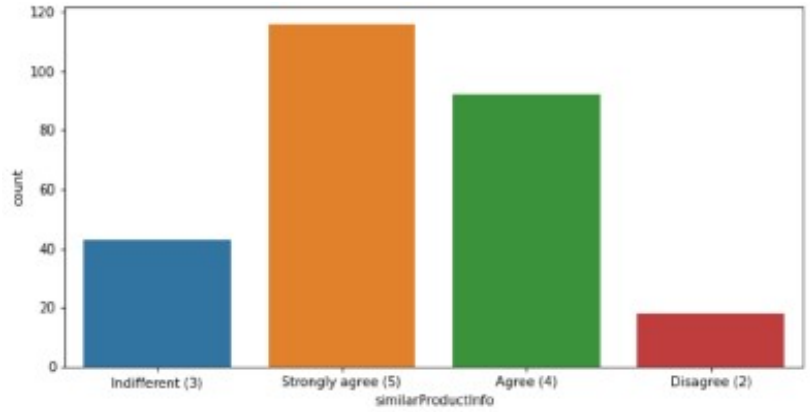

```
more than 15 mins    46.0
6-10 mins            26.0
11-15 mins           17.0
```



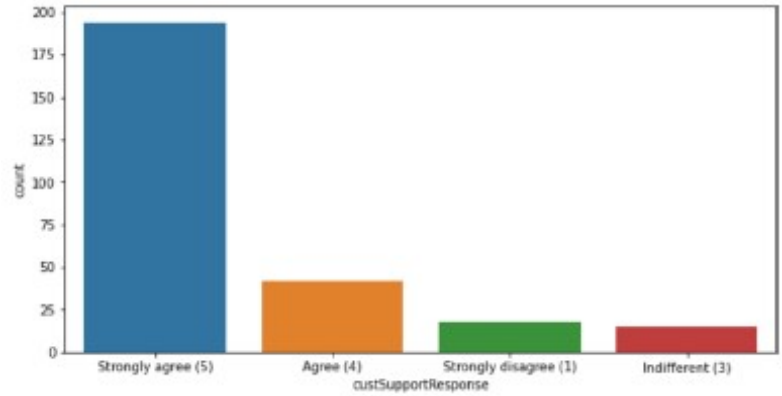

```
Credit/Debit cards                55.0
Cash on delivery (CoD)            28.0
E-wallets (Paytm, Freecharge etc.)  17.0
Name: paymentMode, dtype: float64 2
```



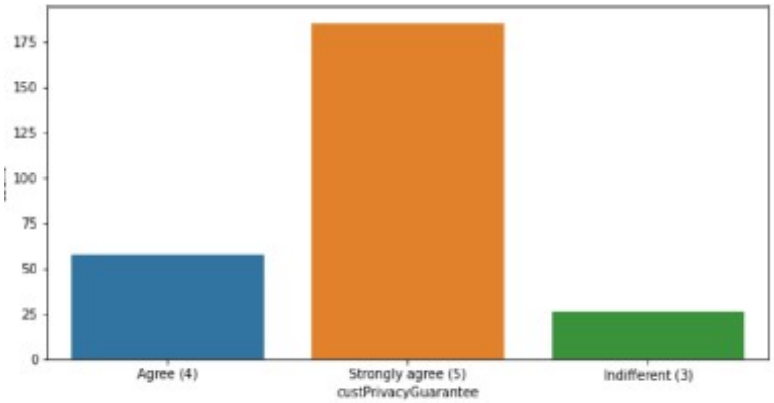

```
Strongly agree (5)       61.0
Agree (4)                30.0
Strongly disagree (1)     7.0
Indifferent (3)           3.0
Name: contentReadability, dtype: float64 2
```



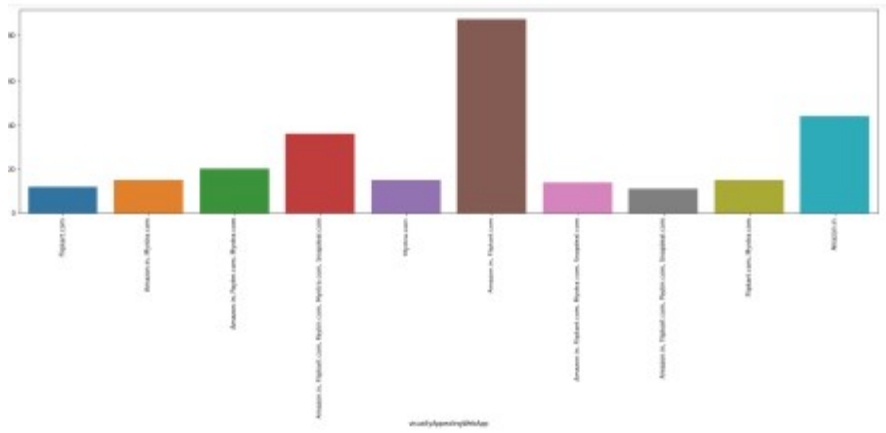

```
Strongly agree (5)    43.0
Agree (4)             34.0
Indifferent (3)       16.0
Disagree (2)           7.0
Name: similarProductInfo, dtype: float64 2
```




```
Strongly agree (5)       72.0
Agree (4)                16.0
Strongly disagree (1)     7.0
Indifferent (3)           6.0
Name: custSupportResponse, dtype: float64 2
```




```
rongly agree (5)    69.0
ree (4)             22.0
different (3)       10.0
me: custPrivacyGuarantee, dtype: float64 2
```



```
Amazon.in, Flipkart.com                                          32.0
Amazon.in                                                        16.0
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com     13.0
Amazon.in, Paytm.com, Myntra.com                                  7.0
Amazon.in, Myntra.com                                             6.0
Myntra.com                                                        6.0
Flipkart.com, Myntra.com                                          6.0
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com                 5.0
Flipkart.com                                                      4.0
Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com                  4.0
Name: visuallyAppealingWebApp, dtype: float64 )
```

By looking into the count plots below are my observations:

- Out of the tota, more number of count are for females than male(more than double).
- Higher number i.e 30% of online shopping is done by the age group of 31-40 years.
- Delhi has maximum percent (22%) which show that in Delhi the online purchase is high as compared to other.
- A city with pin code 201308(Delhi) has maximum online shopping.
- Number of count is high for above 4 years which means people are choosing to shop online since 4 years.
- Number of count is also high for less than 10 times of people who are shopping less than 10 times per year.
- Majority of people use mobile internet for online shopping.
- Majority of customers uses smart phones while doing online shopping.
- Here we see that count contains most of the participants use any other screen size. Also participant's use their 5.5 inches of screen size to shop online which is slightly less than others count.
- Most participants uses Window/windows Mobile to shop online then, Android then, comes IOS/Mac at least.
- Most of the customers uses chrome browser.
- Majority of customers used Search engine to reach their favorite online store for the first time.

- After first visit most of the customers are reaching the online store through Search engine, Via application and Direct Url.
- Most of the customers are exploring the e-retail store more than 15 mins before making purchase decision.
- Maximum customers are using credit/debit cards for their payment.
- The count of abandon (leave without purchase the item) is high for sometimes, which means people who are abandoning sometimes are more.
- Most participants get better alternative offer so, they abandon their shopping cart. And the second main reason to abandon is "Promo code not applicable". Then comes the reasons "Change in Price", "Lack of trust" and "No preferred mode of payment". Here, we can see that we also have a good number of participants who abandon due to "Lack of trust", hence. There is a need to get the trust from the participants.
- Most of the customers are agreeing that the content should be easy to understand on the website.
- People agree that information on similar product to the one highlighted is important for product comparison.
- 70% people believe that Complete information on listed seller and product is important for purchase decision
- Most of the customers agree that the relevant information on listed products must be stated clearly.
- Maximum customers are agreeing that it should be easy to navigate in websites.
- Maximum customers are agreeing that loading and processing speed should be high.
- Maximum customers agree to have user friendly interface of the website.
- Maximum customers agree to have convenient payment methods.
- Maximum customers Trust that the online retail store will fulfill its part of the transaction at the stipulated time.
- Maximum customers agree to have Empathy (readiness to assist with queries) towards the customers.
- Maximum customers agree to have the ability to guarantee the privacy of the customer.
- Maximum customers agree to have Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.).
- Maximum customers agree to have Online shopping gives monetary benefit and discounts.
- Maximum customers agrees to have Enjoyment is derived from shopping online.
- Maximum customers agrees to have Shopping online is convenient and flexible.
- Maximum customers agree that Return and replacement policy of the e-tailer is important for purchase decision.
- Maximum customers agree that Gaining access to loyalty programs is a benefit of shopping online.
- Maximum customers agree that displaying quality Information on the website improves satisfaction of customers.

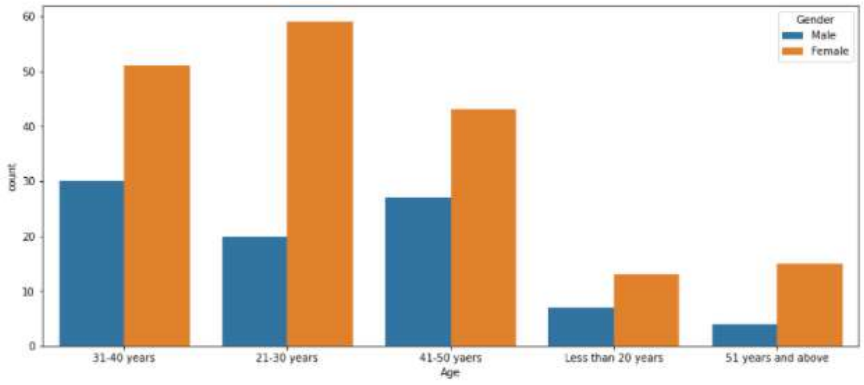
- Maximum customers agree that User derive satisfaction while shopping on a good quality website or application.
- Maximum customers agree that Net Benefit derived from shopping online can lead to user's satisfaction.
- Maximum customers agree that User satisfaction cannot exist without trust.
- Maximum customers agree to have offering a wide variety of listed product in several categories.
- Maximum customers agree to have Provision of complete and relevant product information.
- Maximum customers agree to have monetary savings.
- Maximum customers agree to have the Convenience of patronizing the online retailer.
- Maximum customers agree that Shopping on the website gives you the sense of adventure.
- Maximum customers agree that Shopping on your preferred e-tailer enhances your social status.
- Maximum customers agree that you feel gratification shopping on your favorite e-tailer.
- Maximum customers agree that Shopping on the website helps you fulfill certain roles.
- Maximum customers agree to Getting value for money spent.
- Most of the people are shopping from all available online retailers.
- Most of the people are choosing all retailers for easy to use websites.
- Amazon.com and Flipkart.com has a good visual appeal web-page layout as per the survey.
- Wide variety of product on offer is mostly obtained in Amazon.com and Flipkart.com.
- Maximum count is for Amazon.in, Flipkart.com for having Complete, relevant description information of products.
- Maximum count is for Amazon. in for having Fast loading website speed of website and application.
- Maximum count is for Amazon.in for having Reliable website or application.
- Maximum count is for Amazon.in for having the Quickest to complete purchase.
- Maximum count is for Amazon.in, Flipkart.com for having Availability of several payment options.
- Maximum count is for Amazon.in for having Speedy order delivery.
- Maximum count is for Amazon.in for having Privacy of customers' information.
- Maximum count is for Amazon.in for having Security of customer financial information.
- Maximum count is for Amazon.in for having Perceived Trustworthiness.
- Maximum count is for Amazon.in and Amazon.in, Flipkart.com, Myntra.com, Snapdeal for having Presence of online assistance through multi-channel.
- Maximum count is for Amazon.in for Longer time to get logged in (promotion, sales period).
- Maximum count is for Amazon.in, Flipkart.com for having Longer time in displaying graphics and photos (promotion, sales period).

- Maximum count is for Myntra.com for Late declaration of price (promotion, sales period).
- Maximum count is for choosing Myntra.com for Longer page loading time (promotion, sales period).
- Maximum count is for Snapdeal.com for having limited mode of payment on most products (promotion, sales period).
- Maximum count is for Paytm.com for Longer delivery period.
- Maximum count is for Amazon.in for having Change in website/Application design.
- Maximum count is for Amazon.in,Myntra.com and Snapdeal.com for having Frequent disruption when moving from one page to another.
- Maximum count is for Amazon.in for having Website is as efficient as before.
- Maximum count is for Amazon.in for recommendation.

Univariate Analysis

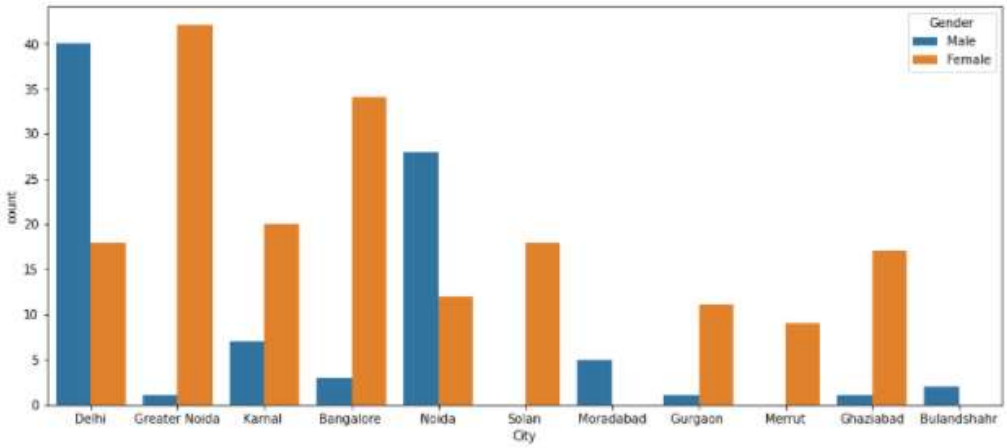
```
In [74]: #Factor plot for Gender
         plt.figure(figsize=[14,6])
         sns.countplot(df['Age'],hue=df['Gender'])

Out[74]: <AxesSubplot:xlabel='Age', ylabel='count'>
```



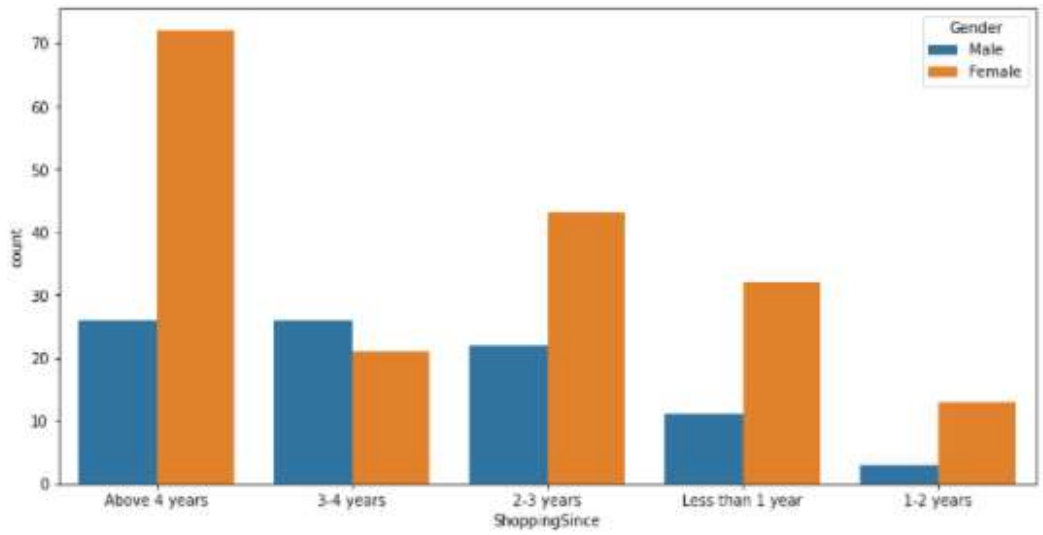

```
In [75]: plt.figure(figsize=[14,6])
         sns.countplot(df['City'],hue=df['Gender'])

Out[75]: <AxesSubplot:xlabel='City', ylabel='count'>
```



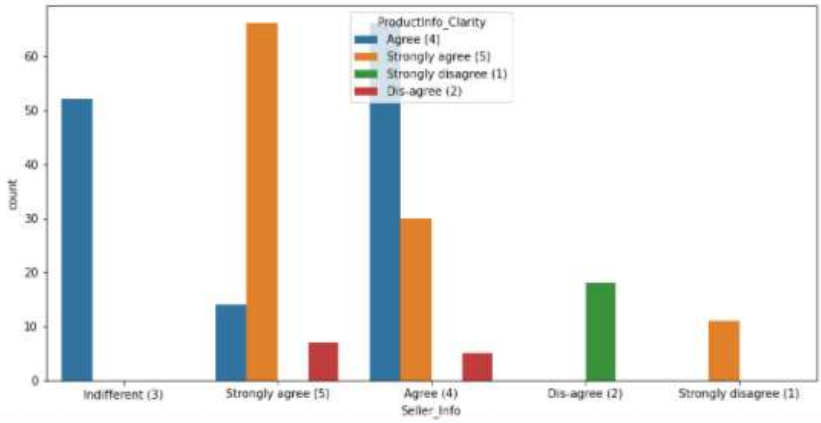
In [76]:
```
plt.figure(figsize=[12,6])
sns.countplot(df['ShoppingSince'],hue=df['Gender'])
```

Out[76]:
```
<AxesSubplot:xlabel='ShoppingSince', ylabel='count'>
```



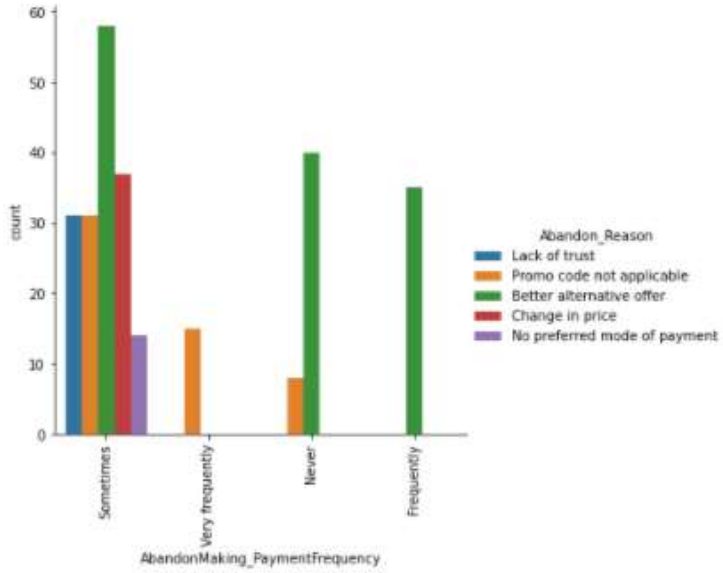
In [81]:
```
#Factor plot for Seller_ProductInfo
plt.figure(figsize=(12,6))
sns.countplot(df['Seller_Info'],hue=df['ProductInfo_Clarity'])
```

Out[81]:
```
<AxesSubplot:xlabel='Seller_Info', ylabel='count'>
```



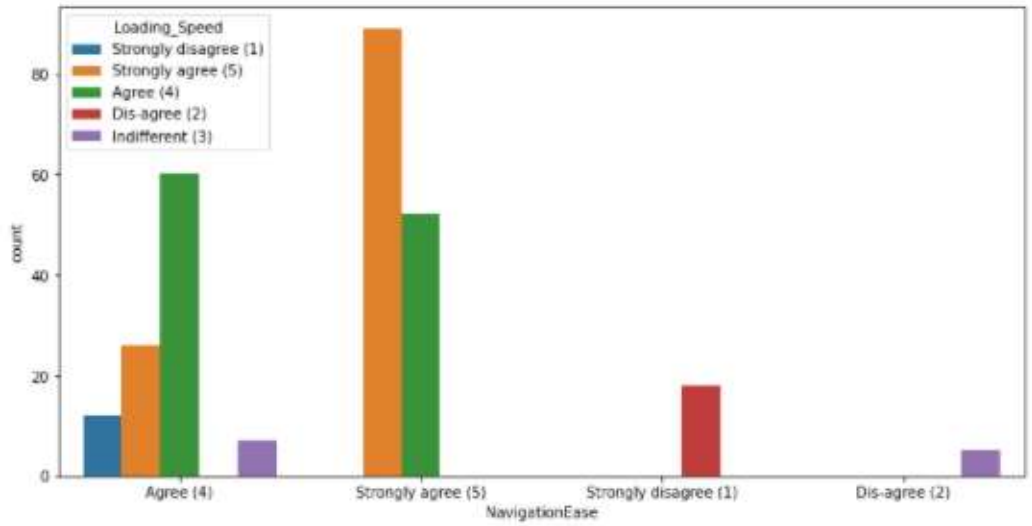
Maximum people agrees to have all relevant information on listed products and agrees with complete information on listed seller and product being offered is important for purchase decision.

In [80]:
```
#Factor plot for Abandon_Frequency
sns.factorplot('AbandonMaking_PaymentFrequency',kind='count',data=df,hue='Abandon_Reason')
plt.xticks(rotation=90);
```



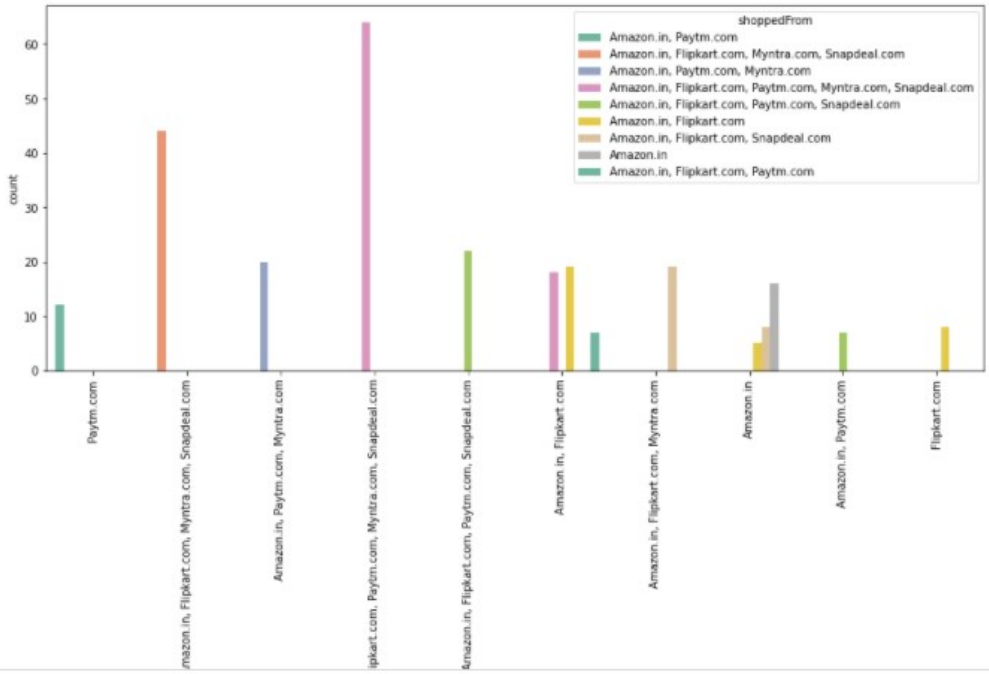
In [83]:
```
plt.figure(figsize=(12,6))
sns.countplot(df['NavigationEase'],hue=df['Loading_Speed'])
```

Out[83]: <AxesSubplot:xlabel='NavigationEase', ylabel='count'>

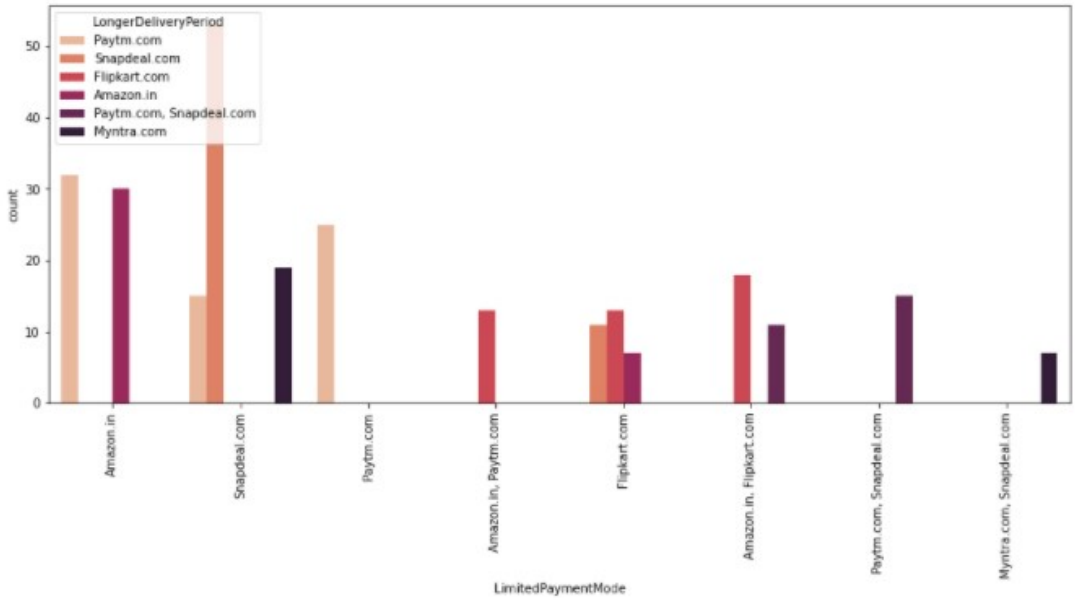


Maximum customers agrees to get easy navigation in website and wants to have loading and processing speed.

[86]:
```
#Count plot for Easy_WebApp
plt.figure(figsize=[15,6])
sns.countplot('EasyWebApp',data=df,hue='shoppedFrom',palette="Set2")
plt.xticks(rotation=90);
```

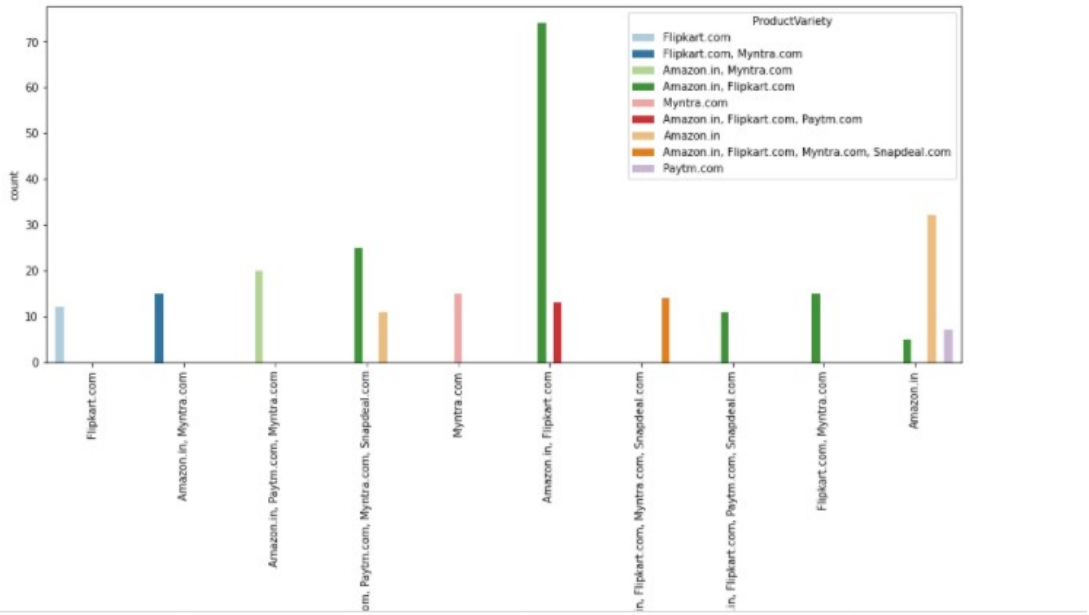


[96]:
```
#Count plot for Limited_PaymentMode
plt.figure(figsize=[15,6])
sns.countplot('LimitedPaymentMode',data=df,hue='LongerDeliveryPeriod',palette="rocket_r")
plt.xticks(rotation=90);
```

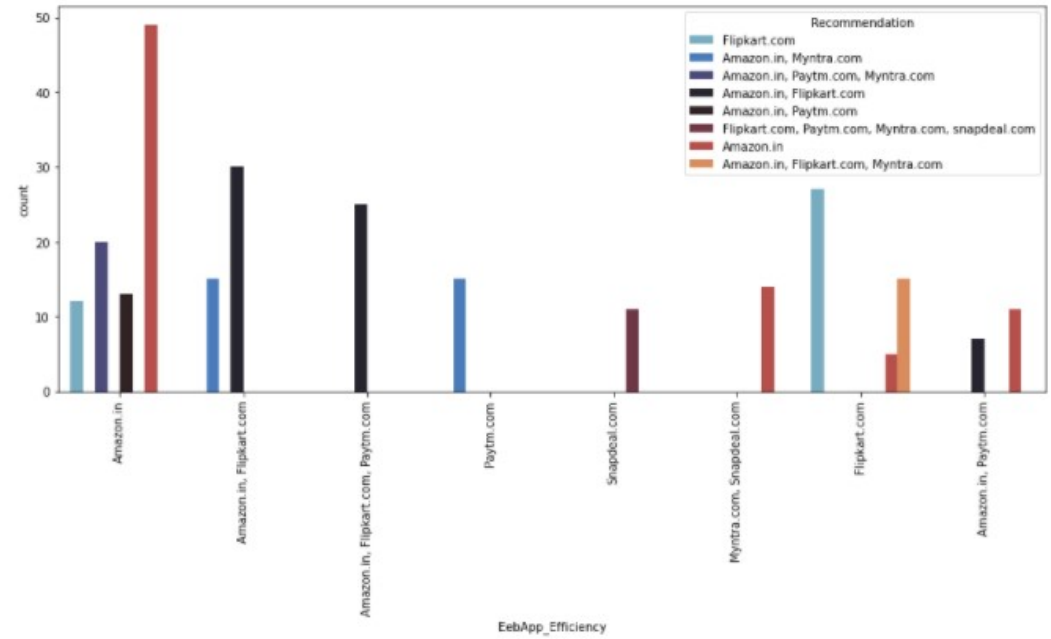


Snapdeal has a drawback of having limited payment modes and also it has a complaint of late delivery.

```
88]: #Count plot for Visually_AppealingWebApp
     plt.figure(figsize=[15,6])
     sns.countplot('VisuallyAppealingWebApp',data=df,hue='ProductVariety',palette="Paired")
     plt.xticks(rotation=90);
```

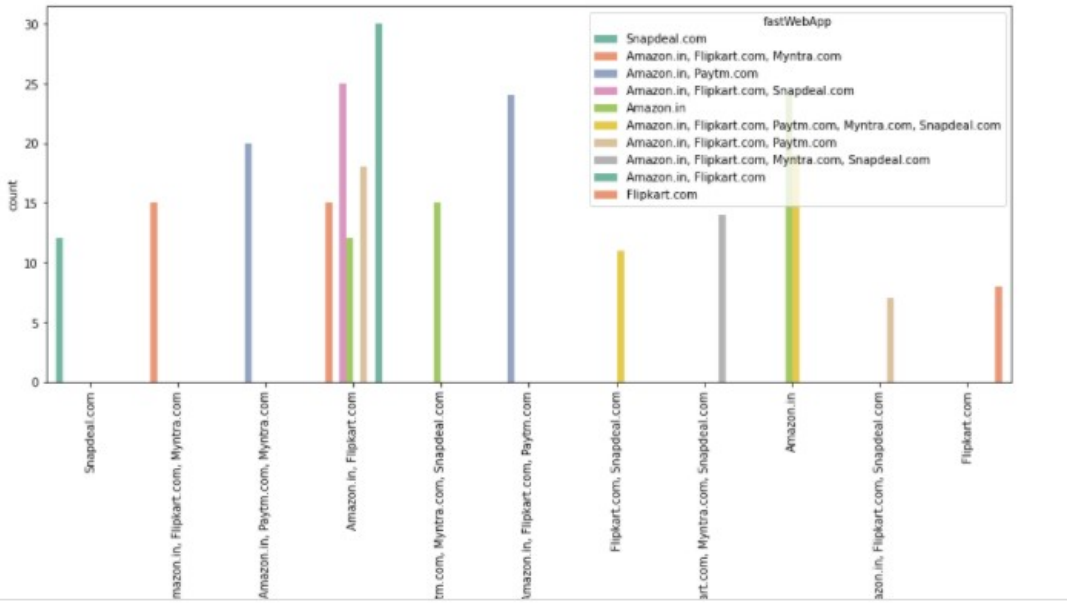



```
100]: #Count plot for WebApp_Efficiency
      plt.figure(figsize=[15,6])
      sns.countplot('EebApp_Efficiency',data=df,hue='Recommendation',palette="icefire")
      plt.xticks(rotation=90);
```

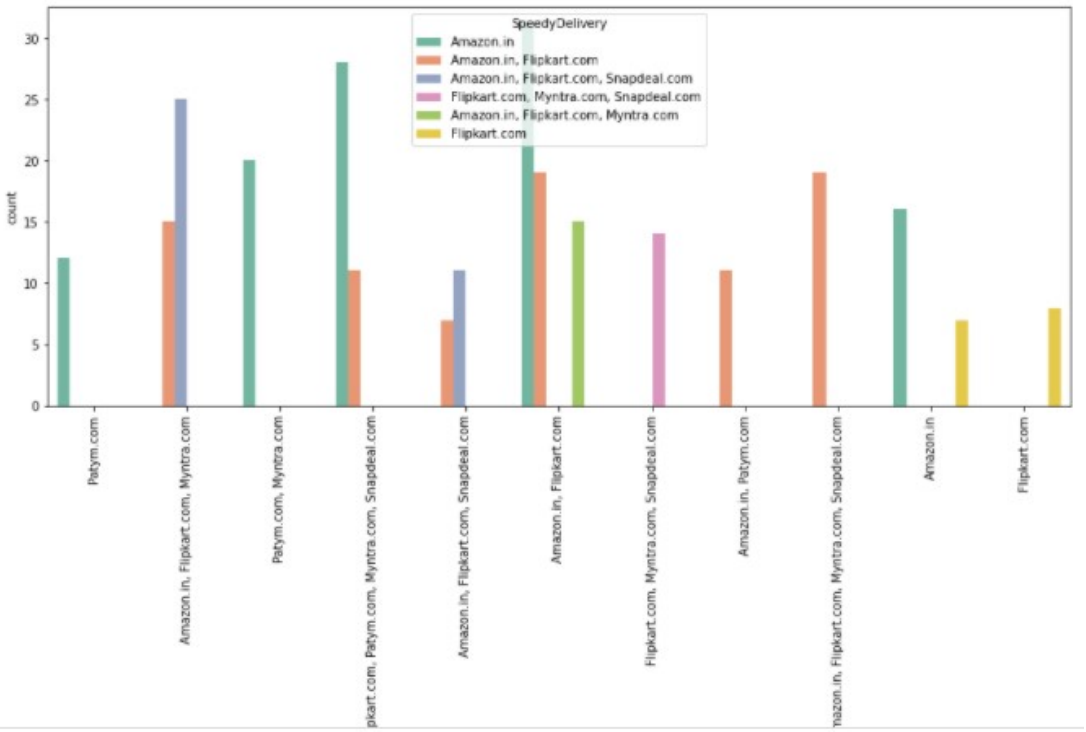


Maximum customers says that Amazon Website efficiency is good and also same is recommended by most of the customer

[91]:
```
#Count plot for Complete_ProductInfo
plt.figure(figsize=[15,6])
sns.countplot('CompleteProductInfo',data=df,hue='fastWebApp',palette="Set2")
plt.xticks(rotation=90);
```



[94]:
```
#Count plot for PaymentOptions_Availability
plt.figure(figsize=[15,6])
sns.countplot('PaymentOptionsAvailability',data=df,hue='SpeedyDelivery',palette="Set2")
plt.xticks(rotation=90);
```



Interpretation of the Results

- Maximum number of customers satisfied and have high retention with Amazon.com and Flipkart.com
- Myntra.com and Snapdeal.com have high rate of attrition

Important Observations on the basis of Customer's perception

- Majority customers believe that content on the website must be easy to read and understand.
- 70% people believe that complete information on listed seller and product is important for purchase decision.
- 90% people believe that all relevant information on listed products must be stated clearly.
- For more than 90% people, ease of navigation of website, loading or processing speed as well as user friendly Interface of the website is important.
- Other important parameters are - Convenient Payment Methods, privacy of customers, return and replacement policy.

Let's Discuss Important Factors for Satisfaction, based on above plotted charts

- Around 20% people don't display quality information on the website which help to improves satisfaction of customers. The other 80% agree that it is important.
- Almost all participants believe that sser derive satisfaction while shopping on a good quality website or application.
- Very high percentage of people believes that net benefit derived from shopping online can lead to user satisfaction.
- Most People believe that user satisfaction cannot exist without trust.

CONCLUSION

AMAZON: Most recommended websites with attractive web-page layout, easy to use, relevant descriptive information, product offers, reliability of website, quickness to complete purchase, trust worthiness.
What can be improved: Takes longer time to login, Late declaration or price during sales and promotion, frequent disruption when moving from one page to another, Limited mode of payment on most of products.

FLIPKART: This is the 2nd most recommended website with fast loading page, security of financial information, trust worthiness, several payments modes, website is as efficient as before.
What can be improved: Takes longer time in displaying graphics, late declaration of price during sales and promotion.

PAYTM: Reliability of website, speedy delivery of products, quickness in purchase.
What can be improved: Longer page loading time, longer delivery period, and late declaration of price during sales and promotion.

MYNTRA: Myntra stands on 3rd most recommended websites with easy to use, wild variety of product offers, several payment methods, attractive visual appealing web-page layout.
What can be improved: Relevant information about product, website loading speed, speedy delivery of products, websites is not much efficient as before.

SNAPDEAL: Least recommended website having less page loading time.
What can be improved: Limited mode of payments, frequent disruption while moving from one page to another, longer delivery period, customer's privacy information, reliability of website, offers on product, and must be an attractive web-page layout.

Few ways to improve customer retention can be as follows:

• User-friendly websites: 87% customers agree with user friendly website interface. The online stores should invest heavily in creating user friendly apps and websites, so that the customers do not have to work around much and the overall shopping experience is smooth,& shoppers get what they want faster and without running into unnecessary complexity that can clog up the path to purchase. It should focus on the user experience by providing shopping categories, filters, and comparison capabilities. User-friendliness requires that your website works on all browsers and devices

• Improve return & refund policy: Return policies are an essential feature of any ecommerce website .90% customers agree that return and replacement policy helps them making purchase decision. It is evident from the fact that people cannot actually try & touch the products, they are purchasing before it reaches home and they would want to return or replace in case of dissatisfaction. Online shopping websites should make strategies around easy return and replacement policy if they want to retain their customers. This is another trust-building feature of online selling. , which reassures buyers that if they are unhappy or just need a different size, the brand is there for them.

• Privacy Policy: Being able to guarantee the privacy of the customer:92% Customer agreed to this. Costumers are concerned about the unauthorized access to their data. Building trust with the customers is crucial for any e-commerce website. An e-commerce privacy policy statement makes business more transparent regarding how you collect, manage, and use data from site visitors.

• Displaying quality Information on the website: 90% customers agree all relevant information on listed products must be stated clearly. Content is one of the crucial challenges for any e-commerce website. it's simply not enough to just list a product name and image on a product page and expect the orders to roll in. Compiling a compelling array of product data, whether that's dimensions, MPNs or spec sheets can all help to convert customers better.

• Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.): 90% customers agreed to it. In case one channel is not available, customers can reach out to multiple channels which again are an important factor. Being able to communicate easily can make the difference in both their shopping experience as well as fulfill business goals of online store.

When customers are satisfied with a company or service, there is a high possibility that they will share their experience with other people Therefore it is crucial for E-commerce to take into account their customer satisfaction because this will retain customer loyalty as well as attract potential customers.

Limitations of this work and Scope for Future Work:

- 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 then likes to shop in the particular websites regularly.

- Getting feedbacks from the customers is also one of the important things 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..

References

1. Al-Dmour, H., Alshuraideh, M., Salehih, S.: A study of Jordanians' television viewers habits. Life Sci. J. 11(6), 161–171 (2014)

2. Alshurideh, M.: The factors predicting students' satisfaction with universities' healthcare clinics' services: a case-study from the Jordanian higher education sector. Dirasat. Adm. Sci. 161(1524), 1–36 (2014)

3. Preikschas, M.W., Cabanelas, P., Rüdiger, K., Lampón, J.F.: Value co-creation, dynamic capabilities and customer retention in industrial markets. J. Bus. Ind. Mark. 32(3), 409–420 (2017)