

PROJECT REPORT ON:

“CUSTOMER RETENTION”

1. Content:

1. Content
2. Introduction
3. Understanding problem statement
4. Importing Dataset and important libraries
5. Preprocessing of data
6. Feature Engineering
7. Data Visualization
8. Conclusion

2. Introduction:

A). What is Customer Retention?

Customer retention means the process of maintaining or keeping customers once you have acquired them. It's all the activities that a company must do in order to keep their customers around. The goal is to build a long-lasting relationship between the brand and consumers. Once a customer becomes loyal to your brand, not only he will buy more from you than a normal customer but he'll spread good words about your business, increase your reputation.

B). Why do we need customer retention ?

- Lower Marketing Costs.
- Repeat Purchases means repeat profit.
- Word-Of-Mouth Advertising.
- Gain Valuable Feedback.
- Sell at premium Prices.
- It's easier and less expensive to retain customer than to acquire them.

3. Understanding problem statement:

The data is collected from the Indian online shoppers. Results indicate the e-retail success factors, which are very much critical for customer satisfaction. There are two sheets (one is detailed) and second is

encoded in the excel file. You may use any of them by extracting in separate excel sheet. The number of column(s) is more than 47.

The features in the dataset are as follows:

- 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 Option?
- 16 How frequently do you abandon (selecting an items and leaving without making payment) your shopping cart?
- 17 Why did you abandon the “Bag”, “Shopping Cart”?
- 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
- 28 Being able to guarantee the privacy of the customer
 - 1. 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
- 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

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)

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?

Above are all the feature of or dataset, factors that effects customer retention.

4. Importing Dataset and Important Libraries:

```

In [57]: 1 # Importing necessary libraries
          2 import numpy as np
          3 import pandas as pd
          4 import seaborn as sns
          5 import matplotlib.pyplot as plt
          6 import warnings
          7 warnings.filterwarnings('ignore')

In [58]: 1 #To print all columns
          2 pd.set_option('display.max_columns',None)
          3 pd.set_option('display.max_rows',None)

In [59]: 1 df= pd.read_excel(r"C:\Users\hp\Dropbox\PC\Desktop\FIIP Robo internship\Customer_retention_dataset\customer_retention_data.xlsx")
          2 df
          3
Out[59]:

```

	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 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? (inches)	10 What is the operating system (OS) of your device? (inches)	11 What browser do you run on your device to access the website? (inches)	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? (inches)	14 How much time do you explore the e-retail store before making a purchase decision? (inches)
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

As seen on previous page important libraries has been imported and Dataset has also been imported and printed as shown in previous page

5. Preprocessing of data:

- I. First i checked the shape of dataset for better understanding of data.(From which i find out there are total of 71 feature and 269 no of entries(row) in each columns)
- II. Then I checked different feature's present in our dataset and there datatype(now I have clear idea about the datatype of each features)
- III. Next i checked for null values if any present in our dataset(And I found out there is no null value present in our dataset)
- IV. Then I checked the unique value each feature contains and there counts(Now I have clear idea about the unique value and the counts of each features)
- V. Then I proceeded further and did some feature engineering

6. Feature engineering :

A) From above preprocessing and observation I discovered that there are some unique values which are duplicated/repeated or kind of similar ,so I decided to merge them for better understanding and analysis of given dataset:

(As can be seen further in next page, that I did some feature engineering and merge the duplicated if any in any column)

The columns in which I merge the duplicates are as follows:

1. Column 6-7 (index (5-6))
2. Column 16-47 (index (15 - 46))

Grouping Duplicate entries:

1 Grouping duplicate entries so it become easier for us to analyse our data as its a very large dataset.

[illegible]

B) After merging duplicates I noticed that dtype of all the features are "Object Data Type" except 1 feature i.e "4 What is the Pin Code of where you shop online from?" so I change the data type of this Feature , so that it become easy for us to analyse the data, code for changing data type is as follows:

```
[df['4 What is the Pin Code of where you shop online from?'] = df['4  
What is the Pin Code of where you shop online from?'].astype("object")]
```

C) I noticed from looking at the dataset that the Title of all the feature are so long , and it's taking extra time to read and understand them, so I renamed all the feature into it's shorter and easy to understand version, which are as follows:

- 1 Gender of respondent ---> Gender
- 2 How old are you? ---> Age
- 3 Which city do you shop online from? ---> Shopping_City
- 4 What is the Pin Code of where you shop online from? ---> PinCode
- 5 Since How Long You are Shopping Online ? ---> Shopping_Since

- 6 How many times you have made an online purchase in the past 1 year? ---> Shopping_Frequency
- 7 How do you access the internet while shopping on-line? ---> Internet_Access
- 8 Which device do you use to access the online shopping? ---> Device_Used
- 9 What is the screen size of your mobile device? ---> Screen_Size
- 10 What is the operating system (OS) of your device? ---> Operating_System
- 11 What browser do you run on your device to access the website? ---> Browser_Used
- 12 Which channel did you follow to arrive at your favorite online store for the first time? ---> Channel_FirstUsed
- 13 After first visit, how do you reach the online retail store? ---> Login_Mode
- 14 How much time do you explore the e- retail store before making a purchase decision? ---> TimeSpent_ForPurchase
- 15 What is your preferred payment Option? ---> Payment_Mode
- 16 How frequently do you abandon (selecting an items and leaving without making payment) your shopping cart?---> Abandon_Frequency
- 17 Why did you abandon the “Bag”, “Shopping Cart”? ---> Abandon_Reason
- 18 The content on the website must be easy to read and understand ---> Content_Readability
- 19 Information on similar product to the one highlighted is important for product comparison ---> Similar_ProductInfo
- 20 Complete information on listed seller and product being offered is important for purchase decision. ---> Seller_ProductInfo
- 21 All relevant information on listed products must be stated clearly ---> ProductInfo_Clarity
- 22 Ease of navigation in website ---> Ease_Navigation
- 23 Loading and processing speed ---> Loading_ProcessingSpeed
- 24 User friendly Interface of the website ---> UserFriendly_Interface
- 25 Convenient Payment methods ---> Conveninet_PaymentMode

- 26 Trust that the online retail store will fulfill its part of the transaction at the stipulated time ---> TimelyFulfilment_Trust
- 27 Empathy (readiness to assist with queries) towards the customers ---> Customer_Empathy
- 28 Being able to guarantee the privacy of the customer ---> CustPrivacy_Guarantee
- 29 Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.) ---> VariousChannel_Responses
- 30 Online shopping gives monetary benefit and discounts ---> Benefit_Discount
- 31 Enjoyment is derived from shopping online ---> Enjoy_OnlineShopping
- 32 Shopping online is convenient and flexible ---> Convenience_Flexibility
- 33 Return and replacement policy of the e-tailer is important for purchase decision ---> Returns_ReplacementPolicy
- 34 Gaining access to loyalty programs is a benefit of shopping online ---> Loyalty_ProgramAccess
- 35 Displaying quality Information on the website improves satisfaction of customers ---> QualityInfo_Satisfaction
- 36 User derive satisfaction while shopping on a good quality website or application ---> WebsiteQuality_Satisfaction
- 37 Net Benefit derived from shopping online can lead to users satisfaction ---> NetBenefit_Satisfaction
- 38 User satisfaction cannot exist without trust ---> User_Trust
- 39 Offering a wide variety of listed product in several category --> Product_SeveralCategory
- 40 Provision of complete and relevant product information ---> Relevant_ProductInfo
- 41 Monetary savings ---> Monetary_Savings
- 42 The Convenience of patronizing the online retailer ---> Patronizing_Convenience
- 43 Shopping on the website gives you the sense of adventure ---> Adventure_Sense
- 44 Shopping on your preferred e-tailer enhances your social status ---> Enhances_SocialStatus

- 45 You feel gratification shopping on your favorite e-tailer ---> Gratification_Shopping
- 46 Shopping on the website helps you fulfill certain roles ---> Role_Fulfilment
- 47 Getting value for money spent ---> Money_Worthy
- 48 From the following, tick any (or all) of the online retailers you have shopped from ---> Shopped_From
- 49 Easy to use website or application ---> Easy_WebApp
- 50 Visual appealing web-page layout ---> Visually_AppealingWebApp
- 51 Wild variety of product on offer---> Product_Variety
- 52 Complete, relevant description information of products ---> Complete_ProductInfo
- 53 Fast loading website speed of website and application ---> Fast_WebApp
- 54 Reliability of the website or application ---> Reliable_WebApp
- 55 Quickness to complete purchase ---> Quick_Purchase
- 56 Availability of several payment options ---> PaymentOptions_Availability
- 57 Speedy order delivery ---> Fast_Delivery
- 58 Privacy of customers information ---> CustInfo_Privacy
- 59 Security of customer financial information ---> FinancialInfo_Security
- 60 Perceived Trustworthiness ---> Perceived_Trustworthiness
- 61 Presence of online assistance through multi-channel ---> MultiChannel_Assistance
- 62 Longer time to get logged in (promotion, sales period) ---> Long_LoginTime
- 63 Longer time in displaying graphics and photos (promotion, sales period) ---> LongPhoto_DisplayTime
- 64 Late declaration of price (promotion, sales period) ---> LatePrice_Declaration
- 65 Longer page loading time (promotion, sales period) ---> Long_LoadingTime
- 66 Limited mode of payment on most products (promotion, sales period) ---> Limited_PaymentMode
- 67 Longer delivery period ---> Late_Delivery

- 68 Change in website/Application design ---> ChangeWebApp_Design
- 69 Frequent disruption when moving from one page to another ---> Page_Disruption
- 70 Website is as efficient as before ---> WebApp_Efficiency
- 71 Which of the Indian online retailer would you recommend to a friend? ---> Recommendation

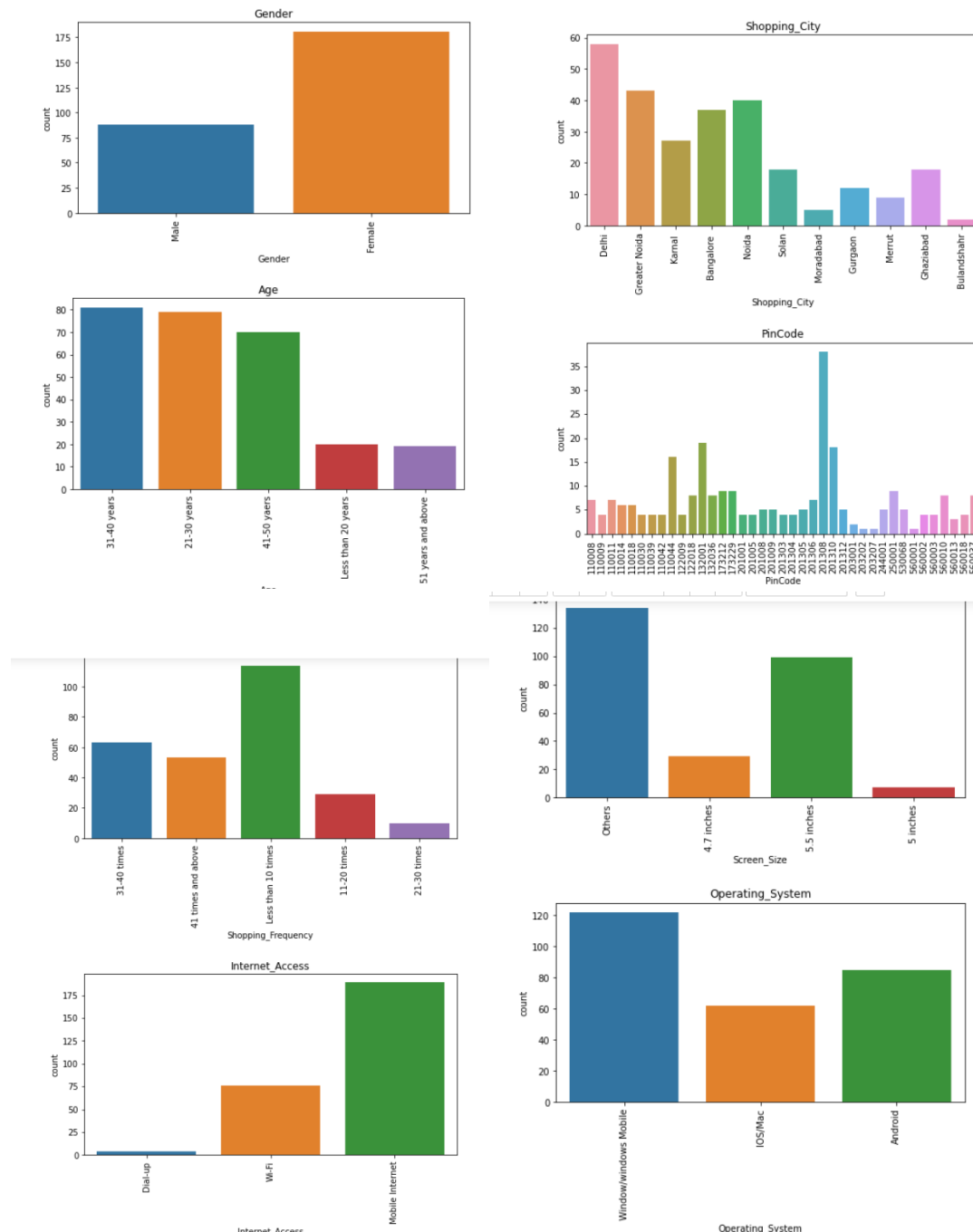
Out[74]:

Gender	Age	Shopping_City	PinCode	Shopping_Since	Shopping_Frequency	Internet_Access	Device_Used	Screen_Size	Operating_System	Browser_Used
Male	31-40 years	Delhi	110009	Above 4 years	31-40 times	Dial-up	Desktop	Others	Window/windows Mobile	Google chrome
Female	21-30 years	Delhi	110030	Above 4 years	41 times and above	Wi-Fi	Smartphone	4.7 inches	IOS/Mac	Google chrome
Female	21-30 years	Greater Noida	201308	3-4 years	41 times and above	Mobile Internet	Smartphone	5.5 inches	Android	Google chrome
Male	21-30 years	Karnal	132001	3-4 years	Less than 10 times	Mobile Internet	Smartphone	5.5 inches	IOS/Mac	Safari
Female	21-30 years	Bangalore	530068	2-3 years	11-20 times	Wi-Fi	Smartphone	4.7 inches	IOS/Mac	Safari

As can be seen above columns has been renamed, now we can proceed further and analyse it.

7. Data Visualiztion

Since all the features are categorical we can use only categorical plotting to get better insight. And particularly I have used bar plot for v



Similarly I counplot each and every feature and the observation that I made from the count plot are as follows:

1. Female customer often prefer to purchase more as compared to Male customer
2. People between 21 to 50 years of age tends shops more.
3. Delhi has maximum count which means in Delhi the online purchase is high.
4. A city with pin code 201308 has maximum online shopping, which means in this area people oftenly buy/purchase products online.

5. The count is high for Above 4 years which means people are choosing to shop online since 4 years.
6. The count is high for less than 10 times. People who are shopping less than 10 times (usually during any festival or may be some special occasion) per year are more in number.
7. Most of the customer uses their mobile internet to shop online, which seems to be more convenient to use.
8. Maximum customers use smartphones while shopping, which is easy to operate.
9. Mobile screen size has maximum count for others, and 2nd most count is for 5.5 which is the most used screen size.
10. Customers with OS windows/mobile are more in number, as it's the most common OS.
11. Most of the customers use Chrome browser, which seems to be best in market and easy to use.
12. Most of the customers used Search engine to reach their favorite online store for the first time, as search engine shows variety of similar product from a number of sites.
13. After first visit most of the customers are reaching the online store through Search engine, Via application and Direct Url.
14. Most of the customers are exploring the e-retail store more than 6 mins before making purchase decision (checking all info and deciding which one to buy usually take time).
15. Maximum customers are using credit/debit cards for their payment (most convenient and easy to use).
16. The count of abandon is high for sometimes, which means people who are abandoning sometimes are more.
17. Most of the customers are abandoning because to get better alternative offers, which seems to be obvious usually we prefer good product at cheaper price.
18. Most of the customers are agreeing that the content should be easy to understand on the website.
19. Most of the customers are agreeing that the Information on similar product to the one highlighted is important for product comparison.
20. Most of the customers are agreeing that the Complete information on listed seller and product being offered is important for purchase decision.

21. Most of the customers agrees that the relevant information on listed products must be stated clearly(as there should be no room for any kind of error or misunderstanding).
22. Maximum customers are agreeing that it should be easy to navigate in websites.
23. Maximum customers are agreeing that loading and processing speed should be high.
24. Maximum customers agrees to have user friendly interface of the website.
25. Maximum customers agrees to have convenient payment methods.
26. Maximum customers Trust that the online retail store will fulfill its part of the transaction at the stipulated time.
27. Maximum customers agrees to have Empathy (readiness to assist with queries) towards the customers.
28. Maximum customers agrees to have the ability to guarantee the privacy of the customer.
29. Maximum customers agrees to have Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.).
30. Maximum customers agrees to have Online shopping gives monetary benefit and discounts.
31. Maximum customers agrees to have Enjoyment is derived from shopping online.
32. Maximum customers agrees to have Shopping online is convenient and flexible.
33. Maximum customers agrees that Return and replacement policy of the e-tailer is important for purchase decision.
34. Maximum customers agrees that Gaining access to loyalty programs is a benefit of shopping online.
35. Maximum customers agrees that Displaying quality Information on the website improves satisfaction of customers.
36. Maximum customers agrees that User derive satisfaction while shopping on a good quality website or application.
37. Maximum customers agrees that Net Benefit derived from shopping online can lead to users satisfaction.
38. Maximum customers agrees that User satisfaction cannot exist without trust.

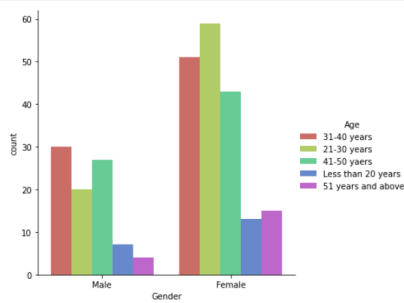
39. Maximum customers agrees to have Offering a wide variety of listed product in several category.
40. Maximum customers agrees to have Provision of complete and relevant product information.
41. Maximum customers agrees to have Monetary savings.
42. Maximum customers agrees to have the Convenience of patronizing the online retailer.
43. Maximum customers agrees that Shopping on the website gives you the sense of adventure.
44. Maximum customers agrees that Shopping on your preferred e-tailer enhances your social status.
45. Maximum customers agrees that You feel gratification shopping on your favorite e-tailer.
46. Maximum customers agrees that Shopping on the website helps you fulfill certain roles.
47. Maximum customers agrees to Getting value for money spent.
48. Most of the people are shopping from all available online retailers.
49. Most of the people are choosing all retailers for easy to use websites.
50. Maximum count is for Amazon.in, Flipkart.com as Visual appealing web-page layout.
51. Maximum count is for Amazon.in, Flipkart.com for having Wild variety of product on offer.
52. Maximum count is for Amazon.in, Flipkart.com for having Complete, relevant description information of products.
53. Maximum count is for Amazon.in for having Fast loading website speed of website and application.
54. Maximum count is for Amazon.in for having Reliable website or application.
55. Maximum count is for Amazon.in for having the Quickest to complete purchase.
56. Maximum count is for Amazon.in, Flipkart.com for having Availability of several payment options.
57. Maximum count is for Amazon.in for having Speedy order delivery.
58. Maximum count is for Amazon.in for having Privacy of customers' information.

59. Maximum count is for Amazon.in for having Security of customer financial information.
60. Maximum count is for Amazon.in for having Perceived Trustworthiness.
61. Maximum count is for Amazon.in and Amazon.in,Flipkart.com,Myntra.com,Snapdeal for having Presence of online assistance through multi-channel.
62. Maximum count is for Amazon.in for Longer time to get logged in (promotion, sales period).
63. Maximum count is for Amazon.in,Flipkart.com for having Longer time in displaying graphics and photos (promotion, sales period).
64. Maximum count is for Myntra.com for Late declaration of price (promotion, sales period).
65. Maximum count is for choosing Myntra.com for Longer page loading time (promotion, sales period).
66. Maximum count is for Snapdeal.com for having Limited mode of payment on most products (promotion, sales period).
67. Maximum count is for Paytm.com for Longer delivery period.
68. Maximum count is for Amazon.in for having Change in website/Application design.
69. Maximum count is for Amazon.in,Myntra.com and Snapdeal.com for having Frequent disruption when moving from one page to another.
70. Maximum count is for Amazon.in for having Website is as efficient as before.
71. Maximum count is for Amazon.in for recommendation

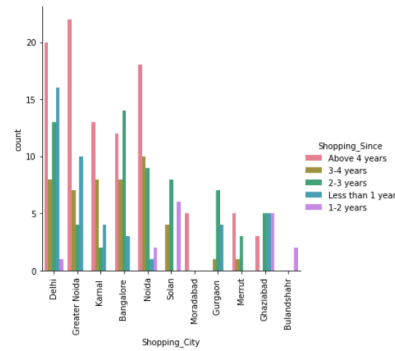
Similarly I did bivariate analysis as shown below:

Bivariate Analysis:

```
In [93]: 1 #Factor plot for Gender
2 sns.factorplot('Gender', kind='count', data=df, hue='Age', palette=
3 plt.xticks(rotation=90);
```



```
In [94]: 1 #Factor plot for Shopping_City
2 sns.factorplot('Shopping_City', kind='count', data=df, hue='Shopping_Since', palette="husl")
3 plt.xticks(rotation=90);
```



I did this for different combination of feature which seem relatable to me and in general and my observations were:

- I. Most of the customers are females with age from 20-50 and they used mobile internet to access with Windows as operating system.
- II. And most of the customers used search engines to get into the platform first time and eventually the download the app most of the time .
- III. Max customers uses there debit/credit cards for there payment.
- IV. Costomer sometimes abandon the product in the cart and main reason being availability of better product & offer.
- V. Almost customers agrees to have Content_Readability, Similar_ProductInfo, Seller_ProductInfo, ProductInfo_Clarity, Ease_Navigation, Loading_ProcessingSpeed, UserFriendly_Interface, Conveninet_PaymentMode, TimelyFulfilment_Trust, Customer_Empathy, CustPrivacy_Guarantee, VariousChannel_Responses, Benefit_Discount, Enjoy_OnlineShopping, Convenience_Flexibility, Returns_ReplacementPolicy, Loyalty_ProgramAccess, QualityInfo_Satisfaction, WebsiteQuality_Satisfaction, NetBenefit_Satisfaction, User_Trust, Product_SeveralCategory, Relevant_ProductInfo, Monetary_Savings, Patronizing_Convenience, Adventure_Sense, Enhances_SocialStatus, Gratification_Shopping, Role_Fulfilment, Money_Worthy.

- VI. Compared to other platforms shopping with Amazon.in and Flipkart.com has maximum benefits rather than drawbacks.

8. Conclusion

There is high risk of customer churn with:

- a) Myntra.com
- b) Snapdeal.com
- c) Paytm.com

High customer retention and satisfaction with:

- a) Amazon.in
- b) Flipkart.com

When customers are satisfied with a company or service, there is a high possibility that they will share their experience with other people Therefore it is crucial for E-commerce to take into account their customer satisfaction because this will retain customer loyalty as well as attract potential customers.

To conclude, having the right customer retention strategy will keep sellers company growing if they know how to take advantage of it. Then customers will find their way back and continue buying stuff from the best company.