E-retail factors for customer activation and retention: A case study from Indian e-commerce customers

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

- **service quality**
- **system quality**
- **information quality**
- **trust and net benefit.**

BUSINESS AND DATAUNDERSTANDNG

- **E-commerce is expanding steadily in the country. Customers have increasing choice of products at the competitive rates. E-commerce is probably creating the biggest revolution in the retail industry, and this trend would continue in the years to come. Retailers should leverage the digital retail channels, which would enable them to spend less money on real estate while reaching out to more customers in tier-2 and tier-3 cities.**
- **Both organised and unorganised retail companies have to work together to ensure better prospects for the overall retail industry, while generating new benefits for their customers. Also the long-term outlook for the industry is positive, supported by rising incomes, favourable demographics, entry of foreign players, and increasing urbanisation. This statistic gives information on the retail m-commerce revenue in India from 2015 to 2020. In 2015, mobile retail e-commerce sales in India amounted to 6.02 billion U.S. dollars and are projected to reach 37.96 billion U.S. dollars in 2020.**

PROBLEM STATEMENT

- 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.

A P P R O A C H

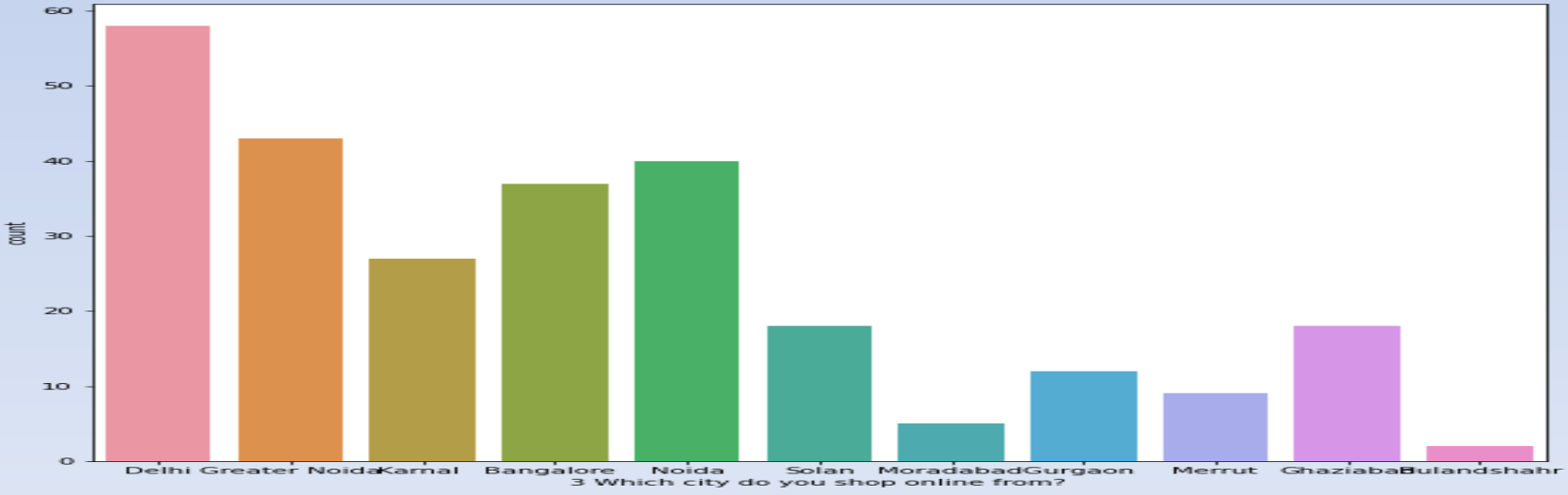
- **We have use Exploratory Data Analysis (EDA) to analyze the patterns present in the data. The analysis will find out variables affecting E-retail factors for customer activation and retention.A case study from Indian e-commerce customers.**
- **We will perform this Exploratory Data Analysis on the sample data provided in order to gain insights, and explain them in simple business terms and represent**

DATA CLEANSING

- Fixing data types of columns in the dataset.
- Checking the datatypes of each coloumn
- Checking null values
- Data info(to disply datadypes of the coloum)
- index #starting no and ending no

Data exploration

- to check shape no of rows, columns
- index #to find starting no and ending n
- statstical summary using discribe
- Plots used to check data coloumn

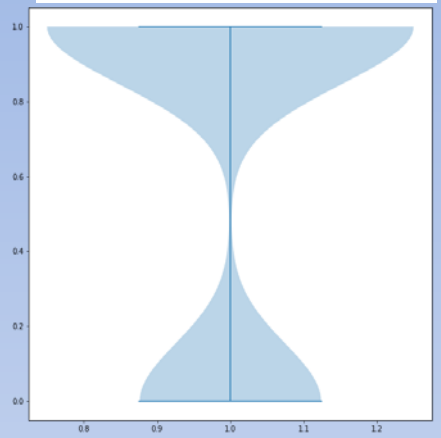


Violin plots with continous data

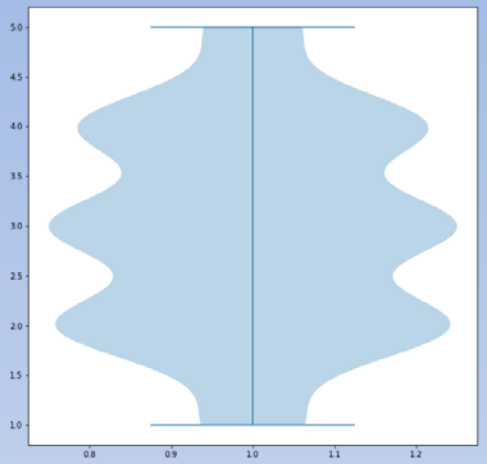
- **Violin plots are used when you want to observe the distribution of numeric data, and are especially useful when you want to make a comparison of distributions between multiple groups. The peaks, valleys, and tails of each group's density curve can be compared to see where groups are similar or different.** are a standard way to graph continuous variables because they show the distribution of the values.

Violin plots with continous data

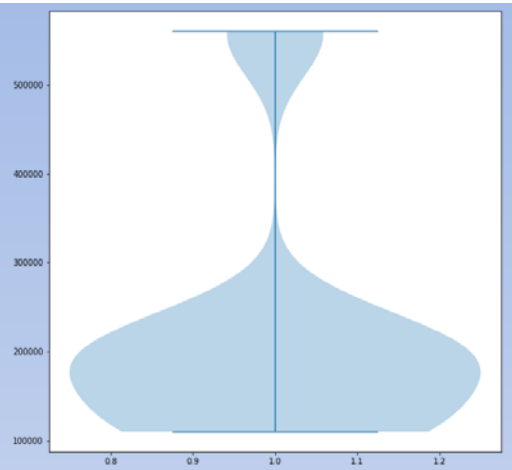
Continous variable 1Gender of respondent

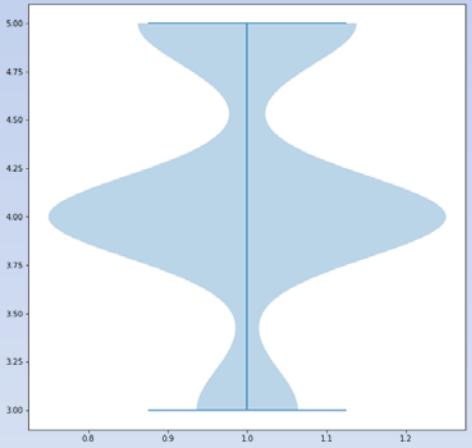


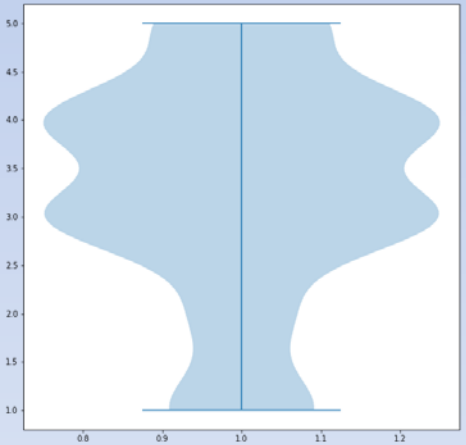
Continous variable 2 How old are you?

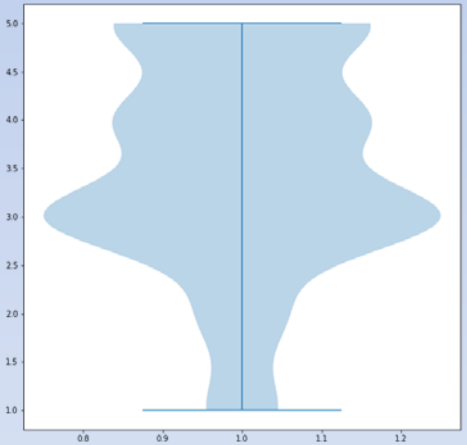


Continous variable 4 What is the Pin Code of where you shop online from?



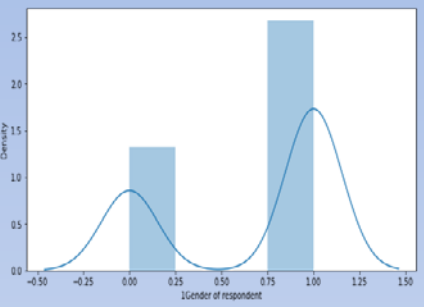


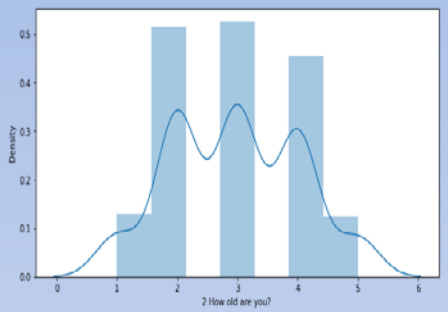


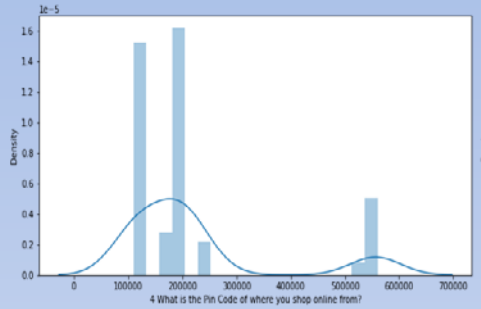


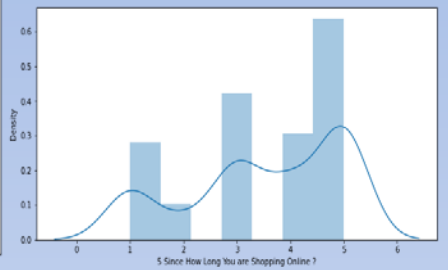
Distribution plots

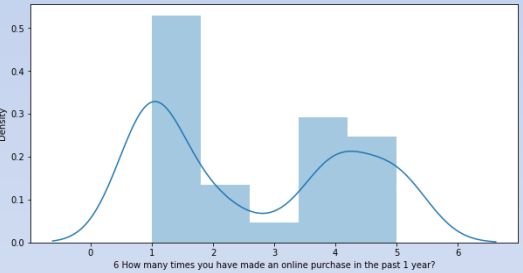
A Distplot or distribution plot, depicts the variation in the data distribution. Seaborn Distplot represents the overall distribution of continuous data variables. The Seaborn module along with the Matplotlib module is used to depict the distplot with different variations in it.

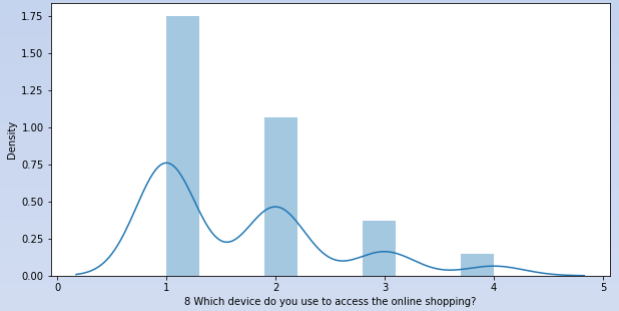






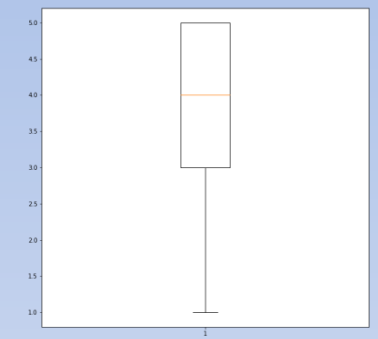
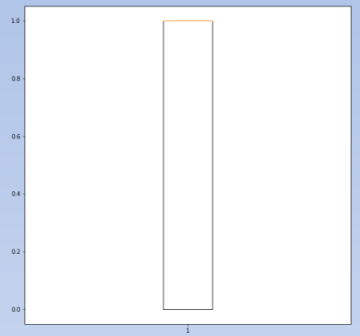
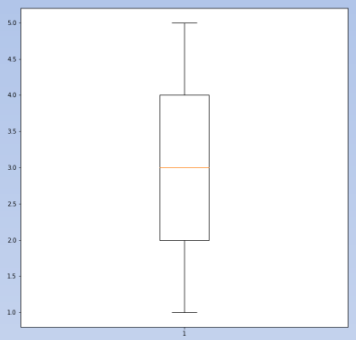
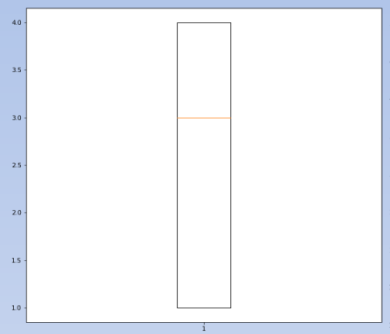
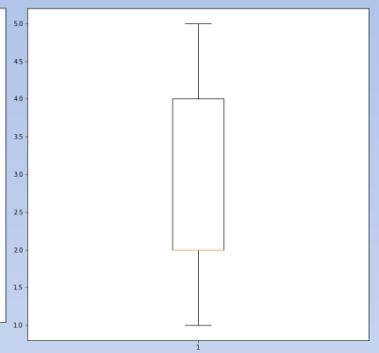
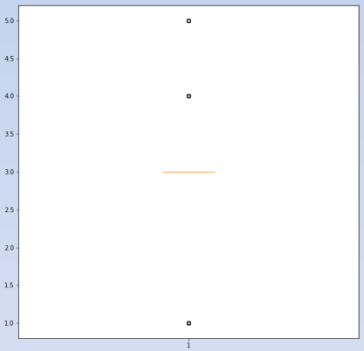






Used Boxplots with contionous data

- A box plot is a graph of **the distribution of a continuous** variable. ... The plot uses a box to show the values that are larger than the first quartile and smaller than the fourth quartile. These are the values that are closest to the center (median) of the values

Amount Related variables have outliers which affects overall analysis

OBSERVATION

#THE ABOVE RESULTS ARE SUGGESTING THERE ARE NO MUCH OUTLIERS IN DATA

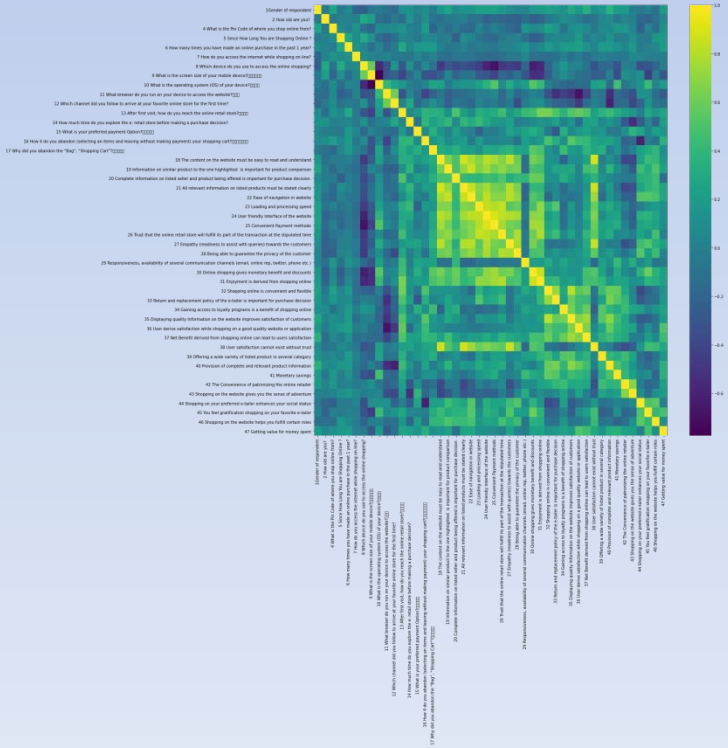
Missing values in data

- **Missing value is the value of blank**. We often meet them when we analyze large size data. Outlier and missing value are also called "abnormal value", "noise", "trash", "bad data" and "incomplete data".

- **Observation:**
- #There are no misssing values in our data

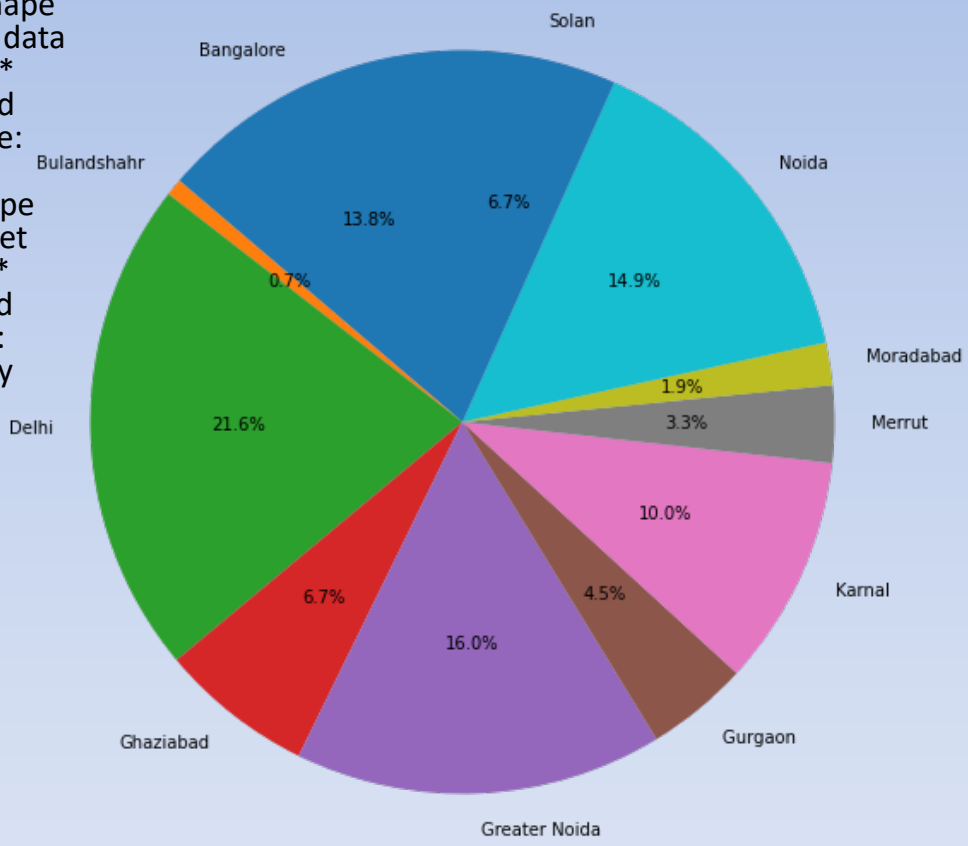
Correlation Analaysis

- If both variables tend **to increase or decrease together**, the coefficient is positive, and the line that represents the correlation slopes upward. If one variable tends to increase as the other decreases, the coefficient is negative, and the line that represents the correlation slopes downward.



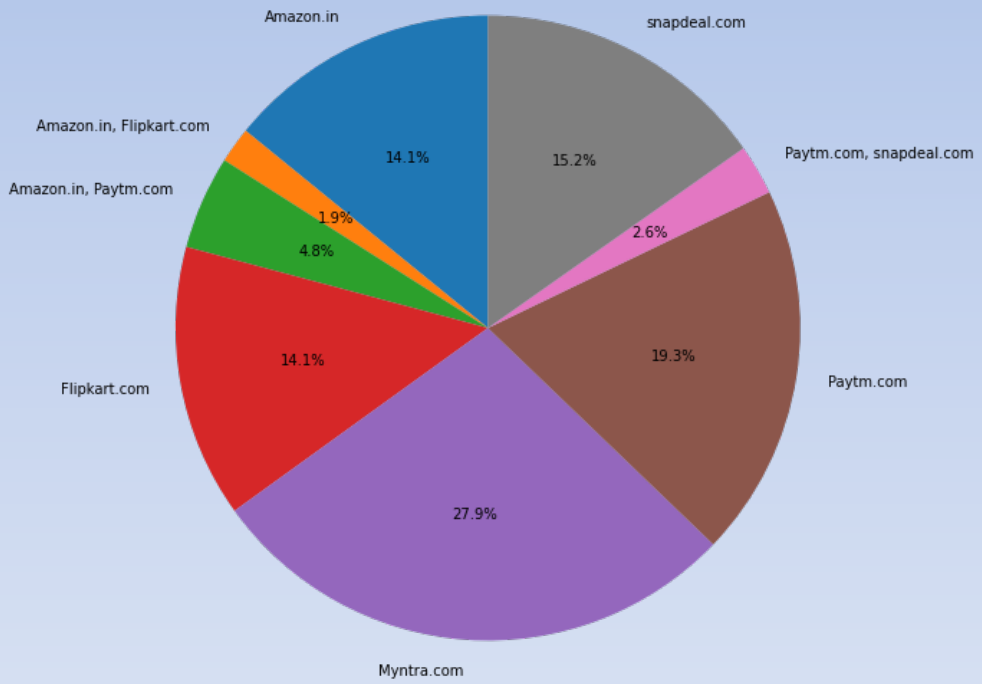
City wise group the data - Groupby

- **Grouping Data for better understanding**
- ********City Name: Bangalore ********* data set shape (37, 71) ********City Name: Bulandshahr ********* data set shape (2, 71) ********City Name: Delhi ********* data set shape (58, 71) ********City Name: Ghaziabad ********* data set shape (18, 71) ********City Name: Greater Noida ********* data set shape (43, 71) ********City Name: Gurgaon ********* data set shape (12, 71) ********City Name: Karnal ********* data set shape (27, 71) ********City Name: Merrut ********* data set shape (9, 71) ********City Name: Moradabad ********* data set shape (5, 71) ********City Name: Noida ********* data set shape (40, 71) ********City Name: Solan ********* data set shape (18, 71)
-



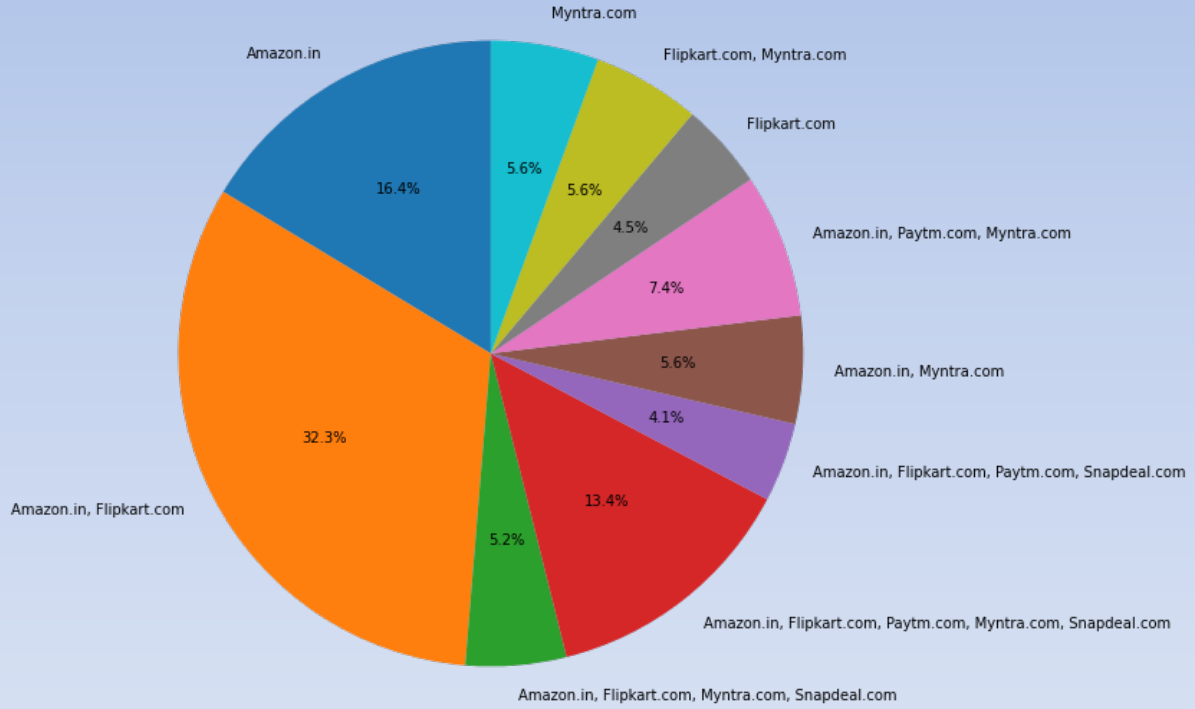
Source wise group the data - Groupby

- In exploratory data analysis, we often would like to analyze data by some categories. In SQL, the GROUP BY statement groups row that has the same category values into summary rows. In Pandas, SQL's GROUP BY operation is performed using the similarly named groupby() method. Pandas' groupby() allows us to split data into separate groups to perform computations for better analysis.



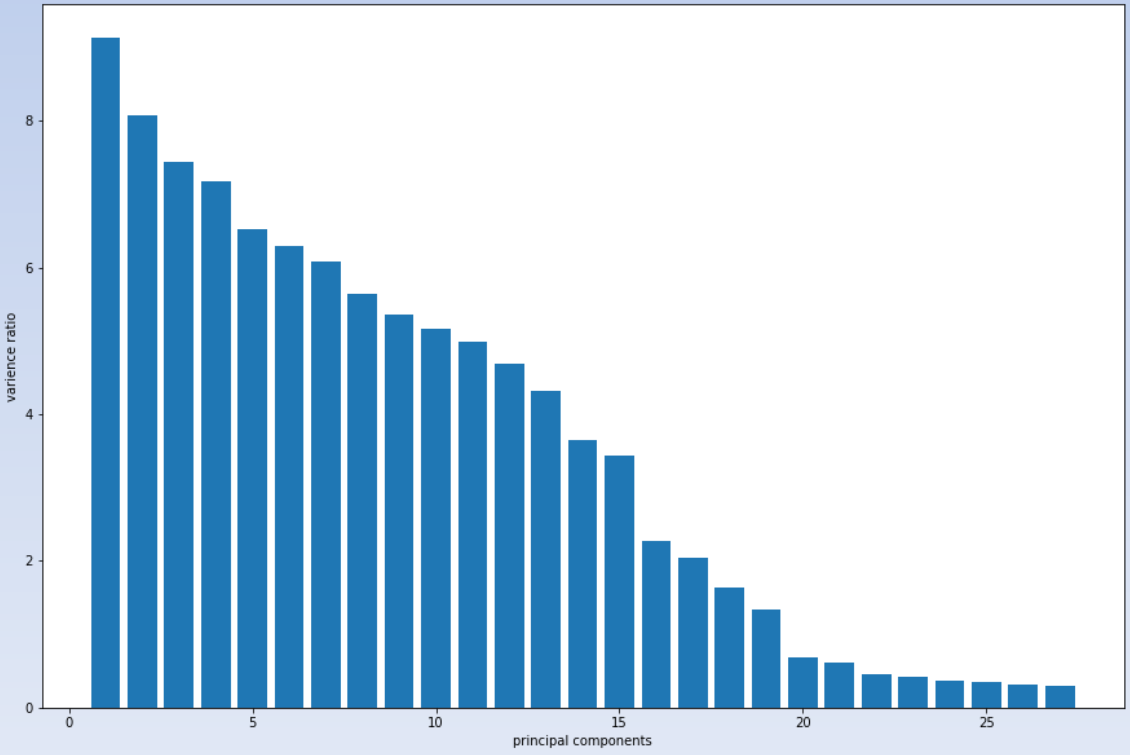
Visual appealing web-page layout wise Grouping the data

- visually appealing website starts with solid design, and a conscious color scheme. The colors that you use resonate with different feelings, and can trigger different responses. When defining your brand, carry that over onto your website to create a holistic experience of you and your company.



Conclusion

As can be seen from the analysis, various factors for selecting online retailers are website design and response which is as found in various review of literature. Because if the response in not quick, the customer will try to look for other websites. Second factor is actual service provided by the online retailer. In addition security while doing online transaction and confidentiality of data is an important factor in the selection of online retailer. Last but not the least; service provided by online retailer in case of any problem is also important factor in India for millennials. As the millennials are young and dynamic, they expect the same type of response from the website



Through PCA reduced the dimension till 27 and captured variance is 98%. This is signficant data diemnsion technique