



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

Submitted by:

KUMAR GOURABH (kumargourabh94@gmail.com)

ACKNOWLEDGMENT

I would like to thank my mentors at Data Trained, who taught me the concepts of Data Analysis, building a machine learning model, and tuning the parameters for best outcomes.

For this particular task, I referred the following websites and articles when stuck:

- <https://towardsdatascience.com/a-common-mistake-to-avoid-when-encoding-ordinal-features-79e402796ab4>
- <https://stackoverflow.com/questions/43590489/gridsearchcv-random-forest-regressor-tuning-best-params>
- <https://www.codegrepper.com/code-examples/delphi/scikit+pca+preserve+column+names+pca+pipeline>
- <https://stackoverflow.com/questions/22984335/recovering-features-names-of-explained-variance-ratio-in-pca-with-sklearn>

I would also like to thank my mentor in Fliprobo, Muskan Vats, for providing me with the dataset and problem statement for performing this wonderful task.

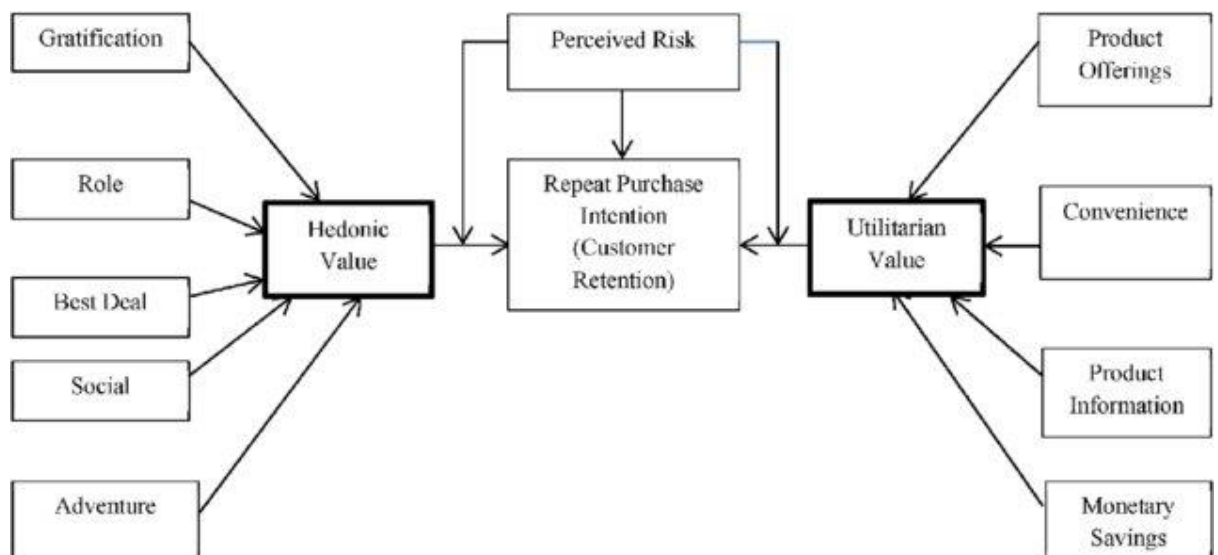
INTRODUCTION

Business Problem Framing

Data Scientists have to apply their analytical skills to give findings and conclusions in detailed data analysis written in jupyter notebook

Conceptual Background of the Domain Problem

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.



Analytical Problem Framing

I will be going about this task in 2 parts:

- First Part: Descriptive Statistics &
- Second Part: Inferential Statistics

Descriptive statistics describes data (for example, a chart or graph) and **inferential statistics** allows us to make predictions ("inferences") from that data. With inferential statistics, we take data from samples and make generalizations about a population.

For example, we might stand in a mall and ask a sample of 100 people if they like shopping at Sears. We could make a bar chart of yes or no answers (that would be descriptive statistics) or we could use our research (and inferential statistics) to reason that around 75-80% of the population (all shoppers in all malls) like shopping at Sears.¹

Data Sources and their formats

There are two sheets (one is detailed) and second is encoded in the excel file. The data is in the form of a Questionnaire - probably Google Forms.

Shape – 269 Rows; 71 Columns

First 5 Rows:

	1 Gender 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?	10 What is the operating system (OS) of your device?	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?	13 After first visit, how do you reach the online retail store?	14 How much time do you explore the e-retail store before making a purchase decision?
0	Male	31-40 years	Delhi	110009	Above 4 years	31-40 times	Dial-up	Desktop	Others	Window/windows Mobile	Google chrome	Search Engine	Search Engine	6-10 mins
1	Female	21-30 years	Delhi	110030	Above 4 years	41 times and above	Wi-Fi	Smartphone	4.7 inches	IOS/Mac	Google chrome	Search Engine	Via application	more than 15 mins
2	Female	21-30 years	Greater Noida	201308	3-4 years	41 times and above	Mobile Internet	Smartphone	5.5 inches	Android	Google chrome	Search Engine	Via application	11-15 mins
3	Male	21-30 years	Karnal	132001	3-4 years	Less than 10 times	Mobile Internet	Smartphone	5.5 inches	IOS/Mac	Safari	Search Engine	Search Engine	6-10 mins
4	Female	21-30 years	Bangalore	530068	2-3 years	11-20 times	Wi-Fi	Smartphone	4.7 inches	IOS/Mac	Safari	Content Marketing	Via application	more than 15 mins

¹<https://www.statisticshowto.com/probability-and-statistics/statistics-definitions/inferential-statistics/>

15 What is your preferred payment Option? t 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 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
E-wallets (Paytm, Freecharge etc.)	Sometimes	Lack of trust	Agree (4)	Indifferent (3)	Indifferent (3)	Agree (4)	Agree (4)	Strongly disagree (1)	Dis-agree (2)	Dis-agree (2)	Disagree (2)	Strongly agree (5)
Credit/Debit cards	Very frequently	Promo code not applicable	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)
E-wallets (Paytm, Freecharge etc.)	Sometimes	Promo code not applicable	Strongly agree (5)	Agree (4)	Agree (4)	Agree (4)	Agree (4)	Agree (4)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)
Credit/Debit cards	Never	Better alternative offer	Agree (4)	Agree (4)	Indifferent (3)	Agree (4)	Strongly agree (5)	Agree (4)	Strongly agree (5)	Strongly agree (5)	Agree (4)	Strongly agree (5)
Credit/Debit cards	Frequently	Better alternative offer	Strongly agree (5)	Indifferent (3)	Indifferent (3)	Agree (4)	Agree (4)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Agree (4)	Strongly agree (5)
28 Being able to guarantee the privacy of the customer	29 Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.)	30 Online shopping gives monetary benefit and discounts	31 Enjoyment is derived from shopping online	32 Shopping online is convenient and flexible	33 Return and replacement policy of the e-tailer is important for purchase decision	34 Gaining access to loyalty programs is a benefit of shopping online	35 Displaying quality Information on the website improves satisfaction of customers	36 User derive satisfaction while shopping on a good quality website or application	37 Net Benefit derived from shopping online can lead to users satisfaction	38 User satisfaction cannot exist without trust	39 Offering a wide variety of listed product in several category	
Agree (4)	Agree (4)	indifferent (3)	Strongly disagree (1)	Dis-agree (2)	Dis-agree (2)	Agree (4)	indifferent (3)	Strongly agree (5)	Agree (4)	Strongly agree (5)	indifferent (3)	
Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	
Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Agree (4)	Strongly agree (5)	
Strongly agree (5)	Agree (4)	Agree (4)	indifferent (3)	indifferent (3)	Agree (4)	indifferent (3)	indifferent (3)	Agree (4)	indifferent (3)	Agree (4)	Agree (4)	
Agree (4)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Dis-agree (2)	Agree (4)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Agree (4)	

40 Provision of complete and relevant product information	41 Monetary savings	42 The Convenience of patronizing the online retailer	43 Shopping on the website gives you the sense of adventure	44 Shopping on your preferred e-tailer enhances your social status	45 You feel gratification shopping on your favorite e-tailer	46 Shopping on the website helps you fulfill certain roles	47 Getting value for money spent	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
indifferent (3)	Disagree (2)	indifferent (3)	Agree (4)	Agree (4)	Strongly agree (5)	Agree (4)	Strongly agree (5)	Amazon.in, Paytm.com	Paytm.com	Flipkart.com	Flipkart.com	Snapdeal.com
Strongly agree (5)	Strongly agree (5)	indifferent (3)	indifferent (3)	indifferent (3)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com	Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com	Amazon.in, Myntra.com	Flipkart.com, Myntra.com	Amazon.in, Flipkart.com, Myntra.com
Strongly agree (5)	Strongly agree (5)	Agree (4)	Agree (4)	Agree (4)	indifferent (3)	indifferent (3)	Agree (4)	Amazon.in, Paytm.com, Myntra.com	Amazon.in, Paytm.com, Myntra.com	Amazon.in, Paytm.com, Myntra.com	Amazon.in, Myntra.com	Amazon.in, Paytm.com, Myntra.com
Agree (4)	Strongly agree (5)	Agree (4)	Agree (4)	Strongly agree (5)	Agree (4)	indifferent (3)	Agree (4)	Amazon.in, Flipkart.com, Paytm.com, Myntra.com...	Amazon.in, Flipkart.com, Paytm.com, Myntra.com...	Amazon.in, Flipkart.com, Paytm.com, Myntra.com...	Amazon.in, Flipkart.com	Amazon.in, Flipkart.com
Agree (4)	Strongly agree (5)	Agree (4)	indifferent (3)	Strongly disagree (1)	Strongly agree (5)	Strongly disagree (1)	Strongly agree (5)	Amazon.in, Flipkart.com, Paytm.com, Myntra.com...	Amazon.in, Flipkart.com, Paytm.com, Myntra.com...	Myntra.com	Myntra.com	Amazon.in, Flipkart.com, Paytm.com, Myntra.com...
Fast loading website speed of website and application	Reliability of the website or application	Quickness to complete purchase	Availability of several payment options	Speedy order delivery	Privacy of customers' information	Security of customer financial information	Perceived Trustworthiness	Presence of online assistance through multi-channel	Longer time to get logged in (promotion, sales period)	Longer time in displaying graphics and photos (promotion, sales period)		
napdeal.com	Paytm.com	Paytm.com	Patym.com	Amazon.in	Amazon.in	Amazon.in	Flipkart.com	Paytm.com	Amazon.in	Amazon.in		
Amazon.in, Flipkart.com, Myntra.com	Myntra.com	Amazon.com, Flipkart.com, Myntra.com	Amazon.in, Flipkart.com, Myntra.com	Amazon.in, Flipkart.com	Myntra.com	Myntra.com	Myntra.com	Amazon.in, Flipkart.com, Myntra.com	Amazon.in, Flipkart.com	Myntra.com		
Amazon.in, Paytm.com	Amazon.in, Paytm.com, Myntra.com	Amazon.com, Paytm.com, Myntra.com	Patym.com, Myntra.com	Amazon.in	Amazon.in	Amazon.in, Paytm.com, Myntra.com	Amazon.in, Myntra.com	Myntra.com	Myntra.com	Myntra.com		
Amazon.in, Flipkart.com, napdeal.com	Amazon.in, Flipkart.com, Paytm.com	Amazon.com, Flipkart.com, Paytm.com	Amazon.in, Flipkart.com, Myntra.com	Amazon.in, Flipkart.com, Snapdeal.com	Amazon.in, Flipkart.com, Myntra.com	Amazon.in, Flipkart.com, Snapdeal.com	Amazon.in, Flipkart.com, Snapdeal.com	Amazon.in, Flipkart.com, Myntra.com, Snapdeal	Snapdeal.com	Myntra.com, Snapdeal.com		
Amazon.in	Amazon.in, Paytm.com, Myntra.com	Amazon.com, Flipkart.com, Paytm.com, Myntra.co...	Amazon.in, Flipkart.com, Patym.com, Myntra.com...	Amazon.in	Amazon.in, Paytm.com	Paytm.com	Amazon.in, Myntra.com	Amazon.in, Myntra.com	Flipkart.com, Paytm.com	Paytm.com		

Late declaration of price (promotion, sales period)	Longer page loading time (promotion, sales period)	Limited mode of payment on most products (promotion, sales period)	Longer delivery period	Change in website/Application design	Frequent disruption when moving from one page to another	Website is as efficient as before	Which of the Indian online retailer would you recommend to a friend?
Flipkart.com	Flipkart.com	Amazon.in	Paytm.com	Flipkart.com	Amazon.in	Amazon.in	Flipkart.com
snapdeal.com	Snapdeal.com	Snapdeal.com	Snapdeal.com	Amazon.in	Myntra.com	Amazon.in, Flipkart.com	Amazon.in, Myntra.com
Myntra.com	Myntra.com	Amazon.in	Paytm.com	Paytm.com	Paytm.com	Amazon.in	Amazon.in, Paytm.com, Myntra.com
Myntra.com	Paytm.com	Paytm.com	Paytm.com	Amazon.in, Flipkart.com	Amazon.in, Flipkart.com	Amazon.in, Flipkart.com, Paytm.com	Amazon.in, Flipkart.com
Paytm.com	Paytm.com	Snapdeal.com	Paytm.com	Amazon.in	Snapdeal.com	Paytm.com	Amazon.in, Myntra.com

Checking for Null Values in the dataset

```
# Identifying null values
sns.heatmap(df.isnull())
plt.axis('off')
plt.show()
```



The above heatmap shows there are no null values in the collected dataset.

1.Descriptive

As there are a huge number of features, it makes sense to divide the data into these categories:

1. categorical_columns = This will contain demographic features and other person specific features like age, gender, browser]
2. rating_columns = This will contain all questions which have answers in the form of (Agree, Strongly Agree, Indifferent etc.)
3. ecommerce_columns = This will contain all questions which have answers in the form of (amazon, flipkart, myntra etc.)

```
# categorical_columns can be located from column index 0 to 16
categorical_columns=df.iloc[:,17].columns
categorical_columns=categorical_columns.to_list()

# rating_columns can be located from column index 17 to 46
rating_columns=df.iloc[:,17:47].columns
rating_columns=rating_columns.to_list()

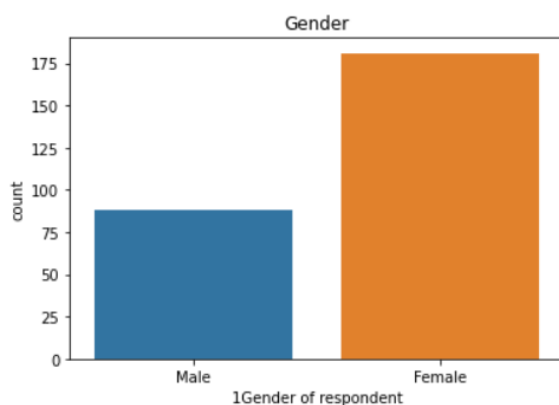
# ecommerce_columns can be located from column index 47 and above
ecommerce_columns=df.iloc[:,47:].columns
ecommerce_columns=ecommerce_columns.to_list()
```

Let's have a look at the overall demographics of the collected sample

```
print(df['1Gender of respondent'].value_counts())
sns.countplot(x=df['1Gender of respondent'])
plt.title('Gender')
```

```
Female    181
Male      88
Name: 1Gender of respondent, dtype: int64

Text(0.5, 1.0, 'Gender')
```

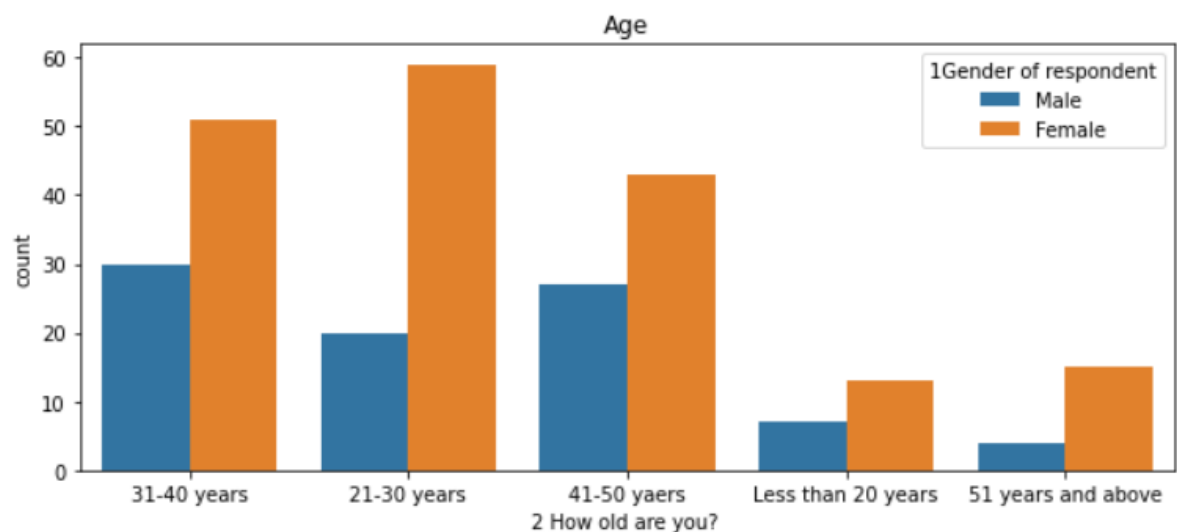


The Results of the analysis has around 50% more Female participants as compared to Males


```
print(df['2 How old are you? '].value_counts())
plt.figure(figsize=(10,4))
sns.countplot(x=df['2 How old are you? '],hue = df['1Gender of respondent'])
plt.title('Age')
```

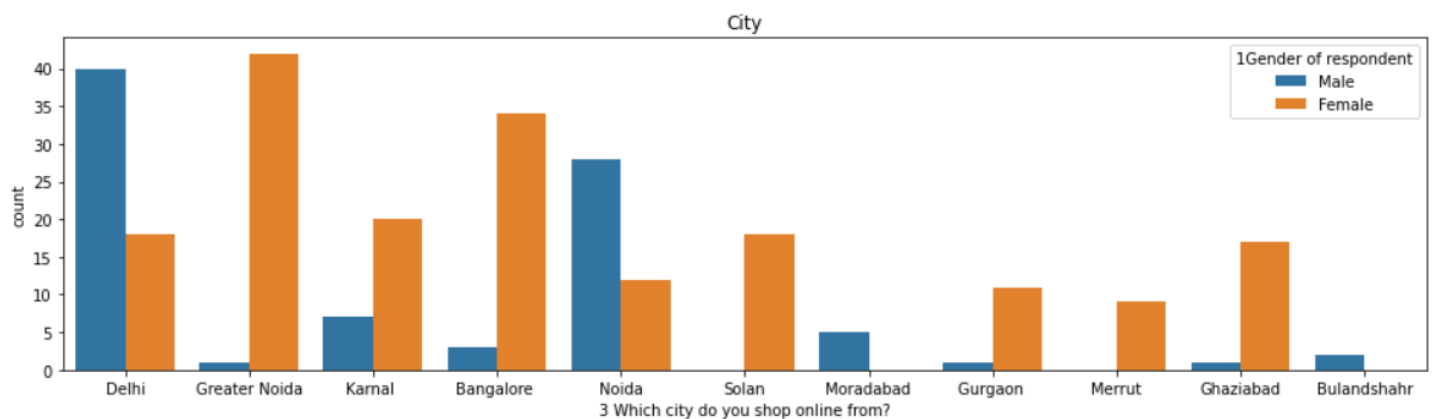
```
31-40 years      81
21-30 years      79
41-50 yaers      70
Less than 20 years  20
51 years and above  19
Name: 2 How old are you? , dtype: int64
```

```
Text(0.5, 1.0, 'Age')
```



Spread Based on City

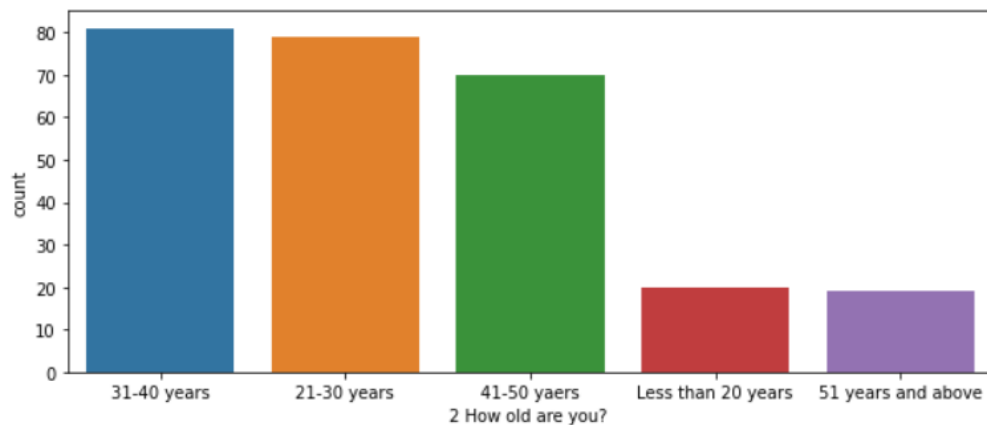
```
Text(0.5, 1.0, 'City')
```



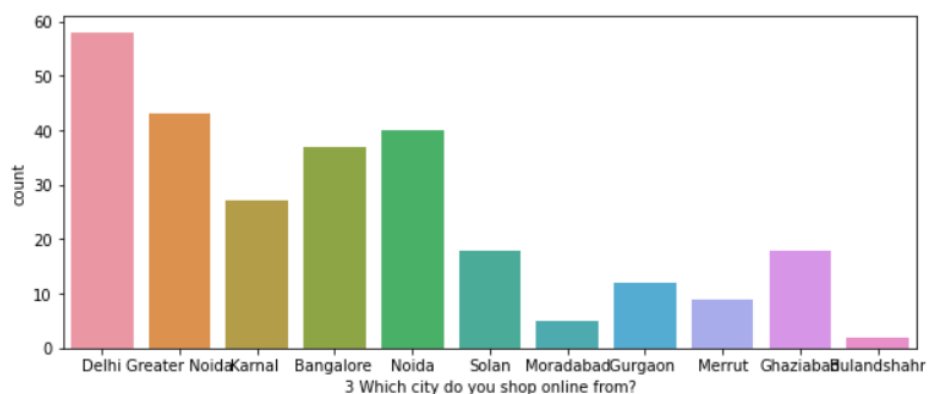
Univariate Analysis

1. Age

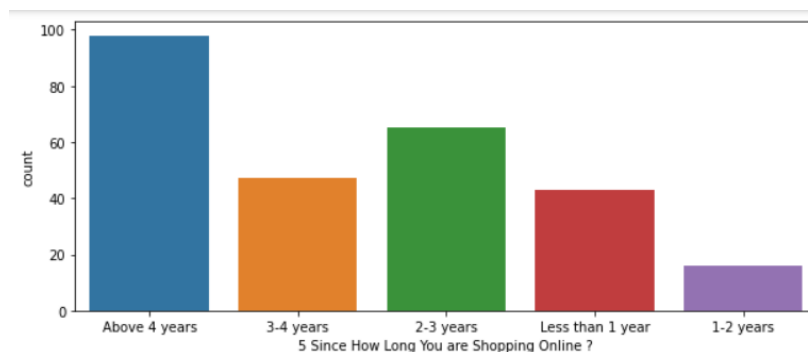
```
31-40 years      81
21-30 years      79
41-50 yaers      70
Less than 20 years  20
51 years and above  19
Name: 2 How old are you? , dtype: int64
```



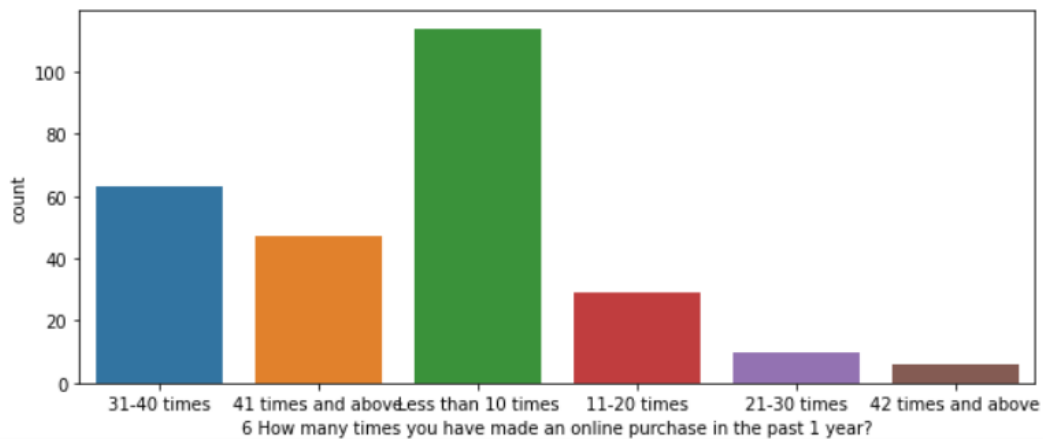
2. City



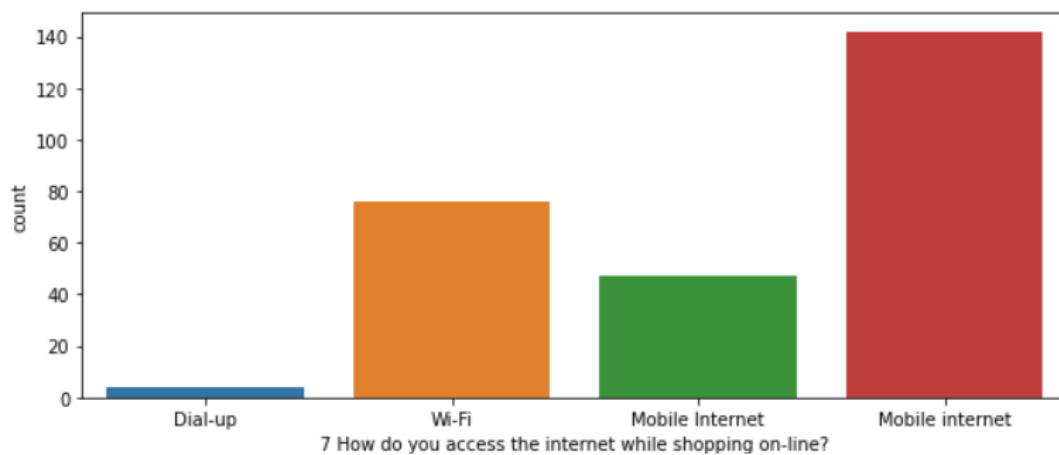
3. Time since online shopping



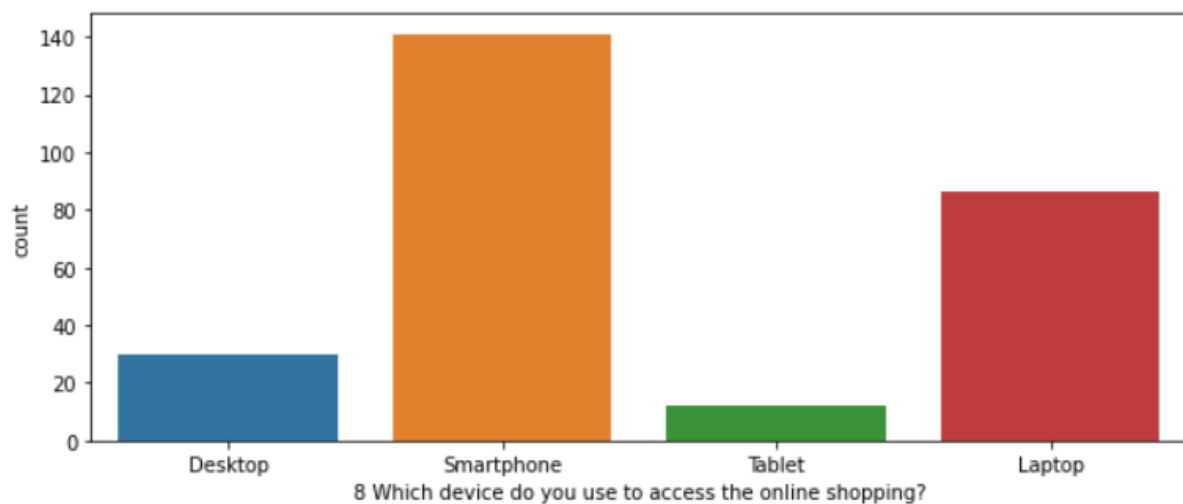
4. How many times you have made an online purchase in the past 1 year?



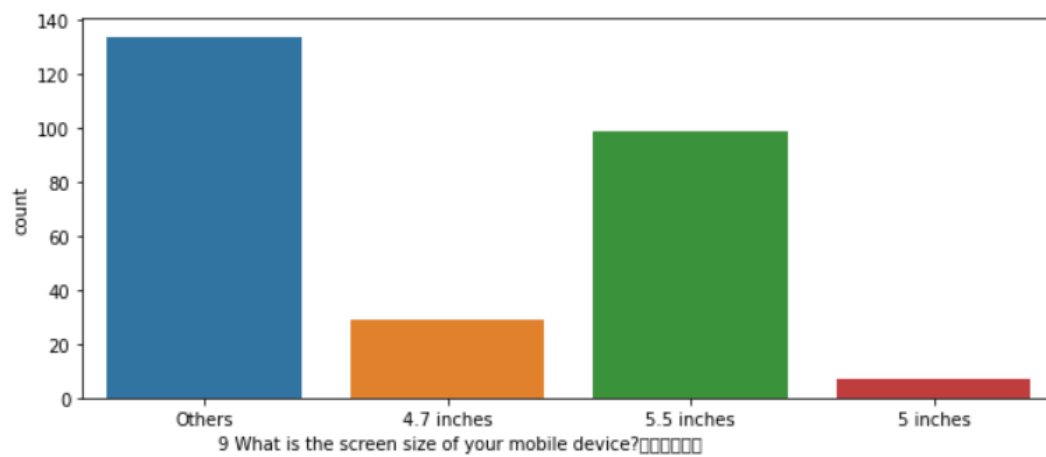
5. How do you access the internet while shopping on-line?



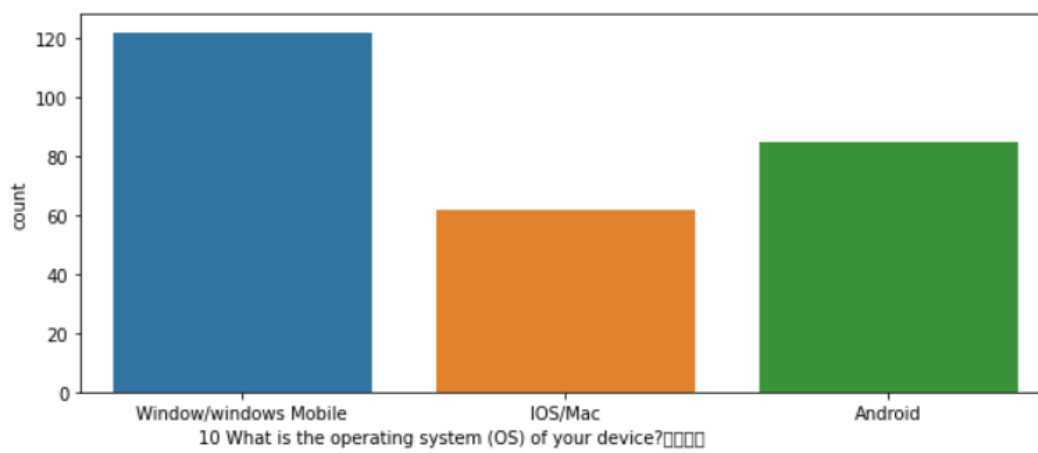
6. Which device do you use to access the online shopping?



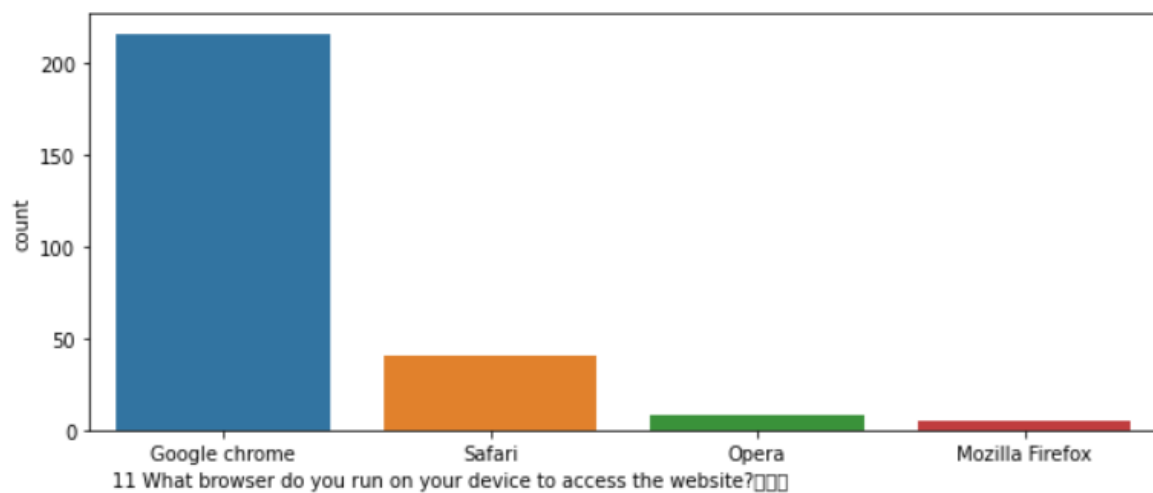
7. What is the screen size of your mobile device?



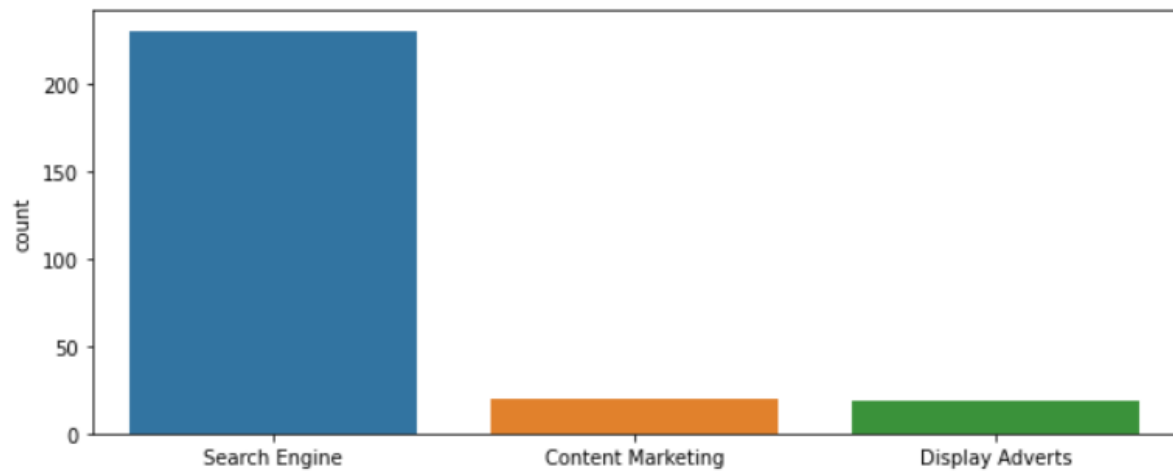
8. What is the operating system (OS) of your device?



9. What browser do you run on your device to access the website?

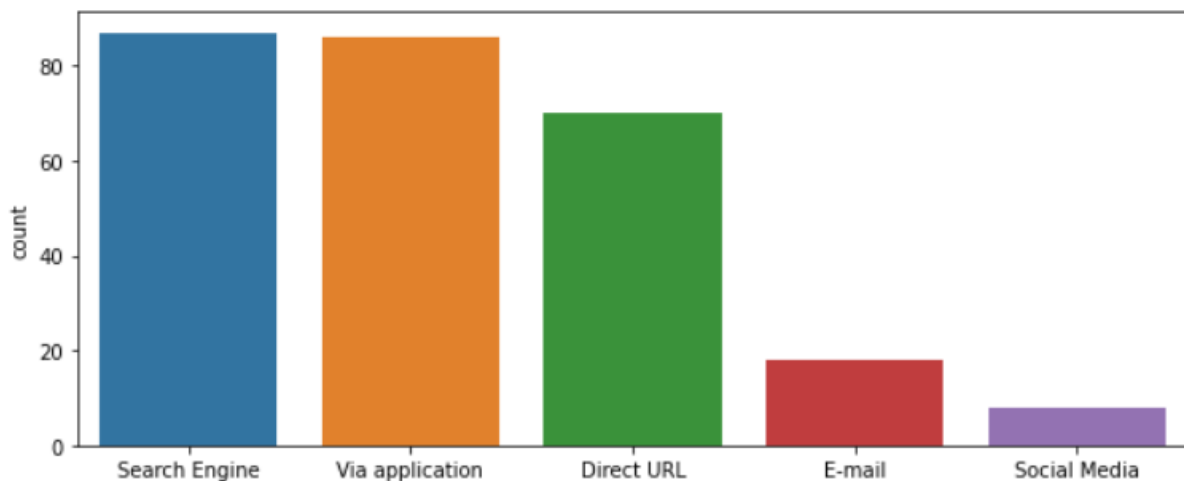


10. Which channel did you follow to arrive at your favorite online store for the first time?



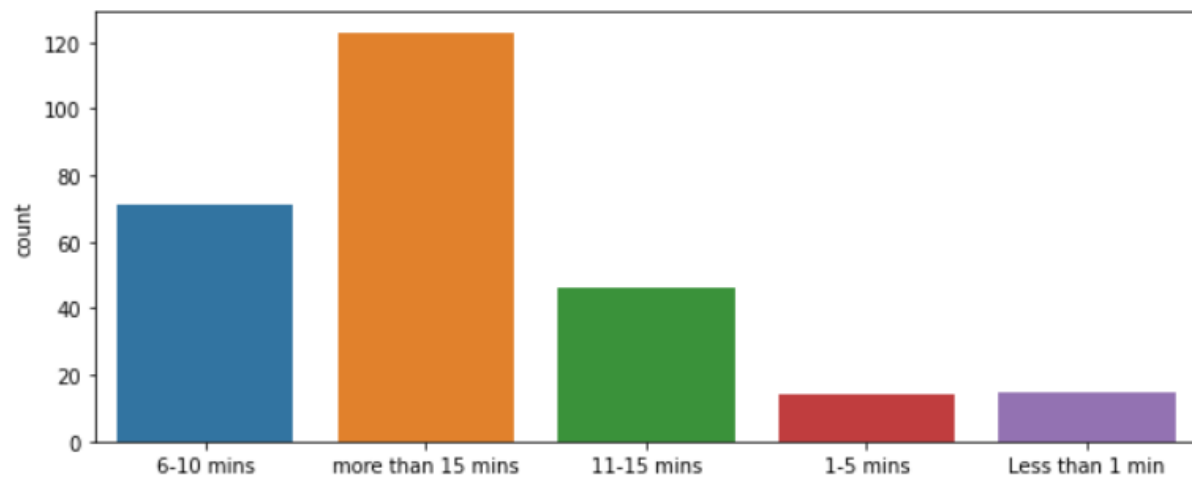
12 Which channel did you follow to arrive at your favorite online store for the first time?

11. After first visit, how do you reach the online retail store?



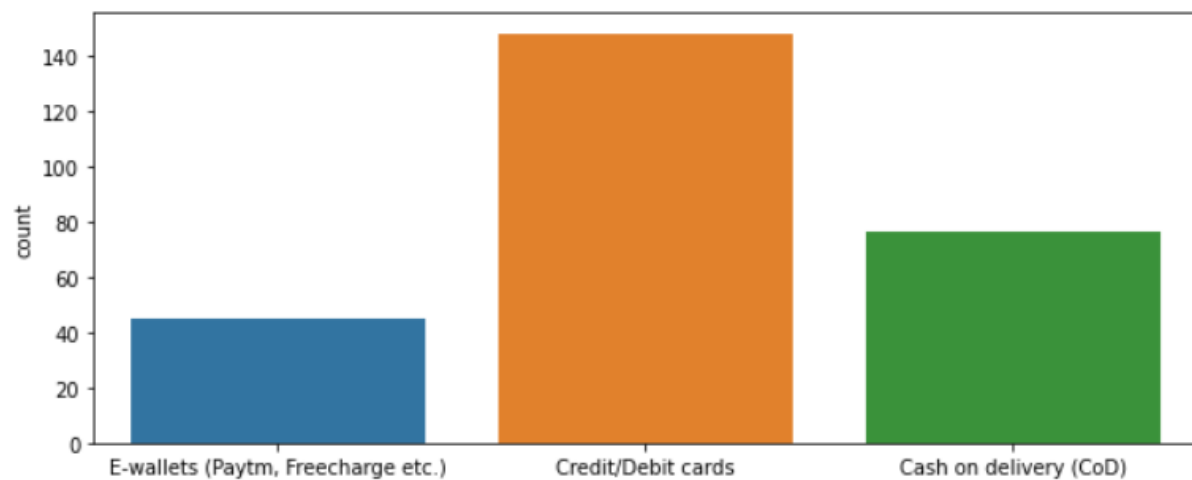
13 After first visit, how do you reach the online retail store?

12. How much time do you explore the e- retail store before making a purchase decision?



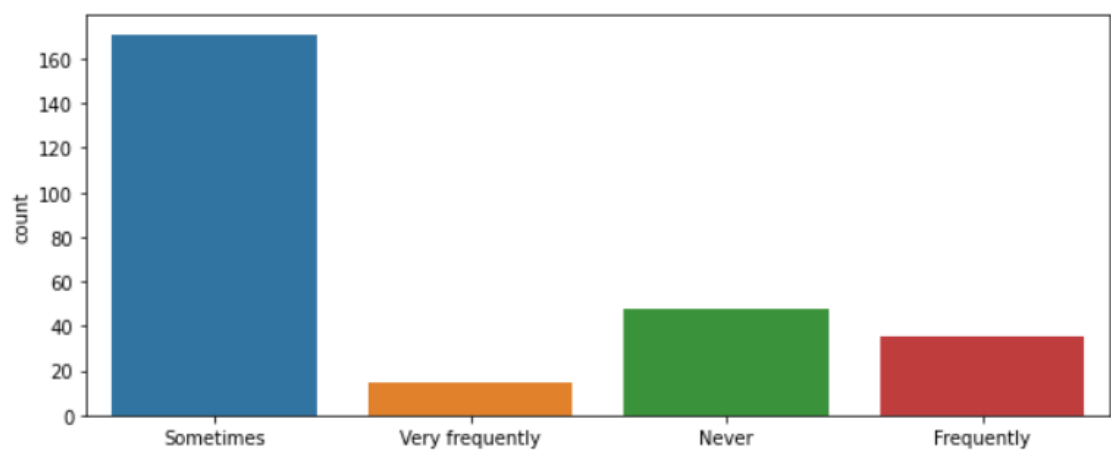
14 How much time do you explore the e- retail store before making a purchase decision?

13. What is your preferred payment Option?



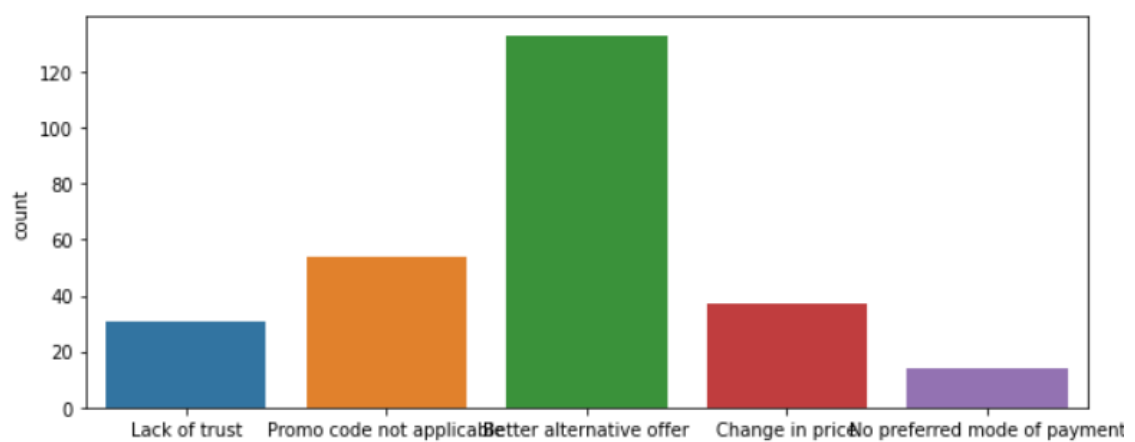
15 What is your preferred payment Option?□□□□□

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



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

15. Why did you abandon the "Bag", "Shopping Cart"?



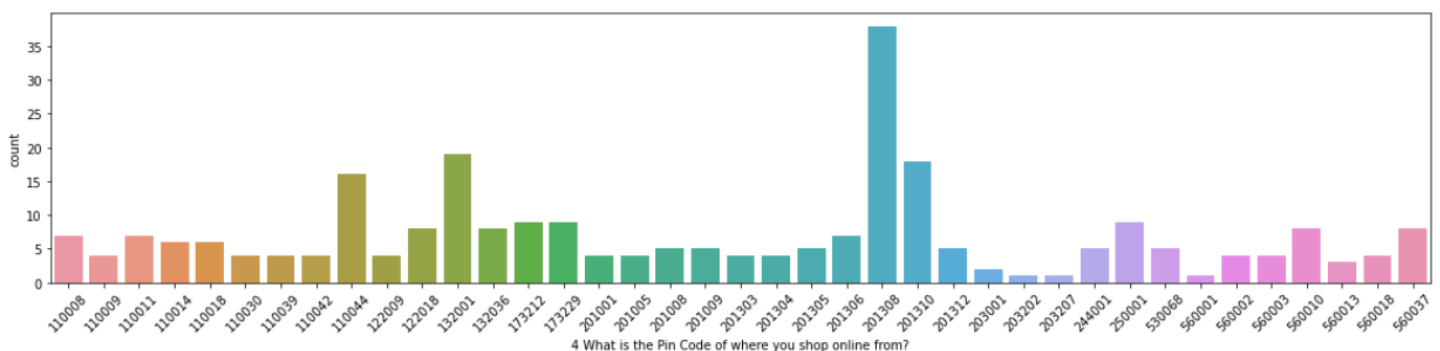
Why did you abandon the "Bag", "Shopping Cart"?□□□□□

Some Important Conclusions Based on the initial Questions

1. Age: 85% of the people are within age group 20-50 years. Very small percentage of people are less than 20 or more than 50 years old.
2. Delhi, Greater Noida, Noida and Bangalore have the maximum participants.
3. Maximum number of participants have been using Online Shopping platforms for more than 4 years.
4. Purchases in last 1 year: Maximum People have made purchases less than 10 times. Next in line are 30-40 and 40+
5. Most people use Mobile internet during shopping on-line.
6. Very few people use desktops or tablets; while most are using Smartphones and laptops.
7. A large percentage of people use Google Chrome.
8. Search Engines are the most used channel which guide people to their favourite online store.
9. Most people take purchase decision after spending more than 15 minutes.
10. Credit/Debit cards are the most used mode of payment.
11. Maximum people abandon the items in cart 'sometimes' as opposed to never or frequently.
12. The biggest reason to abandon is 'Better alternative offer'.

We'll Now have a look at how the above Features are varying based on the Location(Pin-Code)

```
#Let's see the counts for pin codes
plt.figure(figsize=(20,4))
sns.countplot(x=df['4 What is the Pin Code of where you shop online from?'])
plt.xticks(rotation=45)
plt.show()
```



Pin Code 201308: Has the Most Number of buyers, closely Followed by 132001, 201310 and 110044.

201308 and 201310: These belong to Gr. Noida, Gautham Budha Nagar. As this area consists of the Knowledge Park, where a lot of Students are available, this may have led to the high count.

132001 - Karnal, Haryana

110044 - South Delhi, Badarpur Area

Let's now have a look at the Rating based parameters

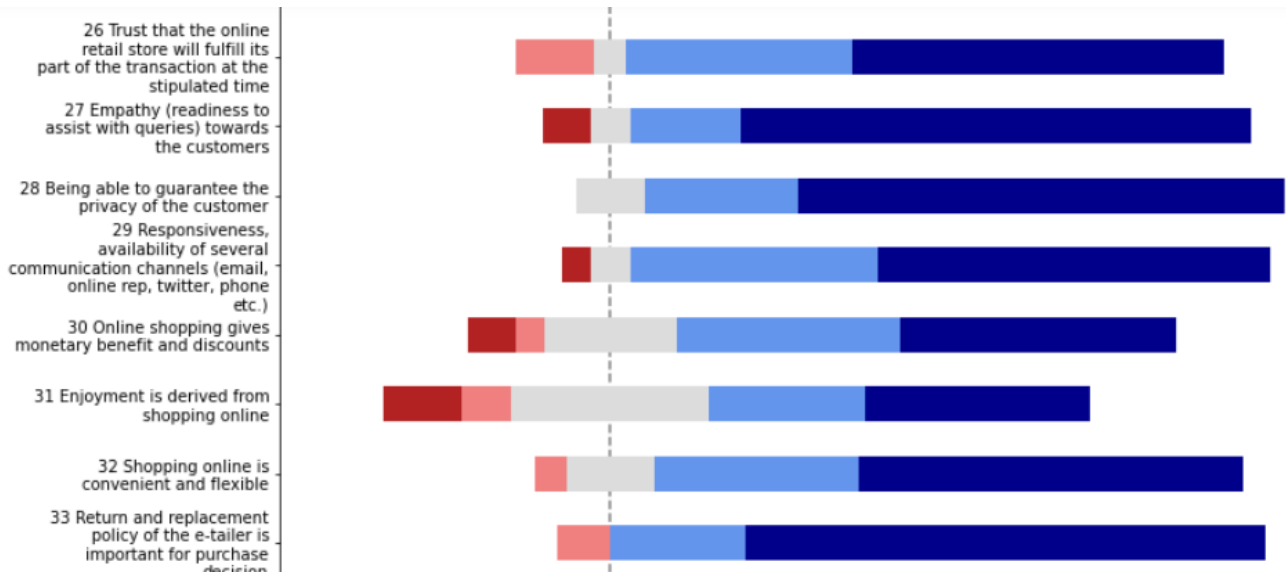
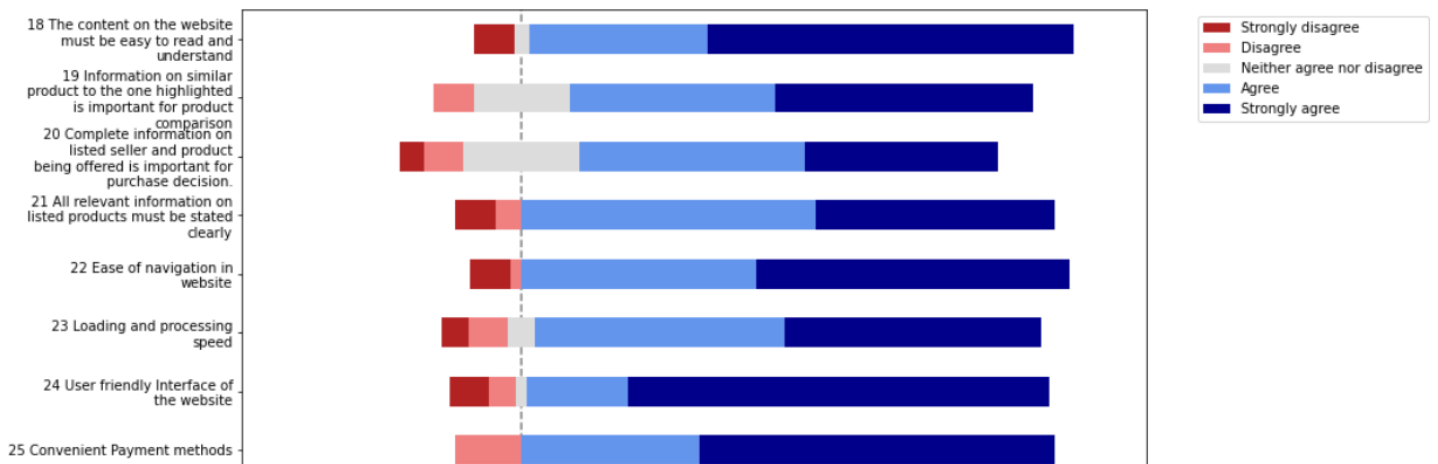
```
#It makes more sense to use the coded data for this part
# Loading Dataset
df_coded = pd.read_csv('customer_retention_dataset_coded.csv')
#pip install plot-likert
dict = {1: 'Strongly disagree', 2: 'Disagree', 3: 'Neither agree nor disagree', 4: 'Agree', 5: 'Strongly agree'}

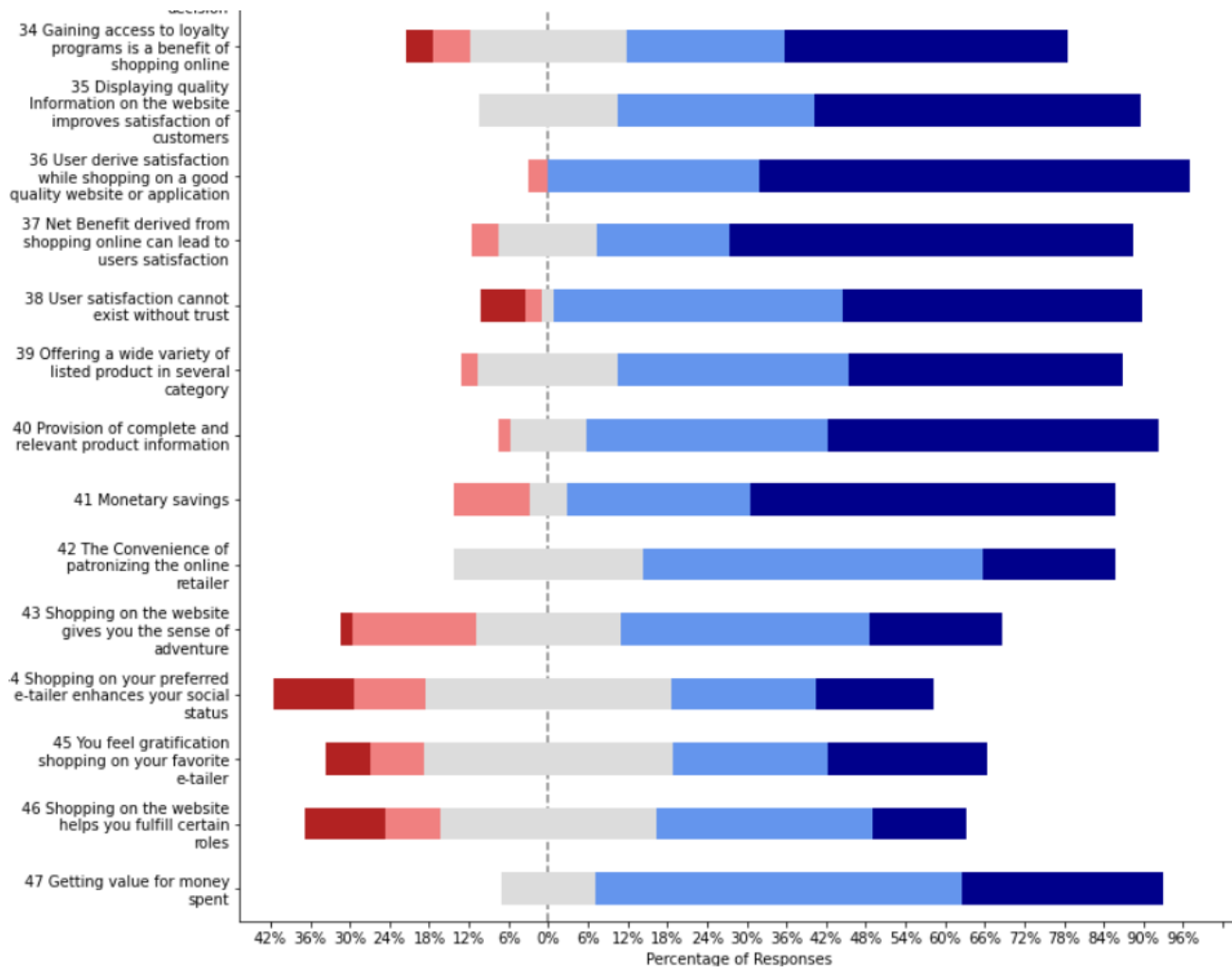
df_likert=df_coded.copy()
# Remap the values of the dataframe
for col in rating_columns:
    df_likert.replace({col: dict},inplace=True)

import plot_likert

plot_likert.plot_likert(df_likert[rating_columns], plot_likert.scales.agree, plot_percentage=True,figsize=(12,24))

<AxesSubplot:xlabel='Percentage of Responses'>
```





Important Observations on the basis of Customer's perception

1. A large majority of people believe that content on the website must be easy to read and understand.
2. People agree that Information on similar product to the one highlighted is important.
3. 70% people believe that Complete information on listed seller and product is important for purchase decision.
4. 90% people believe that All relevant information on listed products must be stated clearly
5. For more than 90% people, ease of navigation of website, loading or processing speed as well as User friendly Interface of the website is important.
6. Other important parameters are - Convenient Payment Methods, privacy of customers, responsiveness, availability of several Communication channels, return and replacement policy.

Satisfaction related Features

1. Around 20% people don't that Displaying quality Information on the website improves satisfaction of customers. The other 80% agree that it is important.
2. Almost all participants believe that User derive satisfaction while shopping on a good quality website or application.
3. Very high percentage of people believe that Net Benefit derived from shopping online can lead to users satisfaction.
4. Most People believe that User satisfaction cannot exist without trust.

Now, we'll see the company name specific parameters

Questions and their responses in frequencies (Not displaying all graphs, Displaying Frequency Counts below)

1.

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 shopped from;

2.

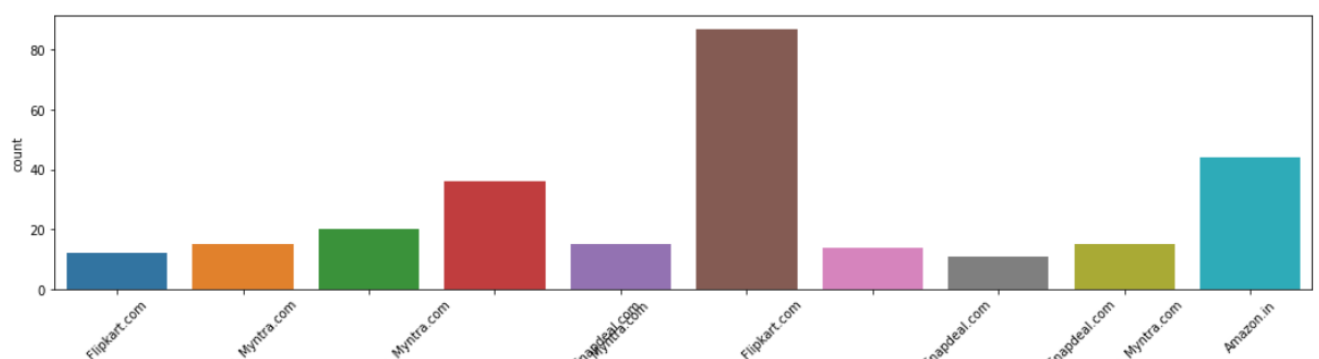
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com	64
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com	44
Amazon.in, Flipkart.com	44
Amazon.in	29
Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com	22
Amazon.in, Paytm.com, Myntra.com	20
Amazon.in, Flipkart.com, Myntra.com	19
Paytm.com	12
Flipkart.com	8
Amazon.in, Paytm.com	7

Name: Easy to use website or application, dtype: int64

3.

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
Amazon.in, Myntra.com	15
Flipkart.com, Myntra.com	15
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



4.

Amazon.in, Flipkart.com	130
Amazon.in	43
Amazon.in, Myntra.com	20
Flipkart.com, Myntra.com	15
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

5.

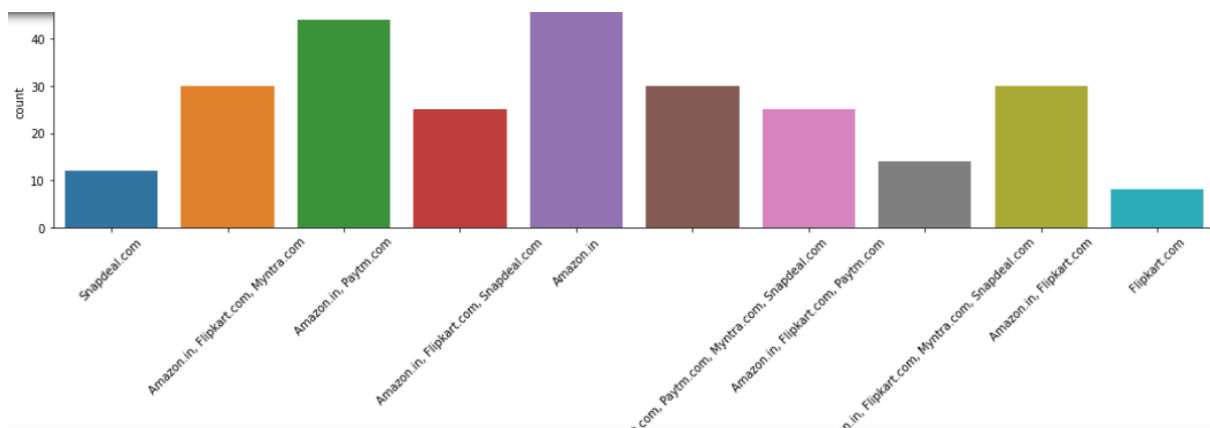
Amazon.in, Flipkart.com	100
Amazon.in	43
Amazon.in, Flipkart.com, Paytm.com	24
Amazon.in, Paytm.com, Myntra.com	20
Amazon.in, Flipkart.com, Myntra.com	15
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com	15
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com	14
Snapdeal.com	12
Flipkart.com, Snapdeal.com	11
Flipkart.com	8
Amazon.in, Flipkart.com, Snapdeal.com	7

Name: Complete, relevant description information of products, dtype

6.

Amazon.in	51
Amazon.in, Paytm.com	44
Amazon.in, Flipkart.com, Myntra.com	30
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.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:



7.

Amazon.in	61
Amazon.in, Flipkart.com	50
Amazon.in, Flipkart.com, Paytm.com	36
Amazon.in, Paytm.com, Myntra.com	35
Amazon.in, Flipkart.com, Snapdeal.com	18
Flipkart.com	15
Myntra.com	15
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com	14
Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com	13
Paytm.com	12
Name: Reliability of the website or application, dtype:	

8.

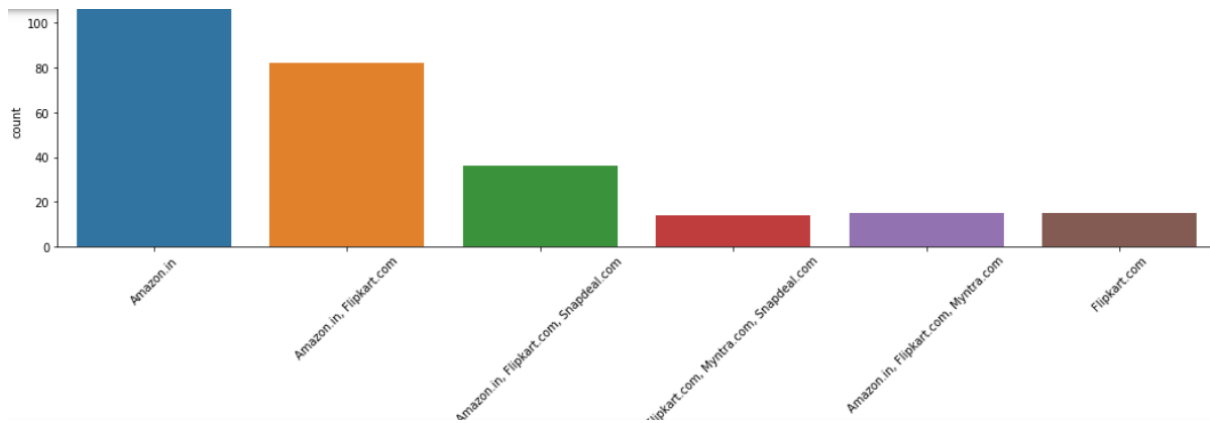
Amazon.com	66
Amazon.com, Flipkart.com, Paytm.com	47
Amazon.com, Flipkart.com	37
Amazon.com, Flipkart.com, Myntra.com	30
Paytm.com	25
Amazon.com, Paytm.com, Myntra.com	20
Flipkart.com	15
Amazon.com, Flipkart.com, Paytm.com, Myntra.com, Snapdeal	15
Flipkart.com, Myntra.com, Snapdeal	14
Name: Quickness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

9.

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	

10.

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



11.

```
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
Flipkart.com 15
Amazon.in, Paytm.com 15
Myntra.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
```

12.

```
Amazon.in 51
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com 42
Flipkart.com 33
Amazon.in, Flipkart.com, Snapdeal.com 25
Amazon.in, Flipkart.com 24
Amazon.in, Paytm.com, Myntra.com 20
Amazon.in, Snapdeal.com 19
Paytm.com 15
Myntra.com 15
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com 14
Amazon.in, Flipkart.com, Paytm.com 11
Name: Security of customer financial information, dtype: int64
```

13.

```
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
```

14.

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	

15.

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)	

16.

Amazon.in, Flipkart.com	60
Amazon.in	39
Myntra.com	35
Snapdeal.com	34
Myntra.com, Snapdeal.com	25
Flipkart.com, Snapdeal.com	19
Flipkart.com	15
Paytm.com	15
Amazon.in, Myntra.com, Snapdeal.com	14
Amazon.in, Paytm.com	13
Name: Longer time in displaying graphics and photos (promotion, sales period)	

17.

Myntra.com	75
Paytm.com	52
snapdeal.com	41
Flipkart.com	38
Amazon.in	38
Amazon.in, Paytm.com	13
Paytm.com, snapdeal.com	7
Amazon.in, Flipkart.com	5
Name: Late declaration of price (promotion, sales period)	

18.

Myntra.com	61
Paytm.com	59
Flipkart.com	32
Snapdeal.com	23
Amazon.in, Flipkart.com	18
Amazon.in	16
Paytm.com, Snapdeal.com	15
Amazon.in, Snapdeal.com	14
Amazon.in, Paytm.com	13
Flipkart.com, Snapdeal.com	11
Amazon.in, Paytm.com, Myntra.com	7
Name: Longer page loading time (promotion, sales period)	

19.

Snapdeal.com	87
Amazon.in	62
Flipkart.com	31
Amazon.in, Flipkart.com	29
Paytm.com	25
Paytm.com, Snapdeal.com	15
Amazon.in, Paytm.com	13
Myntra.com, Snapdeal.com	7
Name: Limited mode of payment on most products (promotion, sales period)	

20.

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

21.

Amazon.in	53
Myntra.com	52
Snapdeal.com	49
Paytm.com	39
Flipkart.com	26
Amazon.in, Flipkart.com	25
Myntra.com, Snapdeal.com	14
Flipkart.com, Snapdeal.com	11
Name: Frequent disruption when moving from one page to another	

22.

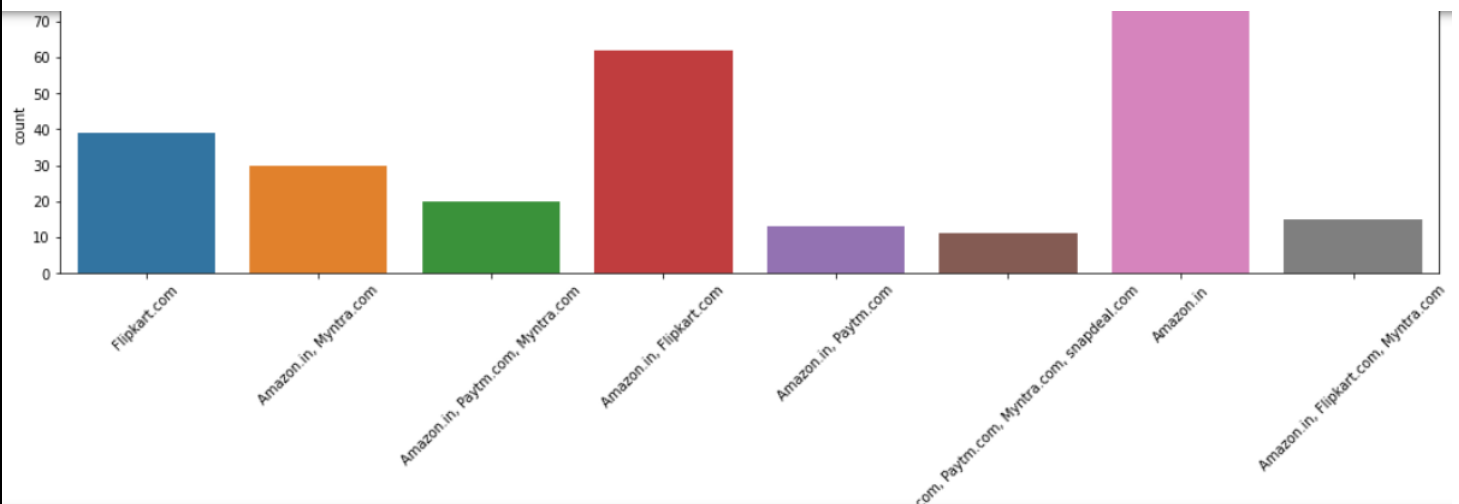
Amazon.in	94
Flipkart.com	47
Amazon.in, Flipkart.com	45
Amazon.in, Flipkart.com, Paytm.com	25
Amazon.in, Paytm.com	18
Paytm.com	15
Myntra.com, Snapdeal.com	14
Snapdeal.com	11

Name: Website is as efficient as before,

23.

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?



Observations from Multiple Options based Questions

1. Maximum people have shopped from these 5 companies - Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com.
2. Most people find Easy to use website or application are - Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com Also, Amazon.com and Flipkart.com are the major choices.
3. In terms of Visual appealing web-page layout also, Amazon.com and Flipkart.com seem to take the lead.
4. Talking about Wide variety of product on offer: Amazon.com and Flipkart.com are the major choices.

5. Complete, relevant description information of products: Once again, maximum people have chosen to go with mazon.com and Flipkart.com
6. Fast loading website speed of website and application: Amazon seems to take the lead in this category, although paytm and Flipkart are not far behind.
7. Reliability of the website or application: Amazon seems to take the lead in this category too, although Flipkart and paytm are not far behind.
8. Quickness to complete purchase: Amazon seems to take the lead in this category too, although Flipkart and paytm are not far behind.
9. Availability of several payment options: Here, Amazon and Flipkart, both are the favourites. Although a lot of people also tend to go towards Myntra
10. Speedy order delivery: Amazon seems to take the lead in this category too, followed by Flipkart.
11. Privacy of customers' information: Amazon has a good reputation for maintaining privacy, followed by Flipkart
12. Security of customer financial information: Here other than Amazon and Flipkart, Paytm.com, Myntra.com, Snapdeal.com are also trusted by a lot of people. This shows that all compamnies pay special attention to security.
13. Perceived Trustworthiness: Amazon and Flipkart are winners here as well. Although Myntra.com and Snapdeal.com are also not far behind.
14. Presence of online assistance through multi-channel: Amazon.in, Flipkart.com, Myntra.com, Snapdeal
15. Longer time to get logged in (promotion, sales period): Amazon, paytm
16. Longer time in displaying graphics and photos (promotion, sales period): Amazon.in, Flipkart.com
17. Late declaration of price: Myntra, Paytm, Snapdeal. These companies should work on this area to improve.
18. Longer page loading time (promotion, sales period): Myntra and Paytm have bad feedback in this. Flipkart should also have a look into it, as it is at the 3rd place.
19. Limited mode of payment on most products (promotion, sales period): Snapdeal is the most voted answer for this.
20. Longer delivery period: Paytm and Snapdeal need to shorten their delivery time.
21. Change in website/Application design: Amazon.in
22. Frequent disruption when moving from one page to another: Amazon.in
23. Website is as efficient as before: Amazon.in
24. Which of the Indian online retailer would you recommend to a friend?: Amazon/Flipkart

Conclusion of Descriptive Analysis

Comparing the Customer's Perceptions and the Company's performance we can conclude that the Companies likely to have

High Customer Satisfaction and Retention:

1. Amazon.com
2. Flipkart.com

High Risk of Customer Churn:

1. Myntra.com
2. Snapdeal.com

Let us move to Inferential Statistics now

2. Inferential Statistics

First Hypothesis

H0: There is no association between gender and e-retail company (Gender and e-retail company are independent)

H1: There is an association between gender and e-retail company

```
# The 2 Columns needed are 1st and Last
dataset_table=pd.crosstab(df['1Gender of respondent'],
                          df['Which of the Indian online retailer would you recommend to a friend?'])
dataset_table
```

Which of the Indian online retailer would you recommend to a friend?	Amazon.in	Amazon.in, Flipkart.com	Amazon.in, Flipkart.com, Myntra.com	Amazon.in, Myntra.com	Amazon.in, Paytm.com	Amazon.in, Paytm.com, Myntra.com	Flipkart.com	Flipkart.com, Paytm.com, Myntra.com, snapdeal.com
1Gender of respondent								
Female	47	42	14	22	11	16	28	1
Male	32	20	1	8	2	4	11	10

```
#Observed Values
Observed_Values = dataset_table.values
print("Observed Values :-\n",Observed_Values)
```

```
Observed Values :-
[[47 42 14 22 11 16 28  1]
 [32 20  1  8  2  4 11 10]]
```

```
Expected_Values=val[3]
```

```
Expected_Values
```

```
array([[53.15613383, 41.71747212, 10.0929368 , 20.18587361,  8.7472119 ,
        13.45724907, 26.24163569,  7.40148699],
       [25.84386617, 20.28252788,  4.9070632 ,  9.81412639,  4.2527881 ,
        6.54275093, 12.75836431,  3.59851301]])
```

```
no_of_rows=len(dataset_table.iloc[0:2,0])
no_of_columns=len(dataset_table.iloc[0,0:8])
ddof=(no_of_rows-1)*(no_of_columns-1)
print("Degree of Freedom:-",ddof)
alpha = 0.05
```

Degree of Freedom:- 7

```
from scipy.stats import chi2
chi_square=sum([(o-e)**2./e for o,e in zip(Observed_Values,Expected_Values)])
chi_square_statistic=chi_square[0]+chi_square[1]
```

```
print("chi-square statistic:-",chi_square_statistic)
```

chi-square statistic:- 2.1852257166231297

```
critical_value=chi2.ppf(q=1-alpha,df=ddof)
print('critical_value:',critical_value)
```

critical_value: 14.067140449340169

```
#p-value
p_value=1-chi2.cdf(x=chi_square_statistic,df=ddof)
print('p-value:',p_value)
print('Significance level: ',alpha)
print('Degree of Freedom: ',ddof)
```

p-value: 0.9488810780740886

Significance level: 0.05

Degree of Freedom: 7

```
if chi_square_statistic>=critical_value:
    print("Reject H0,There is an association between gender and e-retail company recommended to a friend")
else:
    print("Failed to Reject H0,There is no association between gender and e-retail company recommended to a friend")

if p_value<=alpha:
    print("Reject H0,There is an association between gender and e-retail company recommended to a friend")
else:
    print("Failed to Reject H0,There is no association between gender and e-retail company recommended to a friend")
```

Result: Failed to Reject H0, there is no association between gender and e-retail company recommended to a friend.

Second Hypothesis

H0: There is no association between age and e-retail company that person would recommend to a friend(age and e-retail company are independent)

H1: There is an association between age and e-retail company that person would recommend to a friend

Which of the Indian online retailer would you recommend to a friend?	Amazon.in	Amazon.in, Flipkart.com	Amazon.in, Flipkart.com, Myntra.com	Amazon.in, Myntra.com	Amazon.in, Paytm.com	Amazon.in, Paytm.com, Myntra.com	Flipkart.com	Flipkart.com, Paytm.com, Myntra.com, snapdeal.com
2 How old are you?								
21-30 years	18	24	5	13	0	8	8	3
31-40 years	16	18	3	6	8	3	27	0
41-50 yaers	27	14	0	7	5	6	4	7
51 years and above	6	5	7	0	0	0	0	1
Less than 20 years	12	1	0	4	0	3	0	0

```
#Observed Values
Observed_Values = dataset_table.values
print("Observed Values :-\n",Observed_Values)
```

```
Observed Values :-
[[18 24  5 13  0  8  8  3]
 [16 18  3  6  8  3 27  0]
 [27 14  0  7  5  6  4  7]
 [ 6  5  7  0  0  0  0  1]
 [12  1  0  4  0  3  0  0]]
```

```
val=stats.chi2_contingency(dataset_table)
```

```
Expected_Values=val[3]
Expected_Values
```

```
array([[23.20074349, 18.20817844,  4.40520446,  8.81040892,  3.81784387,
        5.87360595, 11.4535316 ,  3.23048327],
 [23.78810409, 18.66914498,  4.51672862,  9.03345725,  3.91449814,
        6.02230483, 11.74349442,  3.31226766],
 [20.55762082, 16.133829 ,  3.90334572,  7.80669145,  3.38289963,
        5.20446097, 10.14869888,  2.86245353],
 [ 5.57992565,  4.37918216,  1.05947955,  2.11895911,  0.91821561,
        1.41263941,  2.75464684,  0.77695167],
 [ 5.87360595,  4.60966543,  1.11524164,  2.23048327,  0.96654275,
        1.48698885,  2.89962825,  0.81784387]])
```

```
no_of_rows=len(dataset_table.iloc[0:5,0])
no_of_columns=len(dataset_table.iloc[0,0:8])
ddof=(no_of_rows-1)*(no_of_columns-1)
print("Degree of Freedom:-",ddof)
alpha = 0.05
```

```
Degree of Freedom:- 28
```

```
chi_square=sum([(o-e)**2./e for o,e in zip(Observed_Values,Expected_Values)])
chi_square_statistic=chi_square[0]+chi_square[1]
```

```
print("chi-square statistic:-",chi_square_statistic)
```

```
chi-square statistic:- 17.219333961318835
```

```
critical_value=chi2.ppf(q=1-alpha,df=ddof)
print('critical_value:',critical_value)
```

```
critical_value: 41.33713815142739
```

```
#p-value
p_value=1-chi2.cdf(x=chi_square_statistic,df=ddof)
print('p-value:',p_value)
print('Significance level: ',alpha)
print('Degree of Freedom: ',ddof)
print('p-value:',p_value)
```

```
p-value: 0.9441297677059306
Significance level: 0.05
Degree of Freedom: 28
p-value: 0.9441297677059306
```

```
if chi_square_statistic>=critical_value:
    print("Reject H0,There is an association between age and e-retail company that person would recommend to a friend")
else:
    print("Failed to Reject H0, Therefore, There is no association between age and e-retail company that person would recommend")

if p_value<=alpha:
    print("Reject H0,There is an association between age and e-retail company that person would recommend to a friend")
else:
    print("Failed to Reject H0, Therefore, There is no association between age and e-retail company that person would recommend")
```

Result: Failed to Reject H0, Therefore, there is no association between age and e-retail company that person would recommend.