FLIP ROBO

CUSTOMER RETENSTION CASE STUDY PROJECT

Submitted by :-

K.Vineeth Nair

ACKNOWLEDGMENT

I would like to thank Flip Robo Technologies for providing me with the opportunity to work on this project from which I have learned a lot. Some of the reference sources are as follows:

➢ Stack Overflow
➢ Medium.com
➢ scikit-learn.org
➢ Python official documentation

Problem statement:

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.

Problem definition:

In this project, a dataset was provided containing the details of the participants of a survey, along with their online shopping experiences, preferences, and opinions regarding various ecommerce websites.
The Dataset was first checked for null values, and then the various feature columns were analyzed. Exploratory Data analysis was conducted to investigate the relationships that existed between the columns, using various visualization techniques.
The dataset was worked with to study and understand with several perceived risks helped to understand Customer retention and loyalty to various ecommerce websites.

Data preparation cleaning:

- Reading the CSV file and doing initial statistical analysis (shape, values etc)

- Data Pre-processing: Reading the uniques values for each column and removing those which won't be significant in the analysis further.

- Create a new data frame to proceed with the analysis further

EDA Concluding Remarks:

- Find patterns of data through visualization and reveal the hidden trends from data.

- Using both matplotlib and seaborn library to visualize the data.

- Finding relationships between features.

- Analysing both the numerical and the categorical columns separately.

HARDWARE AND SOFTWARE REQUIREMENTS AND TOOLS USED

HARDWARE: ACER ASPIRE E15
SOFTWARE:
- Jupyter Notebook (Anaconda 3) – Python 3
- Microsoft Office 365 Package

Libraries used:

```python
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

from sklearn.linear_model import LogisticRegression
from sklearn.svm import SVC
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn.neighbors import KNeighborsClassifier

from sklearn.preprocessing import LabelEncoder, StandardScaler
from sklearn.metrics import accuracy_score,confusion_matrix,classification_report,roc_curve,roc_auc_score
from sklearn.model_selection import train_test_split,GridSearchCV,KFold,cross_val_score
from statsmodels.stats.outliers_influence import variance_inflation_factor
from sklearn.utils import resample
from scipy.stats import zscore
import warnings
warnings.filterwarnings('ignore')
```

About the Dataset:

```python
In [2]: # Loading Dataset
        df = pd.read_excel('customer_retention_dataset.xlsx')
        df.head()
```

Out[2]:

	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\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	S
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

Exploratory Data Analysis (EDA)

Firstly to understand about the dataset, we should know how much data is there in the dataset.

```
# Checking the dimensions of this dataset.
df.shape
```

```
(269, 71)
```

- Data set contains 269 rows and 71 columns.

```
# checkng the columns present in the dataset.
df.columns
```

```
Index(['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\t                                                              ',
       '10 What is the operating system (OS) of your device?\t\t\t\t\t                                                              ',
       '11 What browser do you run on your device to access the website?\t\t\t
',
       '12 Which channel did you follow to arrive at your favorite online store for the first time?
',
       '13 After first visit, how do you reach the online retail store?\t\t\t\t
',
       '14 How much time do you explore the e- retail store before making a purchase decision?
',
       '15 What is your preferred payment Option?\t\t\t\t\t\t
',
       '16 How frequently do you abandon (selecting an items and leaving without making payment) your shopping cart?\t\t\t\t\t
\t\t                                                              ',
       '17 Why did you abandon the "Bag", "Shopping Cart"?\t\t\t\t\t
',
       '18 The content on the website must be easy to read and understand',
       '19 Information on similar product to the one highlighted  is important for product comparison',
       '20 Complete information on listed seller and product being offered is important for purchase decision.',
       '21 All relevant information on listed products must be stated clearly',
       '22 Ease of navigation in website', '23 Loading and processing speed',
       '24 User friendly Interface of the website',
       '25 Convenient Payment methods',
       '26 Trust that the online retail store will fulfill its part of the transaction at the stipulated time',
       '27 Empathy (readiness to assist with queries) towards the customers',
       '28 Being able to guarantee the privacy of the customer',
```

- There is space and characters present in the dataset.

```
# Remove tabs,space and digits present in column names.
from string import digits
df.columns = df.columns.str.replace('\t','')
remove_digits = str.maketrans('', '', digits)
df.columns = df.columns.str.translate(remove_digits)
df.columns = df.columns.str.strip()
```

- Tabs and Spaces between the column headings are successfully removed.

```
# Cheching the datatypes of all the columns.
df.dtypes
```

```
Gender of respondent                                            object
How old are you?                                                object
Which city do you shop online from?                             object
What is the Pin Code of where you shop online from?              int64
Since How Long You are Shopping Online ?                        object
                                                                  ...
Longer delivery period                                          object
Change in website/Application design                            object
Frequent disruption when moving from one page to another        object
Website is as efficient as before                               object
Which of the Indian online retailer would you recommend to a friend?    object
Length: 71, dtype: object
```

- Seems that most of the columns are object variables present in the dataset.
- only 1 int i.e, for column "What is the Pin Code of where you shop online from? "

```
# For getting the overview of the dataset
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 269 entries, 0 to 268
Data columns (total 71 columns):
 #   Column                                                      Non-Null Coun
t  Dtype
---  ------                                                      ------------
-  -----
 0   Gender of respondent                                        269 non-null
object
 1   How old are you?                                            269 non-null
object
 2   Which city do you shop online from?                         269 non-null
object
 3   What is the Pin Code of where you shop online from?         269 non-null
int64
 4   Since How Long You are Shopping Online ?                    269 non-null
object
 5   How many times you have made an online purchase in the past  year?    269 non-null
object
```

- There are 1 int and 70 object variables present in the dataset.

```
# checking the null values in the dataframe
df.isnull().sum()
```

```
Gender of respondent                                                      0
How old are you?                                                          0
Which city do you shop online from?                                       0
What is the Pin Code of where you shop online from?                       0
Since How Long You are Shopping Online ?                                  0
                                                                         ..
Longer delivery period                                                    0
Change in website/Application design                                      0
Frequent disruption when moving from one page to another                  0
Website is as efficient as before                                         0
Which of the Indian online retailer would you recommend to a friend?      0
Length: 71, dtype: int64
```

- Seems there is no missing value present in the dataset.

```
df.nunique()
```

```
Gender of respondent                                                      2
How old are you?                                                          5
Which city do you shop online from?                                      11
What is the Pin Code of where you shop online from?                      39
Since How Long You are Shopping Online ?                                   5
                                                                         ..
Longer delivery period                                                    6
Change in website/Application design                                      7
Frequent disruption when moving from one page to another                  8
Website is as efficient as before                                         8
Which of the Indian online retailer would you recommend to a friend?      8
Length: 71, dtype: int64
```

- All the Variables or Features are Categorical type.

```
# Summary statistics
df.describe()
```

	What is the Pin Code of where you shop online from?
count	269.000000
mean	220465.747212
std	140524.341051
min	110008.000000
25%	122018.000000
50%	201303.000000
75%	201310.000000
max	560037.000000

- As we have only 1 int variable, we have got description of one column only.

Data Visualization

```python
plt.figure(figsize=(15,5),facecolor='white')
sns.countplot(df['Gender of respondent'])
plt.show()
df['Gender of respondent'].value_counts() , print(df['Gender of respondent'].value_counts(normalize=True))
```
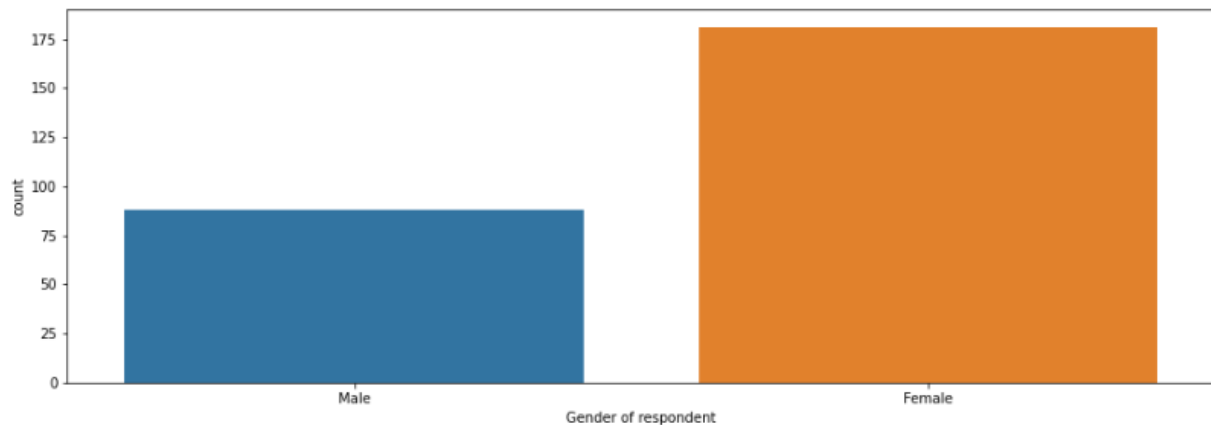


```
Female    0.672862
Male      0.327138
Name: Gender of respondent, dtype: float64

(Female    181
 Male       88
 Name: Gender of respondent, dtype: int64,
 None)
```
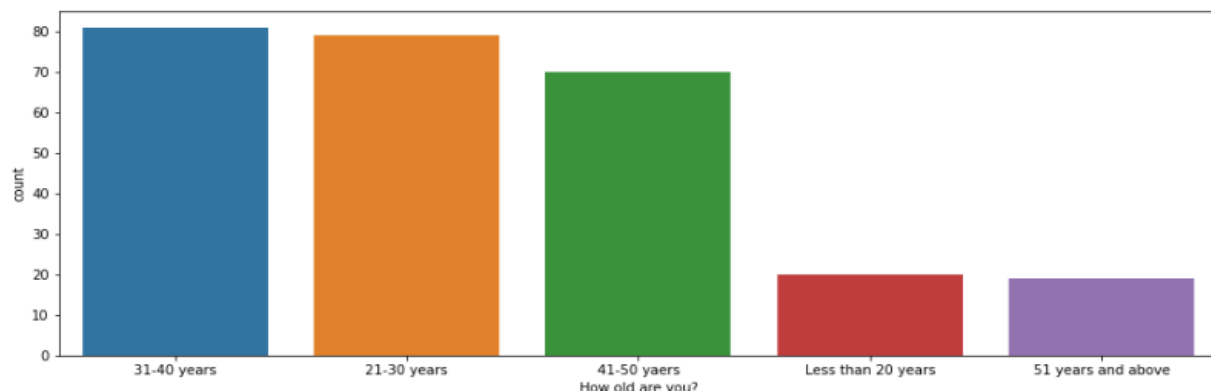
As we can See Majority (181 count/ 67.3%) are Female compare to Male (88 count/ 32.7%).

```python
df['How old are you?'].value_counts()
```

```
1    81
0    79
2    70
4    20
3    19
Name: How old are you?, dtype: int64
```

```python
plt.figure(figsize=(15,5),facecolor='white')
sns.countplot(df['How old are you?'])
```

```
<AxesSubplot:xlabel='How old are you?', ylabel='count'>
```


- Maximum count is 81 between the age group of 31-40 years and Minimum count is 19 of the age group of 51 years and above.

```
plt.figure(figsize=(15,5),facecolor='white')
sns.countplot(hue='How old are you?',x='Gender of respondent',data=df)
plt.show()
```
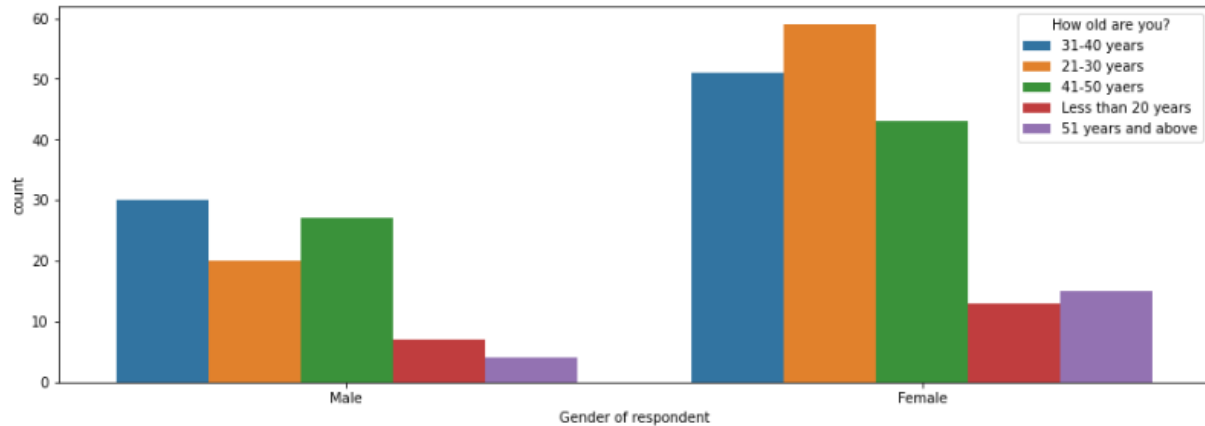

- Maximum count in male group is between the age group of 31-40 years and Minimum count is of 51 years and above.
- Maximum count in female group is between the age group of 21-30 years and Minimum count is less than 20 years.

```
df['Which city do you shop online from?'].value_counts()
```

```
Delhi            58
Greater Noida    43
Noida            40
Bangalore        37
Karnal           27
Solan            18
Ghaziabad        18
Gurgaon          12
Merrut            9
Moradabad         5
Bulandshahr       2
Name: Which city do you shop online from?, dtype: int64
```

```
plt.figure(figsize=(15,5),facecolor='white')
sns.countplot(df['Which city do you shop online from?'], palette="Set1")
```

```
<AxesSubplot:xlabel='Which city do you shop online from?', ylabel='count'>
```
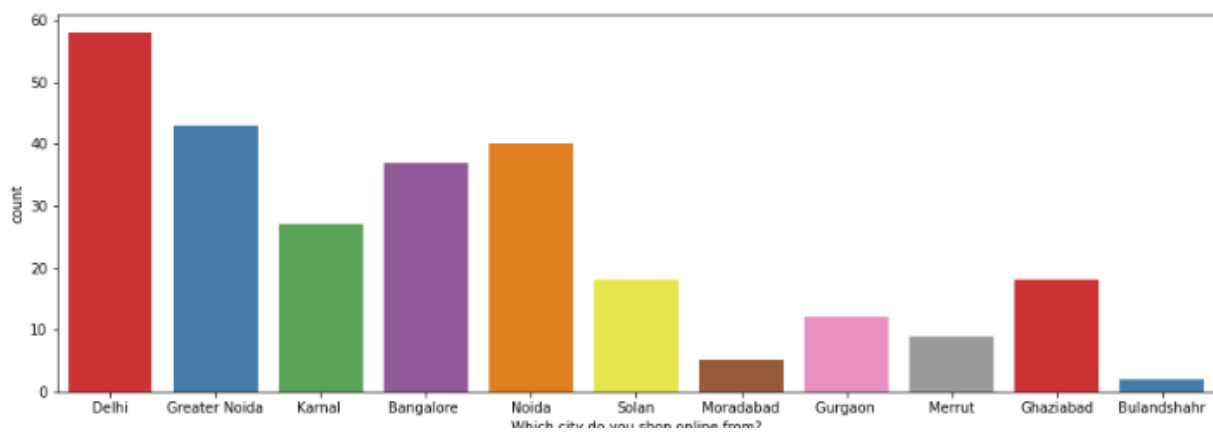

- Maximum count is 58 from Delhi and Minimum count is 2 from Bulandshahr.

```
plt.figure(figsize=(15,5),facecolor='white')
sns.countplot(df['Which city do you shop online from?'], hue=df['Gender of respondent'])
```

<AxesSubplot:xlabel='Which city do you shop online from?', ylabel='count'>
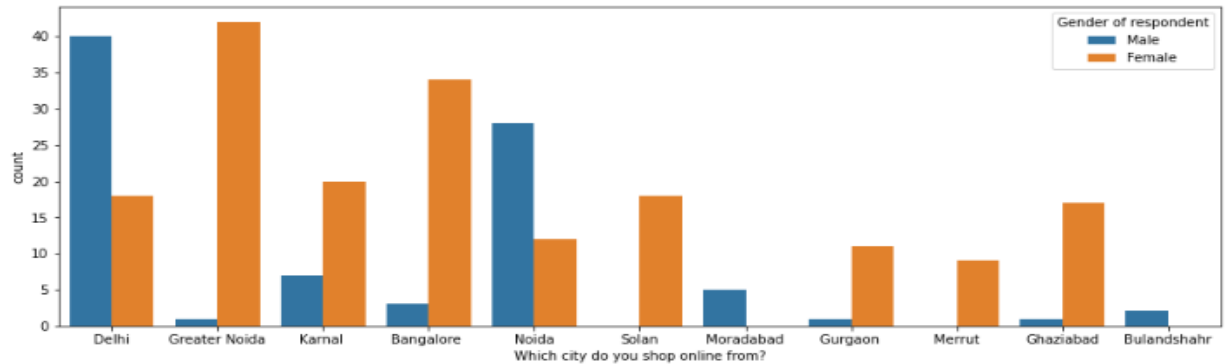
- In Delhi mostly Male prefer to shop online whereas females from Greater Noida prefer to shop online.

```
plt.figure(figsize=(15,5))
sns.countplot(df['Which city do you shop online from?'],hue=df['How old are you?'])
```

<AxesSubplot:xlabel='Which city do you shop online from?', ylabel='count'>
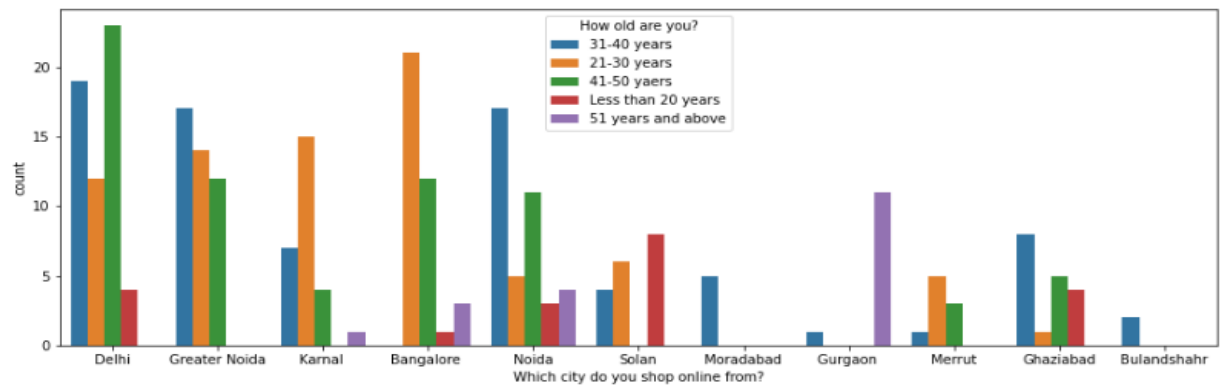
- Customers between 31-40 and 41-50 years of age prefer to shop online compare to other age group in Delhi.
- Customers between 21-30 years of age prefer to shop online compare to other age group in Bangalore.

```
plt.figure(figsize=(15,5))
sns.countplot(df['Since How Long You are Shopping Online ?'],hue=df['How old are you?'])
```

<AxesSubplot:xlabel='Since How Long You are Shopping Online ?', ylabel='count'>
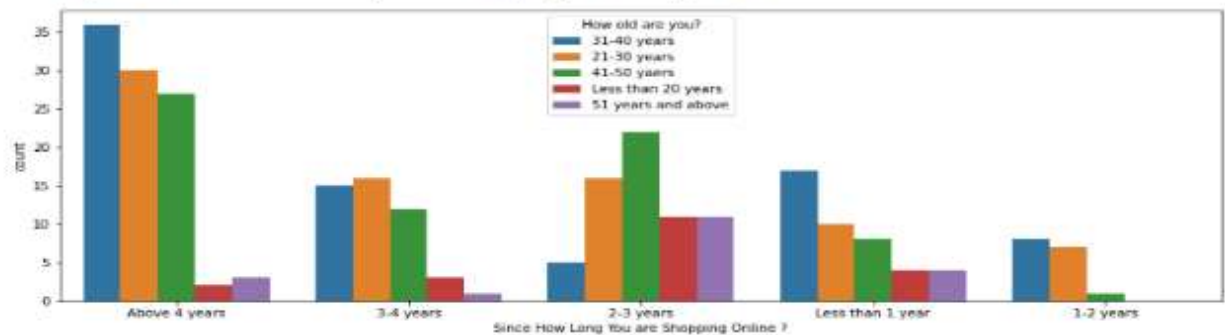
- Customers between 21-30, 31-40 and 41-50 years of age are shopping online for more than 4 Years.

```
plt.figure(figsize=(15,5),facecolor='white')
sns.countplot(df['How do you access the internet while shopping on-line?'], palette="Set1")
```

<AxesSubplot:xlabel='How do you access the internet while shopping on-line?', ylabel='count'>
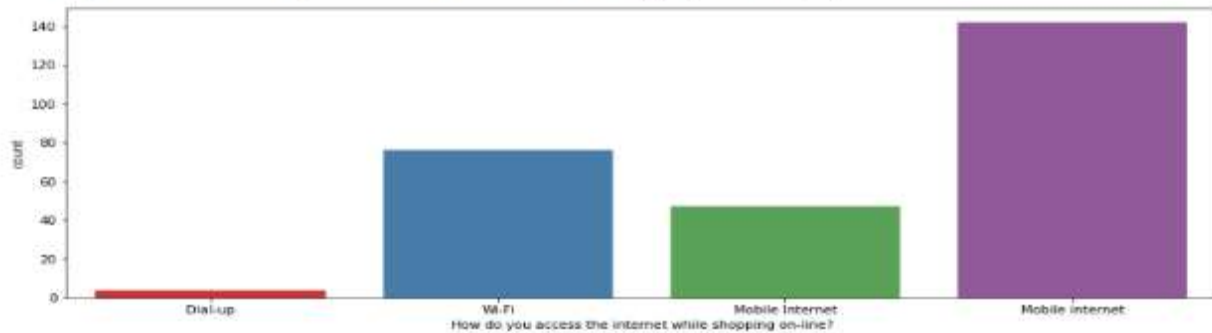
- Customers mostly use Mobile internet and Wi-Fi than Dial up for online shopping.

```
plt.figure(figsize=(15,5),facecolor='white')
sns.countplot(df['Which device do you use to access the online shopping?'], palette="Set1")
```

<AxesSubplot:xlabel='Which device do you use to access the online shopping?', ylabel='count'>
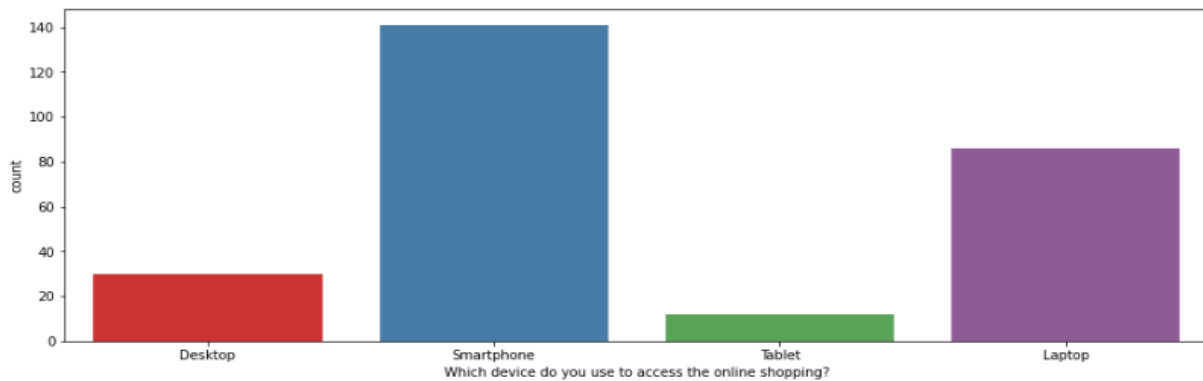
- Customers use Smartphone and laptop for online shopping than Desktop and Tablet.

```
plt.figure(figsize=(15,5),facecolor='white')
sns.countplot(df['Which device do you use to access the online shopping?'], hue=df['How old are you?'])
```

<AxesSubplot:xlabel='Which device do you use to access the online shopping?', ylabel='count'>
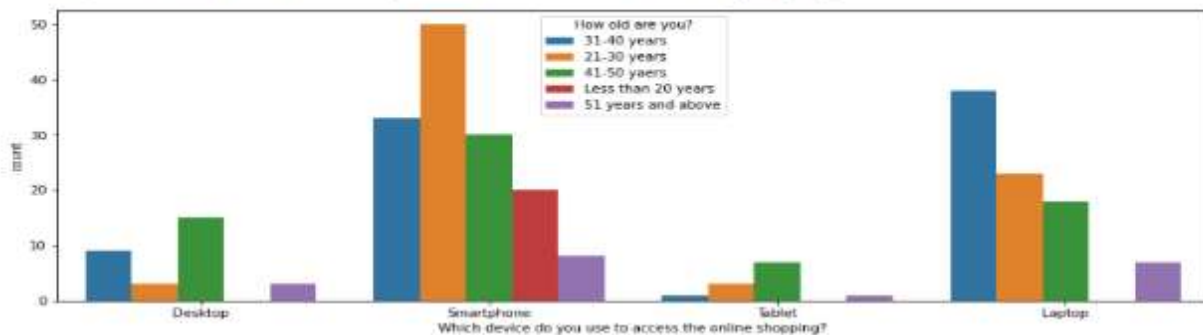
- Majority Customers using Smartphone for online shopping are between 21-30 years of age group.
- Whereas using Laptop for online shopping are between 31-40 years of age group.

```
plt.figure(figsize=(15,5),facecolor='white')
sns.countplot(df['What is the operating system (OS) of your device?'], hue=df['which device do you use to access the online shopp
```

<AxesSubplot:xlabel='What is the operating system (OS) of your device?', ylabel='count'>
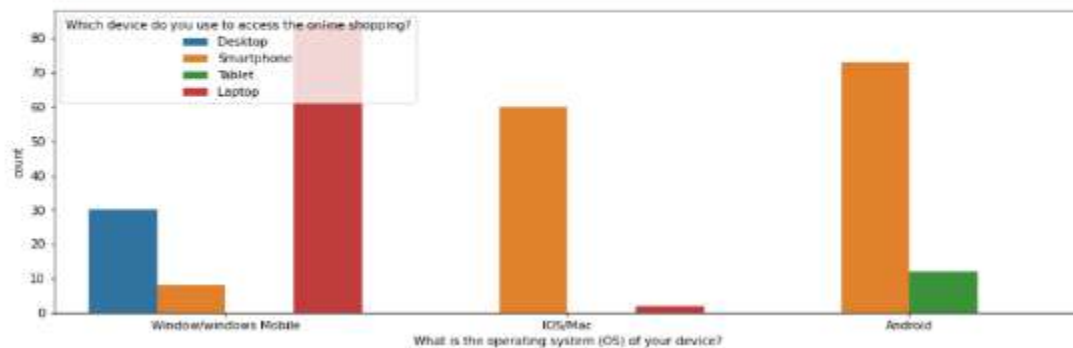
- Majority Customers using Smartphone for online shopping are Andriod or IOS/Mac operating System users.
- Whereas using Laptop users are using Windows operating System.

```
plt.figure(figsize=(15,5),facecolor='white')
sns.countplot(df['What is the operating system (OS) of your device?'], hue=df['How old are you?'])
```

<AxesSubplot:xlabel='What is the operating system (OS) of your device?', ylabel='count'>
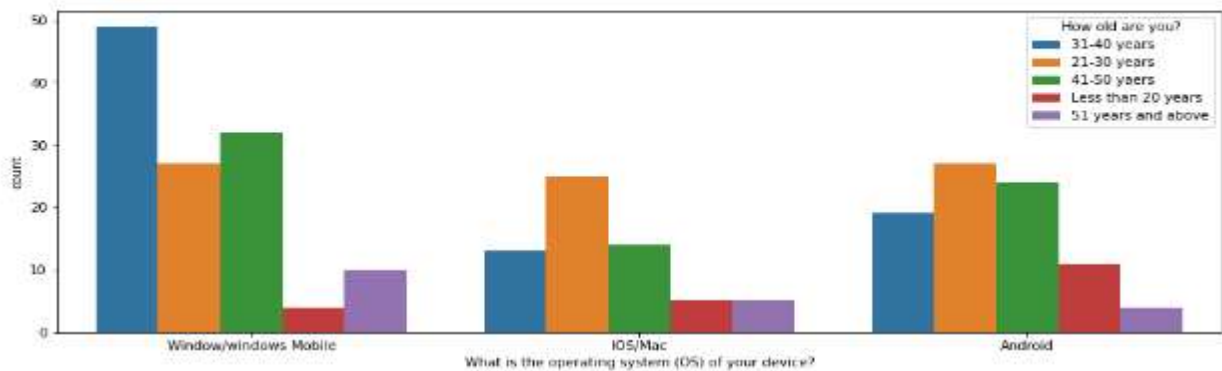
- Majority Customers using Windows operating System for online shopping are of 31-40 years of age group.

```
plt.figure(figsize=(15,5),facecolor='white')
sns.countplot(df['What browser do you run on your device to access the website?'], palette="Set1")
```

<AxesSubplot:xlabel='What browser do you run on your device to access the website?', ylabel='count'>
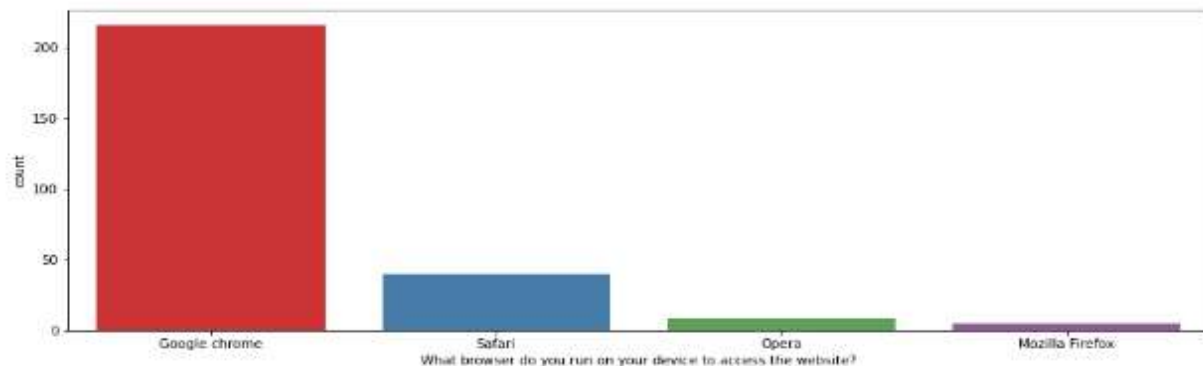

- Majority Customers using Google Chrome as web browser for online shopping.

```
plt.figure(figsize=(15,5),facecolor='white')
sns.countplot(df['What browser do you run on your device to access the website?'], hue=df['Which device do you use to access the
```

`<AxesSubplot:xlabel='What browser do you run on your device to access the website?', ylabel='count'>`
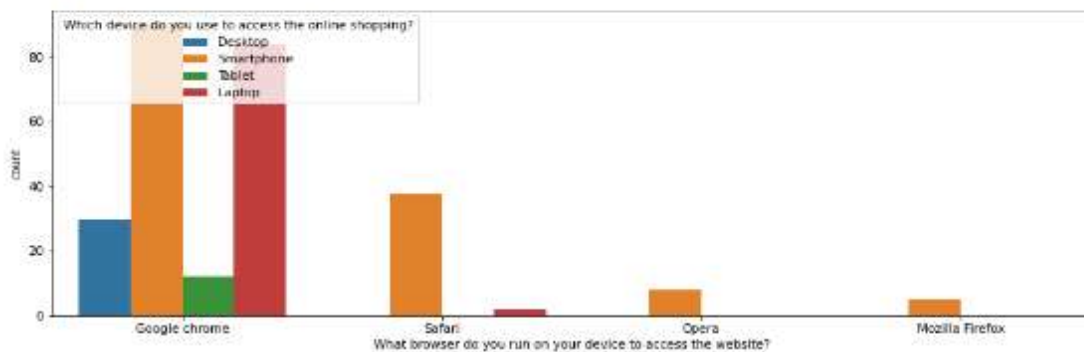

- Majority Customers using Google Chrome as web browser use Smartphone and Laptop for online shopping.

```
plt.figure(figsize=(15,5),facecolor='white')
sns.countplot(df['After first visit, how do you reach the online retail store?'], palette="Set1")
```

`<AxesSubplot:xlabel='After first visit, how do you reach the online retail store?', ylabel='count'>`
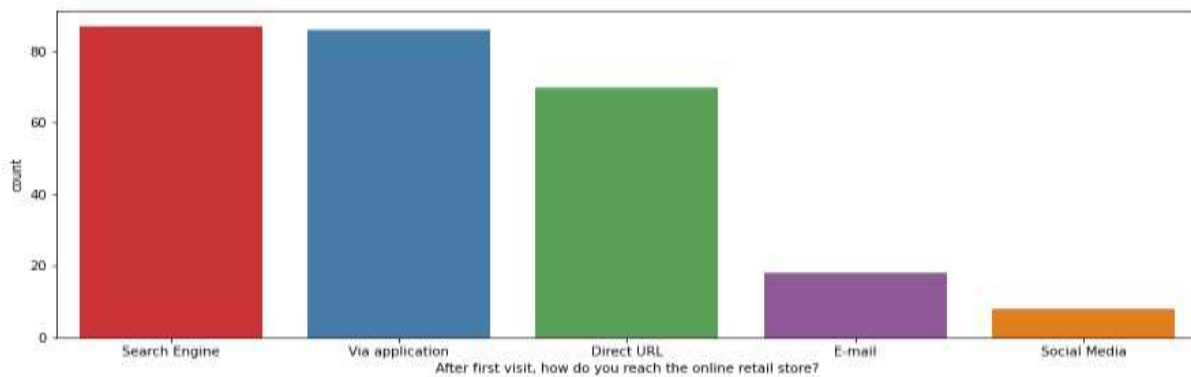

- Majority Customers reach the online retail store through Search Engine or Via application or Direct URL.

```
plt.figure(figsize=(15,5),facecolor='white')
sns.countplot(df['How much time do you explore the e- retail store before making a purchase decision?'], palette="Set1")
```

`<AxesSubplot:xlabel='How much time do you explore the e- retail store before making a purchase decision?', ylabel='count'>`
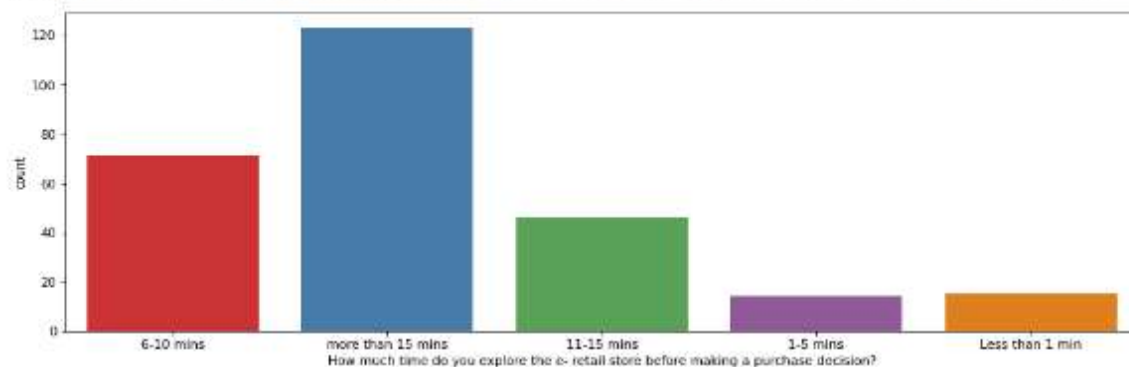
- Majority Customers explore for more than 15- mins before making a purchase decision.

```
plt.figure(figsize=(15,5),facecolor='white')
sns.countplot(df['What is your preferred payment Option?'], palette="Set1")
```

<AxesSubplot:xlabel='What is your preferred payment Option?', ylabel='count'>
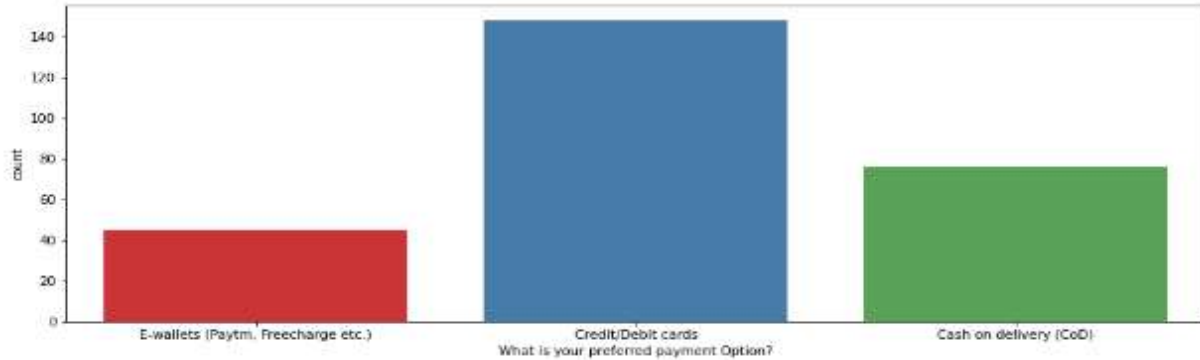
- Majority Customers prefer for Card/Debit cards as payment option for online shopping.

```
plt.figure(figsize=(15,5))
print(df['How frequently do you abandon (selecting an items and leaving without making payment) your shopping cart?'].value_count
sns.countplot(df['How frequently do you abandon (selecting an items and leaving without making payment) your shopping cart?'])
```

Sometimes   171
Never   48
Frequently   35
Very frequently 15
Name: How frequently do you abandon (selecting an items and leaving without making payment) your shopping cart?, dtype: int64

<AxesSubplot:xlabel='How frequently do you abandon (selecting an items and leaving without making payment) your shopping car
t?', ylabel='count'>
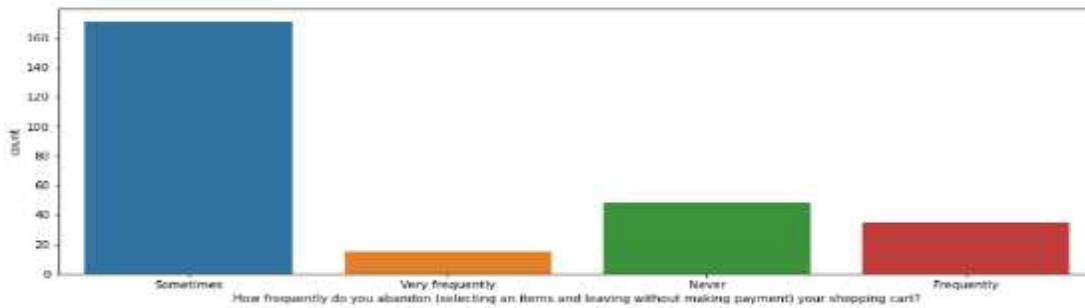

```
plt.figure(figsize=(15,5))
print(df['The content on the website must be easy to read and understand'].value_counts() )
sns.countplot(df['The content on the website must be easy to read and understand'])
```

Strongly agree (5)   164
Agree (4)   80
Strongly disagree (1) 18
Indifferent (3)   7
Name: The content on the website must be easy to read and understand, dtype: int64

<AxesSubplot:xlabel='The content on the website must be easy to read and understand', ylabel='count'>
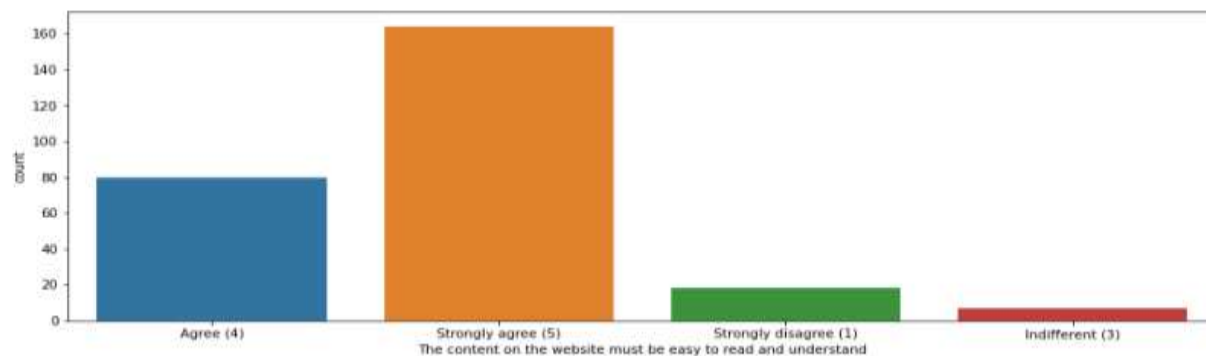

```

```python
plt.figure(figsize=(15,5))
print(df['Complete information on listed seller and product being offered is important for purchase decision.'].value_counts())
sns.countplot(df['Complete information on listed seller and product being offered is important for purchase decision.'])
```

```
Agree (4)                101
Strongly agree (5)        87
Indifferent (3)           52
Dis-agree (2)             18
Strongly disagree (1)     11
Name: Complete information on listed seller and product being offered is important for purchase decision., dtype: int64
```

```
<AxesSubplot:xlabel='Complete information on listed seller and product being offered is important for purchase decision.', ylabel='count'>
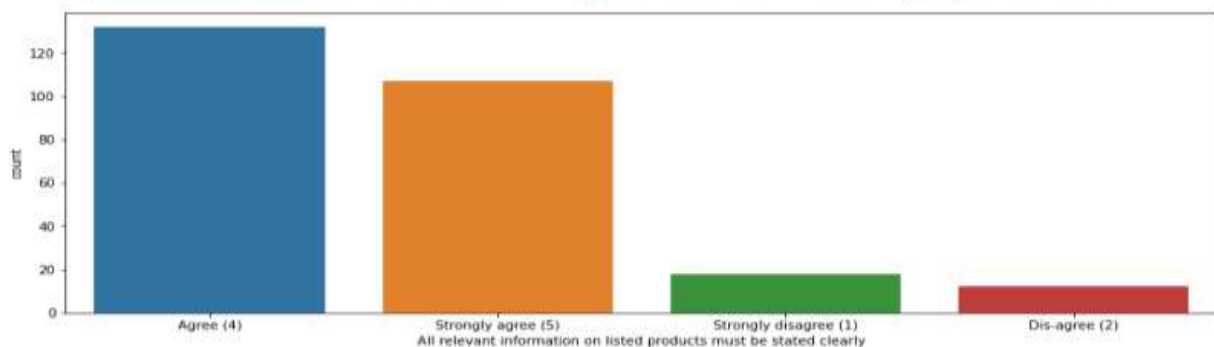```
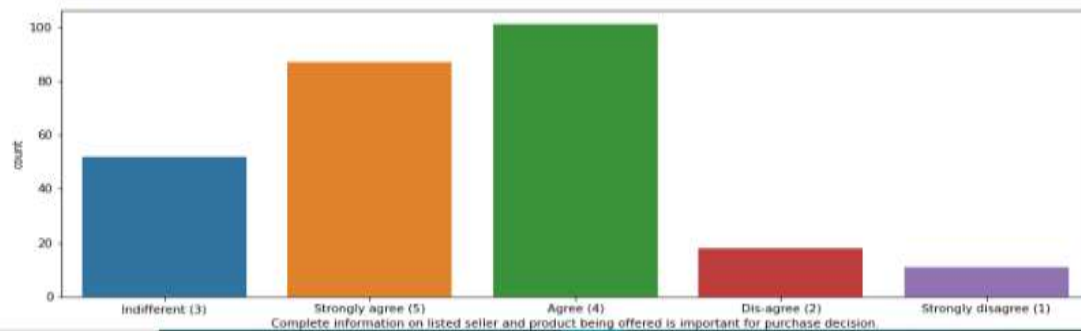


```python
plt.figure(figsize=(15,5))
print(df['All relevant information on listed products must be stated clearly'].value_counts())
sns.countplot(df['All relevant information on listed products must be stated clearly'])
```

```
Agree (4)                132
Strongly agree (5)       107
Strongly disagree (1)     18
Dis-agree (2)             12
Name: All relevant information on listed products must be stated clearly, dtype: int64
```

```
<AxesSubplot:xlabel='All relevant information on listed products must be stated clearly', ylabel='count'>
```



```python
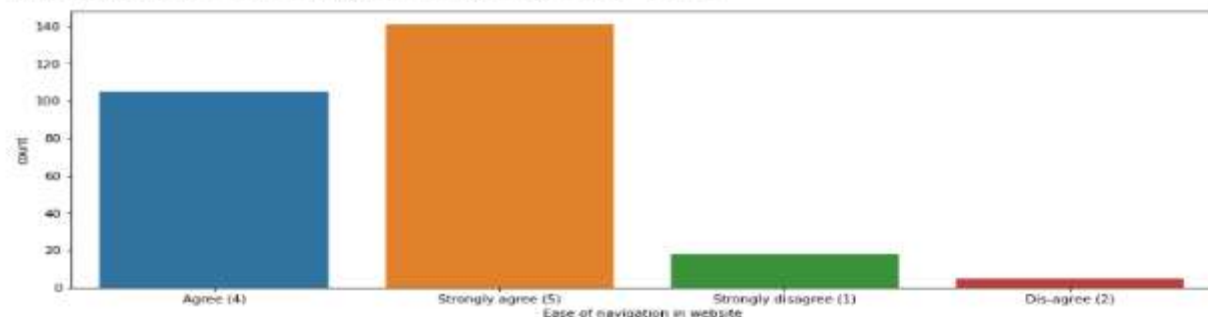plt.figure(figsize=(15,5))
print(df['Ease of navigation in website'].value_counts())
sns.countplot(df['Ease of navigation in website'])
```

```
Strongly agree (5)       141
Agree (4)                105
Strongly disagree (1)     18
Dis-agree (2)              5
Name: Ease of navigation in website, dtype: int64
```

```
<AxesSubplot:xlabel='Ease of navigation in website', ylabel='count'>
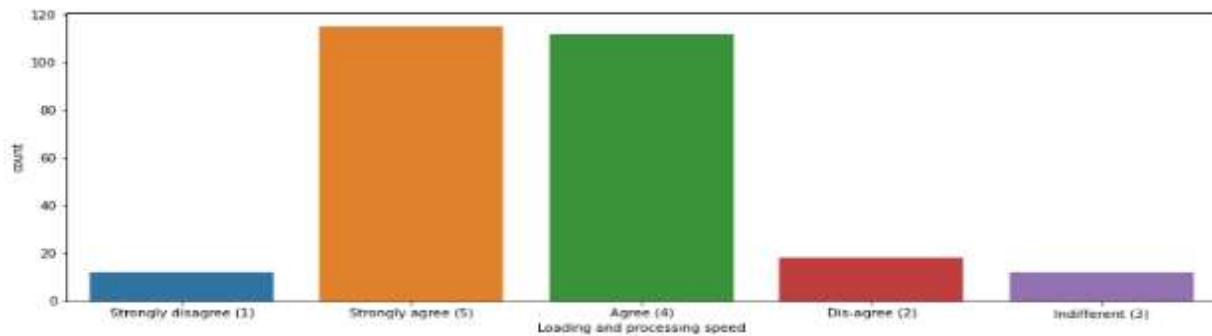```

```
plt.figure(figsize=(15,5))
print(df['Loading and processing speed'].value_counts())
sns.countplot(df['Loading and processing speed'])
```

```
Strongly agree (5)        115
Agree (4)                 112
Dis-agree (2)              18
Indifferent (3)            12
Strongly disagree (1)      12
Name: Loading and processing speed, dtype: int64
```

```
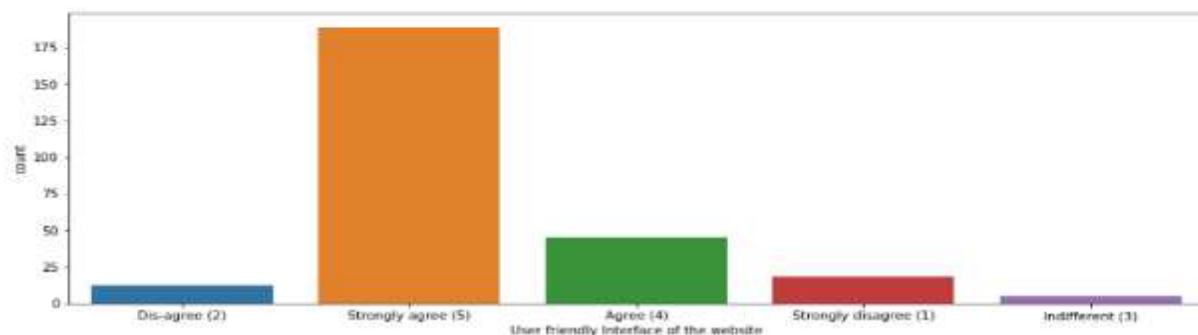<AxesSubplot:xlabel='Loading and processing speed', ylabel='count'>
```



```
plt.figure(figsize=(15,5))
print(df['User friendly Interface of the website'].value_counts())
sns.countplot(df['User friendly Interface of the website'])
```

```
Strongly agree (5)        189
Agree (4)                  45
Strongly disagree (1)      18
Dis-agree (2)              12
Indifferent (3)             5
Name: User friendly Interface of the website, dtype: Int64
```

```
<AxesSubplot:xlabel='User friendly Interface of the website', ylabel='count'>
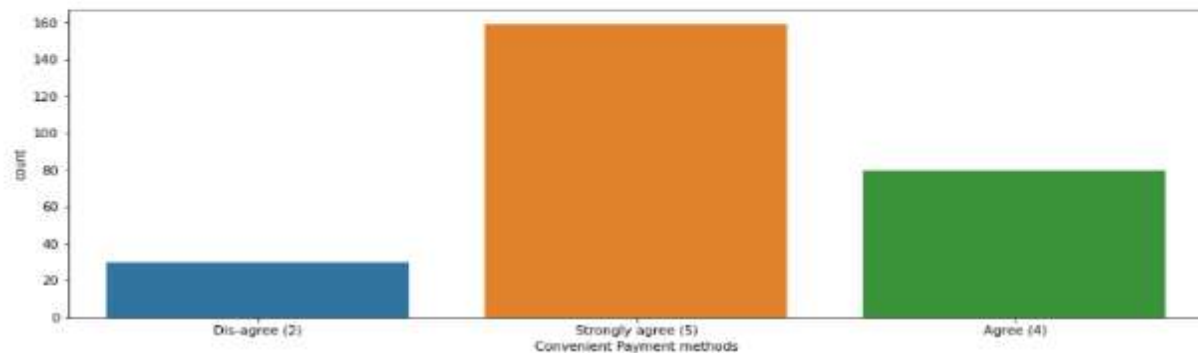```



```
plt.figure(figsize=(15,5))
print(df['Convenient Payment methods'].value_counts())
sns.countplot(df['Convenient Payment methods'])
```

```
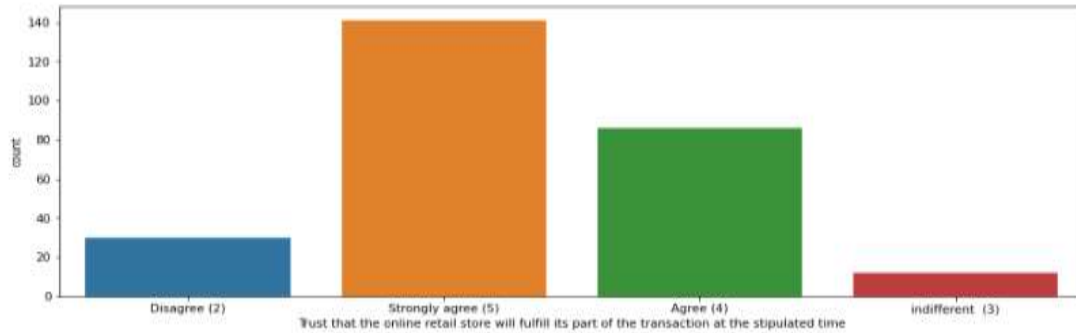Strongly agree (5)        159
Agree (4)                  80
Dis-agree (2)              30
Name: Convenient Payment methods, dtype: Int64
```

```
<AxesSubplot:xlabel='Convenient Payment methods', ylabel='count'>
```

```
plt.figure(figsize=(15,5))
print(df['Trust that the online retail store will fulfill its part of the transaction at the stipulated time'].value_counts())
sns.countplot(df['Trust that the online retail store will fulfill its part of the transaction at the stipulated time'])
```

```
Strongly agree (5)    141
Agree (4)              86
Disagree (2)           30
indifferent  (3)       12
Name: Trust that the online retail store will fulfill its part of the transaction at the stipulated time, dtype: int64
```

```
<AxesSubplot:xlabel='Trust that the online retail store will fulfill its part of the transaction at the stipulated time', ylabel='count'>
```



```
plt.figure(figsize=(15,5))
print(df['Empathy (readiness to assist with queries) towards the customers'].value_counts())
sns.countplot(df['Empathy (readiness to assist with queries) towards the customers'])
```

```
Strongly agree (5)       194
Agree (4)                 42
Strongly disagree (1)     18
indifferent (3)           15
Name: Empathy (readiness to assist with queries) towards the customers, dtype: int64
```

```
<AxesSubplot:xlabel='Empathy (readiness to assist with queries) towards the customers', ylabel='count'>
```
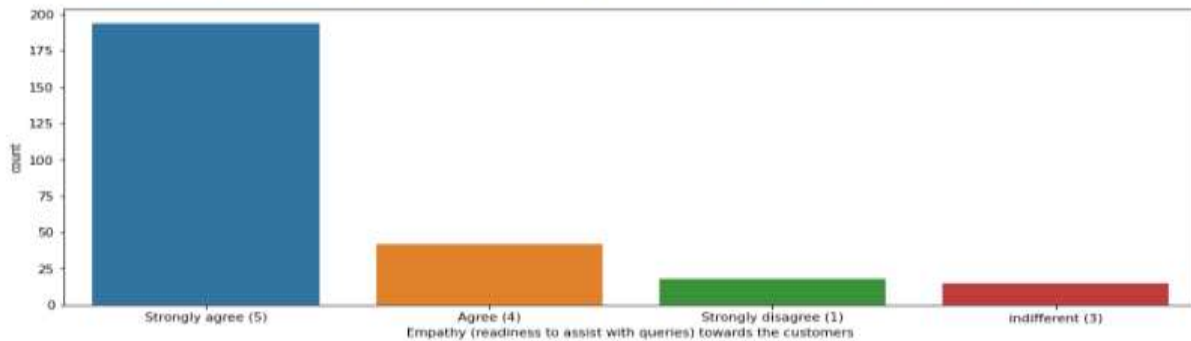


```
plt.figure(figsize=(15,5))
print(df['Being able to guarantee the privacy of the customer'].value_counts())
sns.countplot(df['Being able to guarantee the privacy of the customer'])
```

```
Strongly agree (5)    185
Agree (4)              58
indifferent (3)        26
Name: Being able to guarantee the privacy of the customer, dtype: int64
```

```
<AxesSubplot:xlabel='Being able to guarantee the privacy of the customer', ylabel='count'>
```
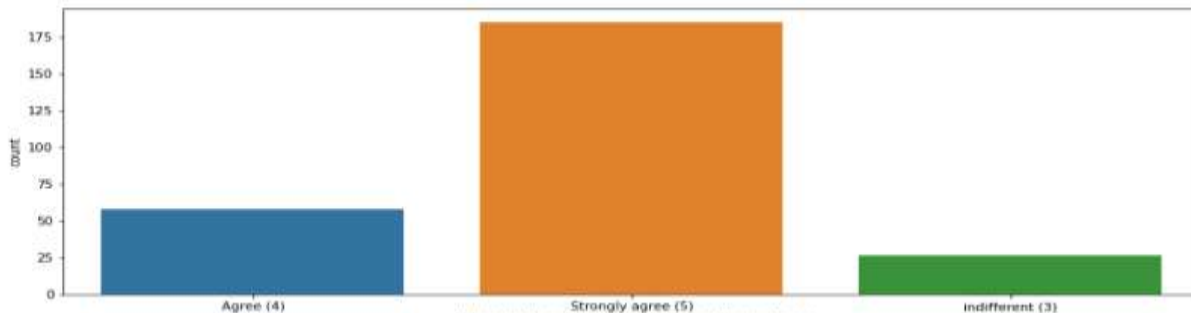
```
plt.figure(figsize=(35,5))
print(df['Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.)'].value_counts
sns.countplot(df['Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.)'])
```

```
Strongly agree (5)       149
Agree (4)                 94
indifferent (3)           15
Strongly disagree (1)     11
Name: Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.), dtype: int64
```

<AxesSubplot:xlabel='Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.)',
ylabel='count'>
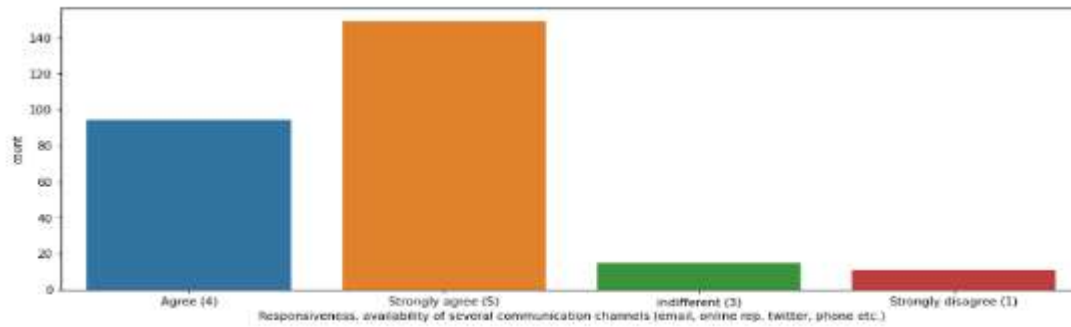

```
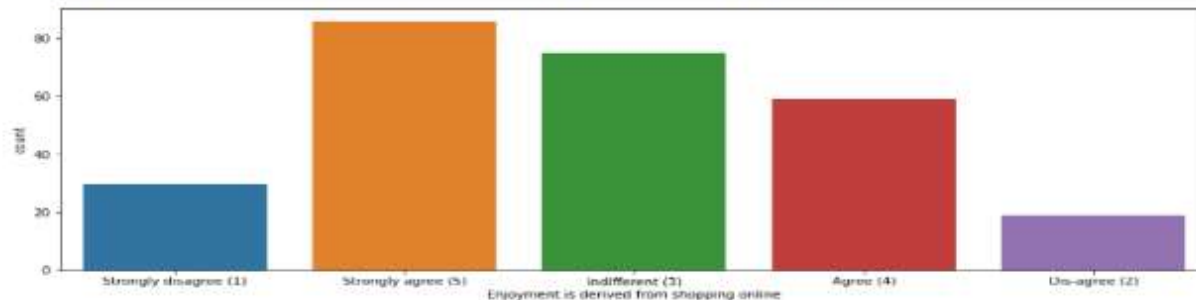plt.figure(figsize=(15,5))
print(df['Online shopping gives monetary benefit and discounts'].value_counts())
sns.countplot(df['Online shopping gives monetary benefit and discounts'])
```

```
Strongly agree (5)       105
Agree (4)                 85
indifferent (3)           50
Strongly disagree (1)     18
Dis-agree (2)             11
Name: Online shopping gives monetary benefit and discounts, dtype: int64
```

<AxesSubplot:xlabel='Online shopping gives monetary benefit and discounts', ylabel='count'>
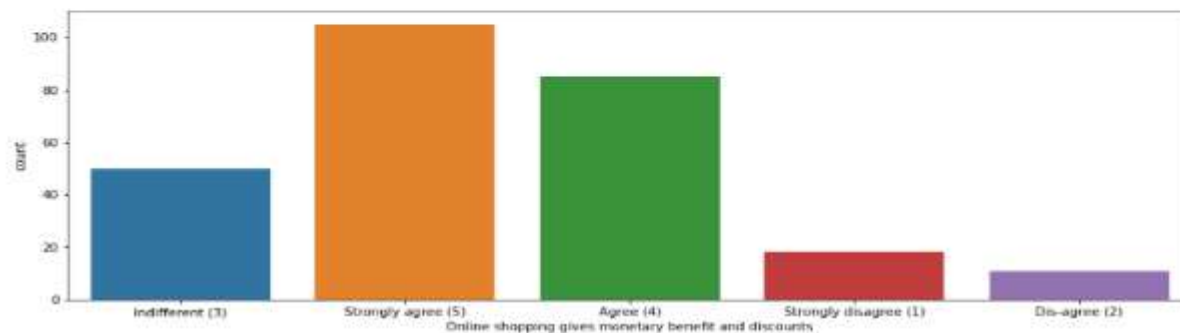

```
plt.figure(figsize=(15,5))
print(df['Enjoyment is derived from shopping online'].value_counts())
sns.countplot(df['Enjoyment is derived from shopping online'])
```

```
Strongly agree (5)       86
indifferent (3)          75
Agree (4)                59
Strongly disagree (1)    30
Dis-agree (2)            19
Name: Enjoyment is derived from shopping online, dtype: int64
```

<AxesSubplot:xlabel='Enjoyment is derived from shopping online', ylabel='count'>

```
plt.figure(figsize=(15,5))
print(df['Shopping online is convenient and flexible'].value_counts())
sns.countplot(df['Shopping online is convenient and flexible'])
```

```
Strongly agree (5)     146
Agree (4)               78
indifferent (3)         33
Dis-agree (2)           12
Name: Shopping online is convenient and flexible, dtype: int64
```

```
<AxesSubplot:xlabel='Shopping online is convenient and flexible', ylabel='count'>
```
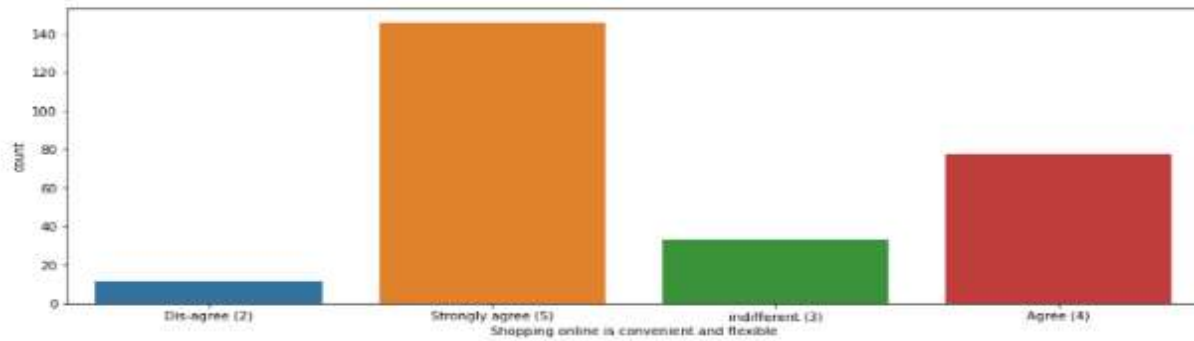


```
plt.figure(figsize=(15,5))
print(df['Return and replacement policy of the e-tailer is important for purchase decision'].value_counts())
sns.countplot(df['Return and replacement policy of the e-tailer is important for purchase decision'])
```

```
Strongly agree (5)     198
Agree (4)               51
Dis-agree (2)           20
Name: Return and replacement policy of the e-tailer is important for purchase decision, dtype: int64
```

```
<AxesSubplot:xlabel='Return and replacement policy of the e-tailer is important for purchase decision', ylabel='count'>
```
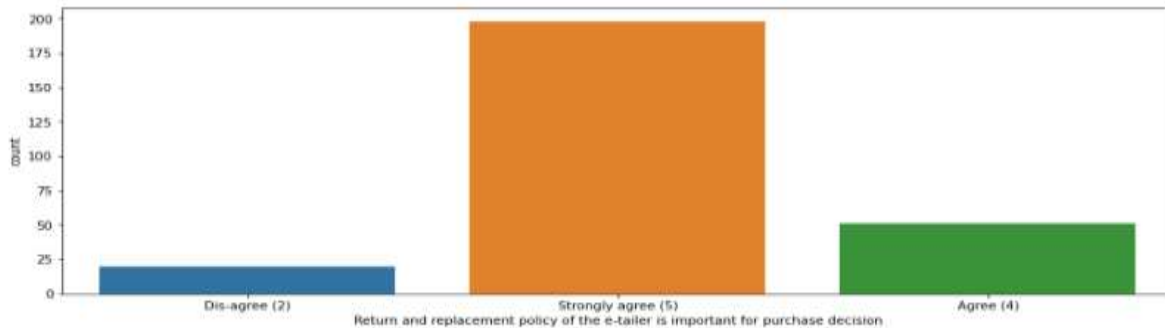


```
plt.figure(figsize=(15,5))
print(df['User derive satisfaction while shopping on a good quality website or application'].value_counts())
sns.countplot(df['User derive satisfaction while shopping on a good quality website or application'])
```

```
Strongly agree (5)     175
Agree (4)               86
Dis-agree (2)            8
Name: User derive satisfaction while shopping on a good quality website or application, dtype: int64
```

```
<AxesSubplot:xlabel='User derive satisfaction while shopping on a good quality website or application', ylabel='count'>
```
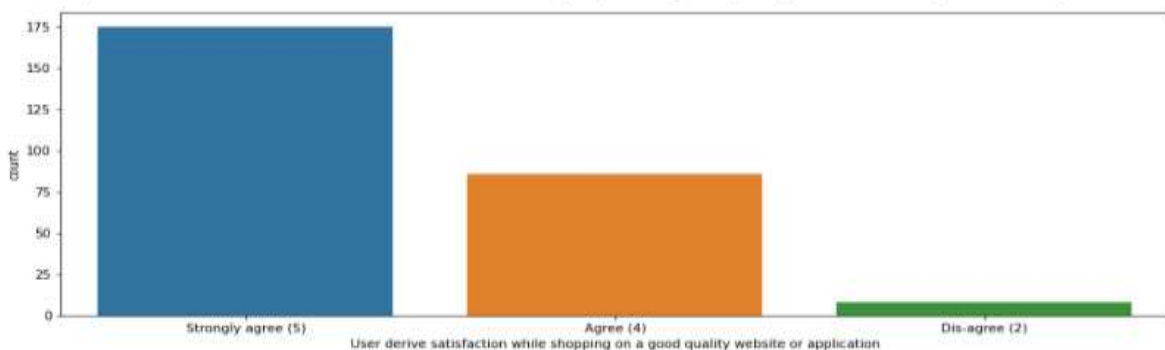
```
plt.figure(figsize=(15,5))
print(df['Monetary savings'].value_counts())
sns.countplot(df['Monetary savings'])
```

```
Strongly agree (5)    148
Agree (4)              75
Disagree (2)           31
indifferent (3)        15
Name: Monetary savings, dtype: int64
```

```
<AxesSubplot:xlabel='Monetary savings', ylabel='count'>
```


- Majority Customers have strongly agreed for Monetary Savings.

```
plt.figure(figsize=(15,5))
plt.xticks(rotation=50)
print(df['Easy to use website or application'].value_counts())
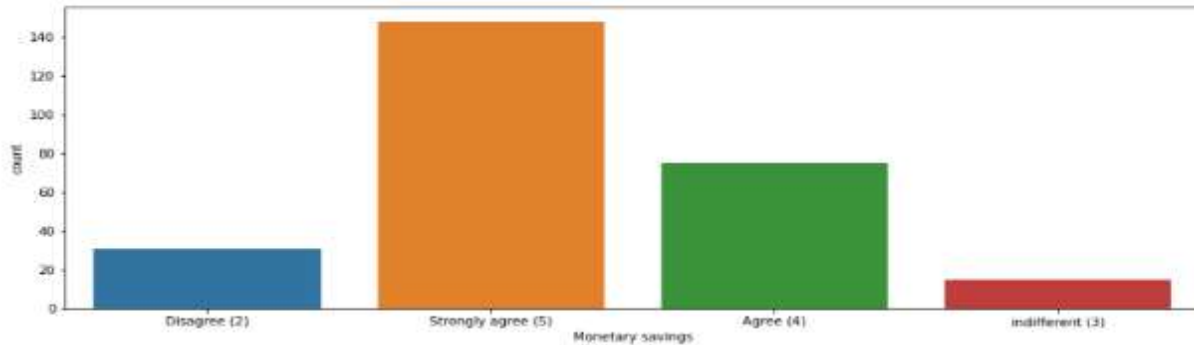sns.countplot(df['Easy to use website or application'])
```

```
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com    64
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com              44
Amazon.in, Flipkart.com                                        44
Amazon.in                                                      29
Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com              22
Amazon.in, Paytm.com, Myntra.com                              20
Amazon.in, Flipkart.com, Myntra.com                           19
Paytm.com                                                     12
Flipkart.com                                                   8
Amazon.in, Paytm.com                                           7
Name: Easy to use website or application, dtype: int64
```

```
<AxesSubplot:xlabel='Easy to use website or application', ylabel='count'>
```
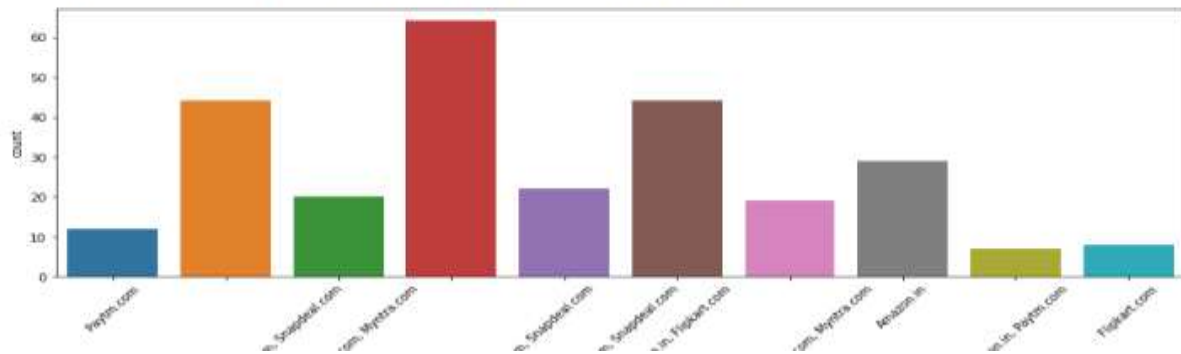

- Majority Customers have strongly agreed that Amazon.in is the easy to use website or application with second place for Flipkart.com and rest following.

```
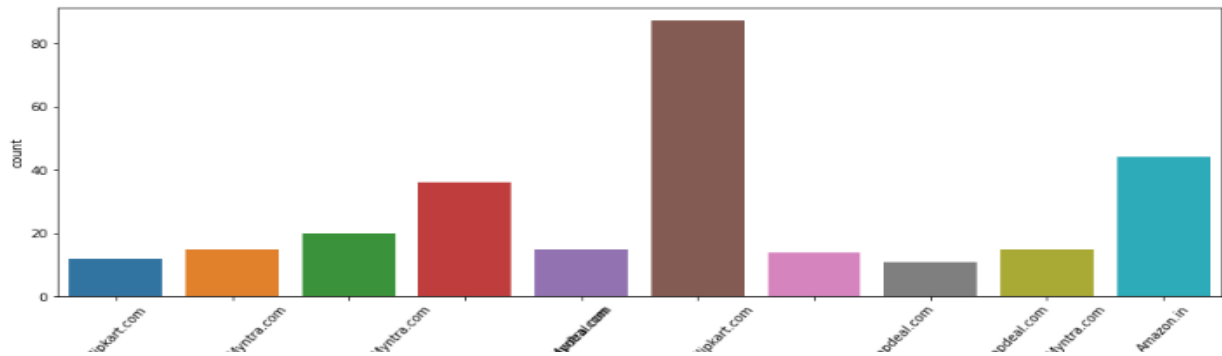plt.figure(figsize=(15,5))
plt.xticks(rotation=50)
print(df['Visual appealing web-page layout'].value_counts())
sns.countplot(df['Visual appealing web-page layout'])
```

```
Amazon.in, Flipkart.com                                              87
Amazon.in                                                            44
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com         36
Amazon.in, Paytm.com, Myntra.com                                     20
Myntra.com                                                           15
Amazon.in, Myntra.com                                                15
Flipkart.com, Myntra.com                                             15
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com                    14
Flipkart.com                                                         12
Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com                     11
Name: Visual appealing web-page layout, dtype: int64
```

```
<AxesSubplot:xlabel='Visual appealing web-page layout', ylabel='count'>
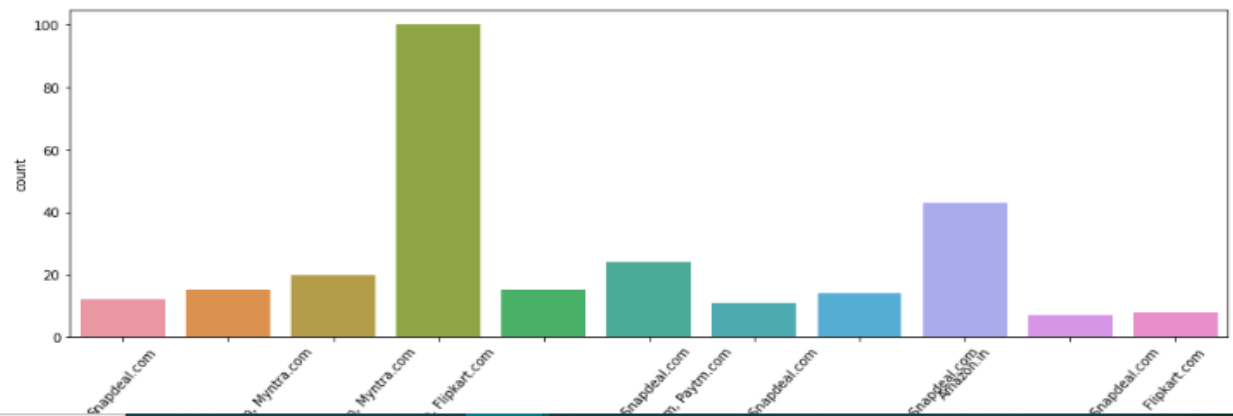```



- Majority Customers have strongly agreed that Amazon.in and Flipkart.com are the visual appealing web-page layout and rest following.

```
plt.figure(figsize=(15,5))
plt.xticks(rotation=50)
print(df['Complete, relevant description information of products'].value_counts())
sns.countplot(df['Complete, relevant description information of products'])
```

```
Amazon.in, Flipkart.com                                              100
Amazon.in                                                            43
Amazon.in, Flipkart.com, Paytm.com                                   24
Amazon.in, Paytm.com, Myntra.com                                     20
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com         15
Amazon.in, Flipkart.com, Myntra.com                                  15
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com                    14
Snapdeal.com                                                         12
Flipkart.com, Snapdeal.com                                           11
Flipkart.com                                                         8
Amazon.in, Flipkart.com, Snapdeal.com                                7
Name: Complete, relevant description information of products, dtype: int64
```

```
<AxesSubplot:xlabel='Complete, relevant description information of products', ylabel='count'>
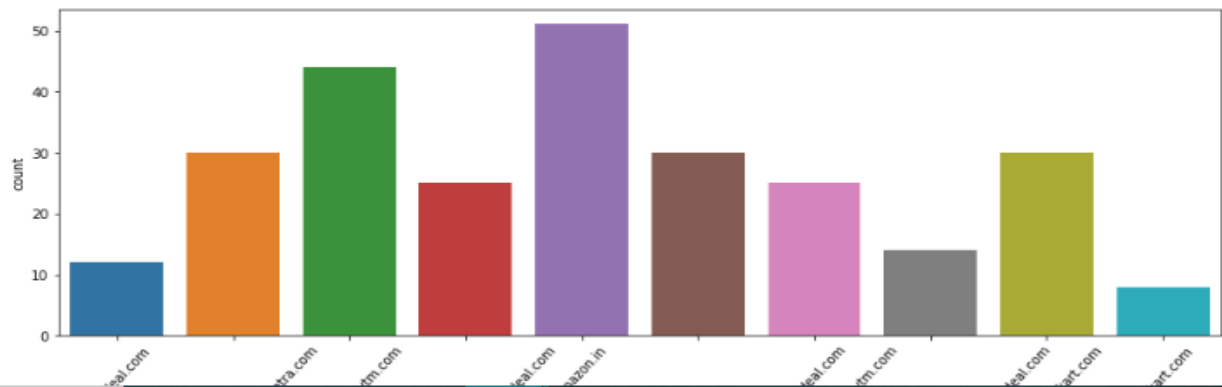```



- Majority Customers have strongly agreed that Amazon.in and Flipkart.com have the Complete, relevant description information of products and rest following.

```
plt.figure(figsize=(15,5))
plt.xticks(rotation=50)
print(df['Fast loading website speed of website and application'].value_counts())
sns.countplot(df['Fast loading website speed of website and application'])
```

```
Amazon.in                                                       51
Amazon.in, Paytm.com                                            44
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com    30
Amazon.in, Flipkart.com, Myntra.com                             30
Amazon.in, Flipkart.com                                         30
Amazon.in, Flipkart.com, Snapdeal.com                           25
Amazon.in, Flipkart.com, Paytm.com                              25
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com               14
Snapdeal.com                                                    12
Flipkart.com                                                     8
Name: Fast loading website speed of website and application, dtype: int64
```

```
<AxesSubplot:xlabel='Fast loading website speed of website and application', ylabel='count'>
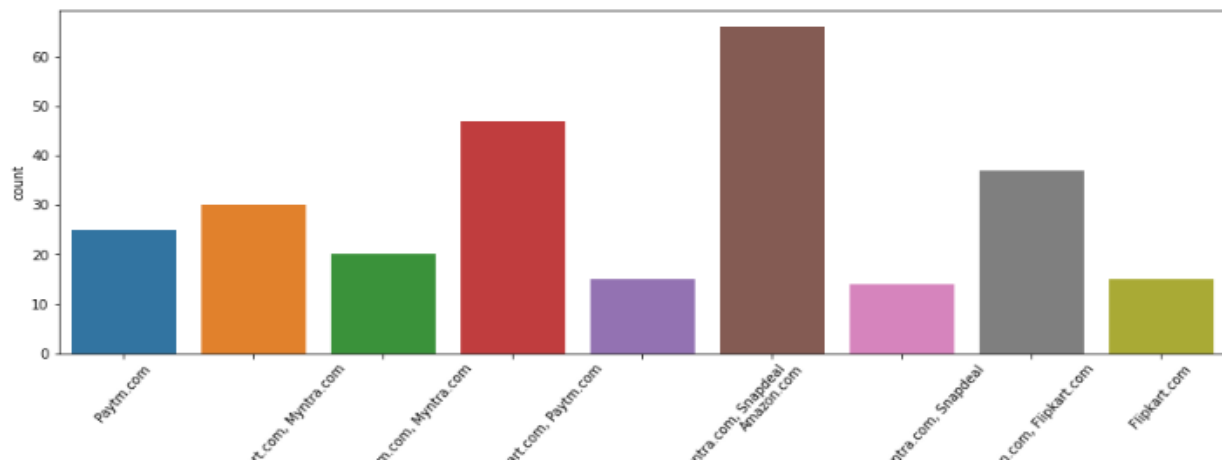```



- Majority Customers have strongly agreed that Amazon.in have the Fast loading website speed of website and application and rest following.

```
plt.figure(figsize=(15,5))
plt.xticks(rotation=50)
print(df['Quickness to complete purchase'].value_counts())
sns.countplot(df['Quickness to complete purchase'])
```

```
Amazon.com                                                      66
Amazon.com, Flipkart.com, Paytm.com                             47
Amazon.com, Flipkart.com                                        37
Amazon.com, Flipkart.com, Myntra.com                            30
Paytm.com                                                       25
Amazon.com, Paytm.com, Myntra.com                               20
Flipkart.com                                                    15
Amazon.com, Flipkart.com, Paytm.com, Myntra.com, Snapdeal       15
Flipkart.com, Myntra.com, Snapdeal                              14
Name: Quickness to complete purchase, dtype: int64
```

```
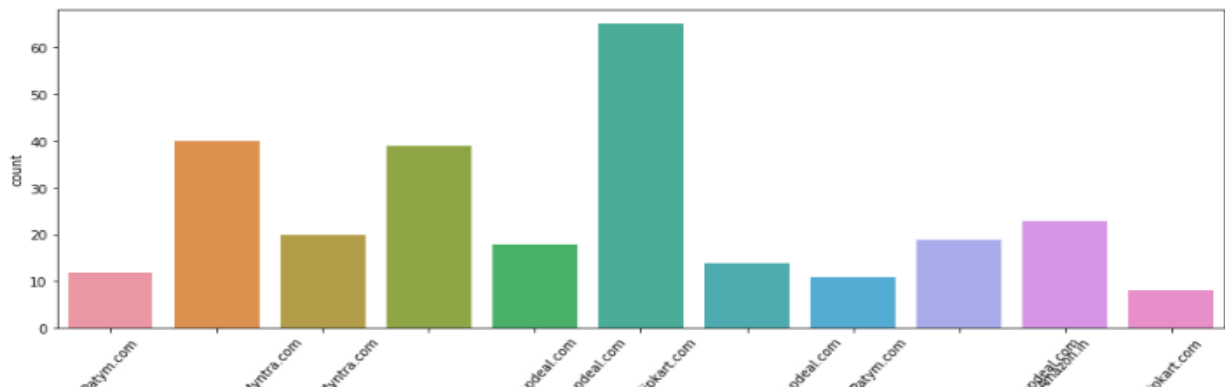<AxesSubplot:xlabel='Quickness to complete purchase', ylabel='count'>
```


- Majority Customers have strongly agreed that Amazon.in and Flipkart.com have the Quickness to complete purchase and rest following.

```
plt.figure(figsize=(15,5))
plt.xticks(rotation=50)
print(df['Availability of several payment options'].value_counts())
sns.countplot(df['Availability of several payment options'])
```

```
Amazon.in, Flipkart.com                                      65
Amazon.in, Flipkart.com, Myntra.com                          40
Amazon.in, Flipkart.com, Patym.com, Myntra.com, Snapdeal.com 39
Amazon.in                                                    23
Patym.com, Myntra.com                                        20
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com            19
Amazon.in, Flipkart.com, Snapdeal.com                        18
Flipkart.com, Myntra.com, Snapdeal.com                       14
Patym.com                                                    12
Amazon.in, Patym.com                                         11
Flipkart.com                                                  8
Name: Availability of several payment options, dtype: int64
```

```
<AxesSubplot:xlabel='Availability of several payment options', ylabel='count'>
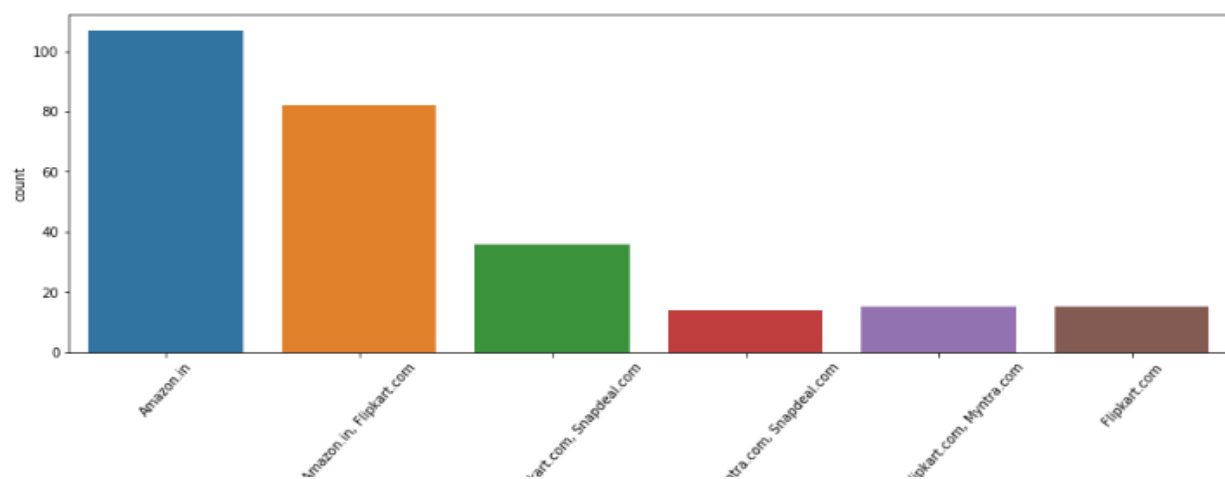```



- Majority Customers have strongly agreed that Amazon.in and Flipkart.com have the Availability of several payment options and rest following.

```
plt.figure(figsize=(15,5))
plt.xticks(rotation=50)
print(df['Speedy order delivery'].value_counts())
sns.countplot(df['Speedy order delivery'])
```

```
Amazon.in                                      107
Amazon.in, Flipkart.com                         82
Amazon.in, Flipkart.com, Snapdeal.com           36
Amazon.in, Flipkart.com, Myntra.com             15
Flipkart.com                                    15
Flipkart.com, Myntra.com, Snapdeal.com          14
Name: Speedy order delivery, dtype: int64
```

```
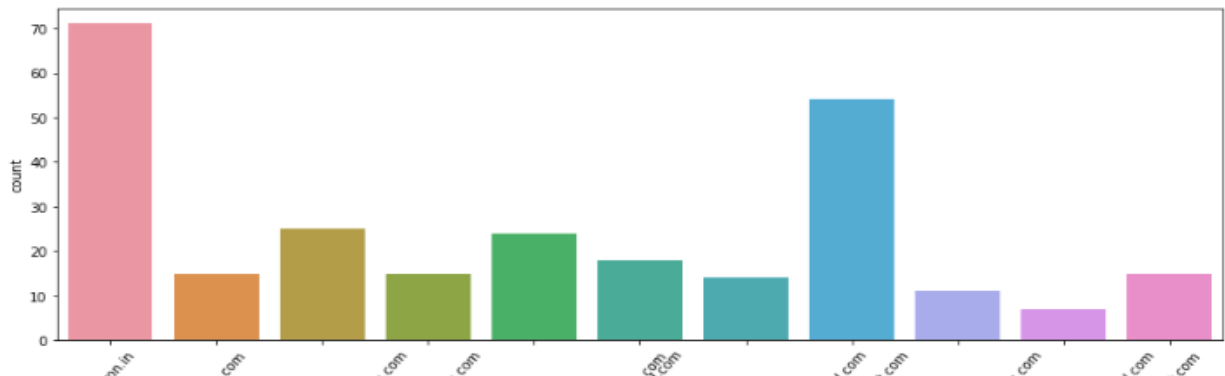<AxesSubplot:xlabel='Speedy order delivery', ylabel='count'>
```


- Majority Customers have strongly agreed that Amazon.in have the Speedy order delivery and rest following.

```
plt.figure(figsize=(15,5))
plt.xticks(rotation=50)
print(df['Privacy of customers' information'].value_counts())
sns.countplot(df['Privacy of customers' information'])
```

```
Amazon.in                                                      71
Amazon.in, Flipkart.com                                        54
Amazon.in, Flipkart.com, Myntra.com                            25
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com   24
Paytm.com                                                      18
Amazon.in, Paytm.com                                           15
Myntra.com                                                     15
Flipkart.com                                                   15
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com              14
Amazon.in, Flipkart.com, Paytm.com                             11
Amazon.in, Flipkart.com, Snapdeal.com                           7
Name: Privacy of customers' information, dtype: int64
```

`<AxesSubplot:xlabel='Privacy of customers' information', ylabel='count'>`

- Majority Customers have strongly agreed that Amazon.in have Privacy of customers and rest following.

```
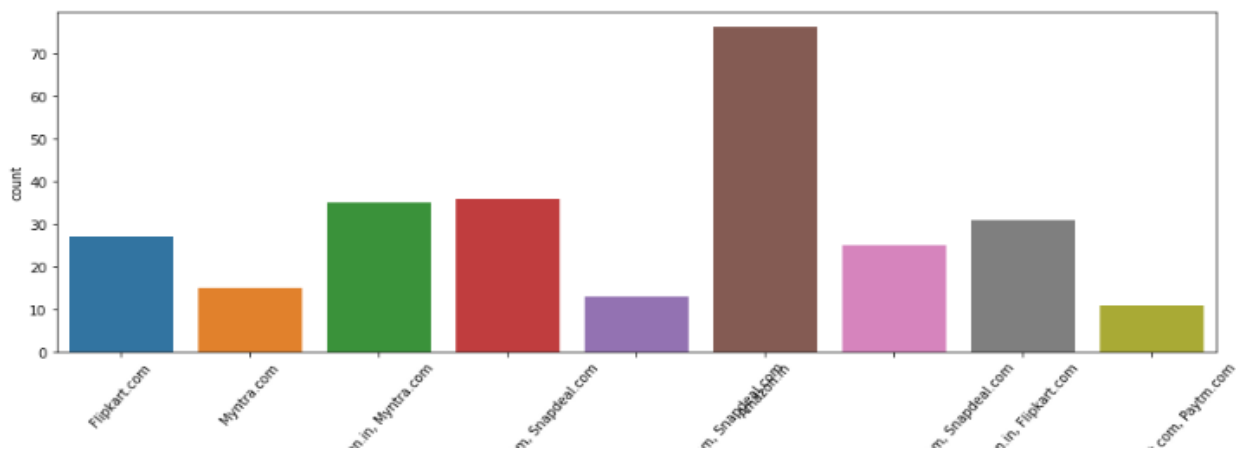plt.figure(figsize=(15,5))
plt.xticks(rotation=50)
print(df['Perceived Trustworthiness'].value_counts())
sns.countplot(df['Perceived Trustworthiness'])
```

```
Amazon.in                                                      76
Amazon.in, Flipkart.com, Snapdeal.com                          36
Amazon.in, Myntra.com                                          35
Amazon.in, Flipkart.com                                        31
Flipkart.com                                                   27
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com              25
Myntra.com                                                     15
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com   13
Amazon.in, Flipkart.com, Paytm.com                             11
Name: Perceived Trustworthiness, dtype: int64
```

`<AxesSubplot:xlabel='Perceived Trustworthiness', ylabel='count'>`

- Majority Customers have strongly agreed that Amazon.in have the Perceived Trust worthiness and rest following.

```
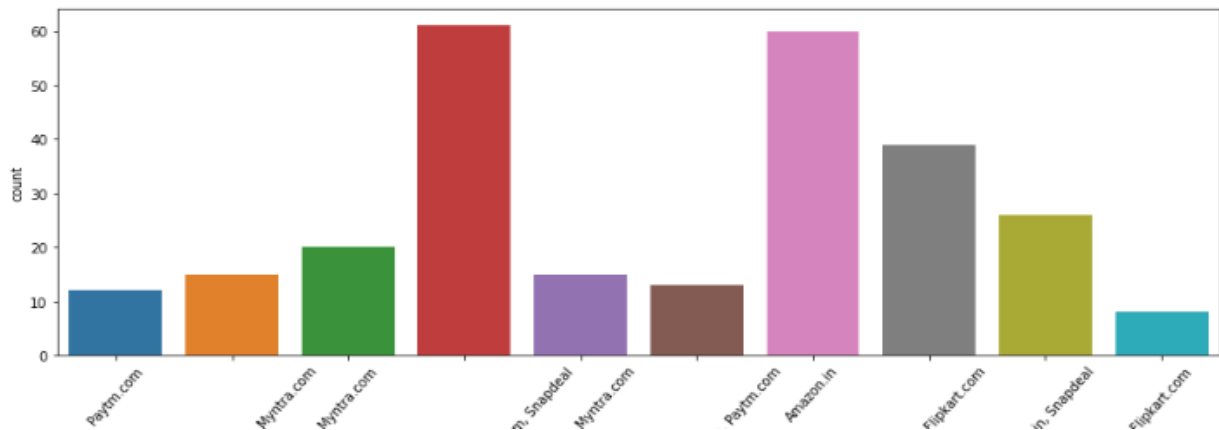plt.figure(figsize=(15,5))
plt.xticks(rotation=50)
print(df['Presence of online assistance through multi-channel'].value_counts())
sns.countplot(df['Presence of online assistance through multi-channel'])
```

```
Amazon.in, Flipkart.com, Myntra.com, Snapdeal      61
Amazon.in                                          60
Amazon.in, Flipkart.com                            39
Amazon.in, Snapdeal                                26
Myntra.com                                         20
Amazon.in, Myntra.com                              15
Amazon.in, Flipkart.com, Myntra.com                15
Amazon.in, Flipkart.com, Paytm.com                 13
Paytm.com                                          12
Flipkart.com                                        8
Name: Presence of online assistance through multi-channel, dtype: int64

<AxesSubplot:xlabel='Presence of online assistance through multi-channel', ylabel='count'>
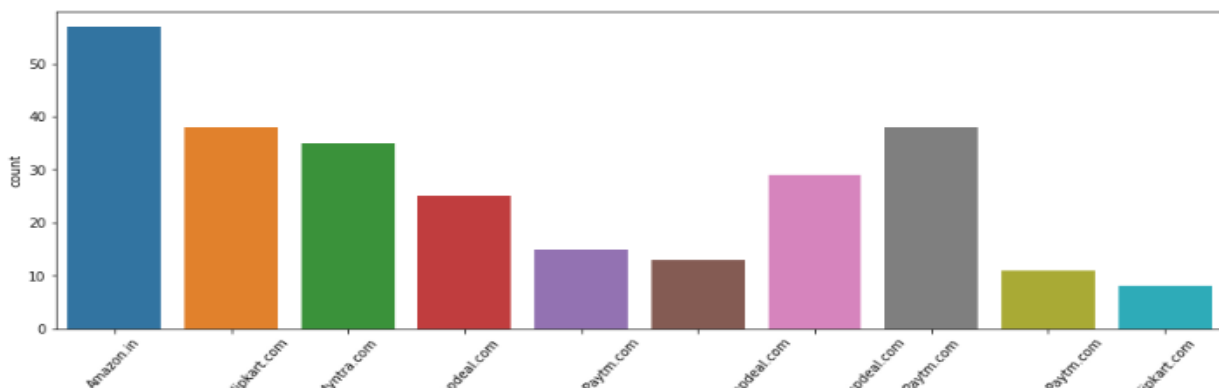```



- Majority Customers have strongly agreed that Amazon.in and Flipkart.com have Presence of online assistance through multi-channel and rest following.

```
plt.figure(figsize=(15,5))
plt.xticks(rotation=50)
print(df['Longer time to get logged in (promotion, sales period)'].value_counts())
sns.countplot(df['Longer time to get logged in (promotion, sales period)'])
```

```
Amazon.in                                           57
Paytm.com                                           38
Amazon.in, Flipkart.com                             38
Myntra.com                                          35
Amazon.in, Flipkart.com, Snapdeal.com               29
Snapdeal.com                                        25
Flipkart.com, Paytm.com                             15
Flipkart.com, Paytm.com, Snapdeal.com               13
Amazon.in, Paytm.com                                11
Flipkart.com                                         8
Name: Longer time to get logged in (promotion, sales period), dtype: int64

<AxesSubplot:xlabel='Longer time to get logged in (promotion, sales period)', ylabel='count'>
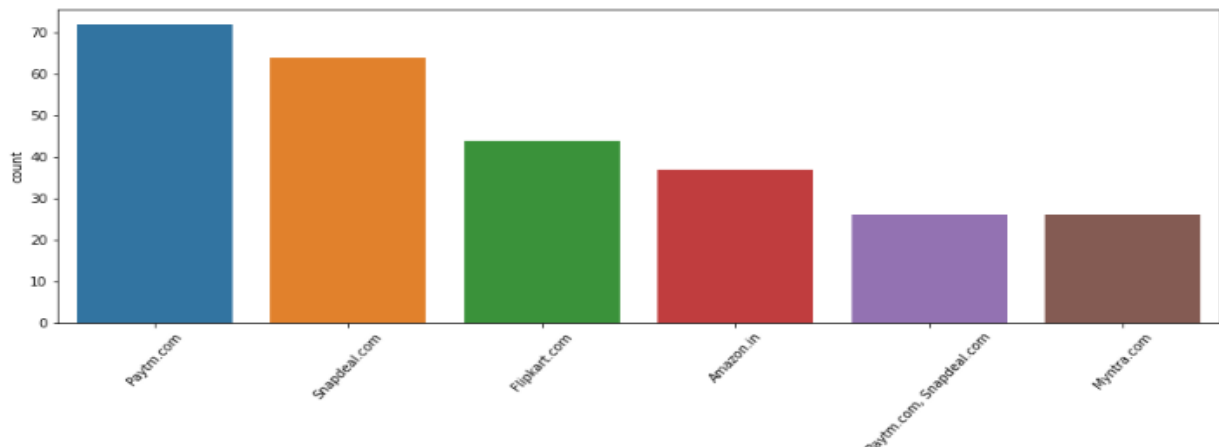```



- Customers have strongly agreed that Amazon.in have the longer time to get logged in and rest following.

```
plt.figure(figsize=(15,5))
plt.xticks(rotation=50)
print(df['Longer delivery period'].value_counts())
sns.countplot(df['Longer delivery period'])
```

```
Paytm.com                  72
Snapdeal.com               64
Flipkart.com               44
Amazon.in                  37
Myntra.com                 26
Paytm.com, Snapdeal.com    26
Name: Longer delivery period, dtype: int64
```

```
<AxesSubplot:xlabel='Longer delivery period', ylabel='count'>
```


- Majority Customers have strongly agreed that Paytm.in and snapdeal.com have the Longer delivery period and rest following.

```
plt.figure(figsize=(15,5))
plt.xticks(rotation=50)
print(df['Change in website/Application design'].value_counts())
sns.countplot(df['Change in website/Application design'])
```

```
Amazon.in                    96
Paytm.com                    63
Amazon.in, Flipkart.com      45
Myntra.com                   30
Flipkart.com                 20
Snapdeal.com                  8
Flipkart.com, Myntra.com      7
Name: Change in website/Application design, dtype: int64
```

```
<AxesSubplot:xlabel='Change in website/Application design', ylabel='count'>
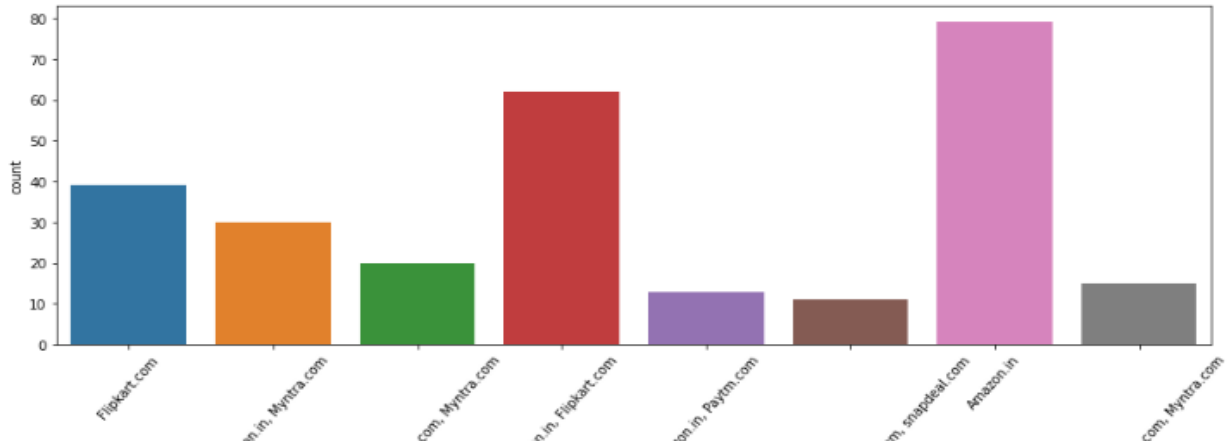```
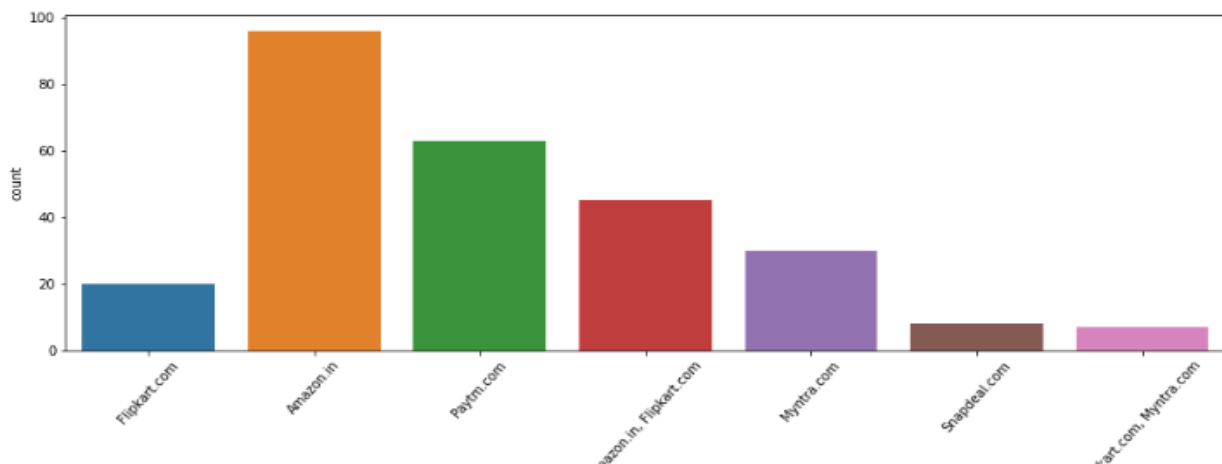


- Majority Customers have strongly agreed that Amazon.in have the Change in website/application design and rest following.

```
plt.figure(figsize=(15,5))
plt.xticks(rotation=50)
print(df['Which of the Indian online retailer would you recommend to a friend?'].value_counts())
sns.countplot(df['Which of the Indian online retailer would you recommend to a friend?'])
```

```
Amazon.in                                             79
Amazon.in, Flipkart.com                               62
Flipkart.com                                          39
Amazon.in, Myntra.com                                 30
Amazon.in, Paytm.com, Myntra.com                      20
Amazon.in, Flipkart.com, Myntra.com                   15
Amazon.in, Paytm.com                                  13
Flipkart.com, Paytm.com, Myntra.com, snapdeal.com     11
Name: Which of the Indian online retailer would you recommend to a friend?, dtype: int64
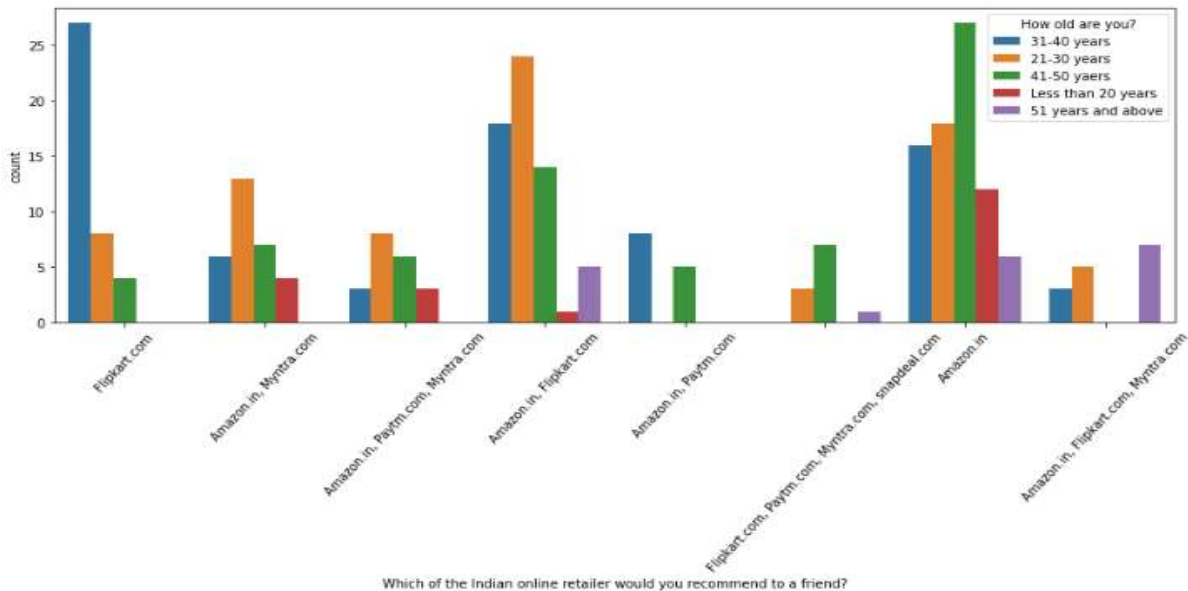```

```
<AxesSubplot:xlabel='Which of the Indian online retailer would you recommend to a friend?', ylabel='count'>
```


- Majority Customers have strongly agreed that Amazon.in is the Indian online retailer which they would recommend to their friends and rest following.

```
plt.figure(figsize=(15,5),facecolor='white')
plt.xticks(rotation=50)
sns.countplot(df['Which of the Indian online retailer would you recommend to a friend?'], hue=df['How old are you?'])
```

```
<AxesSubplot:xlabel='Which of the Indian online retailer would you recommend to a friend?', ylabel='count'>
```


- Majority Customers of 41-50 age group have strongly agreed that Amazon.in is the Indian online retailer which they would recommend to their friends and rest following.
- Majority Customers of 31-40 age group have strongly agreed that Flipkart.com is the Indian online retailer which they would recommend to their friends and rest following.

```
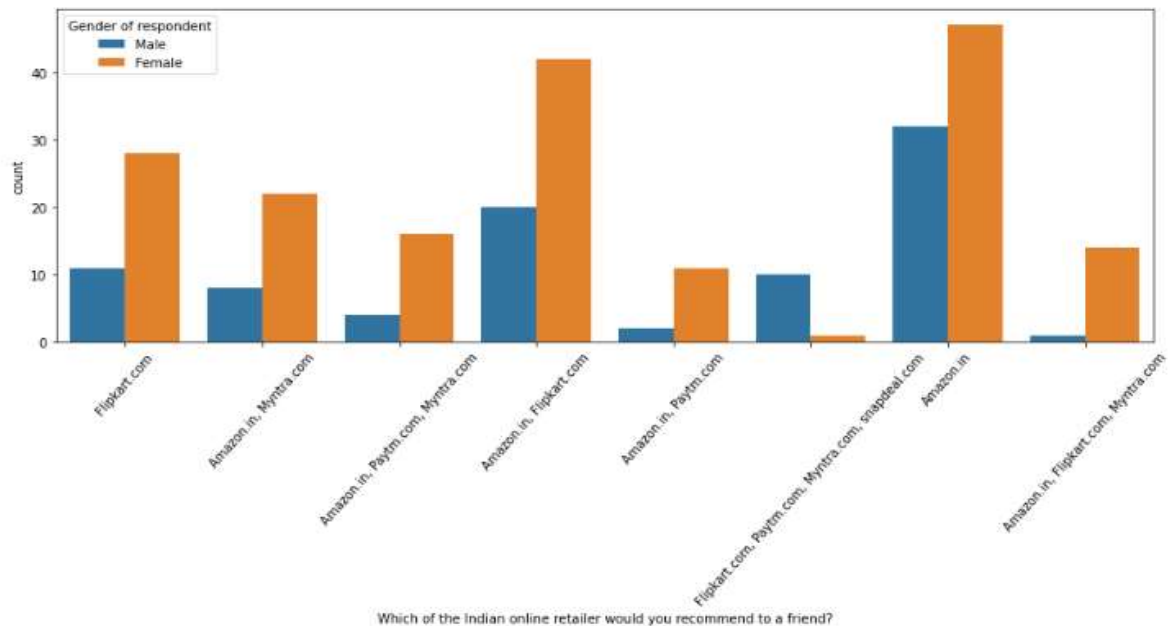plt.figure(figsize=(15,5),facecolor='white')
plt.xticks(rotation=50)
sns.countplot(df['Which of the Indian online retailer would you recommend to a friend?'], hue=df['Gender of respondent'])
```

<AxesSubplot:xlabel='Which of the Indian online retailer would you recommend to a friend?', ylabel='count'>

- Majority Female Customers have strongly agreed that Amazon.in is the Indian online retailer which they would recommend to their friends and rest following.

```
plt.figure(figsize=(15,5))
plt.xticks(rotation=50)
print(df['Wild variety of product on offer'].value_counts())
sns.countplot(df['Wild variety of product on offer'])
```

```
Amazon.in, Flipkart.com                              130
Amazon.in                                             43
Amazon.in, Myntra.com                                 20
Myntra.com                                            15
Flipkart.com, Myntra.com                              15
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com     14
Amazon.in, Flipkart.com, Paytm.com                    13
Flipkart.com                                          12
Paytm.com                                              7
Name: Wild variety of product on offer, dtype: int64
```

<AxesSubplot:xlabel='Wild variety of product on offer', ylabel='count'>
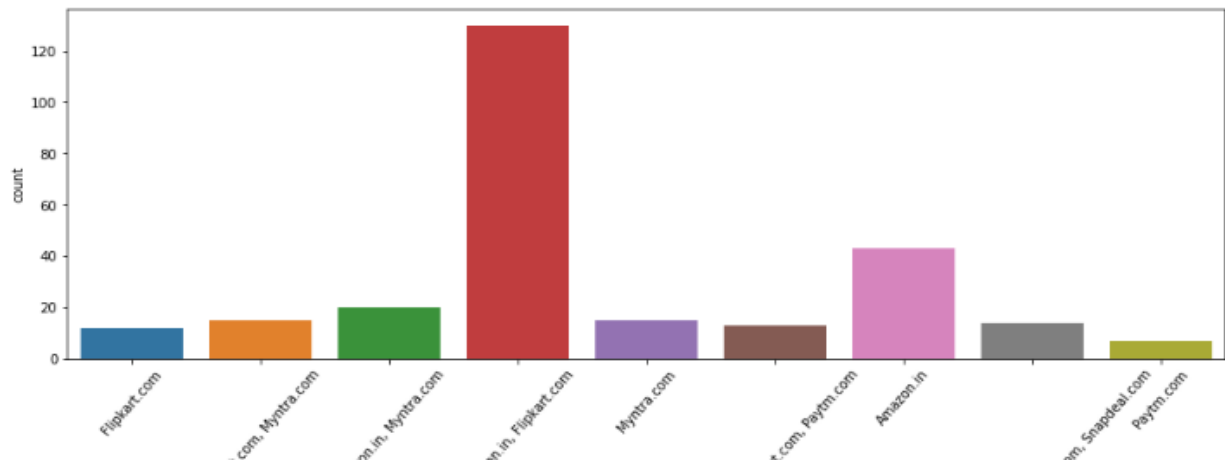
- Majority Customers have strongly agreed that Amazon.in and flipkart.com have Wild variety of product on offer and rest following.

Outcomes from the data analysis:

- 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.
- 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 main 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.
- 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.
- Getting feedbacks from the customers is also on of the important thing to improve the sales of the company. The retailer 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 there is 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.

Conclusions:

- The endeavor of this study is to identify the motivating factors towards online shopping and in which we come to know the customers likely to shop more.
- 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.
- 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.

As in the final plot in which user were asked which online retailer they would recommend to a friend in this Amazon.in topped the list because it is providing all the features that users want. Website is efficient and it is fast loading, it gives complete, relevant description and information of products. It is reliable and quick to complete the purchase. Amazon give speedy delivery to its customers and there are several payment options available on the website. It provides online assistance through multi channels. Providing good deals on products. Its website have visual appealing webpage layout and they offer wide variety of products and its application is easy to use. Lastly the main thing why user recommend it is because of its Trustworthiness and its robust Security in protecting customer financial information and their Privacy information.
These all features make it top the list of recommended online retailer .Providing these features it is retaining its customers.

KEY FINDINGS AND CONCLUSIONS OF THE STUDY
Company with higher customer satisfaction and retention:
1. Amazon.com

2. Flipkart.com

Company with high risk of customer churn:
1. Snapdeal.com