FLIP ROBO

Customer Retention

Submitted by:

RAHUL RANJAN

ACKNOWLEDGMENT

I am using this opportunity to express my gratitude to everyone who supported me throughout the course of this project. I am thankful for their aspiring guidance, invaluably constructive criticism and friendly advice during the project work. I am sincerely grateful to them for sharing their truthful and illuminating views on a number of issues related to the project.
I wish to thank, all the faculties in data trained academy as this project utilized knowledge gained from every course that formed the Data science program.

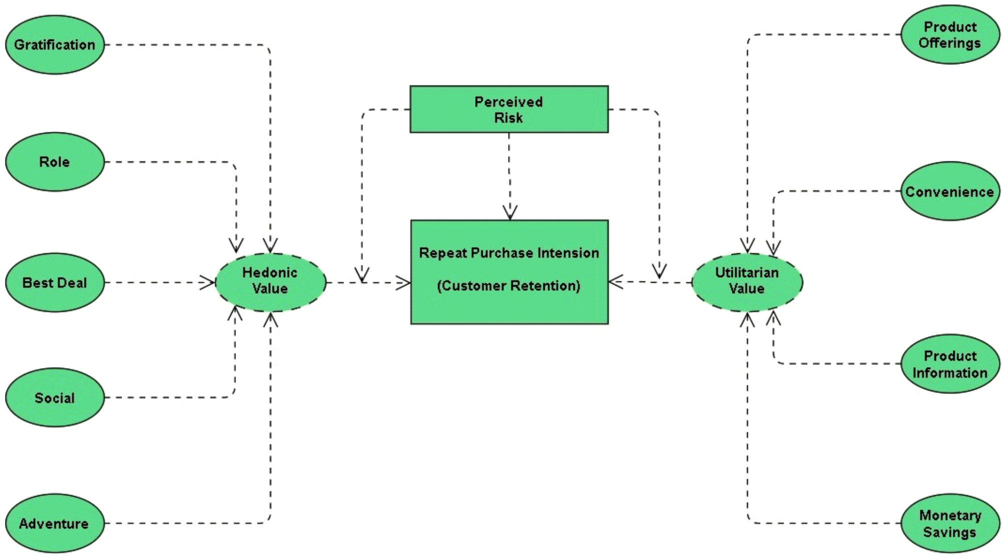
INTRODUCTION

Objective of the study

The objective of the project is to apply analytical skills to give findings and conclusions in detailed data analysis of E-retail factors for customer activation and retention.

Business Model

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 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



Literature Survey

Hedonic and utilitarian shopping values

A consumer's behaviour is a result of motives, attitudes and values and may manifest into purchase and consumption behaviour. Westbrook and Black (1985) posit that some shopping motives are utilitarian in nature whereas others are hedonic. The utilitarian and hedonic values have been the focus of much interest and research (Hirschman and Holbrook, 1982; Batra and Ahtola, 1991; Babin et al, 1994; Wang et al, 2000; Millan and Howard, 2007; Teller et al, 2008). Consumer values have been broadly termed as utilitarian (Bloch and Bruce, 1984; Batra and Ahtola, 1991; Engel et al, 1993; Babin et al, 1994) which are more task oriented in nature and hedonic which are related to entertainment and fun-seeking behaviour (Bellenger et al, 1976). Bloch and Richins (1983) postulate that hedonic values are characterized by heightened arousal, excitement, adventure and entertainment. Shopping behaviour provides excitement whereas the consumer interacts with the store environment and gives cues while they examine products (Maclnnis and Price, 1987) that may be perceived as enjoyment.

Consumers with strong hedonic values may not be satisfied with the functional aspects of shopping and may look for pleasurable stimulants (Fischer and Arnold, 1990; Wang et al, 2000). The hedonic values are related to gratification of the senses enhanced through experiences of pleasure, entertainment, fantasy and playfulness (Hirschman and Holbrook, 1982; Babin et al, 1994). The consumer values have been defined in terms of being intrinsic and extrinsic; the extrinsic values are related to the functional attributes of shopping, and are mainly 'utilitarian' in nature. The intrinsic values signify the 'enjoyment, fun and leisure' motives (Babin et al, 1994). The utilitarian values are based upon rational and analytical information processing whereas the hedonic values comprise of arousal of the senses (Holbrook and Hirschman, 1982; Hirschman, 1983; Fischer and Arnold, 1990) and self gratification.

The traditional shopping behaviours of product acquisition and consumption may no longer explain the shopping 'experience' the consumers seek when they go to a store or a mall. They look beyond mere assortment of products and functional attributes. Babin et al (1994) state that most consumption activities must combine both utilitarian and hedonic attributes and their absence may not reflect the totality of shopping experience (Bloch and Richins, 1983). Research in the past few years has recognized the pivotal role hedonic values play in shopping and how they add to the emotional value (Langrehr, 1991; Babin et al, 1994; Roy, 1994).

EDA Steps

1.Checking the missing values

2.Checking for numerical columns

3.Checking for the distribution of numerical variables

4.Checking for Categorical variables

5.Types of categorical variables

6.detecting outliers

Analytical Problem Framing

```
In [6]: data.head()
```

Out[6]:

	Gender	Age	City	Pincode	YearsOfOnlineShopping	PurchaseInPast1Year	InternetAccess	DeviceUsed	MobileScre
0	Male	31-40 years	Delhi	110009	Above 4 years	31-40 times	Dial-up	Desktop	
1	Female	21-30 years	Delhi	110030	Above 4 years	41 times and above	Wi-Fi	Smartphone	4.7
2	Female	21-30 years	Greater Noida	201308	3-4 years	41 times and above	Mobile Internet	Smartphone	5.5
3	Male	21-30 years	Karnal	132001	3-4 years	Less than 10 times	Mobile Internet	Smartphone	5.5

```
In [7]: # Checking shape in our dataset
        data.shape
```

Out[7]: (269, 71)

Data Preprocessing

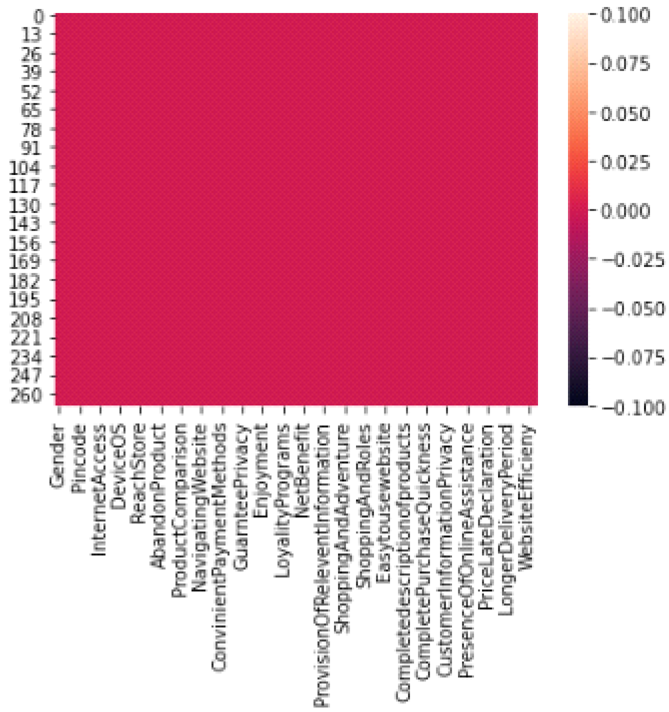
Replacing columns

```
data.columns
```

```
Index(['Gender', 'Age', 'City', 'Pincode', 'YearsOfOnlineShopping',
       'PurchaseInPast1Year', 'InternetAccess', 'DeviceUsed',
       'MobileScreenSize', 'DeviceOS', 'DeviceBrowser', 'Channel',
       'ReachStore', 'ExploreTime', 'PaymentOption', 'AbandonProduct',
       'WhyAbandon', 'WebsiteContent', 'ProductComparison', 'PurchaseDecision',
       'RelevantInfoOnListedProducts', 'NavigatingWebsite', 'Speed',
       'UserFriendlyInterface', 'ConvinientPaymentMethods', 'Trust', 'Empathy',
       'GuarnteePrivacy', 'CommunicationChannelsAvailability',
       'BenifitsAndDiscounts', 'Enjoyment', 'Flexibility',
       'ReturnAndReplacementPolicy', 'LoyalityPrograms',
       'DisplayQualityInformation', 'goodqualitywebsitesatisfaction',
       'NetBenefit', 'UserSatisfactionAndTrust', 'WideVarietyOfProducts',
       'ProvisionOfReleventInformation', 'MonetarySavings',
       'Convenienceofpatronizingonlineretailer', 'ShoppingAndAdventure',
       'ShoppingAndSocialStatus', 'ShoppingAndGratification',
       'ShoppingAndRoles', 'ValueForMoneySpent', 'OnlineRetailersShoppedFrom',
       'Easytousewebsite', 'Visualappealingwebpagelayout',
       'Varietyofproductonoffer', 'Completedescriptionofproducts',
       'FastLoadingWebsiteSpeed', 'WebSiteReliability',
       'CompletePurchaseQuickness', 'PaymentsOptionsAvailability',
       'Speedyorderdelivery', 'CustomerInformationPrivacy',
       'CustomerFinancialInformationSecurity', 'PerceivedTrustworthiness',
       'PresenceOfOnlineAssistance', 'LogInTime', 'DisplayGraphicsTime',
       'PriceLateDeclaration', 'PageLoadingTime', 'LimitedModeOfPayment',
       'LongerDeliveryPeriod', 'WebsiteDesignChange',
       'FrequentDisruptionInMovingFromPageToPage', 'WebsiteEfficieny',
       'IndianOnlineRetailerToRecommend'],
      dtype='object')
```

Dealing with null values

In [16]:
```
sns.heatmap(data.isnull())
```
Out[16]: <AxesSubplot:>



No null values present

Data Info And Description

```
Data columns (total 71 columns):
 #   Column                                 Non-Null Count  Dtype
---  ------                                 --------------  -----
 0   Gender                                 269 non-null    object
 1   Age                                    269 non-null    object
 2   City                                   269 non-null    object
 3   Pincode                                269 non-null    int64
 4   YearsOfOnlineShopping                  269 non-null    object
 5   PurchaseInPast1Year                    269 non-null    object
 6   InternetAccess                         269 non-null    object
 7   DeviceUsed                             269 non-null    object
 8   MobileScreenSize                       269 non-null    object
 9   DeviceOS                               269 non-null    object
 10  DeviceBrowser                          269 non-null    object
 11  Channel                                269 non-null    object
 12  ReachStore                             269 non-null    object
 13  ExploreTime                            269 non-null    object
 14  PaymentOption                          269 non-null    object
 15  AbandonProduct                         269 non-null    object
 16  WhyAbandon                             269 non-null    object
 17  WebsiteContent                         269 non-null    object
 18  ProductComparison                      269 non-null    object
 19  PurchaseDecision                       269 non-null    object
 20  RelevantInfoOnListedProducts           269 non-null    object
 21  NavigatingWebsite                      269 non-null    object
 22  Speed                                  269 non-null    object
 23  UserFriendlyInterface                  269 non-null    object
 24  ConvinientPaymentMethods               269 non-null    object
 25  Trust                                  269 non-null    object
 26  Empathy                                269 non-null    object
 27  GuarnteePrivacy                        269 non-null    object
 28  CommunicationChannelsAvailability      269 non-null    object
 29  BenifitsAndDiscounts                   269 non-null    object
 30  Enjoyment                              269 non-null    object
 31  Flexibility                            269 non-null    object
 32  ReturnAndReplacementPolicy             269 non-null    object
 33  LoyalityPrograms                       269 non-null    object
 34  DisplayQualityInformation              269 non-null    object
 35  goodqualitywebsitesatisfaction         269 non-null    object
 36  NetBenefit                             269 non-null    object
 37  UserSatisfactionAndTrust               269 non-null    object
 38  WideVarietyOfProducts                  269 non-null    object
 39  ProvisionOfReleventInformation         269 non-null    object
 40  MonetarySavings                        269 non-null    object
```

In [12]:
```
data.describe()
```
Out[12]:

	Pincode
count	269.000000
mean	220465.747212
std	140524.341051
min	110008.000000
25%	122018.000000
50%	201303.000000
75%	201310.000000
max	560037.000000

In [13]:
```
data.describe(include='object')
```
Out[13]:

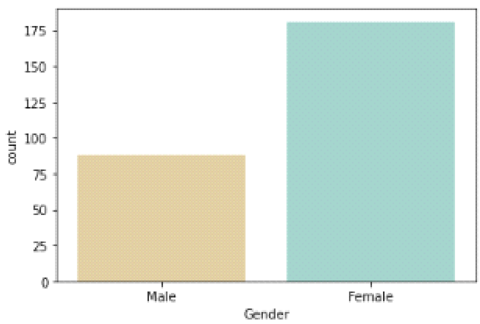
	Gender	Age	City	YearsOfOnlineShopping	PurchaseInPast1Year	InternetAccess	DeviceUsed	MobileScreenSize	DeviceOS	DeviceBrowser	Ch
count	269	269	269	269	269	269	269	269	269	269	
unique	2	5	11	5	6	4	4	4	3	4	
top	Female	31-40 years	Delhi	Above 4 years	Less than 10 times	Mobile internet	Smartphone	Others	Window/windows Mobile	Google chrome	S E
freq	181	81	58	98	114	142	141	134	122	216	

Data Visualization

Univariate Analysis

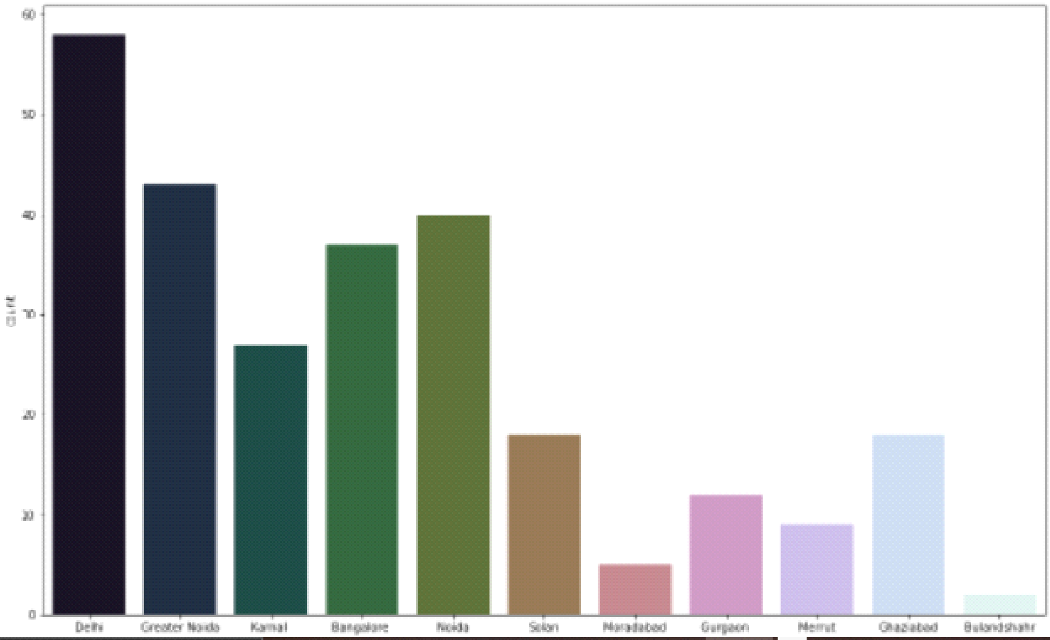
We will use a single characteristic to assess practically all of its attributes in a univariate analysis.

']:
```
# Let's Plot countplot for each columns to check unique values counts.
sns.countplot(data['Gender'],palette='BrBG')
data['Gender'].value_counts()
```
']:
```
Female    181
Male       88
Name: Gender, dtype: int64
```



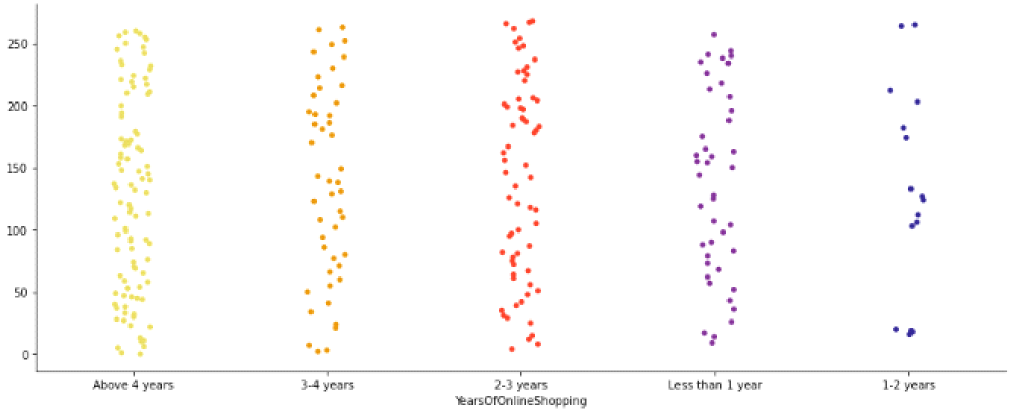
```
# Let's check different cities in our data collection.
plt.figure(figsize=(16,10))
sns.countplot(data['City'],palette='cubehelix')
data['City'].value_counts()
```

```
Delhi             58
Greater Noida     43
Noida             40
Bangalore         37
Karnal            27
Solan             18
Ghaziabad         18
Gurgaon           12
Merrut             9
Moradabad          5
Bulandshahr        2
Name: City, dtype: int64
```



```
]: # Ploting catplot for column 5 Since How Long You are Shopping Online ?
sns.catplot(x='YearsOfOnlineShopping',y= data.index,data=data, palette='CMRmap_r',aspect=2.5)
data['YearsOfOnlineShopping'].value_counts()
```

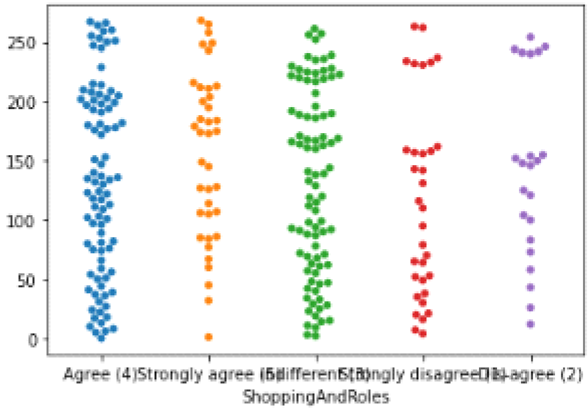
```
]: Above 4 years        98
2-3 years            65
3-4 years            47
Less than 1 year     43
1-2 years            16
Name: YearsOfOnlineShopping, dtype: int64
```



From above catplot we can say that we have some fixed amount of people who do online shop regularly also they are the customers who already shopping from more then 4 years, also very less people who recently doing online shopping.

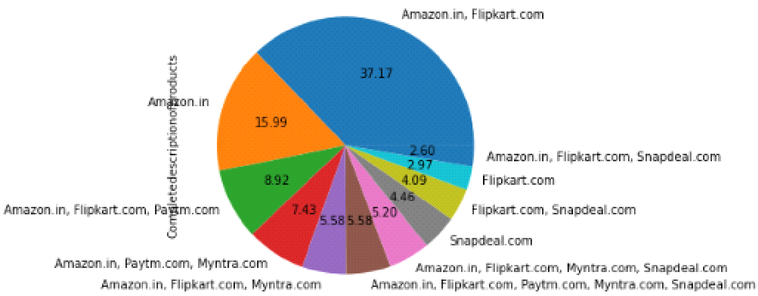
```
]: sns.swarmplot(x= 'ShoppingAndRoles',
                 y= data.index, data= data)
data['ShoppingAndRoles'].value_counts()
```

```
]: Agree (4)                88
indifferent (3)          88
Strongly agree (5)       38
Strongly disagree (1)    33
Dis-agree (2)            22
Name: ShoppingAndRoles, dtype: int64
```



```
]: x = data['Completedescriptionofproducts'].value_counts([0])
   x.plot(kind = 'pie', figsize = (5,5), fontsize=10, autopct = '%.2f')
   data['Completedescriptionofproducts'].value_counts()
```

```
]: Amazon.in, Flipkart.com                                      100
   Amazon.in                                                      43
   Amazon.in, Flipkart.com, Paytm.com                             24
   Amazon.in, Paytm.com, Myntra.com                               20
   Amazon.in, Flipkart.com, Myntra.com                            15
   Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com   15
   Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com              14
   Snapdeal.com                                                   12
   Flipkart.com, Snapdeal.com                                     11
   Flipkart.com                                                    8
   Amazon.in, Flipkart.com, Snapdeal.com                           7
   Name: Completedescriptionofproducts, dtype: int64
```

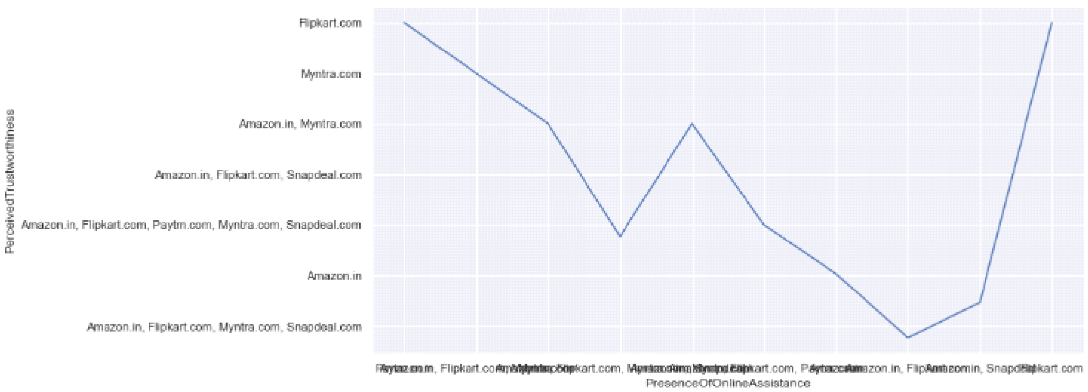



```
]: x = data['FastLoadingWebsiteSpeed'].value_counts([0])
```

Bivariate Analysis:

Bivariate analysis is when we compare data between two attributes that are precisely the same.

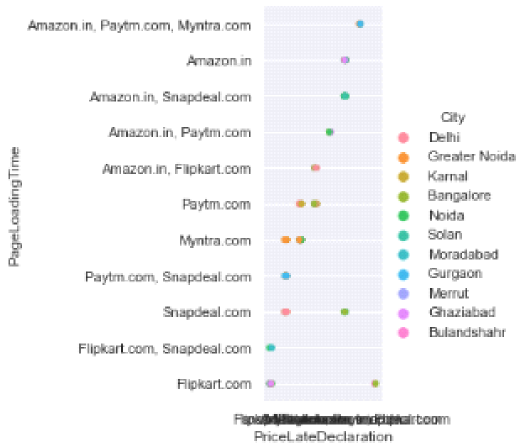
```
]: plt.figure(figsize=(12, 6))
   sns.set_theme(style="darkgrid")
   sns.lineplot(data=data, x="PresenceOfOnlineAssistance", y="PerceivedTrustworthiness", ci=None)
   plt.show()
```



```
]: plt.figure(figsize=(20,5))
   sns.catplot(x= 'PriceLateDeclaration', y= 'PageLoadingTime',hue='City', data=data)
   data['PageLoadingTime'].value_counts()
```

```
]: Myntra.com                        61
   Paytm.com                         59
   Flipkart.com                      32
   Snapdeal.com                      23
   Amazon.in, Flipkart.com           18
   Amazon.in                         16
   Paytm.com, Snapdeal.com           15
   Amazon.in, Snapdeal.com           14
   Amazon.in, Paytm.com              13
   Flipkart.com, Snapdeal.com        11
   Amazon.in, Paytm.com, Myntra.com   7
   Name: PageLoadingTime, dtype: int64

   <Figure size 1440x360 with 0 Axes>
```

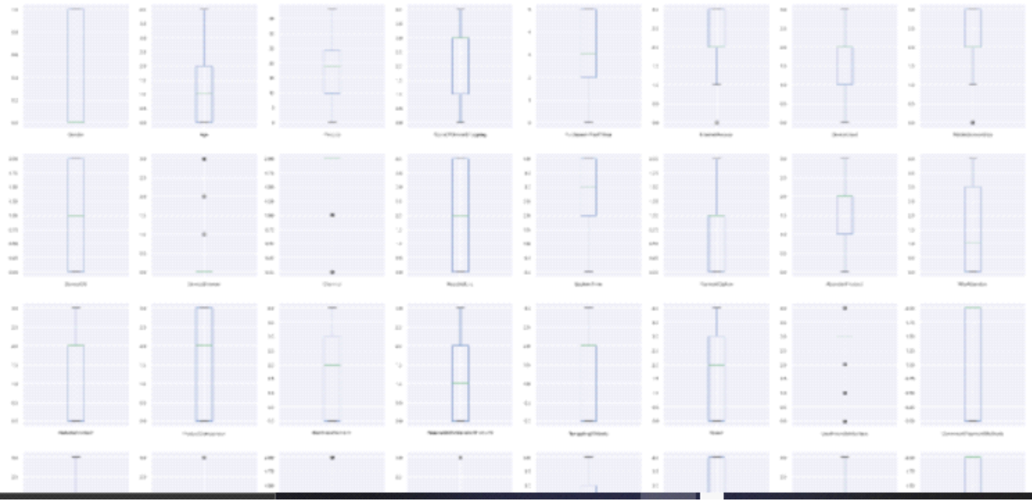


Multivarite Analysis:

We shall compare more than two variables in the multivariate analysis.

```
In [81]: #plotting the boxplot for each column in order to check the outliers
df.plot(kind='box',subplots = True,layout=(10,8),figsize = (40,60))
```

```
Out[81]: Gender                                    AxesSubplot(0.125,0.816017;0.0824468x0.0639831)
Age                                      AxesSubplot(0.223936,0.816017;0.0824468x0.0639...
Pincode                                  AxesSubplot(0.322872,0.816017;0.0824468x0.0639...
YearsOfOnlineShopping                    AxesSubplot(0.421809,0.816017;0.0824468x0.0639...
PurchaseInPast1Year                      AxesSubplot(0.520745,0.816017;0.0824468x0.0639...
                                                                    ...
WebsiteDesignChange                      AxesSubplot(0.322872,0.20178;0.0824468x0.0639831)
FrequentDisruptionInMovingFromPageToPage AxesSubplot(0.421809,0.20178;0.0824468x0.0639831)
WebsiteEfficieny                         AxesSubplot(0.520745,0.20178;0.0824468x0.0639831)
IndianOnlineRetailerToRecommend          AxesSubplot(0.619681,0.20178;0.0824468x0.0639831)
City                                     AxesSubplot(0.718617,0.20178;0.0824468x0.0639831)
Length: 71, dtype: object
```



Using of Label Encoder to encode catagorical data:

It is used to convert

Now Importing Label Encoder to encode catagorical data.

```
]: from sklearn.compose import make_column_transformer     # Importing colum transformer to apply encoding technique.
from sklearn.preprocessing import LabelEncoder    # Importing Lable Encoder to encode catagorical colummns.
```

```
]: le = LabelEncoder()   # Instantsiating Label encoder for converting our catagorical data.
```

```
]: # Applying Label Encoder on selected columns.
df[headers] = df[headers].apply(le.fit_transform)
```

```
]: df.head()
```

]:

	Gender	Age	Pincode	YearsOfOnlineShopping	PurchaseInPast1Year	InternetAccess	DeviceUsed	MobileScreenSize	DeviceOS	DeviceBrowser	Channel	Re
0	1	1	1	3	2	0	0	3	2	0	2	
1	0	0	5	3	3	3	2	0	1	0	2	
2	0	0	23	2	3	1	2	2	0	0	2	
3	1	0	11	2	5	1	2	2	1	3	2	
4	0	0	31	1	0	3	2	0	1	3	0	

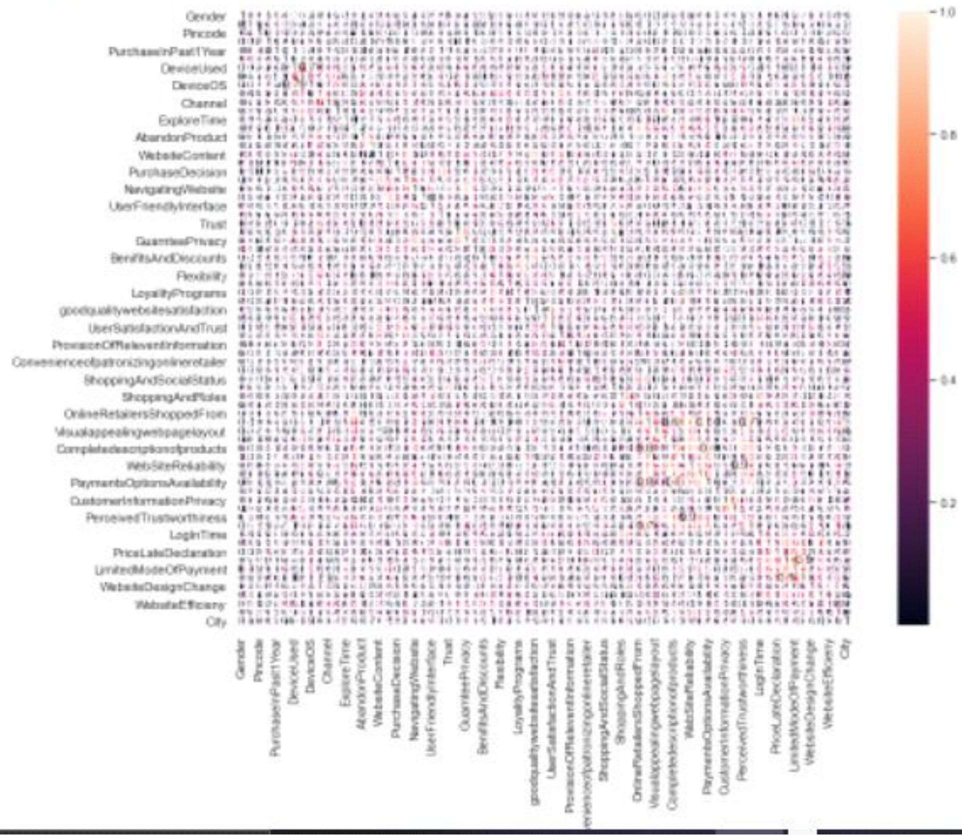
Now we can see all the catagorical columns in our dataset are conveted into Numeric value.

Correlation:

- Correlation: correlation between the available features and customer retention will be evaluated using the Pearson Coefficient Correlation (default correlation method) to identify whether the features have a negative, positive or zero correlation with regards the customer repurchase chances.

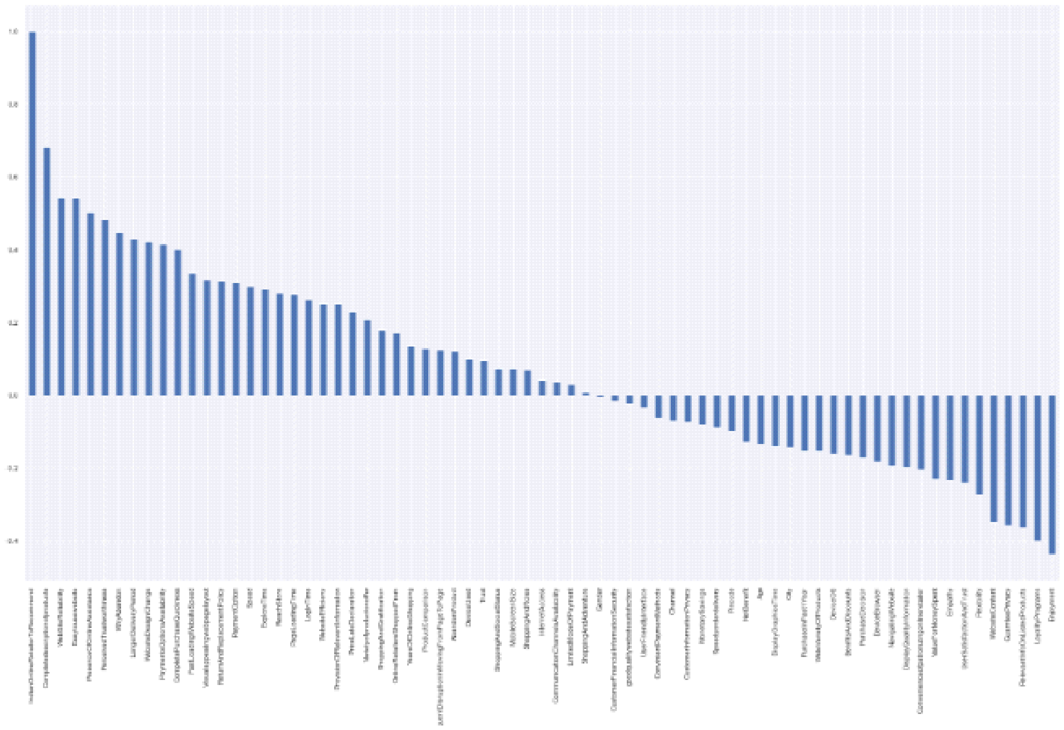
```
# Let's check correlationship between target variable and features variable.
data_corr = df.corr().abs()
plt.figure(figsize=(15,10))
sns.heatmap(data=data_corr,linewidths=1,annot=True, square=True,)
```

```
<AxesSubplot:>
```



[84]:
```
#correlation of independent features with regards the target "Price"
plt.figure(figsize=(28, 16))
correlations =df.corr()['IndianOnlineRetailerToRecommend'].sort_values(ascending=False)
correlations.plot(kind='bar')
```

[84]: `<AxesSubplot:>`



unique values and their correspending counts of each column:

```
---------------------------------------
1     100
0      43
4      24
7      20
2      15
5      15
3      14
10     12
9      11
8       8
6       7
Name: Completedescriptionofproducts, dtype: int64
---------------------------------------
0    51
7    44
1    30
2    30
5    30
4    25
6    25
3    14
9    12
8     8
Name: FastLoadingWebsiteSpeed, dtype: int64
---------------------------------------
0    61
1    50
3    36
6    35
5    18
7    15
8    15
2    14
4    13
9    12
Name: WebSiteReliability, dtype: int64
---------------------------------------
0    66
3    47
1    37
2    30
8    25
5    20
4    15
6    15
7    14
Name: CompletePurchaseQuickness, dtype: int64
---------------------------------------
1     65
2     40
```

```
3    33
1    22
Name: ShoppingAndRoles, dtype: int64
----------------------------------------
0    149
1     82
2     38
Name: ValueForMoneySpent, dtype: int64
----------------------------------------
4    82
2    44
1    32
5    29
6    27
8    20
0    16
7    12
3     7
Name: OnlineRetailersShoppedFrom, dtype: int64
----------------------------------------
4    64
1    44
3    44
0    29
5    22
7    20
2    19
9    12
8     8
6     7
Name: Easytousewebsite, dtype: int64
----------------------------------------
1    87
0    44
3    36
6    20
5    15
8    15
9    15
2    14
7    12
4    11
Name: Visualappealingwebpagelayout, dtype: int64
----------------------------------------
1    130
0     43
4     20
6     15
7     15
2     14
3     13
5     12
```

```
1      5
Name: ProvisionOfReleventInformation, dtype: int64
---------------------------------------
2    148
0     75
1     31
3     15
Name: MonetarySavings, dtype: int64
---------------------------------------
0    138
2     77
1     54
Name: Convenienceofpatronizingonlineretailer, dtype: int64
---------------------------------------
0    101
4     59
2     54
1     50
3      5
Name: ShoppingAndAdventure, dtype: int64
---------------------------------------
4    100
0     59
2     48
3     33
1     29
Name: ShoppingAndSocialStatus, dtype: int64
---------------------------------------
4    101
2     65
0     63
1     22
3     18
Name: ShoppingAndGratification, dtype: int64
---------------------------------------
0    88
4    88
2    38
3    33
1    22
Name: ShoppingAndRoles, dtype: int64
---------------------------------------
0    149
1     82
2     38
Name: ValueForMoneySpent, dtype: int64
---------------------------------------
4    82
2    44
1    32
5    29
```

```
1     20
Name: ReturnAndReplacementPolicy, dtype: int64
---------------------------------------
2    115
0     64
4     64
1     15
3     11
Name: LoyalityPrograms, dtype: int64
---------------------------------------
1    133
0     80
2     56
Name: DisplayQualityInformation, dtype: int64
---------------------------------------
2    175
0     86
1      8
Name: goodqualitywebsitesatisfaction, dtype: int64
---------------------------------------
2    164
0     54
3     40
1     11
Name: NetBenefit, dtype: int64
---------------------------------------
2    122
0    117
3     18
1      7
4      5
Name: UserSatisfactionAndTrust, dtype: int64
---------------------------------------
2    111
0     94
3     57
1      7
Name: WideVarietyOfProducts, dtype: int64
---------------------------------------
2    135
0     98
3     31
1      5
Name: ProvisionOfReleventInformation, dtype: int64
---------------------------------------
2    148
0     75
```

```
1     30
3     12
Name: Trust, dtype: int64
----------------------------------------
1    194
0     42
2     18
3     15
Name: Empathy, dtype: int64
----------------------------------------
1    185
0     58
2     26
Name: GuarnteePrivacy, dtype: int64
----------------------------------------
1    149
0     94
3     15
2     11
Name: CommunicationChannelsAvailability, dtype: int64
----------------------------------------
2    105
0     85
4     50
3     18
1     11
Name: BenifitsAndDiscounts, dtype: int64
----------------------------------------
2    86
4    75
0    59
3    30
1    19
Name: Enjoyment, dtype: int64
----------------------------------------
2    146
0     78
3     33
1     12
Name: Flexibility, dtype: int64
----------------------------------------
2    198
0     51
1     20
Name: ReturnAndReplacementPolicy, dtype: int64
----------------------------------------
2    115
0     64
4     64
1     15
3     11
```

```
3    116
0     92
2     43
1     18
Name: ProductComparison, dtype: int64
---------------------------------------
0    101
3     87
2     52
1     18
4     11
Name: PurchaseDecision, dtype: int64
---------------------------------------
0    132
2    107
3     18
1     12
Name: RelevantInfoOnListedProducts, dtype: int64
---------------------------------------
2    141
0    105
3     18
1      5
Name: NavigatingWebsite, dtype: int64
---------------------------------------
3    115
0    112
1     18
2     12
4     12
Name: Speed, dtype: int64
---------------------------------------
3    189
0     45
4     18
1     12
2      5
Name: UserFriendlyInterface, dtype: int64
---------------------------------------
2    159
0     80
1     30
Name: ConvinientPaymentMethods, dtype: int64
---------------------------------------
2    141
0     86
1     30
3     12
Name: Trust, dtype: int64
---------------------------------------
1    194
```

```
1      5
Name: DeviceBrowser, dtype: int64
---------------------------------------
2    230
0     20
1     19
Name: Channel, dtype: int64
---------------------------------------
2    87
4    86
0    70
1    18
3     8
Name: ReachStore, dtype: int64
---------------------------------------
4    123
2     71
1     46
3     15
0     14
Name: ExploreTime, dtype: int64
---------------------------------------
1    148
0     76
2     45
Name: PaymentOption, dtype: int64
---------------------------------------
2    171
1     48
0     35
3     15
Name: AbandonProduct, dtype: int64
---------------------------------------
0    133
4     54
1     37
2     31
3     14
Name: WhyAbandon, dtype: int64
---------------------------------------
2    164
0     80
3     18
1      7
Name: WebsiteContent, dtype: int64
---------------------------------------
3    116
0     92
2     43
1     18
```

```
26     2
27     1
28     1
32     1
Name: Pincode, dtype: int64
---------------------------------------
3    98
1    65
2    47
4    43
0    16
Name: YearsOfOnlineShopping, dtype: int64
---------------------------------------
5    114
2     63
3     47
0     29
1     10
4      6
Name: PurchaseInPast1Year, dtype: int64
---------------------------------------
2    142
3     76
1     47
0      4
Name: InternetAccess, dtype: int64
---------------------------------------
2    141
1     86
0     30
3     12
Name: DeviceUsed, dtype: int64
---------------------------------------
3    134
2     99
0     29
1      7
Name: MobileScreenSize, dtype: int64
---------------------------------------
2    122
0     85
1     62
Name: DeviceOS, dtype: int64
---------------------------------------
0    216
3     40
2      8
1      5
Name: DeviceBrowser, dtype: int64
---------------------------------------
2    230
0     20
```

```
]: #unique values and their corresponding counts of each column:
for col in df.columns:
    print(df[col].value_counts())
    print('---------------------------------------')
```

```
0    181
1     88
Name: Gender, dtype: int64
---------------------------------------
1    81
0    79
2    70
4    20
3    19
Name: Age, dtype: int64
---------------------------------------
23    38
11    19
24    18
8     16
30     9
14     9
13     9
38     8
12     8
35     8
10     8
22     7
0      7
2      7
3      6
4      6
29     5
17     5
18     5
21     5
31     5
25     5
9      4
37     4
34     4
33     4
19     4
7      4
5      4
6      4
20     4
1      4
16     4
15     4
```

```
7     14
Name: CompletePurchaseQuickness, dtype: int64
---------------------------------------
1     65
2     40
4     39
0     23
10    20
3     19
5     18
8     14
9     12
6     11
7      8
Name: PaymentsOptionsAvailability, dtype: int64
---------------------------------------
0    107
1     82
3     36
2     15
4     15
5     14
Name: Speedyorderdelivery, dtype: int64
---------------------------------------
0     71
1     54
2     25
5     24
10    18
7     15
8     15
9     15
3     14
4     11
6      7
Name: CustomerInformationPrivacy, dtype: int64
---------------------------------------
0     51
4     42
8     33
5     25
1     24
6     20
7     19
9     15
10    15
2     14
3     11
Name: CustomerFinancialInformationSecurity, dtype: int64
---------------------------------------
0     76
5     36
```

Correlation values with our target variable

Age	0.309575
Channel	0.215928
City	0.173871
ShoppingAndRoles	0.167532
WideVarietyOfProducts	0.144087
GuarnteePrivacy	0.135745
NetBenefit	0.112002

Convenienceofpatronizingonlineretailer	0.105267
Speedyorderdelivery	0.101356
ShoppingAndGratification	0.079476
NavigatingWebsite	0.078981
AbandonProduct	0.078278
Gender	0.077876
Empathy	0.075615
MobileScreenSize	0.062622
LogInTime	0.062336
DisplayQualityInformation	0.056794
Trust	0.049143
Enjoyment	0.043482
LoyalityPrograms	0.036268
InternetAccess	0.017990
DeviceOS	0.015316
FastLoadingWebsiteSpeed	0.015289
UserFriendlyInterface	0.013640
YearsOfOnlineShopping	0.013315
WebsiteDesignChange	0.007841
Flexibility	0.004376
ShoppingAndSocialStatus	-0.004694
PageLoadingTime	-0.005474
DeviceBrowser	-0.012791
MonetarySavings	-0.021391
WhyAbandon	-0.026906
LimitedModeOfPayment	-0.030008
DeviceUsed	-0.051741
ConvinientPaymentMethods	-0.052107
PresenceOfOnlineAssistance	-0.058209

OnlineRetailersShoppedFrom	-0.068387
ExploreTime	-0.071392
ReachStore	-0.083088
ValueForMoneySpent	-0.087458
DisplayGraphicsTime	-0.096575
PaymentsOptionsAvailability	-0.097651
RelevantInfoOnListedProducts	-0.100875
ProductComparison	-0.107277
CompletePurchaseQuickness	-0.107802
WebsiteEfficieny	-0.124076
FrequentDisruptionInMovingFromPageToPage	-0.127148
LongerDeliveryPeriod	-0.130651
WebsiteContent	-0.139930
ProvisionOfReleventInformation	-0.144277
WebSiteReliability	-0.146351
PurchaseDecision	-0.147507
IndianOnlineRetailerToRecommend	-0.152028
BenifitsAndDiscounts	-0.159635
Easytousewebsite	-0.160976
CustomerInformationPrivacy	-0.164571
PerceivedTrustworthiness	-0.169235
Varietyofproductonoffer	-0.177938
CommunicationChannelsAvailability	-0.190372
CustomerFinancialInformationSecurity	-0.192403
Completedescriptionofproducts	-0.195169
PaymentOption	-0.199719
UserSatisfactionAndTrust	-0.206904
PriceLateDeclaration	-0.209112
Visualappealingwebpagelayout	-0.223378

goodqualitywebsitesatisfaction -0.226581

Speed -0.261066

Pincode -0.304554

ReturnAndReplacementPolicy -0.310908

ShoppingAndAdventure -0.318657

See screenshot below for all correlations:

```
]: # To display all value of rows
   pd.set_option("display.max_rows", None)
   # let's calculate the featues Correlation values with our target variable.
   df.drop('Target',axis=1).corrwith(df.Target ).sort_values(ascending=False)
```

```
]: Age                                       0.309575
   Channel                                   0.215928
   City                                      0.173871
   ShoppingAndRoles                          0.167532
   WideVarietyOfProducts                     0.144087
   GuarnteePrivacy                           0.135745
   NetBenefit                                0.112002
   Convenienceofpatronizingonlineretailer    0.105267
   Speedyorderdelivery                       0.101356
   ShoppingAndGratification                  0.079476
   NavigatingWebsite                         0.078981
   AbandonProduct                            0.078278
   Gender                                    0.077876
   Empathy                                   0.075615
   MobileScreenSize                          0.062622
   LogInTime                                 0.062336
   DisplayQualityInformation                 0.056794
   Trust                                     0.049143
   Enjoyment                                 0.043482
   LoyalityPrograms                          0.036268
   InternetAccess                            0.017990
   DeviceOS                                  0.015316
   FastLoadingWebsiteSpeed                   0.015289
   UserFriendlyInterface                     0.013640
   YearsOfOnlineShopping                     0.013315
   WebsiteDesignChange                       0.007841
   Flexibility                               0.004376
   ShoppingAndSocialStatus                  -0.004694
   PageLoadingTime                          -0.005474
   DeviceBrowser                            -0.012791
   MonetarySavings                          -0.021391
   WhyAbandon                               -0.026906
   LimitedModeOfPayment                     -0.030008
   DeviceUsed                               -0.051741
   ConvinientPaymentMethods                 -0.052107
   PresenceOfOnlineAssistance               -0.058209
   OnlineRetailersShoppedFrom               -0.068387
```

```
DeviceUsed                                  -0.051741
ConvinientPaymentMethods                    -0.052107
PresenceOfOnlineAssistance                  -0.058209
OnlineRetailersShoppedFrom                  -0.068387
ExploreTime                                 -0.071392
ReachStore                                  -0.083088
ValueForMoneySpent                          -0.087458
DisplayGraphicsTime                         -0.096575
PaymentsOptionsAvailability                 -0.097651
RelevantInfoOnListedProducts                -0.100875
ProductComparison                           -0.107277
CompletePurchaseQuickness                   -0.107802
WebsiteEfficieny                            -0.124076
FrequentDisruptionInMovingFromPageToPage    -0.127148
LongerDeliveryPeriod                        -0.130651
WebsiteContent                              -0.139930
ProvisionOfReleventInformation              -0.144277
WebSiteReliability                          -0.146351
PurchaseDecision                            -0.147507
IndianOnlineRetailerToRecommend             -0.152028
BenifitsAndDiscounts                        -0.159635
Easytousewebsite                            -0.160976
CustomerInformationPrivacy                  -0.164571
PerceivedTrustworthiness                    -0.169235
Varietyofproductonoffer                     -0.177938
CommunicationChannelsAvailability           -0.190372
CustomerFinancialInformationSecurity        -0.192403
Completedescriptionofproducts               -0.195169
PaymentOption                               -0.199719
UserSatisfactionAndTrust                    -0.206904
PriceLateDeclaration                        -0.209112
Visualappealingwebpagelayout                -0.223378
goodqualitywebsitesatisfaction              -0.226581
Speed                                       -0.261066
Pincode                                     -0.304554
ReturnAndReplacementPolicy                  -0.310908
ShoppingAndAdventure                        -0.318657
dtype: float64
```

Observations

- Female customer often prefer to purchase more as compared to Male customer
- People between 21 to 50 years of age tends shops more.
- Dataset appears to have approximately 67% of Female respondents & 32 % of Male
- Delhi has maximum count which means in Delhi the online purchase is high.
- Mobile screen size has maximum count for others,and 2nd most count is for 5.5 which is the most used screen size.
- Customers with OS windows/windows mobile are more in number,as it's the most common OS.
- Most of the customers uses chrome browser, which seems to be best in market and easy to use.

- Most of the customers used Search engine to reach there favorite online store for the first time , as search engine shows variety of similar product from n-number of sites.
- After first visit most of the customers are reaching the online store through Search engine, Via application and Direct Url.
- Most of the customers are exploring the e-retail store more than 6 mins before making purchase decision (checking all info and deciding which one to buy usually take time).
- Maximum customers are using credit/debit cards for there payment(most convenient and easy to use).
- Amazon and flipkart is performing best in all given parameter

CONCLUSION

After the detailed analysis of problem I found the following conclusions:

- Frequency of Females shoping is high so making them satisfied will help the sellers to get more business.
- Loyal customers prefer buying and tend to spend more money on shopping in your store. Statistics show that engaged consumers purchase more frequently. It is necessary to hear customer feedback because most of them are valuable feedbacks.
- Sometimes customer feedback is the best marketing strategy. They are frequent customers so they will know which areas of your business may well be improved. If their feedback is approved, they will extremely excite and support your company with their best ability.
- It found that Amazon and Flipkart are standing best out in the market by using ethical, reasonable business strategies
- The repeat purchase intention (loyalty) positively, Structural equation model has been presented on the primary data collected from the Indian online shoppers.
- Here as an conclusion part I found that using dead old strategies for retailers will effect customer retention.