

**CUSTOMER RETENTION PROJECT**

****

Jasmine kaur

Data science intern, internship 18

**INTRODUCTION**  
Electronic commerce, or commonly known as e-commerce is defined as online trading of products and services. E – Commerce operates using World Wide Web for at least one part of the transaction. There can be other mediums like e-mail, phone calls, etc.

E- Commerce includes the following:

* Online website that shows the product and services along with their description in order to generate sales directly.
* There are marketplaces that give platform to other business firms to sell their products and services. Thus, this includes third party involvement in order to make business-to-business or consumer-to-consumer sales.
* This also incorporates business-to-business sales.
* Internet marketing is used to inform and generate sales from prospective buyers.

According to India Brand Equity Foundation:

1. India’s e-commerce order volume increased by 36% in the last quarter of 2020, with the personal care, beauty and wellness segment being the largest beneficiary.
2. E-commerce in India is at an all-time high in 2021 owing to the country’s 780.27 million internet users as of May 2021.
3. There are certain policies that support the online business such as:

* 100% FDI is allowed in B2B e-commerce.
* 100% FDI under the automated route is permitted in the marketplace model of E-commerce.

1. Increased digital literacy acts as a fuel in the investment in e-commerce firms as more and more traders find this attractive to switch or expand their business to online platform.

BENEFITS OF ONLINE SHOPPING:

1. *CONVINIENCE*: The online shopping provides the ease of shopping while having the convenience of sitting at home. Everything is just a click away. It provides the convenience of shopping 24 hours without waiting in the long queues.
2. *PRICE COMPARISON*: It is possible to compare the prices and offers available from all the different websites. Thus, in the end online shopping helps in getting the better prices.
3. *MORE VARIETY*: A much greater range of goods and services are available in online shopping. This much variety is not possible when has to go from shop to shop physically.
4. *NO CROWD*: There is a problem of facing crowd and chaos during the festival season. This problem can be solved using online platforms. The problem of traffic can be avoided with this. It gives the leisure and easiness and helps avoiding the noisy environment.
5. *DELIVER GIFTS DIRECTLY*: In the digitalized world, sending gifts has become one of the easiest steps as it can be delivered directly to the desired place.

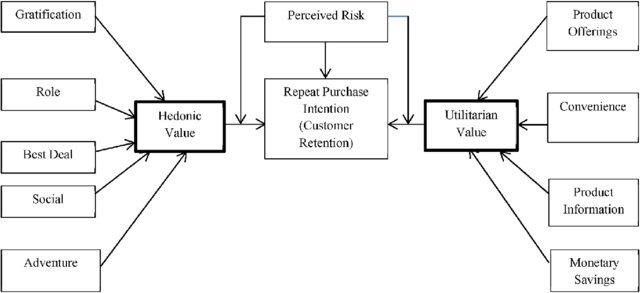
**INDIAN E-COMMERCE SCENARIO:**

Everything from daily grocery to the most expensive phones and laptops is available online. The number of B2B transaction has increased drastically in Indian market. These websites have become daily necessity of people. Also, these platforms act as advertising other brands and thus increases the online sales collectively.

Online platforms of shopping in India are experiencing a period where there is an increase in the number of online sales like never before. The top market players in India belong to telecommunications sector, FMCG sector, services, and mobile sector. High amount of profit earned is then spent on research and development to reach at the new heights of success. On the other side, consumers also enjoy benefits like free gifts, or special offers while making online sales.

**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.



**OBJECTIVE:**

* To analyse the data collected from the Indian online shoppers.
* To bring out the insights from the data which is very critical to understand the factors that influence customer satisfaction?
* This will lead to success of the e-retail business.
* To reveal the factors that leads to satisfaction of the consumer.

**ASSUMPTIONS AND TOOLS USED:**

* Assumed that Mobile Internet and Mobile internet are the same feature, so combined the both.
* Combined 41 times and above with 42 times and above in the how many times you shopped last year.

**Hardware and Software Requirements**:-

Hardware:4GB RAM, 64Bit O/S.

Software Tools: Python 3, Google Colaboratory, MS-Excel, Notepad.

**ANALYTICAL PROBLEM FRAMEWORK:**

In this project, input data has both kinds of data types i.e. object as well as numerical. We have performed the EDA with the help of univariate and multivariate data analysis tools. Along with that various statistical findings have helped us understand the data in a better and more appropriate manner. We checked for various characteristic of the data such as correlation, skewness, missing values and outliers.

**STATISTICAL SUMMARY:**

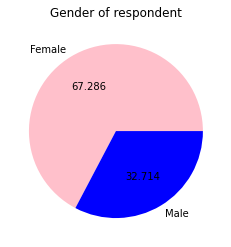
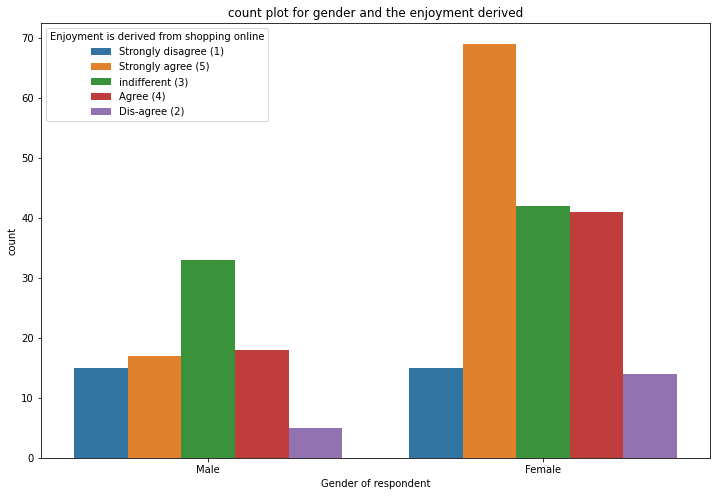
* The dataset has 269 rows and 71 columns
* There is no null value in the data
* There is presence of outliers and skewness in the data
* Some of the features are of object data type while others are numerical

**SEGREGATION OF FEATURES:**

The features are segregated on the basis of:

1. Personal information
2. Repeat purchase behavior of the customer
3. Unsatisfactory service

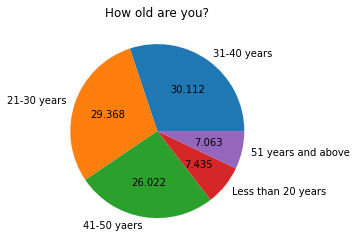
**PRESENTATION OF DATA AND ANALYSIS:**The first feature that we considered is **gender**. There are different patterns that depend upon the gender of the respondent. This is because there are generally females like to shop more and this also happened in the survey. The number of female respondents is twice as compared to the male respondents. Also, females enjoy shopping more than that of males. This can be seen in the histogram below.

The next factor which is very important while considering the online shopping trend is the **age group** of the respondents. There is a certain pattern that follows in every age group like their tendency to shop and to use internet etc. The pie-chart below shows that:

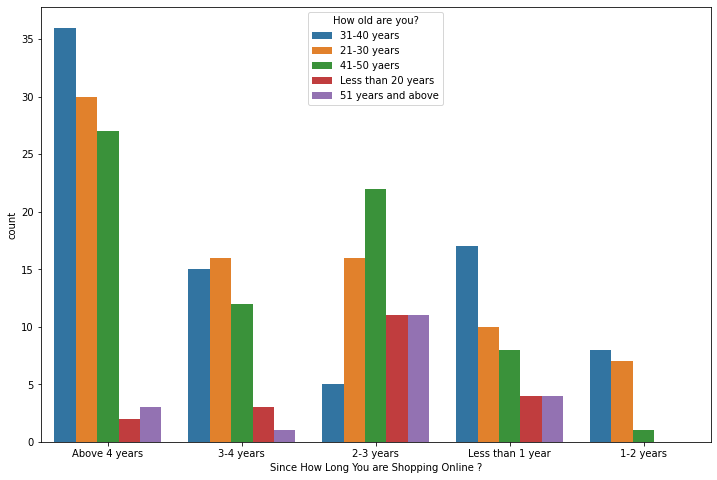
* Most of the respondents are in the age group 31-40 years.
* This is followed by the age group 21-30 years.
* The third highest number of respondents is in the age group 41-50 years.

 This is then followed by teenagers and then the senior citizens

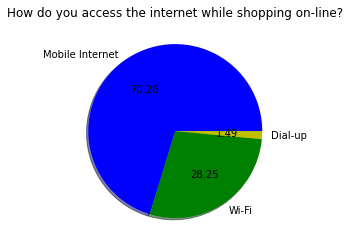


There is also a trend that specifies that there is a **relation between the age group and number of years one is shopping online**. The histogram below shows the following patterns:

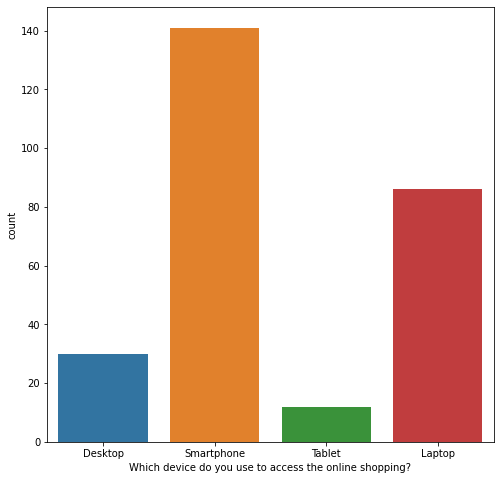
* Highest number of people who have been shopping online lies between the age group 20-50 years.
* The graph represents that the highest number of teenagers and senior citizen are shopping online from 2-3 years.
* There are no teenagers and senior citizens who are shopping online from 1-2 years in the survey data.



Another important feature is **how the respondents use internet and on what device**. Around 70% of the people use mobile internet to shop online. The pie chart is shown below:

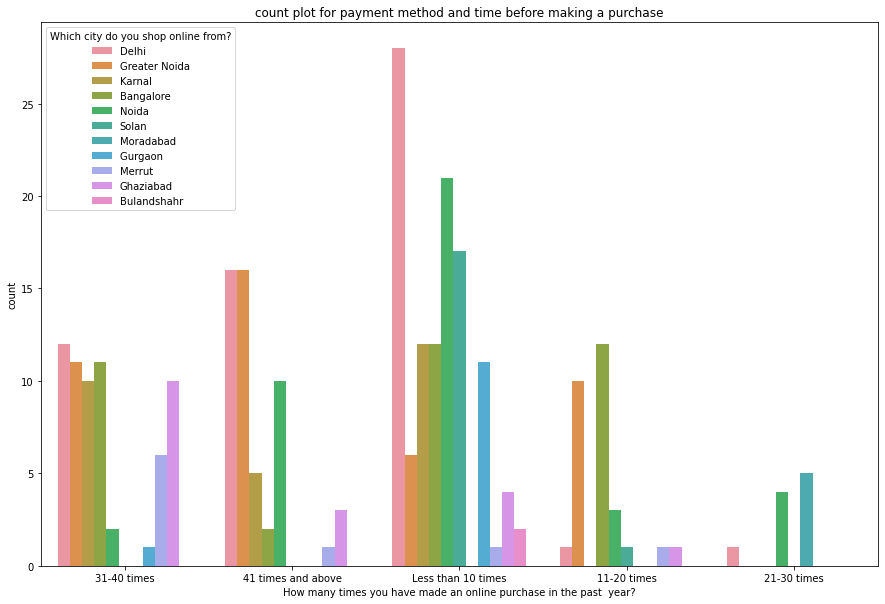


People use different devices while shopping. The distribution of the same is as follows:



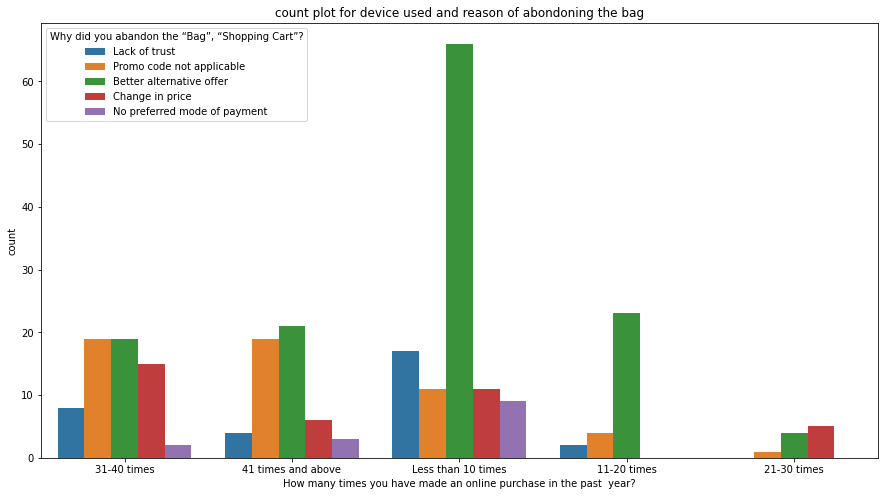
* The highest number of people use Smartphone for doing online shopping.
* The second highest used device is laptop.
* Some people use Desktop to make online purchases.
* Vey few of them use tablet to do online shopping.

The **city** defines various factors about the online shopping trend.



* Most of the people have purchased less than 10 times in the past year.
* Mostly, those who purchased between 11 to 20 times, stay in Noida or greater Noida.
* Very few people purchased 21-30 times, those who purchase stay in Delhi, Noida or Gurgaon.
* Those who purchased 31-40 times are generally from Delhi, Greater Noida, Karnal, Bangalore, and Ghaziabad.
* Most people who purchased more than 41 times are either from Delhi or Greater Noida. Also, some of them are from Bangalore or Karnal or Merrut or Bulandshahr.

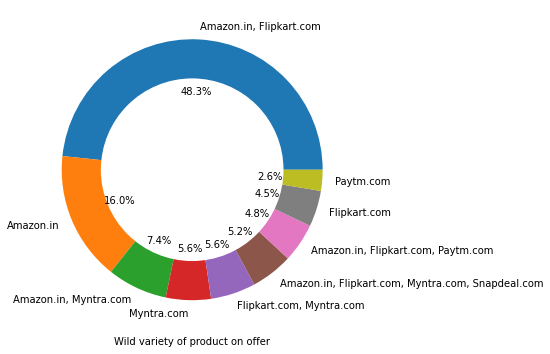
Those who surf online have the tendency to put the products they like in the cart and then rather than purchasing them, they abandon the cart. There are various reasons that leads to this behavior of the customers like availability of better alternative offer or lack of trust etc. The histogram below shows the **trend between the number of times a person has shopped online and the reasons why they abandoned the cart/bag**.



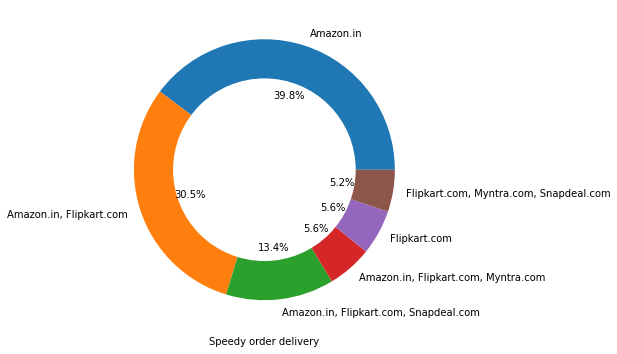
* Those who purchased less than 10 times in the past year mostly abandoned their bags because they were able to see some better alternative offer. In this category other reasons for abandoning is quite less as compared to the better alternative offer.
* Those who purchased 11-20 times, abandoned their bags because of the better alternative offer. Negligible percentage of people abandoned due to Lack of trust and promo code not applicable. There is no other reason for abandoning bags that purchased 11-20 times.
* Those who purchased 21-30 times, abandoned their bags the least. Still, those who did, the reasons include promo code not applicable, better alternative offer, change in price.
* Those who purchase 31 times to 40 times and 41 times and above have almost the same distribution for the reasons of abandoning the bag. It is just those who made 31-40 purchases, abandoned their bags more than those who made more than 41 purchases due to the lack of trust and change in price.

There are various factors that explain the retention rate of the customers. The pie charts and histograms depict some of those factors:

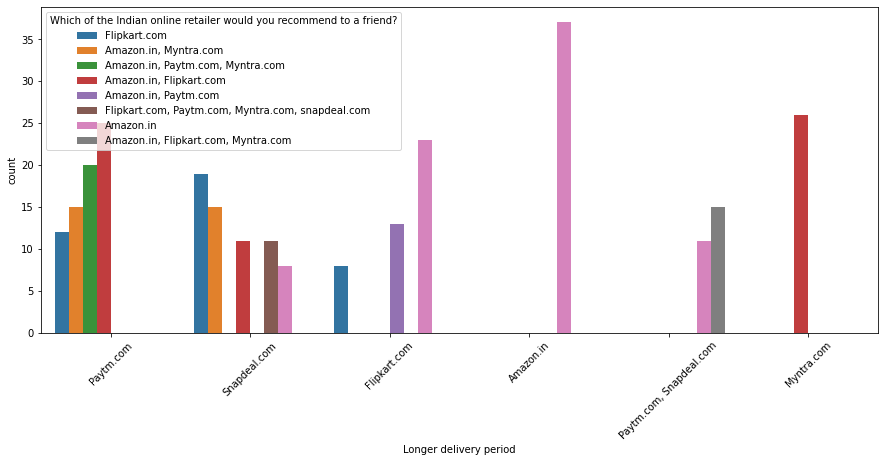
WIDE VARIETY OF PRODUCTS ON OFFER:



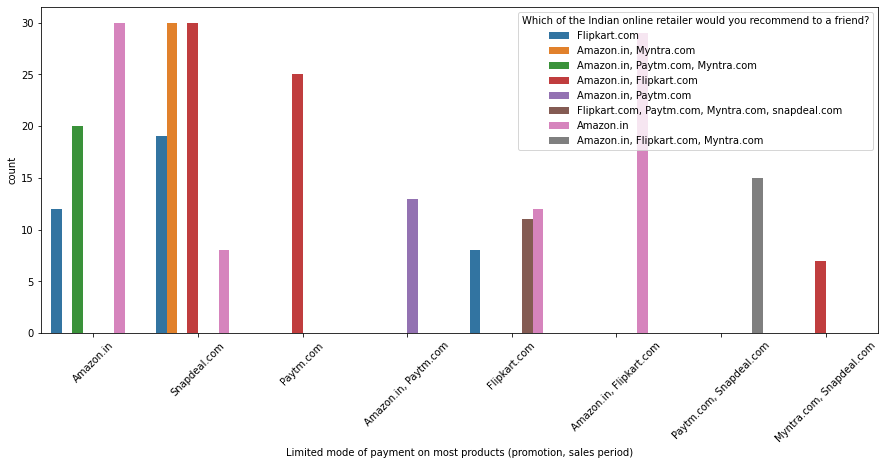
SPEEDY DELIVERY:



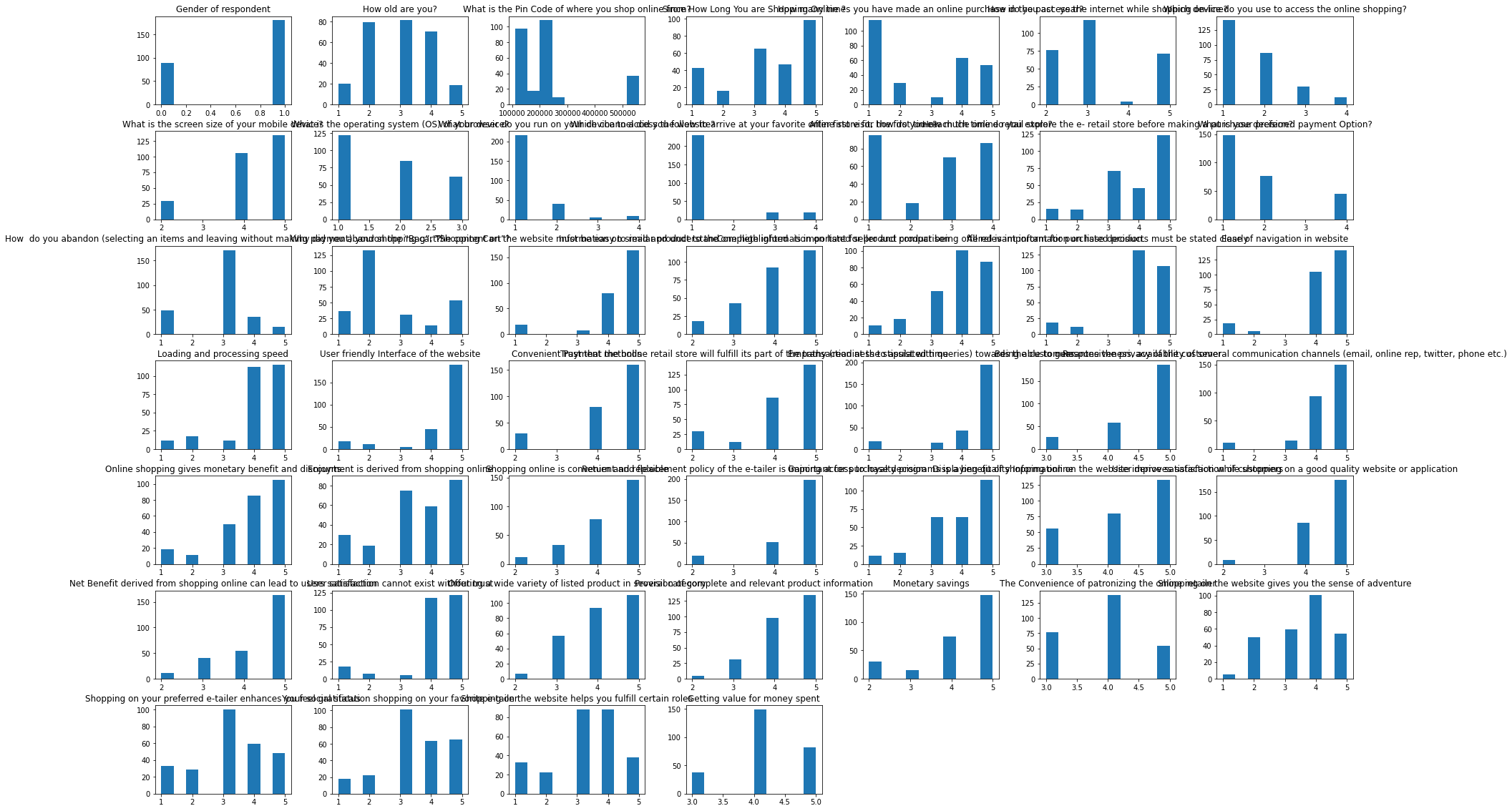
LONGER DELIVERY TIME:



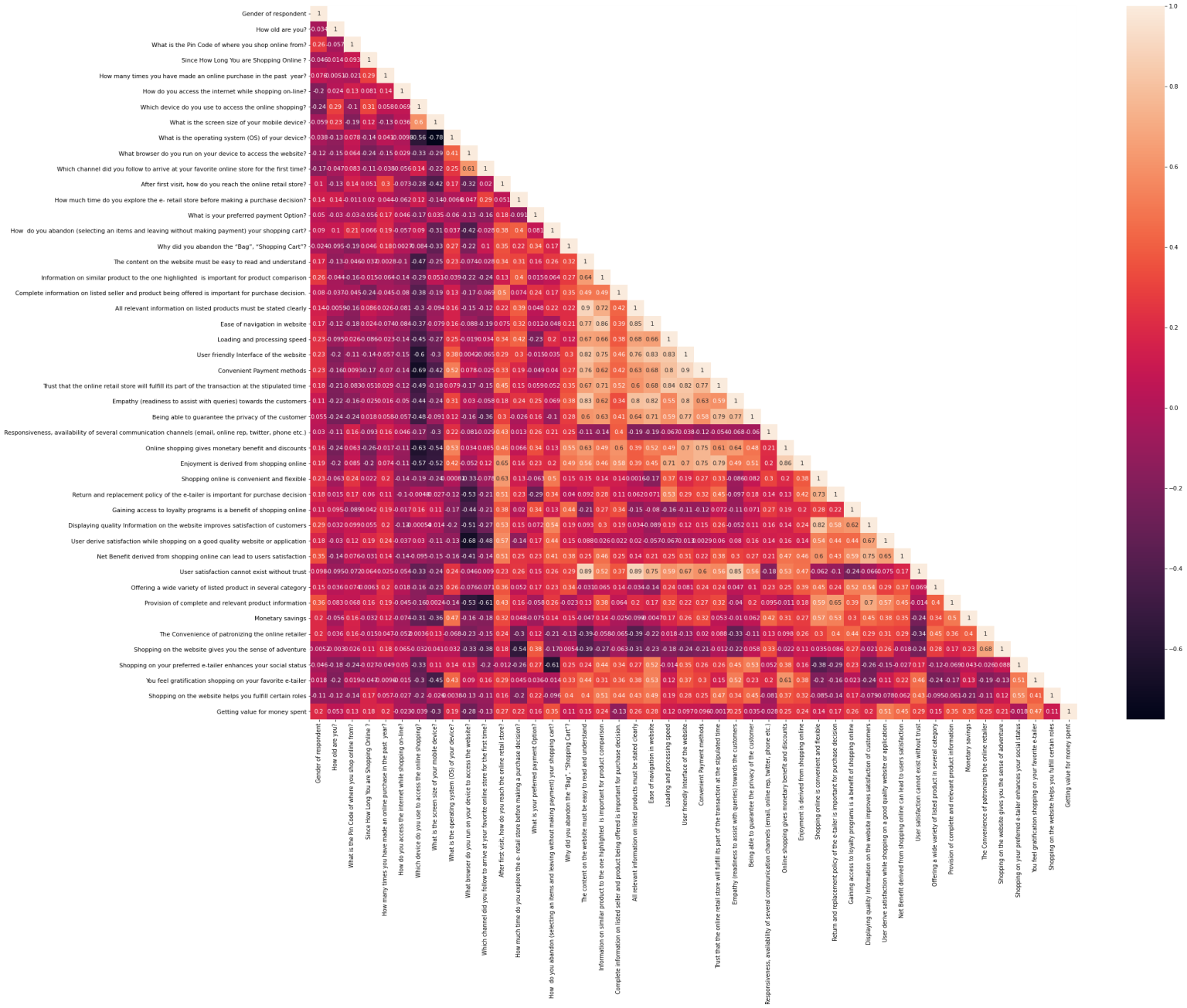
LIMITED MODE OF PAYMENTS AVAILABLE:



THE DISTRIBUTION OF ALL THE VARIABLES CORRESPONDING TO ALL THE OTHER VARIABLES:



HEATMAP OF THE DATASET:



There are various features which are highly correlated with other and thus there exist an evidence of multicollinearity in the data. Also, the distribution of the data shows that there exist skewness and outliers in the data.

**CONCLUSION**

The findings from the project can help e-commerce websites to extent their business:

1. The costing of products, reliability and brand of the company along with the return policies play an important role in buying and retention behavior of the customers.
2. The online platform gives the customer the benefit of comparing the prices, thus being competitive in the market allows the firm to stay and make good number of sales otherwise the customers may switch to the other websites.
3. The trust on the website is very important as the stores are not present physically and the customers don’t know the sellers personally, so in order to influence to customer to purchase from their website, it is very important to have a bridge of trust between the two parties.
4. In online shopping, customer purchases the goods on the basis of a photograph. The photograph can misleading or different from the expectation of the customers. If the return policy is very rigid, the customer will hesitate to buy, no matter how much that person likes the product. Thus, it is important to have flexible return policies.
5. The payment method determines the purchase behavior. Some customers prefer cash on delivery while other prefers digital payment. The availability of both the methods caters a wider range of customer.
6. Last, but not the least, the quality of the product and service must be up to the mark in order to make the buyer a permanent customer.

Trend according to the data available:

1. Some websites are preferred more than the others. Amazon is the most preferred website, which is followed by the Flipkart. These two companies are most trusted and are thus realiable.
2. Snapdeal and PayTM have more sellers from tier 2 and tier 3 cities while Amazon and Flipkart have more sellers from Tier 1 cities.
3. The top companies have most lenient return policies and they take less time to process a return.