Documentation

Project- Customer Retention

1. **Importing the required libraries-:**

- import numpy as np
- import pandas as pd
- import matplotlib.pyplot as plt
- import seaborn as sns
- %matplotlib inline
- import warnings
- warnings. filterwarnings('ignore')

2**. Importing dataset:**- The dataset is having 269 rows and 71columns.

3. **Columns** -:

'1Gender of respondent',
'2 How old are you? ',
'3 Which city do you shop online from?',
'4 What is the Pin Code of where you shop online from?',
'5 Since How Long You are Shopping Online ?',
'6 How many times you have made an online purchase in the past 1year?',
'7 How do you access the internet while shopping on-line?',
'8 Which device do you use to access the online shopping?',
'9 What is the screen size of your mobile device?\t\t\t\t\t\t ',
'10 What is the operating system (OS) of your device?\t\t\t\t ',
'11 What browser do you run on your device to access the website?\t\t\t ',
'12 Which channel did you follow to arrive at your favorite online store for the first time? ',
'13 After first visit, how do you reach the online retail store?\t\t\t\t ' ,
'14 How much time do you explore the e- retail store before making a purchase decision? ',
'15 What is your preferred payment Option?\t\t\t\t\t ',
'16 How frequently do you abandon (selecting an items and leaving without making payment) your shopping cart?\t\t\t\t\t\t\t ',
'17 Why did you abandon the “Bag”, “Shopping Cart”?\t\t\t\t\t ',

'18 The content on the website must be easy to read and understand',
'19 Information on similar product to the one highlighted is important for product comparison ',
'20 Complete information on listed seller and product being offered is important for purchase decision.',
'21 All relevant information on listed products must be stated clearly',
'22 Ease of navigation in website', '23 Loading and processing speed',
'24 User friendly Interface of the website',
'25 Convenient Payment methods',
'26 Trust that the online retail store will fulfill its part of the transaction at the stipulated time',
'27 Empathy (readiness to assist with queries) towards the customers',
'28 Being able to guarantee the privacy of the customer',
'29 Responsiveness, availability of several communication channels (email, online rep, twitter , phone etc.)',
'30 Online shopping gives monetary benefit and discounts',
'31 Enjoyment is derived from shopping online',
'32 Shopping online is convenient and flexible',
'33 Return and replacement policy of the e-tailer is important for purchase decision',
'34 Gaining access to loyalty programs is a benefit of shopping online',
'35 Displaying quality Information on the website improves satisfaction of customers',
'36 User derive satisfaction while shopping on a good quality website or application',
'37 Net Benefit derived from shopping online can lead to users satisfaction',
'38 User satisfaction cannot exist without trust',
'39 Offering a wide variety of listed product in several category',
'40 Provision of complete and relevant product information',
'41 Monetary savings',
'42 The Convenience of patronizing the online retailer',
'43 Shopping on the website gives you the sense of adventure',
'44 Shopping on your preferred e-tailer enhances your social status',
'45 You feel gratification shopping on your favorite e-tailer',
'46 Shopping on the website helps you fulfill certain roles',
'47 Getting value for money spent',
'From the following, tick any (or all) of the online retailers you have shopped from; ',
'Easy to use website or application',
'Visual appealing web-page layout', 'Wild variety of product on offer',
'Complete, relevant description information of products',
'Fast loading website speed of website and application',
'Reliability of the website or application',
'Quickness to complete purchase',
'Availability of several payment options', 'Speedy order delivery ',
'Privacy of customers' information',
'Security of customer financial information',
'Perceived Trustworthiness',
'Presence of online assistance through multi-channel',
'Longer time to get logged in (promotion, sales period)',
'Longer time in displaying graphics and photos (promotion, sales period)',
'Late declaration of price (promotion, sales period)',
'Longer page loading time (promotion, sales period)',
'Limited mode of payment on most products (promotion, sales period)',

'Longer delivery period', 'Change in website/Application design',
'Frequent disruption when moving from one page to another',
'Website is as efficient as before',
'Which of the Indian online retailer would you recommend to a friend?'],

4. **NULL**-: The dataset has no null values.

5**. EDA** :-

Visualisation of Data (categorical columns)

Cities

Delhi	58
Greater Noida	43
Noida	40
Bangalore	37
Karnal	27
Solan	18
Ghaziabad	18
Gurgaon	12
Merrut	9
Moradabad	5
Bulandshahr	2

- There are total 11 cities from which customers are doing online shopping and
- Delhi is having maximum no. of customers.

Duration of Online Shopping

Above 4 years	98
2-3 years	65
3-4 years	47
Less than 1 year	43
1-2 years	16

- Maximum customers are doing online transactions for more than 4 years.

Transactions made in past 1 year

Less than 10 times	114
31-40 times	63
41 times and above	47
11-20 times	29
21-30 times	10
42 times and above	6

- Maximum users made online transactions less than 10 times

Transaction Devices

Smartphone	141
Laptop	86
Desktop	30
Tablet	12

- Users do most of the transactions through smartphones

6. Dropping pin code column from dataset.

7. Encoding columns from 1 to 17 using Label Encoder as the nature of data are categorically nominal.

8.Encoding columns from 18 to 47 using Ordinal Encoder as the nature of data are following some order.
9. Plotting heatmap to check correlation among the features.

10.Standaradizing data to perform PCA.

11.Applying PCA for dimension reduction as number of columns are large.

- Imported PCA from sklearn. decomposition.
- Train and transformed data using PCA
- Plotted scree plot to get the optimum number of Principal Components.