
# Customer Retention Analysis

Submitted by:

Yatika Taneja

# 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.

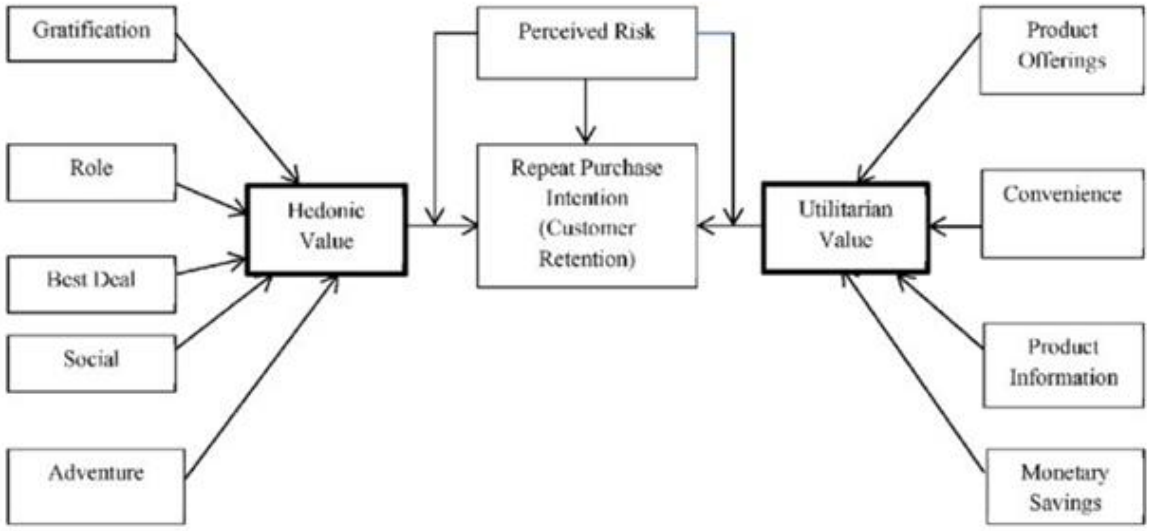
# INTRODUCTION

## 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 (MacInnis 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).

# Analytical Problem Framing

## Data Sources

The sample data is provided to us from our client database. The dataset has 269 observations and 71 features.

Dataset:

| | Gender | Age | City | Pincode | YearsOfOnlineShopping | PurchaseInPast1Year | InternetAccess | DeviceUsed | MobileScreenSize | DeviceOS | ... | Log |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | Male | 31-40 years | Delhi | 110009 | Above 4 years | 31-40 times | Dial-up | Desktop | Others | Window/windows Mobile | ... | Am |
| 1 | Female | 21-30 years | Delhi | 110030 | Above 4 years | 41 times and above | Wi-Fi | Smartphone | 4.7 inches | IOS/Mac | ... | Ama Flipk |
| 2 | Female | 21-30 years | Greater Noida | 201308 | 3-4 years | 41 times and above | Mobile Internet | Smartphone | 5.5 inches | Android | ... | Myn |
| 3 | Male | 21-30 years | Karnal | 132001 | 3-4 years | Less than 10 times | Mobile Internet | Smartphone | 5.5 inches | IOS/Mac | ... | Snapde |
| 4 | Female | 21-30 years | Bangalore | 530068 | 2-3 years | 11-20 times | Wi-Fi | Smartphone | 4.7 inches | IOS/Mac | ... | Flipka Pay |

5 rows × 71 columns

# Data Preprocessing

Replacing columns

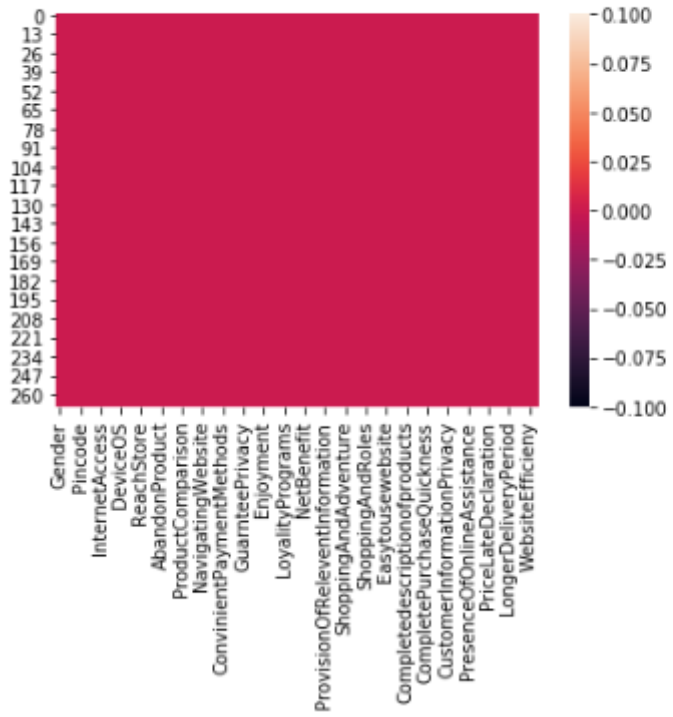
```
df1.columns
```

```
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

```
sns.heatmap(df.isnull())
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x116b9b10a88>
```




No null values present

# Data Info And Description

```
 #   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
```

```
df1.describe()
```

| | 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 |

```
df1.describe(include='object')
```

| | Gender | Age | City | YearsOfOnlineShopping | PurchaseInPast1Year | InternetAccess | DeviceUsed | MobileScreenSize | DeviceOS | DeviceBrowser | ... |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 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 | ... |
| freq | 181 | 81 | 58 | 98 | 114 | 142 | 141 | 134 | 122 | 216 | ... |

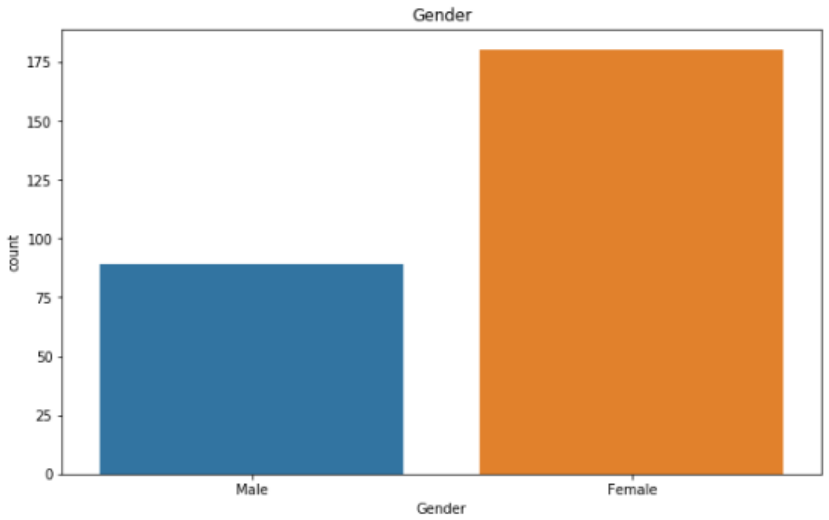
4 rows × 70 columns

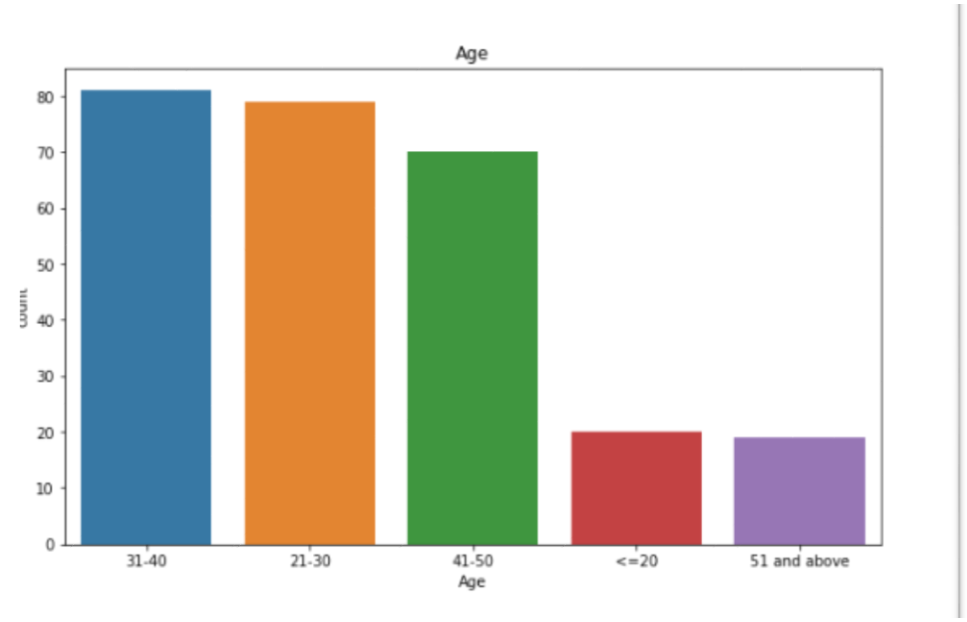
# Data Visualization

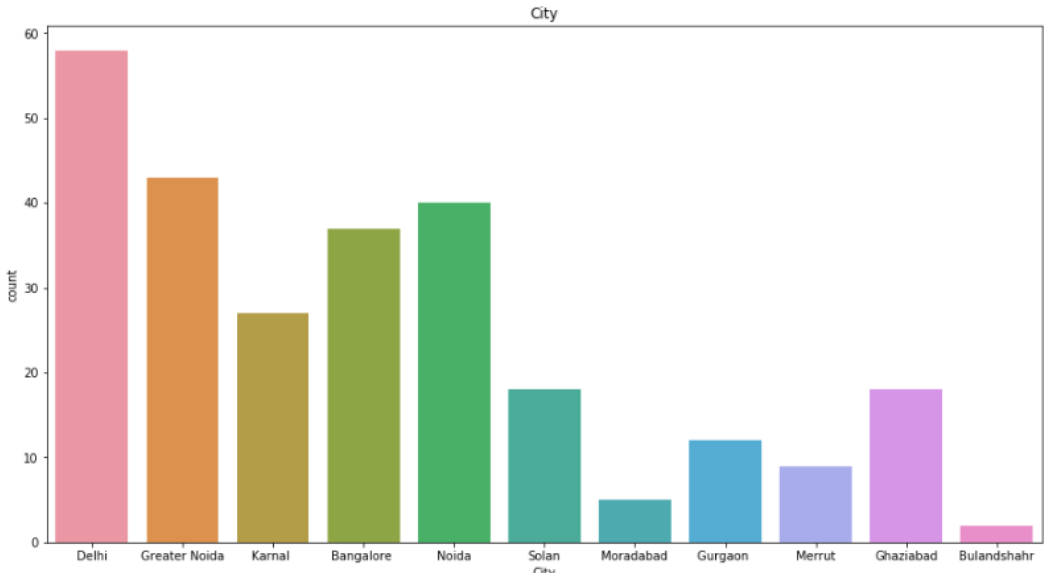
## 1. CountPlot

```
plt.figure(figsize=(10,6))
plt.title("Gender")
sns.countplot(x='Gender',data = df1)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x116ba14cf88>
```






**FOR ALL VARIABLES:**

```
categorical_feats = df1.dtypes[df1.dtypes == "object"].index
print("Number of Categorical features: ", len(categorical_feats))
```

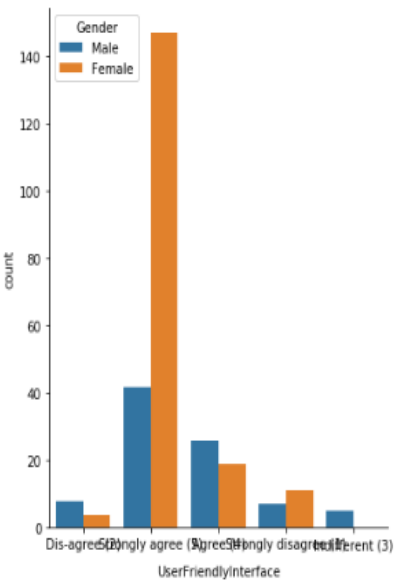
```
Number of Categorical features:  70
```

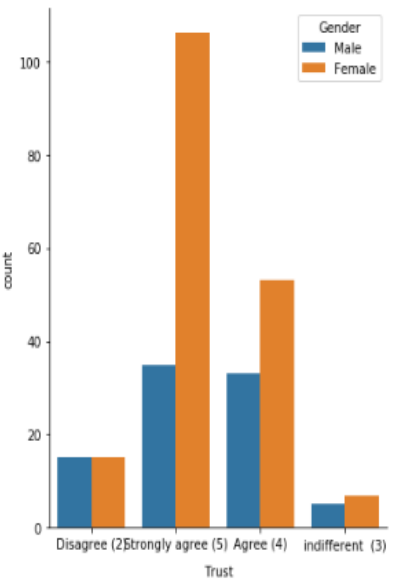
```
li_cat_feats = list(categorical_feats)
nr_rows = 35
nr_cols = 2

fig, axs = plt.subplots(nr_rows, nr_cols, figsize=(nr_cols*8,nr_rows*10))

for r in range(0,nr_rows):
    for c in range(0,nr_cols):
        i = r*nr_cols+c
        if i < len(li_cat_feats):
            ax1=sns.countplot(x=li_cat_feats[i], data=df1, ax = axs[r][c])
            ax1.set_xlabel(li_cat_feats[i], labelpad=10)

sns.despine()
```

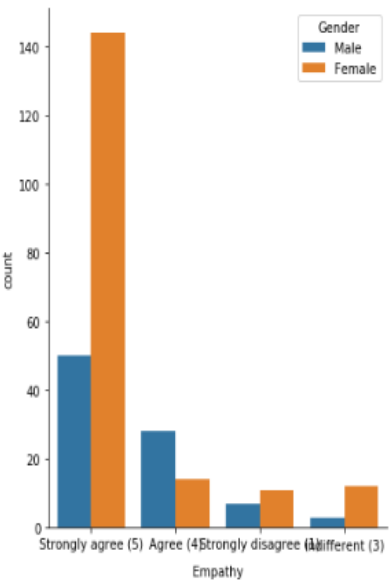
## 2. Relationship of Hedonic Values with Gender and Age

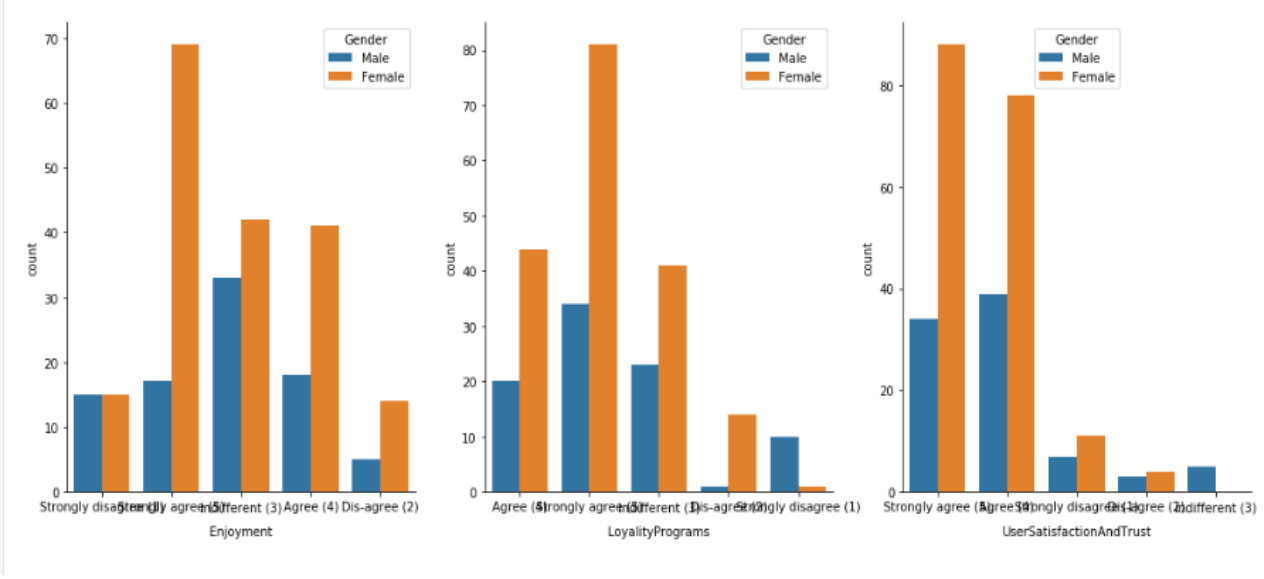
**Hedonic Variables:**

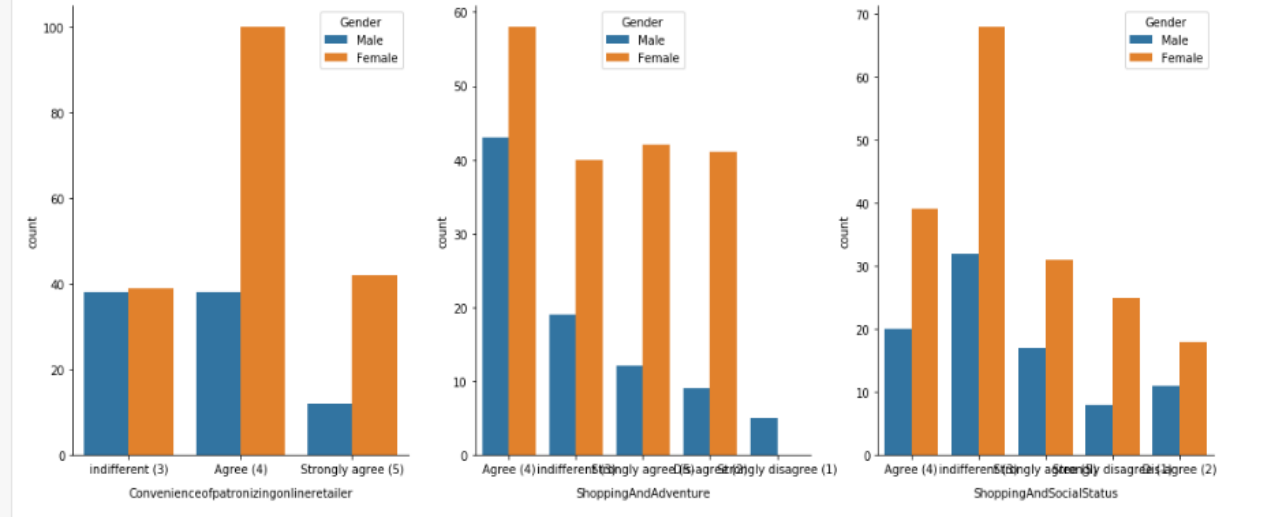
- 'UserFriendlyInterface
- 'Trust
- Empathy
- Enjoyment
- LoyalityPrograms
- 'UserSatisfactionAndTrust
- 'Convenienceofpatronizingonlineretailer
- ShoppingAndAdventure'
- ShoppingAndSocialStatus
- ShoppingAndGratification'
- ShoppingAndRoles
- ValueForMoneySpent

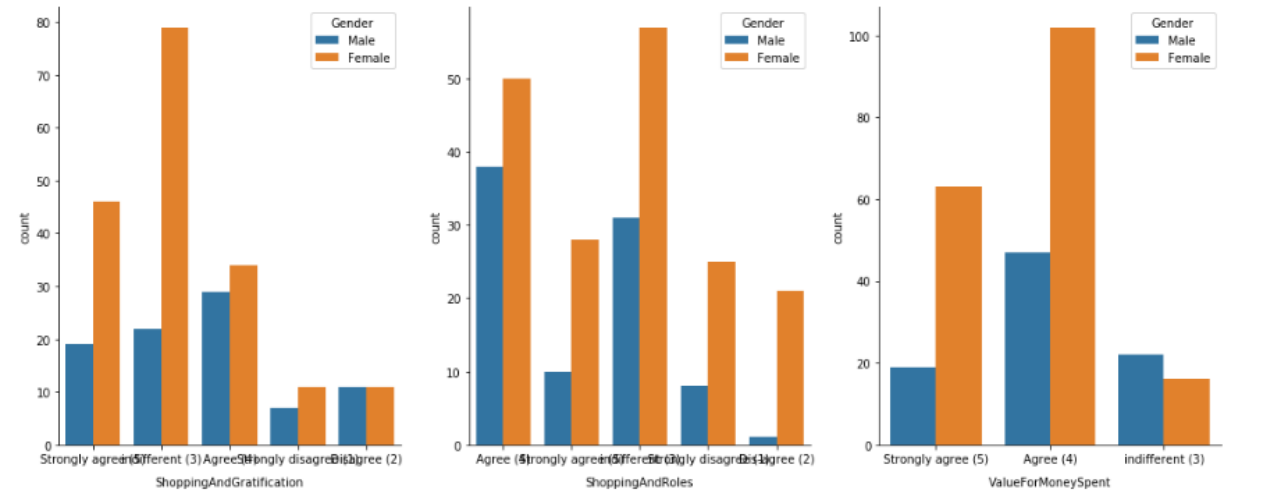


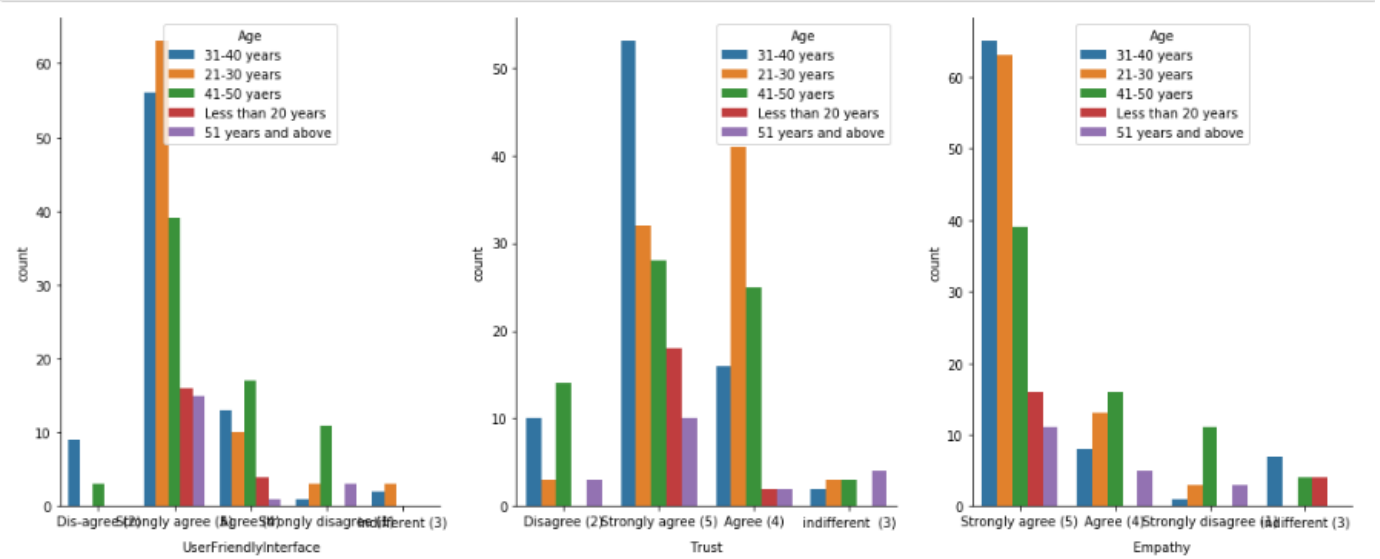



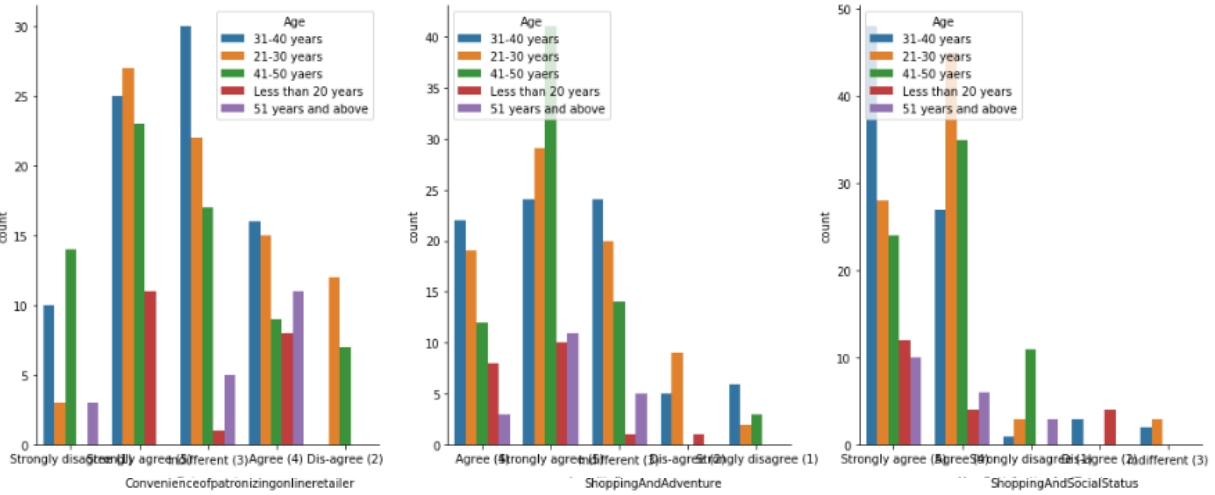



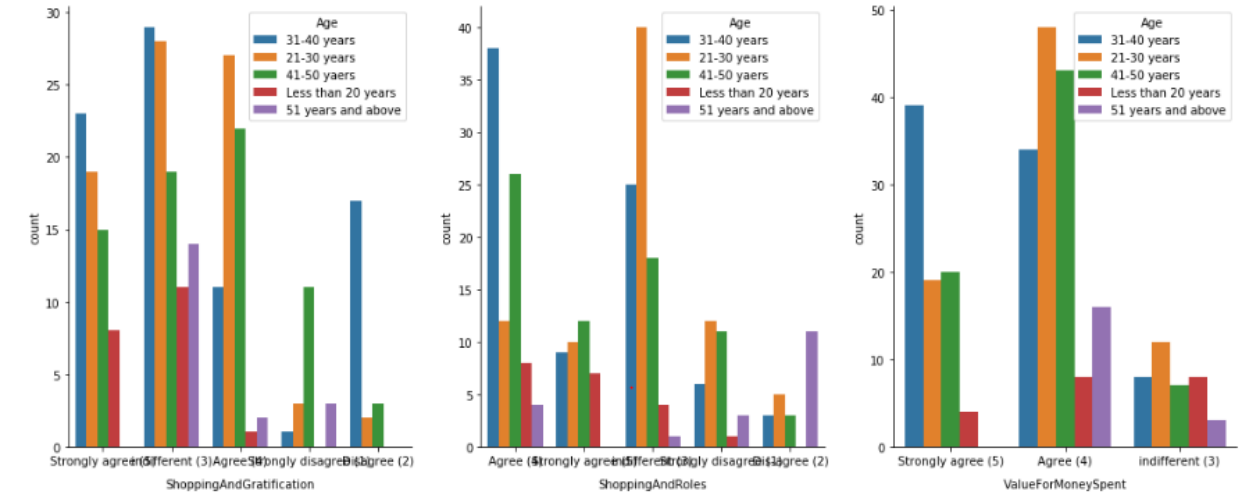

Gender
Male
Female
count
Enjoyment
LoyalityPrograms
UserSatisfactionAndTrust


Gender
Male
Female
count
indifferent (3)
Agree (4)
Strongly agree (5)
Convenienceofpatronizingonlineretailer
Agree (4)
Strongly agree (5)
ShoppingAndAdventure
Agree (4)
ShoppingAndSocialStatus


Gender
Male
Female
count
Strongly agree (5)
indifferent (3)
Agree (4)
ShoppingAndGratification
Agree (4)
Strongly agree (5)
ShoppingAndRoles
Strongly agree (5)
Agree (4)
indifferent (3)
ValueForMoneySpent



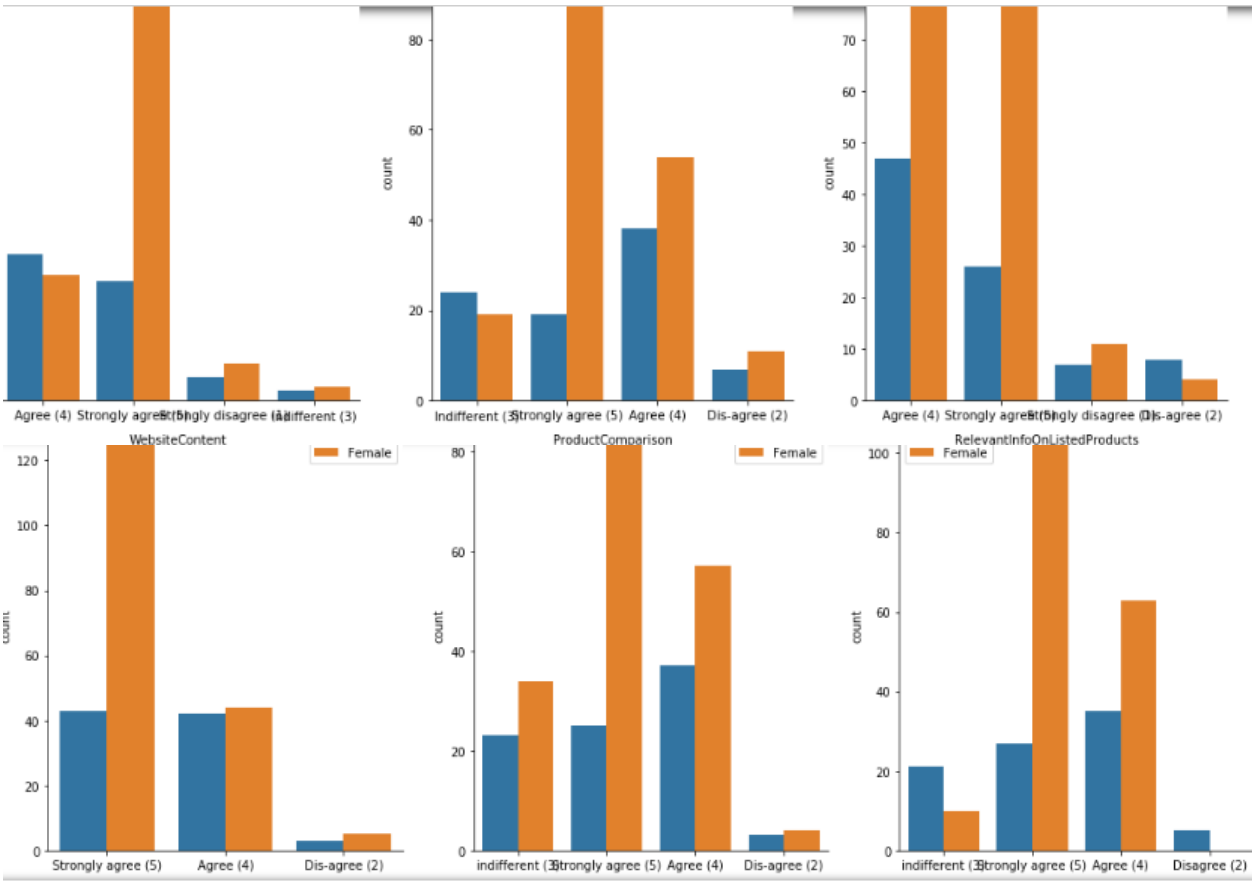




## 3. Relationship of Utilitarian value with Age and Gender

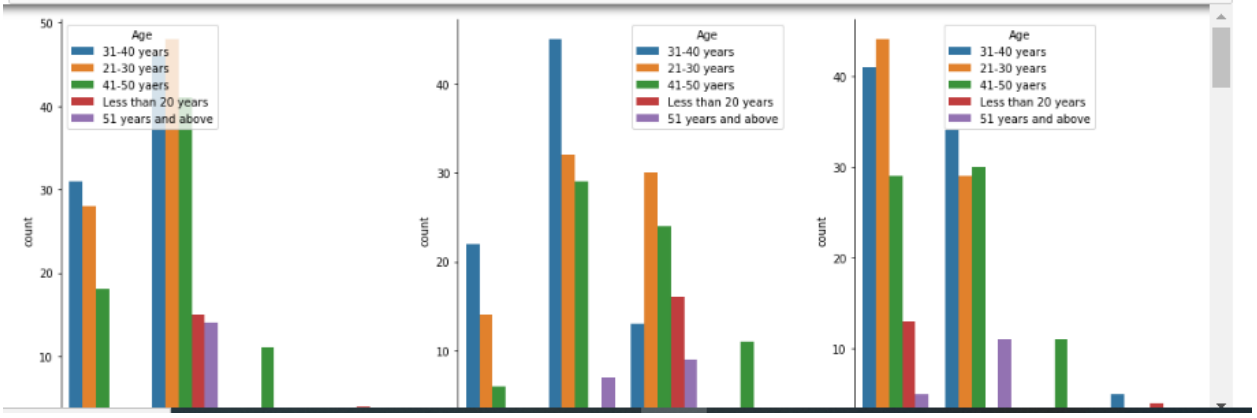
### Utilitarian Variables:

- WebsiteContent
- ProductComparison
- RelevantInfoOnListedProducts
- NavigatingWebsite
- ConvinientPaymentMethods
- GuarnteePrivacy
- CommunicationChannelsAvailability
- Flexibility
- DisplayQualityInformation

- goodqualitywebsitesatisfaction
- WideVarietyOfProducts
- ProvisionOfReleventInformation
- MonetarySavings




```
nr_rows = 5
nr_cols = 3

fig, axs = plt.subplots(nr_rows, nr_cols, figsize=(nr_cols*6,nr_rows*8))

for r in range(0,nr_rows):
    for c in range(0,nr_cols):
        i = r*nr_cols+c
        if i < len(Utilitarian_value):
            ax1=sns.countplot(x=Utilitarian_value[i], hue='Age',data=df1, ax = axs[r][c])
            ax1.set_xlabel(Utilitarian_value[i], labelpad=10)

sns.despine()
```



# CONCLUSION

**Results of the data analysis performed indicates the e-retail success factors, which are very much critical for customer satisfaction. It is concluded that combination of both utilitarian and hedonistic value are needed to affect the repeat purchase intention positively**