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

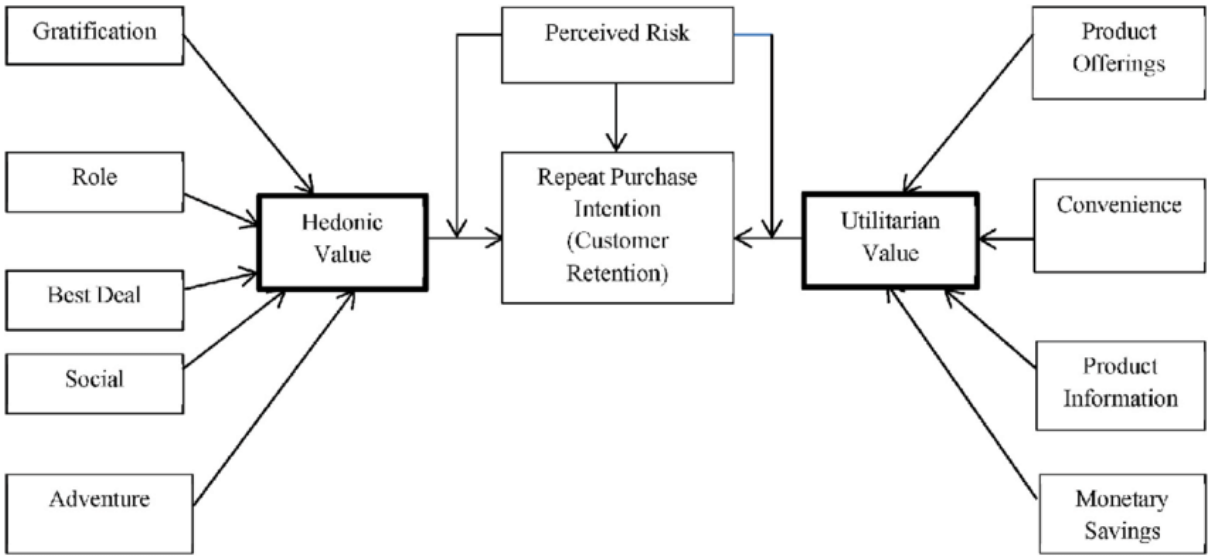
My name is Chaitali Nakade and in this report I will give the information about E-retail factors for customer activation and retention. Customer satisfaction has developed as one of the most important factors that guarantee the success of online store.

We can see now a days people are more attracted towards online shopping because e-retail shopping provides them a huge verity of products and services according to there choices. It will save our time as well as our transportation charges too because most of the online shopping portal gives us free delivery at our door steps. After market research we found that Indian E-retail market experience a very good demand of services and products since last six to seven years. In this report we will see the factors which highly affect the services and demands in online shopping.

PROBLEM DEFINITION: -

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.



In above diagram we can see the factors which very important for any online services. If any online shopping platform wants to attract the customers for repeat purchase of products from same site than they need to give the best quality of services like they have to provide best deals, social marketing as much as possible, product information, very good product information and so many things.

Now, let check the data set and analyse it with some important factors.

DATA ANALYSIS: -

Importing the Libraries:

```
#Import libreries

import pandas as pd #Data processing
import numpy as np #Linear algebra
from sklearn.preprocessing import StandardScaler #resize the distribution of values
from sklearn.linear_model import LinearRegression #algorithm
from sklearn.model_selection import train_test_split #estimate the performance
import statsmodels.api as sm
import matplotlib.pyplot as plt #data visualization
import seaborn as sn  #data visualization
import pickle #model saving

import warnings
warnings.filterwarnings("ignore")
```

All these libraries are used for model building and data processing. Uses of all libraries are mentioned over there.

Getting the Dataset:

```
#Import the dataset

data= pd.read_excel(r"C:\Users\Chaitali Nakade\Downloads\Customer_retention_dataset--1-\Customer_retention_dataset\customer_retention_dataset.xlsx")
data.head()
```

	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	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
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

The complete dataset is imported in variable name data. Now, we can analyse the data.

EXPLORATORY DATA ANALYSIS(EDA): -

Now, let's check what will be the important factors for successful online shopping portal and which factor affects the demand in it.

It is used to analyse and examine data sets and review their main features, frequently using data visualization methods.

This dataset is having 71 columns. We can not analyse it directly. So, we need to put following instruction through which we can analyse the all-columns details.

```
pd.set_option('display.max_rows', None)
```

Shape of data with number of rows and columns

```
data.shape
```

```
(269, 71)
```

There are 269 rows and 71 columns are present in dataset.

Basic information of Dataset:

```
data.info()
```

Output:

```
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
 55  Availability of several payment options
269 non-null    object
 56  Speedy order delivery
269 non-null    object
 57  Privacy of customers' information
269 non-null    object
 58  Security of customer financial information
269 non-null    object
 59  Perceived Trustworthiness
269 non-null    object
 60  Presence of online assistance through multi-channel
269 non-null    object
 61  Longer time to get logged in (promotion, sales period)
269 non-null    object
 62  Longer time in displaying graphics and photos (promotion, sales period)
269 non-null    object
 63  Late declaration of price (promotion, sales period)
269 non-null    object
 64  Longer page loading time (promotion, sales period)
269 non-null    object
 65  Limited mode of payment on most products (promotion, sales period)
269 non-null    object
 66  Longer delivery period
269 non-null    object
 67  Change in website/Application design
269 non-null    object
 68  Frequent disruption when moving from one page to another
269 non-null    object
 69  Website is as efficient as before
269 non-null    object
 70  Which of the Indian online retailer would you recommend to a friend?
269 non-null    object
dtypes: int64(1), object(70)
memory usage: 149.3+ KB
```

- With this instruction we will get the basic information like is there any null value present in whole dataset, datatypes of all column and non-null values of column and column names. As shown in output there is no null value values in whole dataset.
- Simultaneously we can find here only one column is having integer data type and other 70 columns are having object data type.

```
data.describe()
```

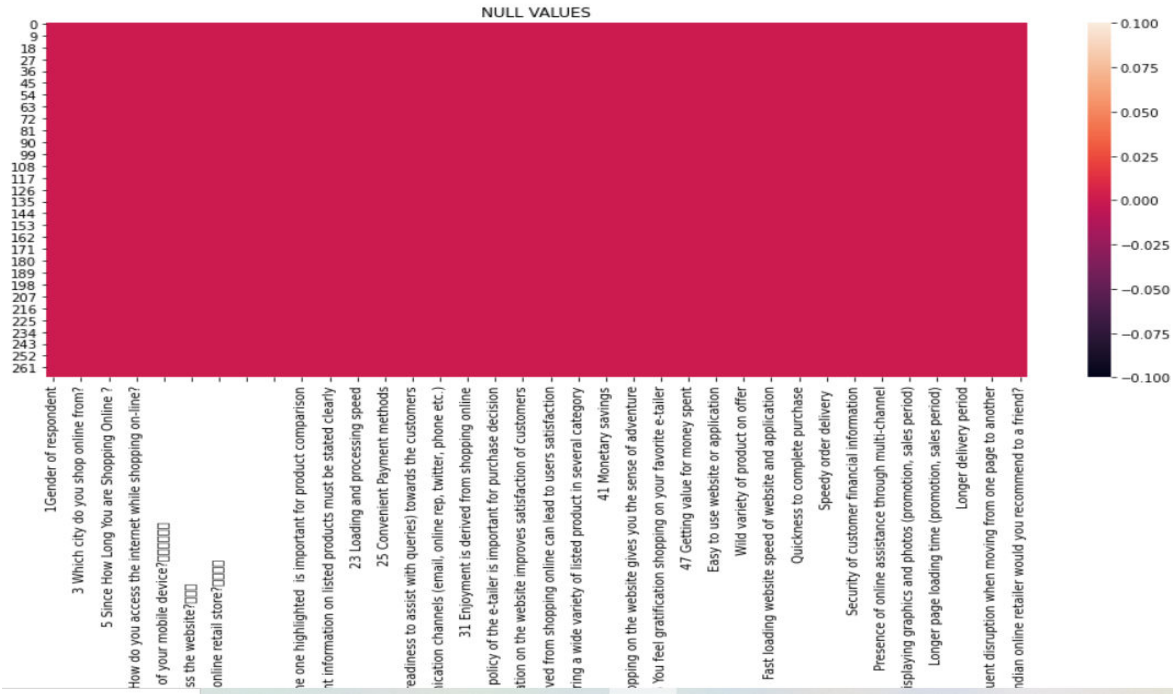
	4 What is the Pin Code of where you shop online from?
count	269.000000
mean	220465.747212
std	140524.341051
min	110008.000000
25%	122018.000000
50%	201303.000000
75%	201310.000000
max	560037.000000

We can see the detailed description of all the rows and columns which are having numeric data only. the detailed description like count, mean value, standard deviation, minimum and maximum value.

Data Visualization:

Through visualization let's check is there any null value present in dataset or not

```
plt.figure(figsize=[16,6])
sn.heatmap(data.isnull())
plt.title("NULL VALUES")
plt.show()
```



Visualization of null value through heat map. We can see in above plot there is no white lines are present in full graph which means there is no null value present in whole dataset.

Now see the visualization through **count plot:**

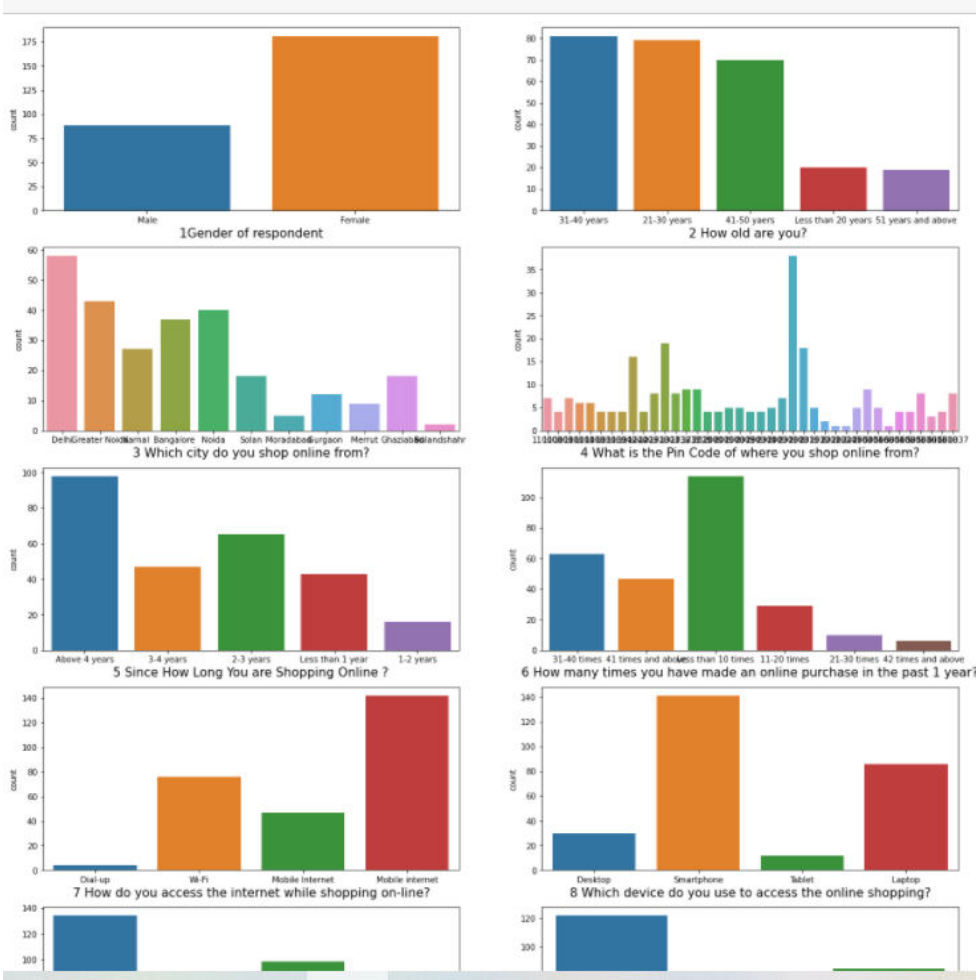
Visualization (Uni Variate analysis)

Uni variate analysis works with only one variable, hence it is called uni variate.

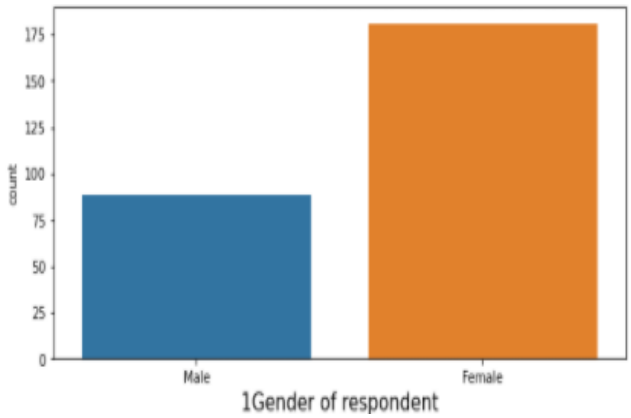
```
plt.figure(figsize = (20,50), facecolor = 'white')
plotnumber = 1

for column in data[0:20]:
    if plotnumber <= 20: # as we see there are eight columns in the data
        ax = plt.subplot(10,2,plotnumber)
        sn.countplot(data[column])
        plt.xlabel(column,fontsize=15)

    plotnumber += 1
plt.show()
```

A countplot basically **counts the classes and returns a quantity of their occurrences.**

First, I will plot only 20 columns out of 71. So, I can analyse it properly. The output of count plot will be like this;

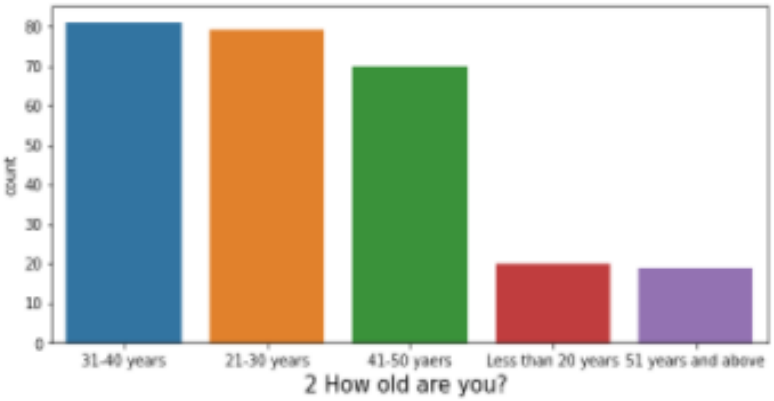


1. Gender of respondent: -



after visualizing the data, we conclude that females are more likely do the shopping as compare to males

2. How old are you?

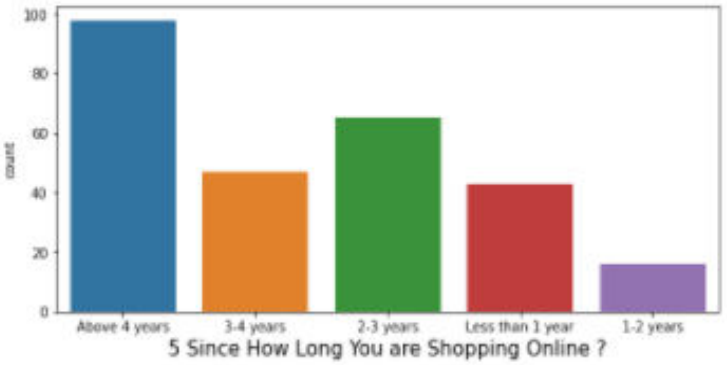


Majority of people whose age are lies in between 21 to 4o years are usually do the more shopping with the peak at age 31 to 40years.

3. Which city do you online from? :-

Delhi people are more likely to do more shopping as compare to other cities people.

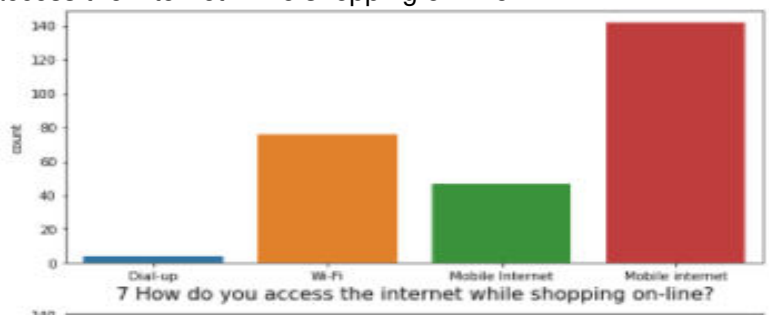
4. Since how long you are shopping online?



5 Since How Long You are Shopping Online ?

More customers are doing shopping since last 4 years with peak of 95.

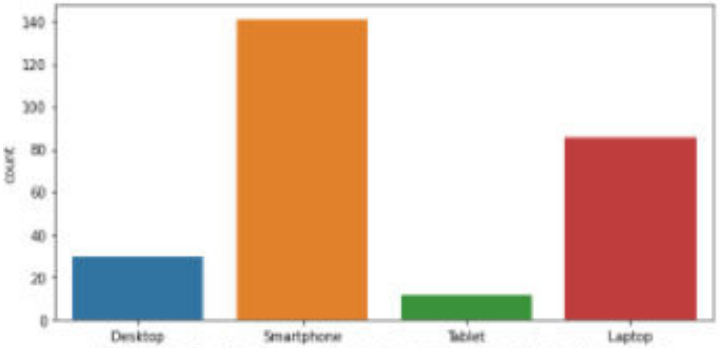
5. How do you access the internet while shopping on-line?



7 How do you access the internet while shopping on-line?

Mobile internet is priority of customers while doing online shopping

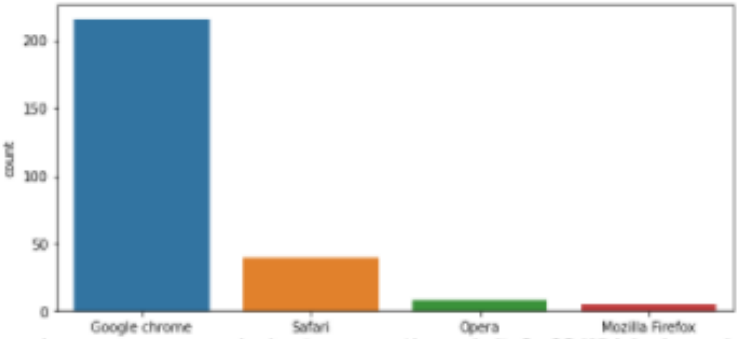
6. Which device do you use to access the online shopping?



8 Which device do you use to access the online shopping?

Maximum people are likely to do shopping through smartphones and tablet is in least priority to do shopping.

7. What browser do you run on your device to access the website? :-



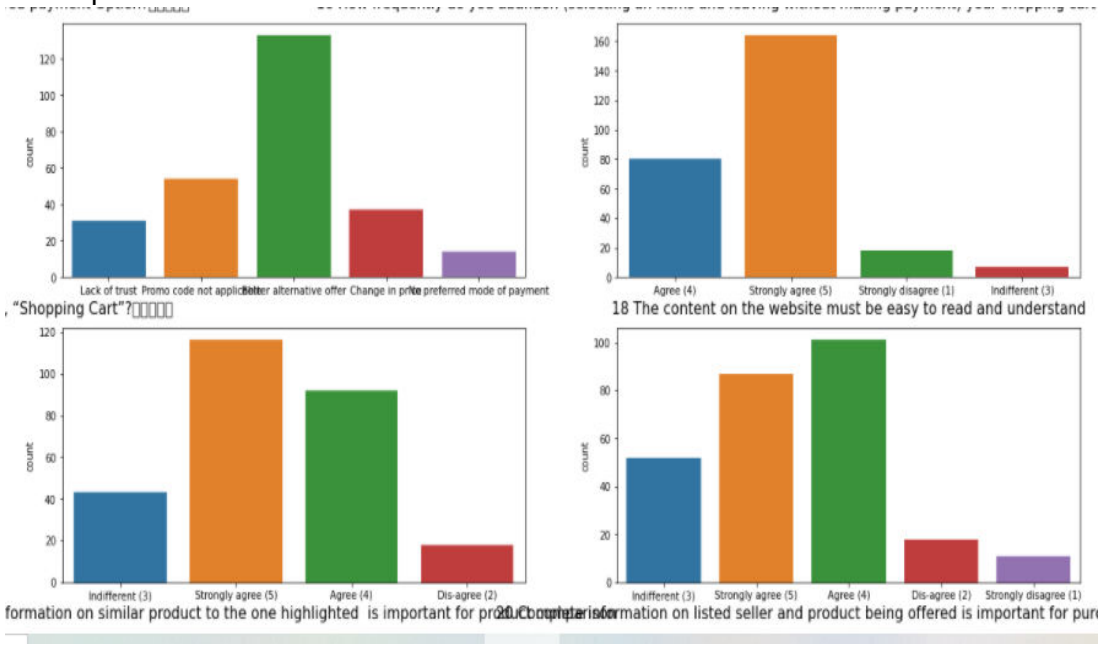
11 What browser do you run on your device to access the website?□□12 Which channel di

Google chrome are more likely to used by customer for online shopping and mozilla firefox is having least priority.

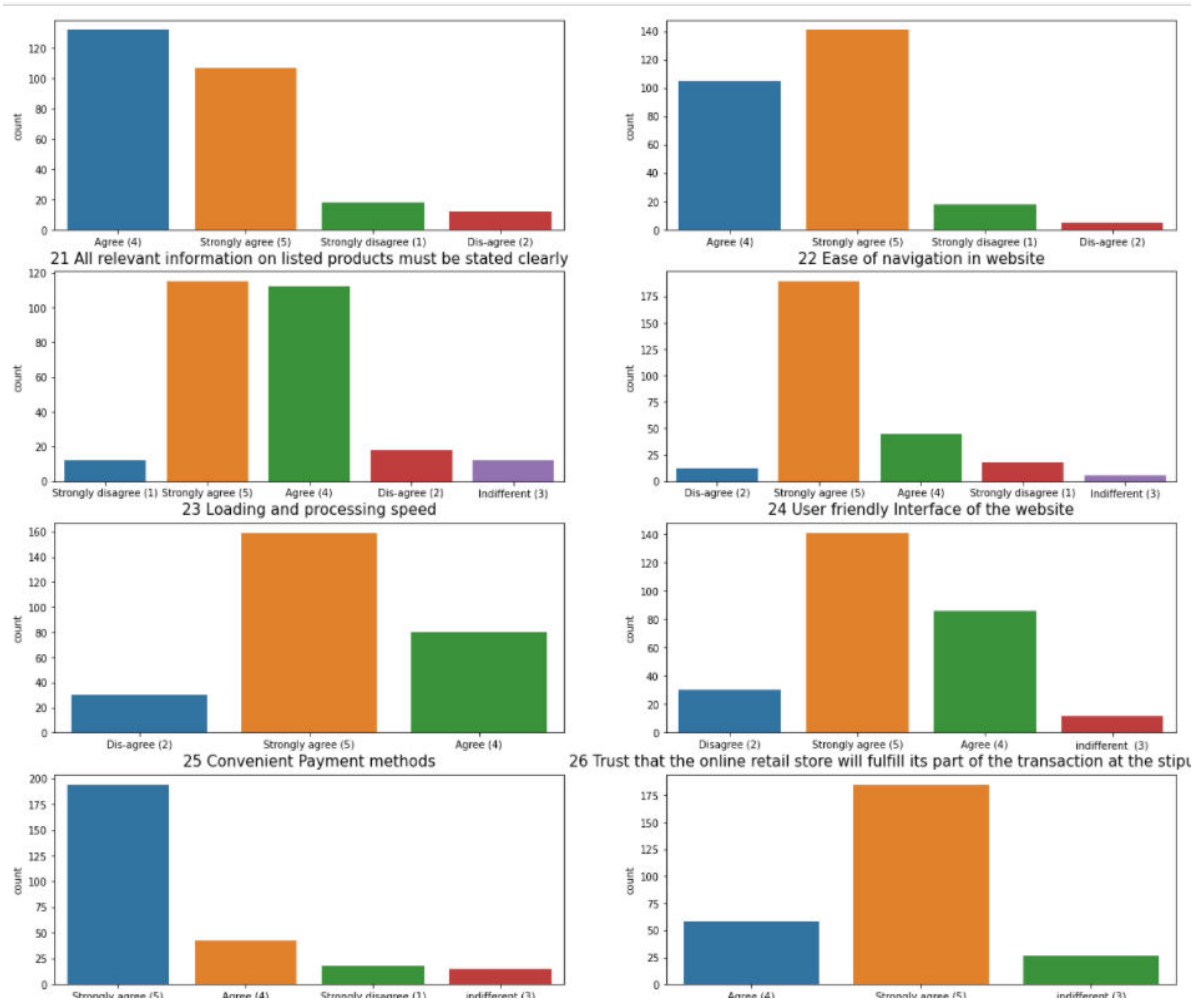
8. Preferred payment Option: -

majority of customers are doing payments through credit or debit cards

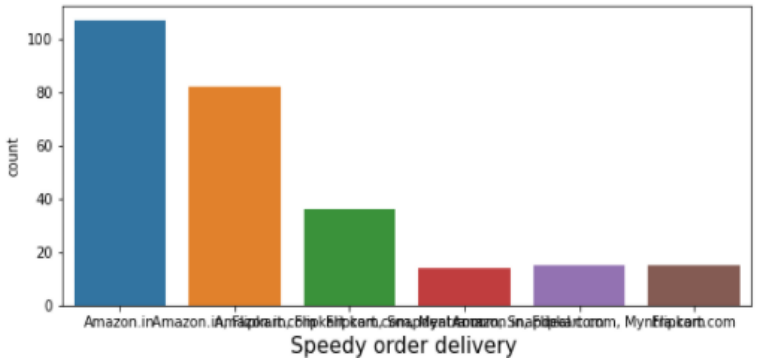
9. Majority of customers are agree/ strongly agreed for website content and highlight of similar products.



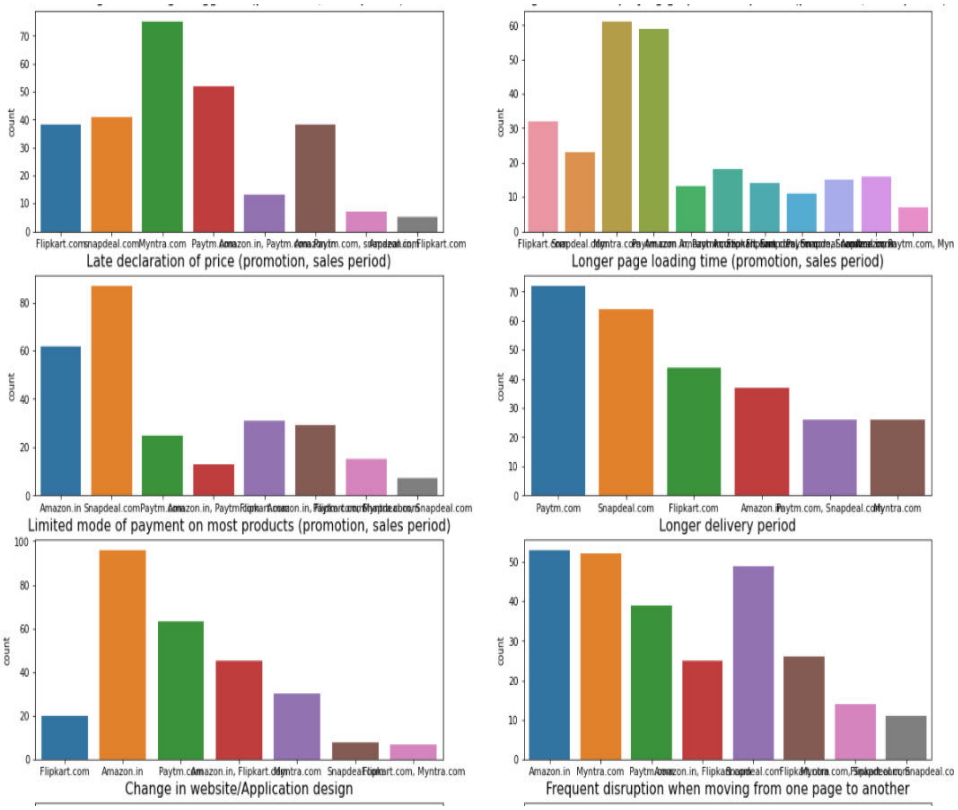
10. Very less customers are disagreeing with the valuable content regarding effective onl ine shopping and rest are either strongly agree or agree for that.



11. Speedy order delivery: -



As shown in diagram amazon.in is proving fastest delivery compare to other



- Amazon.in is having more privacy as compare to other web sites
- Amazon.in is having very secure financial services as compare to other websites
- Amazon.in is taking too much to time to log in as compare to other websites
- Myntra.com is declare very late price compare to other
- majority of customers are facing loading of pages in websites and that websites are myntra and amazon with peak of myntra.com
- Limited mode of payment on most products are having in snapdeal.com
- We can see paytm.com and snapdeal.com taking longer delivery period with peak of paytm.com
- change in design are more likely to see in amazin.in

12. Which of the Indian online retailer would you recommend to a friend?

as we can see majority of customer are likely to prefer and suggest amazon and myn tra with peak of amazon.in.

After analysing the data, we can find that the important factor which affects the onl ine shopping. Customer satisfaction one of the most important factor in any platfo rm and e-retail shoppers gives priority to it. We can see people more attracted tow ards the online shopping because of all these significant factors. The market of on line shopping will automatically increase if we improve the Quality of services and this factor.

Thank you