# Customer Retention Case-study.

* E-retail factors for customer activation and retention in the Indian online Retail Sector.



Project Report

* Submitted by : Nilotpal Mukherjee

**ACKNOWLEDGMENT**

The dataset was provided by FilpRobo technologies. The dataset mainly deals with the information [**E-retail factors for customer activation and retention: A case study from Indian e-commerce customers**](https://www.researchgate.net/publication/346412647_E-retail_factors_for_customer_activation_and_retention_An_empirical_study_from_Indian_e-commerce_customers)**.**The project was done using python libraries using jupyter notebook. There libraries used for the processing of data and its analysis are numpy, pandas, sklearn . Along with the help from professionals at datatrained for running various algorithms also from my SME at FlipRobo.

**Introduction**

**Problem statement**:

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.

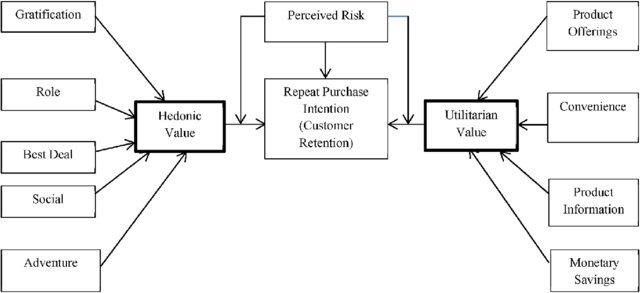
The dataset provided has to be analysed and properly understood so that success factors can be obtained.

**Booming of E-commerce in India**

With development of the internet sector the e-commerce industry in India has grown by leaps and bounds to such an extent that it boosts the acceleration of commerce and finance. With more time this sector promises more growth and with India boasting of a population of over 1.3 billion people, making it a huge potential for faster a bigger growth.

Having such a diverse population will put India in the driving seat for deciding customer satisfaction will help draw more companies thereby propelling the growth of the country even further.

As India has a huge human resource count so it is natural to expect it is the human resource that will decide the growth of the companies providing service. For the companies too is necessary to gain access to all of the country’s human resource as it provides for a better opportunity to gain entry to a huge customer base which in turn yields to bigger profits. But in order to gain customer the companies must also provide value along with the service. In today’s world the value lies in making the experience of customers better, by providing extra benefit along with service, by providing utilitarian values like offers where customers can save money, other product offerings/ information, etc. The companies must also provide hedonic values to its customers like bet deals, a sense of adventure, social value of the product, etc. If these criteria are met with then it will result in better customer retention and at the same increase the customer base. This relationship can be best explained with the help of the following diagram:



So exploring these factors will be help to provide a better insight into the behaviours of the customer and also their preference. It will also help explore details regarding the workings of the companies and overall a better experience can be provided to make the process of online hopping more enjoyable. Also, as India has a very young population it ill also helps expand the customer base.

Analytical Problem Framing

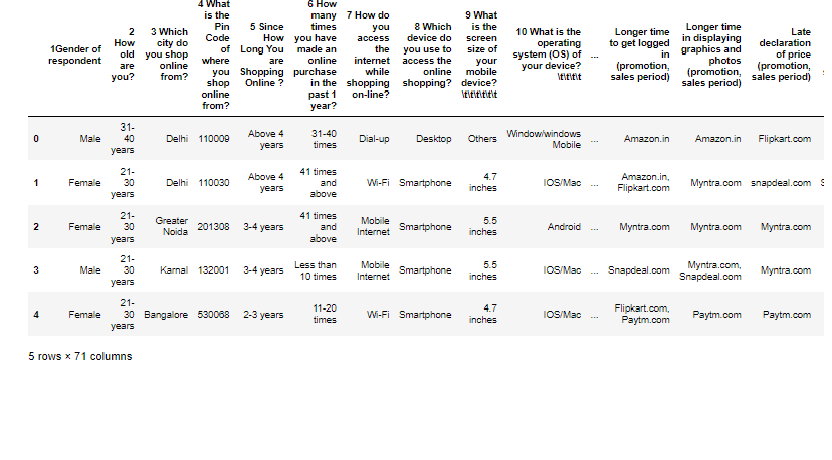
* **Analysis of problem dataset:**

The dataset was provided was mostly categorical data analysis was done mainly by exploring the relationships in between the various independent factors by using univariate analysis, bivariate analysis and multi variate analysis. Some of the columns had overlapping data that had to be extrapolated and multiple columns had to be created. Missing values had to be handled. And data cleaning procedures had to implemented. As the dataset as categorical it had no outliers.

* Dataset formats and tools used:

The dataset contained 269 – rows and 71 columns of which the datatype was object.

The dataset was without missing value(nan- values.)



It had the following columns:

'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 ',

'11 What browser do you run on your device to access the website?\t\t\t ',

'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?\t\t\t\t ',

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

'15 What is your preferred payment Option?\t\t\t\t\t ',

'16 How frequently do you abandon (selecting an items and leaving without making payment) your shopping cart?\t\t\t\t\t\t\t ',

'17 Why did you abandon the “Bag”, “Shopping Cart”?\t\t\t\t\t ',

'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',

'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?',

* Tools used :

import pandas as pd

import numpy as np

import seaborn as sns

import matplotlib.pyplot as plt

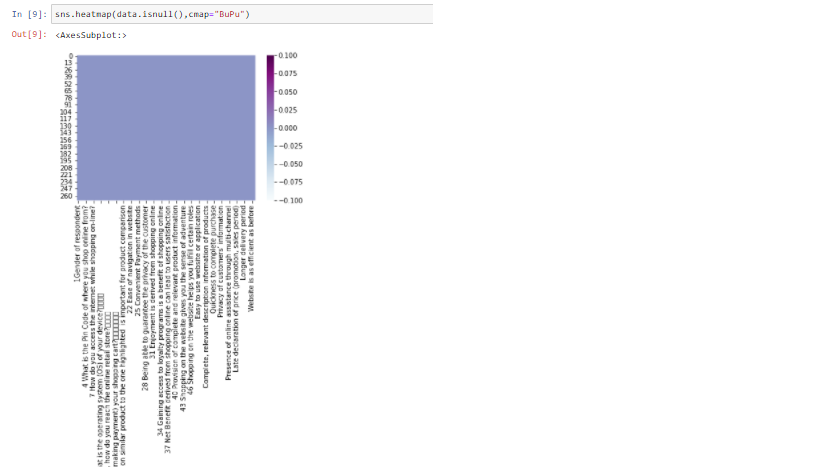
%matplotlib inline

import warnings

warnings.filterwarnings('ignore')

* Visualizations :

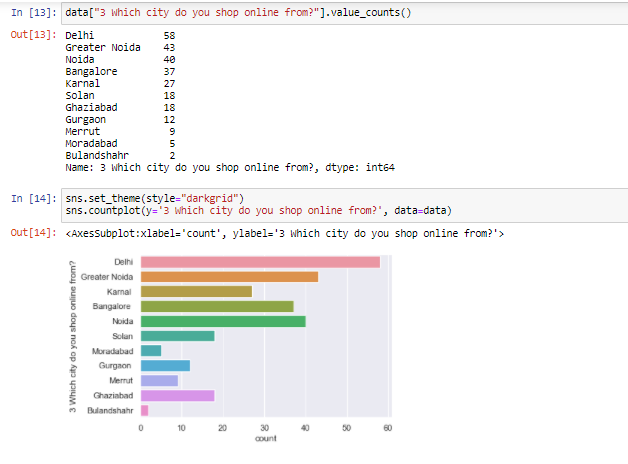
The data visualization as achieved using the matplotlib and the seaborn libraries some of which are as follows:



The heatmap represents the lack of missing values in the dataset.



The plot shows the count of the Gender (male/ female) shopping from the online retail stores. The number of female were higher (181) than that in comparison to the male (88).



Shows that Delhi is the city with highest number of online shoppers. With a count of 58.

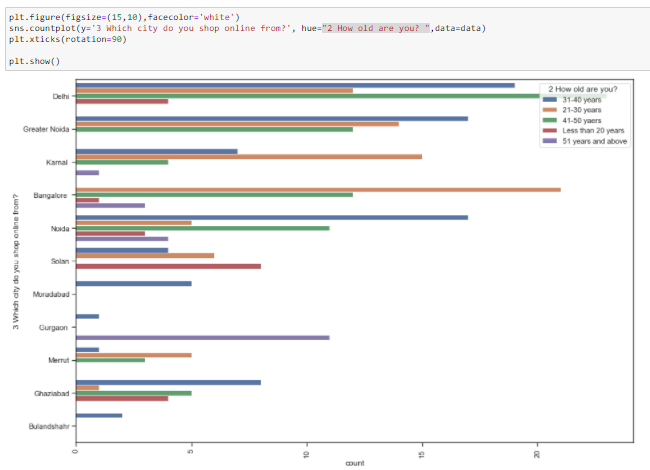


As it can be see that all the cities have female shoppers with Greater Noida having the most number of female followed by Bangalore, whereas The highest number of male shoppers are from Delhi followed by Noida. However, least number female shoppers are from Meerut whereas the least number of male hoppers are from Greater Noida, Gurgaon and Ghaziabad with Merrut not having any male shopper.



As it can be seen there are 5 different groups of shopper based on there experience with shopping which is as follows:

* Above 4 years
* 3-4 years
* 2-3 years
* 1-2 years
* Less than 1 year.

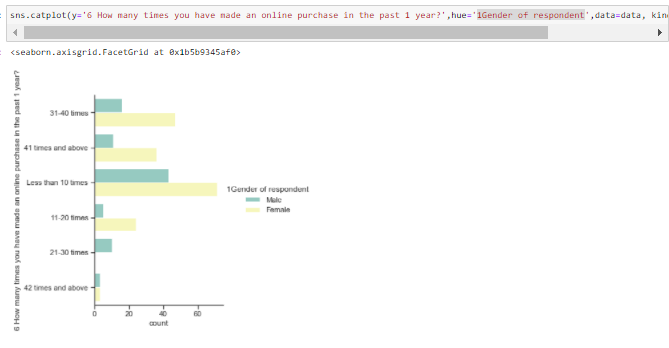


of the cities Noida has the most diverse age groups shopping online.

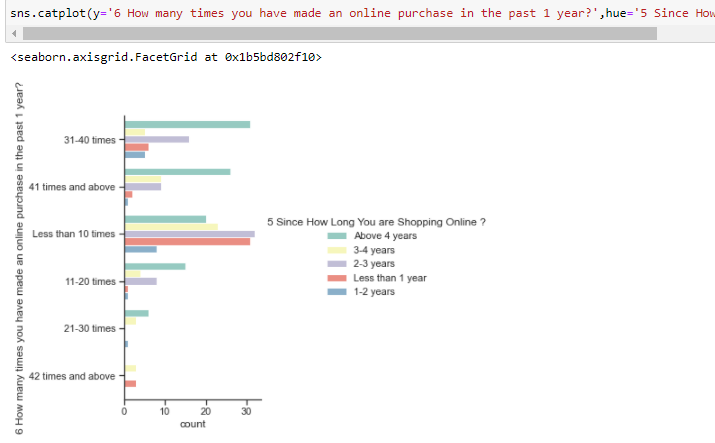
The age group of 31- 40 years and 21- 30 years are most in number while Delhi has the highest 41-50 years age group followed by 31-40 years where as Bangalore has the highest 21-30 years shopping online.

Of the all the cities Bulandshahr has the lowest number of people shopping online and it has the specific age group 31-40 years.

Gurgaon has the highest number of 51 years shopping online.

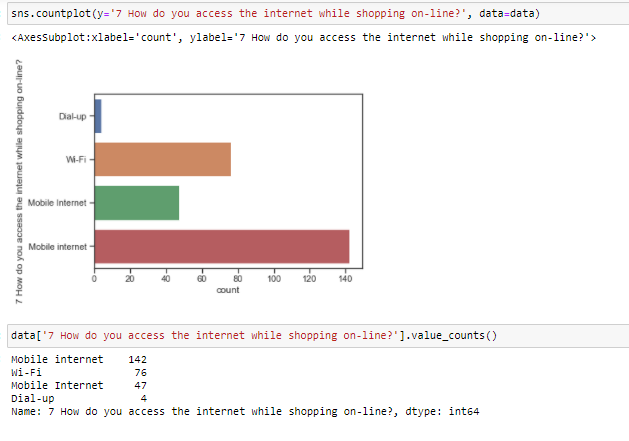


The above plot shows that in all the count of shopping online for the past one year the number of female are more in number than in comparison to the male.

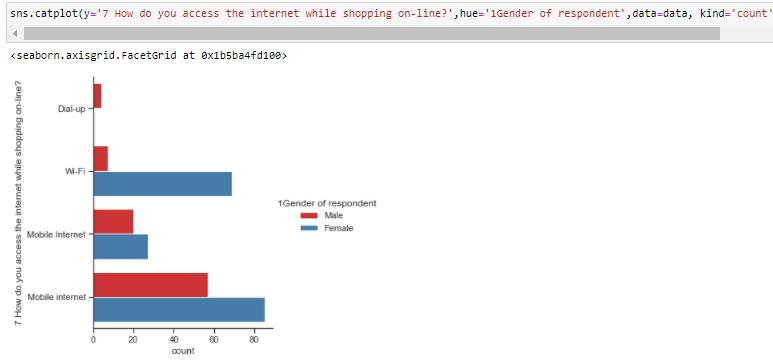


The plot shows that highest number of online purchases made was 41 times and then followed by 31- 40 times and for both cases the group of people with the highest online shopping experience(above 4 years) are responsible for hoppin fthe most number of times.

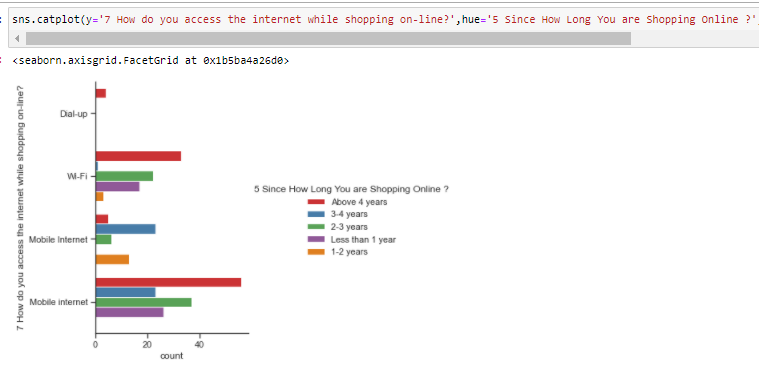
* It also shows that the highest number of times that online shopping was done was 42 times and it had two specific groups of people :
  + 1st group new to online shopping doing shopping for the past 1 year
    - 2nd group have been shopping online for the past 3-4 years



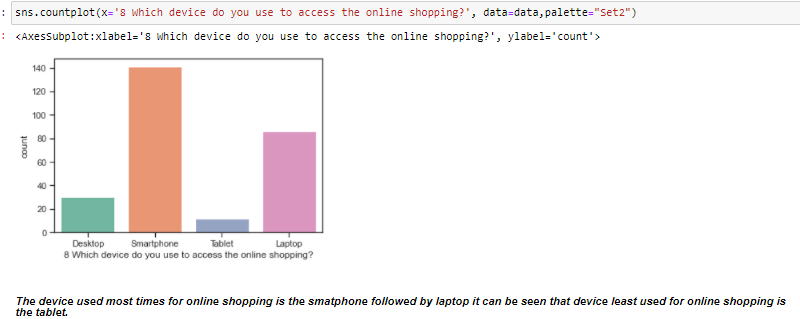
It can be seen that mobile internet is the most popular way of accessing the internet followed by WiFi.



Of the two most acceptable ways to shop online for both cases females are the majority users. It might hint on the fact women are more into shopping online based on the data.

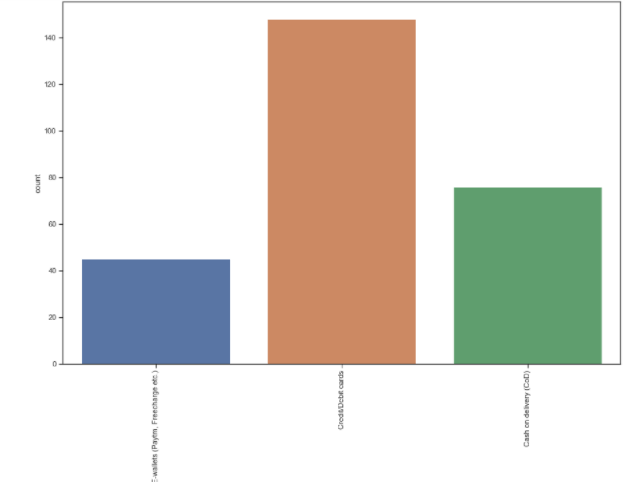


IT CAN BE SEEN THAT FOR PEOPLE WITH AN EXPERIENCE OF ABOVE 4 YEARS MOST PREFER THE MOBILE INTERNET THEN FOLLOWED BY wIFI.

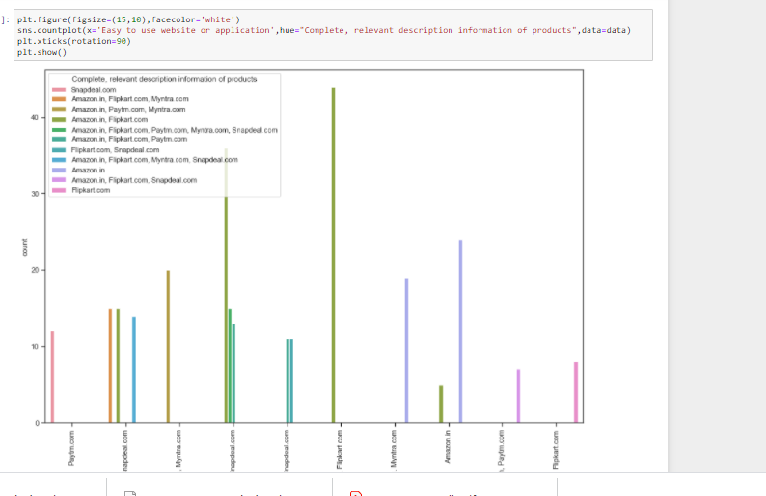


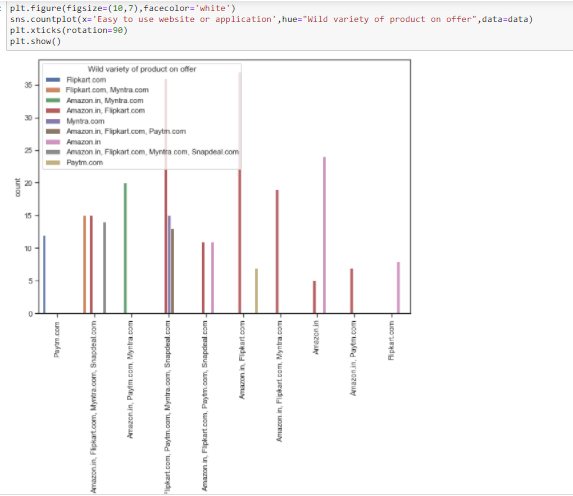


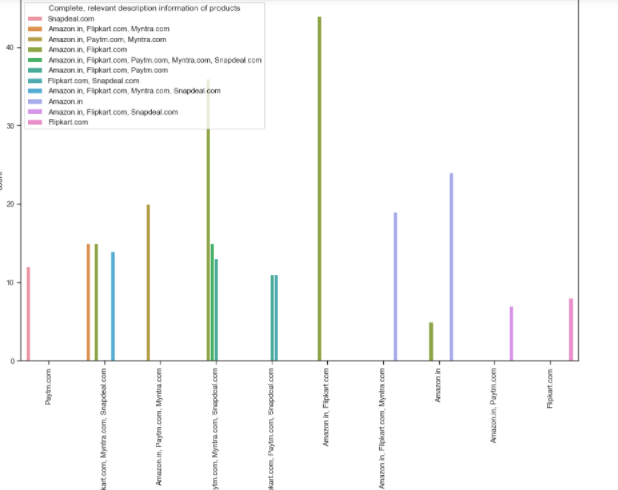
The two most used ways for accessing the online retail store is by using the search engine and via application.



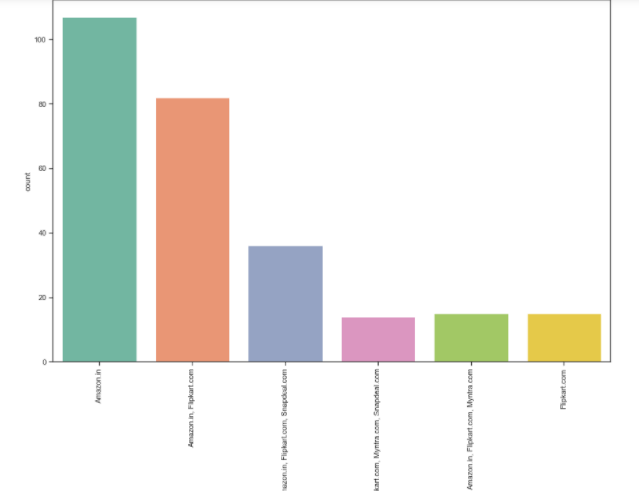
MOst expected payments option is credit/debit cards followed by COD and the least used one is E-wallets.





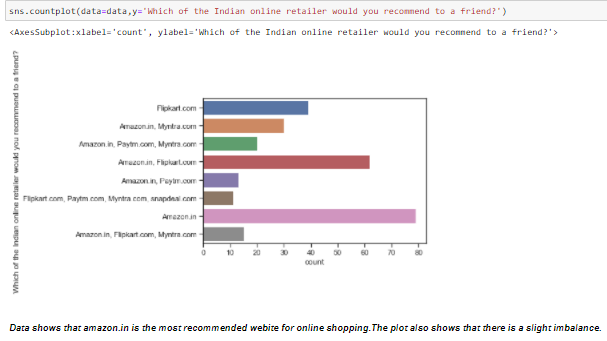


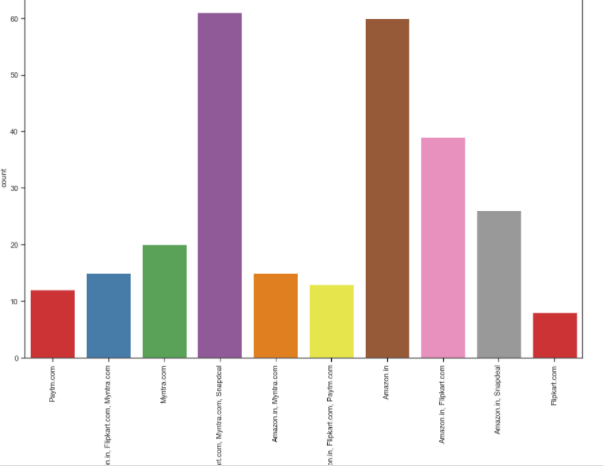
The above plots explore relationships like –“Easy to use website or application”, “Complete, relevant description information of products”, “Wild variety of product on offer” with the online retailers and in all the cases Amazon.in was the top-most in all aspect.



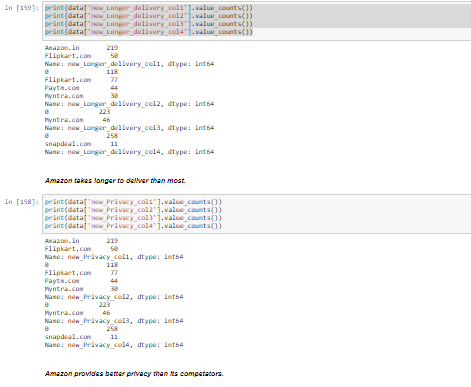
The plot shows that Amazon is the best for making speedy deliveries.

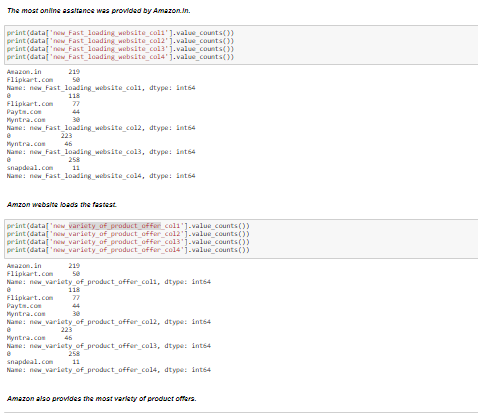
So on exploring features likes these we get an insight into the wants of customers and what makes Amazon.in better than its peers in the business which is evident by the next slide.

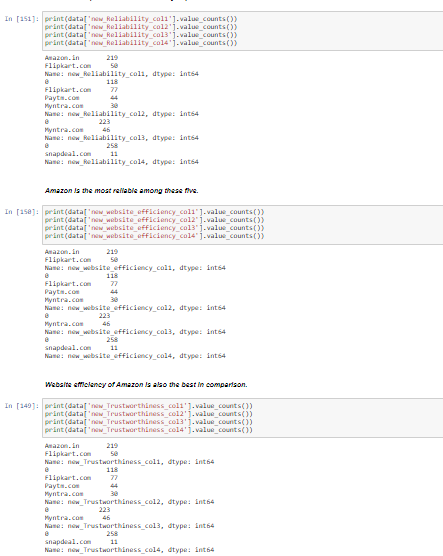




The plot is showing the relationships between the online retailers and assistance provided by them to the users (online). Here again Amazon.in provides the most assistance.







Of the feedbacks from 269 people 219 like shopping from Amazon.in and 127 people like hopping from Flipkart.com , 76 people liked shopping from Myntra.com , 44 people from Paytm.com and only 11 from Snapdeal.com making the website least preferred by the people.

It can be understood that people like to switch in between websites and shop specially in between Amazon.in and Flipkart.com.

The dataset was able to show why Amazon was the favourite of the users as it was most recommended online retailer among users. As for various cases like "trustworthiness" , "website efficiency", " Reliability", "variety\_of\_product\_offer", etc it was seen that the customers preferred Amazon.in(as 219 people preferred shopping from it) more in comparison to the other online retailers. After Aamazon.in, Flipkart .com was the second most preferred website as voted for by 127 people.

Conclusion:

The dataset contains a detailed information about the e-commerce industry of India with the passage of time multiple Online-retailers have come to create a highly efficient market place where product can be bought an sold from the safe confines of ones' home. The dataset helps provide light on some of the important factors that help define the company to the customers which in turn helps to build a bigger better customer base which is loyal. The dataset dive into a lot of information starting from the Gender of the users to their preferences of using different devices for accessing the internet, the browsers that they use, the screen size of their device , preferences about the website , delivery of products and also the loading speed , how much time it takes to log in ,etc. It actually goes on to show how well the data collection has been done . It also shows how much attention to detail has been given about the customer satisfaction. One thing that is clear from the dataset provided is the fact all the online retailers mentioned namely

- Amazon.in

- Flipkart.com

- Myntra.com

- Paytm.com

- Snapdeal.com

care deeply about providing a better experience to the shoppers and want to make it very much to the liking of its users.

What was interesting to observe was that the more a retailer is able provide a better experience to it's shoppers the more hoppers are drawn to it. It could be seen that various people from different geography and age groups were shopping online and mostly the shoppers were female.

It could also be seen that Amazon.in was able to perform with high frequency in all the categories that dealt with providing better customer experience.

The dataset was able to show why Amazon was the favorite of the users as it was most recommended online retailer among users.

It can be seen how providing hedonic and utilitarian values can provide a better environment that can usher more customers and also create a proper retention channel. From the analysis done I can report to the best of my knowledge that Amazon.in as an online retailer is providing better comfort than its peers and the feedback of the users is advocative the fact. So for other online retailers to improve there scores in comparison to Amazon they have improve in certain areas where they are lacking . The best part about the dataset is that it helps to provide pointers as to the places where the other companies can perform better. At the end in can be said that such progress can be brought about by paying close attention to detail and by getting continuous , periodic feedbacks from the customers and implementing changes.