FLIP ROBO

Customer Retention-Case Study

Submitted by:

Durgadhar Pathak

Internship 15

ACKNOWLEDGMENT

The internship opportunity I have with Flip Robo Technologies is a great chance for learning and professional development. I am also grateful to our SME Mr. Sajid Choudhary and Ms. Seema Verma for their valuable and constructive suggestions during the planning and development of this project. Their quick support and references helped a lot in building this project. Also, I am very thankful to our SMEs & Flip Robo Team for understanding technical issue faced by me and provide quick resolution & providing enough time for submission.

Also, I am thankful to DT support Team for their continuous effort to resolve our queries during project building.

Research papers that helped me in this project was as follows: -

1- https://www.researchgate.net/publication/236204251_Customer_Retention_Model_-_A_Case_Study_of_a_Branded_Passenger_Car
2- https://kth.diva-portal.org/smash/get/diva2:1223723/FULLTEXT01.pdf
3- http://norma.ncirl.ie/2923/1/larajanescarff.pdf

Articles that helped me in this project was as follows:

1- https://www.sciencedirect.com/science/article/pii/S187704281406039X
2- https://www.forbes.com/sites/shamahyder/2021/04/09/how-to-win-the-retention-game-a-case-study-with-lingopie/
3- https://www.tandfonline.com/doi/full/10.1080/23311975.2020.1738200

References:

1- https://machinelearningmastery.com/
2- https://scikit-learn.org/stable/
3- https://www.geeksforgeeks.org/machine-learning/
4- https://pandas.pydata.org/
5- https://www.datacamp.com/
6- https://www.ibm.com/cloud/learn/machine-learning
7- https://www.oreilly.com/library/view/web-scraping-with/9781491985564/
8- http://analytikk.com/case-studies/customer-retention-case-study/

INTRODUCTION

- ## Business Problem Framing

  Customer retention: - Customer retention refers to a company's ability to turn customers into repeat buyers and prevent them from switching to a competitor. It indicates whether your product and the quality of your service please your existing customers. It's also the lifeblood of most subscription-based companies and service providers.

- ## Conceptual Background of the Domain Problem

  - Customer retention benefits: -
  - 1- Cost savings: Customer retention is generally more cost-effective than acquiring first-time customers.
  - 2- Positive word of mouth marketing: Loyal customers are more likely to tell their friends and family about your brand.
  - 3- A better bottom line: Increasing retention rates by just 5 percent can increase revenue by 25 percent to 95 percent.

- # Customer retention Strategy

  Customer retention Strategy: -Customer retention strategies are the processes and initiatives businesses put in place to build customer loyalty and improve customer lifetime value. Customer retention is different from customer acquisition or lead generation. It focuses on customers who have already signed up for a service or purchased a product from you.

  But retaining customers is about more than just transactions—it's about relationships. Research shows that customers view their relationships with brands similarly to their relationships with friends. Customers like brands that are reliable, authentic, and aware of what matters to them.

  If we want to create an effective customer retention strategy, it starts with understanding what our customers value and providing a product that matches that. Focusing on buyer relationships with your existing customers will help in boost their brand loyalty. These shoppers will continue to choose our brand even when presented with other options. With this loyal base, our brand will be more likely to weather volatile markets.

- ## Review of Literature

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

    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.
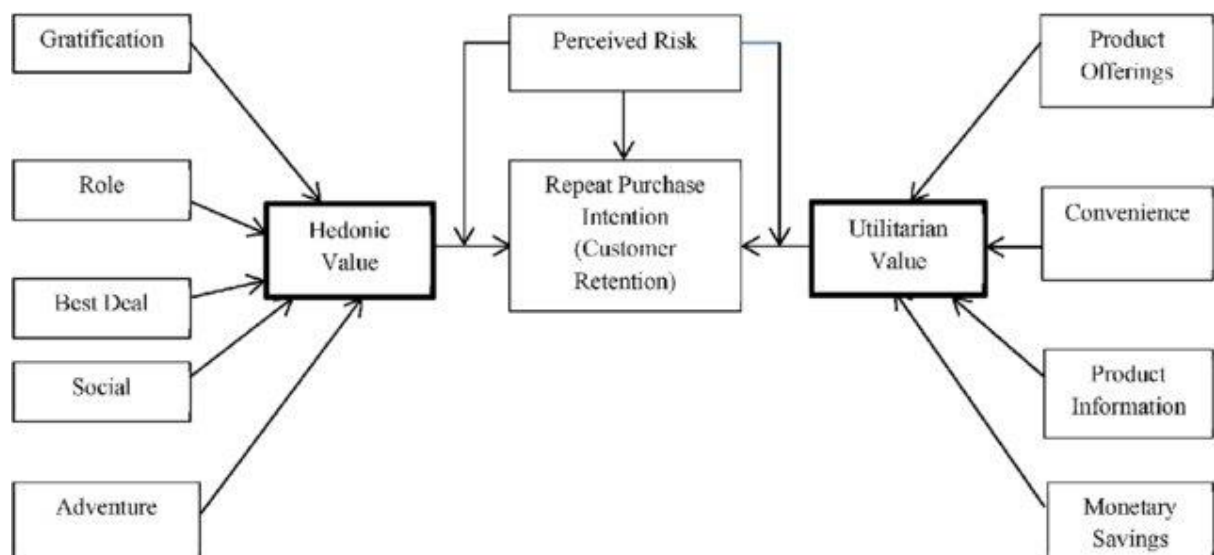
- # Motivation for the Problem Undertaken

    Since e-commerce growth has been significant so far at each industry for the last couple of years, understanding the customer who purchase your product and their journey post order in the platform plays a crucial role for retaining a customer. From the dataset we got the feedback of each of the parameters for a customer and through the data analysis each significant parameter has been dig down and understand the deeper meaning of it. This project helps me understand the objective of customer feedback which at enterprise level has a greater focus.

Analytical Problem Framing

- ## Mathematical/ Analytical Modelling of the Problem

We have to apply our analytical skills to give findings, conclusions and detailed data analysis and write these all on Jupyter notebook. Only data analysis is required by client.

Statistical Summary: summary statistics is used to summarize set of observations, in order to communicate the largest amount of information as simply as possible. It includes central Tendency, dispersion, skewness, variance, range, deviation etc.

	Gender	Age	city	Pin Code	How Long You are Shopping Online	How many times in the past 1 year	How do you access the internet	device	screen size	operating system	...	Longer time to get logged in	Longer time in displaying graphics and photos	Late declaration of price	Longer page loading time	pay
count	269	269	269	269	269	269	269	269	269	269	...	269	269	269	269	
unique	2	5	11	39	5	6	4	4	4	3	...	10	10	8	11	
top	Female	31-40 years	Delhi	201308	Above 4 years	Less than 10 times	Mobile internet	Smartphone	Others	Window/windows Mobile	...	Amazon.in	Amazon.in, Flipkart.com	Myntra.com	Myntra.com	Snap
freq	181	81	58	38	98	114	142	141	134	122	...	57	60	75	61	

Correlation: After seeing many correlated values we can say that many columns have correlation values and dropping some of these will be better for our dataset.

Skewness: If the skewness is between -0.5 and 0.5, then dataset is fairly symmetrical and symmetrical distribution will have a skewness of zero. So accordingly, we are removing skewness using NumPy mathematical functions log & square transform.

Missing Data: -

Observation: Dataset has no missing values.

Data Sources and their formats

Dataset provided by client is in *.csv format. There are two sheets (one is detailed) and second is encoded in the excel file. We can use any of them. The number of column(s) is more than 47.

The data is collected from the Indian online shoppers. Results indicate the e-retail success factors, which are very much critical for customer satisfaction.

Dataset: - Different columns are as: -

#	Column	Non-Null Count	Dtype
0	Gender	269 non-null	object
1	Age	269 non-null	object
2	city	269 non-null	object
3	Pin Code	269 non-null	int64
4	How Long You are Shopping Online	269 non-null	objects
5	How many times in the past 1 year	269 non-null	objects
6	How do you access the internet	269 non-null	object
7	device	269 non-null	object
8	screen size	269 non-null	object
9	operating system	269 non-null	object
10	browser	269 non-null	object
11	channel	269 non-null	object
12	After first visit	269 non-null	object
13	How much times	269 non-null	object
14	payment Option	269 non-null	object
15	How frequently do you abandon	269 non-null	object
16	Why did you abandon	269 non-null	object
17	content on the website	269 non-null	object
18	similar product	269 non-null	object
19	Complete information	269 non-null	object
20	relevant information	269 non-null	object
21	Ease of navigation	269 non-null	object
22	Loading and processing speed	269 non-null	object
23	User friendly	269 non-null	object
24	Convenient Payment methods	269 non-null	object
25	Trust that the online retail store	269 non-null	object

26	Empathy towards the customers	269 non-null	object
27	privacy of the customer	269 non-null	object
28	Responsiveness and availability	269 non-null	object
29	monetary benefit and discounts	269 non-null	object
30	Enjoyment is derived from shopping online	269 non-null	object
31	Shopping online is convenient and flexible	269 non-null	object
32	Return and replacement policy	269 non-null	object
33	Gaining access to loyalty	269 non-null	object
34	quality Information on the website	269 non-null	object
35	User derive satisfaction	269 non-null	object
36	Net Benefit	269 non-null	object
37	User satisfaction cannot exist without trust	269 non-null	object
38	Offering a wide variety	269 non-null	object
39	Provision of complete and relevant product information	269 non-null	object
40	Monetary savings	269 non-null	object
41	The Convenience of patronizing the online retailer	269 non-null	object
42	sense of adventure	269 non-null	object
43	enhances your social status	269 non-null	object
44	gratification shopping on your favourite e-tailer	269 non-null	object
45	Shopping on the website helps you fulfil certain roles	269 non-null	object
46	Getting value for money spent	269 non-null	object
47	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 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
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	269 non-null	object
62	Longer time in displaying graphics and photos	269 non-null object	
63	Late declaration of price	269 non-null	object
64	Longer page loading time	269 non-null	object
65	Limited mode of payment on most products	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	

- # Data Pre-processing Done

    The data pipeline starts with collecting the data and ends with communicating the results & Data pre-processing is a data mining technique which is used to transform the raw data in a useful and efficient format. It involves 4 steps: - Cleaning, Formatting, Scaling, and Normalization. While building a machine learning model, if we haven't done any pre-processing like correcting outliers, handling different formats, normalization and scaling of data, or feature engineering, we might end up considering those 1% of results that are false. Some steps performed in this project are: -

- Column wise Empty cell analysis done & found no missing values.
- Checked unique values in each column to explore dataset more deeply.

- **Feature Encoding and Normalization**: - Features came in a variety of format, e.g., integers, floating numbers, string values, etc. So, we Checked Concise Summary of our Data Frame and we have noticed that 8 columns have object (str or mix str) data type. This was a challenge for us as these features cannot be directly used for training. To prevent regression biases towards certain features, we dropped some of the irrelevant features and make sure that we haven't lose important data.

Data Visualization: Data visualization is the graphical representation of information and data. Different visualizations for our input and output features are as-
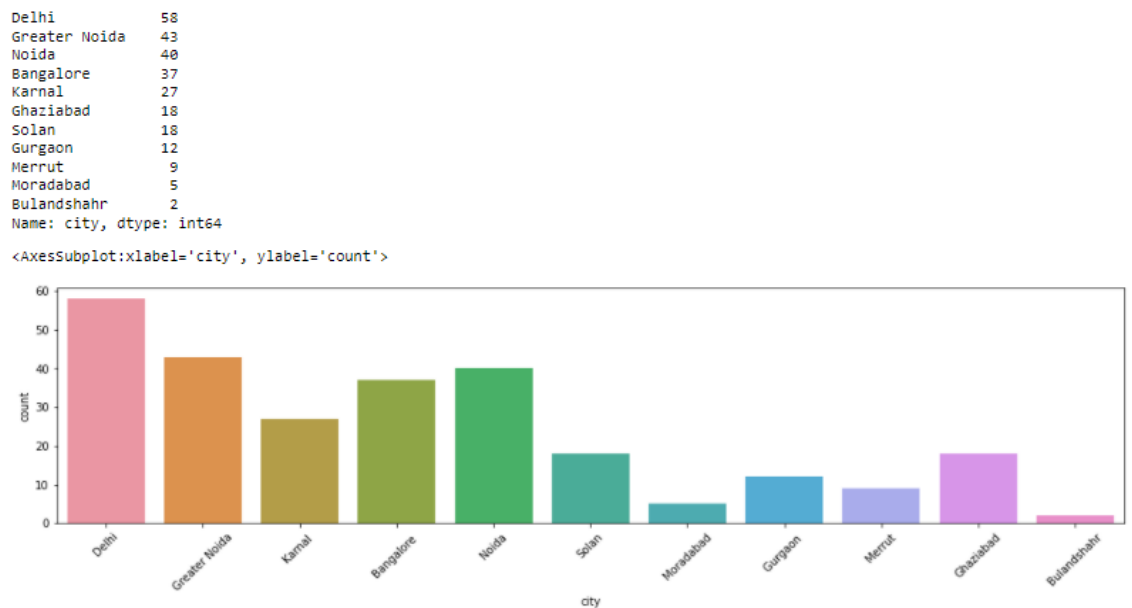
Customer Related Information: -

1- Count plot for " Gender" feature: -

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

Gender
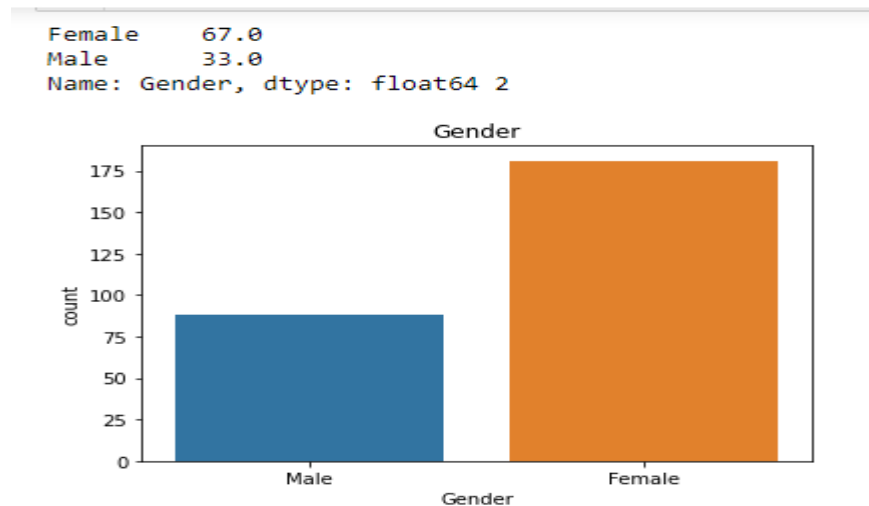
Observation:

Out of total customers, 67% customers are female.

2- Count plot for " city" feature: -

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

```
<AxesSubplot:xlabel='city', ylabel='count'>
```


Observations:

1- People from Delhi city shops the most online, followed by greater Noida,Noida and Bangalore.

2- least online shopping is in bulandshahar.

3- Count plot for "Age" feature: -

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

```
<AxesSubplot:xlabel='Age', ylabel='count'>
```
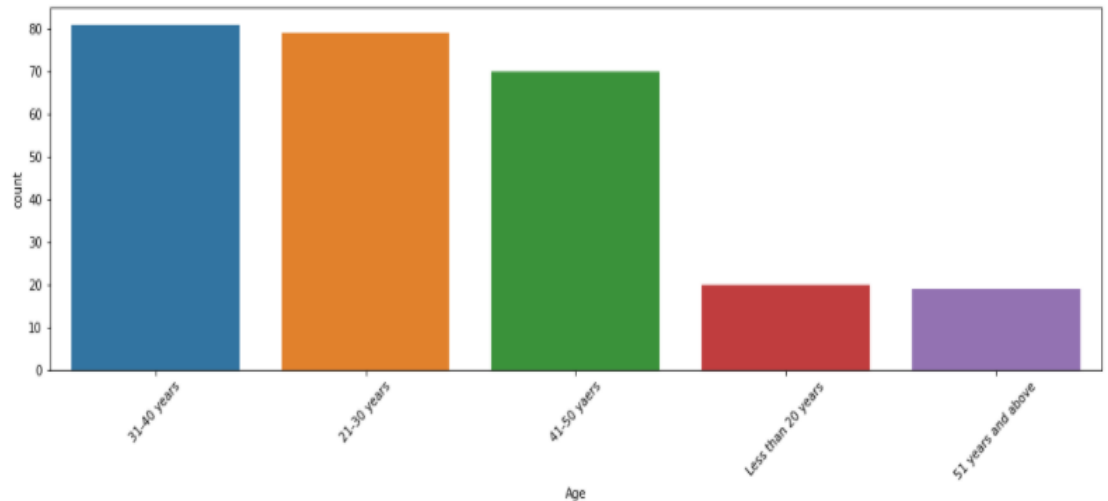

Observations:

1-Customers aged between 31 to 40 years, use to do maximum online shopping.

2-Customers aged between 21 to 30 years, are 2nd highest online shoppers.

4- Line plot for "How do you access the internet" feature: -

```
Mobile internet   142
Wi-Fi              76
Mobile Internet    47
Dial-up             4
Name: How do you access the internet, dtype: int64
```

```
<AxesSubplot:>
```
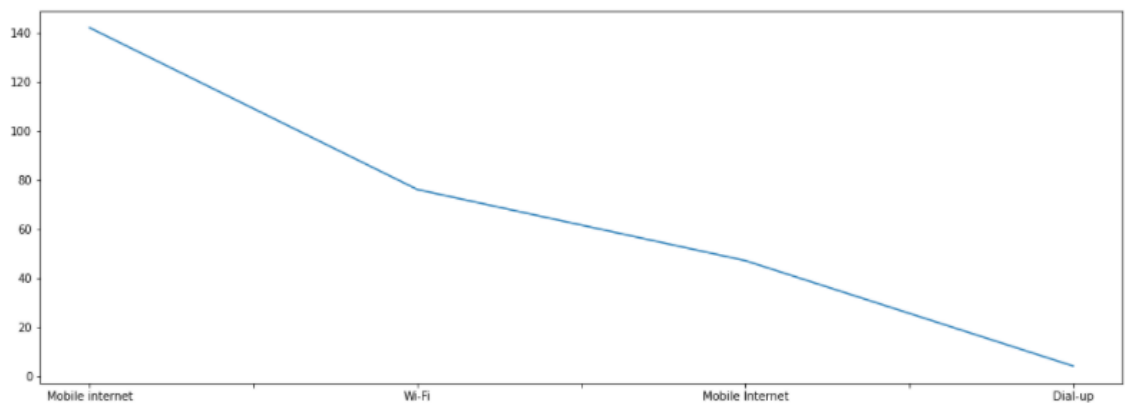

Observations:

1- Maximum online shoppers use Mobile internet.

2- Very few shoppers use Dial-up connection for shopping.

5- Count plot for "device" feature: -

```
Smartphone     141
Laptop          86
Desktop         30
Tablet          12
Name: device, dtype: int64
```

```
<AxesSubplot:xlabel='device', ylabel='count'>
```
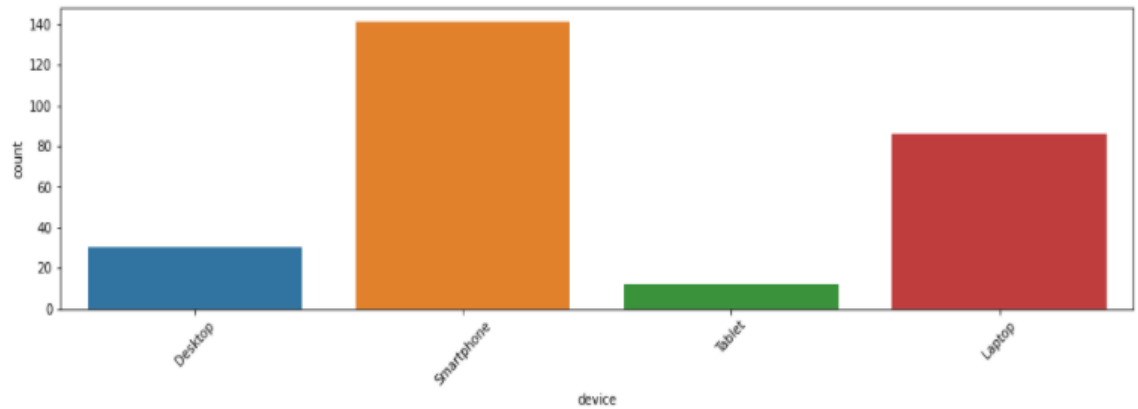

Observations:

1- Maximum customers use Smartphone for shopping purpose.

2- Very few customers use Tablet for shopping purpose.

6- Line plot for How Long You are Shopping Online" feature: -

```
Above 4 years      98
2-3 years          65
3-4 years          47
Less than 1 year   43
1-2 years          16
Name: How Long You are Shopping Online, dtype: int64
```

```
<AxesSubplot:>
```
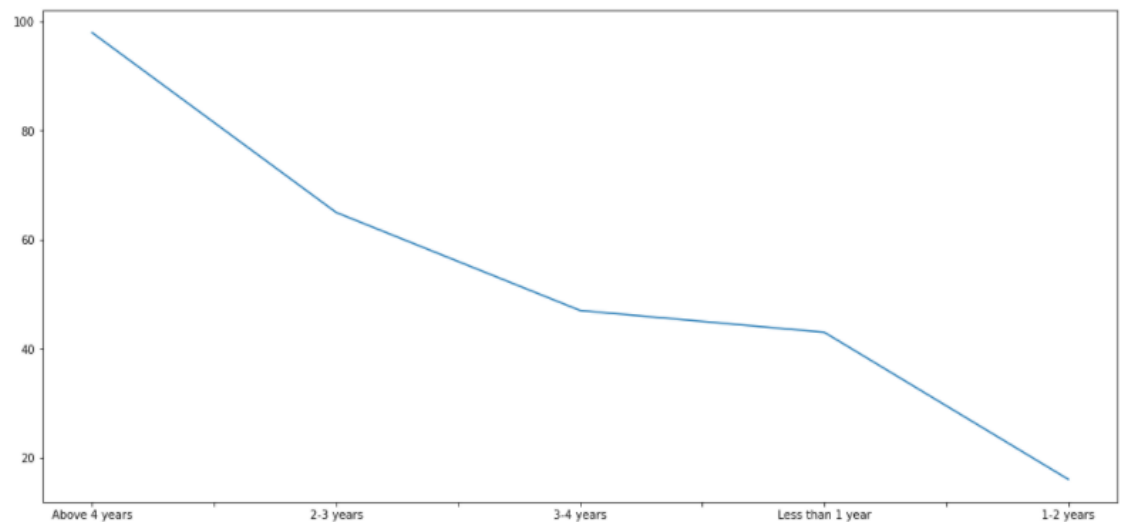

Observations:

1- Shoppers having tenure more than 4 years are maximum and we can say that they are potential & reliable customers.

2- From above Plot and calculation, we can say that if a customer shops for a longer time then definitely he will become a potentail customer.

7- Count plot for "operating system" feature: -

```
Window/windows Mobile    122
Android                   85
IOS/Mac                   62
Name: operating system, dtype: int64

<AxesSubplot:xlabel='operating system', ylabel='count'>
```
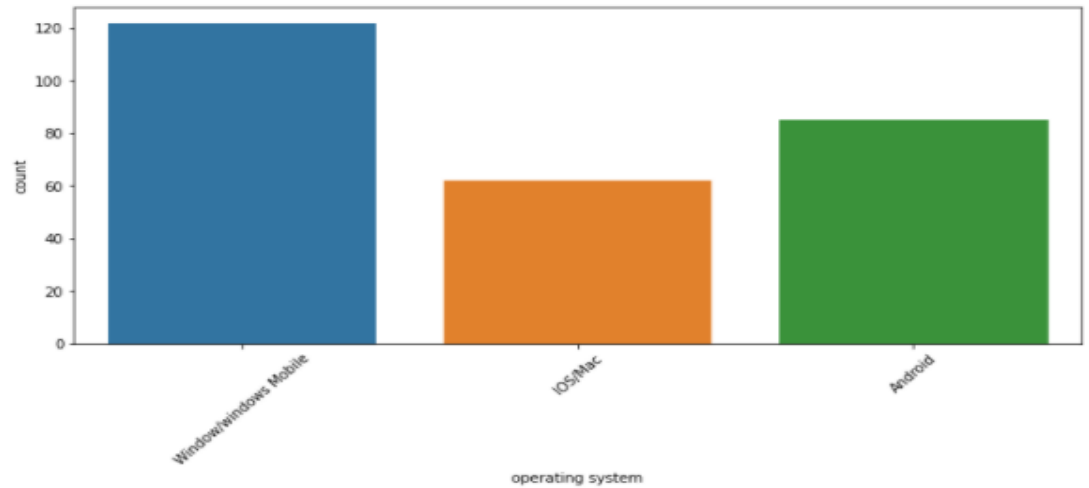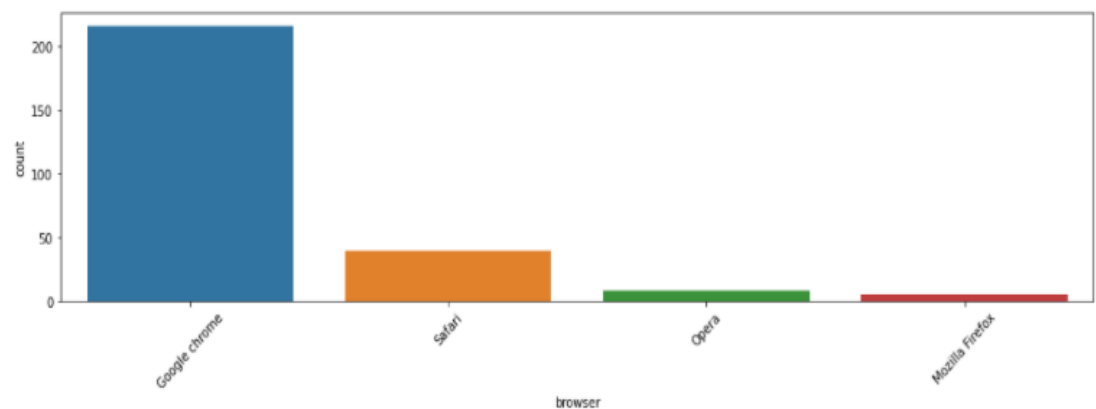

Observations:

1- Maximum shoppers use window o/s.

2- Mac users are less as compared to other o/s.

3- Iphone/Mac o/s is costly as compared to windows & android so it might be the reason of less mac users.

8- Count plot for "browser" feature: -

```
Google chrome    216
Safari            40
Opera              8
Mozilla Firefox    5
Name: browser, dtype: int64

<AxesSubplot:xlabel='browser', ylabel='count'>
```


Observations:

1- 80% shoppers use Google chrome browser. Reason behind high number of google chrome users may be it's user friendliness, security features,user interface,extensions,translation features & stability.

2- Near around only 2% shoppers use Mozilla Firefox.

9- Count plot for "channel" feature: -

```
Search Engine      230
Content Marketing   20
Display Adverts     19
Name: channel, dtype: int64

<AxesSubplot:xlabel='channel', ylabel='count'>
```
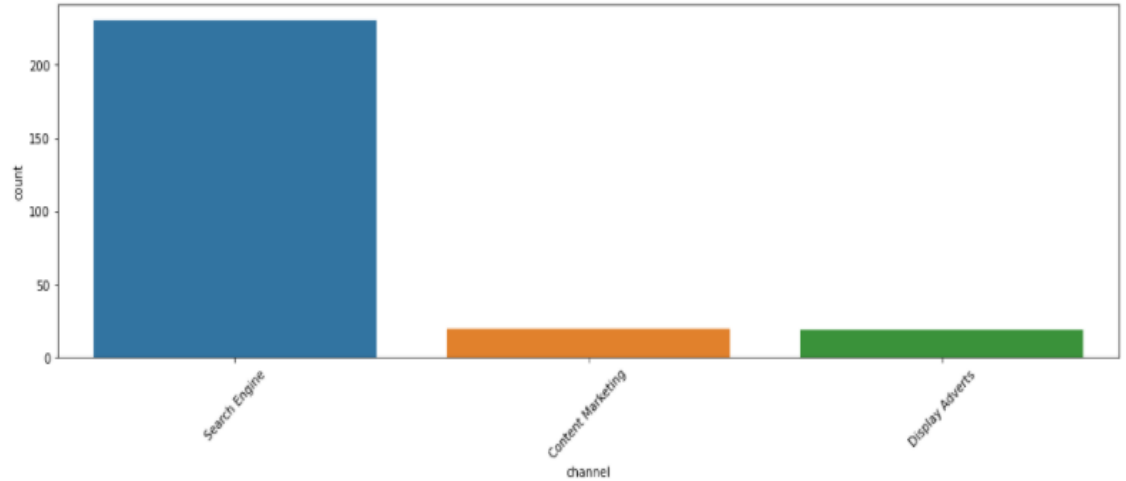

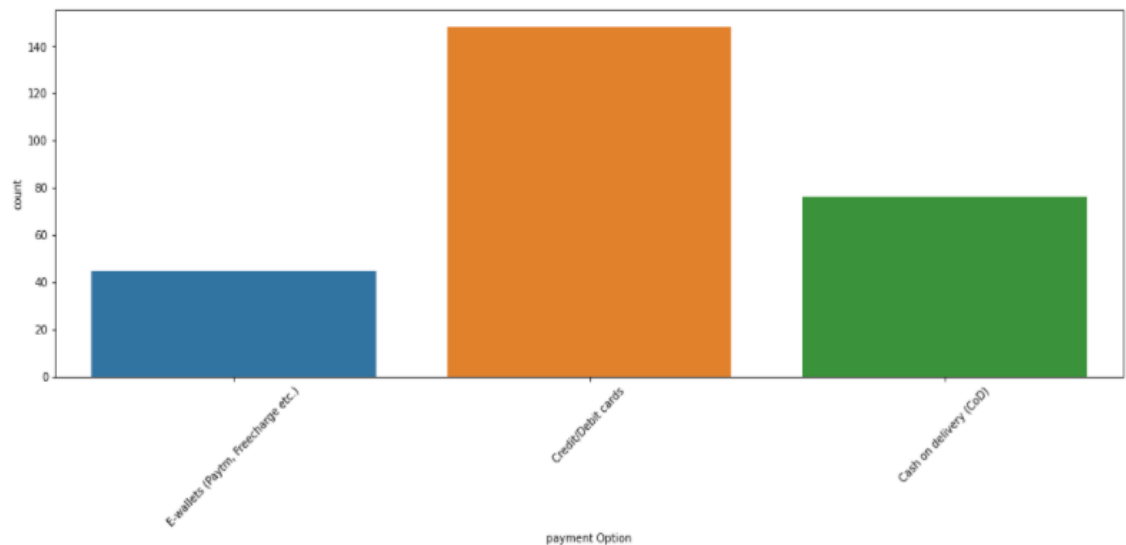Observations:

1- Maximum shoppers use Search Engine channel while shopping because of user easiness and multiple option exploration.

2- Only 7% shopper use Display Adverts channel may be due to increase of online shopping portals.

10- Count plot for "payment Option" feature: -

```
Credit/Debit cards                148
Cash on delivery (CoD)             76
E-wallets (Paytm, Freecharge etc.)  45
Name: payment Option, dtype: int64

<AxesSubplot:xlabel='payment Option', ylabel='count'>
```
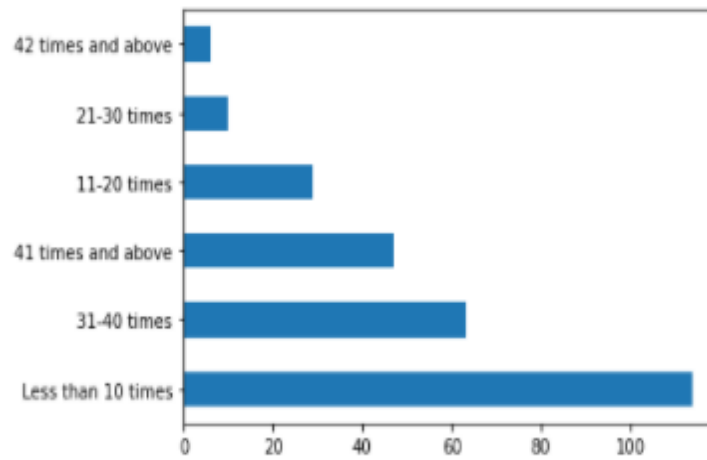

Observations:

1- Maximum customers use Credit/Debit cards for payment transactions.

2- Only 16-17% customers use E-wallet for payment transactions.

11- Count plot for "How many times in the past 1 year" feature using bars: -

`<AxesSubplot:>`

Observations:

1- In last one year, maximum number of customers shopped for less than 10 times.

2- Frequency of customers shopping more than 30 is less but these customers are potential & regular customers.

12- Line plot for "Why did you abandon" feature: -

```
Better alternative offer       133
Promo code not applicable       54
Change in price                 37
Lack of trust                   31
No preferred mode of payment    14
Name: Why did you abandon, dtype: int64
```
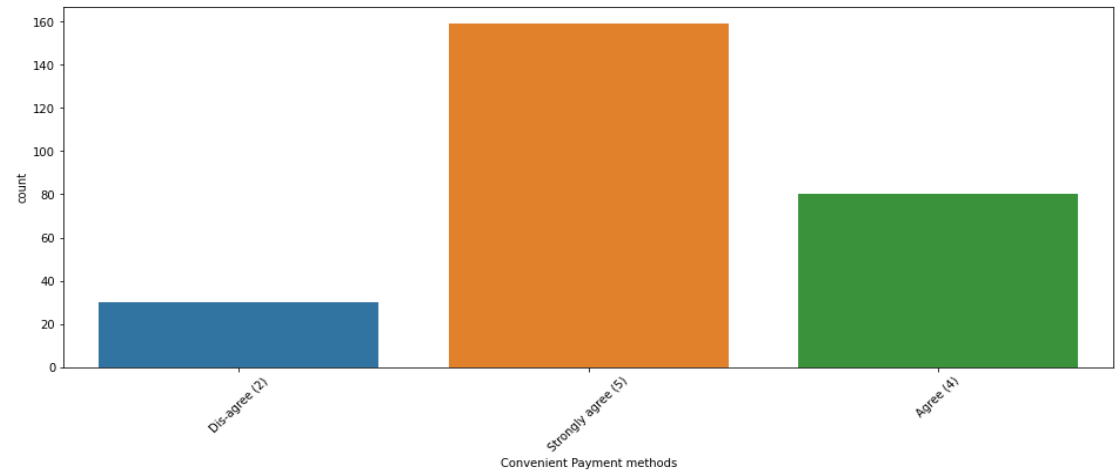
`<AxesSubplot:>`
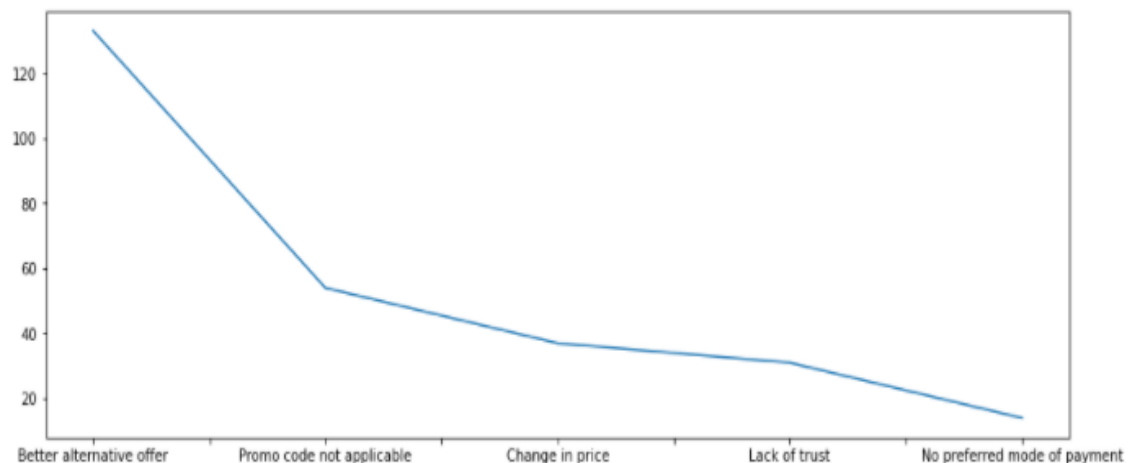
Observations:

1- Maximum near around 50% of customers, abandon the purchase due to Better alternative offer.

2- No preferred mode of payment has not too much (only 5%) impact on abandon.

Customer Feedback Analysis: -

1- Count plot for "Convenient Payment methods" feature: -

```
Strongly agree (5)    159
Agree (4)              80
Dis-agree (2)          30
Name: Convenient Payment methods, dtype: int64

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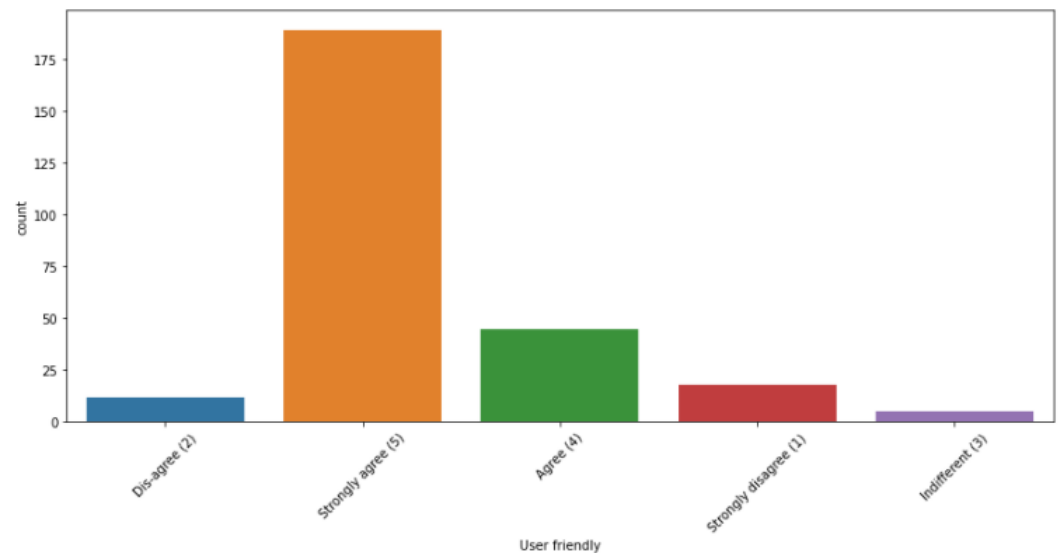

Observations:

1- Maximum near around 59% of customers are strongly agree about Convenient Payment methods.

2- Only 11% of customers are dis-agree on Convenient Payment methods.

2- Count plot for "User friendly" feature: -

```
Strongly agree (5)       189
Agree (4)                 45
Strongly disagree (1)     18
Dis-agree (2)             12
Indifferent (3)            5
Name: User friendly, dtype: int64

<AxesSubplot:xlabel='User friendly', ylabel='count'>
```
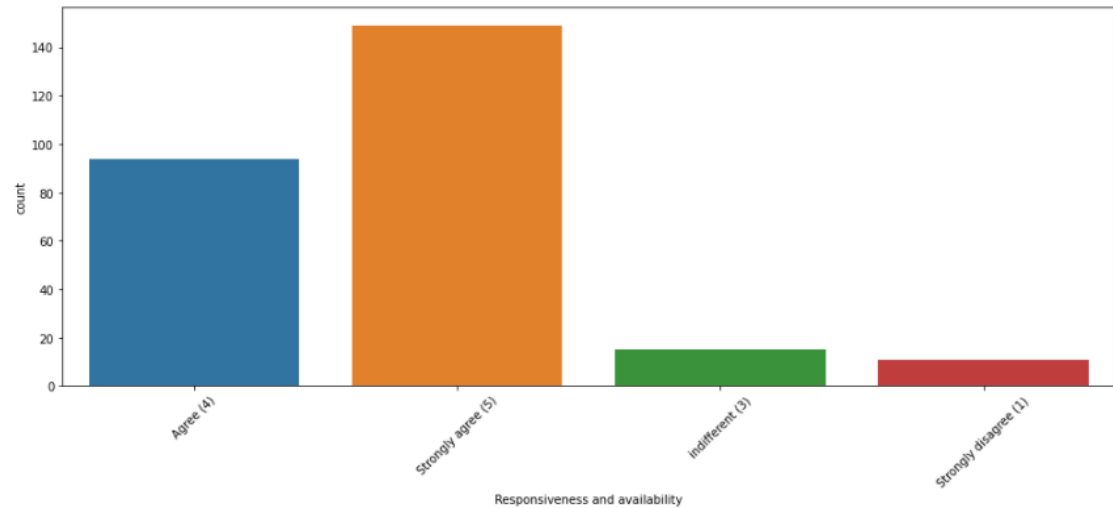

Observations:

1- Maximum near around 70% of customers are strongly agree on user friendliness.

2- Only 4.5% of customers are disagree on user friendliness and only near around 2% customers have indifferent opinion on user friendliness.

3- Count plot for "Responsiveness and availability" feature: -

```
Strongly agree (5)      149
Agree (4)                94
indifferent (3)          15
Strongly disagree (1)    11
Name: Responsiveness and availability, dtype: int64

<AxesSubplot:xlabel='Responsiveness and availability', ylabel='count'>
```
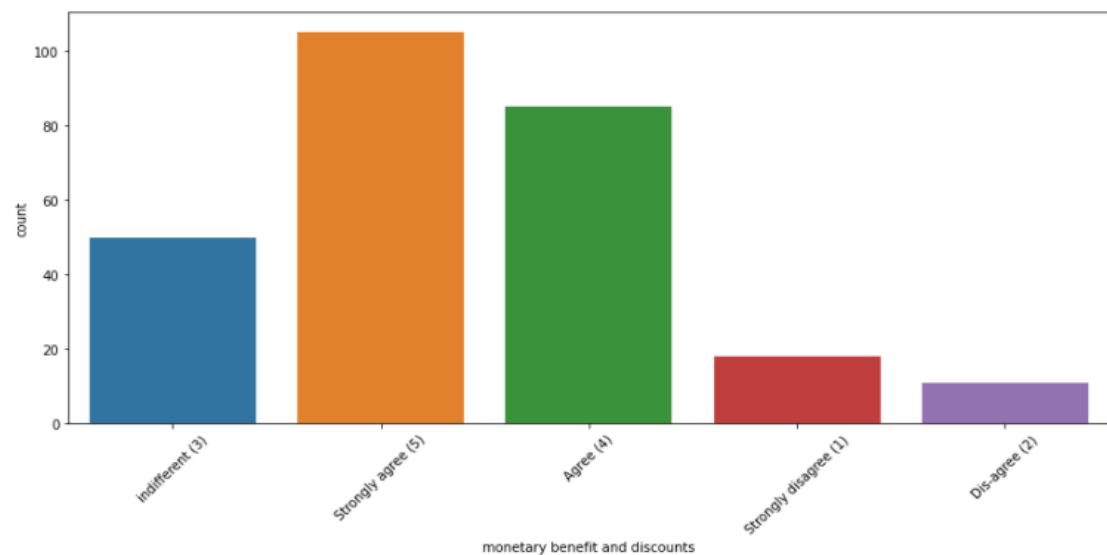

Observations:

1- Maximum near around 55% of customers are strongly agree on responsiveness and availability.

2- Only 4% of customers are strongly disagree on responsiveness and availability.

4- Count plot for "monetary benefit and discounts" feature: -

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

: <AxesSubplot:xlabel='monetary benefit and discounts', ylabel='count'>
```
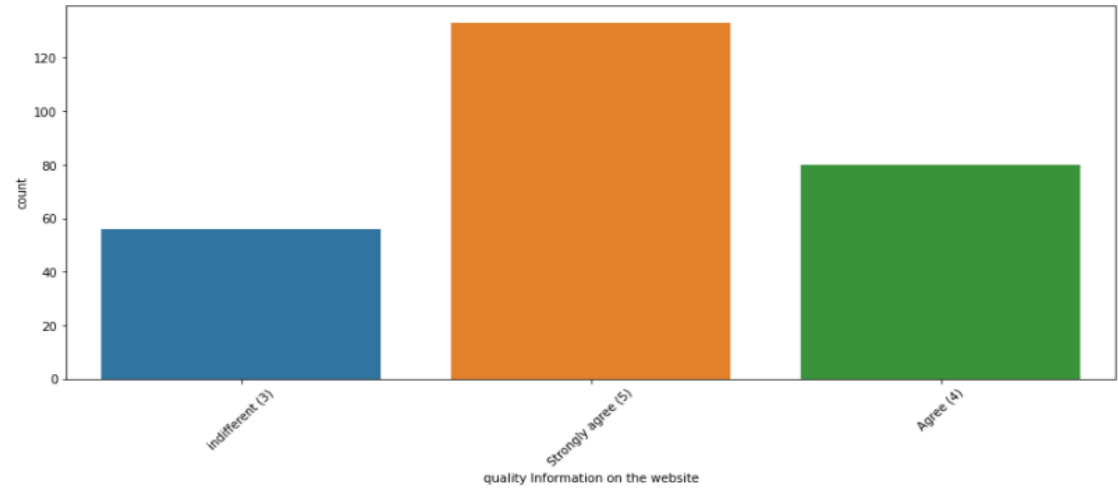

Observations:

1- Maximum near around 39% of customers are strongly agree with monetary benefit and discounts.

2- Only 4% of customers are disagree with monetary benefit and discounts.

5- Count plot for "quality Information on the website" feature: -

```
Strongly agree (5)    133
Agree (4)              80
indifferent (3)        56
Name: quality Information on the website, dtype: int64

<AxesSubplot:xlabel='quality Information on the website', ylabel='count'>
```
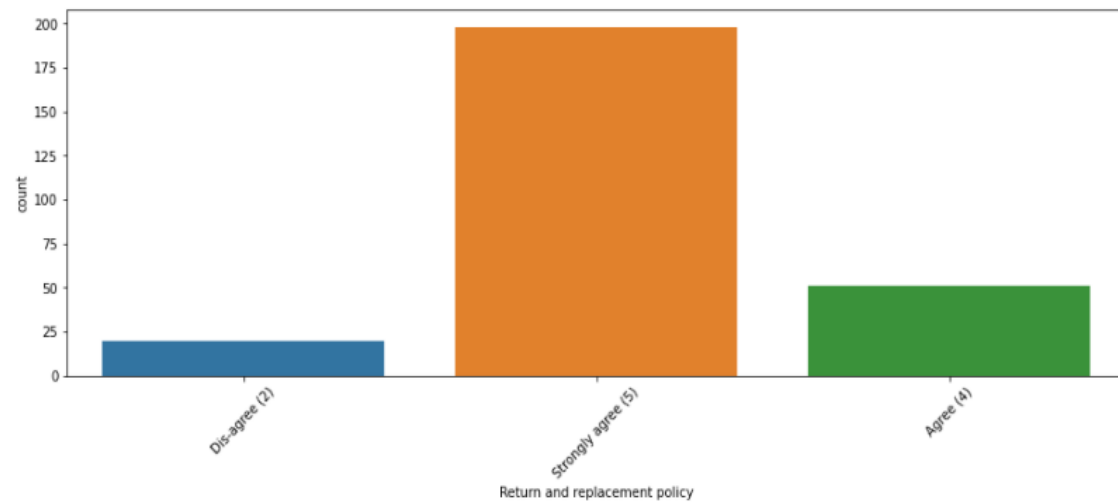

Observations:

1- Maximum near around 50% of customers are strongly agree about quality Information on the website.

2- Only 20.81% of customers have indifferent opinion about quality Information on the website.

6- Count plot for "Return and replacement policy" feature: -

```
Strongly agree (5)    198
Agree (4)              51
Dis-agree (2)          20
Name: Return and replacement policy, dtype: int64

<AxesSubplot:xlabel='Return and replacement policy', ylabel='count'>
```
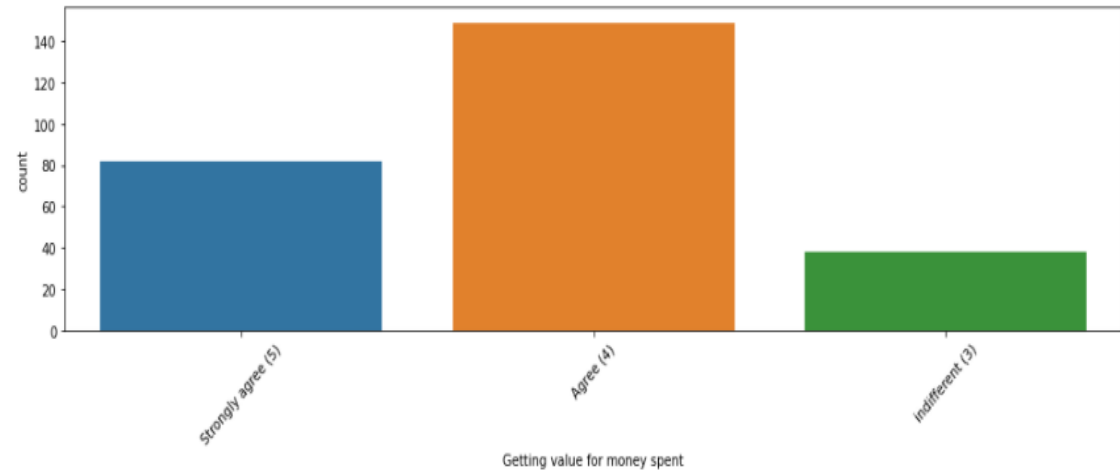

Observations:

1- Maximum near around 73.60% of customers are strongly agree towards Return and replacement policy.

2- Only 7.43% of customers are disagree towards Return and replacement policy.

7- Count plot for "Getting value for money spent" feature:

```
Agree (4)             149
Strongly agree (5)     82
indifferent (3)        38
Name: Getting value for money spent, dtype: int64

<AxesSubplot:xlabel='Getting value for money spent', ylabel='count'>
```
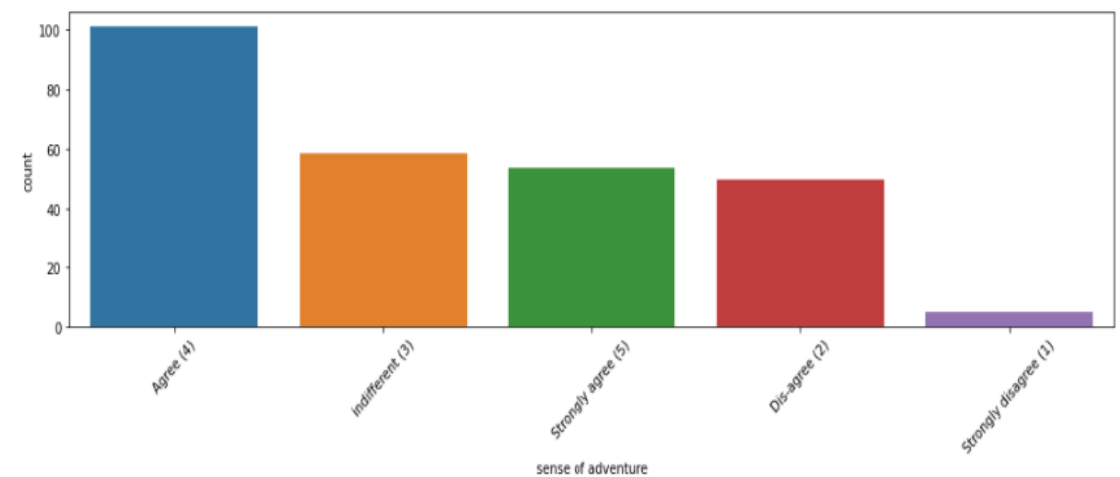

Observations:

1- Maximum near around 55.39% of customers are agree with Getting value for money spent.

2- Only 14.12% of customers have indifferent opinion about Getting value for money spent.

8- Count plot for "sense of adventure" feature:

```
Agree (4)             101
indifferent (3)        59
Strongly agree (5)     54
Dis-agree (2)          50
Strongly disagree (1)   5
Name: sense of adventure, dtype: int64

<AxesSubplot:xlabel='sense of adventure', ylabel='count'>
```
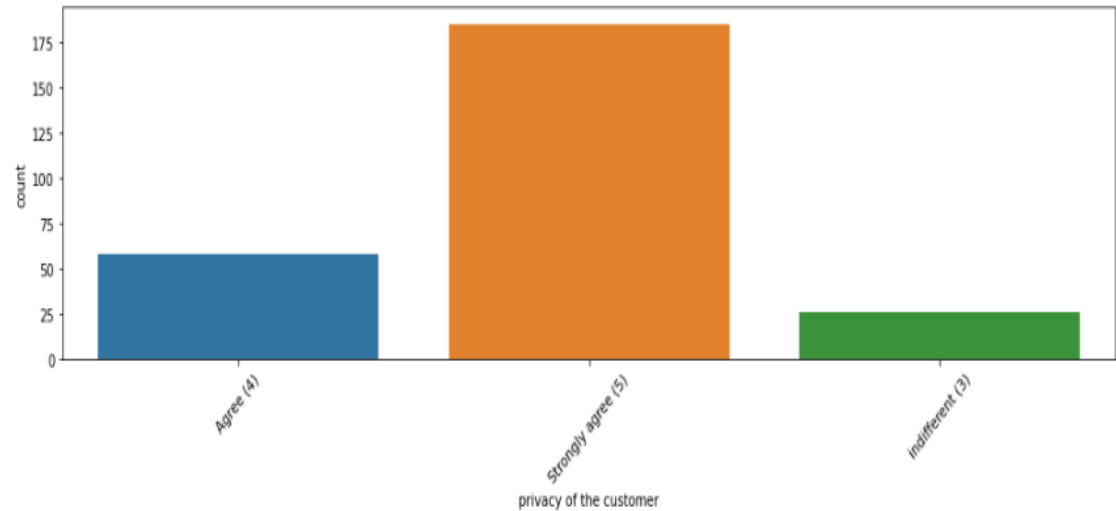

Observations:

1- Maximum near around 37.54% of customers are agree on sense of adventure.

2- Only near around 2% of customers are strongly disagree on sense of adventure.

9- Count plot for "privacy of the customer" feature:

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

<AxesSubplot:xlabel='privacy of the customer', ylabel='count'>
```


Observations:

1- Maximum near around 68.77% of customers are strongly agree with privacy policy.

2- Only near around 9.66% of customers have indifferent opinion on privacy policy.

10- Count plot for "Net Benefit" feature:

```
Strongly agree (5)    164
Agree (4)              54
indifferent (3)        40
Dis-agree (2)          11
Name: Net Benefit, dtype: int64

<AxesSubplot:xlabel='Net Benefit', ylabel='count'>
```
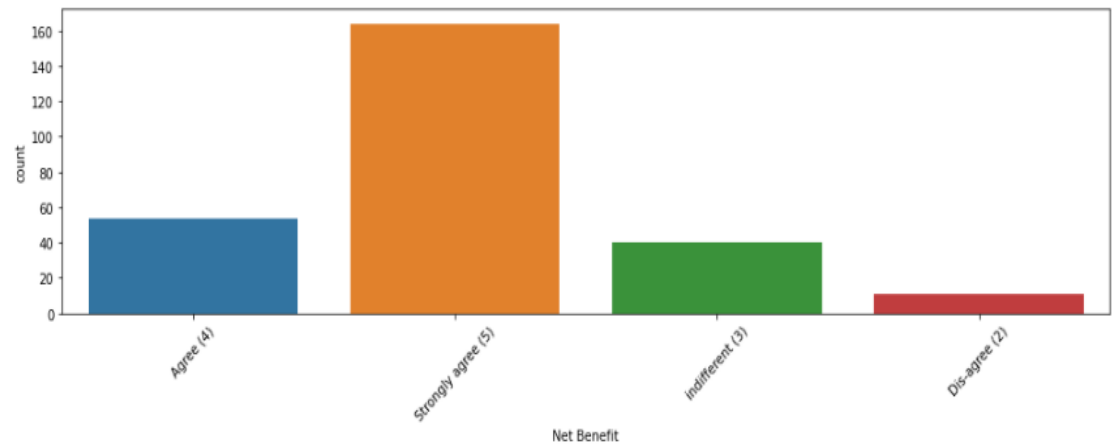
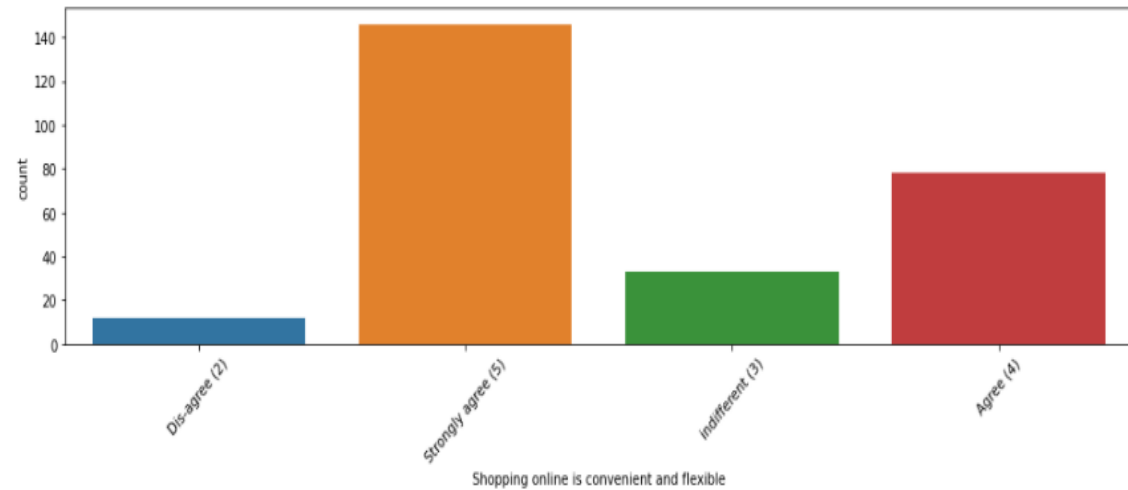

Observations:

1- Maximum near around 61% of customers are strongly agree on achievement of net benefit.

2- Only near around 4% of customers are disagree on achievement of net benefit.

11- Count plot for "Shopping online is convenient and flexible" feature:

```
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'>
```


Observations:

1- Maximum near around 54% of customers are strongly agree with convienece and flexibility of online Shopping.

2- Only near around 4.5% of customers are disagree with convienece and flexibility of online Shopping.

12- Count plot for "Gaining access to loyalty" feature:

```
Strongly agree (5)    115
Agree (4)              64
indifferent (3)        64
Dis-agree (2)          15
Strongly disagree (1)  11
Name: Gaining access to loyalty, dtype: int64

: <AxesSubplot:xlabel='Gaining access to loyalty', ylabel='count'>
```
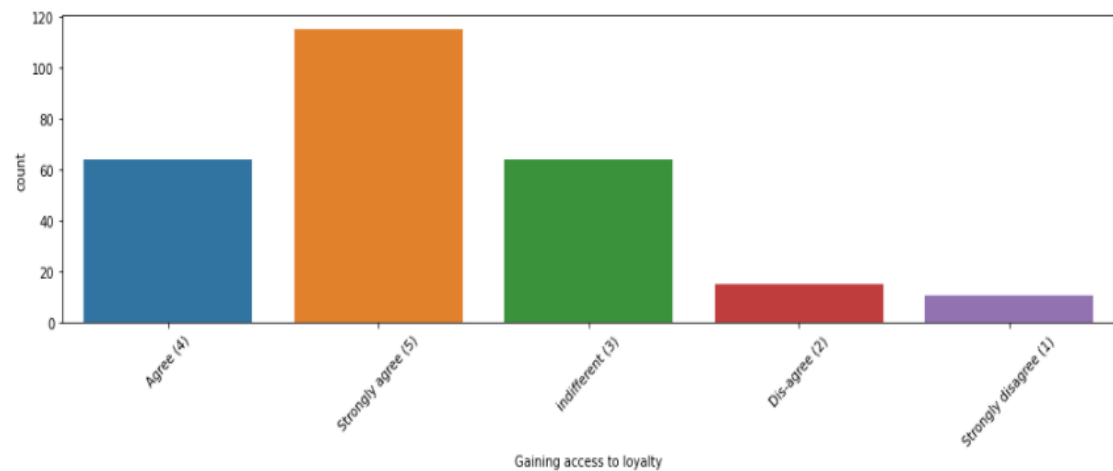

Observations:

1- Maximum near around 43% of customers are strongly agree by gaining access to loyalty.

2- Only near around 4% of customers are strongly disagree by gaining access to loyalty.

13- Count plot for "User satisfaction cannot exist without trust" feature:

```
Strongly agree (5)        122
Agree (4)                 117
Strongly disagree (1)      18
Dis-agree (2)               7
indifferent (3)             5
Name: User satisfaction cannot exist without trust, dtype: int64

<AxesSubplot:xlabel='User satisfaction cannot exist without trust', ylabel='count'>
```
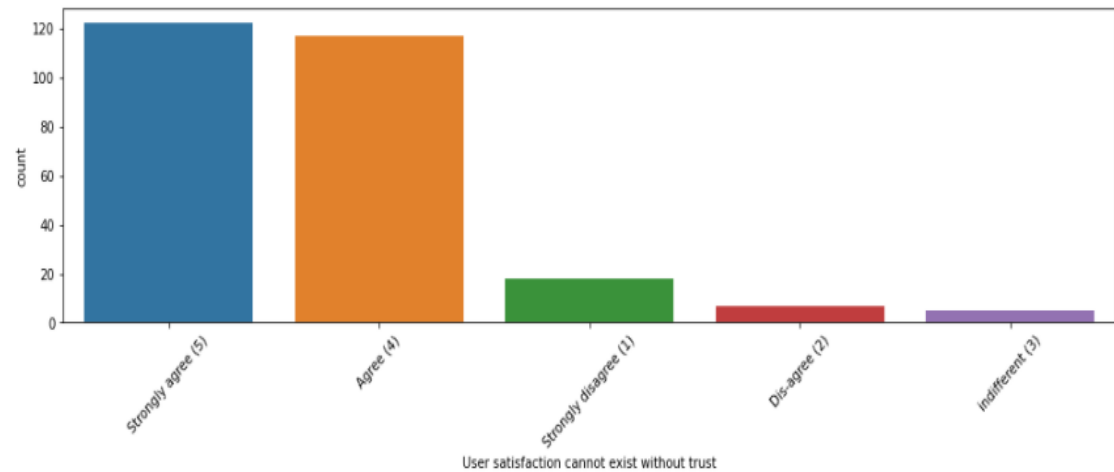

Observations:

1- Maximum near around 45.35% of customers are strongly agree with user satisfaction cannot exist without trust.

2- Only near around 2% of customers have indifferent opinion about user satisfaction and trust.

14- Count plot for " Provision of complete and relevant product information" feature:

```
Strongly agree (5)     135
Agree (4)               98
indifferent (3)         31
Disagree (2)             5
Name: Provision of complete and relevant product information, dtype: int64

<AxesSubplot:xlabel='Provision of complete and relevant product information', ylabel='count'>
```
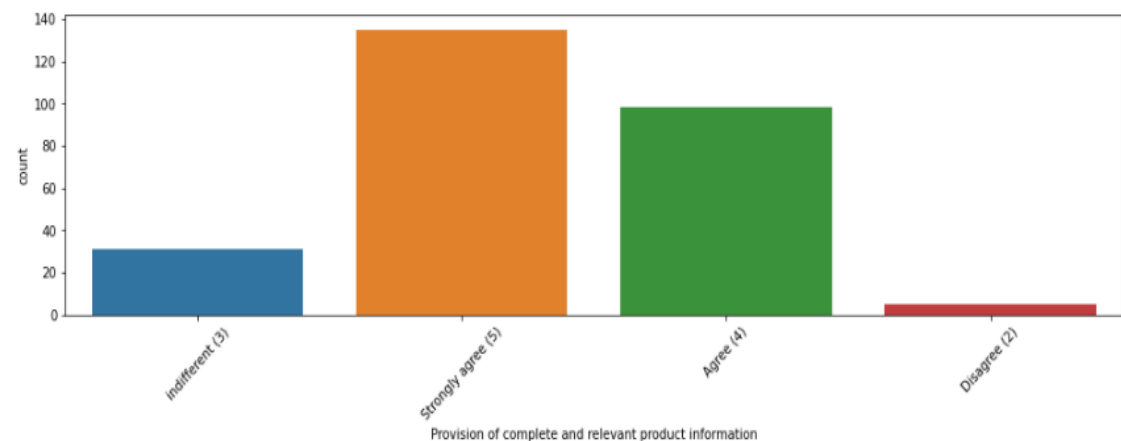

Observations:

1- Maximum near around 50% of customers are strongly agree with Provision of complete and relevant product information.

2- Only near around 2% of customers are disagree with Provision of complete and relevant product information.

15- Count plot for "The Convenience of patronizing the online retailer" feature:

```
Agree (4)           138
indifferent (3)      77
Strongly agree (5)   54
Name: The Convenience of patronizing the online retailer, dtype: int64

<AxesSubplot:xlabel='The Convenience of patronizing the online retailer', ylabel='count'>
```


The Convenience of patronizing the online retailer
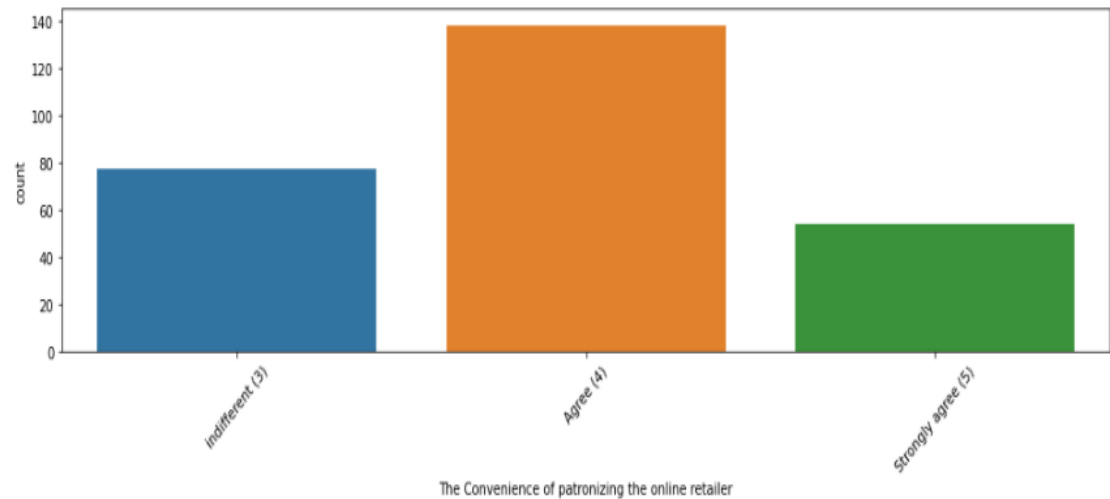
Observations:

1- Maximum near around 51% of customers are agree with the convenience of patronizing the online retailer.

2- Only near around 29% of customers have indifferent opinion about the convenience of patronizing the online retailer.

16- Count plot for "gratification shopping on your favourite e-tailer" feature:

```
indifferent (3)        101
Strongly agree (5)      65
Agree (4)               63
Disagree (2)            22
Strongly disagree (1)   18
Name: gratification shopping on your favorite e-tailer, dtype: int64

<AxesSubplot:xlabel='gratification shopping on your favorite e-tailer', ylabel='count'>
```


gratification shopping on your favorite e-tailer
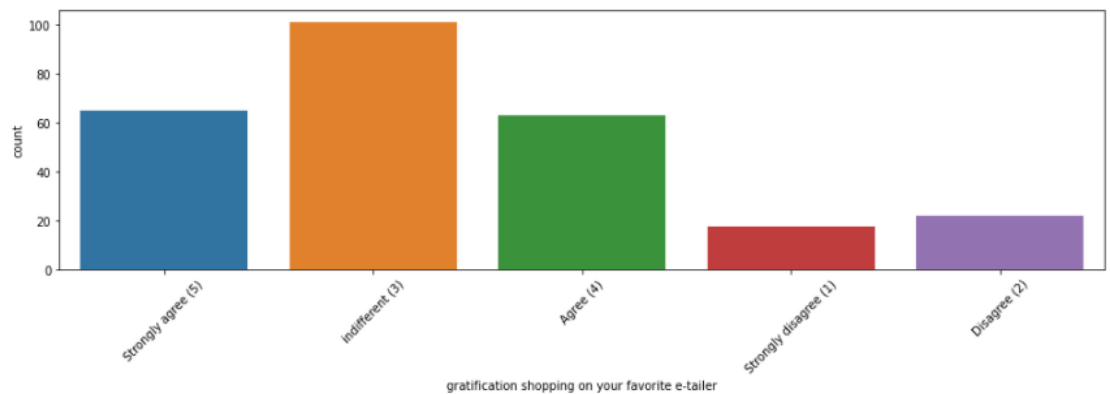
Observations:

1- Maximum near around 37.54% of customers have indifferent opinion about gratification shopping on favorite e-tailer.

2- Near around 23% of customers are agree with gratification shopping on favorite e-tailer.

3- Only near around 7% of customers are strongly disagree with gratification shopping on favorite e-tailer.

17- Count plot for "enhances your social status" feature:

```
indifferent (3)        100
Agree (4)               59
Strongly agree (5)      48
Strongly disagree (1)   33
Dis-agree (2)           29
Name: enhances your social status, dtype: int64

<AxesSubplot:xlabel='enhances your social status', ylabel='count'>
```


Observations:

1- Maximum near around 37% of customers have indifferent opinion about social status enhancement.
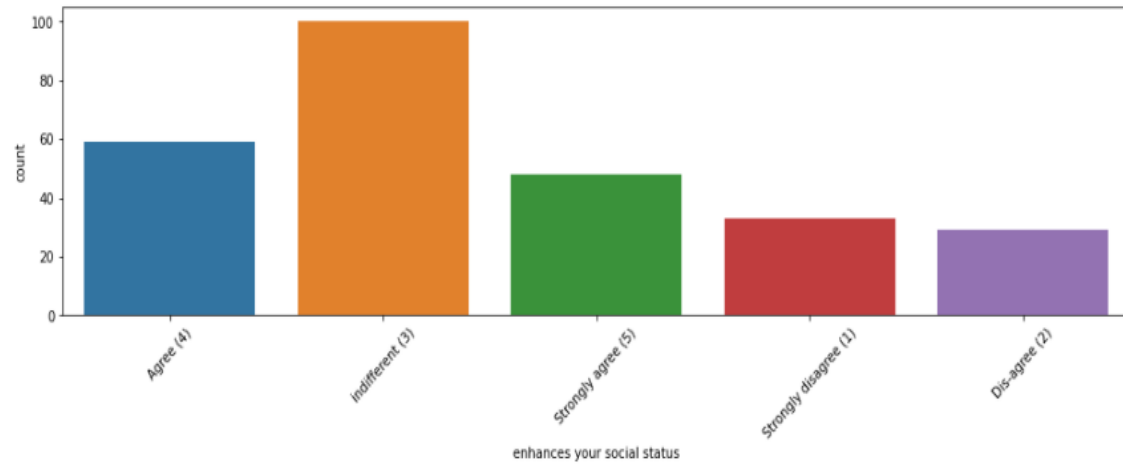
2- Near around 22% of customers are agree with social status enhancement.

3- Only near around 11% of customers are disagree with social status enhancement.

18- Count plot for "Shopping on the website helps you fulfil certain roles" feature:

```
Agree (4)               88
indifferent (3)         88
Strongly agree (5)      38
Strongly disagree (1)   33
Dis-agree (2)           22
Name: Shopping on the website helps you fulfill certain roles, dtype: int64

<AxesSubplot:xlabel='Shopping on the website helps you fulfill certain roles', ylabel='count'>
```
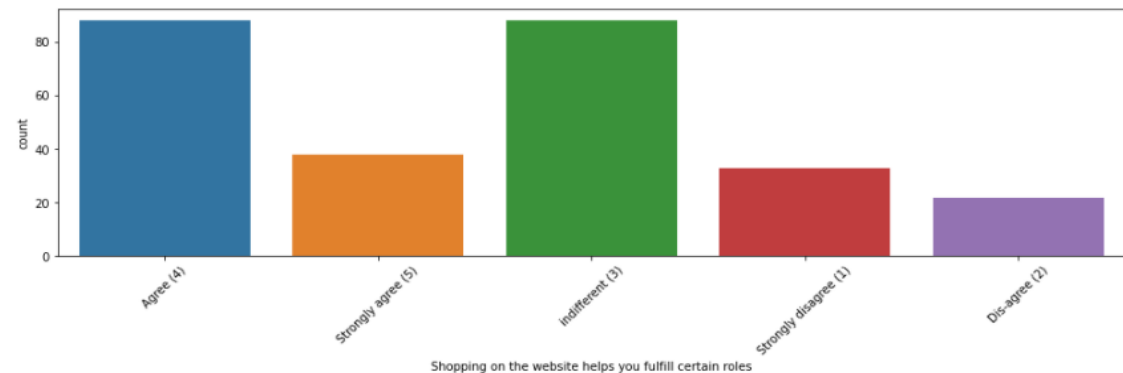

Observations:

1- Maximum near around 33% of customers are agree with Shopping on those website which helps us fulfill certain roles.

2- Also, Near around 33% of customers have indifferent opinion about Shopping on those website which helps us fulfill certain roles.

3- Only near around 8% of customers are disagree with Shopping on those website which helps us fulfill certain roles.

Websites/online Shopping Platforms Analysis: -

1- Bar plot for "Easy to use website or application" feature:

```
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com      64
Amazon.in, Flipkart.com                                           44
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.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
```
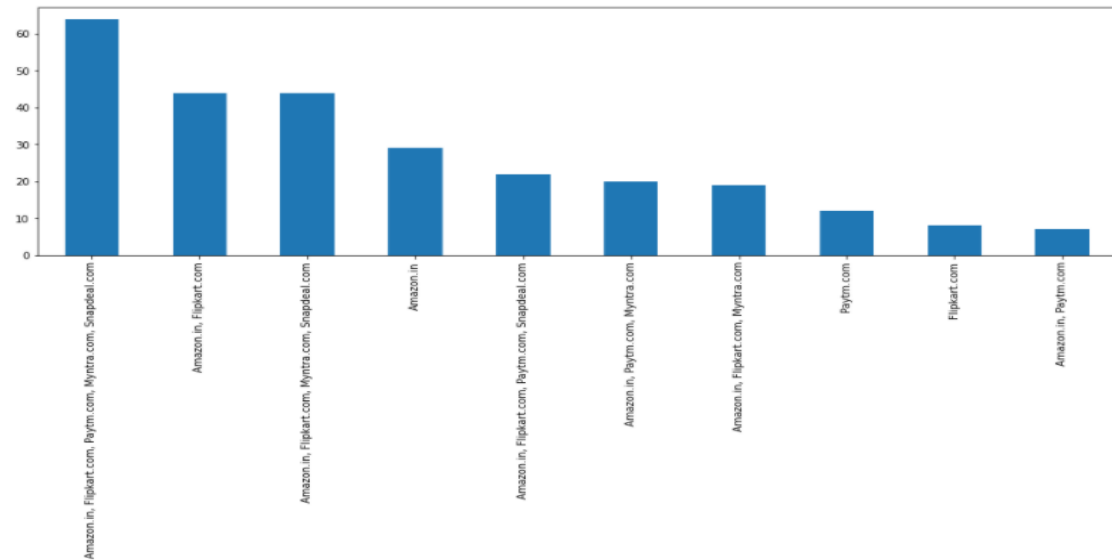

Observations:

1- Maximum near around 23.80% of customers are using Amazon.in, Flipkart.com, Paytm.com, Myntra.com & Snapdeal.com.

2- Due to user friendly interface and support system, these websites are most popular among customers.

3- Amazon.in is the most popular website as per above graph.

2- Bar plot for "Wild variety of product on offer" feature:

```
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

<AxesSubplot:>
```
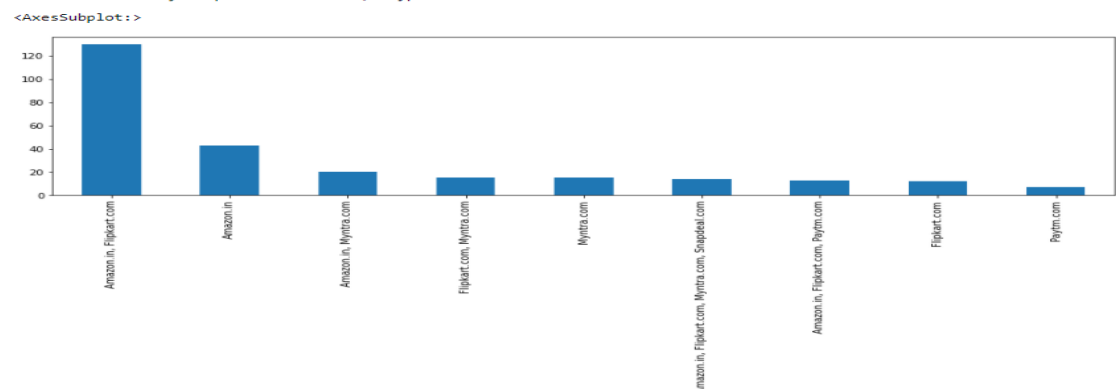
Observations:

1- Maximum near around 48.32% of customers are using Amazon.in & Flipkart.com because these websites have Wild variety of product on offer.

2- Amazon.in is the most popular website in this analysis also.

3- Bar plot for "Reliability of the website or application" feature:

```
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: int64
```

<AxesSubplot:>
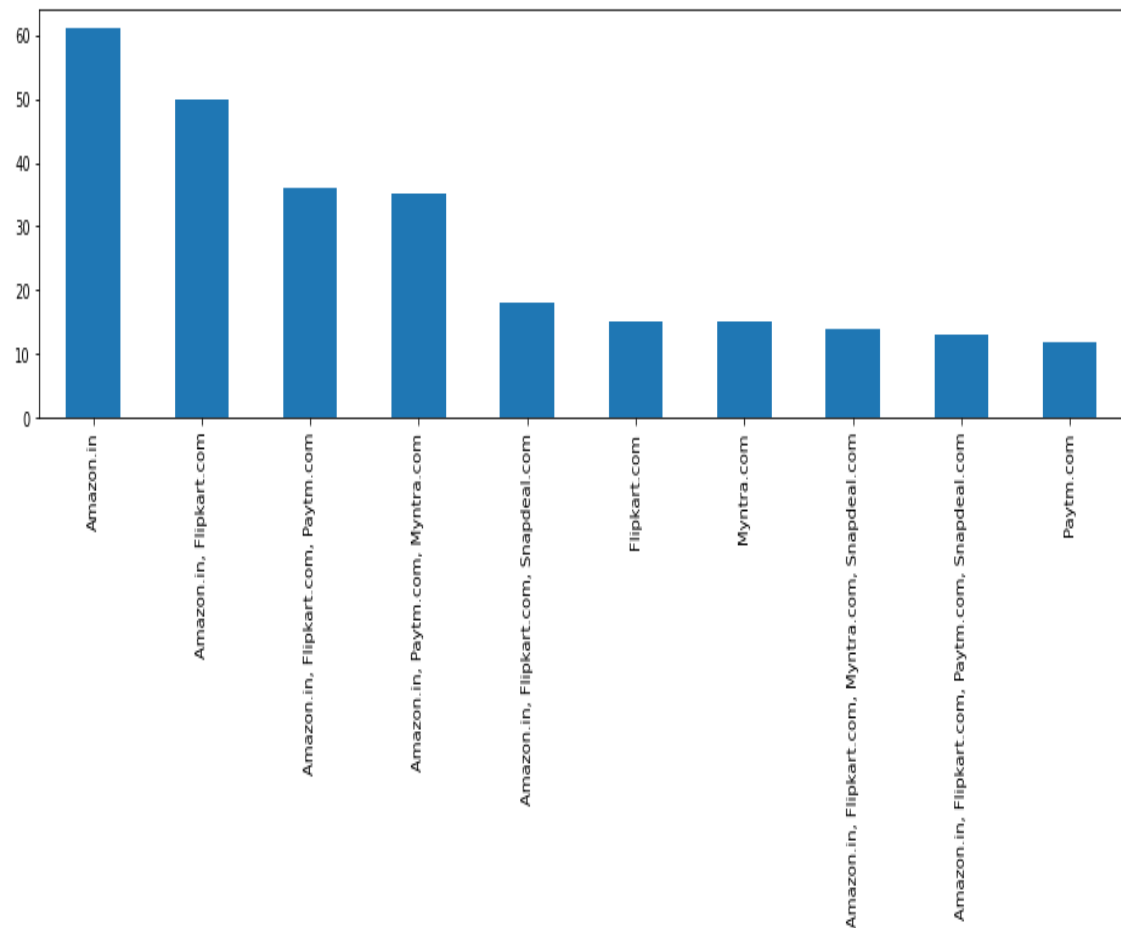
Observations:

1- Maximum near around 22.67% of customers are using Amazon.in due to reliability of the website or application.

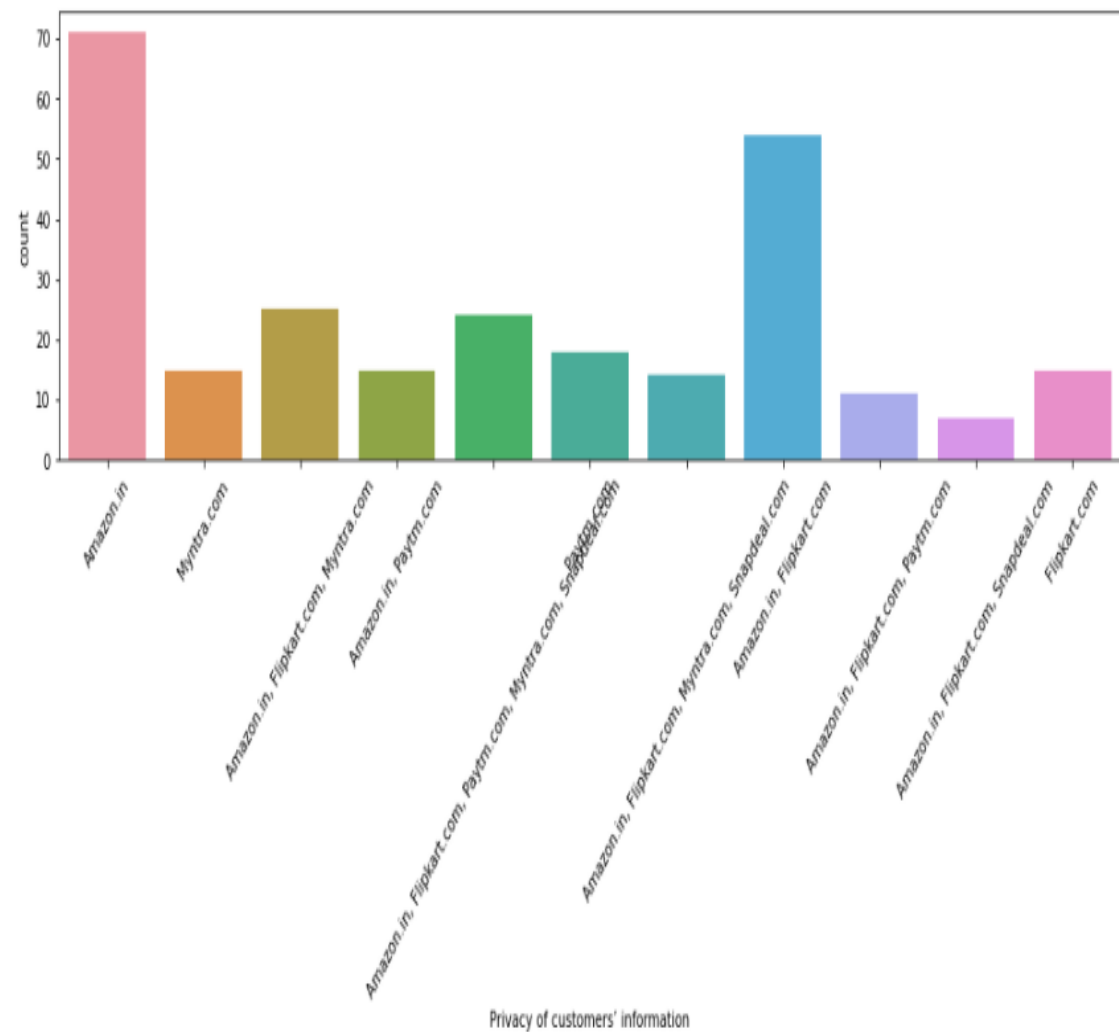2- Amazon.in is the most popular website in this analysis also.

4- Count plot for "Privacy of customers' information" feature:

```
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
Flipkart.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
```

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

Observations:

1- Maximum near around 26.39% of customers are using Amazon.in due to privacy of customers' information.

2- Amazon.in is the most popular website in this analysis also.

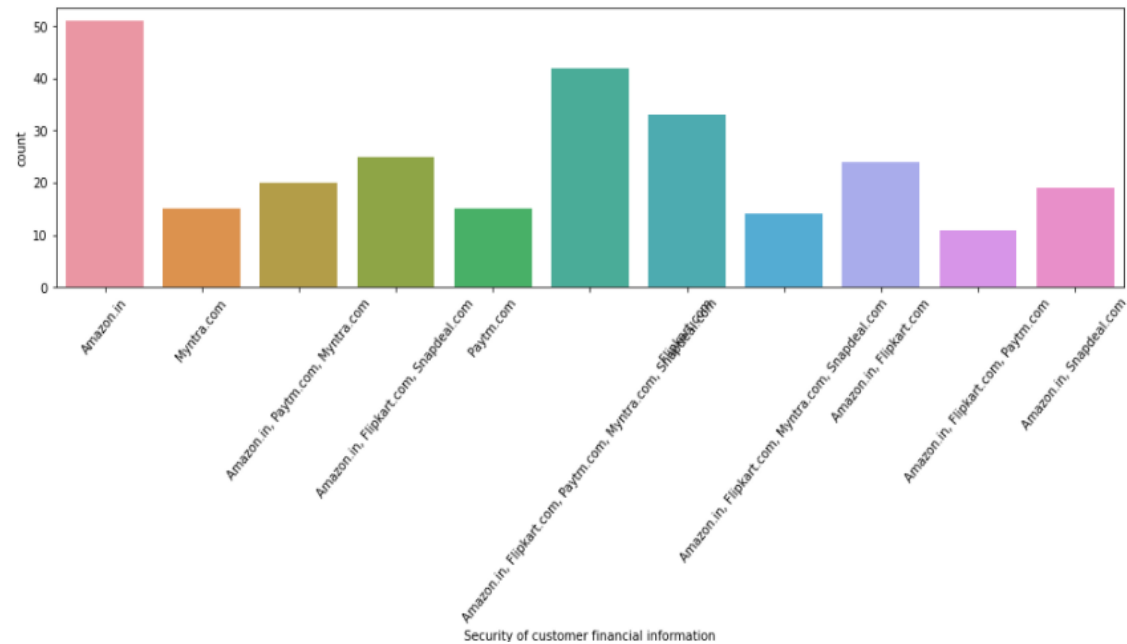## 5- Count plot for "Security of customer financial information" feature:

```
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
```

`<AxesSubplot:xlabel='Security of customer financial information', ylabel='count'>`

Observations:

1- Maximum near around 19% of customers are using Amazon.in due to Security of customer financial information.

2- Amazon.in is the most popular website in this analysis also.

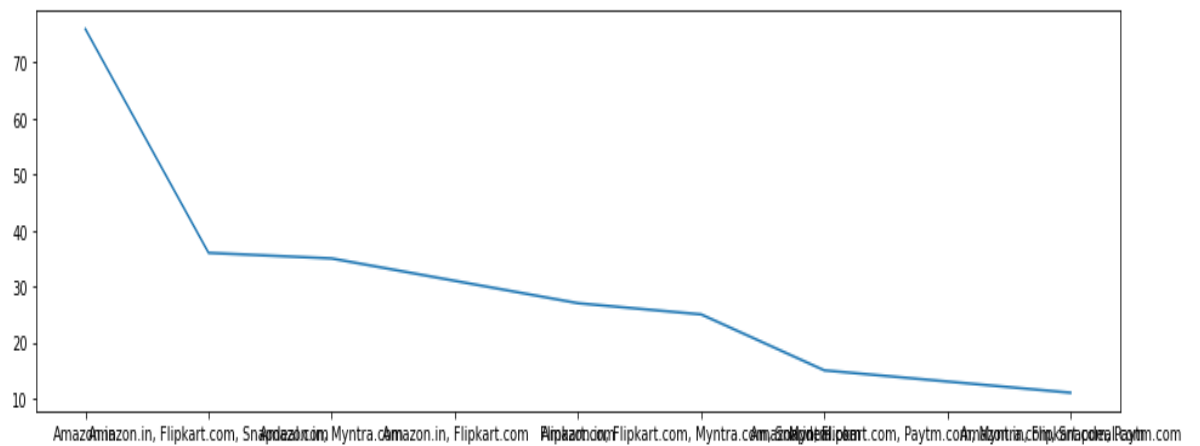## 6- Line plot for "Perceived Trustworthiness" feature:

```
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
```
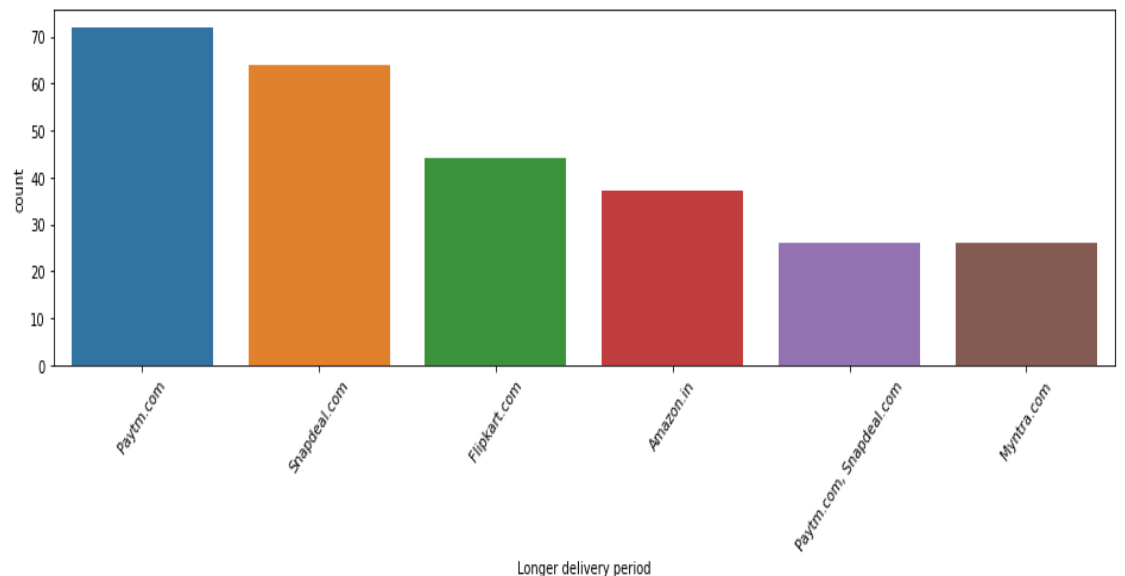
`<AxesSubplot:>`

Observations:

1- Maximum near around 28% of customers are using Amazon.in due to Perceived Trustworthiness.

2- Amazon.in is the most popular website in this analysis also.

7- Count plot for "Longer delivery period" feature:

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

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

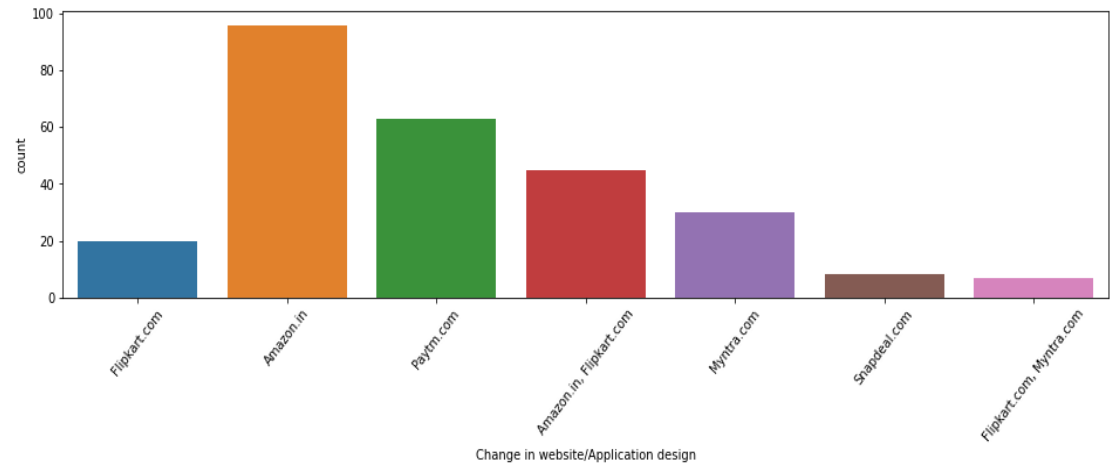
Observations:

1- Paytm.com has longest delivery period in this list.

2- Myntra.com has minimum delivery period in this list.

8- Count plot for "Change in website/Application design" feature:

```
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'>
```
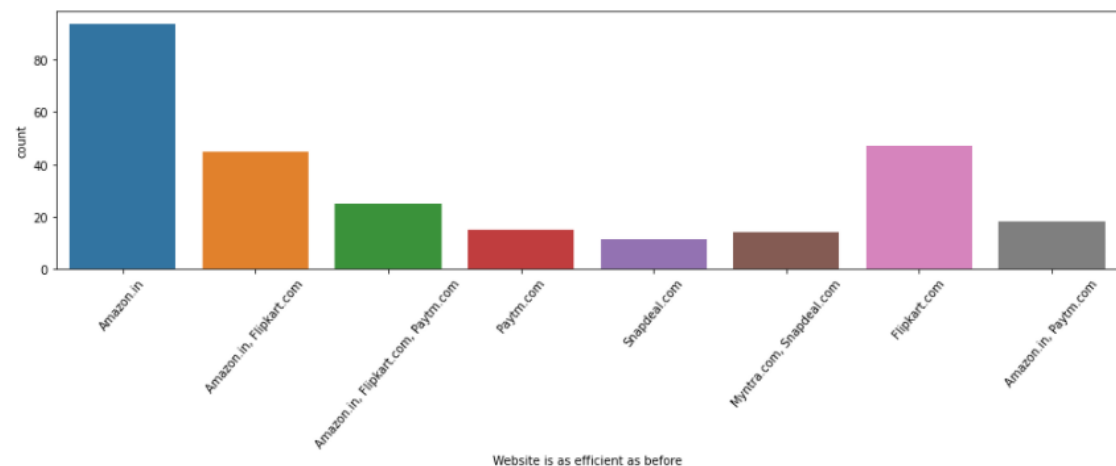

Observations:

1- Maximum near around 35.68% of customers are using Amazon.in due to Change in website/Application design.

2- Amazon.in is the most popular website in this analysis also.

9- Count plot for "Website is as efficient as before" feature:

```
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, dtype: int64

<AxesSubplot:xlabel='Website is as efficient as before', ylabel='count'>
```
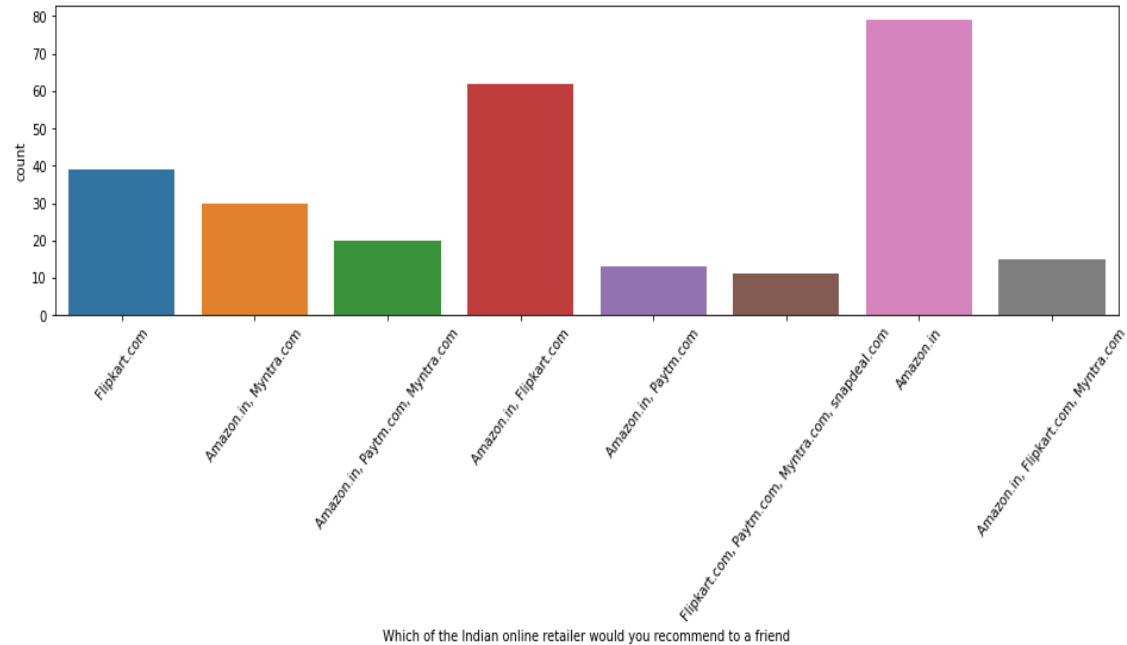

Observations:

1- Maximum near around 34.94% of customers are using Amazon.in website due to it's good efficiency.

2- Amazon.in is the most popular website in this analysis also.

10- Count plot for "Which of the Indian online retailer would you recommend to a friend" feature:

```
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'>`

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

Observations:

1- Maximum near around 29.37% of customers are using Amazon.in website due to its recommendation facilities.

2- Amazon.in is the most popular website in this analysis also.

• Software & Hardware Requirements and Tools Used-

- We have used Python and Jupyter Notebook to compute the majority of this case study. Packages and libraries used in this project are: -

- Pandas (data analysis)

- NumPy (matrix computation)

- Matplotlib (Visualization)

- Seaborn (visualization)

- Warnings (filter warnings) & etc.

- **Software Tools used in this project are as**: -

- 1- Anaconda3 (64-bit)

- 2- Jupyter Notebook 6.1.4

- 3- Python 3.8

- 4- MS-Office 2019 (Excel, Word, Power point)

- 5- Notepad

- 6- Google Chrome Web Browser

- **Hardware used**: - 4GB RAM, Intel I3 Processor.

- **System Software**: 64Bit O/S Windows 10(x64-based processor).

Libraries & Packages used:

We have used Python and Jupyter Notebook to compute the majority of this project. For analysis, visualization, statistics & evaluation, we have used these: -

1- Pandas (data analysis)
2- NumPy (matrix computation)
3- Matplotlib (Visualization)
4- Seaborn (visualization)
5- Warnings (filter warnings) & etc. Microsoft Excel (for calculations and Data Handling).

Steps used in this project: -

1- Define the Problem

2- Data Gathering

3- Data Cleaning

4- Data Exploration and Visualization

5- Descriptive Analysis

6- Finding conclusions in detailed data analysis & etc.

Conclusions on basis of Customer Related Information: -

1) In people doing online shopping Females are more as compared to Males.

2) People in the age group from 21 to 50 are doing more online shopping as compared to other age groups.

3) Delhi Topped in the city list where more number of people doing online shopping.

4) Mostly people are using mobile internet while shopping online.

5) Mostly People are doing online shopping through smartphone.

6) People having windows as operating system (OS) have maximum count.

7) Mostly people are choosing google chrome as browser to access the website

8) People are mostly use search engine to arrive at their favourite online store for the first time.

9) After first visit people are using application or search engine to access website.

Conclusions From Customer Feedback Analysis: -

o Website has user friendly interface.

o Payment method Should be convenient.

o Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc) is necessary.

o Online shopping gives monetary benefit and discounts to the customers.

o Displaying quality Information on the website improves satisfaction of customers.

o Return and replacement policy of the e-tailer is important for purchase decision.

o Doing online shopping give them value of money spend.

o Shopping on the website give them sense of adventure.

o Being able to guarantee the privacy of the customer

o User derive satisfaction while shopping on a good quality website or application.

o Shopping online is convenient and flexible.

- Gaining access to loyalty programs is a benefit of shopping online.

- User satisfaction exist when they have trust one website.

- There should be provision of complete and relevant product information website.

- They feel gratification shopping on your favourite e-tailer.

- The Convenience of patronizing the online retailer.

- User have indifferent opinion that shopping on their preferred e-tailer enhances their social status.

- User have mixed reaction on whether shopping on website helps them to fulfil certain roles some agree it some are indifferent.

Conclusions from Detailed Data Analysis: -

- As in the final count 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.

- They are some cons like the amazon website topping the list In Frequent disruption when moving from one page to another this con company should see and improve it to give overall best experience to the users.

||Thank you||