

# **Indian E-Commerce Analysis and Customer Retention**

## **Abstract-**

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. By hypothesising that a combination of both utilitarian value and hedonistic values are needed to affect the repeat purchase intention (loyalty) positively, Structural equation model has been presented on the primary data collected from the Indian online shoppers. Results indicate the e-retail success factors, which are very much critical for customer satisfaction. By increasing the utilitarian value and hedonistic values derived by the customers, customer satisfaction and hence the customers repeat purchase intention can be increased significantly.

## **Understanding The Dataset**

This Dataset contains data of Indian E-commerce website that has the user data such as gender, age, location and 71 other columns of data about the activity of the user in the website and which website they would recommend to their friend. We could analyse and find which factors affect the recommendation of the users. So we have a problem of dimensionality reduction. PCA could be used to isolate the columns that affect change at a higher level when it comes to user's website recommendation to his/her friends.

In this project we perform dimensionality reduction, to find the factors for user's recommendation of websites.

## **Exploratory Data Analysis**

We can analyse the categorical data of E-Commerce website listed below

- 1) Location
- 2) Gender of user
- 3) Usage Time
- 4) Recommendation of user
- 5) Preferred Payment Option
- 6) Discovery of website/link source

- **Location-**

Location of user can be plotted by getting the latitude and longitude of the City using Nominatim and it can be plotted using Geo Pandas.

We only need the city names to perform a left join with the city data in the original dataset.

- **Gender of user-**

we can use various categorical variables such as gender ,age on these to see how it varies with the region.We encode the categorical variables as Geo-pandas takes only numbers as input.

- **Usage Time-**

we can see that users who are using shopping websites for more than 4 years are the majority and we can plot this data to categorize them by location.

- **Recommendation of user-**

We can analyse which Online retailer the user favors and recommends to his friend.we plot the data to examine the location influence.

While North has a mixed recommendation of all websites south remains to have flipkart,amazon to be the favored website by the Southern Users.

- **Preferred Payment Option-**

we analyse the preferred payment option of users and plot them to understand significance of location on payment methods.

Majority users prefer E-wallet payment methods while some users prefer the other methods.

- **Discovery of website/link source-**

we can analyse the source of visit of users to the shopping website and plot the same on the map.

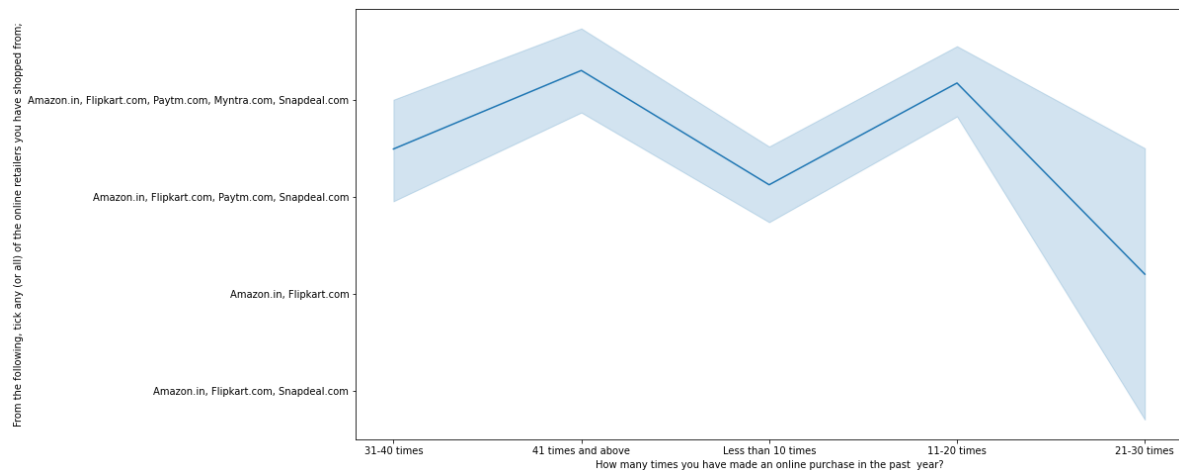
Search Engine appears to be the major source for website visits followed by content marketing.

There is double the number of women than men who have taken this survey. -Most of the people are in their 30's followed by 20's, teenagers and senior citizen are the least in number. -Most of the people belong from delhi, noida and banglore, ambiguity can also be seen as noida has two categories (noida and greater noida) which need to be handled -Most of the

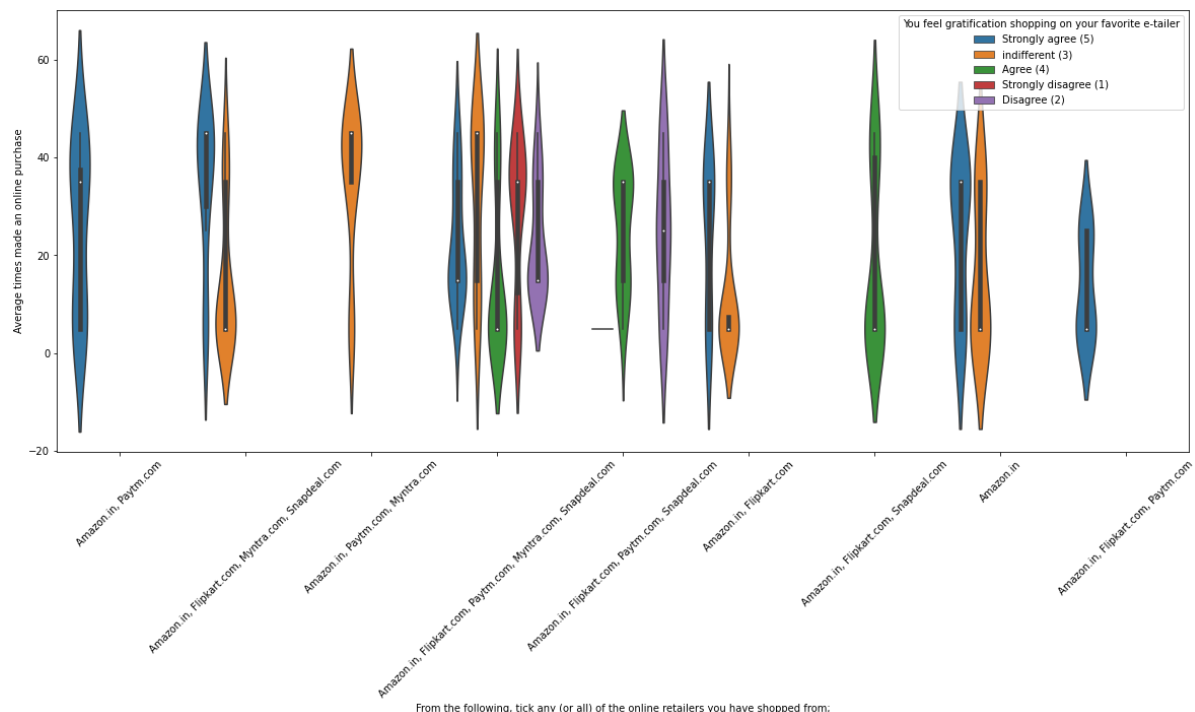
people shopping online have been shopping from a long time. -Majority of people shop online 10 times a year, ambiguity can also be seen for range 42 times and above which needs to be handled

## Analysis on the basis of Various following factors-

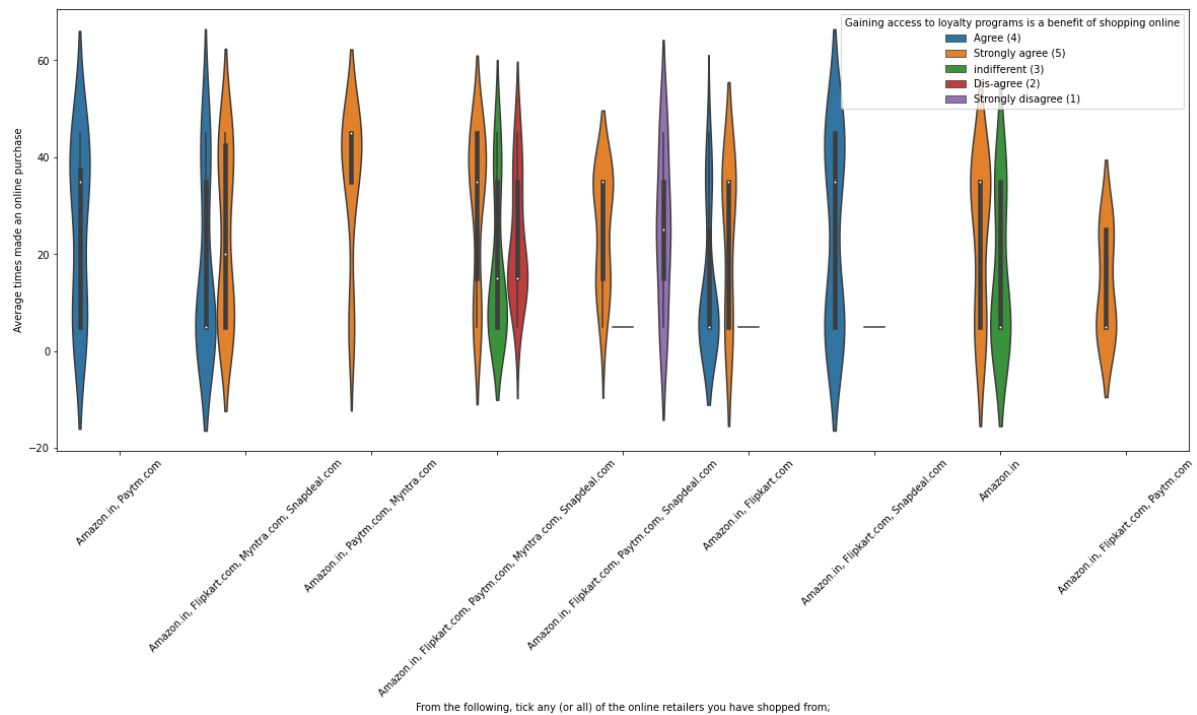
- Intention of Repeat purchase-



- Heavy shoppers who shop more than 41 times a year shop from all the online brands, some of the people who shop for 32-40 and less than 10 times a year seem to exclude myntra. People shop from Amazon and flipkart whatever be the case.

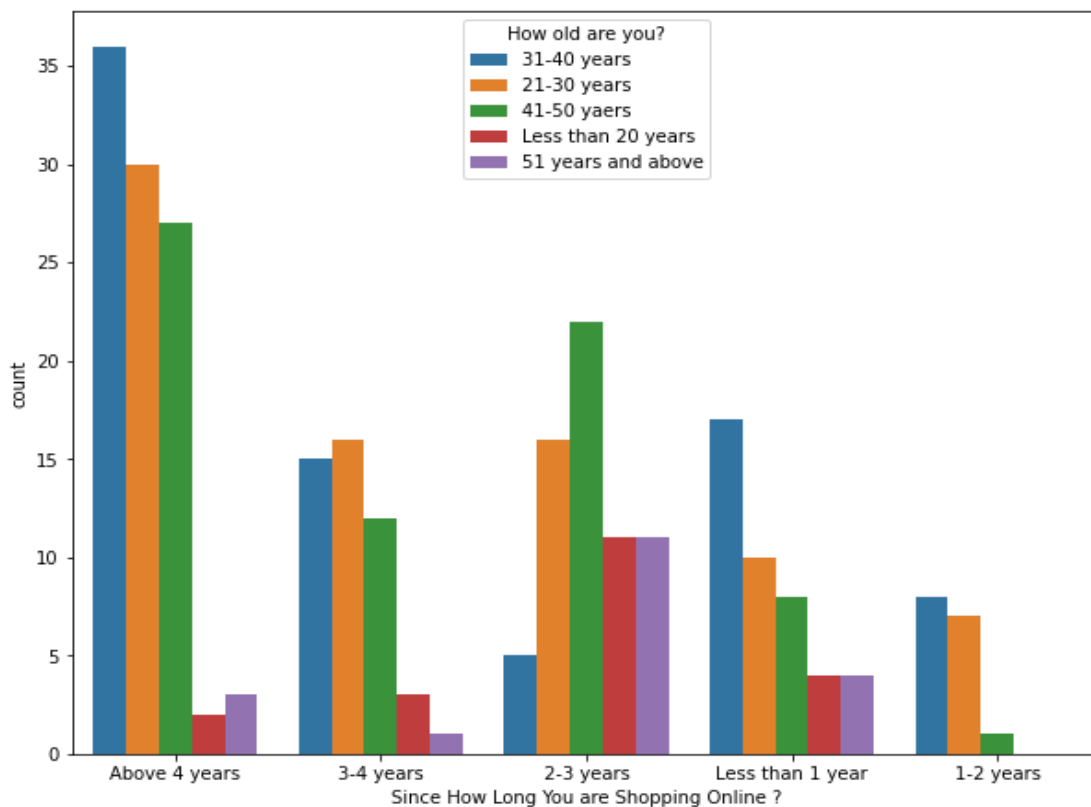


- Almost all the people who have shopped from amazon, flipkart and paytm are satisfied. People who shop from a more number of online brands dosen't seem to be satisfied.

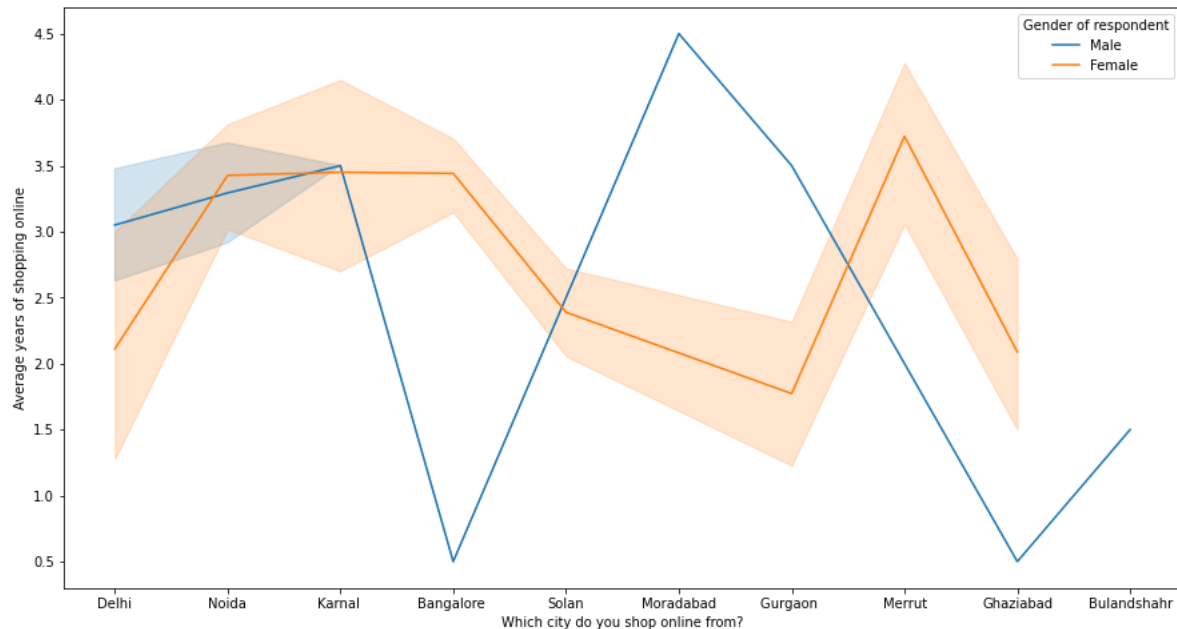


- People shopping from amazon and paytm are getting benefits from the loyalty points, flipkart and sanpdeal also seem to give such benefits but people who shop from almost everywhere disagree with this statement too

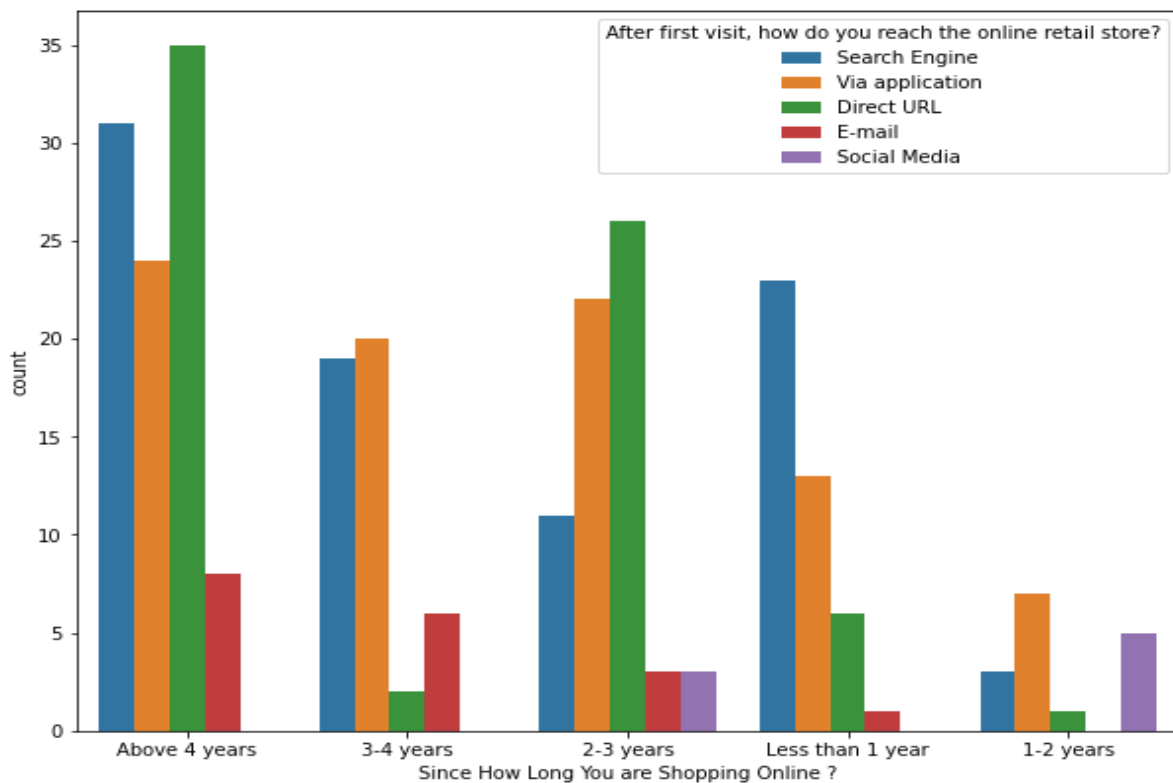
- **Online Retailing-**



- Highest number of people have been shopping online for above 4 years except for the age group below 20 years and above 50 years. People who are shopping online for 1-2 years does not include teenagers and elder people.

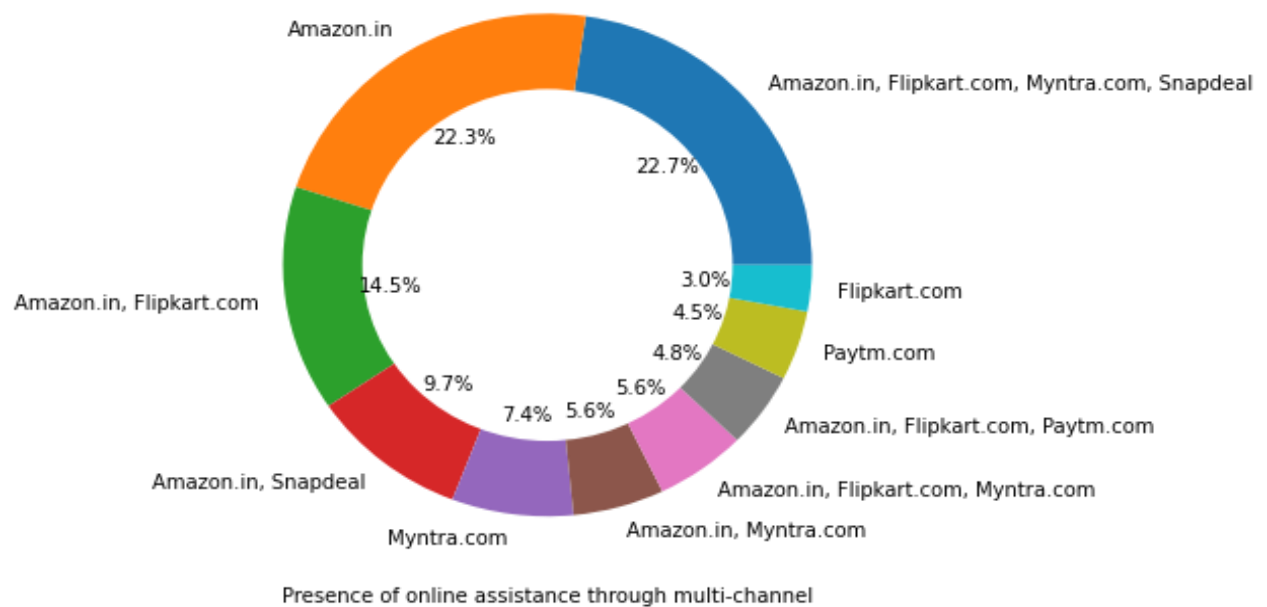


- In lines, we can see that density of female customers is more than male. Men living in banglore and ghaziabad shop have shopped online for less than 1 year. Highest number of men shopping online belong from delhi and noida, while men from moradabad have been shopping online for the longest. Women from meerut and noida have shopped the longest.

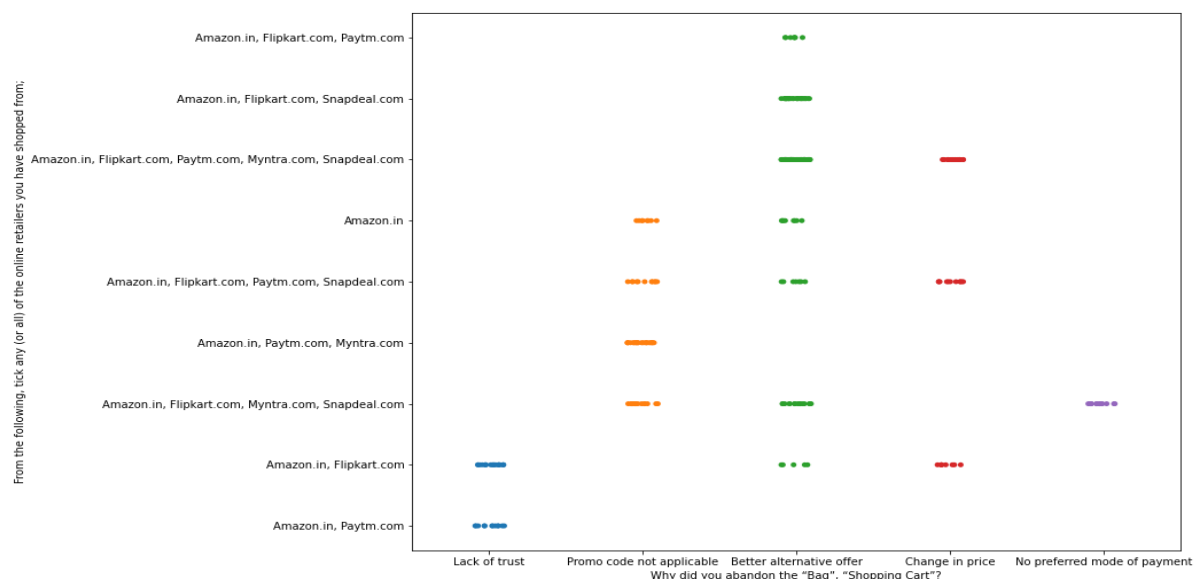


- Even though people who are shopping online for more than 3 years donot use the application rather use search engine and direct url's in large number which indicates that online brands should update all their platforms rather than just application.

- **Brand image-**



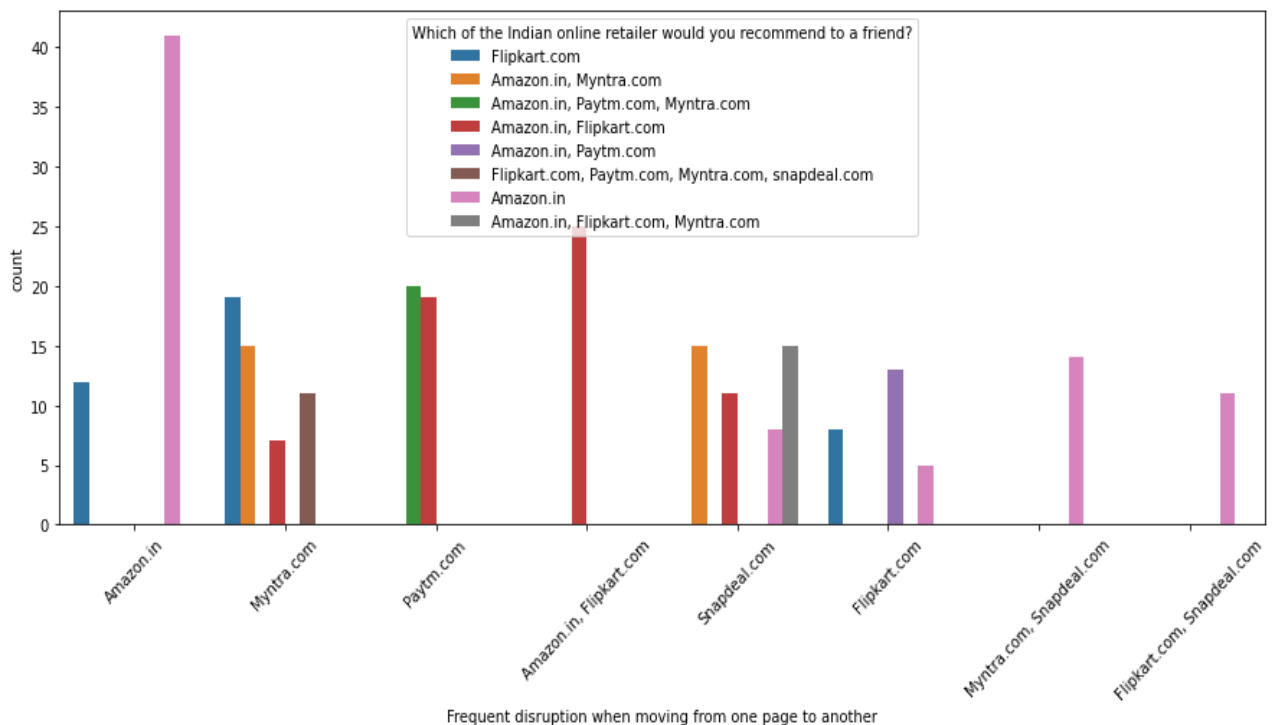
- Amazon, Flipkart have been had the highest votes for having all the positive points and have maintained a very good brand image followed by paytm and the myntra.



- We can clearly see that most of the time people abandon the bag is beacuse they get a better alternative offer or promo code not applicable. There is also lack of trust seen in amazon, flipkart and paytm by some people.

- **Loyalty-**

- Loyal customers are those who keep using the same brand even if it is not good as other brands.



- Customers seem to be more loyal to amazon, flipkart and paytm as even though many of them have given negative remarks about them still they would recommend these platforms to their friend

## **Principal Component Analysis**

To analyse which factors affect the recommendation of user we use PCA. we start by dropping the target column from the training data.

we use ordinal encoder to encode the dataset to numerical variables as PCA takes only numeric variable.

## **Conclusion**

The cost of the product, the reliability of the E-commerce company and the return policies all play an equally important role in deciding the buying behaviour of online customers. The cost is an important factor as it was the basic criteria used by online retailers to attract customers. The reliability of the E-commerce company is also important, as it is even required in offline retail. It is important because customers are paying online, so they need to be sure of security of the online transaction. The return policies are important because in online retail customer does not get to feel the product. Thus, he wants to be sure that it will be possible to return the product if he does not like it in real. Whereas, the logistics factor, which included Cash on

delivery option, One day delivery and the quality of packaging plays a secondary role in this process though these are Must-be-quality. This is so because these all does not interfere with the real product and people believe that this is the basic value that E-commerce websites provide.

All the websites were not equally preferred by online customers. Amazon was the most preferred followed by Flipkart. This can be explained easily by previous result that we got. These two companies are most trusted in the industry and hence, have a huge reliability. Also, the sellers listed on these websites are generally from Tier 1 cities as compared to Snapdeal and PayTM which have more sellers from tier 2 and 3 cities. Also, these websites have the most lenient return policies as compared to others and also the time required to process a return is low for these.