**PROJECT REPORT**

ON

**CUSTOMER RETENTION**

**OBJECTIVE** :-

As we all are aware from current scenario, how online shopping takeover the market of offline shopping. But there is also heavy competition b/w companies to take the market. The biggest challenge have to face that “Customer Retention”. Because in current situation customers have too many options to switch between. Now the challenge is how to add new customers, and how to retain old customers as well.

Our objective behind analysis of the data is to remove out the factors that affect the customer retention.

**MOTIVATION** :-

Deciding on which factor have to work for effective customer retention is never a easy task. It needs so much practice to found out the proper solution, whether it is customer service, delivery,payment,application etc. There is always be a dilemma which factor to choose or not.

**DATA SET** :-

We use “customer\_retention\_dataset” .

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

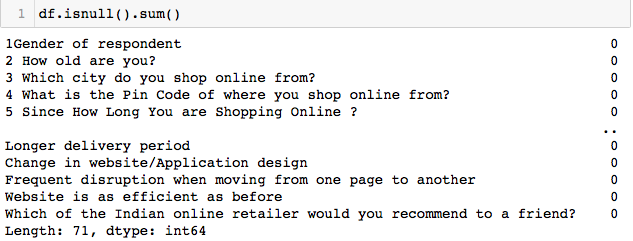
Screen Shot 2022-04-15 at 3.17.11 PM

**Variables in data set:-**

|  |
| --- |
| **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?** |
| **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?** |
| **16 How 4 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** |
| **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?** |

The Data Set has 269 rows and 71 variables/columns .

**NULL VALUES :-**

****

The Data Set has no null values present in the columns. So there is no need to perform any operation to fill the null values.

**DATA TYPES PRESENT IN DATA SET :-**

**Screen Shot 2022-04-15 at 3.20.49 PM**

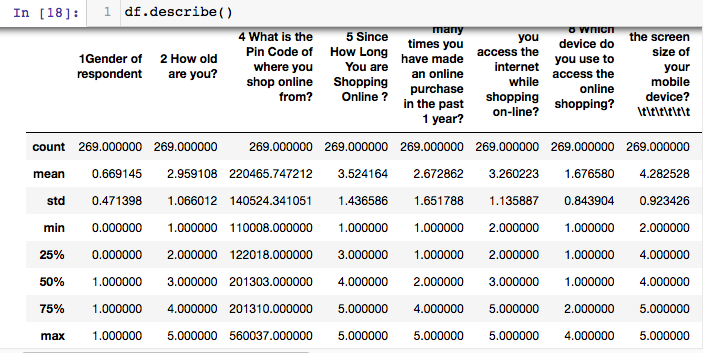
There are two types of data types present in data set.

Int

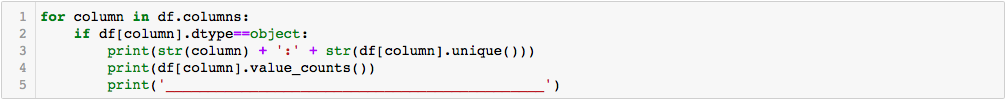
Object

There are 46 int type and 5 object type variables are present in data set.

**STATISTICAL DESCRIPTION :-**

****

**CHECKING UNIQUE VALUES IN ALL FEATURES :-**

****

The unique values are the values which are returns in order of appearance.

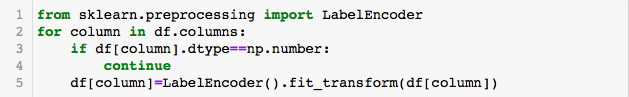
**ENCODING THE FEATURES USING LABEL ENCODER :-**

There were two types of features present in data.

Categorical and Neumerical.

The primary reason we need to convert categorical columns into neumerical columns so that machine learning algorithm can understand it

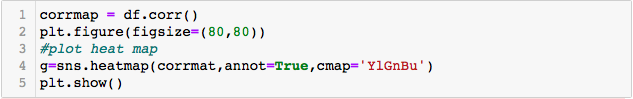
**Label Encoding** is a popular encoding technique for handling categorical variables. In this technique, each label is assigned a unique integer based on alphabetical ordering.

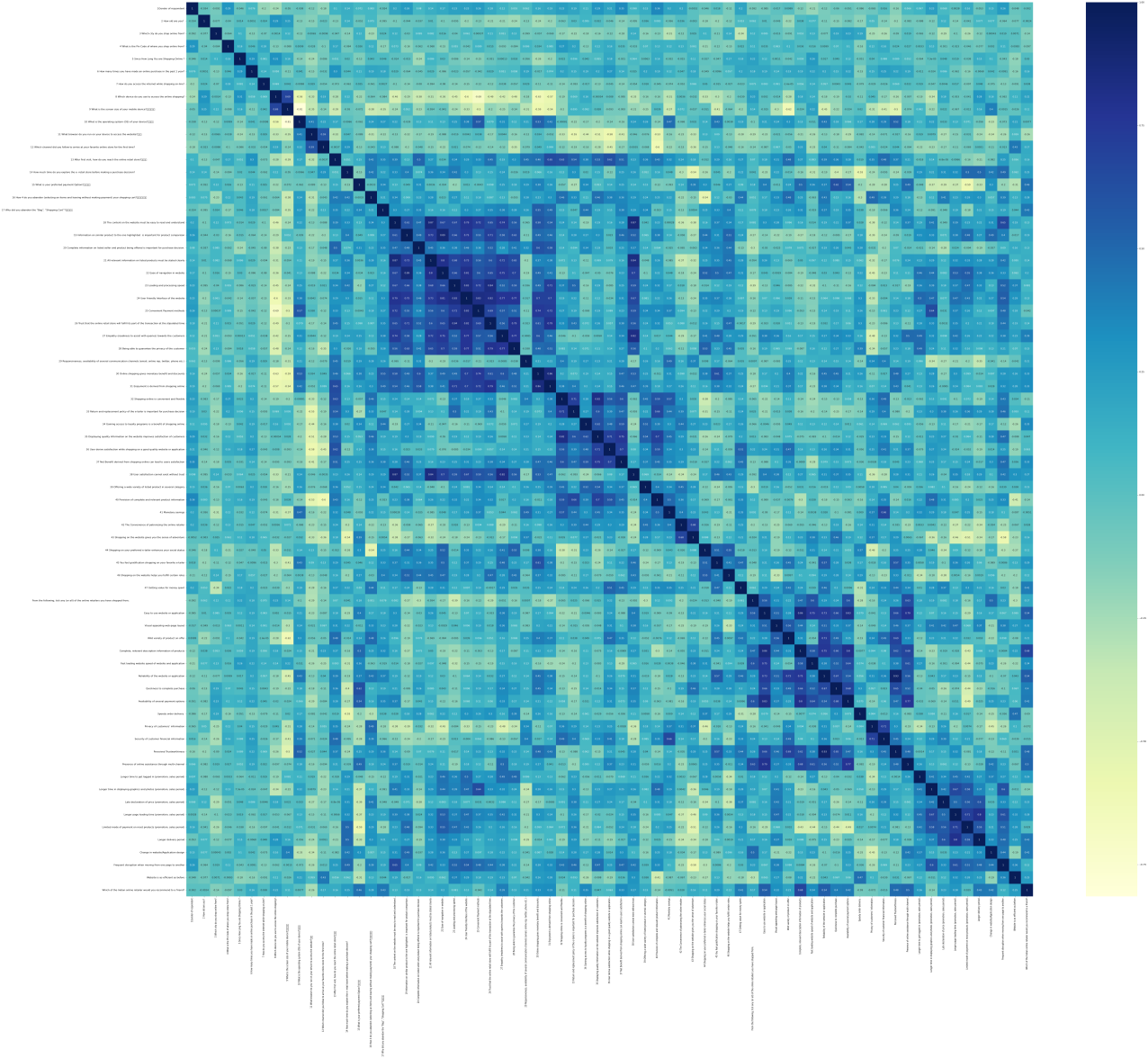


**CORRELATION B/W VARIABLES :-**

**Screen Shot 2022-04-15 at 3.55.50 PM**

Correlation shows the mutual relation between variables. By correlation we can see that how variables are correlated with each other , how much each variable affect other variables.

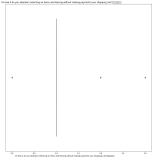
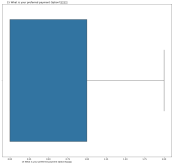
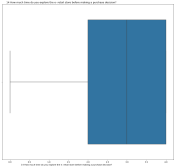
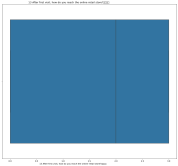
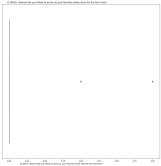
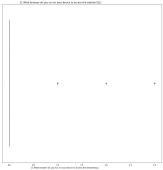
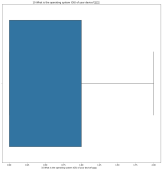
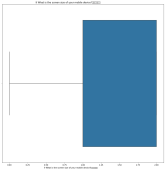
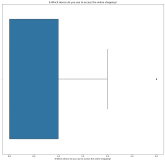
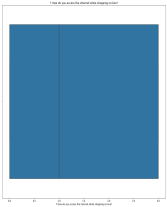
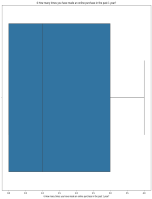
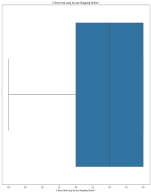
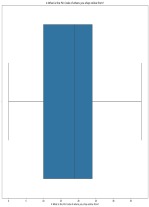
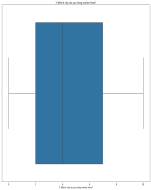
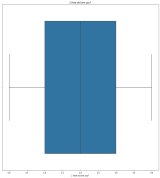
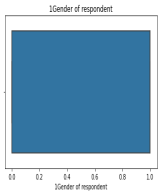




According to the heat map we can see that some of the areas of map are dark, which shows the high correlation between variables.

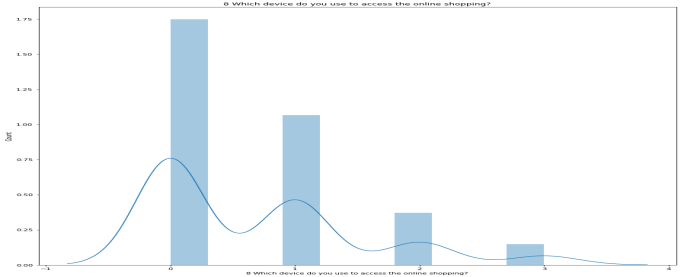
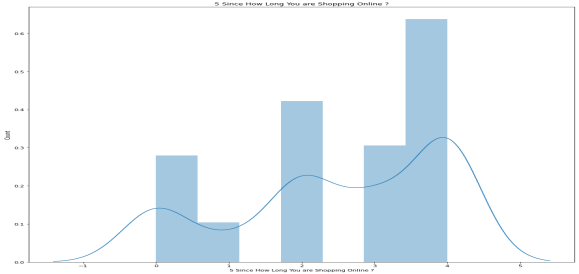
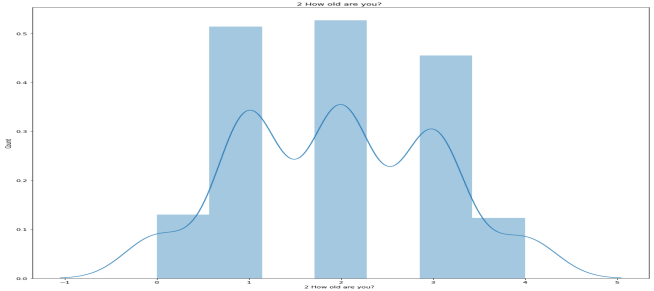
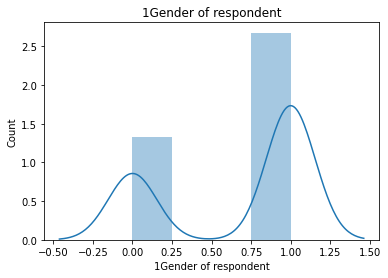
**CHECKING OUTLIERS IN DATASET :-**

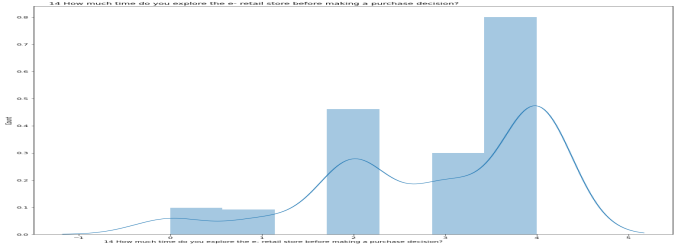
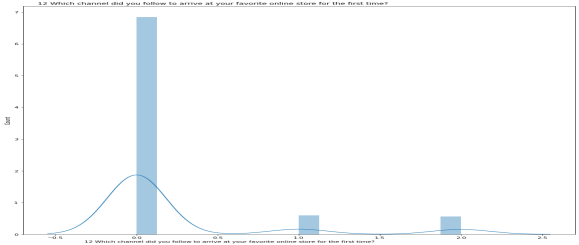
An outlier is a single [data point](https://www.techtarget.com/whatis/definition/data-point) that goes far outside the average value of a group of statistics. Outliers may be exceptions that stand outside individual samples of [populations](https://www.techtarget.com/whatis/definition/population) as well. In a more general context, an outlier is an individual that is markedly different from the norm in some respect.

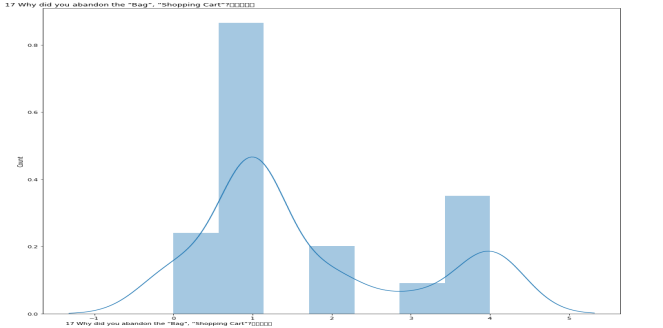
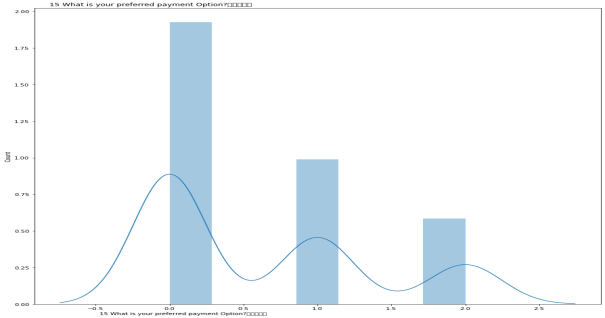


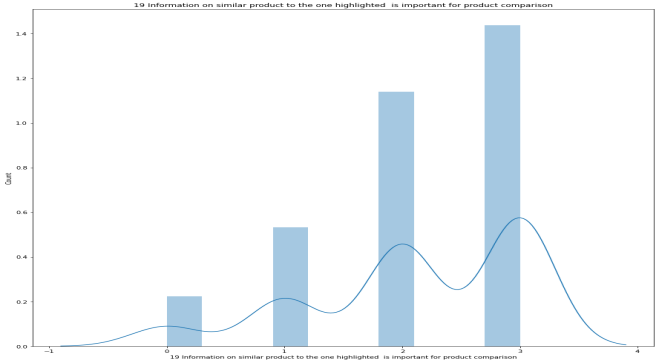
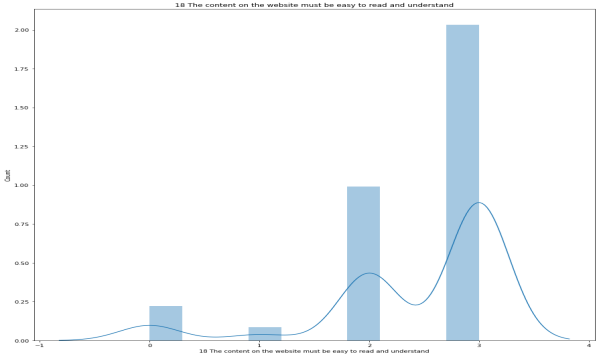
As we can see there are some outliers presents in data , but that are present in features that dos not affect the target variable directly. So we can drop that columns.

**SOME DISTRIBUTION OF DATA :-**

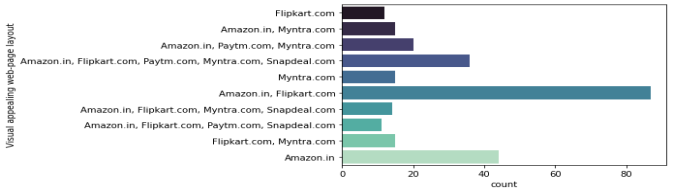
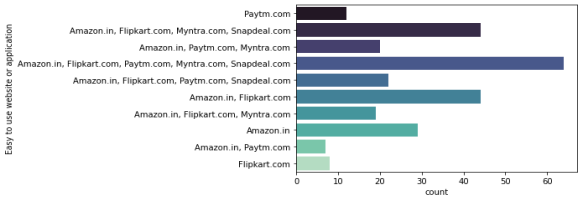


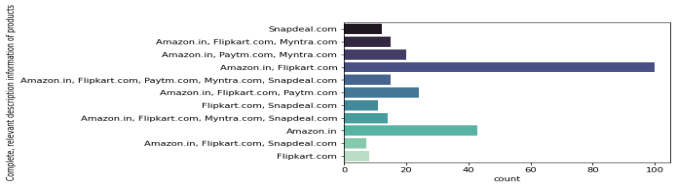
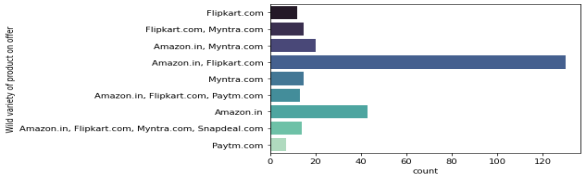




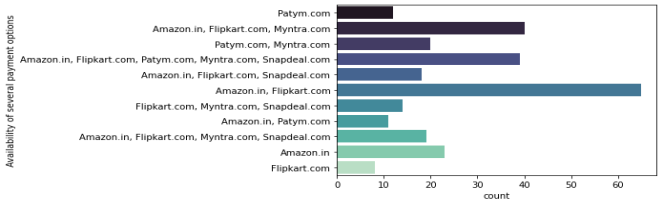
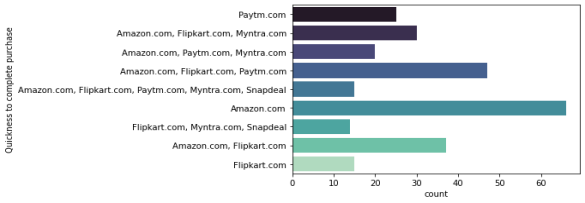


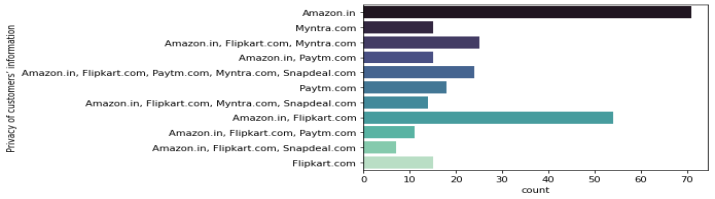
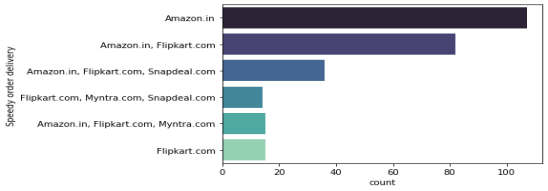
**VISUAL ANALYSYS OF SOME OTHER FACTORS :-**

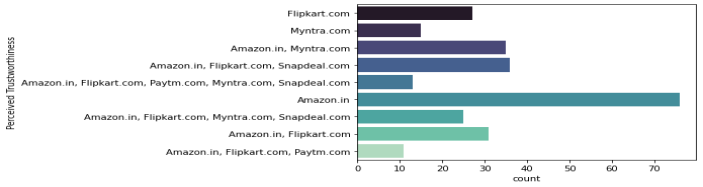
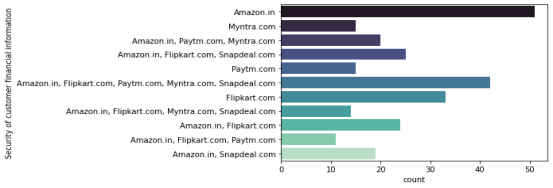


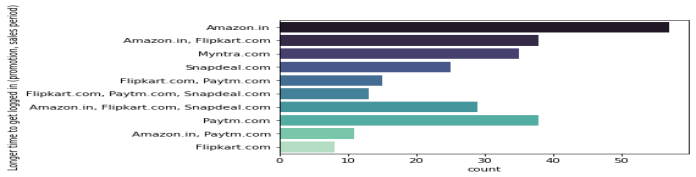
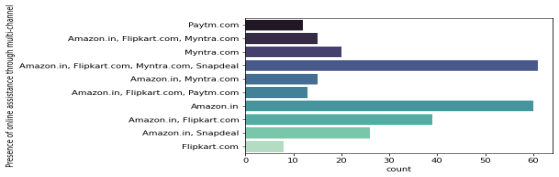


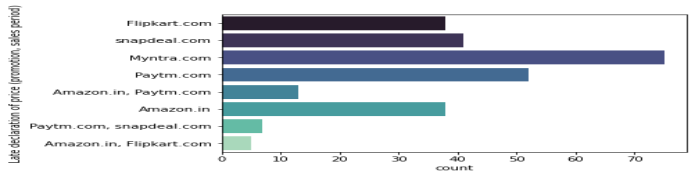
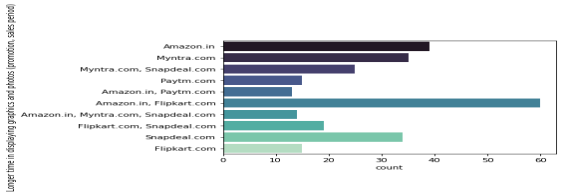


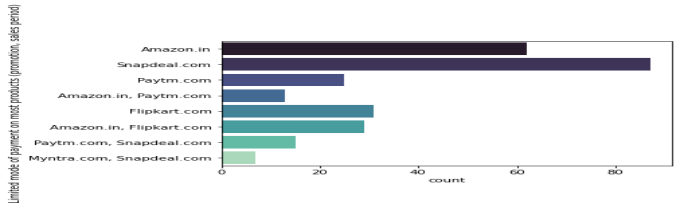
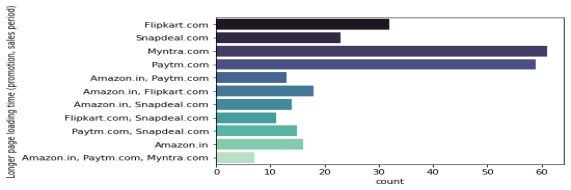


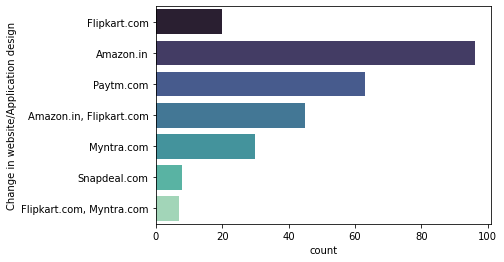
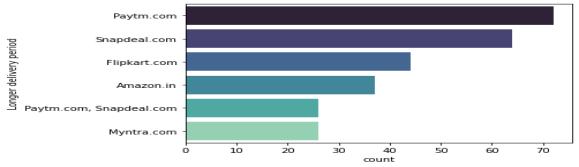


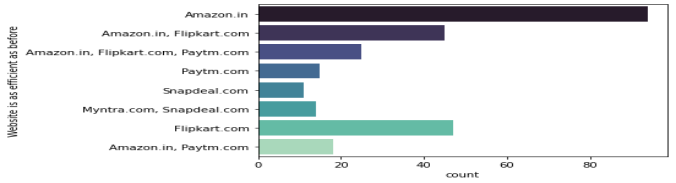
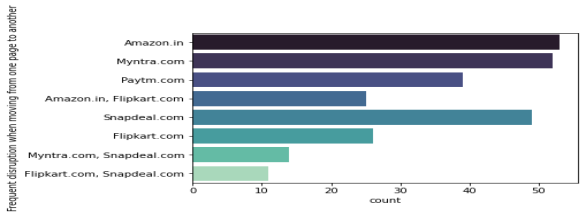


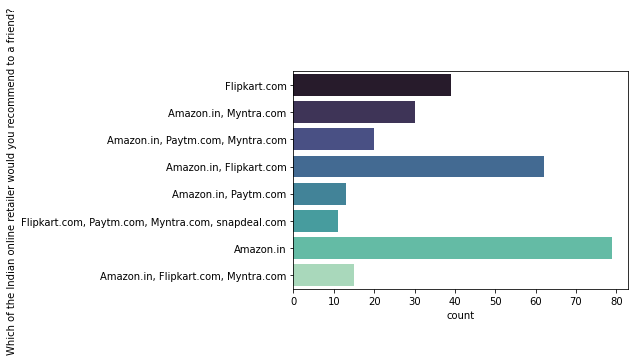












As we can see in above countplot graphs the “Amazon” is working very well in most of the areas.

But areas like website i.e. Login Time and smooth Transition from one page to another page in website or application page have to be consider as a weak point.

Other companies are also doing well but they have to improve their services in many areas .

Now we can say for better customer retention the companies should have target female customers , customers age group of 21-30 , customers who have 3-4 years of shopping history.

On the other hand they should have also improve their service related area like app improvement,website improvement,customer care services, better payment options,delivery options and delivery time,privacy of customer's personal data and financial data,variety of offers,variety of listed products,better information of products.

Better the customer's experience, better the customer retention.