Report on Exploratory Data Analysis (EDA) for E-retail factors for customer activation and retention study from Indian e-commerce customers.

By: Satu Vinayak Pole

Intern – Flip Robo Technologies

Introduction:

There have been many models/theories proposed out 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- i.e. products quality and how customers queries are handled;
- System quality- how good the web page is and how it handles on different kind of devices while serving customers;
- Information quality- is quality of products information, their sellers and delivery information;
- Trust- is a major factor in shopping online as customers don't know the sellers yet they purchase plus customers provide their valuable information to e-com websites and in turn trusts e-com websites to keep them for themselves.
- Net benefit- are usually monitory benefits received to customers after buying online.This also includes EMI options while purchasing and discounts and very large inventory of products to select from.

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

Problem understanding :

Various online shoppers were given a questionnaire where they filled their personal details along with some suggestions and answers related to online shopping. Then they were provided with 5 e-retail store names from which they had to choose/tick what they feel based on the question provided to them.

They were further asked which is their favourite e-store and which store would they recommend to their friends and families. This data was further compiled and provided to us to do data analysis on it and based on that provide a solution as which e-store is better and its reasons.

Exploratory Data Analysis for the data provided:

The dataset is a result of form provided to various e-com customers who shops atleast <10 times a year. They were allowed to tick mark on the questionnaire based on their experience with different kinds of e-com websites. The data is limited to only 5 websites namely Amazon, Flipkart, Snapdeal, Myntra and PayTm. The customers were asked about their experience with different shopping websites and what changes/needs they have for these websites to make their shopping experience better.

The provided dataset is categorical in nature with some columns having 2 categories to some having more than 5 categories. The dataset was evaluated using Jupyter notebook using panda, numpy and seaborn libraries. I have used only those columns that I assumed to be related to customer retention and new customer activation.

The shape of the dataset is (269,71) which means it had 269 rows and 71 columns. This 71 columns consisted of 70 independent variables (features) and 1 dependant variable (label). This also means, total of 269 people answered this questionnaire of 71 questions.

Then the dataset was checked for null values but after checking it was confirmed that the dataset didn't have any null values. This means all persons filled the questionnaire fully without keeping any empty spaces

Then I tried to describe the dataset. Describing the dataset means we get to know the mean of data, count of observations in data, standard deviation from mean, minimum and maximum values in that all columns. Describe function does not work for categorical columns. Since this dataset is categorical, hence no description was attained. The pincode data was misinterpreted as being it a categorical column it was interpreted as continuous data.

Further, The dataset was checked if there were any duplicates. It was found out that there were 166 duplicate entries in dataset.If we train the model with duplicate dataset entries then model created with them has less accuracy hence I removed the duplicates. Now, the shape of dataset after removing duplicates reduced to 103 rows and 71 columns.

The total of 66 females and 37 males responded to survey from different locations from India .Most shoppers are within age limit of 31-50 years. Older people(age>51) and younger people(age<20) are either less prone to buying online or they don't order much online. Most of the shoppers are shopping online for more than 4 years while very less people are new shoppers I.e. since past 1-2 years.

While most of the people used their mobile data to access websites but many of them arrived to their now-favourite e-stores via search engines and not via content marketing or advertisements. Afterwards most people visited e-store via mobile application and search engines. People mostly used either bank cards or cash on delivery option for purchasing items online.

Many people agrees to the fact that content on the website must be easy to understand to buy any item.This also includes complete seller information, product information being listed properly, convenient payment methods, query resolving capability using multiple channels such as sms, email, app,etc., also the return and replacement policy from seller as per product. Customers also strongly agrees that the e-com website should guarantee the privacy of customer data and websites should offer wide variety of products in several categories. Customers also agrees that shopping online gives them monitory benefits such as bank card EMI's, offers using coupon code, etc. in turn getting value for money.

Most people said that Amazon and Flipkart has wide variety of products on offer then Myntra, Snapdeal and lastly Paytm. Myntra being a fashion and lifestyle website does sell only fashion and accessories products.

As per customers rating Amazon and flipkart have wide variety of products available with complete and relevant information about their products that includes speedy delivery. They have a reliable website and/or mobile application and has are trusted by customers for their privacy safekeeping. Also they have various kinds of online channels that helps customer with their pre-shopping to post-shopping queries.

Conclusion:

With keeping all above mentioned points in mind , I would recommend Amazon as best shopping e-commerce website followed by Flipkart. Myntra is only limited to fashion and lifestyle brands while amazon and flipkart deals in every products that a person needs. I would not recommend Snapdeal and Paytm since as per customers survey they fit at the bottom of the chart that includes less products with slow application/ web page loading and slow delivery.