Case Study on

E-Retail Factors for Customer Activation

And Retention

By

VAISHNAVI BALAJI

Internship-31

SME- Gulshana Choudhary

Flip Robo Technologies

#39, NGEF Lane, 2nd Floor, Suite No.1759, Indiranagar, Bangalore, Karnataka, India 560 038

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Introduction

In the era of Meta Verse and flying cars, online shopping was the initial evolution since the boom of internet. As globalization keeps increasing and necessity of employability increases, similarly people find it convenient to purchase online rather that going to the shop and facing the difficulty of transport, traffic and pollution. Online shopping seems to be very beneficial when it comes to senior citizens and single care takers at home who require groceries or any other urgent commodities.

Just as much the traditional shopping focusses on customer satisfaction, it is as much in fact more important when the shopping is done online. Owing to the insecurity the customer has regarding placing an order through an app or a website, it is very critical the process from placing an order to the delivery is seamless and ensures maximum customer satisfaction and encourages the customer for repurchase. Hence, on analyzing the difference answers given by Indian customers who shop online, we can get a brief understanding as to which feature encourages them to purchase and which doesn't thereby increase the feedback to online shopping websites to work towards customer satisfaction.

1.2 Problem Statement

The dataset is a response of e-commerce customers to various questions. The purpose of this dataset is to recognize the features that tend to encourage the ability of a customer's repeat purchase, which gives a good amount of helpful information to the online brands to improve their profits and design their online impact tactfully.

1.3 Aim

Analyze the e-retail success factors which play a major role in customer satisfaction.

1.4 Objectives

(i)   Perform Exploratory Analysis on the data set.

(ii)   Segregate the dataset based on customer's personal information.

(iii)   Analyze the relationship between the features based on the factors of Intention of Repeat Purchase, Online Retailing, Brand Image and Loyalty.

(iv)   Encoding the categorical features and training a model.

(v)   Testing the model and evaluating using the metrics.

CHAPTER 2

EXPLORATORY DATA ANALYSIS

Summary of the given dataset is as follows-

- The data is collected from Indian Online shoppers.

- Dataset consists of 269 rows and 71 columns.

- Except the column "What is the Pin Code of where you shop online from?", all the other columns are of object data type, which gives us 1 continuous type column and 70 categorical type columns.

- There are no null values in the data set.

- The number of women shoppers are more than double of men shoppers.

Gender of respondent
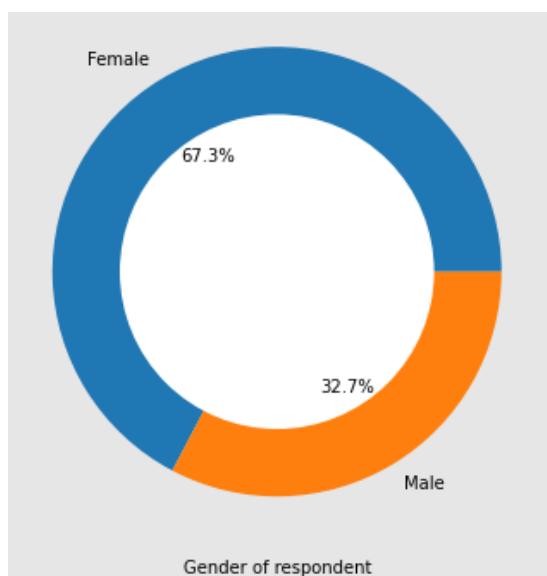
- Majority age group of shoppers is between 20-50 years.

How old are you?
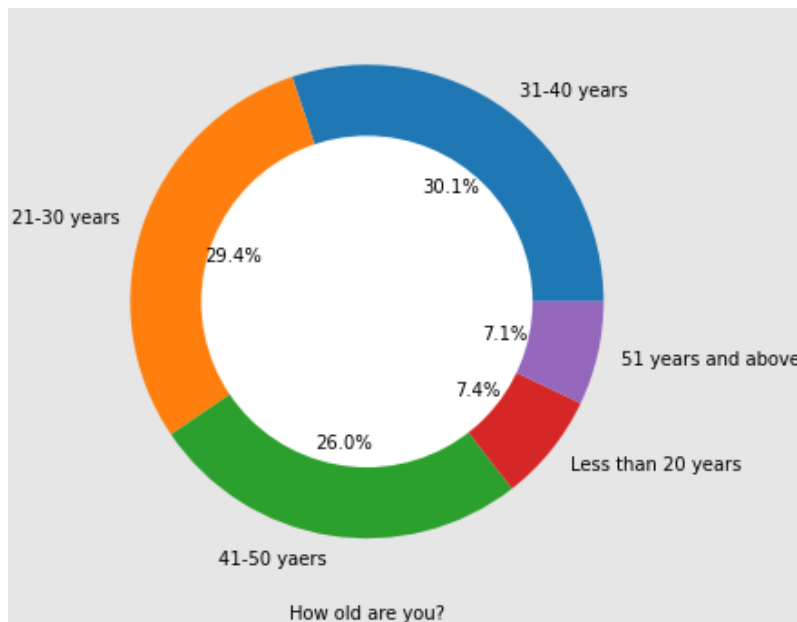
- The shoppers majorly reside from Delhi, Noida followed by Bangalore, least number of shoppers are from Meerut and Moradabad.



Which city do you shop online from?
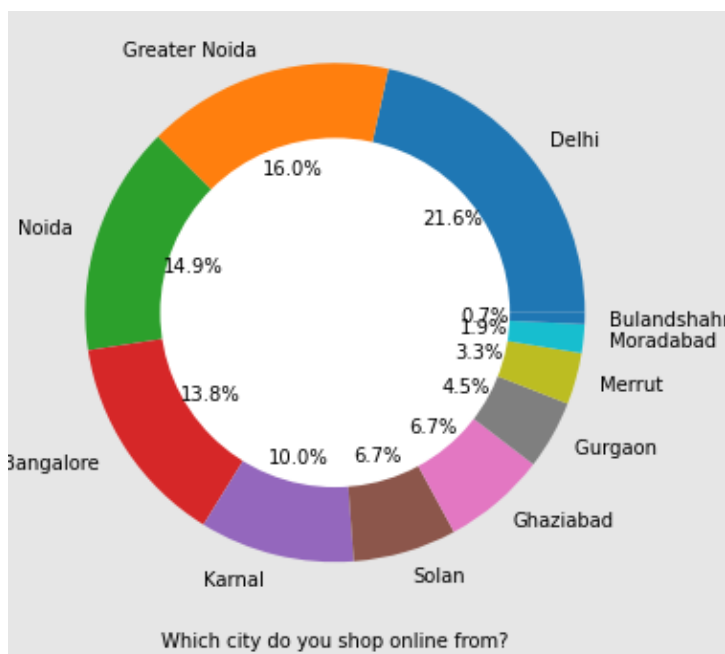
- Most of the shoppers have been shopping for more than 4 years.

- Majority of the customers have made almost 10 online purchases in the past year.

- There are large number of customers who access the shopping app/website using mobile data.

- Most of the customer's device from which they access the shopping app or website has a screen size of almost 6 inches. The range of values of screen size is from 4.7 inches to 5.5 inches.

- High number of shopper's device has an Operating System of Windows which indicates that customers shop from their computer systems or laptops.

- The browser which shoppers commonly use is Google Chrome.

- Shoppers usually access the brand website through the search engine for example- Google, Bing.

- Even after being familiar with the shopping website, customer prefers to visit again through search engine as well as using the application.

- Shoppers tend to explore the shopping website/app for more than 15 minutes before deciding to purchase.

Graphical Analysis of Categorical Columns Distributions
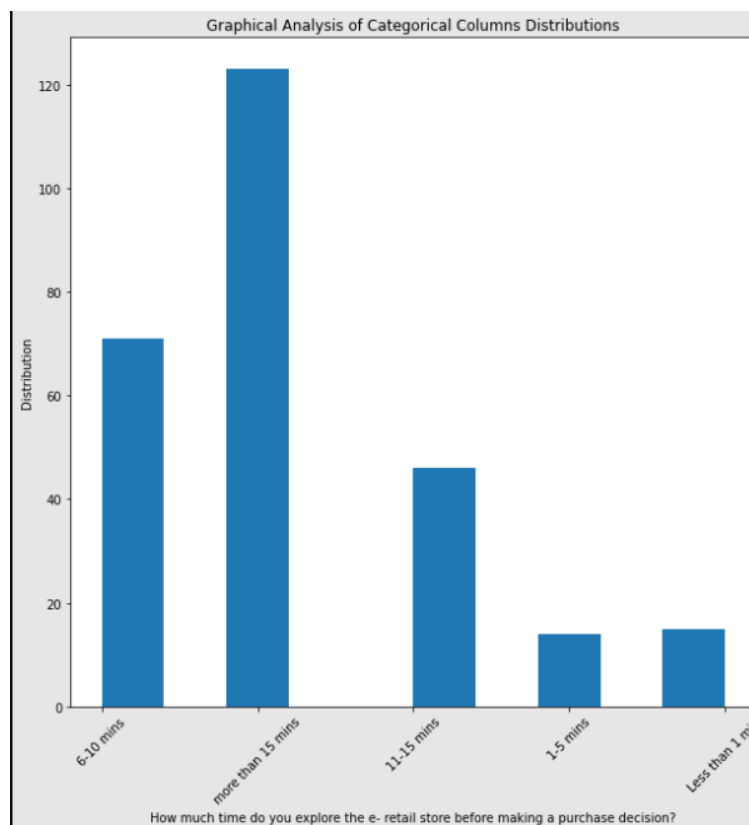
- Huge number of customers prefer the mode of payment as Credit/Debit Card followed by Cash on Delivery. Very few customers prefer e-Wallets assuming the fact that many shoppers would not have activated their e-wallets with their bank accounts.
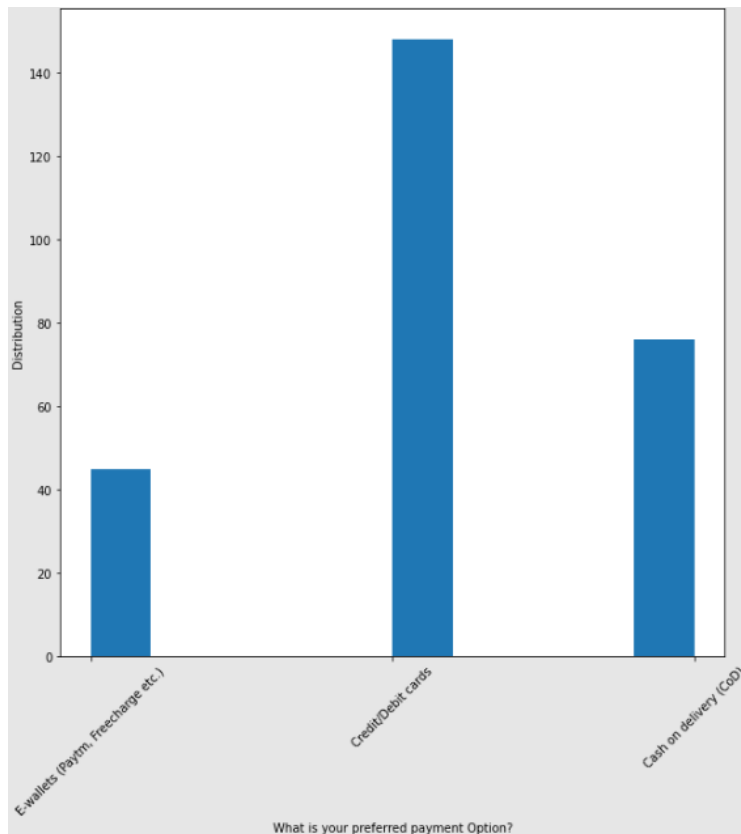
- Many shoppers abandon their cart rarely, very few customers abandon their cart frequently.
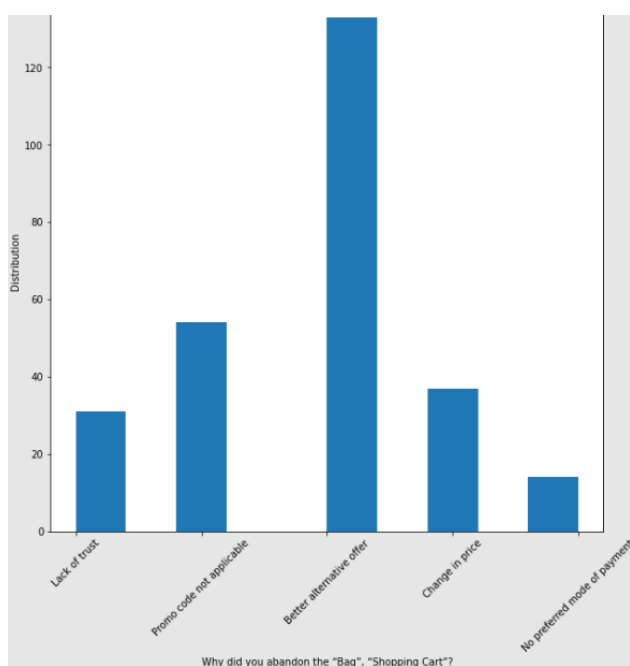
- The customers who abandon the cart are mainly due to finding better offer alternatively or promo code not being applied.

- Customers strongly believe that the content should be easy to read and understand, there are some customers who do not agree as well.

- It is beneficial for shoppers when similar product's information is also easily viewable for ease in comparison, there are a minority group of people who do not believe the same.

- The number of shoppers who strongly agree that there should be clear information of product and seller are less than shoppers who just seem to moderately agree to it.

- Shoppers require the app or website from which they shop to have easy navigation and user-friendly interface.

- It is very important for the customers to have convenient payment methods and complete the transaction in the given turn-around-time to have a positive impact on the customer to initiate the tendency of repurchasing, and from the analysis the customers strongly agree to the same.
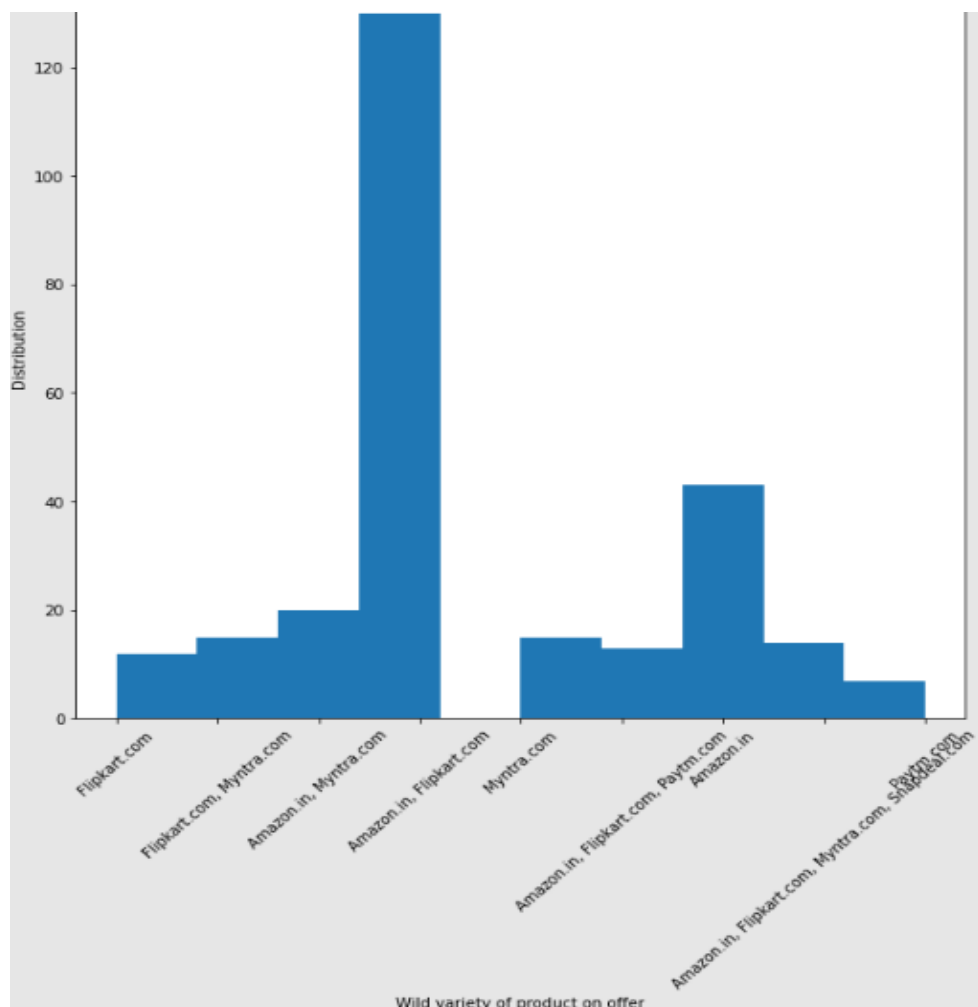

Quickness to complete purchase
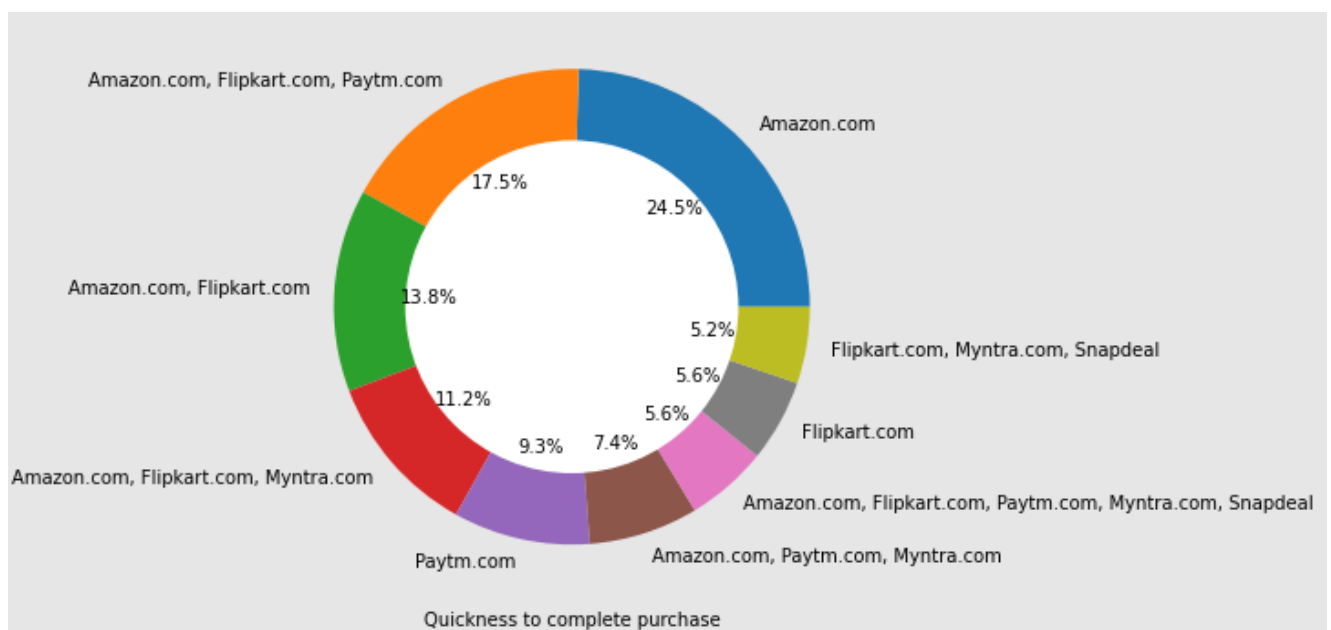
- Just like how in stores and malls, there are representatives who help out the customer's concerns, it is critical for online brands to have a strong support for their customer's issues and shoppers strongly vouch for the same.

- Shoppers strongly believe that their information and details should be private and not misused in any sense.

- From the dataset we can conclude that when an online brand gives the customer a token of monetary benefit/cashback/discount, it gives a sense of encouragement to shop more.

- There are many customers who find a sense of gratification after shopping online but there are many customers too who do not believe the same owing to the fact that they shop online out of necessity and nothing more.

- Many customers shop online due to the reason that it is flexible and convenient.

- Shoppers strongly agree that the return and replacement policy is very important when purchasing from an online brand, owing to the decrease in proximity between the customer and the product, the customer has not physically seen the product while placing the order hence he would like to know whether he can return the product if he does not like it after receiving it.

- When an online brand provides access to loyalty programs for the customer, it encourages the customer to be in touch with the brand which increases the chance of another purchase, but there are few customers too who are not quite interested in loyalty programs.

- Many customers strongly believe that the net benefit they get by online shopping gives them satisfaction which is a huge benefit for the brand and satisfaction does not exist if shopper does not have trust.

- In the given data, customers strongly feel that it is beneficial when they get a wide variety of products in categories which helps them to make a well-informed purchase.

- After media being so powerful, the shopper is well aware of a product price at different online brands as well as a store or mall. Customer finds it very important to know in which option he is getting a high monetary saving.

- Many customers also have a sense to support the online brands, but there are customers too who are not concerned with the same.

- There are customers who are not concerned with the fact that the shopping on their preferred online brand might improve their social status.

- Many customers shop online assuming out of necessity and are not concerned with gratification. Few shoppers find a sense of gratification when they shop from their favorite brand

.
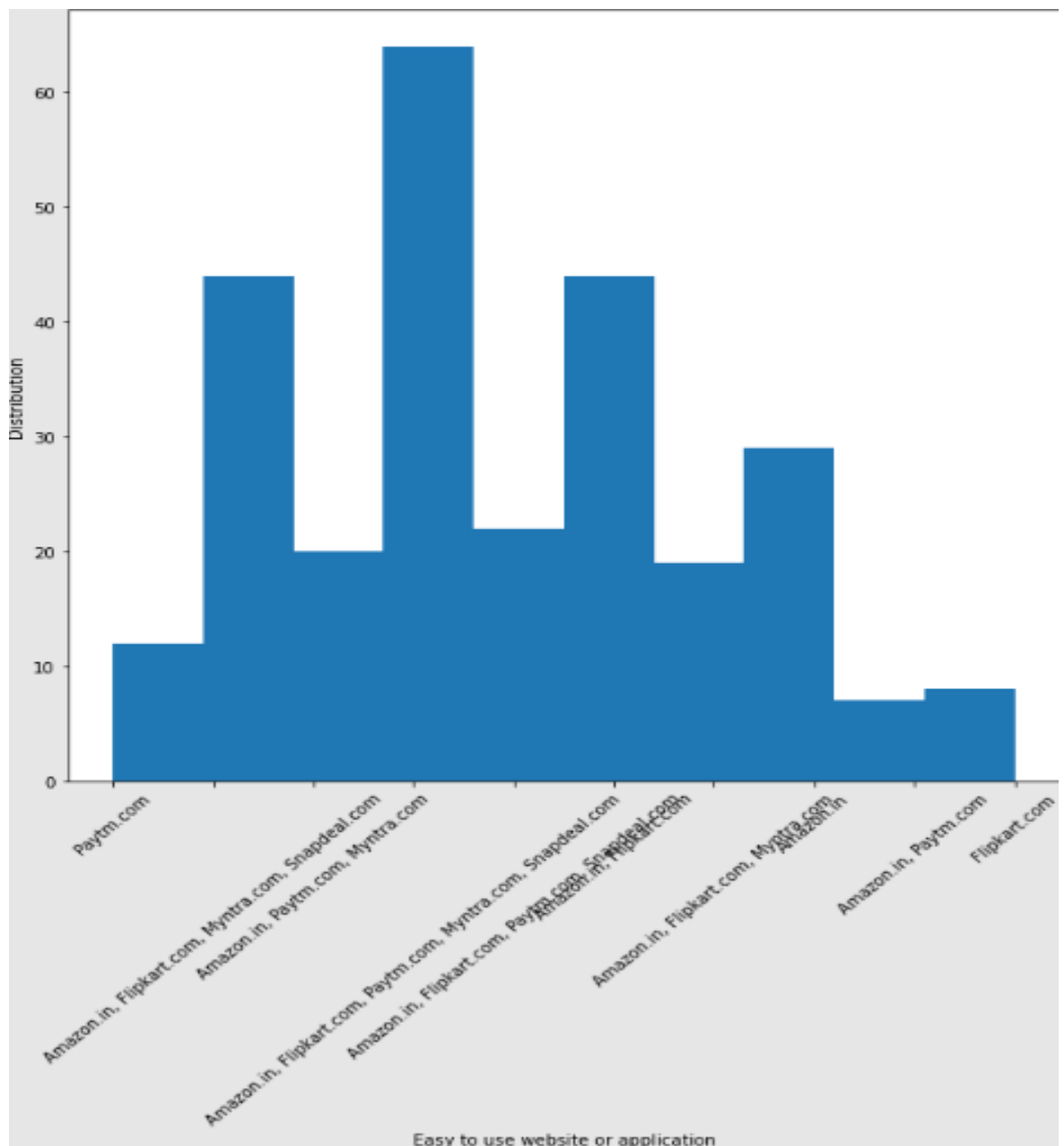
- Achieving a value for the money spent by the customer is important to increase customer satisfaction.

- Most of the customers have shopped from online brands like Amazon, Paytm and Myntra.

- Amazon and Flipkart seem to be very easy to use.

- Amazon and Flipkart seem to have a visual appealing page layout.

- Amazon and Flipkart have a huge variety of products to offer.

- Amazon and Flipkart have a complete relevant description of products.

- Amazon, Flipkart and Paytm have a fast-loading website.

- Quickness of purchase is high in Amazon, Flipkart, Myntra and Paytm.

- With respect to payment options, Amazon, Flipkart and Myntra seem to have several payment options

- Amazon and Flipkart have high speed in delivery.

- Customers are loyal to Amazon, Flipkart and Paytm even though they have negative remarks about them, they would still recommend these brands to their friends.


Which of the Indian online retailer would you recommend to a friend?
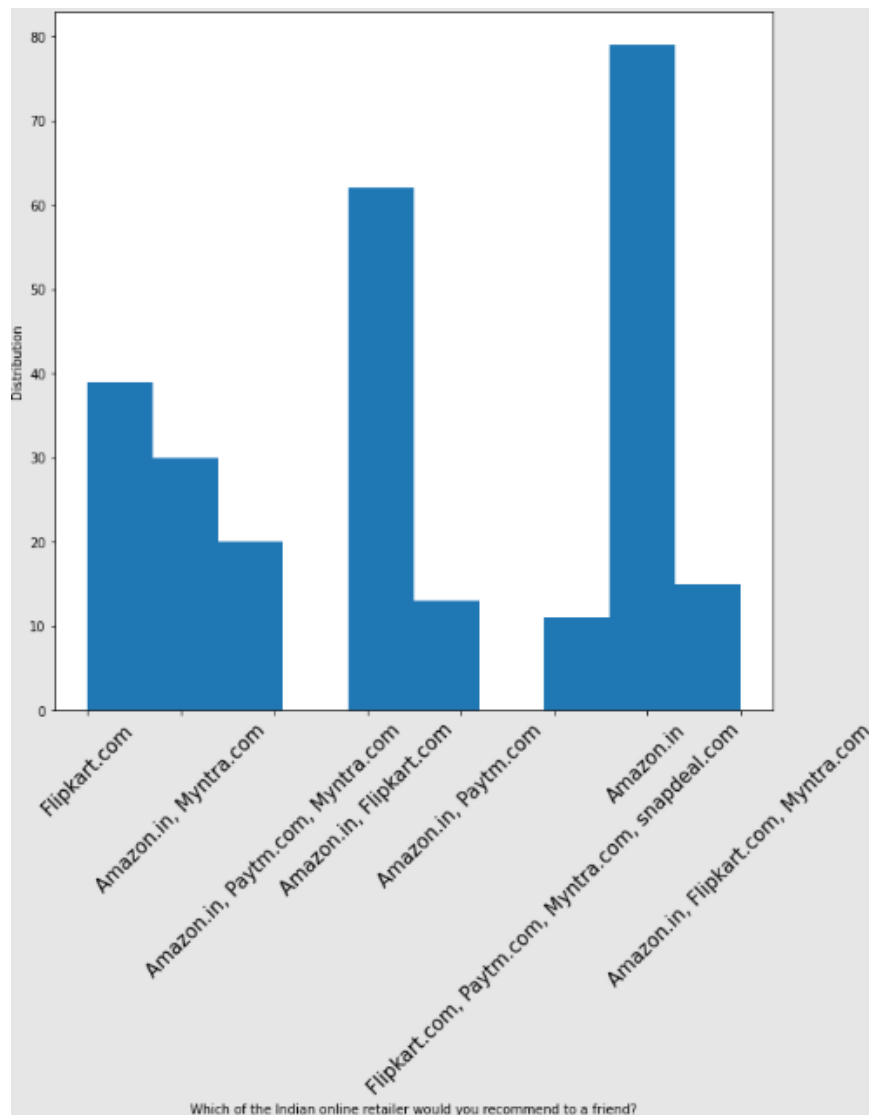
CHAPTER 3

TRAINING AND TESTING MODEL

- From the complete dataset, the features and target are segregated and from the features the categorical columns are encoded and scaled using Min Max Scaler.

- Since there are 70 features, it is important to understand the important features in model development.

- For feature selection Chi 2 test is implemented and below is the result for top 10 highest score features:

Features	Score
Why did you abandon the "Bag", "Shopping Cart"?	75.754028
Loading and processing speed	59.810983
Shopping on the website gives you the sense of adventure	59.253569
What browser do you run on your device to access the website	57.171099
Change in website/Application design	55.301526
Visual appealing web-page layout	54.245760
'Limited mode of payment on most products	53.269266
Longer time to get logged in	48.222655
Longer time in displaying graphics and photos	48.130643
Wild variety of product on offer	47.605973

- Further ahead, Principal Component Analysis is used to understand the weight of variance of different features. The features after scaling using Min Max Scaler is passed through the PCA object.

- The attribute of PCA which is explained_variance_ratio_ is used to plot and understand the number of features needed to retain the information.

Number of Features needed to explain variance
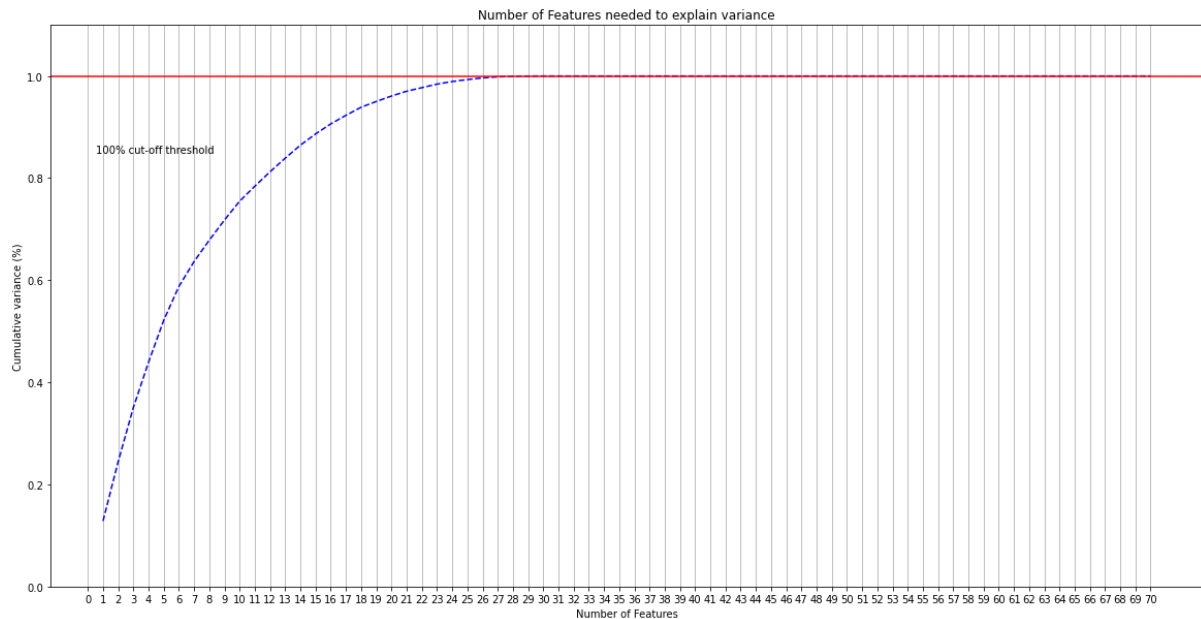
- Using 29 features as the parameter value for n_components, PCA transformation is applied on the features.

- The features and target are further split into training and test data and for training the model, Random Forest Classifier Algorithm is used.

- An accuracy score of 1.0 is achieved for both training data and test data which denotes there is no overfitting or underfitting and the model has used the training information well to predict the target.

- On applying cross-validation also the score obtained is 1.0

- Classification Report for the model is:

	precision	recall	f1-score	support
Amazon.in	1.00	1.00	1.00	21
Amazon.in, Flipkart.com	1.00	1.00	1.00	26
Amazon.in, Flipkart.com, Myntra.com	1.00	1.00	1.00	7
Amazon.in, Myntra.com	1.00	1.00	1.00	10
Amazon.in, Paytm.com	1.00	1.00	1.00	2
Amazon.in, Paytm.com, Myntra.com	1.00	1.00	1.00	6
Flipkart.com	1.00	1.00	1.00	12
Flipkart.com, Paytm.com, Myntra.com, snapdeal.com	1.00	1.00	1.00	5
accuracy			1.00	89
macro avg	1.00	1.00	1.00	89
weighted avg	1.00	1.00	1.00	89

CHAPTER 4

RESULTS AND CONCLUSION

The results of the analysis suggest the following outcome which can assist online brands to extend their business.

- The price of the product, reliability of the online brand and the return policies all play an important role in predicting the buying intention of the customer. The price is an important factor as it is a basic criteria used by online brands to attract customers. Reliability is just as much important as it is in offline retail if not more since the customer are paying online without physically seeing the product in real-time. Since, the customer is not seeing the product, return policy is also extremely important so that if he does not like the product after receiving it, he can return it.

- With respect to the logistics factor, Cash on Delivery, delivery speed plays a secondary role in this process though they are important.

- All the brands are not preferred by shoppers. Most preferred brand is Amazon followed by Flipkart. Amazon and Flipkart are most trusted and reliable by the customers.

- It is important for brands to keep their websites updated since there are many shoppers who shop from website too after the app.

- When brands give loyalty programs access to customers, it incentivizes the customer's shopping experience.

- Though shoppers have negative reviews for brands like Amazon and Flipkart, they would still recommend these brands to their friends and family owing to the reliability.

- From the given dataset, models are developed using many algorithms- Logistic Regression, Decision Tree Classifier, Random Forest Classifier, K-Neighbors Classifier, Support Vector Classifier, Gaussian NB. After comparing the results, the models which are efficient are-Logistic Regression model, Random Forest Classifier model, K-Neighbors Classifier model, Support Vector Classifier model.

- In the above models the parameter which is the target is "Which of the Indian online retailer would you recommend to a friend?", which is considered as a major factor in analyzing Customer Satisfaction.

- The cross-validation score of 0.99 is achieved.