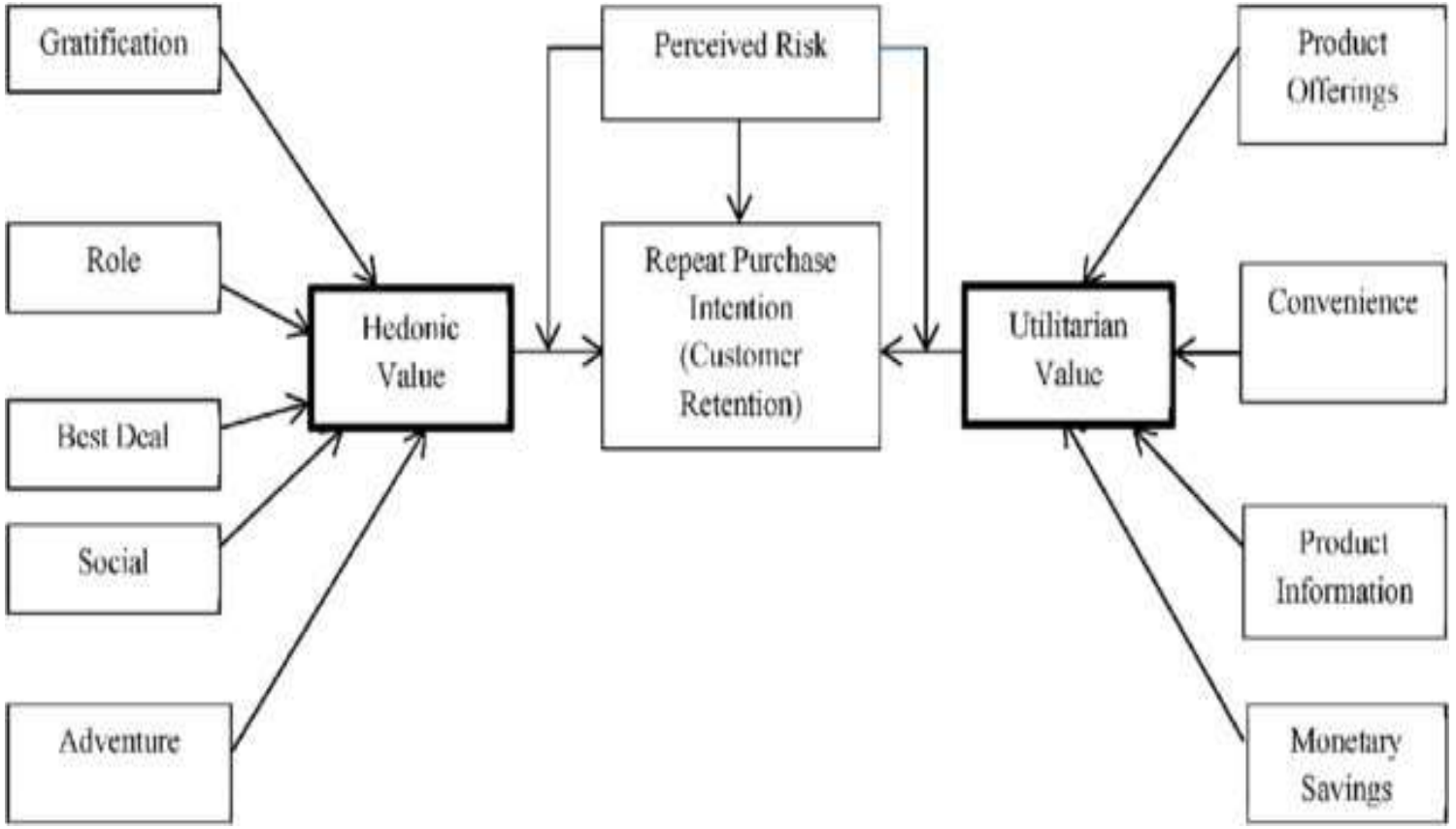
CUSTOMER RETENTION CASE STUDY

Introduction

- **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.
-
- The data is collected from
- **E-retail factors for customer activation and retention: A case study from Indian e-commerce customers**


Gratification
Role
Best Deal
Social
Adventure
Hedonic Value
Perceived Risk
Repeat Purchase Intention (Customer Retention)
Utilitarian Value
Product Offerings
Convenience
Product Information
Monetary Savings

- Table of Contents:
- Exploratory Data Analysis
- Checking type of data
  - Checking columns of the data
  - Checking shape of the data
  - Check there are null values present in the data

  ✓ if present handle it,

  ✓if it is a numerical column then fill it with mean() of that column

  ✓If it is a categeorical column fill it with mode() of that column

- Check the correlation of the data
- Check the description of the data,which tells statistical information about the data like,count, min,max,standard deviation, and the quartiles data
- Check the outliers present in the data, which tells some of the datapoints are outside the whiskers so, we have to handle it
- Check the skewness of the data

➢ **Data visualization**:

Here we have visualized the data by univariate,By variate/multivariate analysis

- Countplot
- Scatter plot
- pair plot

Data Sources and their formats

We got the data in the form of excel sheet

And read that excel file into jupyter notebook i.e loading our data

Import the necessary libraries

- import pandas as pd
- import numpy as np
- import matplotlib.pyplot as plt
- import seaborn as sns
- import warnings
- warnings.filterwarnings('ignore')

And load the excel file

- df=pd.read_excel(r'C:\Users\polasasuresh\Downloads\Customer_retention_dataset-\customer.xls')

df

- #check the columns present in the dataset
- df.columns

Gives the total number of columns present in the data. Here total 71 columns present

```
In [6]: 1 # Let's check datset information
        2 df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 269 entries, 0 to 268
Data columns (total 71 columns):
 #   Column
Non-Null Count  Dtype
---  ------
--------------  -----
 0   1Gender of respondent
269 non-null    object
 1   2 How old are you?
269 non-null    object
 2   3 Which city do you shop online from?
269 non-null    object
 3   4 What is the Pin Code of where you shop online from?
269 non-null    int64
 4   5 Since How Long You are Shopping Online ?
269 non-null    object
 5   6 How many times you have made an online purchase in the past 1 year?
269 non-null    object
 6   7 How do you access the internet while shopping on-line?
269 non-null    object
 7   8 Which device do you use to access the online shopping?
269 non-null    object
 8   9 What is the screen size of your mobile device?
 11  12 Which channel did you follow to arrive at your favorite online store for the first time?
269 non-null    object
 12  13 After first visit, how do you reach the online retail store?
269 non-null    object
 13  14 How much time do you explore the e- retail store before making a purchase decision?
269 non-null    object
 14  15 What is your preferred payment Option?
269 non-null    object
 15  16 How frequently do you abandon (selecting an items and leaving without making payment) your sh
opping cart?
269 non-null    object
 16  17 Why did you abandon the "Bag", "Shopping Cart"?
269 non-null    object
 17  18 The content on the website must be easy to read and understand
269 non-null    object
 18  19 Information on similar product to the one highlighted  is important for product comparison
269 non-null    object
 19  20 Complete information on listed seller and product being offered is important for purchase dec
ision.
269 non-null    object
 20  21 All relevant information on listed products must be stated clearly
269 non-null    object
 21  22 Ease of navigation in website
269 non-null    object
 22  23 Loading and processing speed
269 non-null    object
 23  24 User friendly Interface of the website
269 non-null    object
```

269 non-null object
26 27 Empathy (readiness to assist with queries) towards the customer
269 non-null object
27 28 Being able to guarantee the privacy of the customer
269 non-null object
28 29 Responsiveness, availability of several communication channels
hone etc.)
269 non-null object
29 30 Online shopping gives monetary benefit and discounts
269 non-null object
30 31 Enjoyment is derived from shopping online
269 non-null object
31 32 Shopping online is convenient and flexible
269 non-null object
32 33 Return and replacement policy of the e-tailer is important for
269 non-null object
33 34 Gaining access to loyalty programs is a benefit of shopping onl
269 non-null object
34 35 Displaying quality Information on the website improves satisfa
269 non-null object
35 36 User derive satisfaction while shopping on a good quality websi
269 non-null object
36 37 Net Benefit derived from shopping online can lead to users sati
269 non-null object
37 38 User satisfaction cannot exist without trust
269 non-null object
38 39 Offering a wide variety of listed product in several category
269 non-null object
41 42 The Convenience of patronizing the online retailer
269 non-null object
42 43 Shopping on the website gives you the sense of adventure
269 non-null object
43 44 Shopping on your preferred e-tailer enhances your social status
269 non-null object
44 45 You feel gratification shopping on your favorite e-tailer
269 non-null object
45 46 Shopping on the website helps you fulfill certain roles
269 non-null object
46 47 Getting value for money spent
269 non-null object
47 From the following, tick any (or all) of the online retailers you have shopped from;
269 non-null object
48 Easy to use website or application
269 non-null object
49 Visual appealing web-page layout
269 non-null object
50 Wild variety of product on offer
269 non-null object
51 Complete, relevant description information of products
269 non-null object
52 Fast loading website speed of website and application
269 non-null object
53 Reliability of the website or application
269 non-null object
54 Quickness to complete purchase
269 non-null object

- **Data Preprocessing**

To clean the data ,here we have to check the null values

- df.isnull().sum()

1Gender of respondent 0

2 How old are you? 0

3 Which city do you shop online from? 0

4 What is the Pin Code of where you shop online from? 0

5 Since How Long You are Shopping Online ? 0

..

Longer delivery period 0 Change in website/Application design 0

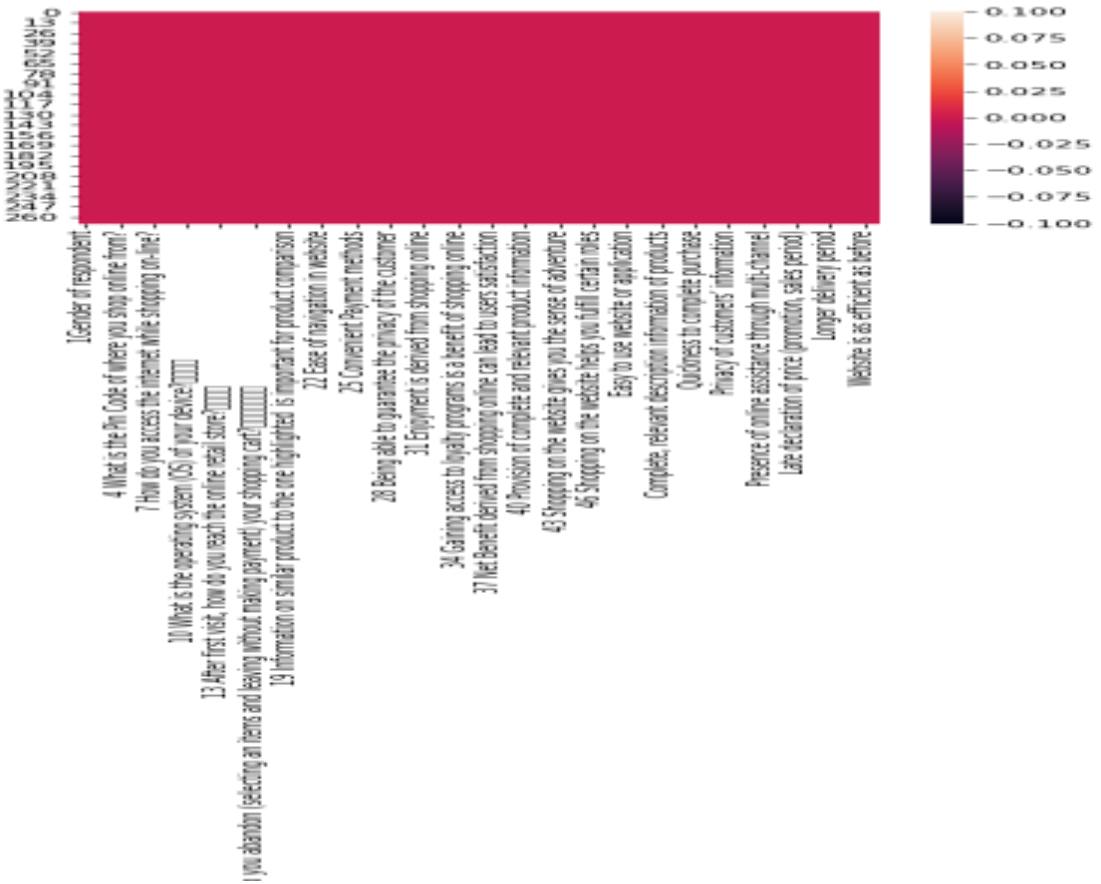
Frequent disruption whenmoving from one page to another 0

Website is as efficient as beforeWhich of the Indian online retailer would you recommend to friend? 0

Length: 71, dtype: int64

- There is no null values present in the data

- We can check with heat map
- sns.heatmap(df.isnull())
- plt.figure(figsize=(14,7))
- plt.show()



- Replacing with proper formats

In [11]:
```
df['7 How do you access the internet while shopping on-line?'].value_counts()
```

Out[11]:
```
Mobile internet    142
Wi-Fi               76
Mobile Internet     47
Dial-up              4
Name: 7 How do you access the internet while shopping on-line?, dtype: int64
```

In [12]:
```
# Replacing characters wiht same information
df['7 How do you access the internet while shopping on-line?']=df['7 How do you access the internet while shopping on-line?'
```

In [17]:
```
# Replace with correct values
for i in websites:

    websites[i]=websites[i].str.replace('Amazon.com','Amazon.in')

    websites[i]=websites[i].str.replace('Snapdeal','Snapdeal.com')
    websites[i]=websites[i].str.replace('snapdeal.com','Snapdeal.com')

    websites[i]=websites[i].str.replace('Snapdeal.com.com','Snapdeal.com')
    websites[i]=websites[i].str.replace('Patym.com','Paytm.com')
```

- **Hardware and Software Tools used**
- There is no hardware used
- Software:Jupyter Notebook (Anaconda) , Microsoft windows,Microsoft Power point
- Languages used: pandas
- Libraries:
    - Pandas
    - Numpy
    - Matplotlib
    - Seaborn

Exploratory Data Analysis

- Checking gender wise data

In [10]:
```
# Let's check data distribution based on Gender

print(cl("Male dataset - {}".format(df[df['1Gender of respondent']=='Male'].shape),color='blue',at
print(cl("Female dataset - {}".format(df[df['1Gender of respondent']=='Female'].shape),color='blue
```

```
Male dataset - (88, 71)
Female dataset - (181, 71)
```

So based on the gender, we can see that our data is not balanced. We have more of Female data.

In [15]:
```
# Let's copy the dataset for websites analysis
websites = df.iloc[:,47:]
```

In [16]:
```
# We have extraced the features which contains websites information
websites
```

Out[16]:

	From the following, tick any (or all) of the online retailers you have shopped from;	Easy to use website or application	Visual appealing web-page layout	Wild variety of product on offer	Complete, relevant description information of products	Fast loading website speed of website and application	Reliability of the website or application	Quickness to complete purchase	Availa of se pay op
0	Amazon.in, Paytm.com	Paytm.com	Flipkart.com	Flipkart.com	Snapdeal.com	Snapdeal.com	Paytm.com	Paytm.com	Patyr
1	Amazon.in, Flipkart.com, Myntra.com,	Amazon.in, Flipkart.com, Myntra.com,	Amazon.in, Myntra.com	Flipkart.com, Myntra.com	Amazon.in, Flipkart.com,	Amazon.in, Flipkart.com,	Myntra.com	Amazon.com, Flipkart.com,	Amaz Flipkar

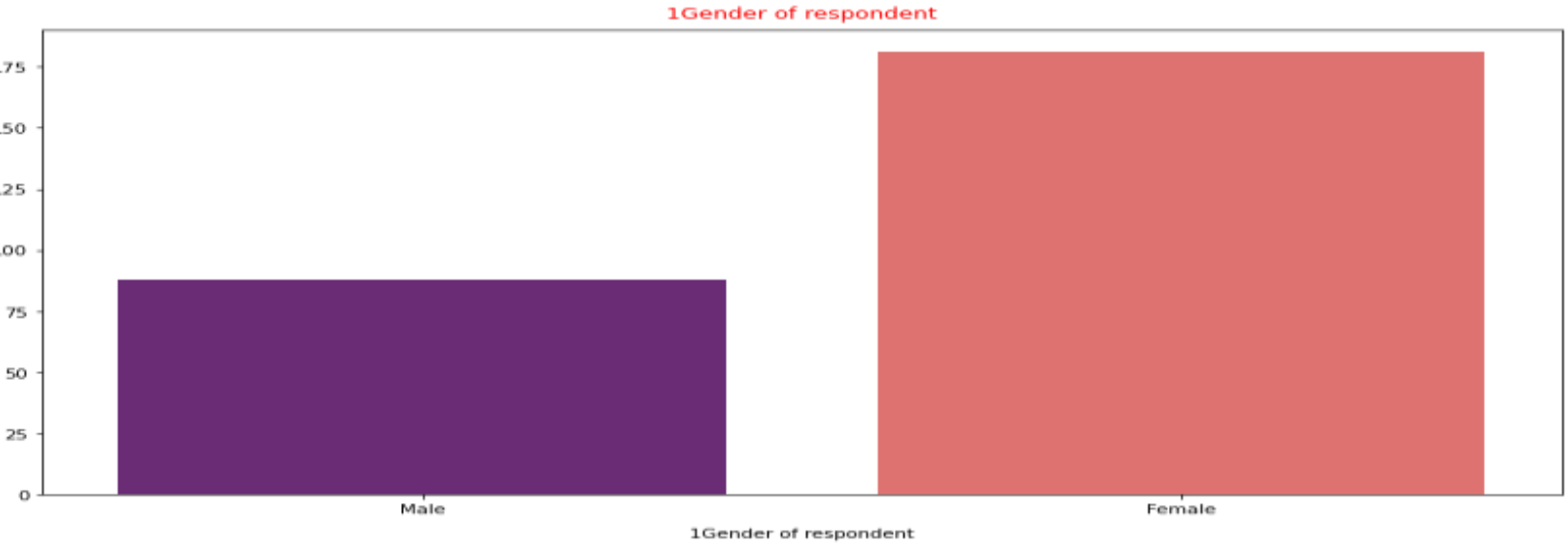
Visualization

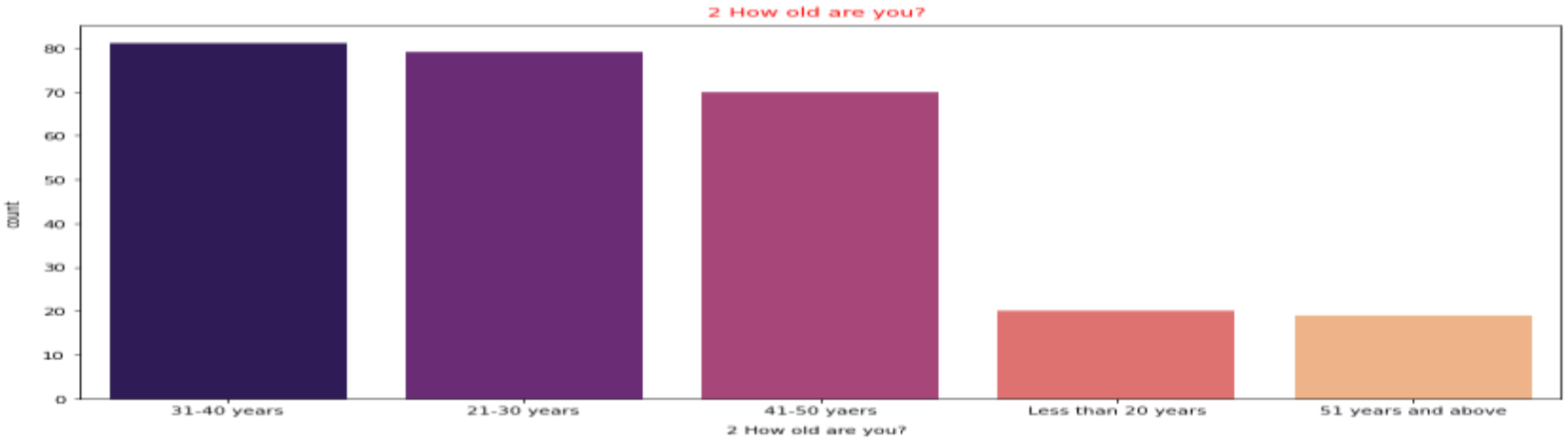
- Created a function for plotting the data

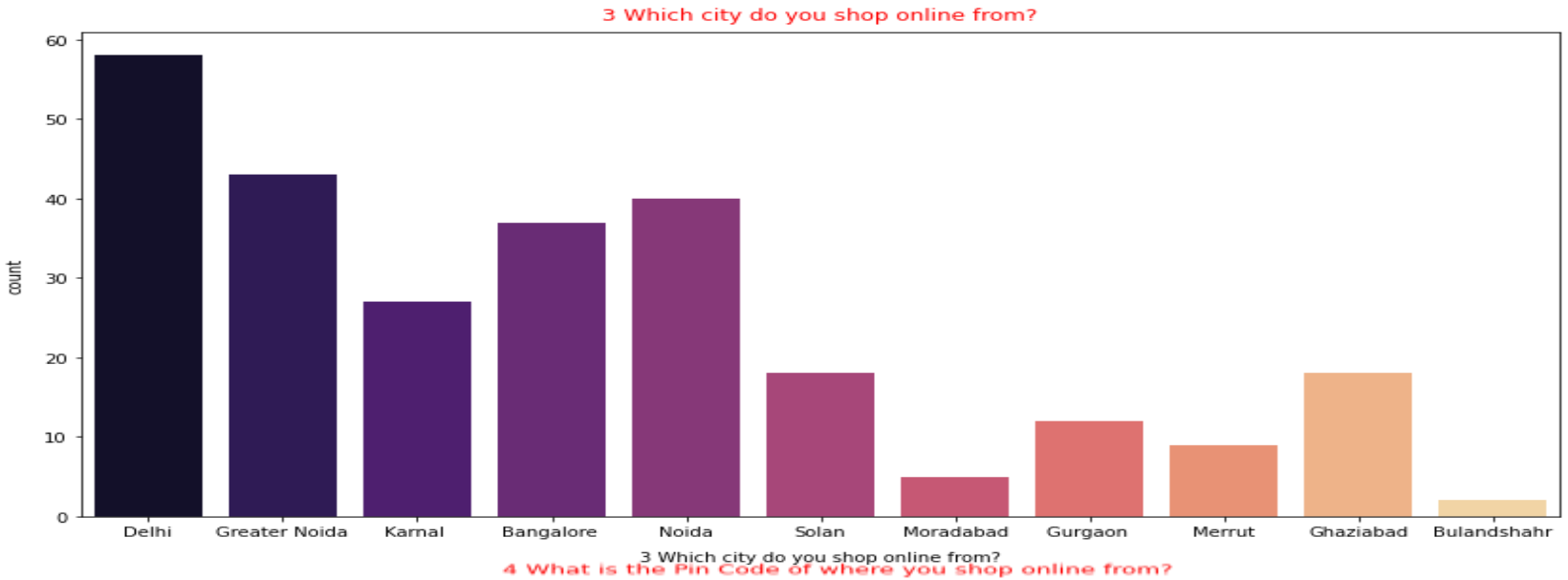
In [13]:

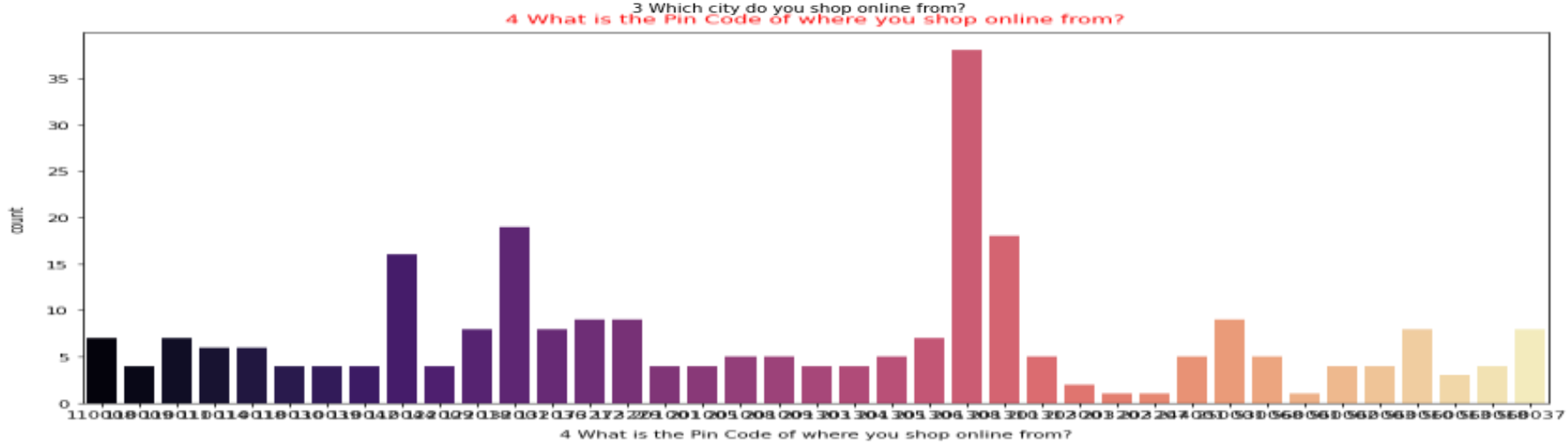
```
# Let's plot countplot for some of the features

for i in df.columns[:17]:
    
    plt.figure(figsize=(10,4))
    sns.countplot(df[i],palette='magma',saturation=0.75)
    plt.title(i)
    plt.setp(plt.title(i,pad=10),  color='red', style='italic')
    plt.setp(plt.xlabel(i,labelpad=10), size='large', color='k', style='italic')
    plt.setp(plt.ylabel("count",labelpad=10), size='large', color='k', style='italic')
    plt.xticks(rotation=90)
    plt.show()
    print(45*"--")
```


1Gender of respondent
175
150
125
100
75
50
25
0
Male
Female
1Gender of respondent

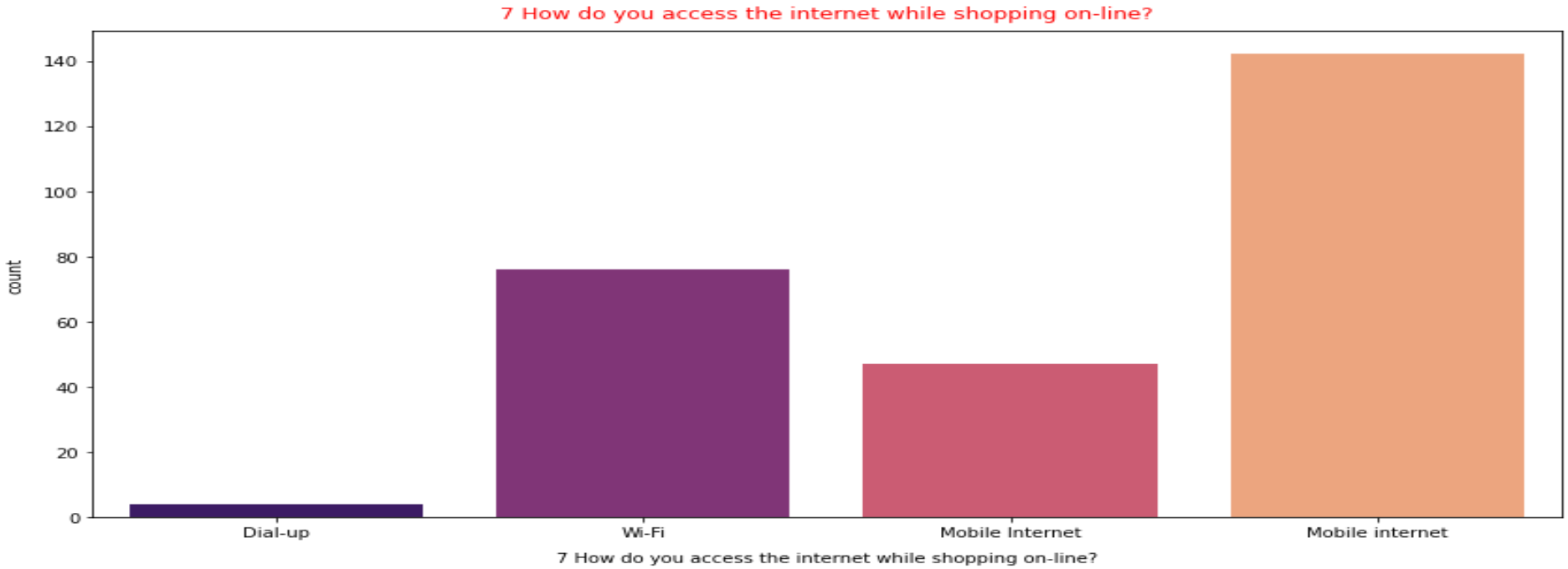

2 How old are you?
count
80
70
60
50
40
30
20
10
0
31-40 years
21-30 years
41-50 yaers
Less than 20 years
51 years and above
2 How old are you?

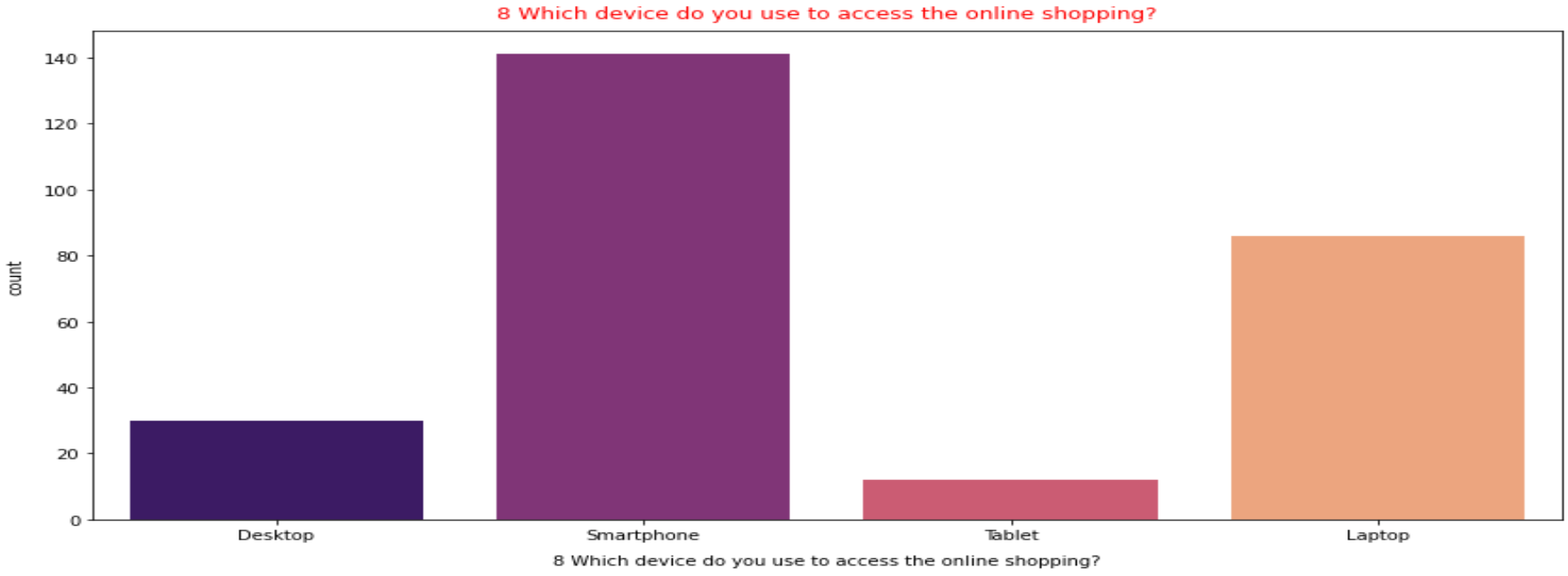

3 Which city do you shop online from?
count
60
50
40
30
20
10
0
Delhi
Greater Noida
Karnal
Bangalore
Noida
Solan
Moradabad
Gurgaon
Merrut
Ghaziabad
Bulandshahr
3 Which city do you shop online from?

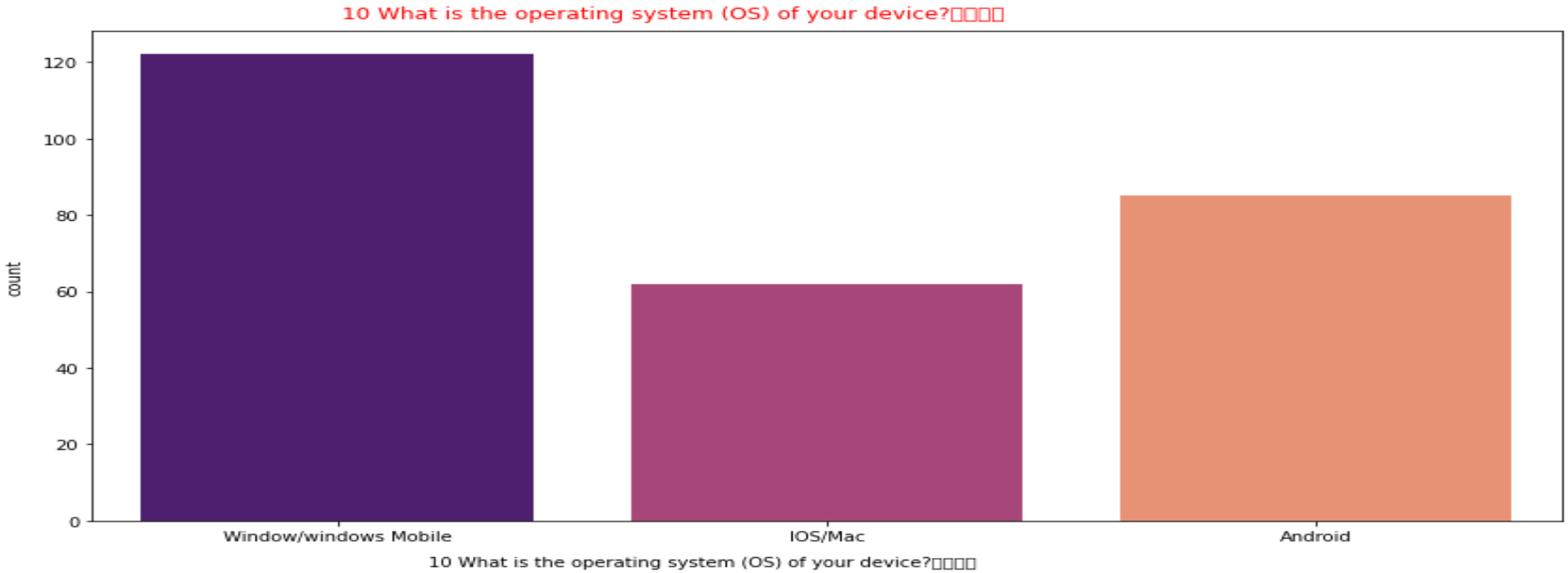

4 What is the Pin Code of where you shop online from?
count
35
30
25
20
15
10
5
0
4 What is the Pin Code of where you shop online from?

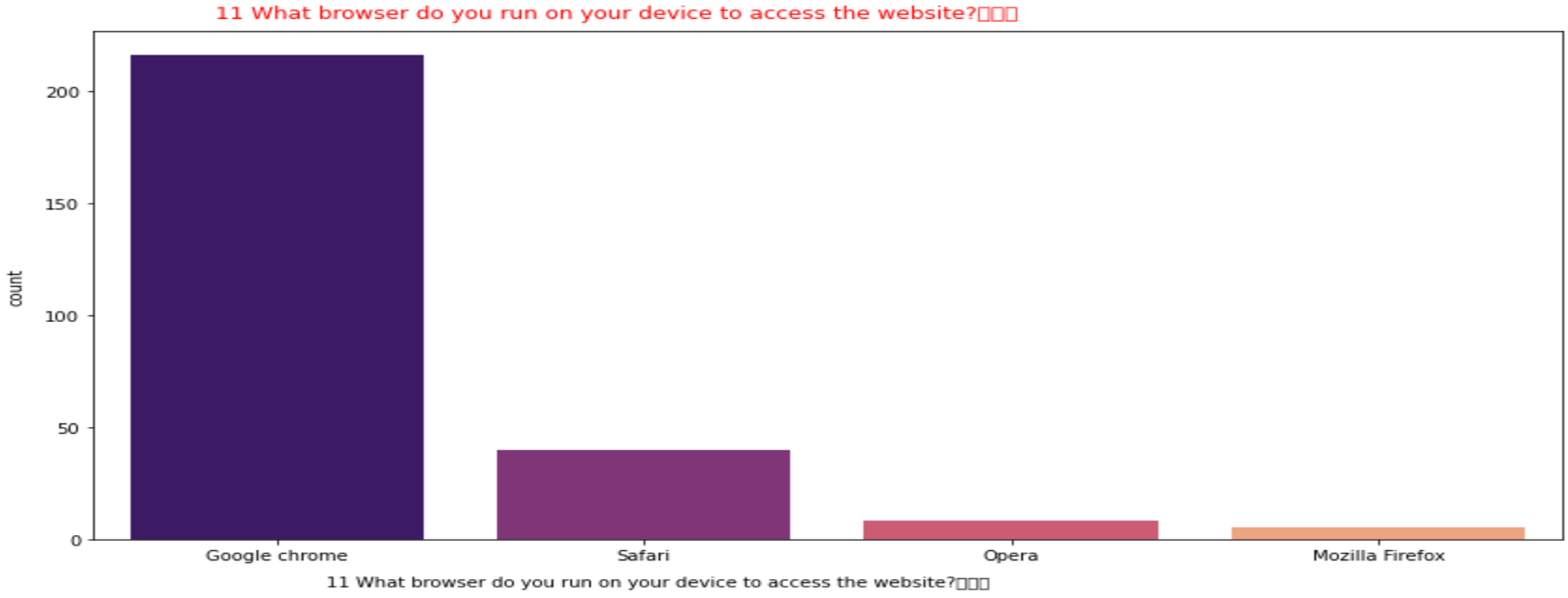

5 Since How Long You are Shopping Online ?
count
100
80
60
40
20
0
Above 4 years
3-4 years
2-3 years
Less than 1 year
1-2 years
5 Since How Long You are Shopping Online ?


6 How many times you have made an online purchase in the past 1 year?
count
100
80
60
40
20
0
31-40 times
41 times and above
Less than 10 times
11-20 times
21-30 times
42 times and above
6 How many times you have made an online purchase in the past 1 year?

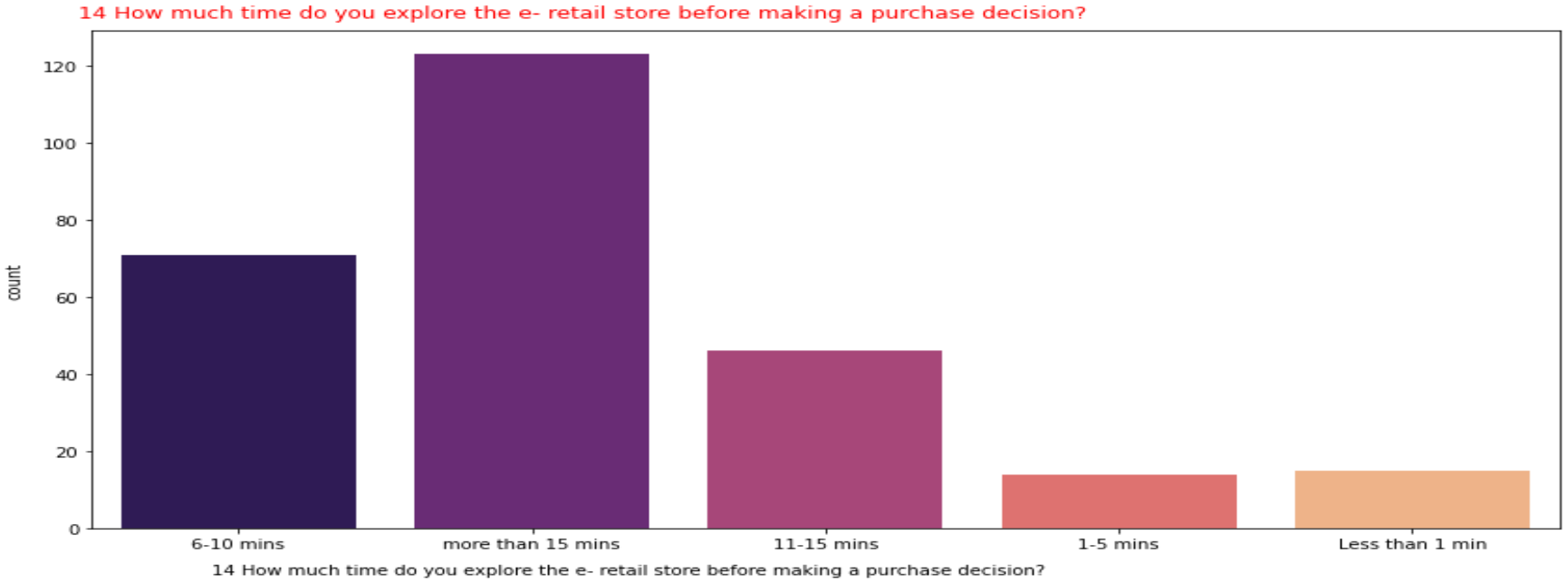

7 How do you access the internet while shopping on-line?
count
140
120
100
80
60
40
20
0
Dial-up
Wi-Fi
Mobile Internet
Mobile internet
7 How do you access the internet while shopping on-line?

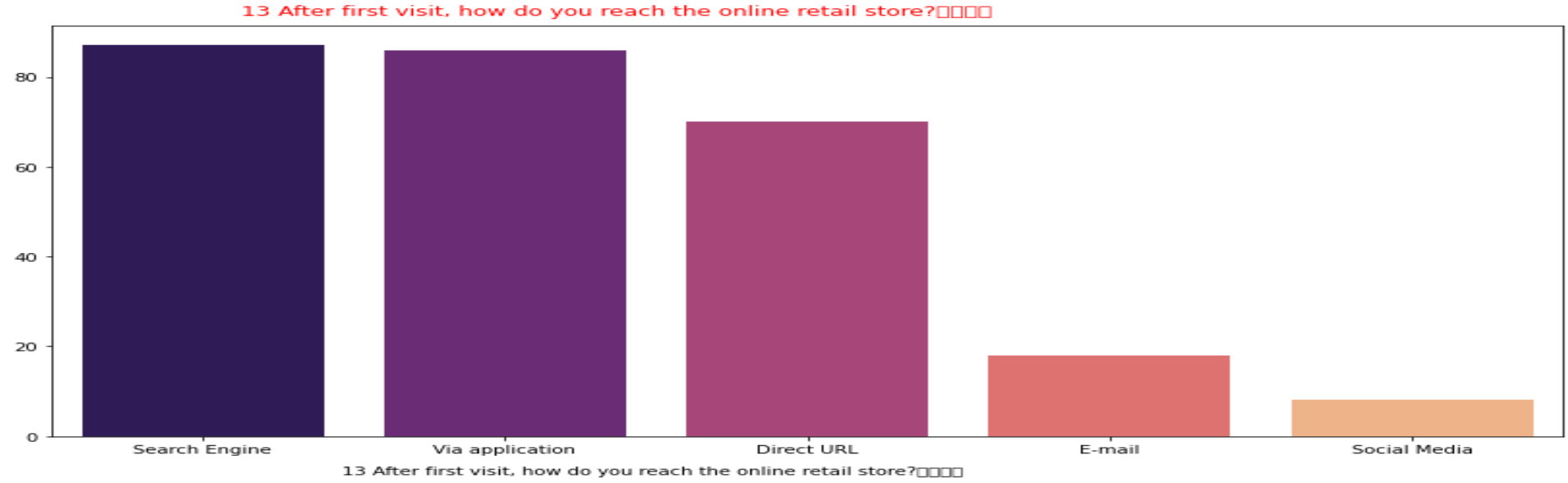

8 Which device do you use to access the online shopping?
count
140
120
100
80
60
40
20
0
Desktop
Smartphone
Tablet
Laptop
8 Which device do you use to access the online shopping?

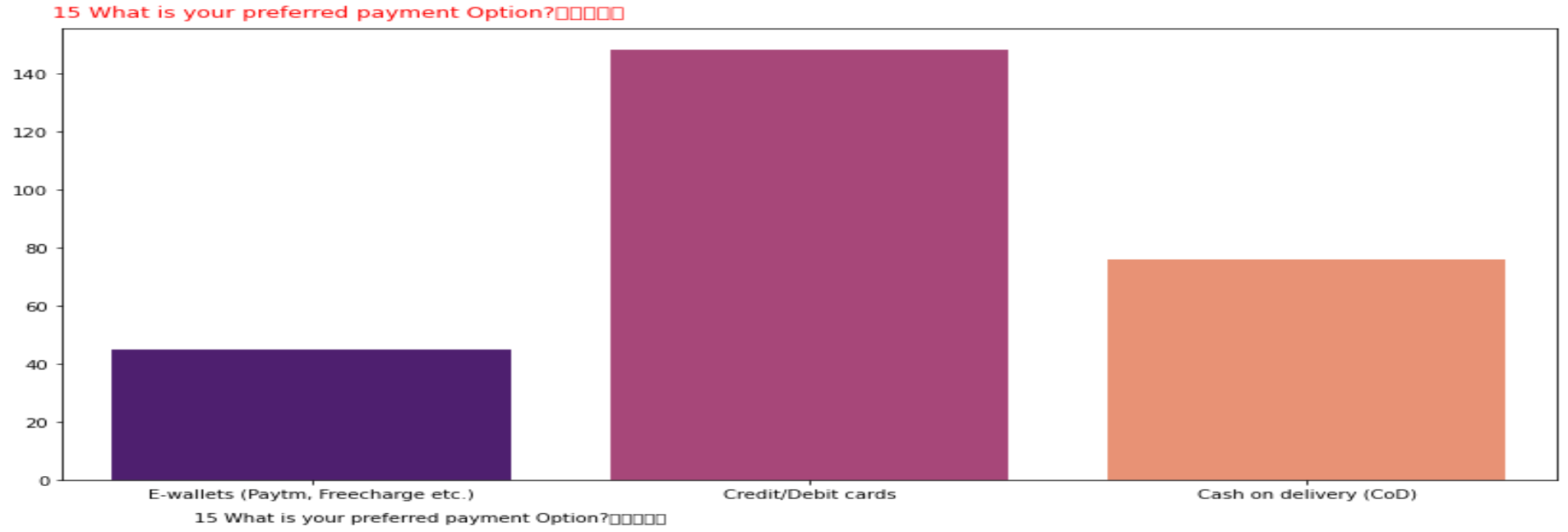

10 What is the operating system (OS) of your device?□□□□
count
120
100
80
60
40
20
0
Window/windows Mobile
IOS/Mac
Android
10 What is the operating system (OS) of your device?□□□□


11 What browser do you run on your device to access the website?□□□
count
200
150
100
50
0
Google chrome
Safari
Opera
Mozilla Firefox
11 What browser do you run on your device to access the website?□□□

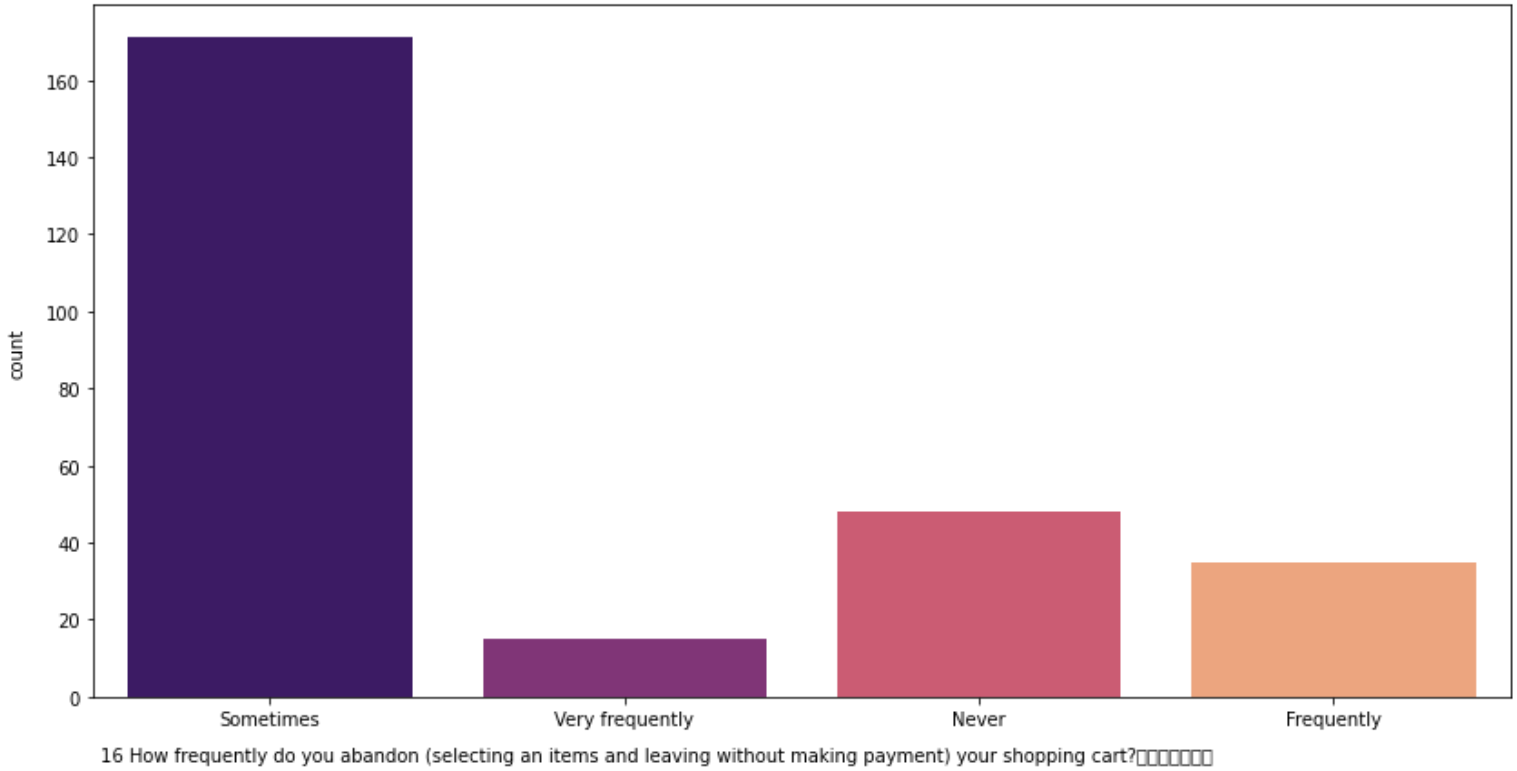

12 Which channel did you follow to arrive at your favorite online store for the first time?
count
200
150
100
50
0
Search Engine
Content Marketing
Display Adverts
12 Which channel did you follow to arrive at your favorite online store for the first time?


14 How much time do you explore the e- retail store before making a purchase decision?
count
120
100
80
60
40
20
0
6-10 mins
more than 15 mins
11-15 mins
1-5 mins
Less than 1 min
14 How much time do you explore the e- retail store before making a purchase decision?


13 After first visit, how do you reach the online retail store?□□□□
80
60
40
20
0
Search Engine
Via application
Direct URL
E-mail
Social Media
13 After first visit, how do you reach the online retail store?□□□□


15 What is your preferred payment Option?□□□□□
140
120
100
80
60
40
20
0
E-wallets (Paytm, Freecharge etc.)
Credit/Debit cards
Cash on delivery (CoD)
15 What is your preferred payment Option?□□□□□

16 How frequently do you abandon (selecting an items and leaving without making payment) your shopping cart?□□□□□□□



1

- observations from countplot:

.females are doing more shopping through online shopping than male.
.In cities like "Delhi, Greater NOida & Bangolore" are doing more shopping throgh onlines may be beacause of their busy schedule or easy/quick delivery
based on time period There are many customers doing online shopping since more than 4 yrs while every year new customer is added.
No.of times online purchase done in last 1 yr is more no of customer done online shopping upto 10 times
Most of the customers are using "Mobile Internet" for online shopping.
Most of the customers doing online shopping through their mobile Phones

- and laptops only.
- The screen size they are using is "others & 5.5 inches". may be the
- Customers are using the scrrens of more than 5.5 screen i.e laptops
- Most of the customers are using the operating system of "window/windows
- mobile"than Android.
- The browser they are using is "Google Chrome“
- "Search Engine" channel is mostly used by the customers to do online
- Shopping
- The cusomers are again using the "search Engine or Application of the
- store" to reach the online retail store.
- most of the customers are spnding morethan 15 mins on e-retail store
- before making purchase.
- most of the customers are doing payments through credit/debit cards
- The customers are abandon the items sometimes because they are getting
- better roducts in less price and good quality

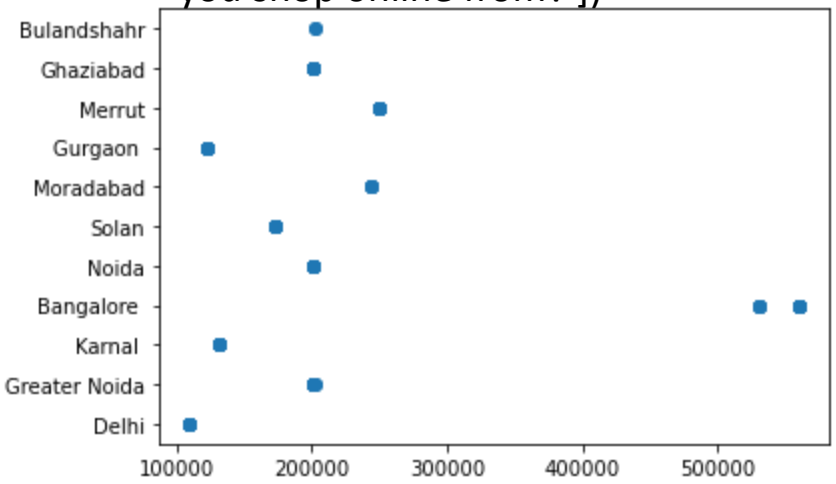
CHECKING CORRELATION

- df.corr()

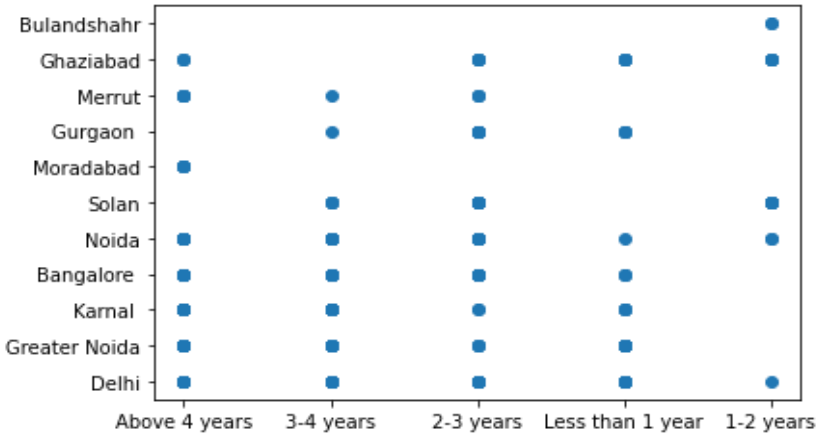
- corr_mat=df.corr()
- plt.figure(figsize=[80,80])
- sns.heatmap(corr_mat,annot=**True**)
- plt.title("CorrelationMatrix")
- plt.savefig('correlation_matrix.jpg')
- plt.show()

Bivariate Analysis

- plt.scatter(df['4 What is the Pin Code of where you shop online from?'],df['3 Which city do you shop online from?'])

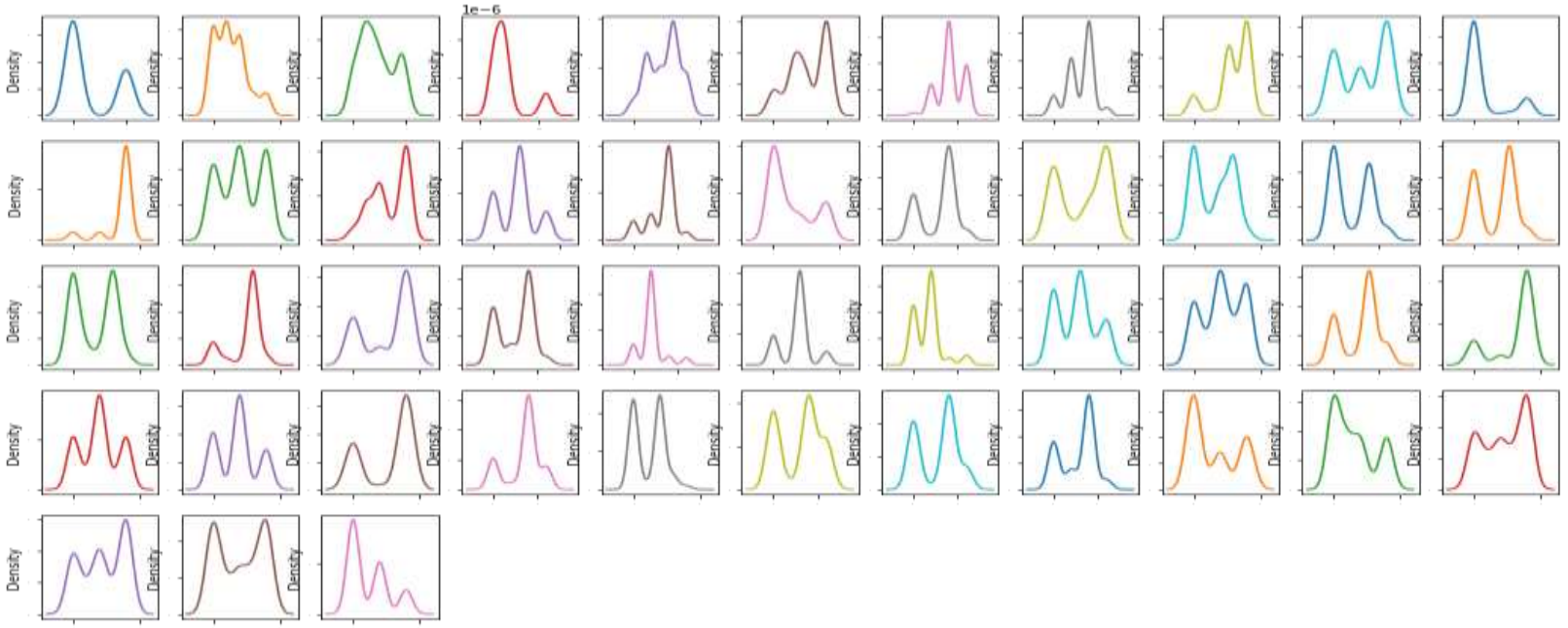


plt.scatter(df['5 Since How Long You are Shopping Online ?'],df['3 Which city do you shop online from?'])



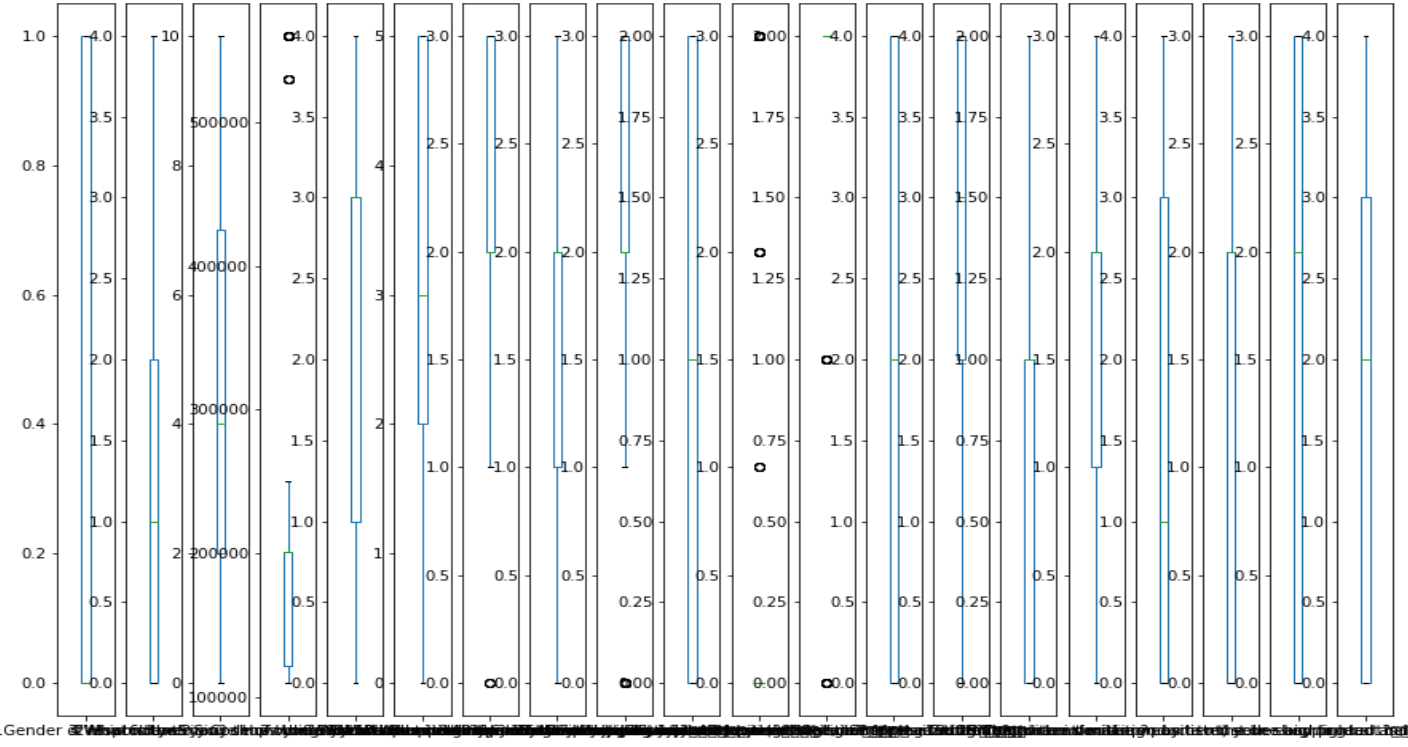
changing the categeorical values into numerical by using OrdinalEncoder

- **from** sklearn.preprocessing **import** OrdinalEncoder enc=OrdinalEncoder()
- **for** i **in** df.columns[:47]:
- **if** df[i].dtypes=='object':
- df[i]=enc.fit_transform(df[i].values.reshape(-1,1))
- df
- **Let's check the data distribution among all the columns**

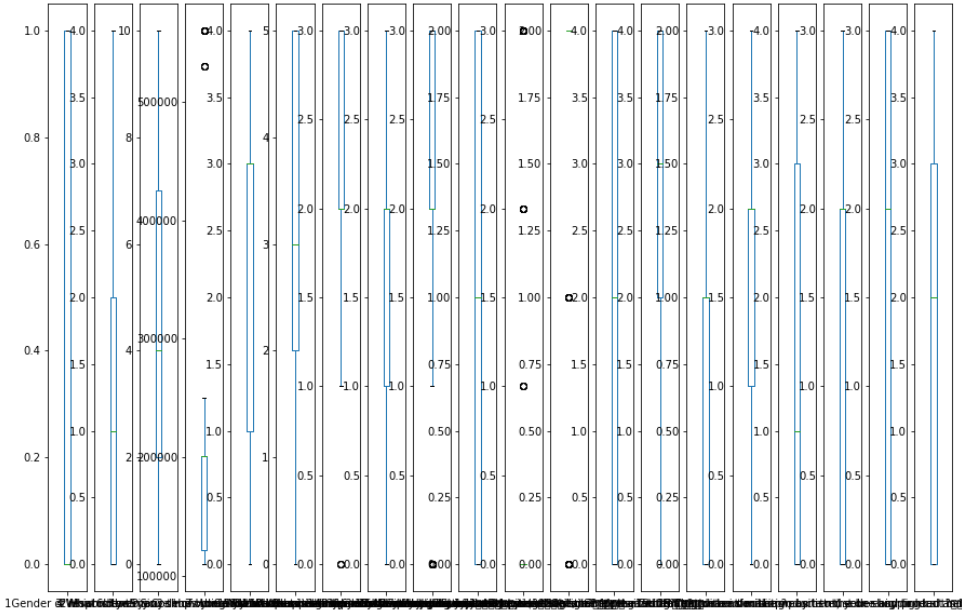


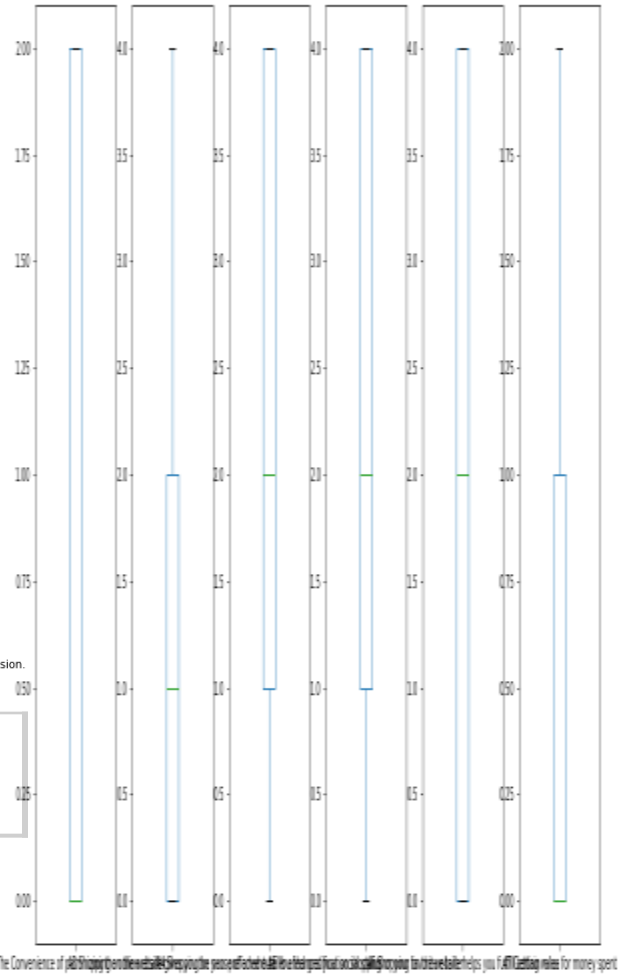
checking outliers

- There are 71 columns in dataset so its not possible to plot each and every columns seperately or plot all together also,so we will print some of the columns first and later ones as second
- *#plotting boxplots for first 20 columns*
- df1=df.iloc[:,:20]
- df1.plot(kind='box',subplots=**True**,figsize=(15,10))
- plt.show()



- df1=df.iloc[:,21:40]
- df1.plot(kind='box',subplots=**True**,figsize=(15,10)) plt.show()





```
df1=df.iloc[:,41:72]

df1.plot(kind='box',subplots=True,figsize=(15,10))
plt.show()
```

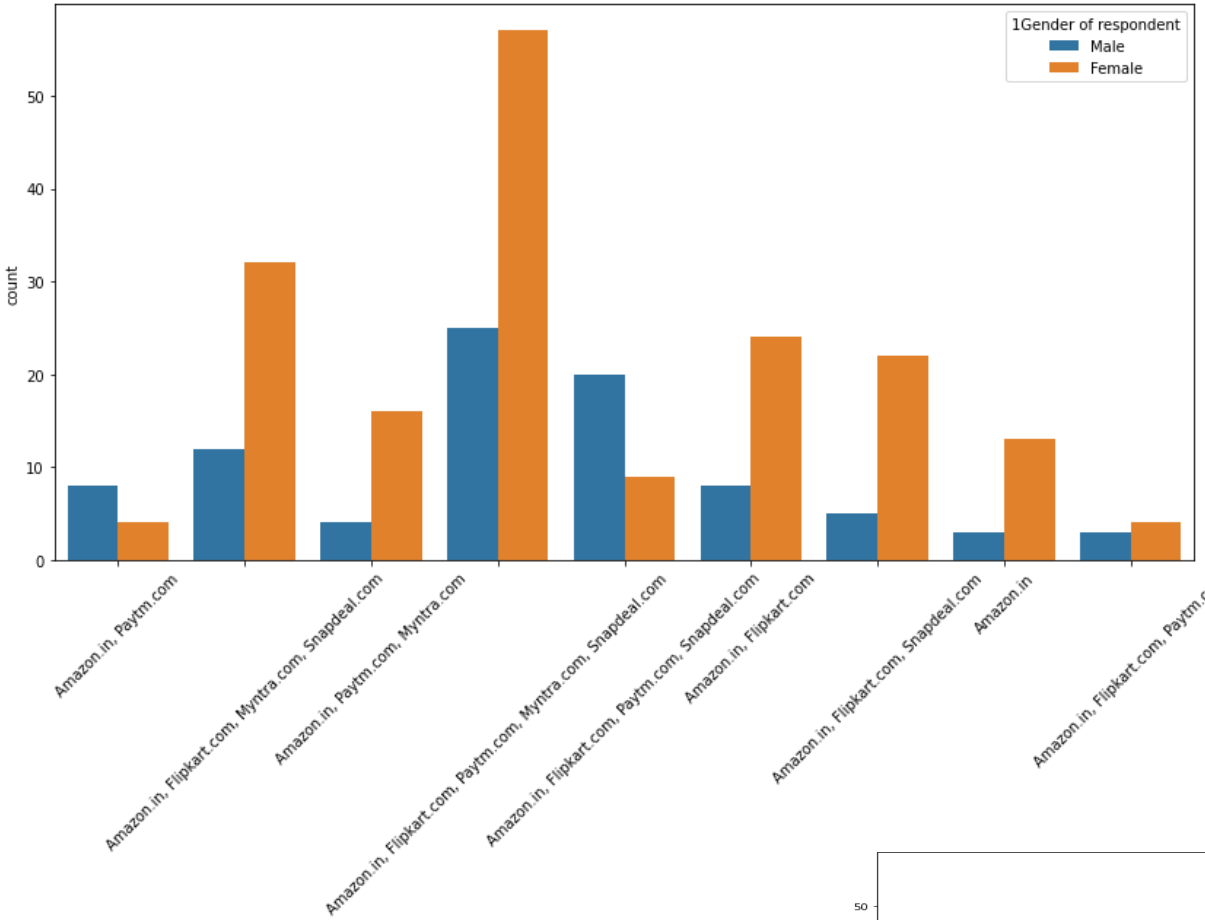
There are no outliers present **in** the data.

- **Analysing the websites data**
- websites=df.iloc[:,47:]
- websites

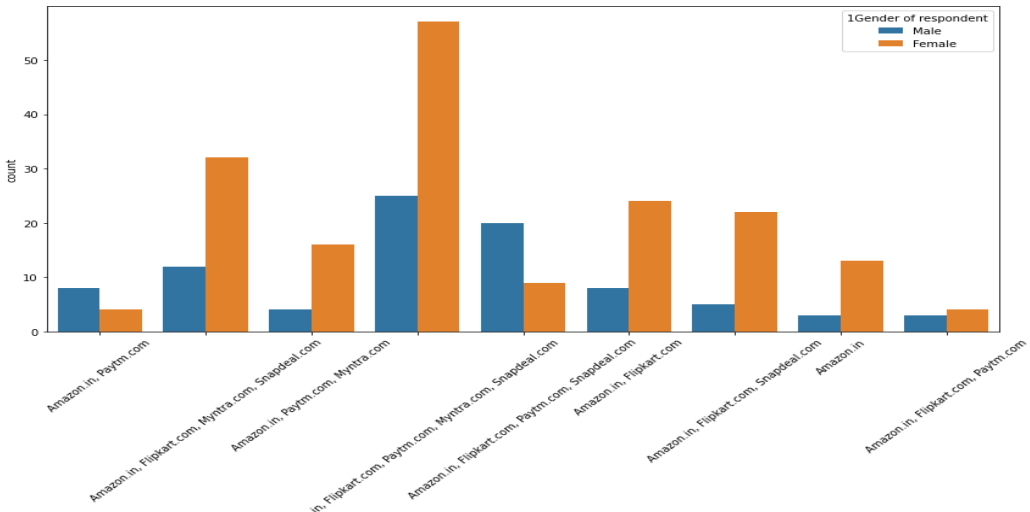
- df['From the following, tick any (or all) of the online retailers you have shopped from; '].value_counts()
-]: Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com 82
- Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com 44
- Amazon.in, Flipkart.com 32
- Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com 29
- Amazon.in, Flipkart.com, Snapdeal.com 27
- Amazon.in, Paytm.com, Myntra.com 20
- Amazon.in 16
- Amazon.in, Paytm.com 12
- Amazon.in, Flipkart.com, Paytm.com 7
- Name: From the following, tick any (or all) of the online retailers you have sh opped from;
- , dtype: int64

- **for** i **in** websites:
- websites[i]=websites[i].str.replace('Amazon.in','Amazon.com')
  websites[i]=websites[i].str.replace('Snapdeal','Snapdeal.com')
- websites[i]=websites[i].str.replace('Snapdeal.com.com','Snapdeal.com')
  websites[i]=websites[i].str.replace('Paytm.com','Paytm.com')
- websites[i]=websites[i].str.replace('Flipkart.com','Flipkart.com') websites

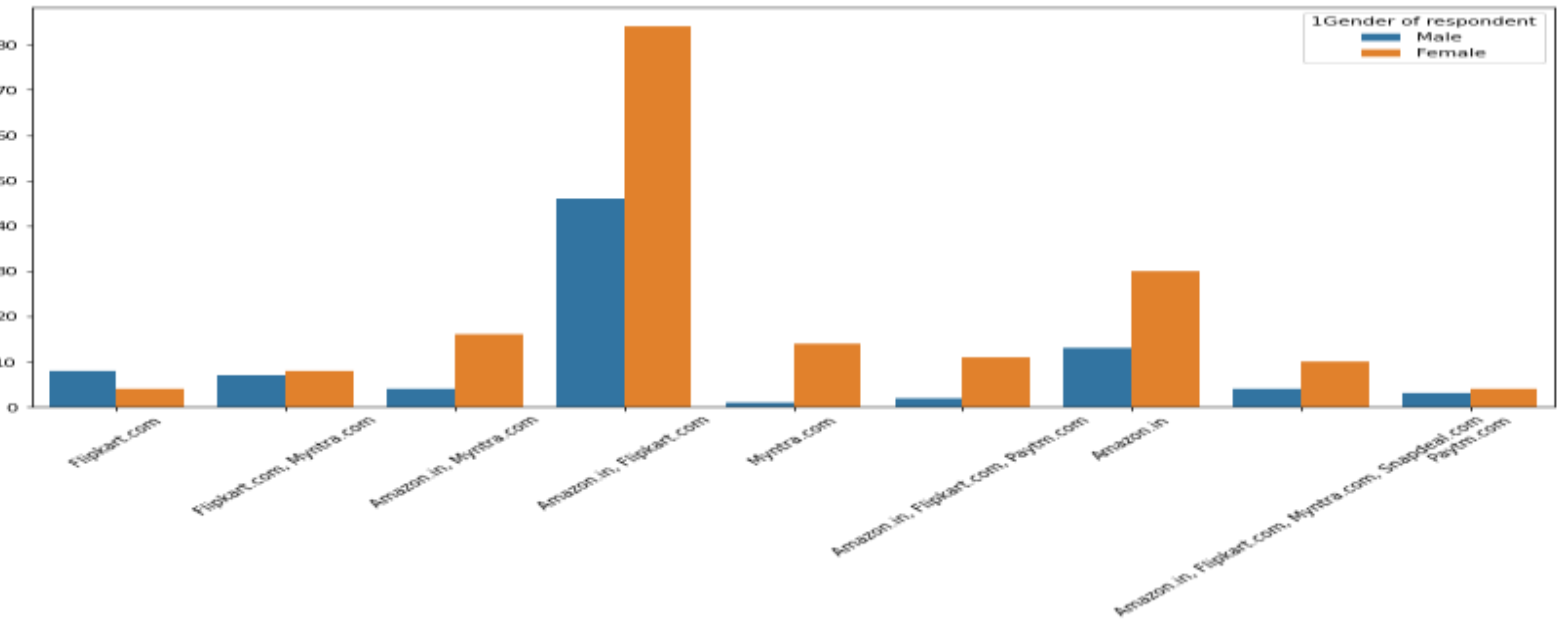
- plt.figure(figsize=(14,7))
- sns.countplot(df['From the following, tick any (or all) of the online retailers you have shopped from; '],hue=df['1Gender of respondent'])
- plt.xticks(rotation=45)
- array([0, 1, 2, 3, 4, 5, 6, 7, 8]), [Text(0, 0, 'Amazon.in, Paytm.com'), Text(1, 0, 'Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com'), Text(2, 0, 'Amazon.in, Paytm.com, Myntra.com'), Text(3, 0, 'Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com'), Text(4, 0, 'Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com'), Text(5, 0, 'Amazon.in, Flipkart.com'), Text(6, 0, 'Amazon.in, Flipkart.com, Snapdeal.com'), Text(7, 0, 'Amazon.in'), Text(8, 0, 'Amazon.in, Flipkart.com, Paytm.com')])
-
-



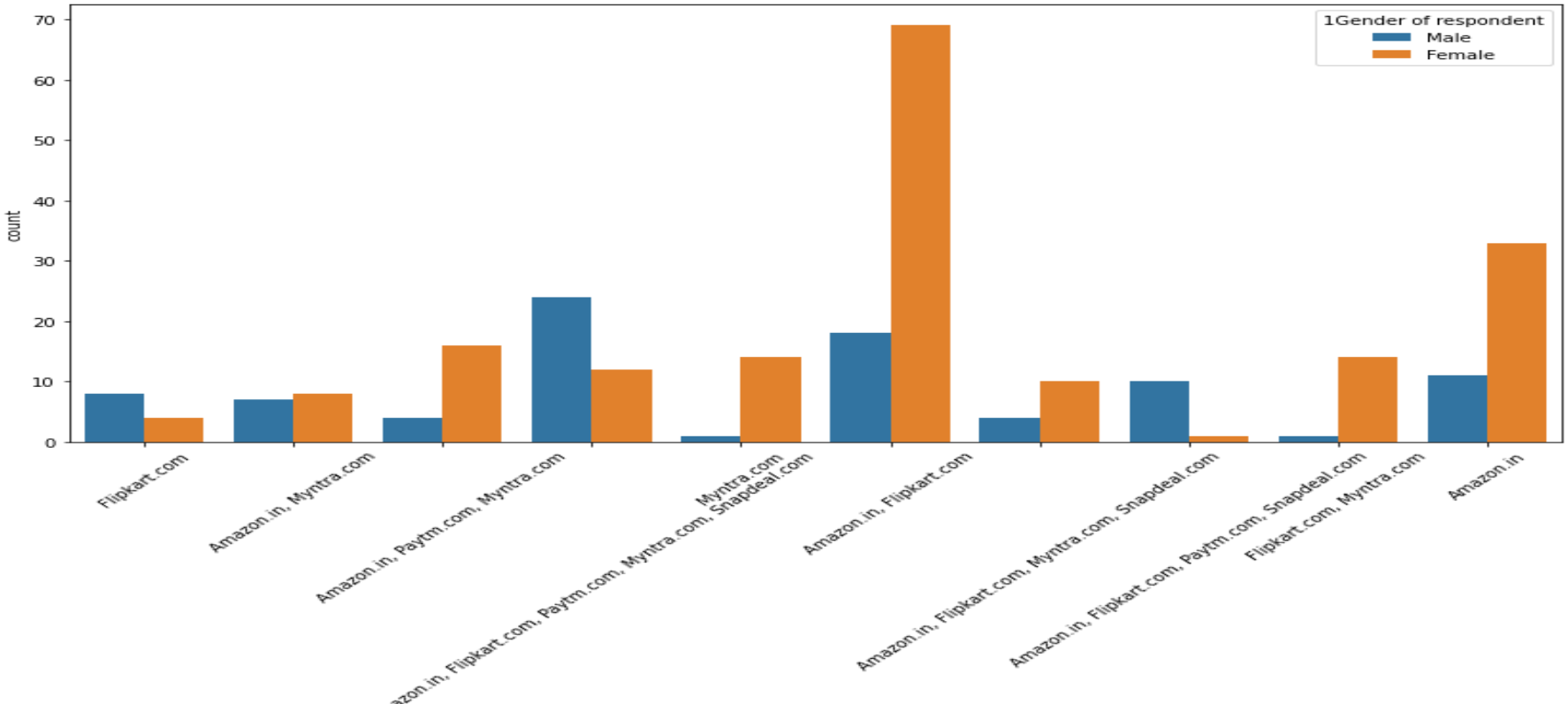
From the following, tick any (or all) of the online retailers you have shopped from;

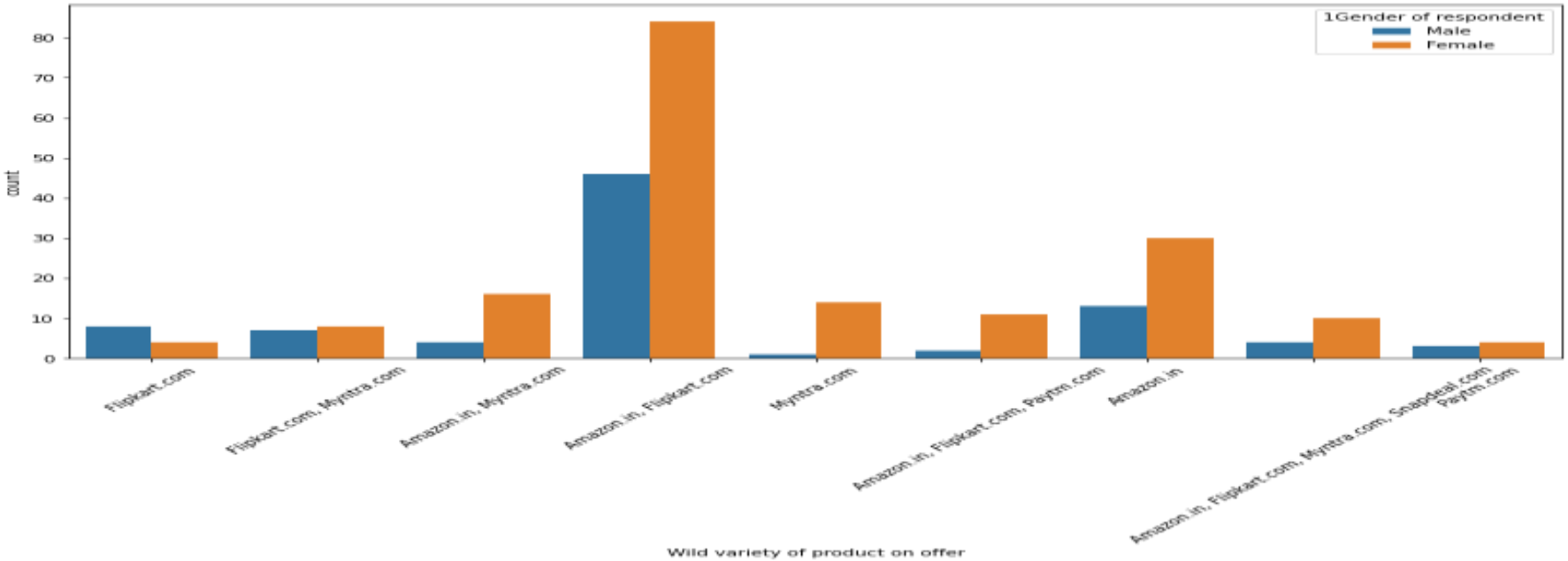


From the following, tick any (or all) of the online retailers you have shopped from;

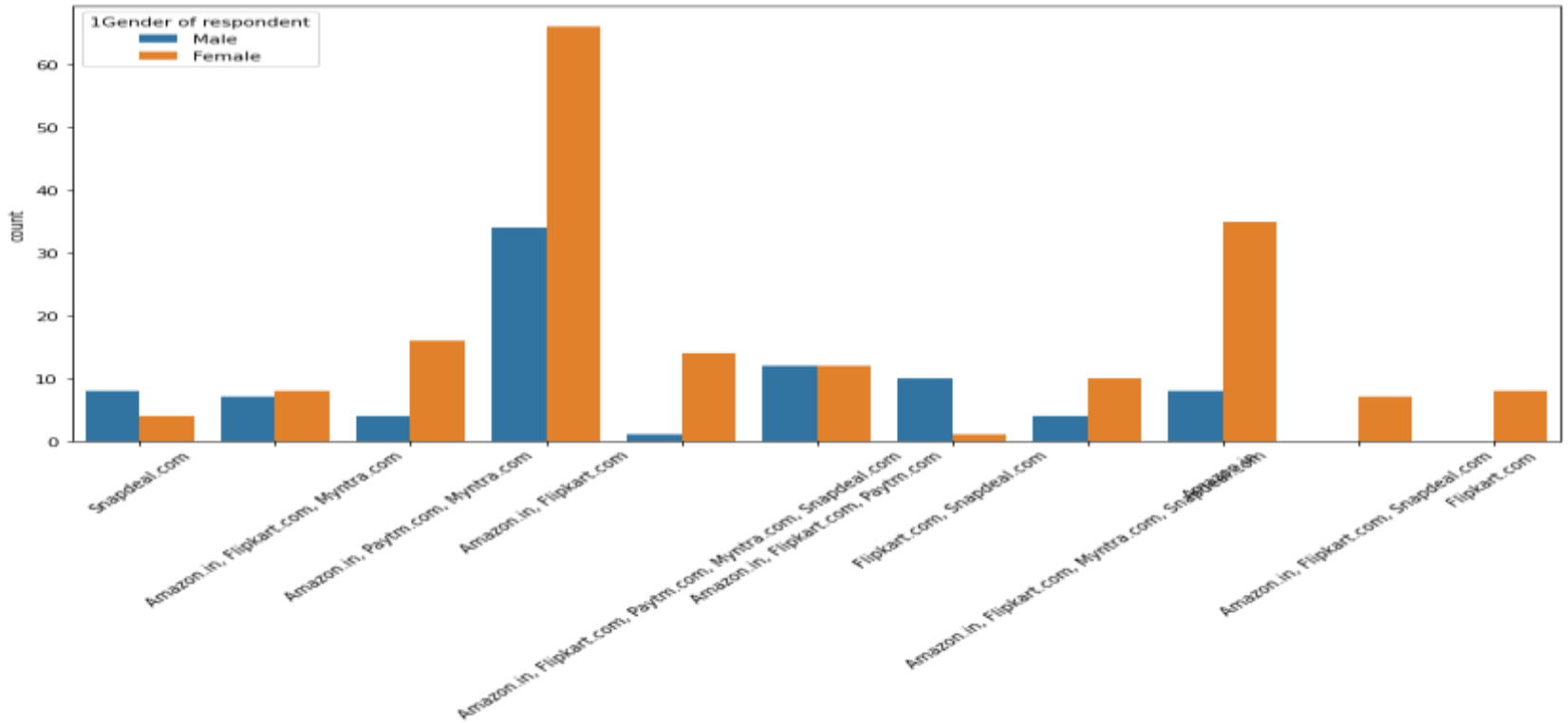


Wild variety of product on offer

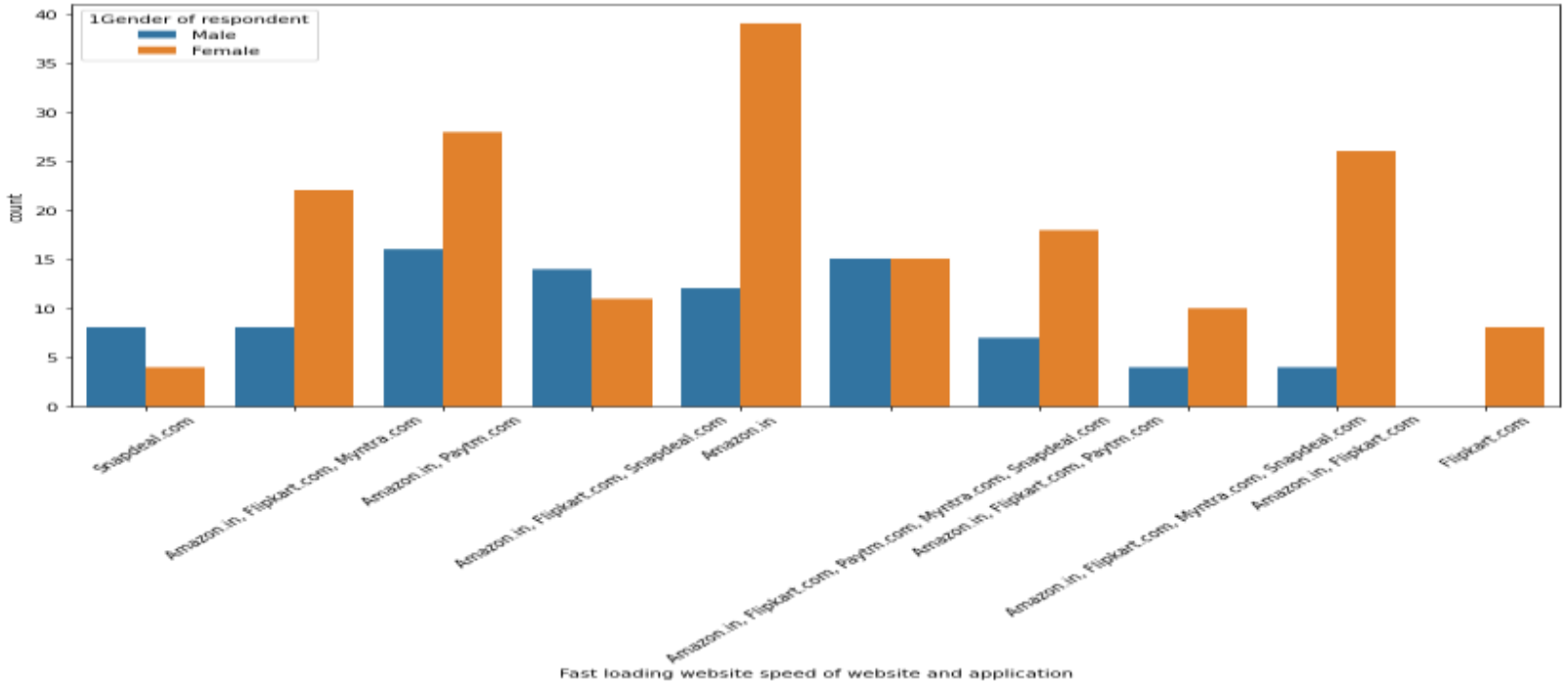




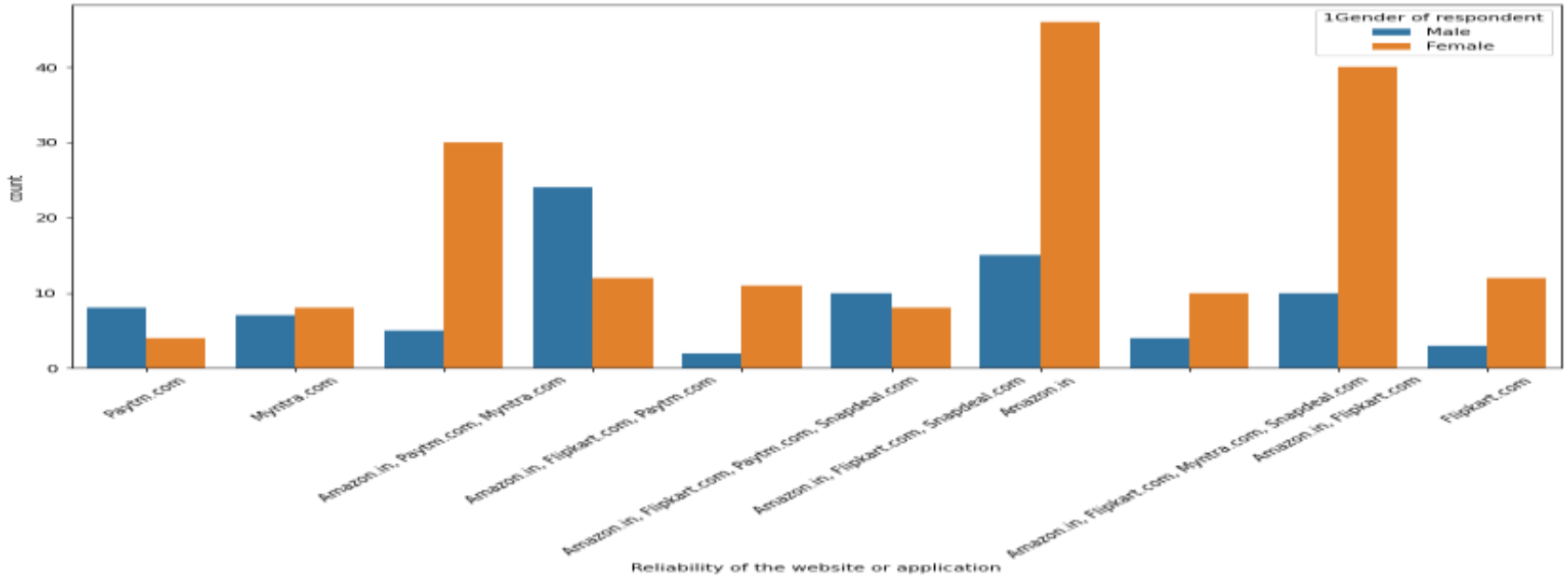
Wild variety of product on offer



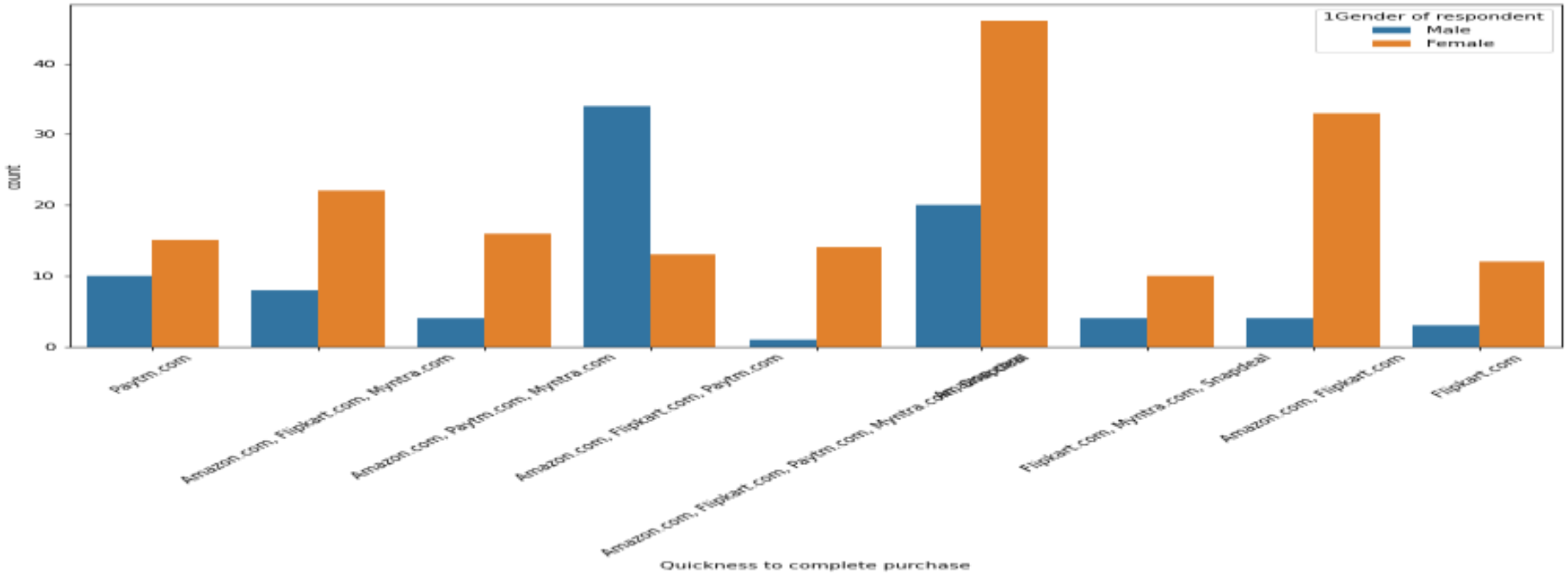
Complete, relevant description information of products

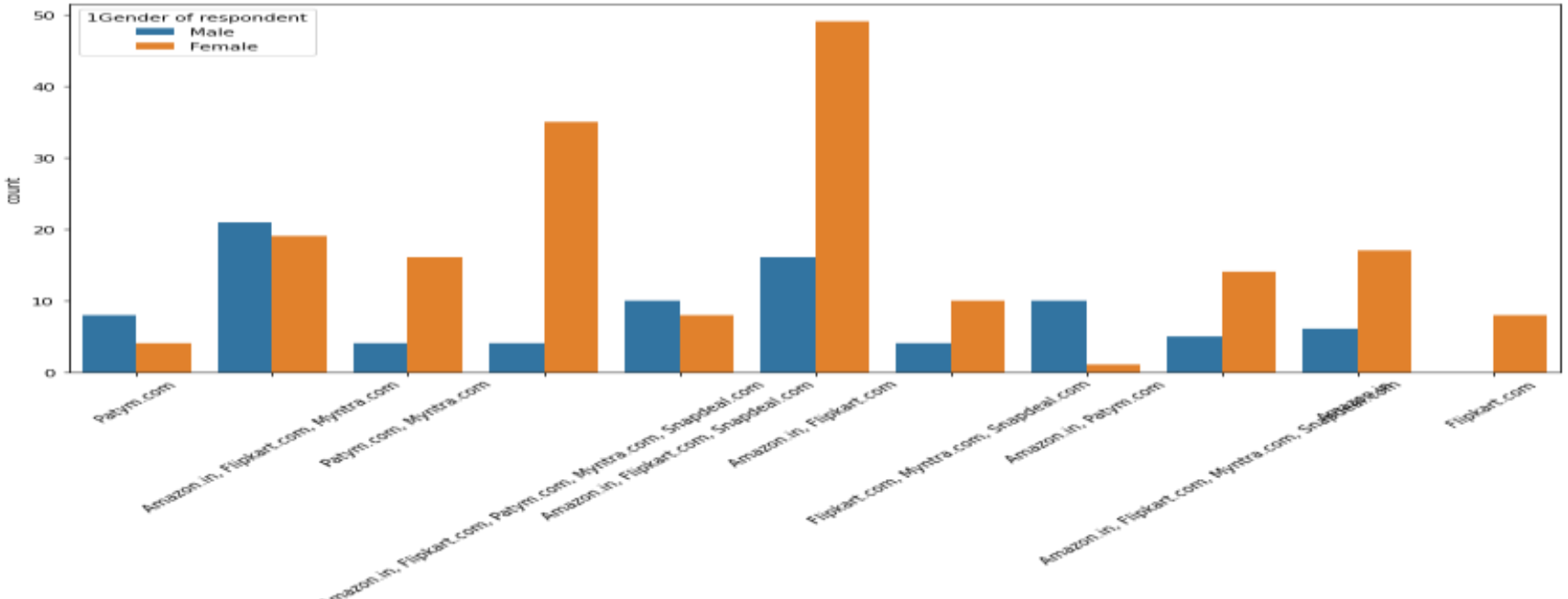


Fast loading website speed of website and application



Reliability of the website or application


1Gender of respondent
Male
Female
count
0
10
20
30
40
Paytm.com
Amazon.com, Flipkart.com, Myntra.com
Amazon.com, Paytm.com, Myntra.com
Amazon.com, Flipkart.com, Paytm.com
Amazon.com, Flipkart.com, Paytm.com, Myntra.com
Flipkart.com, Myntra.com, Snapdeal
Amazon.com, Flipkart.com
Flipkart.com
Quickness to complete purchase


1Gender of respondent
Male
Female
count
0
10
20
30
40
50
Patym.com
Amazon.in, Flipkart.com, Myntra.com
Patym.com, Myntra.com
Amazon.in, Flipkart.com, Patym.com, Myntra.com, Snapdeal.com
Amazon.in, Flipkart.com, Snapdeal.com
Amazon.in, Flipkart.com
Flipkart.com, Myntra.com, Snapdeal.com
Amazon.in, Patym.com
Flipkart.com

- From the following, tick any (or all) of the online retailers you have shopped from; most of the people shopped from Amazon.in,Myntra.com and paid from paytm.com
- most of the people used Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com(64) as Ease to use website or application
- Amazon.in, Flipkart.com are mostly used as Visual appealing web-page layout
- Wild variety of product on offer is on Amazon.in and flipkart.com are the two sites widely used
- Complete, relevant description information of products is given by Amazon.in and Flipkart.com sites only
- Fast loading website speed of website and application is in Amazon.in website
- Most Reliability of the website or application is Amazon.in
- For Quickness to complete purchase females are using Amazon.com,males are using Amazon.com,Flipkart.com, and for puchasing they are using paytm.com
- For Availability of several payment options females are using Amazon.com,Flipkart.com,snapdeal.com and males are using Amazon.com,Flipkart.com,snapdeal.com,Myntra.com
- Privacy of customers' information is more available in Amazon.in and then Flipkart.com and then Myntra.com
- Security of customer financial information is in most in Amazon.in, and then Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com
- Longer time to get logged in (promotion, sales period) is in Amazon.in,Flipkart.com
- Longer page loading time (promotion, sales period)' is much time in myntra.com and then snapdeal.com
- Limited mode of payment on most products (promotion, sales period) is in snapdeal.com
- Longer delivery period is in snapdeal.com
- Frequent disruption when moving from one page to another is in Amazon.com and Myntra.com less in flipkart.com and snapdeal.com
- Website is as efficient as before is Amazon.com
- Which of the Indian online retailer would you recommend to a friend? is Amazon.in and Flipkart.com

Conclusion

- Amazon.com --is the most recomanded website –page layout, easy to use, relevant descriptive information, product offers reliability of website, fastness to purchase,trust worthiness
- →And it 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.com-- is the second most website recommended easy to use,payments easy, fast opening of website,trust worthy. But it display longer time for graphics, security of customer info less, less relaibility.
- Myntra.com—is the Third most website recommended because of its easy access, payment options.
- Paytm.com– speedy deliver of products, quickness in purchase
- Snapdeal.com--- is the least website recommended because of its less payment options, no speedy delivery, longer time to go from one page to another

- Thank you