Data Analysis Applied: E‑Commerce Customer Activation and Retention

Introduction

India has an Internet user base of about 696.77million as of May 2020, about 40% of the population. Despite being the second-largest user base in world, only behind China (650 million, 48% of population), the penetration of e-commerce is low compared to markets like the United States (266 million, 84%), or France (54 M, 81%), but is growing, adding around 6 million new entrants every month. The industry consensus is that growth is at an inflection point.

In India, cash on delivery and credit/debit card payment is the most preferred payment method, accumulating 75% of the e-retail activities. Demand for international consumer products (including long-tail items) is growing faster than in-country supply from authorized distributors and e-commerce offerings. Long tail business strategy allows companies to realize significant profits by selling low volumes of hard-to-find items to many customers, instead of only selling large volumes of a reduced number of popular items. The term was first coined in 2004 by Chris Anderson.

In 2017, the largest e-commerce companies in India were Flipkart, Snapdeal and Amazon. In 2018, Amazon beat Flipkart and was recorded the biggest ecommerce in India in terms of revenue.

Problem definition

Still in such a marketing conditions, it can be tremendously difficult to know exactly how to retain a customer that will actually do well enough to generate revenue. There is so much information and so many ecommerce terms to sift through, so creating a plan of action is necessary to prioritize your resources as you develop your own ecommerce business plan.

This guide offers a streamlined and honest approach to subjects like ecommerce marketing, and provides pointers for sourcing products, along with other tips that are essential for those who are looking to launch a successful online store.

We've covered these things in detail, but this report synthesizes all the things you need for retaining and activating a customer. Without further ado, let's jump in.

Context of the project

   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.

Using the dataset, I have attempted to

1. Import the database
2. Clean the data
3. Analyze the data
4. Standardize the data
5. Sketch the needful insights

Insights of the Data.

   From all the Data available, we can bring out some neat insights or conclusions such as

- ✓ *Customer profiling*
- ✓ *E-Commerce profiling*
- ✓ *E-Commerce Customer Success Importance*
- ✓ *Customer Gratification and Satisfaction Importance*
- ✓ *How Customer Satisfaction promote the Brand?*
- ✓ *How a Customer can be a repeated buyer?*

Tools used

1. Python 3.8
2. NumPy
3. Pandas
4. Matplotlib
5. Seaborn
6. Data science
7. Machine learning

Exploratory Data Analysis.

Exploratory Data Analysis refers to the critical process of performing initial investigations on data so as to discover patterns, to spot anomalies, to test hypothesis and to check assumptions with the help of summary statistics and graphical representations. Post importing the data from a excel file and converting it into a DataFrame I have done the following analysis.
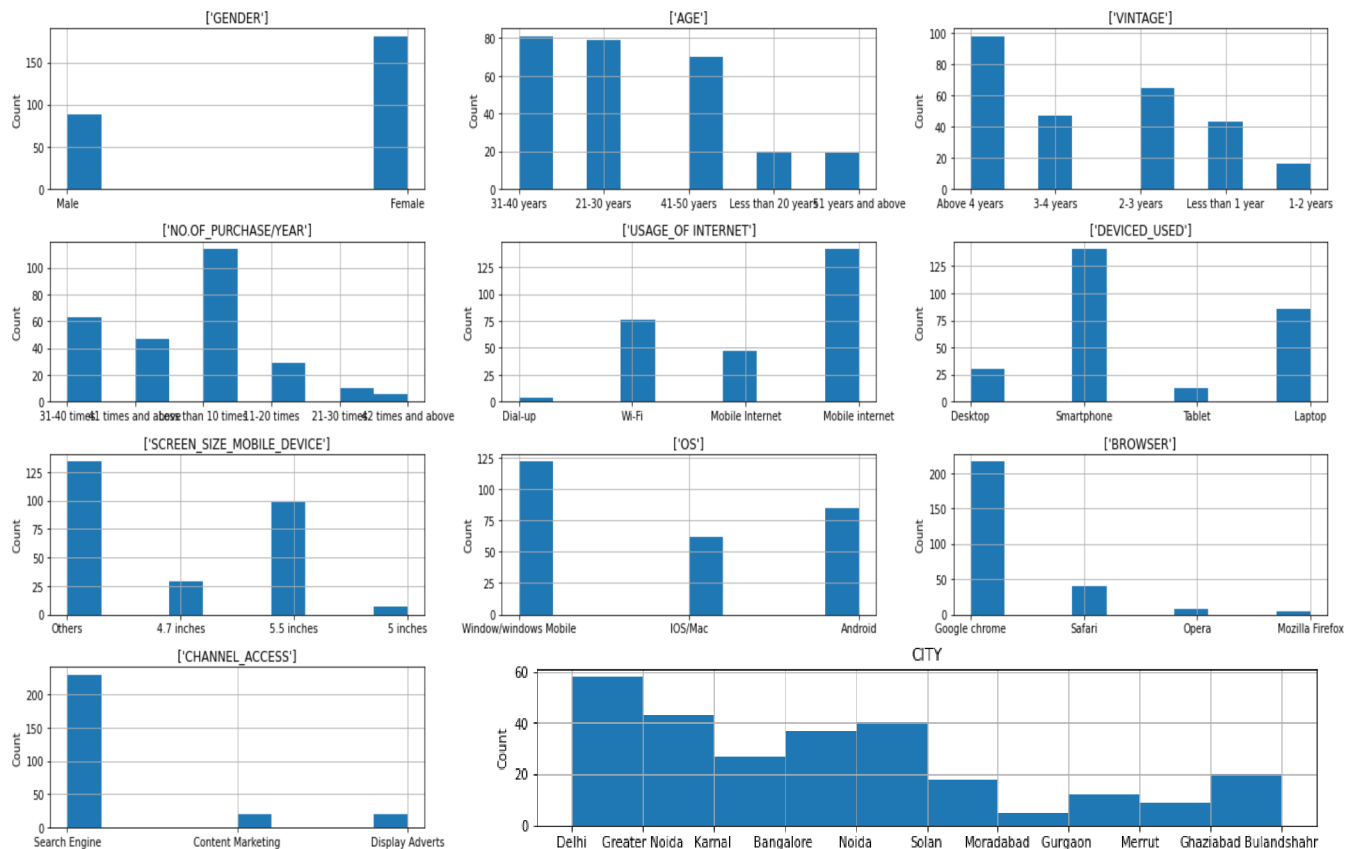
1. Univariate Analysis.
2. Multivariate Analysis.
3. Correlation of the dataset
4. Correlation with the Target variable.

Let's see these Analysis in details as follows.

Univariate Analysis.

I have split the large data into 6 DataFrames to make the analysis easier.

1) Customer and Customer usage details columns.

Key Observations:

a. We can observe that the Female users are higher in number comparatively and highest population of the users are between 21-40 years of age and with the vintage above 4 years.

b. Highest No of purchase done/year is above 10.

c. Customers mostly use mobile internet in smartphones and all or any type of screen size.

d. Operating system is 'Windows' mostly on PC and 'Android' on smart phones and customer uses Chrome and search engines.

So, from above we can say that Female users mostly from Delhi aged between 20-40 years with vintage above 4 years of online shopping shop more with the use of android based smart phones.

2) E-commerce Merchant Website/Application details columns.