CUSTOMER RETENTION

PROJECT REPORT

SUBMITTED BY:-

Joshua Paul

ACKNOWLEDGMENT

I have taken efforts in this project. However, it would not have been possible without the kind support and help of many individuals and organizations. I would like to extend my sincere thanks to all of them. I am highly indebted to Flip Robo Technologies for their guidance and constant supervision as well as for providing necessary information regarding the project & also for their support in completing the project. I want to thank my SME Mr. Shubham Yadav for providing the Dataset and helping us to solve the problem and addressing out our Query in right time. I would like to express my gratitude towards my parents & members of Flip Robo for their kind co-operation and encouragement which help me in completion of this project. I would like to express my special gratitude and thanks to industry persons for giving me such attention and time.

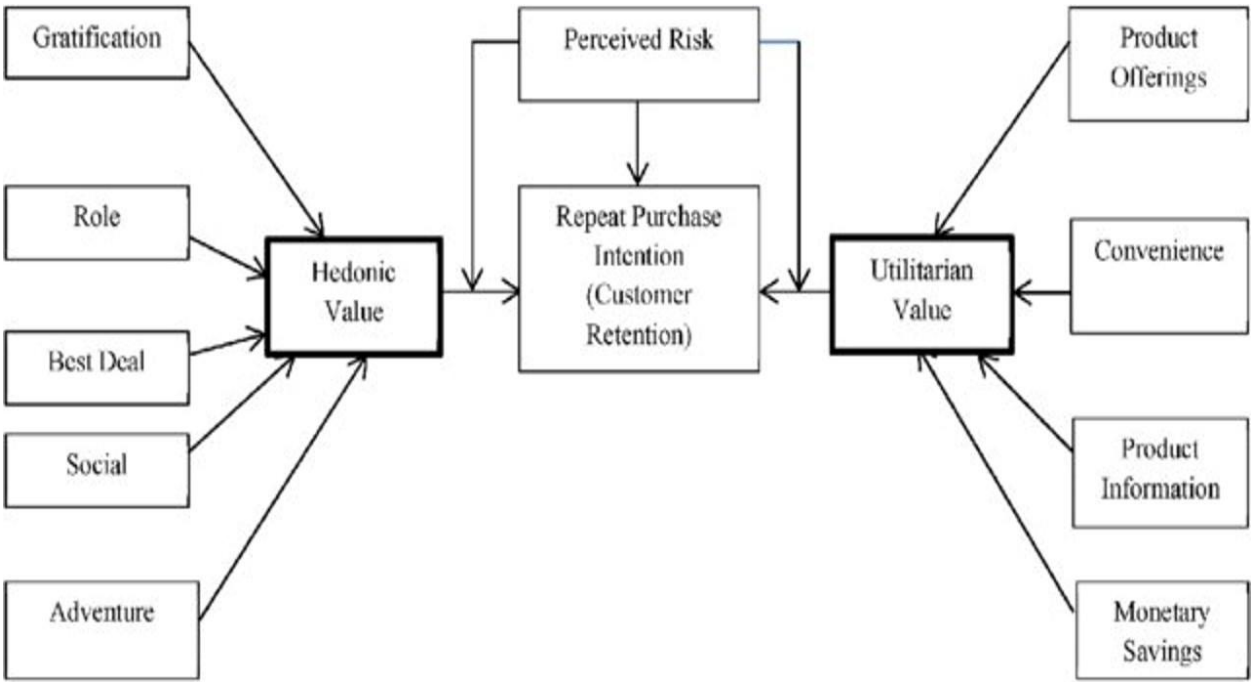
INTRODUCTION

Customer satisfaction has emerged as one of the most important factors that guarantee the success of the 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 online customers repeat purchase intention. The combination of both utilitarian value and hedonistic values is needed to affect the repeat purchase intention (loyalty) positively. The data is collected from Indian online shoppers. Results indicate the e-retail success factors, which are very much critical for customer satisfaction. By increasing the utilitarian value and hedonistic values derived by the customers, customer satisfaction and hence the customer's repeat purchase intention can be increased significantly.

OBJECTIVE

Our major goal in performing this project is to evaluate if people are purchasing items from e-commerce websites, how they gave feedback to these websites based on many positive and negative aspects, and also the details of the users based on factors such as age, gender, and so on.

The Hedonic value consists of factors like Gratification, Role, Best Deal, Social and Adventure. The Utilitarian value consists of factors like Product Offerings, Convenience, Product Information and Monetary Savings. Customer Retention is based on 3 factors, according to the above diagram. They are: Perceived Risk, Hedonic value and Utilitarian value


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

Hardware and Software Requirements

The hardware utilized for this project is a laptop with high-end specifications and a steady internet connection. When it came to the software, I utilized anaconda navigator and Jupyter notebook to conduct my Python programming and analysis. Microsoft Excel is required to use an excel file. In Jupyter notebook, I utilized several Python libraries to complete this project, which I have listed below with appropriate substantiation: 1. Pandas - It is a library that is used to read data, visualize it, and analyze it. 2. NumPy- utilized for dealing with arrays and different mathematical methods. 3. Seaborn- a visualization tool for plotting many sorts of plots. 4. Matplotlib- It provides an object-oriented API for embedding plots into applications. Data sources and their formats. The data was provided to us in an excel file by a highly confidential firm. They had also supplied the problem statement, which explained what they needed from us as well as the criteria that had to be met. Let's look at the data now. I've attached a snapshot to give you an idea of what I'm talking about.

```
In [1]: import pandas as pd
        import numpy as np
```

```
In [4]: df=pd.read_csv('customer_retention_dataset.csv')
```

```
In [5]: df.head()
```

Out[5]:

	1Gender of respondent	2 How old are you?	3 Which city do you shop online from?	4 What is the Pin Code of where you shop online from?	5 Since How Long You are Shopping Online ?	6 How many times you have made an online purchase in the past 1 year?	7 How do you access the internet while shopping on-line?	8 Which device do you use to access the online shopping?	9 What is the screen size of your mobile device? \t\t\t\t\t\t	10 What is the operating system (OS) of your device? \t\t\t\t	...	Longer time to get logged in (promotion, sales period)	Longer time in displaying graphics and photos (promotion, sales period)	Late declaration of price (promotion, sales period)	
0	Male	31-40 years	Delhi	110009	Above 4 years	31-40 times	Dial-up	Desktop	Others	Window/windows Mobile	...	Amazon.in	Amazon.in	Flipkart.com	
1	Female	21-30 years	Delhi	110030	Above 4 years	41 times and above	Wi-Fi	Smartphone	4.7 inches	IOS/Mac	...	Amazon.in, Flipkart.com	Myntra.com	snapdeal.com	S
2	Female	21-30 years	Greater Noida	201308	3-4 years	41 times and above	Mobile Internet	Smartphone	5.5 inches	Android	...	Myntra.com	Myntra.com	Myntra.com	
3	Male	21-30 years	Karnal	132001	3-4 years	Less than 10 times	Mobile Internet	Smartphone	5.5 inches	IOS/Mac	...	Snapdeal.com	Myntra.com, Snapdeal.com	Myntra.com	
4	Female	21-30 years	Bangalore	530068	2-3 years	11-20 times	Wi-Fi	Smartphone	4.7 inches	IOS/Mac	...	Flipkart.com, Paytm.com	Paytm.com	Paytm.com	

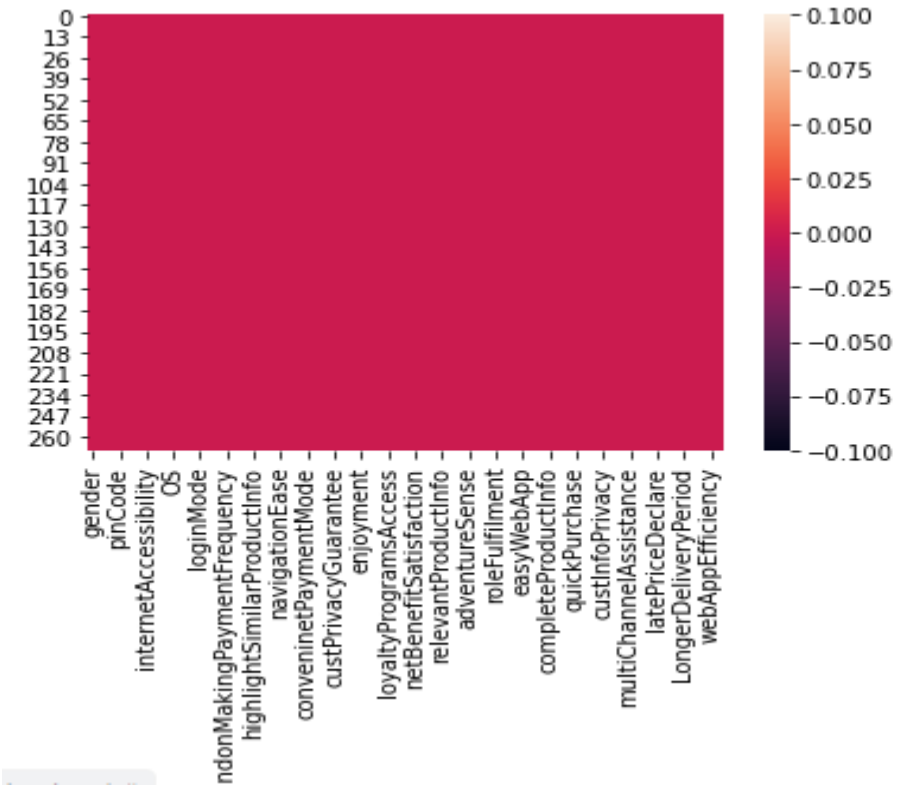
5 rows × 71 columns

we can see that there is a numerical and categorical data are both,We won't face any kind of problem with numerical data ,but categorical data should be convert into numerical as we can not procced with categorical data.

- This dataset has 269 rows and 71 columns.
- Our goal is to explore data insights and conduct in-depth data analysis.

Visualizations

Check for any empty values



Categorizing

```
# categorical_columns are avaialable from column index 0 to 16
categorical_columns=data.iloc[:,:17].columns
categorical_columns=categorical_columns.to_list()
print('categorical_columns shape:', data[categorical_columns].shape)


# rating_columns are avaialable from column index 17 to 46
rating_columns=data.iloc[:,17:47].columns
rating_columns=rating_columns.to_list()
print('\nrating_columns shape:      ', data[rating_columns].shape)


# ecommerce_columns are avaialable from column index 47 and above
ecommerce_columns=data.iloc[:,47:].columns
ecommerce_columns=ecommerce_columns.to_list()
print('\necommerce_columns shape:   ', data[ecommerce_columns].shape)
```

```
categorical_columns shape: (269, 17)

rating_columns shape:      (269, 30)

ecommerce_columns shape:   (269, 24)
```

Exploratory Data Analysis (EDA)

- Categorical columns
- Ratings columns
- Ecommerce columns

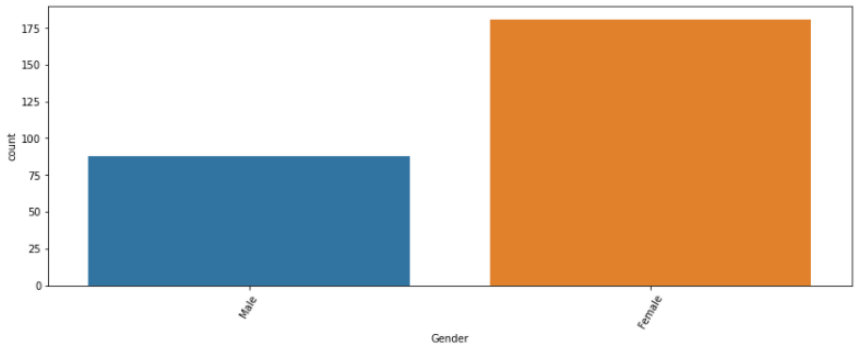
GENDER

In [20]:
```
#Let's see the plot for all the categorical parameters

for col in categorical_columns:
    print('\t'*6,'***',col,'***','\n')
    print(df[col].value_counts())
    plt.figure(figsize=(14,5))
    sns.countplot(x=df[col])
    plt.xticks(rotation=60)
    plt.show()
    print('Observation: \n')
    input('')
    print('-----'*20)
    print('\n')
```

```
                                        *** Gender ***

Female    181
Male       88
Name: Gender, dtype: int64
```



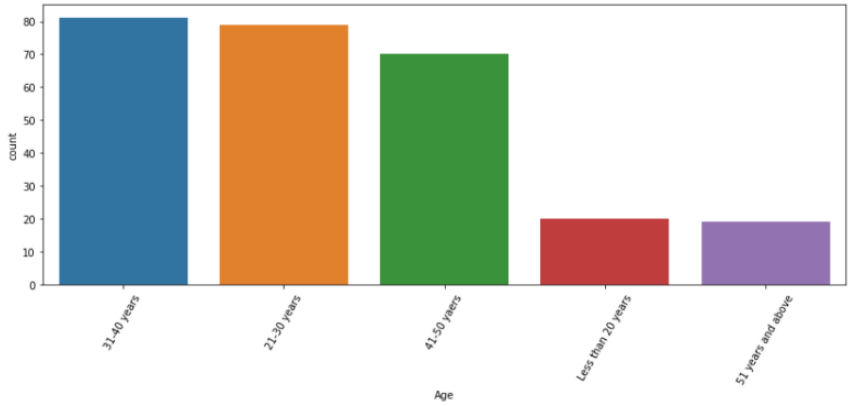

```
Observation:

- Out of the total, more than half customers are female ----------------------------------------------------------------------------------------
```

AGE

```
*** Age ***

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
```



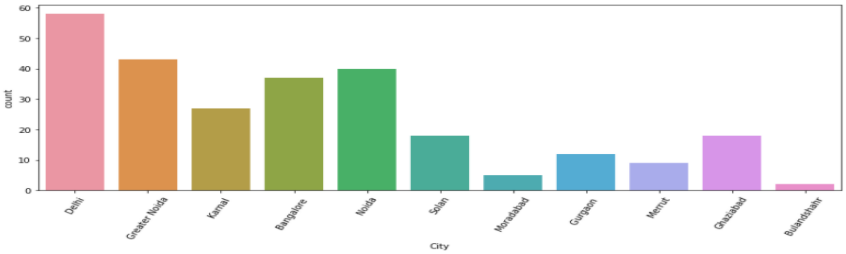
Observation:

- Highest online shopping are done by the Age group of (31-40) years ----------

CUTOMER FROM DIFFERNET CITIES AND PINCODE

```
*** City ***

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

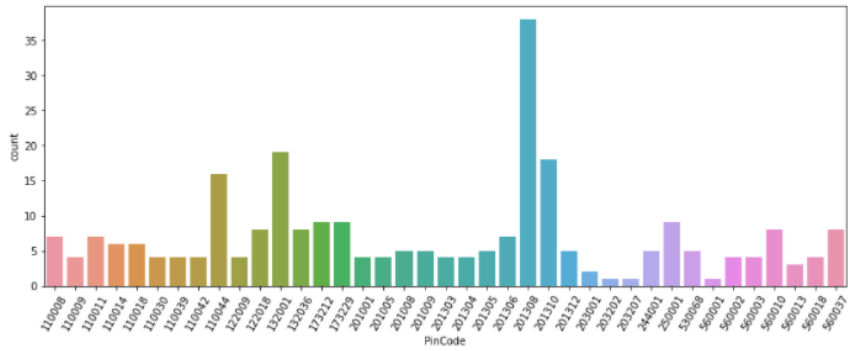


Observation:

- "Delhi" contains highest number of participants then comes "Greater Noida", "Noida", "Bangalore" .Lowest are in "Bulandshahr" ---------

*** PinCode ***

```
201308    38
132001    19
201310    18
110044    16
173229     9
173212     9
250001     9
122018     8
560037     8
132036     8
560010     8
110011     7
110008     7
201306     7
110014     6
110018     6
201008     5
201009     5
201305     5
201312     5
244001     5
530068     5
201005     4
110009     4
110042     4
110039     4
110030     4
201304     4
122009     4
201303     4
560018     4
201001     4
560003     4
560002     4
560013     3
203001     2
203207     1
560001     1
203202     1
Name: PinCode, dtype: int64
```



Observation:

- Highest number of participants are from this pin location: "201306" and rest location also contains more than 2 counts ----------

Internet Accessibility

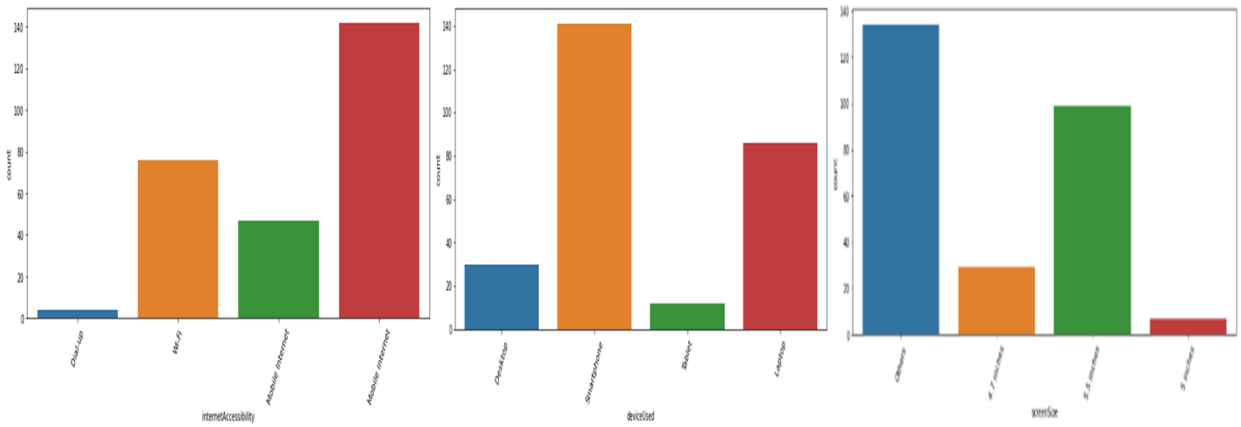
- Most of the participants use their Mobile internet to access products online/shop online

Device Used

- Large number of participants use their Smartphone to shop online and we also have a good number of participants who go with purchases by Laptop

Screen size

- Participants use their 5.5 inches of screen size to shop online largely, and here we can see that in our dataset contains most of the participants use



OPERATING SYSTEM USED

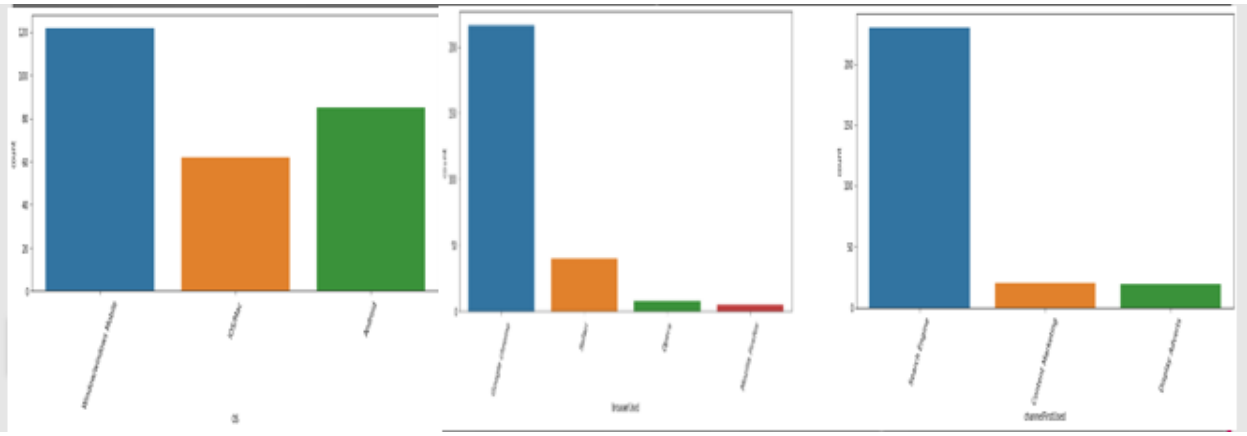
- Most participants uses "Window/windows Mobile" to shop online then, "Android" then, comes "IOS/Mac" at least

Browser Used

- "Google Chrome" browser are widely used for online website access when compared to other browser

channelFirstUsed

- "Search Engine" were widely used to follow to arrive at their favourite online store for the first time. And rest 50-50% "Content Marketing" and "Display Adverts" are used



Ratings Columns Visualization

Rating_columns

In [29]:
```
# Pie charts for some of the features

for i in rating_columns:
    plt.subplots()
    plt.pie(x=df[i].value_counts(), labels=df[i].value_counts().index,
            data=df, shadow=True,
            startangle=60, autopct='%1.1f%%',
            colors=['cyan', 'orange','magenta','limegreen', 'red', 'crimson'],
            wedgeprops = {'linewidth': 4})
    plt.setp(plt.title(i,fontsize=15,color='darkred'),color='blue',style='italic')
    plt.axis('equal')  # Equal aspect ratio ensures that pie is drawn as a circle.
    plt.show()
    print("----------"*10)
    print('\n')
    #break


#Let's also use the coded data for this part

coded_df = pd.read_csv('customer_retention_dataset.csv')
data = pd.read_csv('customer_retention_dataset.csv','datasheet')

rating_cols=df.iloc[:,17:47].columns
rating_cols=rating_cols.to_list()

ratings = {1: 'Strongly disagree', 2: 'Disagree', 3: 'Neither agree nor disagree', 4: 'Agree', 5: 'Strongly agree'}


# Rename the values of the dataframe
for col in rating_cols:
    coded_df.replace({col: ratings},inplace=True)
```

Observations on the basis of Customer's perception

- ❑ Customers mostly believe that the content on the website must be easy to read and understand
- ❑ People agree that information on similar product to the one highlighted is important for product comparison
- ❑ 70% people believe that Complete information on listed seller and product is important for purchase decision.
- ❑ 90% people believe that All relevant information on listed products must be stated clearly
- ❑ For more than 90% of the people believe that the following parameters are important

  - Ease of navigation in website
  - Loading and processing speed
  - User friendly Interface of the website

- Convenient Payment methods
- Trust that the online retail store will fulfill its part of the stipulated time

transaction at the

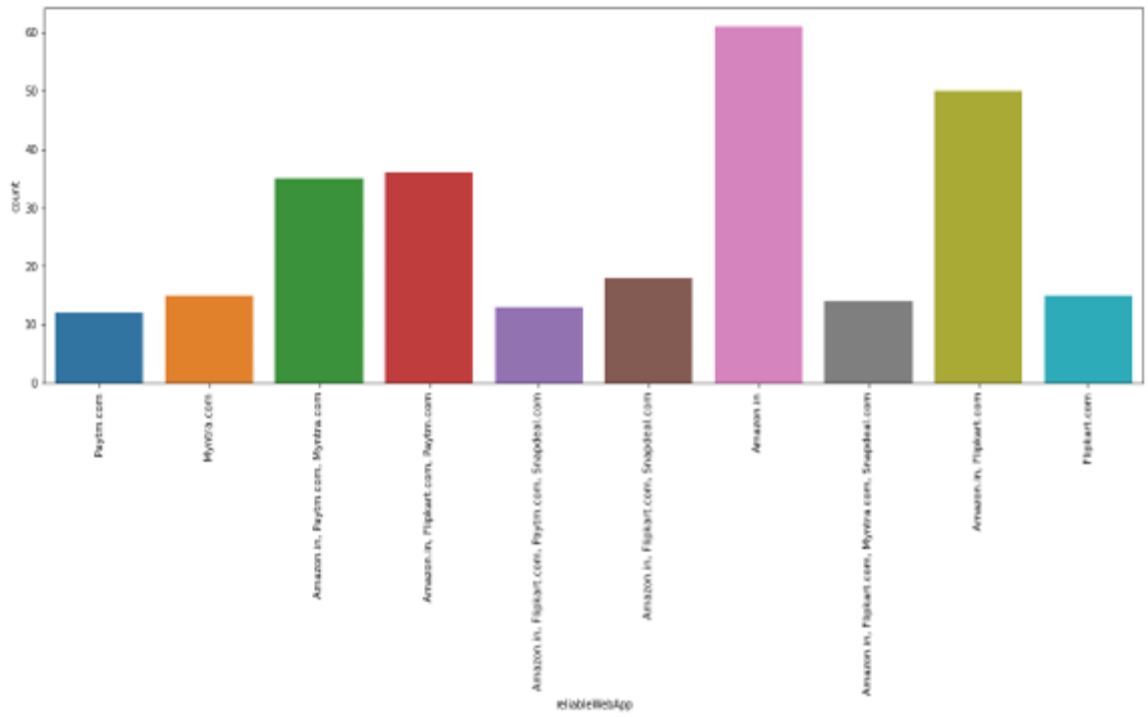
- Empathy (readiness to assist with queries) towards the customers
- Being able to guarantee the privacy of the customer
- Responsiveness, availability of several communication rep, twitter, phone etc.)

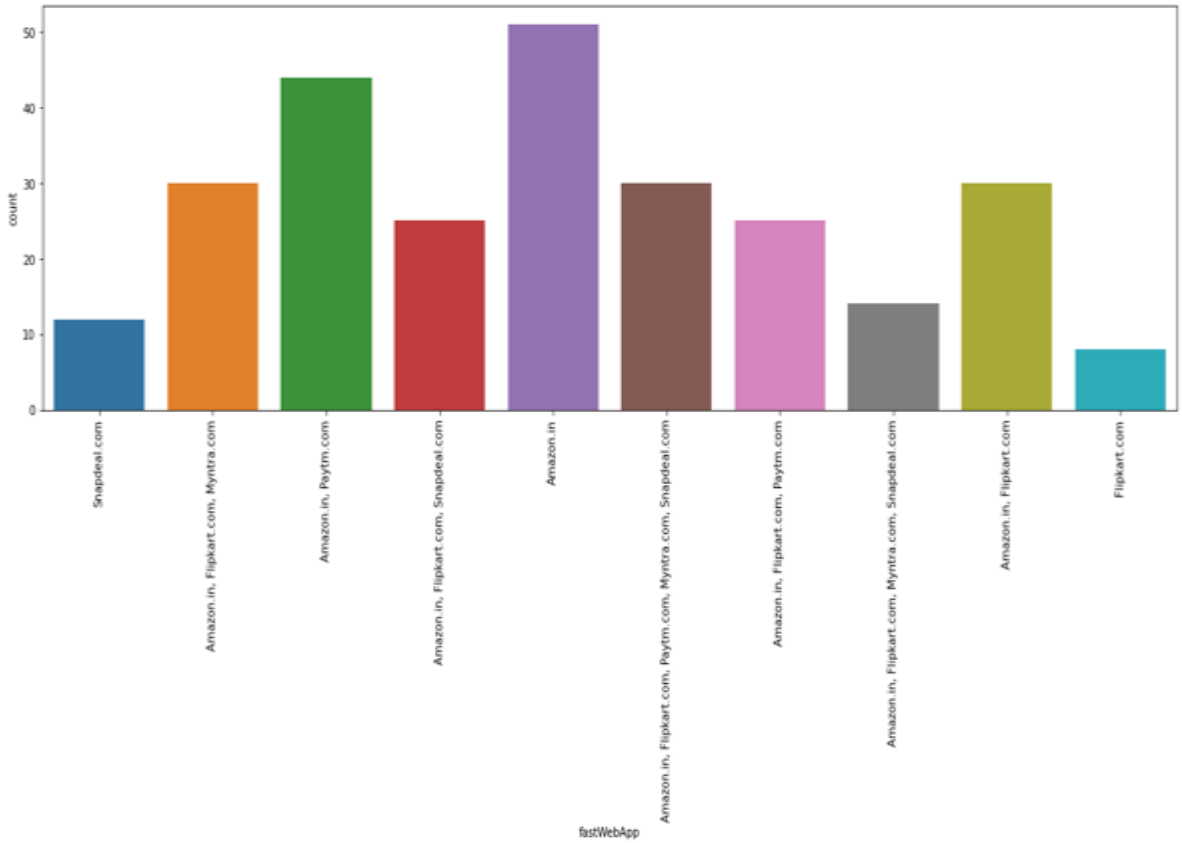
channels (email, online

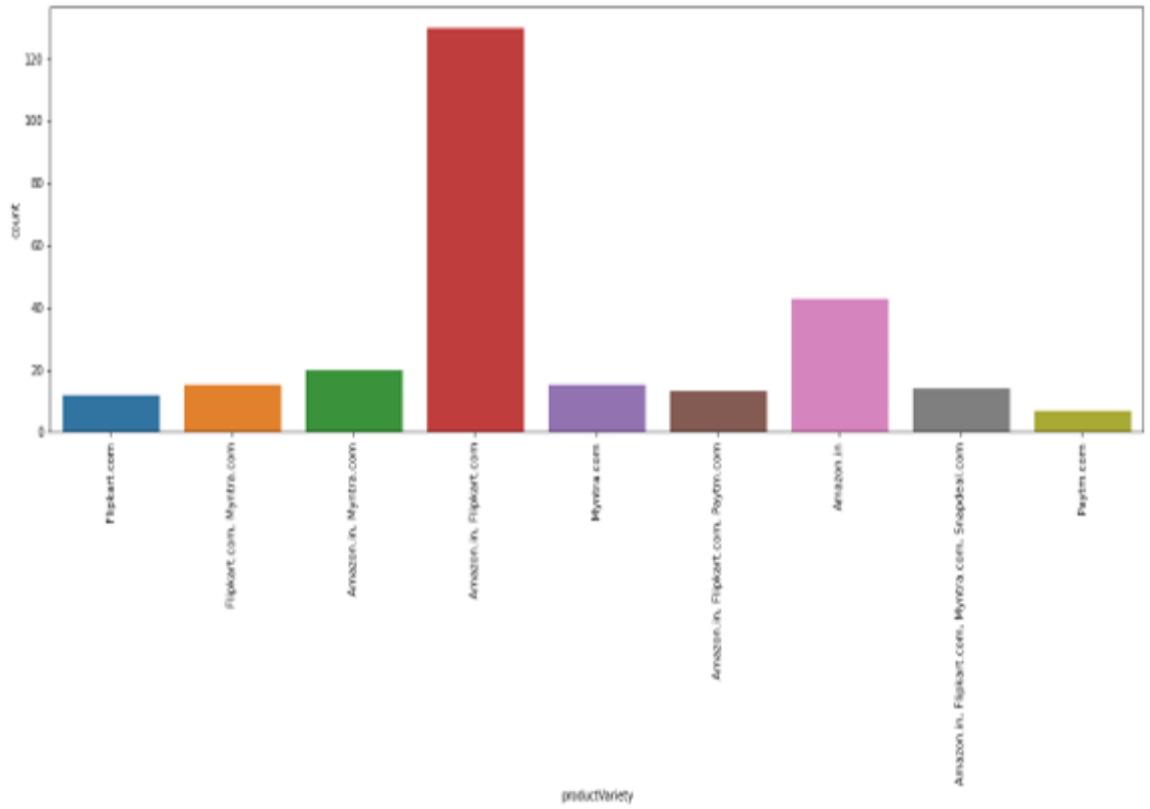
- Online shopping gives monetary benefit and discounts
- Getting value for money spent

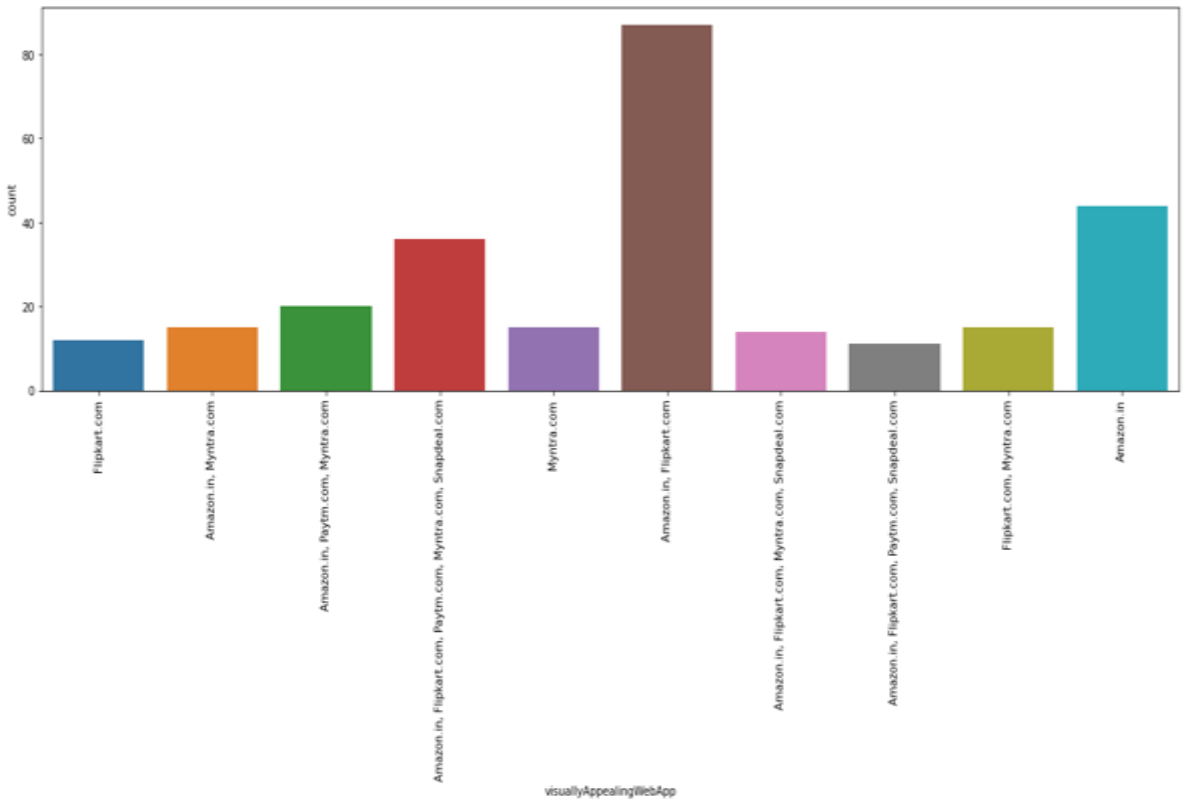
- Around 70% people believe that Enjoyment is derived from shopping online
- 90% people believe that Shopping online is convenient and flexible
- more than 90% of the people believe that Return and replacement policy of the e-tailer is important for purchase decision
- most of the people believe that Gaining access to loyalty programs is a benefit of shopping online

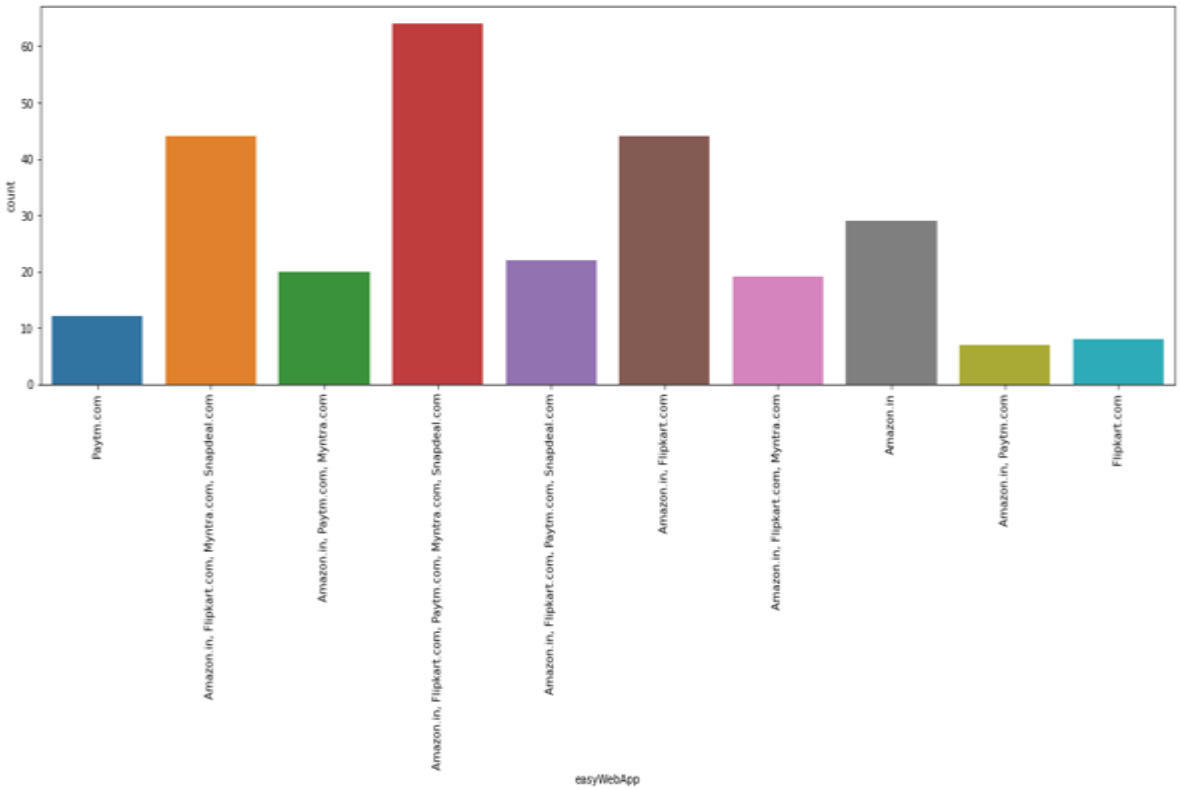
ECOMMERCE COLUMN

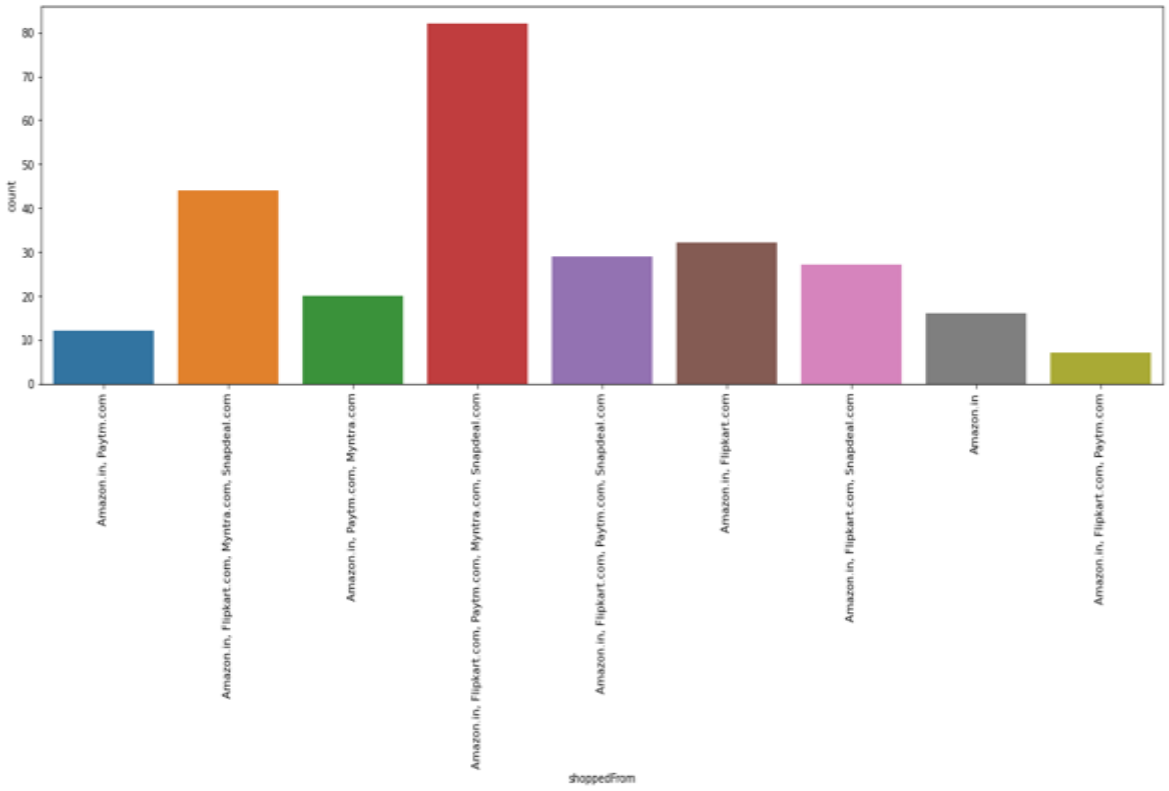





count
0
20
40
60
80
100
120
Flipkart.com
Flipkart.com, Myntra.com
Amazon.in, Myntra.com
Amazon.in, Flipkart.com
Myntra.com
Amazon.in, Flipkart.com, Paytm.com
Amazon.in
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com
Paytm.com
productVariety


count
0
20
40
60
80
Flipkart.com
Amazon.in, Myntra.com
Amazon.in, Paytm.com, Myntra.com
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com
Myntra.com
Amazon.in, Flipkart.com
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com
Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com
Flipkart.com, Myntra.com
Amazon.in
visuallyAppealingWebApp


count
0
10
20
30
40
50
60
Paytm.com
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com
Amazon.in, Paytm.com, Myntra.com
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com
Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com
Amazon.in, Flipkart.com
Amazon.in, Flipkart.com, Myntra.com
Amazon.in
Amazon.in, Paytm.com
Flipkart.com
easyWebApp


count
0
10
20
30
40
50
60
70
80
Amazon.in, Paytm.com
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com
Amazon.in, Paytm.com, Myntra.com
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com
Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com
Amazon.in, Flipkart.com
Amazon.in, Flipkart.com, Snapdeal.com
Amazon.in
Amazon.in, Flipkart.com, Paytm.com
shoppedFrom

SUMMARY

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

High Customer Satisfaction and Retention:

- ✓ Amazon.com
- ✓ Flipkart.com

High Risk of Customer Churn:

- ✓ Myntra.com
- ✓ Snapdeal.com

CONCLUSION

AMAZON: The most recommended websites with attractive web-page layout, easy to use, relevant descriptive information, product offers, reliability of website, quickness to complete purchase, trust worthiness. What can be improved: Takes longer time to login, Late declaration or price during sales and promotion, frequent disruption when moving from one page to another, Limited mode of payment on most of products.

FLIPKART: This is the 2nd most recommended website with fast loading page, security of financial information, trust worthiness, several payments modes, website is as efficient as before. What can be improved: Takes longer time in displaying graphics, late declaration of price during sales and promotion.

PAYTM: Reliability of website, speedy delivery of products, quickness in purchase. What can be improved: Longer page loading time, Longer delivery period, late declaration of price during sales and promotion.

MYNTRA: Myntra stands on 3rd most recommended websites with easy to use, wild variety of product offers, several payment methods, attractive visual appealing web-page layout. What can be improved: Relevant information about product, website loading speed, speedy delivery of products, websites is not much efficient as before.

SNAPDEAL: Least recommended website having less page loading time. What can be improved: Limited mode of payments, frequent disruption while moving from one page to another, Longer delivery period, customer's privacy information, reliability of website, offers on product, and must be an attractive web-page layout.

Thank You