**Problem Statement:-**

Analyzing the data present in the dataset and finding out the factors essential for retaining customers. Also finding out which online shopping platform has provides better service to their customers in terms of good interface, good customers service, privacy, processing speed etc.

**Explanation of dataset:-**

This dataset has 269 rows and 71 columns. From the first 47 rows, we will try to find out, what the customers feel about online shopping, the factors which are helpful in online shopping as per customers. It includes some important columns like gender of respondent, city and pin-code of respondent, since how many years is the customer doing online shopping, screen-size and operating system of device along with other columns.

The columns from 47-71 is basically a comparison between various shopping platforms as per the data given by customers. It includes various columns like, easy to use website or application, wild variety of product on offer, detailed information on products, security of customer financial transaction, methods of payment, longer delivery period etc.

**Explanation of EDA:-**

Assumption in EDA:- We have focused only on customers who are doing online shopping for more than 2 years. Since they are shopping for longer time, we are assuming that their views are essential to know what factors enhance customer retention.

Mostly histograms are plotted to visualize the data. In some cases plotting the histogram was not possible, there we have used DataFrame.value\_counts() to count the data. Here strip plot is preferred instead of scatter plot because of the discrete data present in the columns. We have plotted strip plot between columns which seems to have high correlation.

**Explanation of approach:-**

In columns 47-81 ,the users have given multiple preferences. So each cell in the columns contains names of multiple shopping sites. So it is not possible to draw any conclusion by plotting the histogram of the columns. So we have extracted each shopping site names from the cells in each column and added them in a list and plotted a histogram on the list. Greater the number of times the shopping name appeared in the list, the more the shopping site is preferred by customers.

**Conclusion:-**

1.Customers within age 20-50 has shopped more number of times.

2.Delhi has maximum number of online shoppers. Cities having smaller area and population has less number of online shoppers.

3. Most of the users use mobile internet for online shopping followed by Wi-Fi and dial up connection

4. A huge number of customers have shopped less than 10 times which means they shop online only when they have no alternative.

5.Most customers prefer online shopping because they get value for money.

6.Most shoppers use smart phone for online shopping followed by laptop, desktop, tablet.

7.Most of the online shoppers have windows OS followed by android and ios.

8. Most people arrives first time to an online shopping website through a search engine. While the second time they mostly prefers visiting the site using an app.

9.Most shoppers search more than 15 minutes before making a purchase decision.

10.Most shoppers prefer paying through credit/debit cards.

11.Most customers have abandoned a shopping cart sometime or the other. They mostly abandon cart when they get better alternate offer, when promo-code don’t apply.

12.Customers prefers having information of similar product.

13.Complete information of seller is highly important for customers.

14.Higher processing speed is preferred by most customers.

15.User friendly interface, ease of navigation is highly preferred.

16.Convenient payment methods, trust on online retail store, highly responsive customer service, privacy of customer, return and replacement policy is very essential for retaining customers.

17. Most customers believe that online shopping is convenient and flexible.

18.Also customers prefer online shopping for wide variety.

19. Value for money, monetary savings plays a very important role in retaining customers.