



E-retail factors for customer activation and retention

NAME OF THE PROJECT

Salman Pradhan

Submitted by:

Salman Pradhan

ACKNOWLEDGMENT

I wish to thank Flip Robo Technologies for assigning this project to me and thanks to Mr.Shubham Yadav sir for his valuable technical support on this project.

INTRODUCTION

- **Business Problem Framing**
 - 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.
 - 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 combination of both utilitarian value and hedonistic values are needed to affect the repeat purchase intention (loyalty) positively.
 - We have to analyse the dataset and find useful information from it.

- **Conceptual Background of the Domain Problem**
 - For any digital marketing strategy or digital communications plan, defining a set of objectives and clear strategy behind your plan is essential to the success.
 - So customer retention is nothing but maintaining or growing your existing customer base loyal to your product offering.

- Retention should be at the heart of your digital strategy in maintaining and building communication with your existing customer base.
- **Motivation for the Problem Undertaken**

DataScience help us to make predictions at areas like health sectors, auto industry, education, media etc. For this project we decided to help online websites to add new customers to their online websites or applications and retaining their existing customers by giving them informations like which features are affecting more like monetary saving, product information, product offering, adventure etc.

Analytical Problem Framing

- **Mathematical/ Analytical Modeling of the Problem**

We have to find insights from the dataset, so that we know that customer retention depends on these factors.
- **Data Sources and their formats**

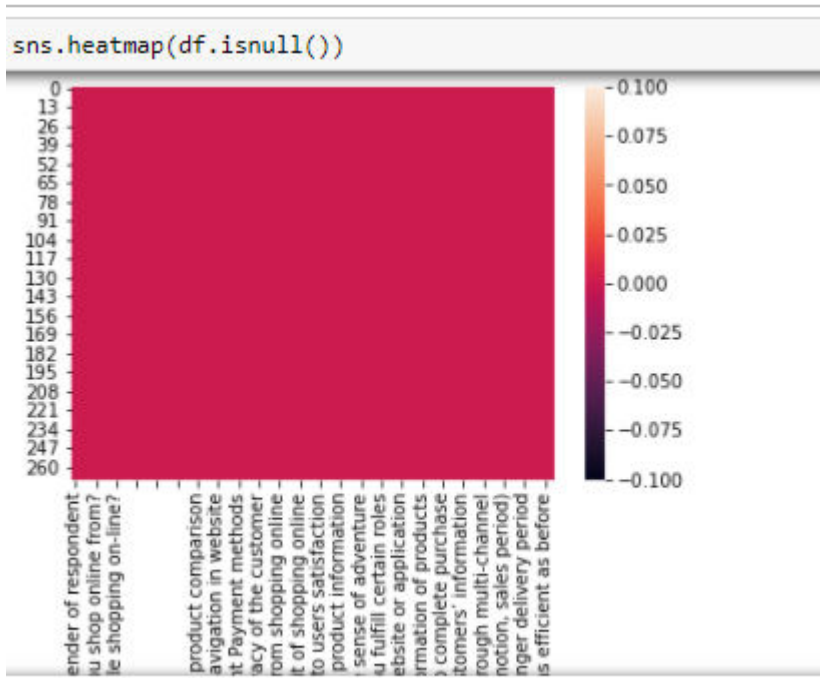
We have one excel file with 2 sheets. One is detailed and one is encoded. I have taken detailed sheet for data analysis.

The dataset has 269 rows and 71 columns.

There is no null values present in the dataset.

```
df=pd.read_excel("C:/Users/Lenovo/OneDrive/Desktop/Customer_retention_dataset/Customer_retention_dataset.xlsx")
pd.pandas.set_option('display.max_columns',100)#used to see all the columns in the dataset
df.head(3)
```

	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?	15 What is your preferred payment method?
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	E-wallet (Paytm)
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	Credit/Debit Card
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	E-wallet (Paytm)



There are no null values present in the dataset

- State the set of assumptions (if any) related to the problem under consideration

As diagram given to us, we found that Gratification, Role, Best Deal, Social, Adventure, Product offering, Convenience, Product information, and monetary Saving are highly dependent on customer retention.

So we compare all this above with all the online shopping sites individually such as Amazon, Flipkart, Myntra, Snapdeal and paytm where sellers are already shopping (i.e. column: 'From the following, tick any (or all) of the online retailers you have shopped from;').

- Hardware and Software Requirements and Tools Used
 1. NumPy: Base n-dimensional array package
 2. Matplotlib: Comprehensive 2D/3D plotting
 3. Seaborn: For plotting graph
 4. Pandas: Data structures and analysis

Importing Libraries

```
|: import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
```

Model/s Development and Evaluation

5. Identification of possible problem-solving approaches (methods)

1. Understand business problem
2. Get data
3. Data Cleaning
4. Data Analysis
5. Visualization

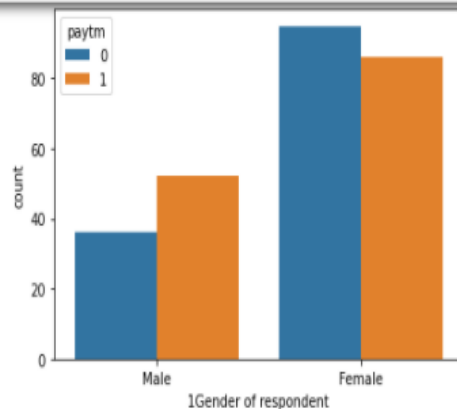
Interpretation of the Results

```
: df24.groupby('1Gender of respondent').sum()  
#Below are sum of how many males and females shop from which shopping sites
```

```
:  
amazon paytm snapdeal flipkart myntra
```

1Gender of respondent					
Female	181	86	124	152	131
Male	88	52	70	81	54

```
: for i in df24.columns:  
    sns.countplot(df24['1Gender of respondent'],hue=df24[i])  
    plt.show()
```



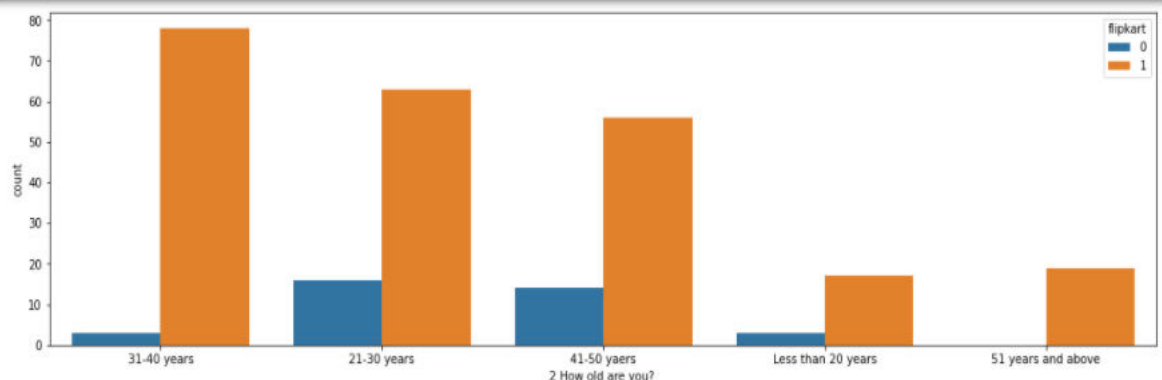
I found that Females are shopping more than males.Both males and females like to shopping from amazon.in than others.

```
: df25.groupby('2 How old are you? ').sum()
```

```
:  
amazon paytm snapdeal flipkart myntra
```

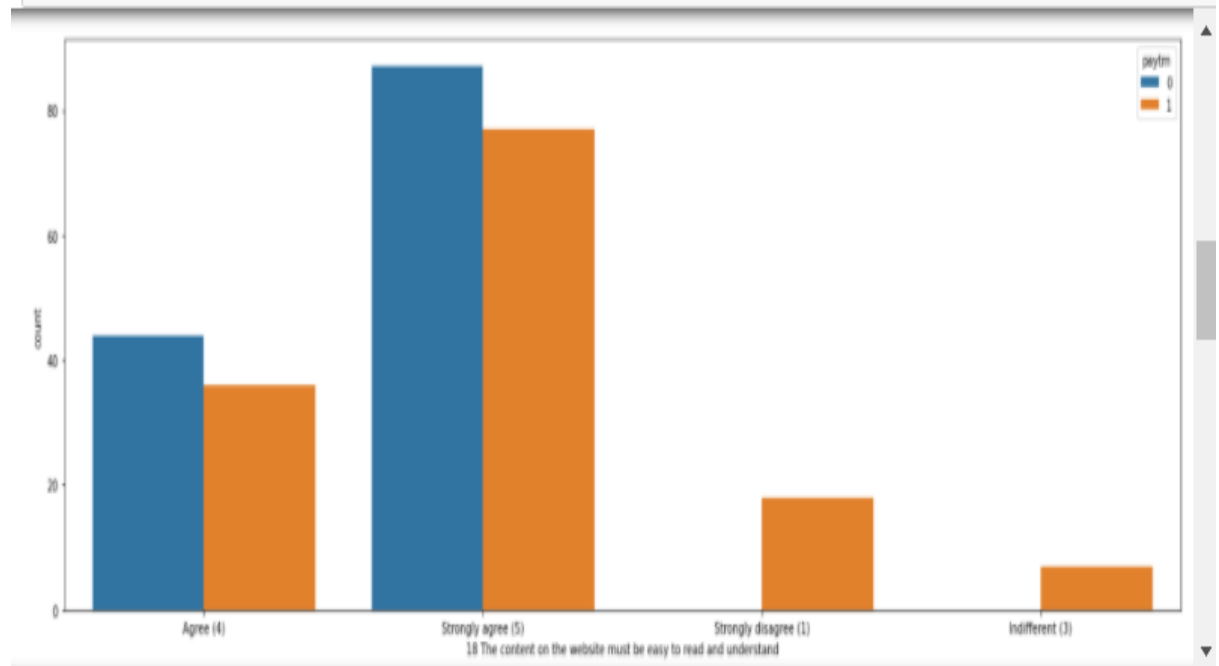
2 How old are you?					
21-30 years	79	37	52	63	55
31-40 years	81	44	55	78	52
41-50 yaers	70	39	55	56	48
51 years and above	19	9	19	19	14
Less than 20 years	20	9	13	17	16

```
: for i in df25.columns:  
    plt.figure(figsize=(18,5))  
    sns.countplot(df25['2 How old are you? '),hue=df25[i])  
    plt.show()
```



Age between 20-50 years shop more than other age.All age seller prefer Amazon as their first choice then flipkart.com.

```
for i in df27.columns:  
    plt.figure(figsize=(22,7))  
    sns.countplot(x='18 The content on the website must be easy to read and understand',hue=i,data=df27)  
    plt.show()
```

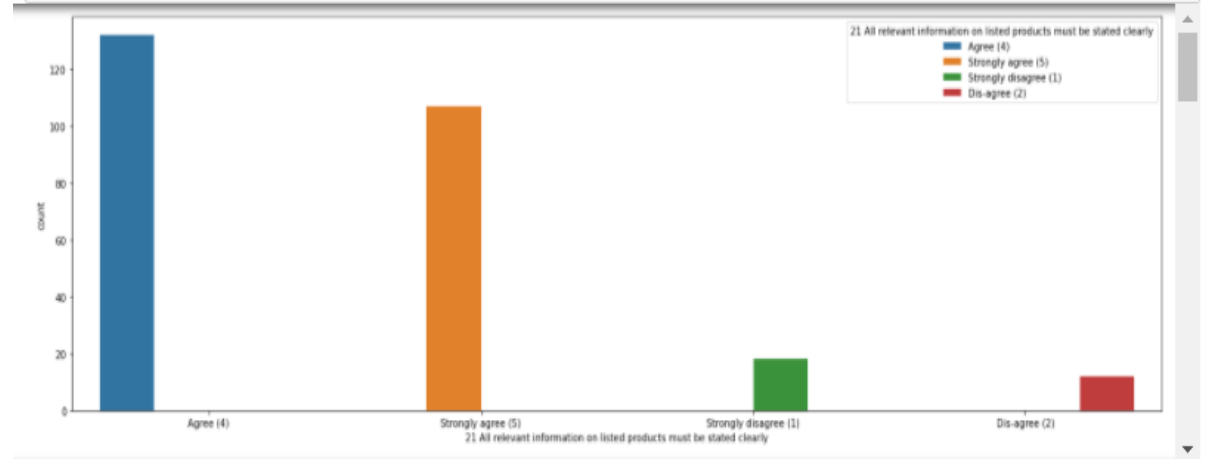


According to most of the seller amazon.in,flipkart.com,snapdeal.com and myntra.com websites are easy to read and understand .

```
df28.groupby('21 All relevant information on listed products must be stated clearly').sum()
```

	amazon	paytm	snapdeal	flipkart	mynta
21 All relevant information on listed products must be stated clearly					
Agree (4)	132	71	85	104	94
Dis-agree (2)	12	7	0	12	0
Strongly agree (5)	107	42	91	99	73
Strongly disagree (1)	18	18	18	18	18

```
for i in df28.columns:
    plt.figure(figsize=(22,7))
    sns.countplot(df28[i],hue=df28[i])
    plt.show()
```

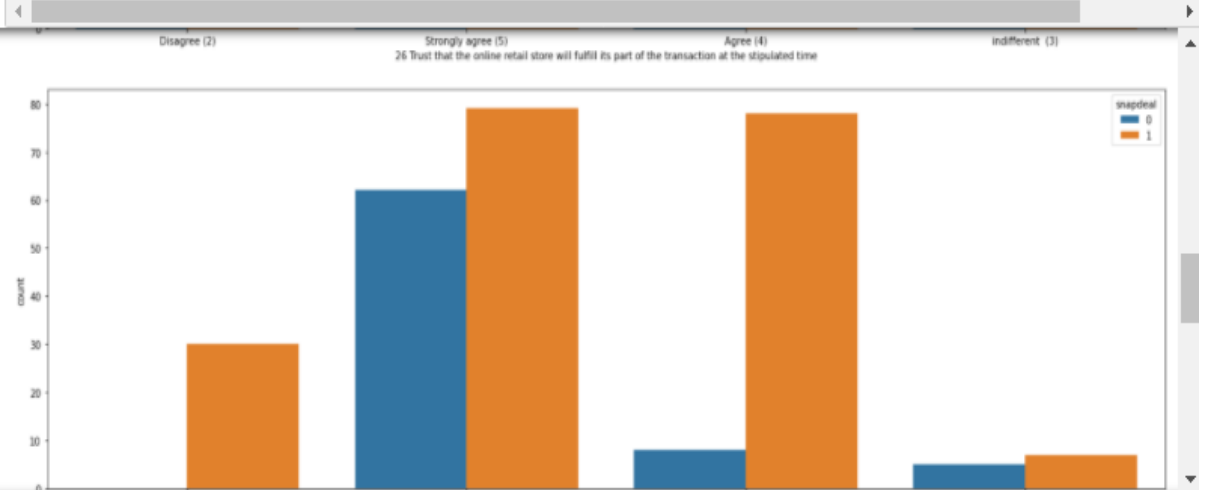


According to most of sellers all relevant information of products stated clearly.


```
df29.groupby('26 Trust that the online retail store will fulfill its part of the transaction at the stipulated time').mean()
```

	amazon	paytm	snapdeal	flipkart	myntra
26 Trust that the online retail store will fulfill its part of the transaction at the stipulated time					
Agree (4)	1.0	0.593023	0.908977	0.908977	0.778070
Disagree (2)	1.0	0.800000	1.000000	1.000000	1.000000
Strongly agree (5)	1.0	0.439716	0.580284	0.801418	0.624113
indifferent (3)	1.0	0.583333	0.583333	1.000000	0.000000

```
for i in df29.columns:
    plt.figure(figsize=(22,7))
    sns.countplot(df29['26 Trust that the online retail store will fulfill its part of the transaction at the stipulated time'],
    plt.show()
```

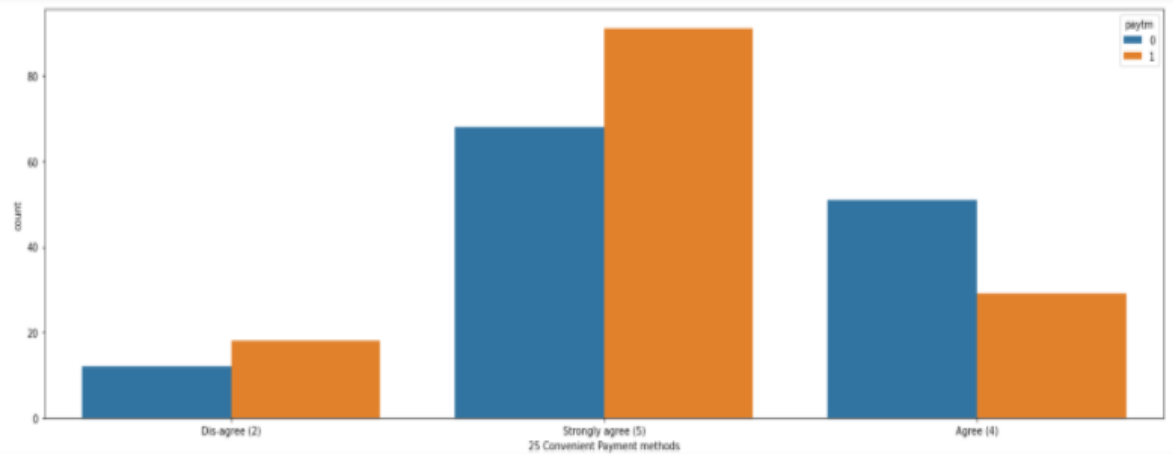


Sellers trust payments on Amazon.com, Flipkart.com, Snadeal.com and myntra.com than paytm.com.

```
df30.groupby('25 Convenient Payment methods').sum()
```

	amazon	paytm	snapdeal	flipkart	mynta
25 Convenient Payment methods					
Agree (4)	80	29	41	72	19
Dis-agree (2)	30	18	30	30	30
Strongly agree (5)	159	91	123	131	136

```
for i in df30.columns:
    plt.figure(figsize=(22,7))
    sns.countplot(df30['25 Convenient Payment methods'],hue=df30[i])
    plt.show()
```

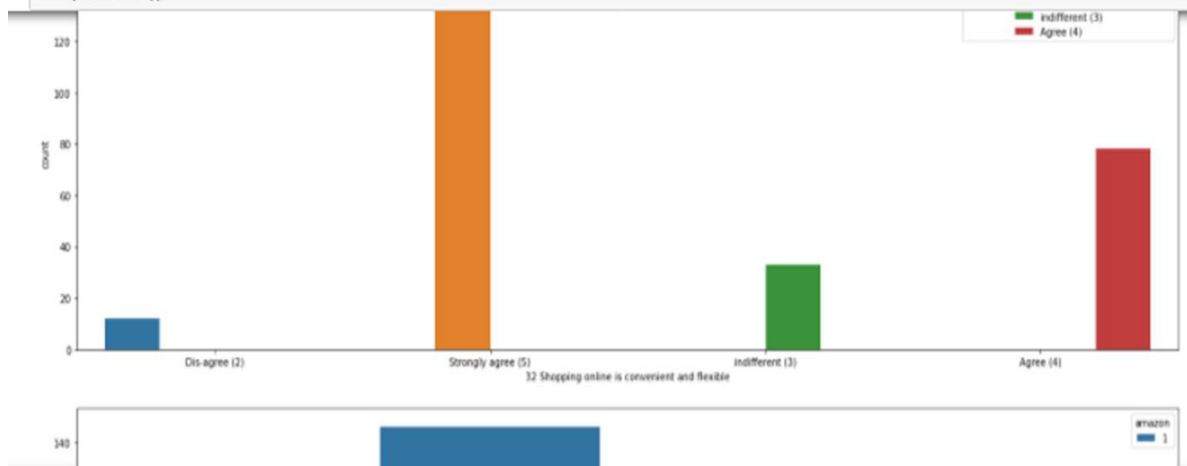


Amazon and flipkart.com have good convenient payment methods than others

```
df31.groupby('32 Shopping online is convenient and flexible').sum()
```

	amazon	paytm	snapdeal	flipkart	mynta
32 Shopping online is convenient and flexible					
Agree (4)	78	36	52	78	23
Dis-agree (2)	12	0	12	12	12
Strongly agree (5)	146	77	105	118	125
indifferent (3)	33	25	25	25	25

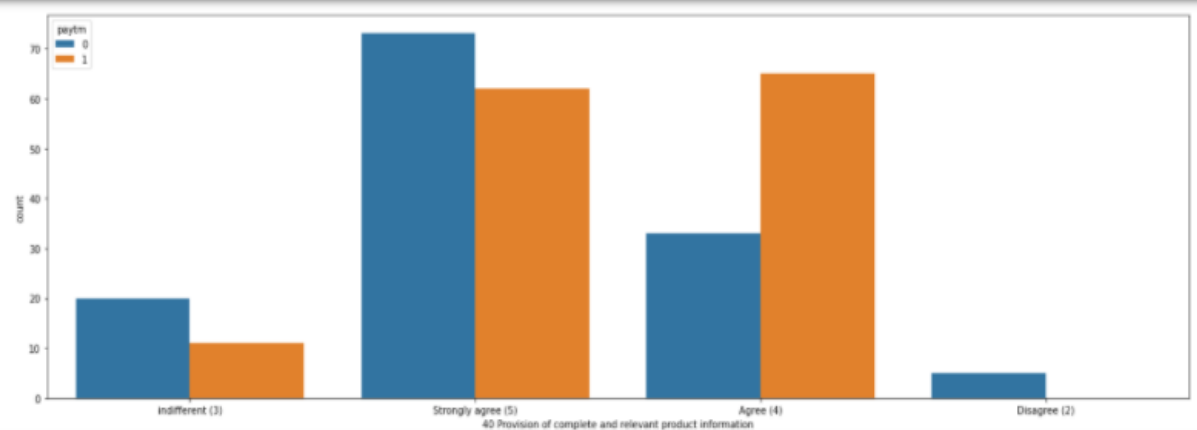
```
for i in df31.columns:
    plt.figure(figsize=(22,7))
    sns.countplot(df31['32 Shopping online is convenient and flexible'],hue=df31[i])
    plt.show()
```



Except 12 sellers ,all sellers told that shopping online is convenient and flexible.

	amazon	paytm	snapdeal	flipkart	mynta
40 Provision of complete and relevant product information					
Agree (4)	98	65	72	98	54
Disagree (2)	5	0	0	5	0
Strongly agree (5)	135	62	99	107	119
indifferent (3)	31	11	23	23	12

```
for i in df32.columns:
    plt.figure(figsize=(22,7))
    sns.countplot(df32['40 Provision of complete and relevant product information'],hue=df32[i])
    plt.show()
```

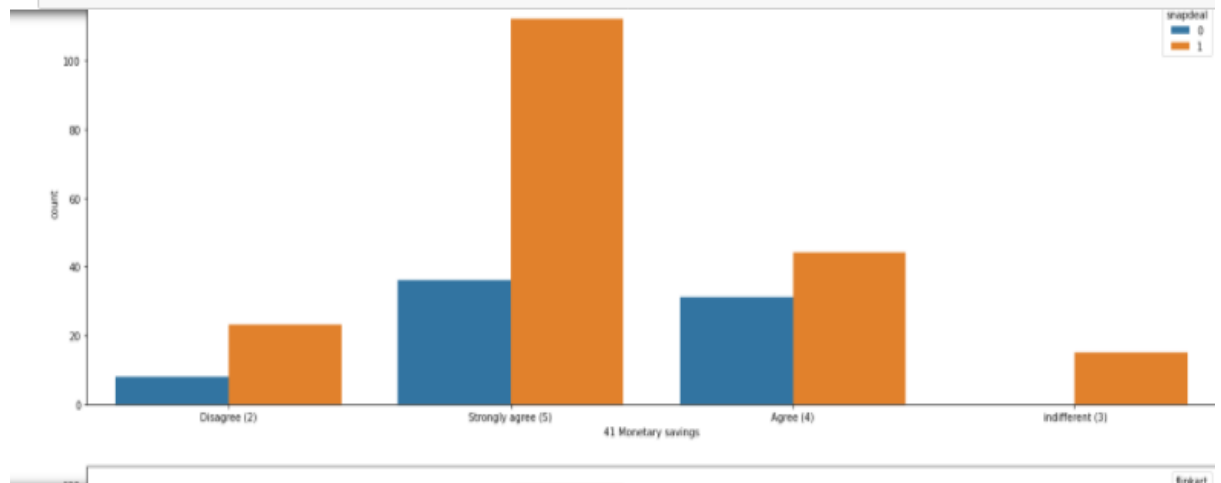


Amazon.in and flipkart.com have mentioned complete and relevant information of product compared to other sites .

```
] df33.groupby('41 Monetary savings').mean()
```

```
]
amazon    paytm    snapdeal    flipkart    mynta
41 Monetary savings
Agree (4)    1.0    0.388887    0.588887    1.000000    0.440000
Disagree (2)  1.0    0.354839    0.741935    0.741935    0.387097
Strongly agree (5)  1.0    0.662162    0.756757    0.810811    0.844595
indifferent (3)  1.0    0.000000    1.000000    1.000000    1.000000
```

```
] for i in df33.columns:
    plt.figure(figsize=(22,7))
    sns.countplot(df33['41 Monetary savings'],hue=df33[i])
    plt.show()
```

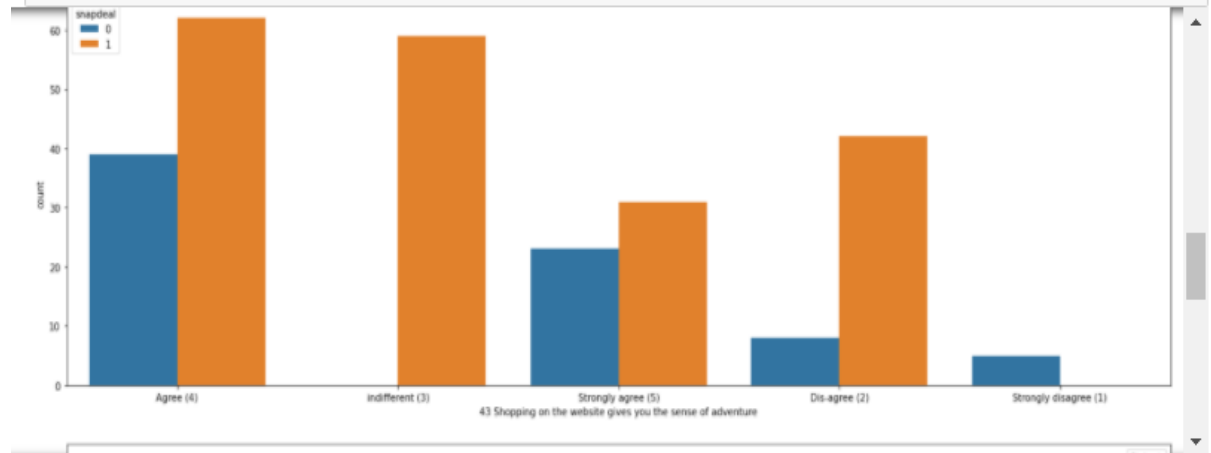


From above we see that buying online is monetary saving. According to customers who shop from amazon, flipkart are getting more monetary saving.

```
: df35.groupby('43 Shopping on the website gives you the sense of adventure').sum()
```

	amazon	paytm	snapdeal	flipkart	mynta
43 Shopping on the website gives you the sense of adventure					
Agree (4)	101	56	62	81	71
Dis-agree (2)	50	0	42	42	42
Strongly agree (5)	54	38	31	46	31
Strongly disagree (1)	5	0	0	5	0
indifferent (3)	59	44	59	59	41

```
: for i in df35.columns:
    plt.figure(figsize=(22,7))
    sns.countplot(df35['43 Shopping on the website gives you the sense of adventure'],hue=df35[i])
    plt.show()
```

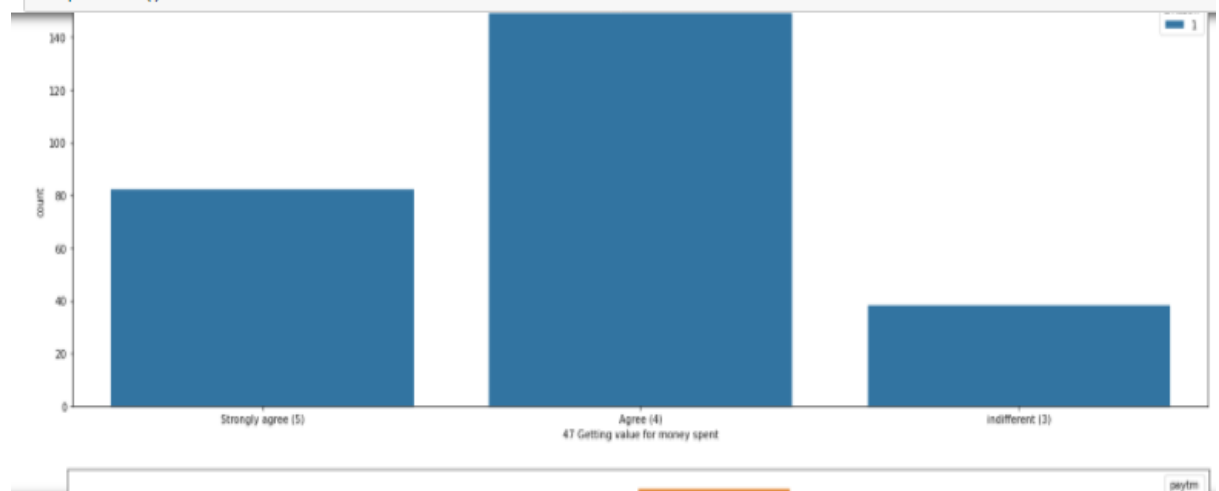


Customers shopping from amazon.in tells that Shopping on the website give the sense of adventure followed by flipkart.com

```
: df36.groupby('47 Getting value for money spent').sum()
```

	amazon	paytm	snapdeal	flipkart	mynta
47 Getting value for money spent					
Agree (4)	149	88	103	129	105
Strongly agree (5)	82	39	66	74	66
indifferent (3)	38	11	25	30	14

```
: for i in df36.columns:
    plt.figure(figsize=(22,7))
    sns.countplot(df36['47 Getting value for money spent'],hue=df36[i])
    plt.show()
```

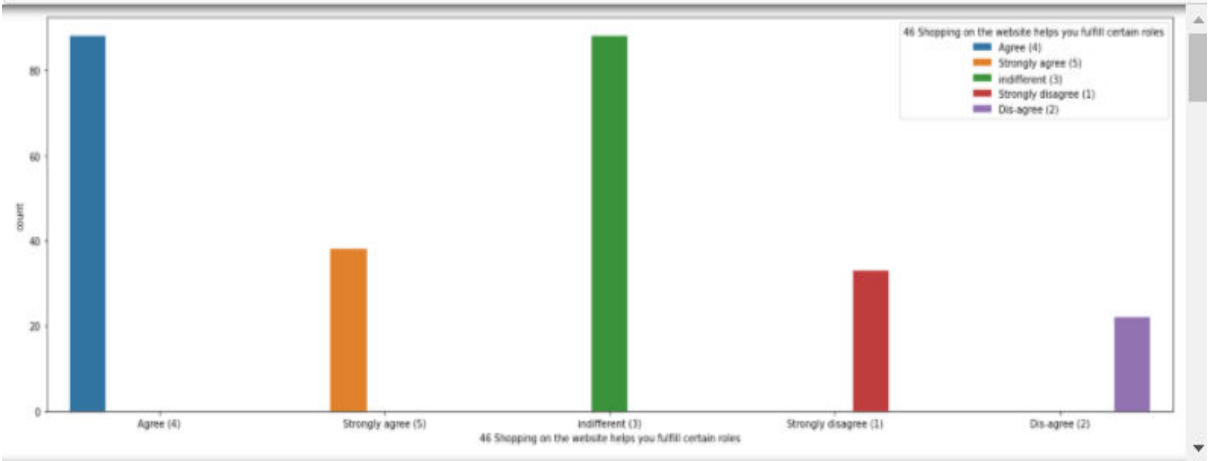


Customers who are buying from online sites all are telling that money spent on buying product from online is value for money

```
df37.groupby('46 Shopping on the website helps you fulfill certain roles').sum()
```

	amazon	paytm	snapdeal	flipkart	mynta
46 Shopping on the website helps you fulfill certain roles					
Agree (4)	88	46	80	80	58
Dis-agree (2)	22	7	22	22	15
Strongly agree (5)	38	7	15	30	15
Strongly disagree (1)	33	33	33	33	33
indifferent (3)	88	45	44	68	64

```
for i in df37.columns:
    plt.figure(figsize=(22,7))
    sns.countplot(df37['46 Shopping on the website helps you fulfill certain roles'],hue=df37[i])
    plt.show()
```



According to all most 75% of customers agreed that shopping on websites helps in fulfilling certain roles.

CONCLUSION

1. Dataset has 269 rows and 71 columns and there is no null value present.
2. We know that Gratification, Role, Best Deal, Social, Adventure, Product offering, Convenience, Product information, and monetary Saving are highly correlated with customer retention. So I compare each of them with from where they are shopping individually.
I conclude that
3. Females are shopping more than males. Both males and females like to shopping from amazon.in than others.

4. Age between 20-50 years shop more than other age. All age seller prefer Amazon as their first choice then flipkart.com.
5. Cities like Bulandshahr, Merrut, Moradabad buy smaller amount of things from ecommerce sites and Cities like Delhi, Greater noida, Noida, Bangalore people are shopping more from online websites and all prefer amazon.in for buying something then flipkart.com.
6. According to most of the seller amazon.in, flipkart.com, snapdeal.com and myntra.com websites are easy to read and understand .
7. According to most of sellers all relevant information of products stated clearly.
8. Sellers trust payments on Amazon.com, Flipkart.com, Snadeal.com and myntra.com than paytm.com.
9. Amazon and flipkart.com have good convenient payment methods than others.
10. Amazon.in and flipkart.com have mentioned complete and relevant information of product compared to other sites .
11. Buying online is monetary saving. According to customers who shop from amazon, flipkart are getting more monetary saving.