EXECUTIVE SUMMARY

Key Operations executed: Exploratory Data Analysis(EDA) and Data Preprocessing(Univariate,Bivariate,Multivariate)

Research Question

E-retail factors for customer activation and retention

Data description

1. From the dataset we have 269 records and 72 features
2. We have a problem of dimensionality(curse of dimensionality)
3. We have too many features and there is a need to perform feature selection to determine customer retention
4. We will definitely drop the columns where the names of the companies are mentioned
5. We are only paying major attention to the factors the lead to customer retention or customer repurchase
6. The dataset set only contains categorical data and no continous data.
7. The target data is categorical meaning that if we are to use a machine learning algorithm it must definitely be a classification algorithm

PROJECT CONTENTS

PROBLEM STATEMENT(customer activation and

retention)

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 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.

Domain Knowledge

This is an important stage in any machine learning process. It involves brainstorming and coming up with as many hypothesis as possible about what could affect the target variable. It facilitates in exploring the data at hand more efficiently and effectively. Domain Knowledge should be done before seeing the data or else we will end up with biased hypotheses. Below are some anticipated assertions on the problem statement.

- There is increased online purchase during weekends and holidays
- online purchases are higher during weekends.
- People love convenience and would use technology if possible
- increased online purchase during festive periods.
- Customer experience,system performance plays a major role in customer activation and retention

Step 1 - ATTRIBUTES OF VARIABLES

1. Gender 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 1 year?
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?
16. How frequently do you abandon (selecting an items and leaving without making payment) your shopping cart?
17. Why did you abandon the “Bag”, “Shopping Cart”?
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

Step 2 - IMPORT ALL NECESSARY LIBRARIES, LOAD DATA SET, DATA DESCRIPTIVE ANALYSIS AND OBSERVATIONS

2.1 - Import all necessary libraries

In [1]:
```
import pandas as pd
import numpy as np
from sklearn.preprocessing import StandardScaler,power_transform,PowerTransformer,LabelE
from sklearn.model_selection import train_test_split,cross_val_score,GridSearchCV, Rando
from sklearn.metrics import accuracy_score, confusion_matrix, roc_curve, roc_auc_score,
from scipy.stats import zscore
from collections import Counter
from sklearn import metrics
import seaborn as sns
import matplotlib.pyplot as plt
import pickle


import warnings
warnings.filterwarnings('ignore')
```

2.2 - Load Dataset

In [2]:
```
#Load dataset
df = pd.read_excel(r'C:\Users\USER\Documents\Fliprobo Technologies\customer_retention_da
df.head()
```

Out[2]:

1Gender of respondent	2 How old are you?	3 Which city do you shop online from?	4 What is the Pin Code of where	5 Since How Long You are Shopping Online ?	6 How many times you have made an online purchase	7 How do you access the internet while	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	10 What is the operating system (OS) o your device' \t\t\t\

				you shop online from?		in the past 1 year?	shopping on-line?			
0	Male	31-40 years	Delhi	110009	Above 4 years	31-40 times	Dial-up	Desktop	Others	Window/windows Mobile
1	Female	21-30 years	Delhi	110030	Above 4 years	41 times and above	Wi-Fi	Smartphone	4.7 inches	IOS/Mac
2	Female	21-30 years	Greater Noida	201308	3-4 years	41 times and above	Mobile Internet	Smartphone	5.5 inches	Androic
3	Male	21-30 years	Karnal	132001	3-4 years	Less than 10 times	Mobile Internet	Smartphone	5.5 inches	IOS/Mac
4	Female	21-30 years	Bangalore	530068	2-3 years	11-20 times	Wi-Fi	Smartphone	4.7 inches	IOS/Mac

5 rows × 71 columns

In [3]:
```
#checking the first five rows
df.head()
```

Out[3]:

	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 1 year?	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\
0	Male	31-40 years	Delhi	110009	Above 4 years	31-40 times	Dial-up	Desktop	Others	Window/windows Mobile
1	Female	21-30 years	Delhi	110030	Above 4 years	41 times and above	Wi-Fi	Smartphone	4.7 inches	IOS/Mac
2	Female	21-30 years	Greater Noida	201308	3-4 years	41 times and above	Mobile Internet	Smartphone	5.5 inches	Androic
3	Male	21-30 years	Karnal	132001	3-4 years	Less than 10 times	Mobile Internet	Smartphone	5.5 inches	IOS/Mac
4	Female	21-30 years	Bangalore	530068	2-3 years	11-20 times	Wi-Fi	Smartphone	4.7 inches	IOS/Mac

5 rows × 71 columns

In [4]:
```
#checking the last five rows
```

```
df.tail()
```

Out[4]:

	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 1 year?	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 operat system (OS your devi \t\t
264	Female	21-30 years	Solan	173212	1-2 years	Less than 10 times	Mobile Internet	Smartphone	5.5 inches	Andı
265	Female	31-40 years	Ghaziabad	201008	1-2 years	31-40 times	Mobile Internet	Smartphone	Others	Andı
266	Female	41-50 yaers	Bangalore	560010	2-3 years	Less than 10 times	Mobile internet	Laptop	Others	Window/wind Mo
267	Female	Less than 20 years	Solan	173229	2-3 years	Less than 10 times	Wi-Fi	Smartphone	5.5 inches	Andı
268	Female	41-50 yaers	Ghaziabad	201009	2-3 years	31-40 times	Mobile Internet	Smartphone	5.5 inches	Andı

5 rows × 71 columns

In [5]:
```
#picking 4 random samples
df.sample(4)
```

Out[5]:

	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 1 year?	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
206	Male	41-50 yaers	Delhi	110008	2-3 years	31-40 times	Mobile internet	Smartphone	5.5 inches	Android
195	Male	31-40 years	Noida	201303	3-4 years	21-30 times	Mobile internet	Smartphone	5.5 inches	Window/windows Mobile
87	Male	41-50 yaers	Delhi	110044	2-3 years	Less than 10 times	Mobile internet	Smartphone	5.5 inches	IOS/Mac
77	Male	Less than	Delhi	110042	3-4 years	42 times and	Mobile internet	Smartphone	4.7 inches	IOS/Mac

20 years	above

4 rows × 71 columns

2.3 - Data Descriptive Analysis

In [6]:
```
#checking the data shape
df.shape
```

Out[6]:
```
(269, 71)
```

In [7]:
```
#checking for nature of datatypes
df.dtypes
```

Out[7]:
```
1Gender of respondent                                                 object
2 How old are you?                                                    object
3 Which city do you shop online from?                                 object
4 What is the Pin Code of where you shop online from?                  int64
5 Since How Long You are Shopping Online ?                            object
                                                                       ...
Longer delivery period                                                object
Change in website/Application design                                  object
Frequent disruption when moving from one page to another              object
Website is as efficient as before                                     object
Which of the Indian online retailer would you recommend to a friend?    object
Length: 71, dtype: object
```

In [8]:
```
#Checking data info
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 269 entries, 0 to 268
Data columns (total 71 columns):
 #   Column

                                     Non-Null Count  Dtype
---  ------

                                     --------------  -----
 0   1Gender of respondent

                                     269 non-null    object
 1   2 How old are you?

                                     269 non-null    object
 2   3 Which city do you shop online from?

                                     269 non-null    object
 3   4 What is the Pin Code of where you shop online from?

                                     269 non-null    int64
 4   5 Since How Long You are Shopping Online ?

                                     269 non-null    object
 5   6 How many times you have made an online purchase in the past 1 year?

                                     269 non-null    object
 6   7 How do you access the internet while shopping on-line?

                                     269 non-null    object
 7   8 Which device do you use to access the online shopping?
```

269 non-null object
8 9 What is the screen size of your mobile device?

269 non-null object
9 10 What is the operating system (OS) of your device?

269 non-null object
10 11 What browser do you run on your device to access the website?

269 non-null object
11 12 Which channel did you follow to arrive at your favorite online store for the first time?

269 non-null object
12 13 After first visit, how do you reach the online retail store?

269 non-null object
13 14 How much time do you explore the e- retail store before making a purchase decision?

269 non-null object
14 15 What is your preferred payment Option?

269 non-null object
15 16 How frequently do you abandon (selecting an items and leaving without making payment) your shopping cart?

269 non-null object
16 17 Why did you abandon the “Bag”, “Shopping Cart”?

269 non-null object
17 18 The content on the website must be easy to read and understand

269 non-null object
18 19 Information on similar product to the one highlighted is important for product comparison

269 non-null object
19 20 Complete information on listed seller and product being offered is important for purchase decision.

269 non-null object
20 21 All relevant information on listed products must be stated clearly

269 non-null object
21 22 Ease of navigation in website

269 non-null object
22 23 Loading and processing speed

269 non-null object
23 24 User friendly Interface of the website

269 non-null object
24 25 Convenient Payment methods

269 non-null object
25 26 Trust that the online retail store will fulfill its part of the transaction at the stipulated time

269 non-null object
26 27 Empathy (readiness to assist with queries) towards the customers

269 non-null object
27 28 Being able to guarantee the privacy of the customer

269 non-null object
28 29 Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.)

```
                                    269 non-null    object
 29  30 Online shopping gives monetary benefit and discounts

                                    269 non-null    object
 30  31 Enjoyment is derived from shopping online

                                    269 non-null    object
 31  32 Shopping online is convenient and flexible

                                    269 non-null    object
 32  33 Return and replacement policy of the e-tailer is important for purchase decision

                                    269 non-null    object
 33  34 Gaining access to loyalty programs is a benefit of shopping online

                                    269 non-null    object
 34  35 Displaying quality Information on the website improves satisfaction of customers

                                    269 non-null    object
 35  36 User derive satisfaction while shopping on a good quality website or application

                                    269 non-null    object
 36  37 Net Benefit derived from shopping online can lead to users satisfaction

                                    269 non-null    object
 37  38 User satisfaction cannot exist without trust

                                    269 non-null    object
 38  39 Offering a wide variety of listed product in several category

                                    269 non-null    object
 39  40 Provision of complete and relevant product information

                                    269 non-null    object
 40  41 Monetary savings

                                    269 non-null    object
 41  42 The Convenience of patronizing the online retailer

                                    269 non-null    object
 42  43 Shopping on the website gives you the sense of adventure

                                    269 non-null    object
 43  44 Shopping on your preferred e-tailer enhances your social status

                                    269 non-null    object
 44  45 You feel gratification shopping on your favorite e-tailer

                                    269 non-null    object
 45  46 Shopping on the website helps you fulfill certain roles

                                    269 non-null    object
 46  47 Getting value for money spent

                                    269 non-null    object
 47  From the following, tick any (or all) of the online retailers you have shopped fro
m;
                                     269 non-null    object
 48  Easy to use website or application

                                    269 non-null    object
 49  Visual appealing web-page layout

                                    269 non-null    object
 50  Wild variety of product on offer
```

```
                                    269 non-null    object
 51  Complete, relevant description information of products

                                    269 non-null    object
 52  Fast loading website speed of website and application

                                    269 non-null    object
 53  Reliability of the website or application

                                    269 non-null    object
 54  Quickness to complete purchase

                                    269 non-null    object
 55  Availability of several payment options

                                    269 non-null    object
 56  Speedy order delivery

                                    269 non-null    object
 57  Privacy of customers' information

                                    269 non-null    object
 58  Security of customer financial information

                                    269 non-null    object
 59  Perceived Trustworthiness

                                    269 non-null    object
 60  Presence of online assistance through multi-channel

                                    269 non-null    object
 61  Longer time to get logged in (promotion, sales period)

                                    269 non-null    object
 62  Longer time in displaying graphics and photos (promotion, sales period)

                                    269 non-null    object
 63  Late declaration of price (promotion, sales period)

                                    269 non-null    object
 64  Longer page loading time (promotion, sales period)

                                    269 non-null    object
 65  Limited mode of payment on most products (promotion, sales period)

                                    269 non-null    object
 66  Longer delivery period

                                    269 non-null    object
 67  Change in website/Application design

                                    269 non-null    object
 68  Frequent disruption when moving from one page to another

                                    269 non-null    object
 69  Website is as efficient as before

                                    269 non-null    object
 70  Which of the Indian online retailer would you recommend to a friend?

                                    269 non-null    object
dtypes: int64(1), object(70)
memory usage: 149.3+ KB
```

In [9]:
```
#Checking data description for all numerical columns
df.describe()
```

Out[9]:

	4 What is the Pin Code of where you shop online from?
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 [10]:
```
#Checking data description for all non-numerical columns
numerics = ['int8','int16','int32','int64','float16','float32','float64']#enlisiting and
categorical_columns=[]#initializing an emty list
features = df.columns.values.tolist()#Pulling all the features(columns) of the dataset a
for col in features:#col is the iterating variable
    if df[col].dtype in numerics:
        continue
    categorical_columns.append(col)
df[categorical_columns].describe()
```

Out[10]:

	1Gender of respondent	2 How old are you?	3 Which city do 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 1 year?	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	10 What is the operating system (OS) of your device? \t\t\t\t	11 ' brc d ru d a web
count	269	269	269	269	269	269	269	269	269	
unique	2	5	11	5	6	4	4	4	3	
top	Female	31-40 years	Delhi	Above 4 years	Less than 10 times	Mobile internet	Smartphone	Others	Window/windows Mobile	G ch
freq	181	81	58	98	114	142	141	134	122	

4 rows × 70 columns

In [11]:
```
#lets check the columns
df.columns
```

Out[11]:
```
Index(['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 1 year?',
       '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 dec
ision?                                                                               ',
       '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 produ
ct 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 a
t 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 decis
ion',
       '34 Gaining access to loyalty programs is a benefit of shopping online',
       '35 Displaying quality Information on the website improves satisfaction of custom
ers',
       '36 User derive satisfaction while shopping on a good quality website or applicat
ion',
       '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 f
rom;                                                                                 ',
       '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?'],
      dtype='object')
```

In [12]:
```
#Checking for null values
df.isnull().sum()
```

Out[12]:
```
1Gender of respondent                                                0
2 How old are you?                                                   0
3 Which city do you shop online from?                                0
4 What is the Pin Code of where you shop online from?                0
5 Since How Long You are Shopping Online ?                           0
                                                                    ..
Longer delivery period                                               0
Change in website/Application design                                 0
Frequent disruption when moving from one page to another             0
Website is as efficient as before                                    0
Which of the Indian online retailer would you recommend to a friend?    0
Length: 71, dtype: int64
```

In [13]:
```
#Checking for null values
df.isnull().sum().sum()
```

Out[13]:
```
0
```

In [14]:
```
df.dtypes
```

Out[14]:
```
1Gender of respondent                                                object
2 How old are you?                                                   object
3 Which city do you shop online from?                                object
4 What is the Pin Code of where you shop online from?                 int64
5 Since How Long You are Shopping Online ?                           object
                                                                      ...
Longer delivery period                                               object
Change in website/Application design                                 object
Frequent disruption when moving from one page to another             object
Website is as efficient as before                                    object
Which of the Indian online retailer would you recommend to a friend?    object
Length: 71, dtype: object
```

In [15]:
```
#Checking the content of each Categorical columns
for column in df:
    if df[column].dtypes=='object':
        print(f"The Value Counts in {column} - ",Counter(df[column]))
```

```
The Value Counts in 1Gender of respondent -  Counter({'Female': 181, 'Male': 88})
The Value Counts in 2 How old are you?  -  Counter({'31-40 years': 81, '21-30 years': 7
9, '41-50 yaers': 70, 'Less than 20 years': 20, '51 years and above': 19})
The Value Counts in 3 Which city do you shop online from? -  Counter({'Delhi': 58, 'Grea
ter Noida': 43, 'Noida': 40, 'Bangalore ': 37, 'Karnal ': 27, 'Solan': 18, 'Ghaziabad':
18, 'Gurgaon ': 12, 'Merrut': 9, 'Moradabad': 5, 'Bulandshahr': 2})
The Value Counts in 5 Since How Long You are Shopping Online ? -  Counter({'Above 4 year
s': 98, '2-3 years': 65, '3-4 years': 47, 'Less than 1 year': 43, '1-2 years': 16})
The Value Counts in 6 How many times you have made an online purchase in the past 1 yea
r? -  Counter({'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})
The Value Counts in 7 How do you access the internet while shopping on-line? -  Counter
({'Mobile internet': 142, 'Wi-Fi': 76, 'Mobile Internet': 47, 'Dial-up': 4})
```

The Value Counts in 8 Which device do you use to access the online shopping? - Counter ({'Smartphone': 141, 'Laptop': 86, 'Desktop': 30, 'Tablet': 12})
The Value Counts in 9 What is the screen size of your mobile device?
- Counter({'Others': 1 34, '5.5 inches': 99, '4.7 inches': 29, '5 inches': 7})
The Value Counts in 10 What is the operating system (OS) of your device?
- Counter({'Window/windows M obile': 122, 'Android': 85, 'IOS/Mac': 62})
The Value Counts in 11 What browser do you run on your device to access the website?
- Counter ({'Google chrome': 216, 'Safari': 40, 'Opera': 8, 'Mozilla Firefox': 5})
The Value Counts in 12 Which channel did you follow to arrive at your favorite online st ore for the first time?
- Counter({'Search Engine': 230, 'Content Marketing': 20, 'Display Adverts': 19})
The Value Counts in 13 After first visit, how do you reach the online retail store?
- Counter ({'Search Engine': 87, 'Via application': 86, 'Direct URL': 70, 'E-mail': 18, 'Social Me dia': 8})
The Value Counts in 14 How much time do you explore the e- retail store before making a purchase decision? -
Counter({'more than 15 mins': 123, '6-10 mins': 71, '11-15 mins': 46, 'Less than 1 min': 15, '1-5 mins': 14})
The Value Counts in 15 What is your preferred payment Option?

- Counter({'Credit/Debit cards': 148, 'Cash on delivery (CoD)': 76, 'E-wallets (Paytm, Freecharge etc.)': 45})
The Value Counts in 16 How frequently do you abandon (selecting an items and leaving wit hout making payment) your shopping cart?
- Counter({'Sometimes': 17 1, 'Never': 48, 'Frequently': 35, 'Very frequently': 15})
The Value Counts in 17 Why did you abandon the “Bag”, “Shopping Cart”?

- Count er({'Better alternative offer': 133, 'Promo code not applicable': 54, 'Change in price': 37, 'Lack of trust': 31, 'No preferred mode of payment': 14})
The Value Counts in 18 The content on the website must be easy to read and understand - Counter({'Strongly agree (5)': 164, 'Agree (4)': 80, 'Strongly disagree (1)': 18, 'Indif ferent (3)': 7})
The Value Counts in 19 Information on similar product to the one highlighted is importa nt for product comparison - Counter({'Strongly agree (5)': 116, 'Agree (4)': 92, 'Indif ferent (3)': 43, 'Dis-agree (2)': 18})
The Value Counts in 20 Complete information on listed seller and product being offered i s important for purchase decision. - Counter({'Agree (4)': 101, 'Strongly agree (5)': 8 7, 'Indifferent (3)': 52, 'Dis-agree (2)': 18, 'Strongly disagree (1)': 11})
The Value Counts in 21 All relevant information on listed products must be stated clearl y - Counter({'Agree (4)': 132, 'Strongly agree (5)': 107, 'Strongly disagree (1)': 18, 'Dis-agree (2)': 12})
The Value Counts in 22 Ease of navigation in website - Counter({'Strongly agree (5)': 1 41, 'Agree (4)': 105, 'Strongly disagree (1)': 18, 'Dis-agree (2)': 5})
The Value Counts in 23 Loading and processing speed - Counter({'Strongly agree (5)': 11 5, 'Agree (4)': 112, 'Dis-agree (2)': 18, 'Strongly disagree (1)': 12, 'Indifferent (3)': 12})
The Value Counts in 24 User friendly Interface of the website - Counter({'Strongly agre e (5)': 189, 'Agree (4)': 45, 'Strongly disagree (1)': 18, 'Dis-agree (2)': 12, 'Indiffe rent (3)': 5})
The Value Counts in 25 Convenient Payment methods - Counter({'Strongly agree (5)': 159, 'Agree (4)': 80, 'Dis-agree (2)': 30})
The Value Counts in 26 Trust that the online retail store will fulfill its part of the t ransaction at the stipulated time - Counter({'Strongly agree (5)': 141, 'Agree (4)': 8 6, 'Disagree (2)': 30, 'indifferent (3)': 12})
The Value Counts in 27 Empathy (readiness to assist with queries) towards the customers - Counter({'Strongly agree (5)': 194, 'Agree (4)': 42, 'Strongly disagree (1)': 18, 'in different (3)': 15})
The Value Counts in 28 Being able to guarantee the privacy of the customer - Counter ({'Strongly agree (5)': 185, 'Agree (4)': 58, 'indifferent (3)': 26})
The Value Counts in 29 Responsiveness, availability of several communication channels (e

mail, online rep, twitter, phone etc.) - Counter({'Strongly agree (5)': 149, 'Agree (4)': 94, 'indifferent (3)': 15, 'Strongly disagree (1)': 11})
The Value Counts in 30 Online shopping gives monetary benefit and discounts - Counter({'Strongly agree (5)': 105, 'Agree (4)': 85, 'indifferent (3)': 50, 'Strongly disagree (1)': 18, 'Dis-agree (2)': 11})
The Value Counts in 31 Enjoyment is derived from shopping online - Counter({'Strongly agree (5)': 86, 'indifferent (3)': 75, 'Agree (4)': 59, 'Strongly disagree (1)': 30, 'Dis-agree (2)': 19})
The Value Counts in 32 Shopping online is convenient and flexible - Counter({'Strongly agree (5)': 146, 'Agree (4)': 78, 'indifferent (3)': 33, 'Dis-agree (2)': 12})
The Value Counts in 33 Return and replacement policy of the e-tailer is important for purchase decision - Counter({'Strongly agree (5)': 198, 'Agree (4)': 51, 'Dis-agree (2)': 20})
The Value Counts in 34 Gaining access to loyalty programs is a benefit of shopping online - Counter({'Strongly agree (5)': 115, 'Agree (4)': 64, 'indifferent (3)': 64, 'Dis-agree (2)': 15, 'Strongly disagree (1)': 11})
The Value Counts in 35 Displaying quality Information on the website improves satisfaction of customers - Counter({'Strongly agree (5)': 133, 'Agree (4)': 80, 'indifferent (3)': 56})
The Value Counts in 36 User derive satisfaction while shopping on a good quality website or application - Counter({'Strongly agree (5)': 175, 'Agree (4)': 86, 'Dis-agree (2)': 8})
The Value Counts in 37 Net Benefit derived from shopping online can lead to users satisfaction - Counter({'Strongly agree (5)': 164, 'Agree (4)': 54, 'indifferent (3)': 40, 'Dis-agree (2)': 11})
The Value Counts in 38 User satisfaction cannot exist without trust - Counter({'Strongly agree (5)': 122, 'Agree (4)': 117, 'Strongly disagree (1)': 18, 'Dis-agree (2)': 7, 'indifferent (3)': 5})
The Value Counts in 39 Offering a wide variety of listed product in several category - Counter({'Strongly agree (5)': 111, 'Agree (4)': 94, 'indifferent (3)': 57, 'Dis-agree (2)': 7})
The Value Counts in 40 Provision of complete and relevant product information - Counter({'Strongly agree (5)': 135, 'Agree (4)': 98, 'indifferent (3)': 31, 'Disagree (2)': 5})
The Value Counts in 41 Monetary savings - Counter({'Strongly agree (5)': 148, 'Agree (4)': 75, 'Disagree (2)': 31, 'indifferent (3)': 15})
The Value Counts in 42 The Convenience of patronizing the online retailer - Counter({'Agree (4)': 138, 'indifferent (3)': 77, 'Strongly agree (5)': 54})
The Value Counts in 43 Shopping on the website gives you the sense of adventure - Counter({'Agree (4)': 101, 'indifferent (3)': 59, 'Strongly agree (5)': 54, 'Dis-agree (2)': 50, 'Strongly disagree (1)': 5})
The Value Counts in 44 Shopping on your preferred e-tailer enhances your social status - Counter({'indifferent (3)': 100, 'Agree (4)': 59, 'Strongly agree (5)': 48, 'Strongly disagree (1)': 33, 'Dis-agree (2)': 29})
The Value Counts in 45 You feel gratification shopping on your favorite e-tailer - Counter({'indifferent (3)': 101, 'Strongly agree (5)': 65, 'Agree (4)': 63, 'Disagree (2)': 22, 'Strongly disagree (1)': 18})
The Value Counts in 46 Shopping on the website helps you fulfill certain roles - Counter({'Agree (4)': 88, 'indifferent (3)': 88, 'Strongly agree (5)': 38, 'Strongly disagree (1)': 33, 'Dis-agree (2)': 22})
The Value Counts in 47 Getting value for money spent - Counter({'Agree (4)': 149, 'Strongly agree (5)': 82, 'indifferent (3)': 38})
The Value Counts in From the following, tick any (or all) of the online retailers you have shopped from;
- Counter({'Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com': 82, 'Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com': 44, 'Amazon.in, Flipkart.com': 32, 'Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com': 29, 'Amazon.in, Flipkart.com, Snapdeal.com': 27, 'Amazon.in, Paytm.com, Myntra.com': 20, 'Amazon.in': 16, 'Amazon.in, Paytm.com': 12, 'Amazon.in, Flipkart.com, Paytm.com': 7})
The Value Counts in Easy to use website or application - Counter({'Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com': 64, 'Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com': 44, 'Amazon.in, Flipkart.com': 44, 'Amazon.in': 29, 'Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com': 22, 'Amazon.in, Paytm.com, Myntra.com': 20, 'Amazon.in, Flipkart.com, Myntra.com': 19, 'Paytm.com': 12, 'Flipkart.com': 8, 'Amazon.in, Paytm.com': 7})
The Value Counts in Visual appealing web-page layout - Counter({'Amazon.in, Flipkart.com': 87, 'Amazon.in': 44, 'Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com':

36, 'Amazon.in, Paytm.com, Myntra.com': 20, 'Amazon.in, Myntra.com': 15, 'Myntra.com': 1
5, 'Flipkart.com, Myntra.com': 15, 'Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com':
14, 'Flipkart.com': 12, 'Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com': 11})
The Value Counts in Wild variety of product on offer - Counter({'Amazon.in, Flipkart.co
m': 130, 'Amazon.in': 43, 'Amazon.in, Myntra.com': 20, 'Flipkart.com, Myntra.com': 15,
'Myntra.com': 15, 'Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com': 14, 'Amazon.in, F
lipkart.com, Paytm.com': 13, 'Flipkart.com': 12, 'Paytm.com': 7})
The Value Counts in Complete, relevant description information of products - Counter
({'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, Flipkar
t.com, Myntra.com, Snapdeal.com': 14, 'Snapdeal.com': 12, 'Flipkart.com, Snapdeal.com':
11, 'Flipkart.com': 8, 'Amazon.in, Flipkart.com, Snapdeal.com': 7})
The Value Counts in Fast loading website speed of website and application - Counter({'A
mazon.in': 51, 'Amazon.in, Paytm.com': 44, 'Amazon.in, Flipkart.com, Myntra.com': 30, 'A
mazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com': 30, 'Amazon.in, Flipkart.c
om': 30, 'Amazon.in, Flipkart.com, Snapdeal.com': 25, 'Amazon.in, Flipkart.com, Paytm.co
m': 25, 'Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com': 14, 'Snapdeal.com': 12, 'Fl
ipkart.com': 8})
The Value Counts in Reliability of the website or application - Counter({'Amazon.in': 6
1, 'Amazon.in, Flipkart.com': 50, 'Amazon.in, Flipkart.com, Paytm.com': 36, 'Amazon.in,
Paytm.com, Myntra.com': 35, 'Amazon.in, Flipkart.com, Snapdeal.com': 18, 'Myntra.com': 1
5, 'Flipkart.com': 15, 'Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com': 14, 'Amazon.
in, Flipkart.com, Paytm.com, Snapdeal.com': 13, 'Paytm.com': 12})
The Value Counts in Quickness to complete purchase - Counter({'Amazon.com': 66, 'Amazo
n.com, Flipkart.com, Paytm.com': 47, 'Amazon.com, Flipkart.com': 37, 'Amazon.com, Flipka
rt.com, Myntra.com': 30, 'Paytm.com': 25, 'Amazon.com, Paytm.com, Myntra.com': 20, 'Amaz
on.com, Flipkart.com, Paytm.com, Myntra.com, Snapdeal': 15, 'Flipkart.com': 15, 'Flipkar
t.com, Myntra.com, Snapdeal': 14})
The Value Counts in Availability of several payment options - Counter({'Amazon.in, Flip
kart.com': 65, 'Amazon.in, Flipkart.com, Myntra.com': 40, 'Amazon.in, Flipkart.com, Paty
m.com, Myntra.com, Snapdeal.com': 39, 'Amazon.in': 23, 'Patym.com, Myntra.com': 20, 'Ama
zon.in, Flipkart.com, Myntra.com, Snapdeal.com': 19, 'Amazon.in, Flipkart.com, Snapdeal.
com': 18, 'Flipkart.com, Myntra.com, Snapdeal.com': 14, 'Patym.com': 12, 'Amazon.in, Pat
ym.com': 11, 'Flipkart.com': 8})
The Value Counts in Speedy order delivery - Counter({'Amazon.in': 107, 'Amazon.in, Fli
pkart.com': 82, 'Amazon.in, Flipkart.com, Snapdeal.com': 36, 'Amazon.in, Flipkart.com, M
yntra.com': 15, 'Flipkart.com': 15, 'Flipkart.com, Myntra.com, Snapdeal.com': 14})
The Value Counts in Privacy of customers' information - Counter({'Amazon.in': 71, 'Amaz
on.in, Flipkart.com': 54, 'Amazon.in, Flipkart.com, Myntra.com': 25, 'Amazon.in, Flipkar
t.com, Paytm.com, Myntra.com, Snapdeal.com': 24, 'Paytm.com': 18, 'Myntra.com': 15, 'Ama
zon.in, Paytm.com': 15, 'Flipkart.com': 15, 'Amazon.in, Flipkart.com, Myntra.com, Snapde
al.com': 14, 'Amazon.in, Flipkart.com, Paytm.com': 11, 'Amazon.in, Flipkart.com, Snapdea
l.com': 7})
The Value Counts in Security of customer financial information - Counter({'Amazon.in':
51, 'Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com': 42, 'Flipkart.com':
33, 'Amazon.in, Flipkart.com, Snapdeal.com': 25, 'Amazon.in, Flipkart.com': 24, 'Amazon.
in, Paytm.com, Myntra.com': 20, 'Amazon.in, Snapdeal.com': 19, 'Myntra.com': 15, 'Paytm.
com': 15, 'Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com': 14, 'Amazon.in, Flipkart.
com, Paytm.com': 11})
The Value Counts in Perceived Trustworthiness - Counter({'Amazon.in': 76, 'Amazon.in, F
lipkart.com, Snapdeal.com': 36, 'Amazon.in, Myntra.com': 35, 'Amazon.in, Flipkart.com':
31, 'Flipkart.com': 27, 'Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com': 25, 'Myntr
a.com': 15, 'Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com': 13, 'Amazon.
in, Flipkart.com, Paytm.com': 11})
The Value Counts in Presence of online assistance through multi-channel - Counter({'Ama
zon.in, Flipkart.com, Myntra.com, Snapdeal': 61, 'Amazon.in': 60, 'Amazon.in, Flipkart.c
om': 39, 'Amazon.in, Snapdeal': 26, 'Myntra.com': 20, 'Amazon.in, Flipkart.com, Myntra.c
om': 15, 'Amazon.in, Myntra.com': 15, 'Amazon.in, Flipkart.com, Paytm.com': 13, 'Paytm.c
om': 12, 'Flipkart.com': 8})
The Value Counts in Longer time to get logged in (promotion, sales period) - Counter
({'Amazon.in': 57, 'Amazon.in, Flipkart.com': 38, 'Paytm.com': 38, 'Myntra.com': 35, 'Am
azon.in, Flipkart.com, Snapdeal.com': 29, 'Snapdeal.com': 25, 'Flipkart.com, Paytm.com':
15, 'Flipkart.com, Paytm.com, Snapdeal.com': 13, 'Amazon.in, Paytm.com': 11, 'Flipkart.c
om': 8})
The Value Counts in Longer time in displaying graphics and photos (promotion, sales peri

```
od) -  Counter({'Amazon.in, Flipkart.com': 60, 'Amazon.in': 39, 'Myntra.com': 35, 'Snapd
eal.com': 34, 'Myntra.com, Snapdeal.com': 25, 'Flipkart.com, Snapdeal.com': 19, 'Paytm.c
om': 15, 'Flipkart.com': 15, 'Amazon.in, Myntra.com, Snapdeal.com': 14, 'Amazon.in, Payt
m.com': 13})
The Value Counts in Late declaration of price (promotion, sales period) -  Counter({'Myn
tra.com': 75, 'Paytm.com': 52, 'snapdeal.com': 41, 'Flipkart.com': 38, 'Amazon.in': 38,
'Amazon.in, Paytm.com': 13, 'Paytm.com, snapdeal.com': 7, 'Amazon.in, Flipkart.com': 5})
The Value Counts in Longer page loading time (promotion, sales period) -  Counter({'Mynt
ra.com': 61, 'Paytm.com': 59, 'Flipkart.com': 32, 'Snapdeal.com': 23, 'Amazon.in, Flipka
rt.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})
The Value Counts in Limited mode of payment on most products (promotion, sales period) -
  Counter({'Snapdeal.com': 87, 'Amazon.in': 62, 'Flipkart.com': 31, 'Amazon.in, Flipkar
t.com': 29, 'Paytm.com': 25, 'Paytm.com, Snapdeal.com': 15, 'Amazon.in, Paytm.com': 13,
'Myntra.com, Snapdeal.com': 7})
The Value Counts in Longer delivery period -  Counter({'Paytm.com': 72, 'Snapdeal.com':
64, 'Flipkart.com': 44, 'Amazon.in': 37, 'Paytm.com, Snapdeal.com': 26, 'Myntra.com': 2
6})
The Value Counts in Change in website/Application design -  Counter({'Amazon.in': 96, 'P
aytm.com': 63, 'Amazon.in, Flipkart.com': 45, 'Myntra.com': 30, 'Flipkart.com': 20, 'Sna
pdeal.com': 8, 'Flipkart.com, Myntra.com': 7})
The Value Counts in Frequent disruption when moving from one page to another -  Counter
({'Amazon.in': 53, 'Myntra.com': 52, 'Snapdeal.com': 49, 'Paytm.com': 39, 'Flipkart.co
m': 26, 'Amazon.in, Flipkart.com': 25, 'Myntra.com, Snapdeal.com': 14, 'Flipkart.com, Sn
apdeal.com': 11})
The Value Counts in Website is as efficient as before -  Counter({'Amazon.in': 94, 'Flip
kart.com': 47, 'Amazon.in, Flipkart.com': 45, 'Amazon.in, Flipkart.com, Paytm.com': 25,
'Amazon.in, Paytm.com': 18, 'Paytm.com': 15, 'Myntra.com, Snapdeal.com': 14, 'Snapdeal.c
om': 11})
The Value Counts in Which of the Indian online retailer would you recommend to a friend?
-  Counter({'Amazon.in': 79, 'Amazon.in, Flipkart.com': 62, 'Flipkart.com': 39, 'Amazon.
in, Myntra.com': 30, 'Amazon.in, Paytm.com, Myntra.com': 20, 'Amazon.in, Flipkart.com, M
yntra.com': 15, 'Amazon.in, Paytm.com': 13, 'Flipkart.com, Paytm.com, Myntra.com, snapde
al.com': 11})
```

Step 2.4 - Observations

The Train Dataset is made up of;

- 269 rows and 72 columns
- 71 features and 1 Target
- We have a problem of dimensionality(curse of dimensionality)
- We have too many features and there is a need to perform feature selection to determine customer retention
- We will definitely drop the columns where the names of the companies are mentioned
- We are only paying major attention to the factors the lead to customer retention or customer repurchase
- The dataset set only contains categorical data and no continous data.
- The target data is categorical meaning that if we are to use a machine learning algorithm it must definitely be a classification algorithm

- we will drop all the columns after the 47th column which contains the names of the online companies choosen.
- We are dropping it because its not important

About the Target/Label:

- How many times you have made an online purchase in the past 1 year?': This is the target variable and has the following properties:
- It has no null values
- Its Datatype is a categorical data

**Since the target is made up of a categorical data hence this is a classification PROBLEM!

Step 3 - EXPLORATORY DATA ANALYSIS(EDA) AND VISUALIZATION

The objectives of the EDA is to ensure;

1. Independent Features will have a normal distribution using Data Transformation to remove skewness
2. Datasets will have the same scale using Normalization and Standardization
3. Column mean will be zero
4. Standard Deviation of the data should be 1

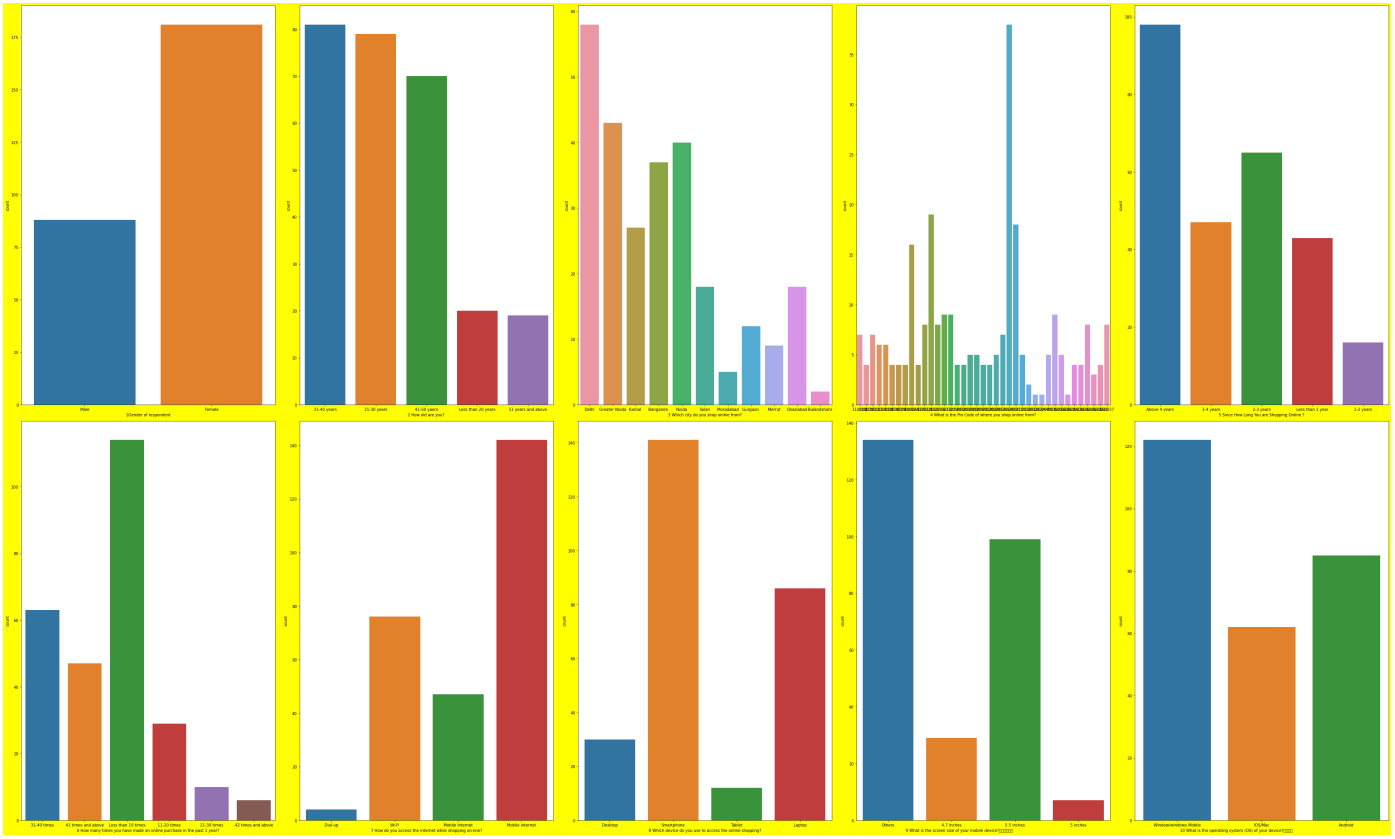
3.1 Count Distribution of Categorical Data - Univariate Analysis

In [16]:
```
#Check for normal distribution and see how data is distributed for every column
plt.figure(figsize=(50,30), facecolor='yellow')#dimension and color
plotnumber = 1 #like a form of counter which we will see its relevance in the 'for loop'

for column in df.iloc[:,0:10]:#triggering your for loop
    if plotnumber<=10:
        ax = plt.subplot(2,5,plotnumber)#2,3 means arrange the graphs in 3-rows and 4-co
        sns.countplot(df[column],data=df)#performing a distribution plot for all the col
        plt.xlabel(column,fontsize=10)#labelling the x-axis with their respctive column
        print(df[column].value_counts())

    plotnumber+=1#incrementals for plot number so as to allow ploting one by one until i
plt.tight_layout()#This is just to show you data in a compact form.
```

```
Female    181
Male       88
Name: 1Gender of respondent, dtype: int64
31-40 years           81
21-30 years           79
41-50 yaers           70
Less than 20 years    20
51 years and above    19
Name: 2 How old are you? , dtype: int64
```

```
Delhi            58
Greater Noida    43
Noida            40
Bangalore        37
Karnal           27
Solan            18
Ghaziabad        18
Gurgaon          12
Merrut            9
Moradabad         5
Bulandshahr       2
Name: 3 Which city do you shop online from?, dtype: int64
201308    38
132001    19
201310    18
110044    16
250001     9
173229     9
173212     9
560010     8
132036     8
122018     8
560037     8
110008     7
110011     7
201306     7
110014     6
110018     6
201305     5
201008     5
201009     5
201312     5
244001     5
530068     5
122009     4
201001     4
560003     4
201304     4
110009     4
201303     4
560002     4
560018     4
110042     4
110030     4
201005     4
110039     4
560013     3
203001     2
203202     1
560001     1
203207     1
Name: 4 What is the Pin Code of where you shop online from?, dtype: int64
Above 4 years       98
2-3 years           65
3-4 years           47
Less than 1 year    43
1-2 years           16
Name: 5 Since How Long You are Shopping Online ?, dtype: int64
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
Name: 6 How many times you have made an online purchase in the past 1 year?, dtype: int6
4
```

```
Mobile internet    142
Wi-Fi               76
Mobile Internet     47
Dial-up              4
Name: 7 How do you access the internet while shopping on-line?, dtype: int64
Smartphone    141
Laptop         86
Desktop        30
Tablet         12
Name: 8 Which device do you use to access the online shopping?, dtype: int64
Others        134
5.5 inches     99
4.7 inches     29
5 inches        7
Name: 9 What is the screen size of your mobile device?\t\t\t\t\t\t
                  , dtype: int64
Window/windows Mobile    122
Android                   85
IOS/Mac                   62
Name: 10 What is the operating system (OS) of your device?\t\t\t\t
                      , dtype: int64
```

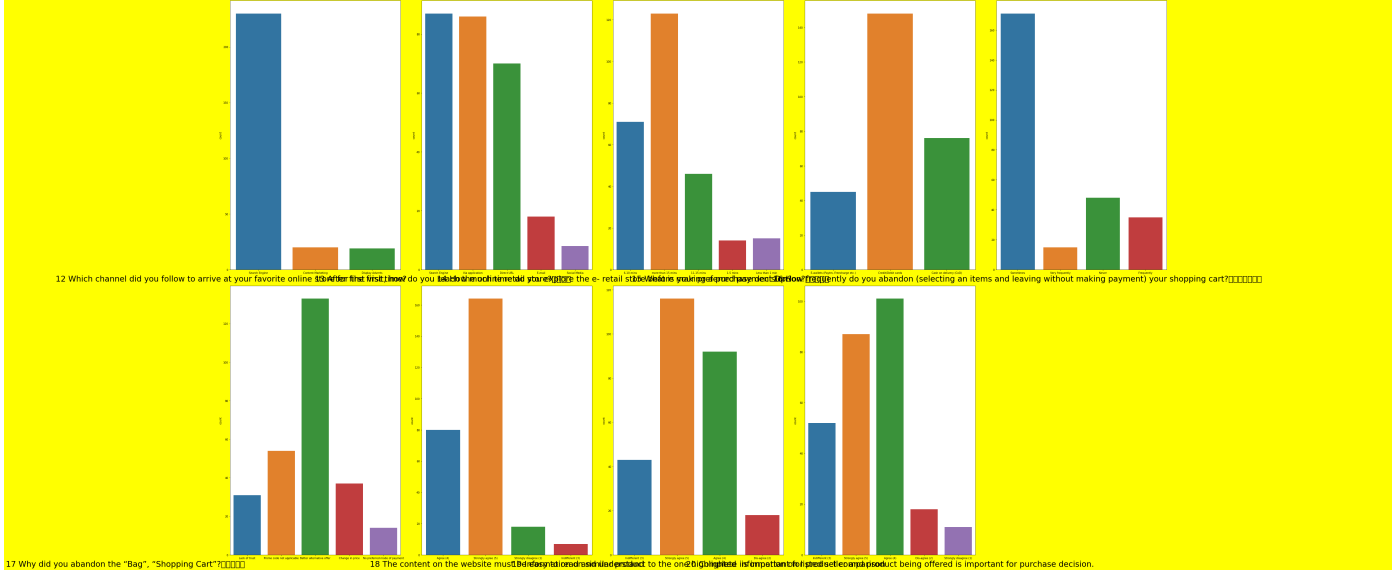


In [17]:

```
#Check for normal distribution and see how data is distributed for every column
plt.figure(figsize=(50,30), facecolor='yellow')#dimension and color
plotnumber = 1 #like a form of counter which we will see its relevance in the 'for loop'

for column in df.iloc[:,11:20]:#triggering your for loop
    if plotnumber<=10:
        ax = plt.subplot(2,5,plotnumber)#2,3 means arrange the graphs in 3-rows and 4-co
        sns.countplot(df[column],data=df)#performing a distribution plot for all the col
        plt.xlabel(column,fontsize=30)#labelling the x-axis with their respctive column
        print(df[column].value_counts())

    plotnumber+=1#incrementals for plot number so as to allow ploting one by one until i
plt.tight_layout()#This is just to show you data in a compact form.
```

```
Search Engine        230
Content Marketing     20
Display Adverts       19
Name: 12 Which channel did you follow to arrive at your favorite online store for the fi
```

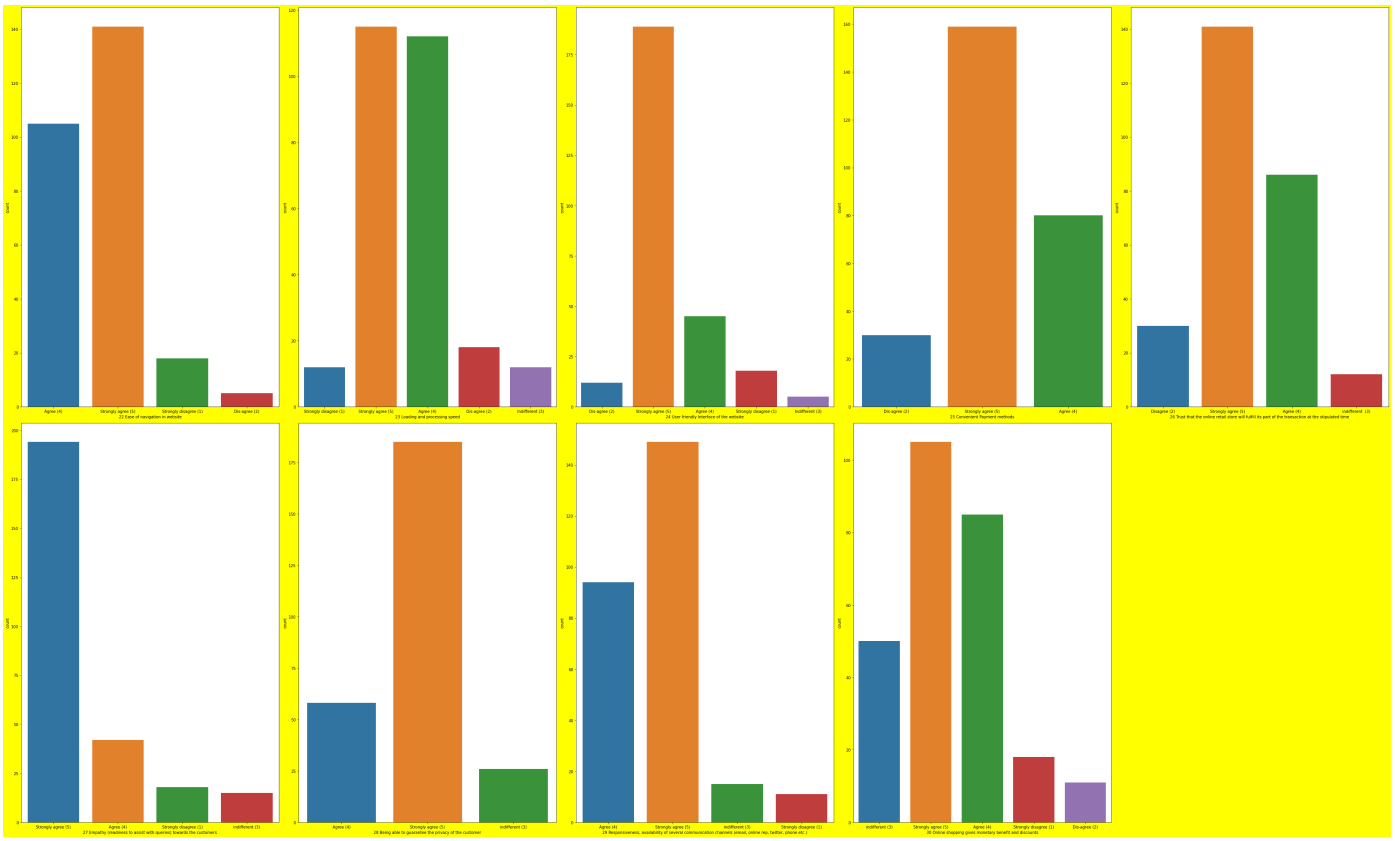
```
rst time?                                                                    , dtype: int
64
Search Engine      87
Via application    86
Direct URL         70
E-mail             18
Social Media        8
Name: 13 After first visit, how do you reach the online retail store?\t\t\t\t
                                          , dtype: int64
more than 15 mins    123
6-10 mins             71
11-15 mins            46
Less than 1 min       15
1-5 mins              14
Name: 14 How much time do you explore the e- retail store before making a purchase decis
ion?                                                                  , dtype: int64
Credit/Debit cards                    148
Cash on delivery (CoD)                 76
E-wallets (Paytm, Freecharge etc.)     45
Name: 15 What is your preferred payment Option?\t\t\t\t\t

        , dtype: int64
Sometimes          171
Never               48
Frequently          35
Very frequently     15
Name: 16 How frequently do you abandon (selecting an items and leaving without making pa
yment) your shopping cart?\t\t\t\t\t\t\t
   , dtype: int64
Better alternative offer        133
Promo code not applicable        54
Change in price                  37
Lack of trust                    31
No preferred mode of payment     14
Name: 17 Why did you abandon the “Bag”, “Shopping Cart”?\t\t\t\t\t

                                          , dtype: int64
Strongly agree (5)       164
Agree (4)                 80
Strongly disagree (1)     18
Indifferent (3)            7
Name: 18 The content on the website must be easy to read and understand, dtype: int64
Strongly agree (5)    116
Agree (4)              92
Indifferent (3)        43
Dis-agree (2)          18
Name: 19 Information on similar product to the one highlighted  is important for product
comparison, dtype: int64
Agree (4)                101
Strongly agree (5)        87
Indifferent (3)           52
Dis-agree (2)             18
Strongly disagree (1)     11
Name: 20 Complete information on listed seller and product being offered is important fo
r purchase decision., dtype: int64
```



In [18]:
```
#Check for normal distribution and see how data is distributed for every column
plt.figure(figsize=(50,30), facecolor='yellow')#dimension and color
plotnumber = 1 #like a form of counter which we will see its relevance in the 'for loop'

for column in df.iloc[:,21:30]:#triggering your for loop
    if plotnumber<=10:
        ax = plt.subplot(2,5,plotnumber)#2,3 means arrange the graphs in 3-rows and 4-co
        sns.countplot(df[column],data=df)#performing a distribution plot for all the col
        plt.xlabel(column,fontsize=10)#labelling the x-axis with their respctive column
        print(df[column].value_counts())

    plotnumber+=1#incrementals for plot number so as to allow ploting one by one until i
plt.tight_layout()#This is just to show you data in a compact form.
```

```
Strongly agree (5)       141
Agree (4)                105
Strongly disagree (1)     18
Dis-agree (2)              5
Name: 22 Ease of navigation in website, dtype: int64
Strongly agree (5)       115
Agree (4)                112
Dis-agree (2)             18
Strongly disagree (1)     12
Indifferent (3)           12
Name: 23 Loading and processing speed, dtype: int64
Strongly agree (5)       189
Agree (4)                 45
Strongly disagree (1)     18
Dis-agree (2)             12
Indifferent (3)            5
Name: 24 User friendly Interface of the website, dtype: int64
Strongly agree (5)    159
Agree (4)              80
Dis-agree (2)          30
Name: 25 Convenient Payment methods, dtype: int64
Strongly agree (5)    141
Agree (4)              86
Disagree (2)           30
indifferent  (3)       12
Name: 26 Trust that the online retail store will fulfill its part of the transaction at
the stipulated time, dtype: int64
Strongly agree (5)       194
Agree (4)                 42
Strongly disagree (1)     18
indifferent (3)           15
Name: 27 Empathy (readiness to assist with queries) towards the customers, dtype: int64
Strongly agree (5)    185
```

```
Agree (4)                58
indifferent (3)          26
Name: 28 Being able to guarantee the privacy of the customer, dtype: int64
Strongly agree (5)       149
Agree (4)                 94
indifferent (3)           15
Strongly disagree (1)     11
Name: 29 Responsiveness, availability of several communication channels (email, online r
ep, twitter, phone etc.), dtype: int64
Strongly agree (5)       105
Agree (4)                 85
indifferent (3)           50
Strongly disagree (1)     18
Dis-agree (2)             11
Name: 30 Online shopping gives monetary benefit and discounts, dtype: int64
```

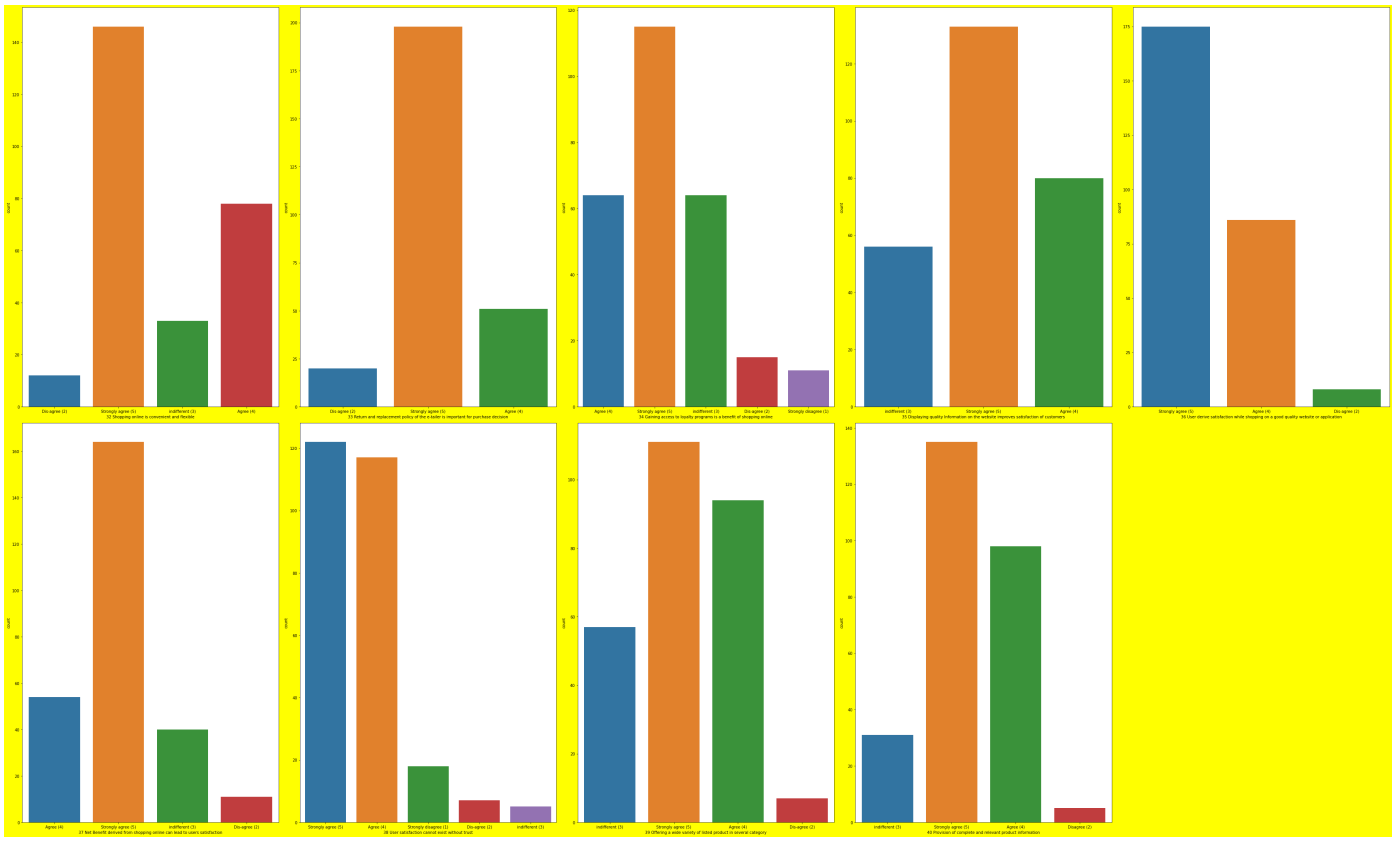


In [19]:

```
#Check for normal distribution and see how data is distributed for every column
plt.figure(figsize=(50,30), facecolor='yellow')#dimension and color
plotnumber = 1 #like a form of counter which we will see its relevance in the 'for loop'

for column in df.iloc[:,31:40]:#triggering your for loop
    if plotnumber<=10:
        ax = plt.subplot(2,5,plotnumber)#2,3 means arrange the graphs in 3-rows and 4-co
        sns.countplot(df[column],data=df)#performing a distribution plot for all the col
        plt.xlabel(column,fontsize=10)#labelling the x-axis with their respctive column
        print(df[column].value_counts())

    plotnumber+=1#incrementals for plot number so as to allow ploting one by one until i
plt.tight_layout()#This is just to show you data in a compact form.
```

```
Strongly agree (5)    146
Agree (4)              78
indifferent (3)        33
Dis-agree (2)          12
Name: 32 Shopping online is convenient and flexible, dtype: int64
Strongly agree (5)    198
Agree (4)              51
Dis-agree (2)          20
Name: 33 Return and replacement policy of the e-tailer is important for purchase decisio
n, dtype: int64
```

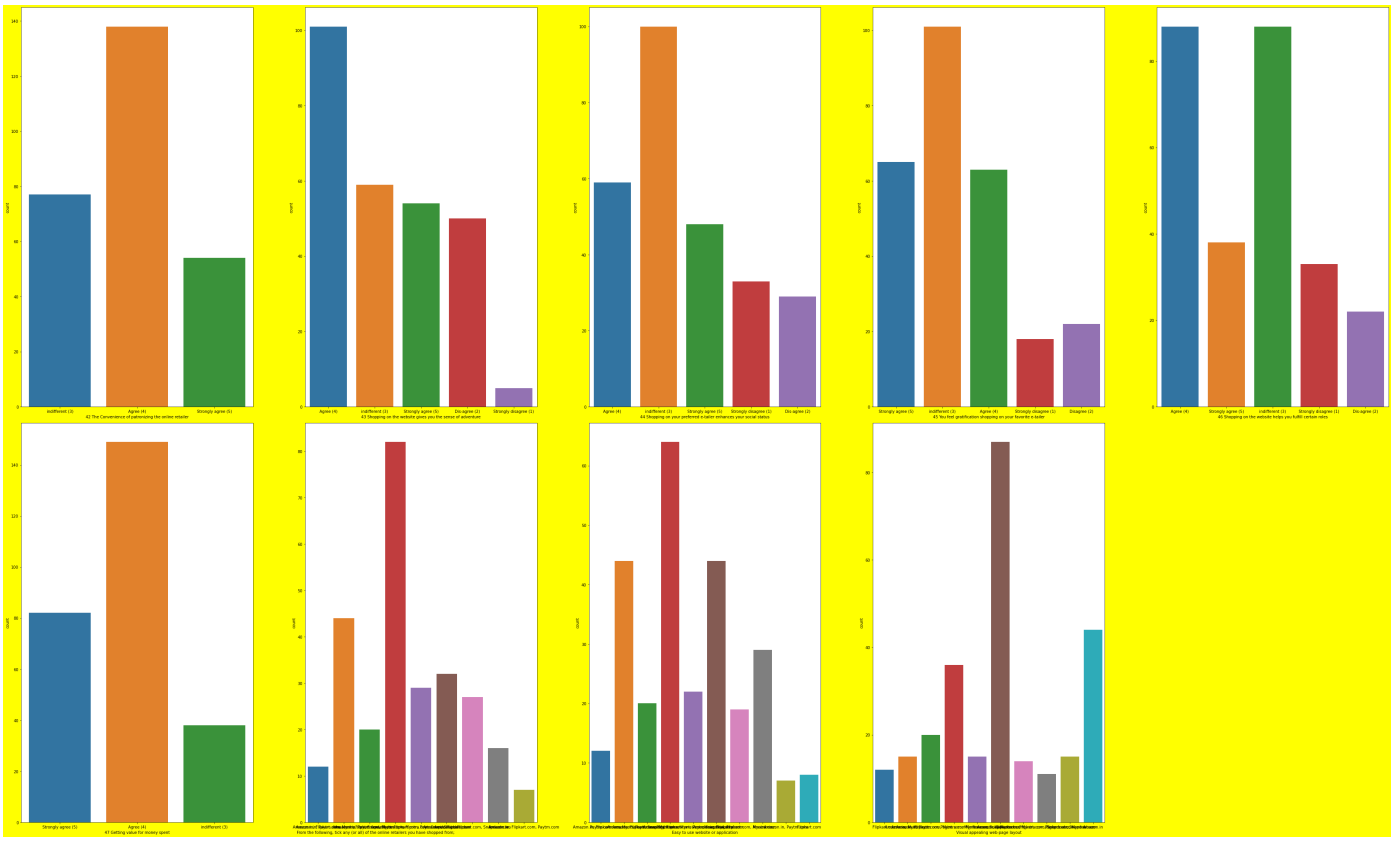
```
Strongly agree (5)       115
Agree (4)                 64
indifferent (3)           64
Dis-agree (2)             15
Strongly disagree (1)     11
Name: 34 Gaining access to loyalty programs is a benefit of shopping online, dtype: int64
Strongly agree (5)    133
Agree (4)              80
indifferent (3)        56
Name: 35 Displaying quality Information on the website improves satisfaction of customers, dtype: int64
Strongly agree (5)    175
Agree (4)              86
Dis-agree (2)           8
Name: 36 User derive satisfaction while shopping on a good quality website or application, dtype: int64
Strongly agree (5)    164
Agree (4)              54
indifferent (3)        40
Dis-agree (2)          11
Name: 37 Net Benefit derived from shopping online can lead to users satisfaction, dtype: int64
Strongly agree (5)       122
Agree (4)                117
Strongly disagree (1)     18
Dis-agree (2)              7
indifferent (3)            5
Name: 38 User satisfaction cannot exist without trust, dtype: int64
Strongly agree (5)    111
Agree (4)              94
indifferent (3)        57
Dis-agree (2)           7
Name: 39 Offering a wide variety of listed product in several category, dtype: int64
Strongly agree (5)    135
Agree (4)              98
indifferent (3)        31
Disagree (2)            5
Name: 40 Provision of complete and relevant product information, dtype: int64
```



In [20]:
```
#Check for normal distribution and see how data is distributed for every column
plt.figure(figsize=(50,30), facecolor='yellow')#dimension and color
plotnumber = 1 #like a form of counter which we will see its relevance in the 'for loop'

for column in df.iloc[:,41:50]:#triggering your for loop
    if plotnumber<=10:
        ax = plt.subplot(2,5,plotnumber)#2,3 means arrange the graphs in 3-rows and 4-co
        sns.countplot(df[column],data=df)#performing a distribution plot for all the col
        plt.xlabel(column,fontsize=10)#labelling the x-axis with their respctive column
        print(df[column].value_counts())

    plotnumber+=1#incrementals for plot number so as to allow ploting one by one until i
plt.tight_layout()#This is just to show you data in a compact form.
```

```
Agree (4)              138
indifferent (3)         77
Strongly agree (5)      54
Name: 42 The Convenience of patronizing the online retailer, dtype: int64
Agree (4)                101
indifferent (3)           59
Strongly agree (5)        54
Dis-agree (2)             50
Strongly disagree (1)      5
Name: 43 Shopping on the website gives you the sense of adventure, dtype: int64
indifferent (3)          100
Agree (4)                 59
Strongly agree (5)        48
Strongly disagree (1)     33
Dis-agree (2)             29
Name: 44 Shopping on your preferred e-tailer enhances your social status, dtype: int64
indifferent (3)          101
Strongly agree (5)        65
Agree (4)                 63
Disagree (2)              22
Strongly disagree (1)     18
Name: 45 You feel gratification shopping on your favorite e-tailer, dtype: int64
Agree (4)                88
indifferent (3)          88
Strongly agree (5)       38
Strongly disagree (1)    33
Dis-agree (2)            22
Name: 46 Shopping on the website helps you fulfill certain roles, dtype: int64
Agree (4)              149
Strongly agree (5)      82
indifferent (3)         38
Name: 47 Getting value for money spent, dtype: int64
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com    82
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com               44
Amazon.in, Flipkart.com                                         32
Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com                29
Amazon.in, Flipkart.com, Snapdeal.com                           27
Amazon.in, Paytm.com, Myntra.com                                20
Amazon.in                                                       16
Amazon.in, Paytm.com                                            12
Amazon.in, Flipkart.com, Paytm.com                               7
Name: From the following, tick any (or all) of the online retailers you have shopped fro
m;                                                                                    , dtype: in
t64
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com    64
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com               44
Amazon.in, Flipkart.com                                         44
Amazon.in                                                       29
Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com                22
Amazon.in, Paytm.com, Myntra.com                                20
Amazon.in, Flipkart.com, Myntra.com                             19
Paytm.com                                                       12
```

```
Flipkart.com                                                          8
Amazon.in, Paytm.com                                                  7
Name: Easy to use website or application, dtype: int64
Amazon.in, Flipkart.com                                              87
Amazon.in                                                            44
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com         36
Amazon.in, Paytm.com, Myntra.com                                     20
Amazon.in, Myntra.com                                                15
Myntra.com                                                           15
Flipkart.com, Myntra.com                                             15
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com                    14
Flipkart.com                                                         12
Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com                     11
Name: Visual appealing web-page layout, dtype: int64
```



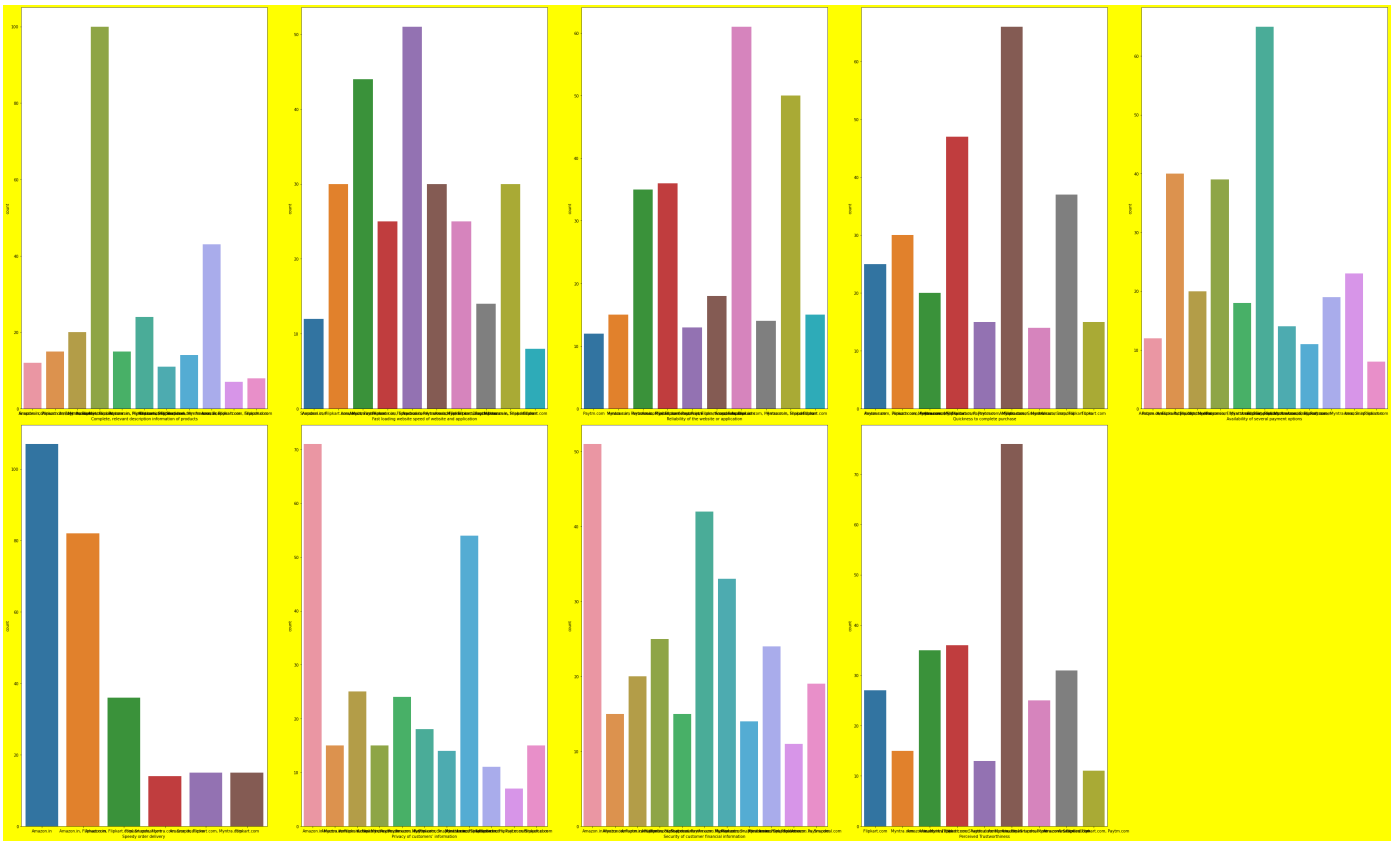
In [21]:

```
#Check for normal distribution and see how data is distributed for every column
plt.figure(figsize=(50,30), facecolor='yellow')#dimension and color
plotnumber = 1 #like a form of counter which we will see its relevance in the 'for loop'

for column in df.iloc[:,51:60]:#triggering your for loop
    if plotnumber<=10:
        ax = plt.subplot(2,5,plotnumber)#2,3 means arrange the graphs in 3-rows and 4-co
        sns.countplot(df[column],data=df)#performing a distribution plot for all the col
        plt.xlabel(column,fontsize=10)#labelling the x-axis with their respctive column
        print(df[column].value_counts())

    plotnumber+=1#incrementals for plot number so as to allow ploting one by one until i
plt.tight_layout()#This is just to show you data in a compact form.
```

```
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: Complete, relevant description information of products, dtype: int64
Amazon.in                                                    51
Amazon.in, Paytm.com                                         44
Amazon.in, Flipkart.com, Myntra.com                          30
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com 30
Amazon.in, Flipkart.com                                      30
Amazon.in, Flipkart.com, Snapdeal.com                        25
Amazon.in, Flipkart.com, Paytm.com                           25
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com            14
Snapdeal.com                                                 12
Flipkart.com                                                  8
Name: Fast loading website speed of website and application, dtype: int64
Amazon.in                                                 61
Amazon.in, Flipkart.com                                   50
Amazon.in, Flipkart.com, Paytm.com                        36
Amazon.in, Paytm.com, Myntra.com                          35
Amazon.in, Flipkart.com, Snapdeal.com                     18
Myntra.com                                                15
Flipkart.com                                              15
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com         14
Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com          13
Paytm.com                                                 12
Name: Reliability of the website or application, dtype: int64
Amazon.com                                                   66
Amazon.com, Flipkart.com, Paytm.com                          47
Amazon.com, Flipkart.com                                     37
Amazon.com, Flipkart.com, Myntra.com                         30
Paytm.com                                                    25
Amazon.com, Paytm.com, Myntra.com                            20
Amazon.com, Flipkart.com, Paytm.com, Myntra.com, Snapdeal    15
Flipkart.com                                                 15
Flipkart.com, Myntra.com, Snapdeal                           14
Name: Quickness to complete purchase, dtype: int64
Amazon.in, Flipkart.com                                      65
Amazon.in, Flipkart.com, Myntra.com                          40
Amazon.in, Flipkart.com, Patym.com, Myntra.com, Snapdeal.com 39
Amazon.in                                                    23
Patym.com, Myntra.com                                        20
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com            19
Amazon.in, Flipkart.com, Snapdeal.com                        18
Flipkart.com, Myntra.com, Snapdeal.com                       14
Patym.com                                                    12
Amazon.in, Patym.com                                         11
Flipkart.com                                                  8
Name: Availability of several payment options, dtype: int64
Amazon.in                                  107
Amazon.in, Flipkart.com                     82
Amazon.in, Flipkart.com, Snapdeal.com       36
Amazon.in, Flipkart.com, Myntra.com         15
Flipkart.com                                15
Flipkart.com, Myntra.com, Snapdeal.com      14
Name: Speedy order delivery , dtype: int64
Amazon.in                                                    71
Amazon.in, Flipkart.com                                      54
Amazon.in, Flipkart.com, Myntra.com                          25
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com 24
Paytm.com                                                    18
Myntra.com                                                   15
Amazon.in, Paytm.com                                         15
Flipkart.com                                                 15
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com            14
Amazon.in, Flipkart.com, Paytm.com                           11
Amazon.in, Flipkart.com, Snapdeal.com                         7
Name: Privacy of customers' information, dtype: int64
Amazon.in                                                    51
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com 42
```

```
Flipkart.com                                                    33
Amazon.in, Flipkart.com, Snapdeal.com                           25
Amazon.in, Flipkart.com                                         24
Amazon.in, Paytm.com, Myntra.com                                20
Amazon.in, Snapdeal.com                                         19
Myntra.com                                                      15
Paytm.com                                                       15
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com               14
Amazon.in, Flipkart.com, Paytm.com                              11
Name: Security of customer financial information, dtype: int64
Amazon.in                                                       76
Amazon.in, Flipkart.com, Snapdeal.com                           36
Amazon.in, Myntra.com                                           35
Amazon.in, Flipkart.com                                         31
Flipkart.com                                                    27
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com               25
Myntra.com                                                      15
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com    13
Amazon.in, Flipkart.com, Paytm.com                              11
Name: Perceived Trustworthiness, dtype: int64
```



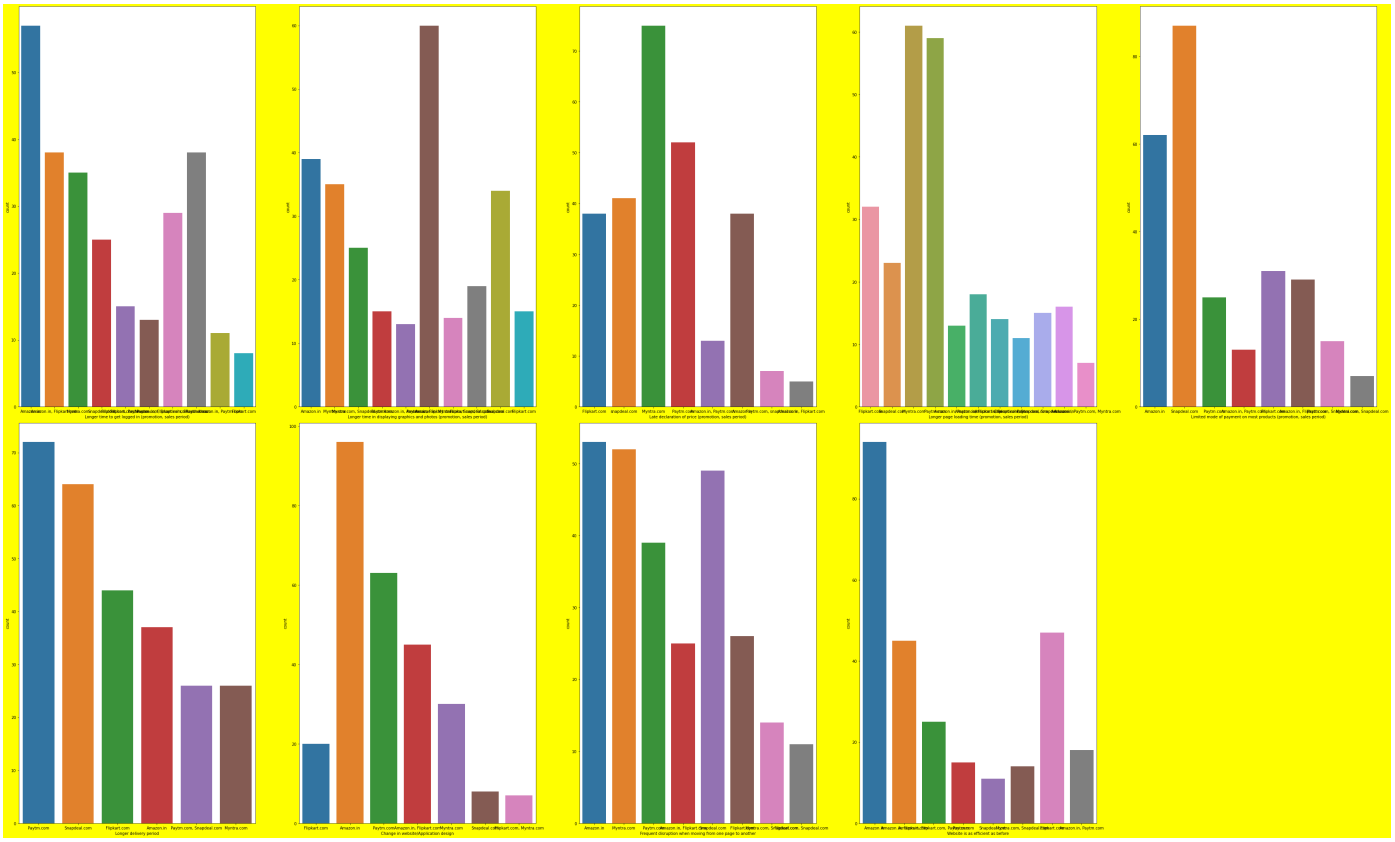
In [22]:

```
#Check for normal distribution and see how data is distributed for every column
plt.figure(figsize=(50,30), facecolor='yellow')#dimension and color
plotnumber = 1 #like a form of counter which we will see its relevance in the 'for loop'

for column in df.iloc[:,61:70]:#triggering your for loop
    if plotnumber<=10:
        ax = plt.subplot(2,5,plotnumber)#2,3 means arrange the graphs in 3-rows and 4-co
        sns.countplot(df[column],data=df)#performing a distribution plot for all the col
        plt.xlabel(column,fontsize=10)#labelling the x-axis with their respctive column
        print(df[column].value_counts())

    plotnumber+=1#incrementals for plot number so as to allow ploting one by one until i
plt.tight_layout()#This is just to show you data in a compact form.
```

```
Amazon.in                                57
Amazon.in, Flipkart.com                  38
Paytm.com                                38
Myntra.com                               35
```

```
Amazon.in, Flipkart.com, Snapdeal.com    29
Snapdeal.com                             25
Flipkart.com, Paytm.com                  15
Flipkart.com, Paytm.com, Snapdeal.com    13
Amazon.in, Paytm.com                     11
Flipkart.com                              8
Name: Longer time to get logged in (promotion, sales period), dtype: int64
Amazon.in, Flipkart.com                60
Amazon.in                              39
Myntra.com                             35
Snapdeal.com                           34
Myntra.com, Snapdeal.com               25
Flipkart.com, Snapdeal.com             19
Paytm.com                              15
Flipkart.com                           15
Amazon.in, Myntra.com, Snapdeal.com    14
Amazon.in, Paytm.com                   13
Name: Longer time in displaying graphics and photos (promotion, sales period), dtype: in
t64
Myntra.com                75
Paytm.com                 52
snapdeal.com              41
Flipkart.com              38
Amazon.in                 38
Amazon.in, Paytm.com      13
Paytm.com, snapdeal.com    7
Amazon.in, Flipkart.com    5
Name: Late declaration of price (promotion, sales period), dtype: int64
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: Longer page loading time (promotion, sales period), dtype: int64
Snapdeal.com               87
Amazon.in                  62
Flipkart.com               31
Amazon.in, Flipkart.com    29
Paytm.com                  25
Paytm.com, Snapdeal.com    15
Amazon.in, Paytm.com       13
Myntra.com, Snapdeal.com    7
Name: Limited mode of payment on most products (promotion, sales period), dtype: int64
Paytm.com                 72
Snapdeal.com              64
Flipkart.com              44
Amazon.in                 37
Paytm.com, Snapdeal.com   26
Myntra.com                26
Name: Longer delivery period, dtype: int64
Amazon.in                  96
Paytm.com                  63
Amazon.in, Flipkart.com    45
Myntra.com                 30
Flipkart.com               20
Snapdeal.com                8
Flipkart.com, Myntra.com    7
Name: Change in website/Application design, dtype: int64
Amazon.in                   53
Myntra.com                  52
```

```
Snapdeal.com                        49
Paytm.com                           39
Flipkart.com                        26
Amazon.in, Flipkart.com             25
Myntra.com, Snapdeal.com            14
Flipkart.com, Snapdeal.com          11
Name: Frequent disruption when moving from one page to another, dtype: int64
Amazon.in                                  94
Flipkart.com                               47
Amazon.in, Flipkart.com                    45
Amazon.in, Flipkart.com, Paytm.com         25
Amazon.in, Paytm.com                       18
Paytm.com                                  15
Myntra.com, Snapdeal.com                   14
Snapdeal.com                               11
Name: Website is as efficient as before, dtype: int64
```



In [23]:
```
#Check for normal distribution and see how data is distributed for every column
plt.figure(figsize=(50,30), facecolor='yellow')#dimension and color
plotnumber = 1 #like a form of counter which we will see its relevance in the 'for loop'

for column in df.iloc[:,71:80]:#triggering your for loop
    if plotnumber<=10:
        ax = plt.subplot(2,5,plotnumber)#2,3 means arrange the graphs in 3-rows and 4-co
        sns.countplot(df[column],data=df)#performing a distribution plot for all the col
        plt.xlabel(column,fontsize=10)#labelling the x-axis with their respctive column
        print(df[column].value_counts())

    plotnumber+=1#incrementals for plot number so as to allow ploting one by one until i
plt.tight_layout()#This is just to show you data in a compact form.
```

```
<Figure size 3600x2160 with 0 Axes>
```

In [24]:
```
df.isnull().sum()
```

Out[24]:
```
1Gender of respondent                                  0
2 How old are you?                                     0
3 Which city do you shop online from?                  0
4 What is the Pin Code of where you shop online from?  0
5 Since How Long You are Shopping Online ?             0
```

```
                                                                    ..
Longer delivery period                                               0
Change in website/Application design                                 0
Frequent disruption when moving from one page to another             0
Website is as efficient as before                                    0
Which of the Indian online retailer would you recommend to a friend?    0
Length: 71, dtype: int64
```

We shall update:

- There are no null values in the dataset

Lets go!

In [25]:
```
#rechecing for null values
df.isnull().sum()
```

Out[25]:
```
1Gender of respondent                                                0
2 How old are you?                                                   0
3 Which city do you shop online from?                                0
4 What is the Pin Code of where you shop online from?                0
5 Since How Long You are Shopping Online ?                           0
                                                                    ..
Longer delivery period                                               0
Change in website/Application design                                 0
Frequent disruption when moving from one page to another             0
Website is as efficient as before                                    0
Which of the Indian online retailer would you recommend to a friend?    0
Length: 71, dtype: int64
```

Observations on Null Update

From the above we can see that there are no null values!

3.2 Encoding

- We cannot perform Exploratory Data Analysis on Non-numerical Data, so we have to first of all convert them to numerical data
- Here we shall adopt Ordinal Encoder to transform the all Non-numerical column into Numerical Data!

In [26]:
```
#initializing....
enc=OrdinalEncoder()
```

In [27]:
```
df
```

Out[27]:

	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 1 year?	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 operat system (OS your devi \t\t
0	Male	31-	Delhi	110009	Above 4	31-40	Dial-up	Desktop	Others	Window/wind

		40 years			years	times				Mo
1	Female	21-30 years	Delhi	110030	Above 4 years	41 times and above	Wi-Fi	Smartphone	4.7 inches	IOS/I
2	Female	21-30 years	Greater Noida	201308	3-4 years	41 times and above	Mobile Internet	Smartphone	5.5 inches	Andı
3	Male	21-30 years	Karnal	132001	3-4 years	Less than 10 times	Mobile Internet	Smartphone	5.5 inches	IOS/I
4	Female	21-30 years	Bangalore	530068	2-3 years	11-20 times	Wi-Fi	Smartphone	4.7 inches	IOS/I
...	...	...	...	...	...	...	...	...	...	
264	Female	21-30 years	Solan	173212	1-2 years	Less than 10 times	Mobile Internet	Smartphone	5.5 inches	Andı
265	Female	31-40 years	Ghaziabad	201008	1-2 years	31-40 times	Mobile Internet	Smartphone	Others	Andı
266	Female	41-50 yaers	Bangalore	560010	2-3 years	Less than 10 times	Mobile internet	Laptop	Others	Window/wind Mo
267	Female	Less than 20 years	Solan	173229	2-3 years	Less than 10 times	Wi-Fi	Smartphone	5.5 inches	Andı
268	Female	41-50 yaers	Ghaziabad	201009	2-3 years	31-40 times	Mobile Internet	Smartphone	5.5 inches	Andı

269 rows × 71 columns

In [28]:
```
df.dtypes
```

Out[28]:
```
1Gender of respondent                                                 object
2 How old are you?                                                    object
3 Which city do you shop online from?                                 object
4 What is the Pin Code of where you shop online from?                  int64
5 Since How Long You are Shopping Online ?                            object
                                                                       ...
Longer delivery period                                                object
Change in website/Application design                                  object
Frequent disruption when moving from one page to another              object
Website is as efficient as before                                     object
Which of the Indian online retailer would you recommend to a friend?  object
Length: 71, dtype: object
```

In [29]:
```
#Tranforming into numerical data
for i in df.columns:
    if df[i].dtypes=='object':
        df[i]=enc.fit_transform(df[i].values.reshape(-1,1))
```

In [30]:
```
df
```

Out[30]:

	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 1 year?	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	...	Lo (p
0	1.0	1.0	2.0	110009	3.0	2.0	0.0	0.0	3.0	2.0	...	
1	0.0	0.0	2.0	110030	3.0	3.0	3.0	2.0	0.0	1.0	...	
2	0.0	0.0	4.0	201308	2.0	3.0	1.0	2.0	2.0	0.0	...	
3	1.0	0.0	6.0	132001	2.0	5.0	1.0	2.0	2.0	1.0	...	
4	0.0	0.0	0.0	530068	1.0	0.0	3.0	2.0	0.0	1.0	...	
...	...	...	...	...	...	...	...	...	...	...	...	
264	0.0	0.0	10.0	173212	0.0	5.0	1.0	2.0	2.0	0.0	...	
265	0.0	1.0	3.0	201008	0.0	2.0	1.0	2.0	3.0	0.0	...	
266	0.0	2.0	0.0	560010	1.0	5.0	2.0	1.0	3.0	2.0	...	
267	0.0	4.0	10.0	173229	1.0	5.0	3.0	2.0	2.0	0.0	...	
268	0.0	2.0	3.0	201009	1.0	2.0	1.0	2.0	2.0	0.0	...	

269 rows × 71 columns

In [31]:
```
df.head()
```

Out[31]:

	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 1 year?	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	...	Long lo (pro
0	1.0	1.0	2.0	110009	3.0	2.0	0.0	0.0	3.0	2.0	...	
1	0.0	0.0	2.0	110030	3.0	3.0	3.0	2.0	0.0	1.0	...	
2	0.0	0.0	4.0	201308	2.0	3.0	1.0	2.0	2.0	0.0	...	
3	1.0	0.0	6.0	132001	2.0	5.0	1.0	2.0	2.0	1.0	...	
4	0.0	0.0	0.0	530068	1.0	0.0	3.0	2.0	0.0	1.0	...	

5 rows × 71 columns

In [32]:
```
df.tail()
```

Out[32]:

	1Gender of respondent	2 How	3 Which	4 What	5 Since How	6 How many	7 How do you	8 Which device do	9 What is the screen	10 What is the	...	Lo

	old are you?	city do you shop online from?	is the Pin Code of where you shop online from?	Long You are Shopping Online ?	times you have made an online purchase in the past 1 year?	access the internet while shopping on-line?	you use to access the online shopping?	size of your mobile device? \t\t\t\t\t\t	operating system (OS) of your device? \t\t\t\t	(p
264	0.0	0.0	10.0	173212	0.0	5.0	1.0	2.0	2.0	0.0	...
265	0.0	1.0	3.0	201008	0.0	2.0	1.0	2.0	3.0	0.0	...
266	0.0	2.0	0.0	560010	1.0	5.0	2.0	1.0	3.0	2.0	...
267	0.0	4.0	10.0	173229	1.0	5.0	3.0	2.0	2.0	0.0	...
268	0.0	2.0	3.0	201009	1.0	2.0	1.0	2.0	2.0	0.0	...

5 rows × 71 columns

In [33]:
```
df.sample(10)
```

Out[33]:

	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 1 year?	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	...	Lo (p
192	1.0	0.0	6.0	132001	2.0	5.0	1.0	2.0	2.0	1.0	...	
221	0.0	0.0	4.0	201310	3.0	2.0	2.0	1.0	3.0	2.0	...	
266	0.0	2.0	0.0	560010	1.0	5.0	2.0	1.0	3.0	2.0	...	
124	0.0	0.0	10.0	173212	0.0	5.0	1.0	2.0	2.0	0.0	...	
28	0.0	2.0	4.0	201306	3.0	2.0	2.0	1.0	3.0	2.0	...	
45	0.0	2.0	0.0	560037	3.0	3.0	2.0	2.0	0.0	1.0	...	
105	0.0	2.0	3.0	201009	1.0	2.0	1.0	2.0	2.0	0.0	...	
131	1.0	2.0	2.0	110018	2.0	5.0	1.0	0.0	3.0	2.0	...	
228	1.0	2.0	2.0	110044	1.0	5.0	2.0	2.0	2.0	1.0	...	
57	1.0	0.0	2.0	110044	4.0	5.0	2.0	1.0	3.0	2.0	...	

10 rows × 71 columns

In [34]:
```
df.describe()
```

Out[34]:

	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	7 How do you access the internet while	8 Which device do you use to access the online shopping?	9 Wh the sc si m

						purchase in the past 1 year?	shopping on-line?		de \t\t\t\
count	269.000000	269.000000	269.000000	269.000000	269.000000	269.000000	269.000000	269.000000	269.00
mean	0.327138	1.330855	4.494424	220465.747212	2.323420	3.237918	2.078067	1.501859	2.25
std	0.470042	1.183774	3.187687	140524.341051	1.176357	1.739331	0.715919	0.751240	0.94
min	0.000000	0.000000	0.000000	110008.000000	0.000000	0.000000	0.000000	0.000000	0.00
25%	0.000000	0.000000	2.000000	122018.000000	1.000000	2.000000	2.000000	1.000000	2.00
50%	0.000000	1.000000	4.000000	201303.000000	3.000000	3.000000	2.000000	2.000000	2.00
75%	1.000000	2.000000	7.000000	201310.000000	3.000000	5.000000	3.000000	2.000000	3.00
max	1.000000	4.000000	10.000000	560037.000000	4.000000	5.000000	3.000000	3.000000	3.00

8 rows × 71 columns

In [35]:
```
df.dtypes
```

Out[35]:
```
1Gender of respondent                                                 float64
2 How old are you?                                                    float64
3 Which city do you shop online from?                                 float64
4 What is the Pin Code of where you shop online from?                   int64
5 Since How Long You are Shopping Online ?                            float64
                                                                        ...
Longer delivery period                                                float64
Change in website/Application design                                  float64
Frequent disruption when moving from one page to another              float64
Website is as efficient as before                                     float64
Which of the Indian online retailer would you recommend to a friend?  float64
Length: 71, dtype: object
```

In [36]:
```
df['6 How many times you have made an online purchase in the past 1 year?']
```

Out[36]:
```
0      2.0
1      3.0
2      3.0
3      5.0
4      0.0
      ...
264    5.0
265    2.0
266    5.0
267    5.0
268    2.0
Name: 6 How many times you have made an online purchase in the past 1 year?, Length: 26
9, dtype: float64
```

Observations on Encoding

Frome the above its crystal clear that all Non-numrical data have been converted(encoded) to Numerical Data!

3.3 - Scatter Plot Check(Bivariate Analysis)

In [37]:
```
#extracting only the features(wihtout the Label) from the original dataset and store in
```

```
features=df.drop(columns=['6 How many times you have made an online purchase in the past
features
```

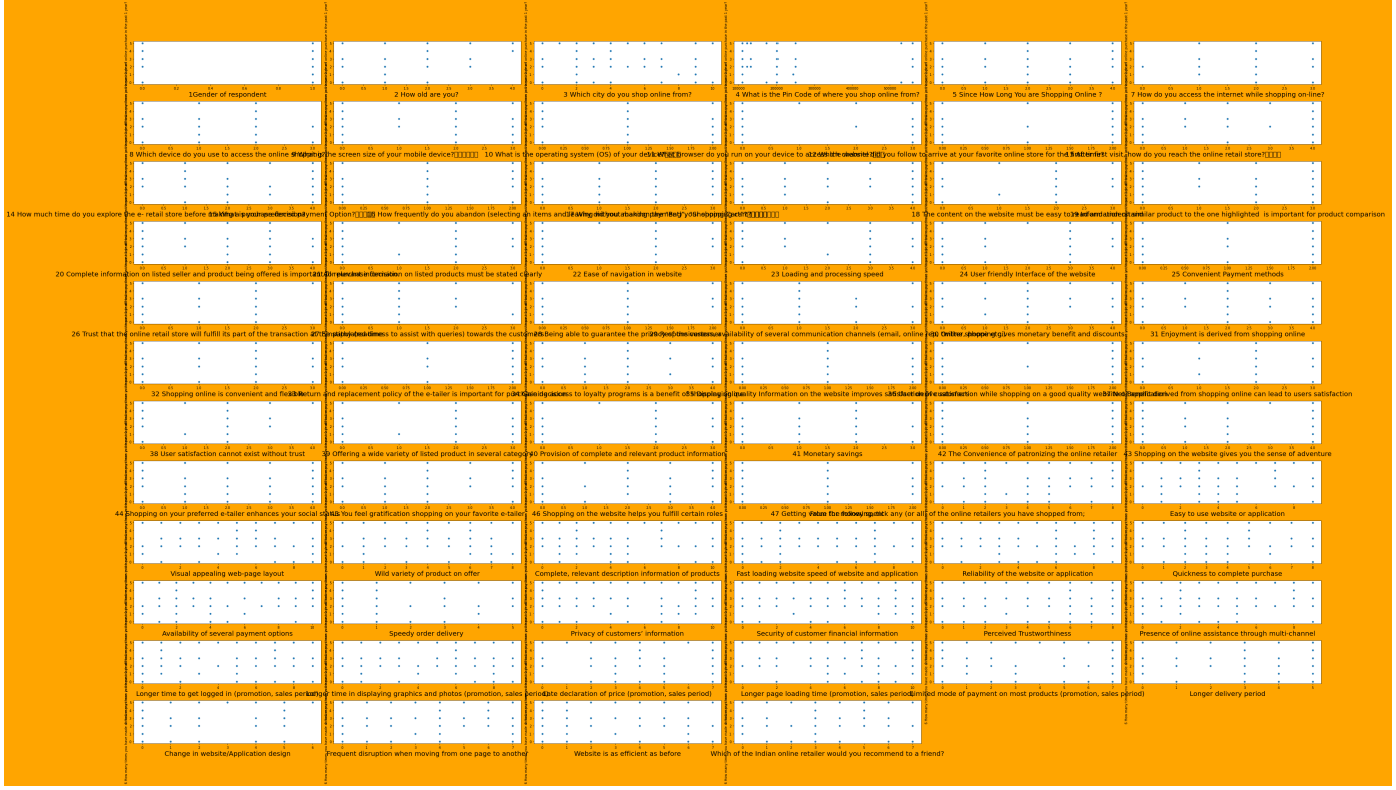
Out[37]:

	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 ?	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	...	Lo (p
0	1.0	1.0	2.0	110009	3.0	0.0	0.0	3.0	2.0	0.0	...	
1	0.0	0.0	2.0	110030	3.0	3.0	2.0	0.0	1.0	0.0	...	
2	0.0	0.0	4.0	201308	2.0	1.0	2.0	2.0	0.0	0.0	...	
3	1.0	0.0	6.0	132001	2.0	1.0	2.0	2.0	1.0	3.0	...	
4	0.0	0.0	0.0	530068	1.0	3.0	2.0	0.0	1.0	3.0	...	
...	...	...	...	...	...	...	...	...	...	...	...	
264	0.0	0.0	10.0	173212	0.0	1.0	2.0	2.0	0.0	2.0	...	
265	0.0	1.0	3.0	201008	0.0	1.0	2.0	3.0	0.0	0.0	...	
266	0.0	2.0	0.0	560010	1.0	2.0	1.0	3.0	2.0	0.0	...	
267	0.0	4.0	10.0	173229	1.0	3.0	2.0	2.0	0.0	0.0	...	
268	0.0	2.0	3.0	201009	1.0	1.0	2.0	2.0	0.0	0.0	...	

269 rows × 70 columns

In [38]:

```
#Check for the relatinship between each features and Target
plt.figure(figsize=(50,30), facecolor='orange')#dimension and line color
plotnumber = 1 #like a form of counter which we will see its relevance in the 'for loop'

for column in features:#triggering your for loop
    if plotnumber<=72:#since there are 3 features
        ax = plt.subplot(12,6,plotnumber)#3,4 means arrange the graphs in 3-rows and 3-c
        sns.scatterplot(x=features[column],y=df['6 How many times you have made an onlin
        plt.xlabel(column,fontsize=20)#labelling the x-axis with their respctive column
    plotnumber+=1#incrementals for plot number so as to allow ploting one by one until i
plt.tight_layout()#This is just to show you data in a compact form.
```



Observations on Scatter Plot

From the above scatter plot we can see a strong relationship between some of the features and the Label(Item_Outlet_Sales).

3.4 - Correlation Check(Collinearity and Multicollinearity)- Multivariate Analysis;

-The dataset is made up of only categorical data so we case of collinearity wont stand because its for continous values

In [39]:
```
#checking for general correlation
df.corr()
```

Out[39]:

	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 1 year?	7 How do you access the internet while shopping on-line?	8 Which device do you use to access the online shopping?	9 W the \t\t'
1Gender of respondent	1.000000	0.046169	0.080912	-0.260365	-0.057096	0.077876	-0.309029	0.061673	0.
2 How old are you?	0.046169	1.000000	0.113712	-0.120704	-0.087847	0.309575	0.255594	0.022383	-0.
3 Which city do you shop online from?	0.080912	0.113712	1.000000	-0.416597	-0.138329	0.173871	-0.010436	0.020650	0.

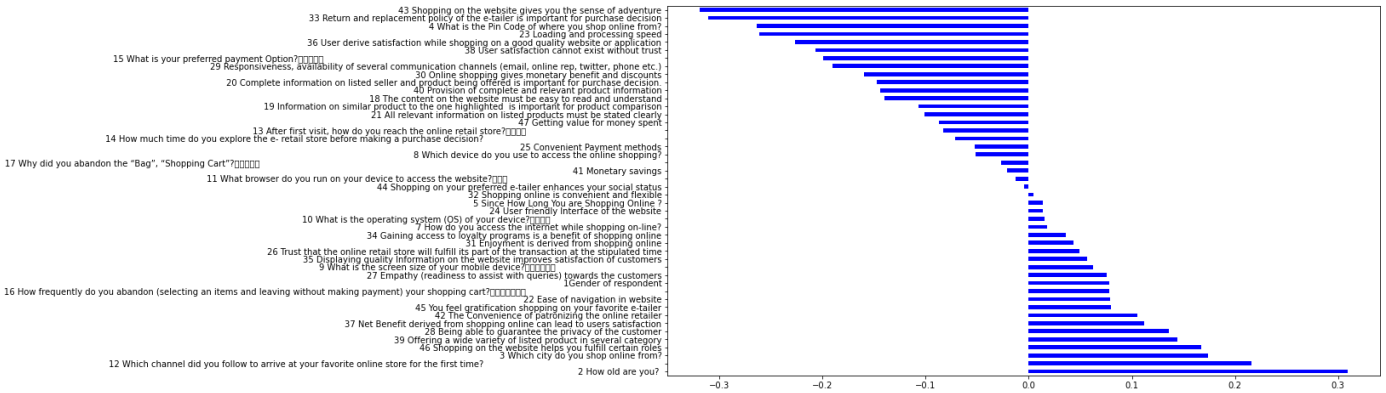
4 What is the Pin Code of where you shop online from?	-0.260365	-0.120704	-0.416597	1.000000	-0.090049	-0.263685	0.005511	-0.004980	-0.
5 Since How Long You are Shopping Online ?	-0.057096	-0.087847	-0.138329	-0.090049	1.000000	0.013315	0.226883	-0.125240	0.
...	...	...	...	...	...	...	...	...	
Longer delivery period	0.060838	-0.156173	-0.123369	-0.078660	0.218641	-0.130651	0.101297	-0.104665	0.
Change in website/Application design	-0.164818	-0.134558	0.000427	-0.058715	0.220347	0.007841	0.147770	-0.052146	0.
Frequent disruption when moving from one page to another	-0.256638	-0.018825	0.019167	0.039936	0.025919	-0.127148	0.349813	0.084876	-0.
Website is as efficient as before	0.055663	-0.008582	0.007117	0.037662	-0.024316	-0.124076	0.266932	0.333868	-0.
Which of the Indian online retailer would you recommend to a friend?	-0.003372	-0.135263	-0.142123	-0.045388	0.136106	-0.152028	0.041129	0.099425	0.

71 rows × 71 columns

In [40]:
```
#checking for correlation between each feature and Label
df.corr()['6 How many times you have made an online purchase in the past 1 year?'].sort_
```

Out[40]:
```
43 Shopping on the website gives you the sense of adventure
                                                                                -0.318657
33 Return and replacement policy of the e-tailer is important for purchase decision
                                                                                -0.310908
4 What is the Pin Code of where you shop online from?
                                                                                -0.263685
23 Loading and processing speed
                                                                                -0.261066
36 User derive satisfaction while shopping on a good quality website or application
                                                                                -0.226581


                                                                                   ...
46 Shopping on the website helps you fulfill certain roles
                                                                                 0.167532
3 Which city do you shop online from?
                                                                                 0.173871
12 Which channel did you follow to arrive at your favorite online store for the first ti
me?                                                                              0.215928
2 How old are you?
                                                                                 0.309575
6 How many times you have made an online purchase in the past 1 year?
                                                                                 1.000000
Name: 6 How many times you have made an online purchase in the past 1 year?, Length: 71,
dtype: float64
```

In [41]:
```
#Plotting a bar chart to visualize correlation through the stand-alone bars
plt.figure(figsize=(15,8))
df.iloc[:,0:47].corr()['6 How many times you have made an online purchase in the past 1
plt.show()
```

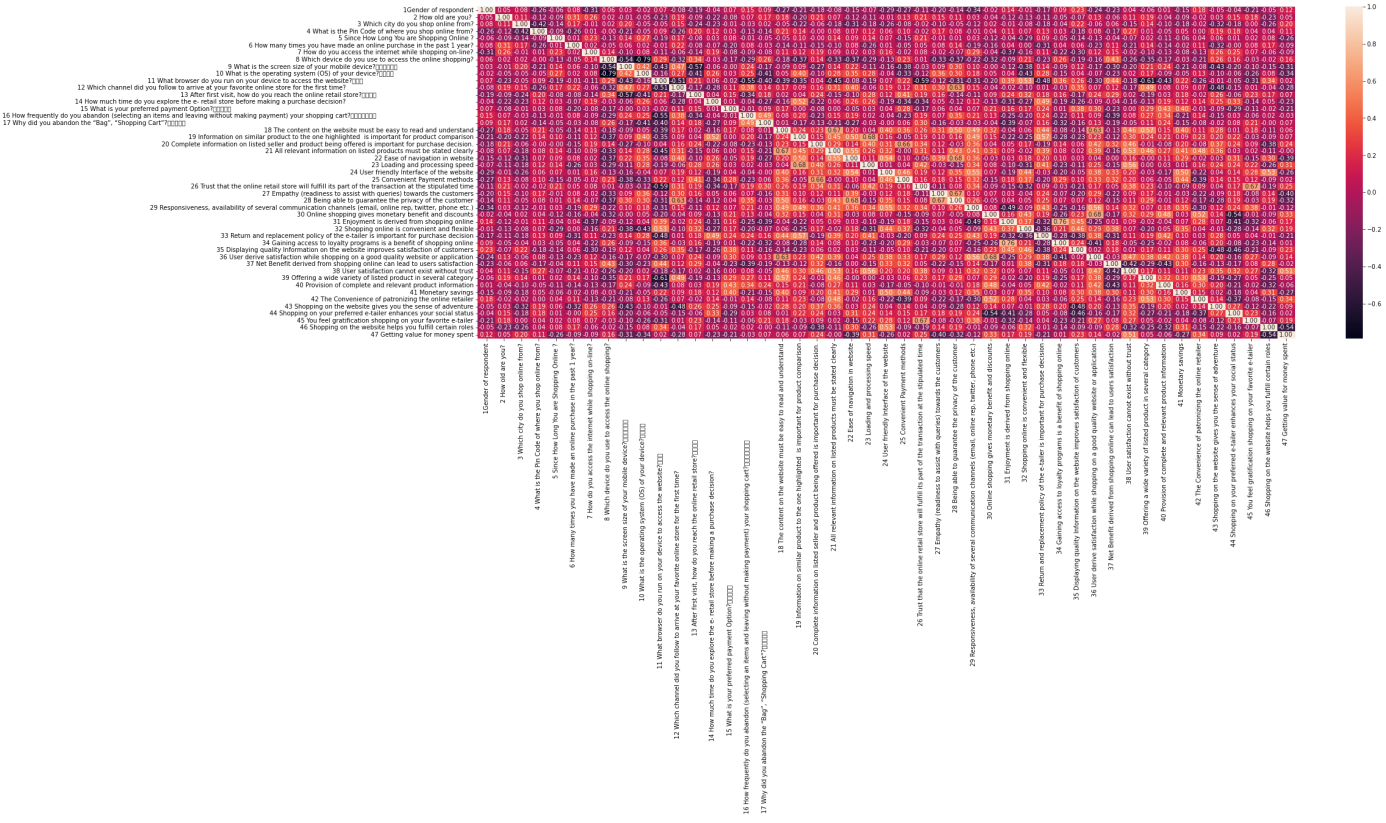


Bivariate analysis between categorical values

In [42]:
```
#Plotting a heatmap to show correlation through the brightness of the map-squares
plt.figure(figsize=(30,10))
sns.heatmap(df.iloc[:,0:47].corr(),annot=True, linewidth=0,linecolor='black',fmt='.2f')
```

Out[42]:

```
<AxesSubplot:>
```



Bivariate analysis on the categorical values which is features vs label

In [43]:
```
df.iloc[:,0:47].corr()['6 How many times you have made an online purchase in the past 1
```

Out[43]:

```
43 Shopping on the website gives you the sense of adventure

                                                    -0.318657
33 Return and replacement policy of the e-tailer is important for purchase decision

                                                    -0.310908
4 What is the Pin Code of where you shop online from?
```

-0.263685
23 Loading and processing speed

-0.261066
36 User derive satisfaction while shopping on a good quality website or application

-0.226581
38 User satisfaction cannot exist without trust

-0.206904
15 What is your preferred payment Option?\t\t\t\t\t

-0.199719
29 Responsiveness, availability of several communication channels (email, online rep, tw
itter, phone etc.)
-0.190372
30 Online shopping gives monetary benefit and discounts

-0.159635
20 Complete information on listed seller and product being offered is important for purc
hase decision.
-0.147507
40 Provision of complete and relevant product information

-0.144277
18 The content on the website must be easy to read and understand

-0.139930
19 Information on similar product to the one highlighted is important for product compa
rison
-0.107277
21 All relevant information on listed products must be stated clearly

-0.100875
47 Getting value for money spent

-0.087458
13 After first visit, how do you reach the online retail store?\t\t\t\t

-0.083088
14 How much time do you explore the e- retail store before making a purchase decision?

-0.071392
25 Convenient Payment methods

-0.052107
8 Which device do you use to access the online shopping?

-0.051741
17 Why did you abandon the “Bag”, “Shopping Cart”?\t\t\t\t\t

-0.026906
41 Monetary savings

-0.021391
11 What browser do you run on your device to access the website?\t\t\t

-0.012791
44 Shopping on your preferred e-tailer enhances your social status

-0.004694
32 Shopping online is convenient and flexible

0.004376
5 Since How Long You are Shopping Online ?

0.013315
24 User friendly Interface of the website

0.013640
10 What is the operating system (OS) of your device?\t\t\t\t

0.015316
7 How do you access the internet while shopping on-line?

0.017990
34 Gaining access to loyalty programs is a benefit of shopping online

0.036268
31 Enjoyment is derived from shopping online

0.043482
26 Trust that the online retail store will fulfill its part of the transaction at the st
ipulated time
0.049143
35 Displaying quality Information on the website improves satisfaction of customers

0.056794
9 What is the screen size of your mobile device?\t\t\t\t\t\t

0.062622
27 Empathy (readiness to assist with queries) towards the customers

0.075615
1Gender of respondent

0.077876
16 How frequently do you abandon (selecting an items and leaving without making payment)
your shopping cart?\t\t\t\t\t\t\t
0.078278
22 Ease of navigation in website

0.078981
45 You feel gratification shopping on your favorite e-tailer

0.079476
42 The Convenience of patronizing the online retailer

0.105267
37 Net Benefit derived from shopping online can lead to users satisfaction

0.112002
28 Being able to guarantee the privacy of the customer

0.135745
39 Offering a wide variety of listed product in several category

0.144087
46 Shopping on the website helps you fulfill certain roles

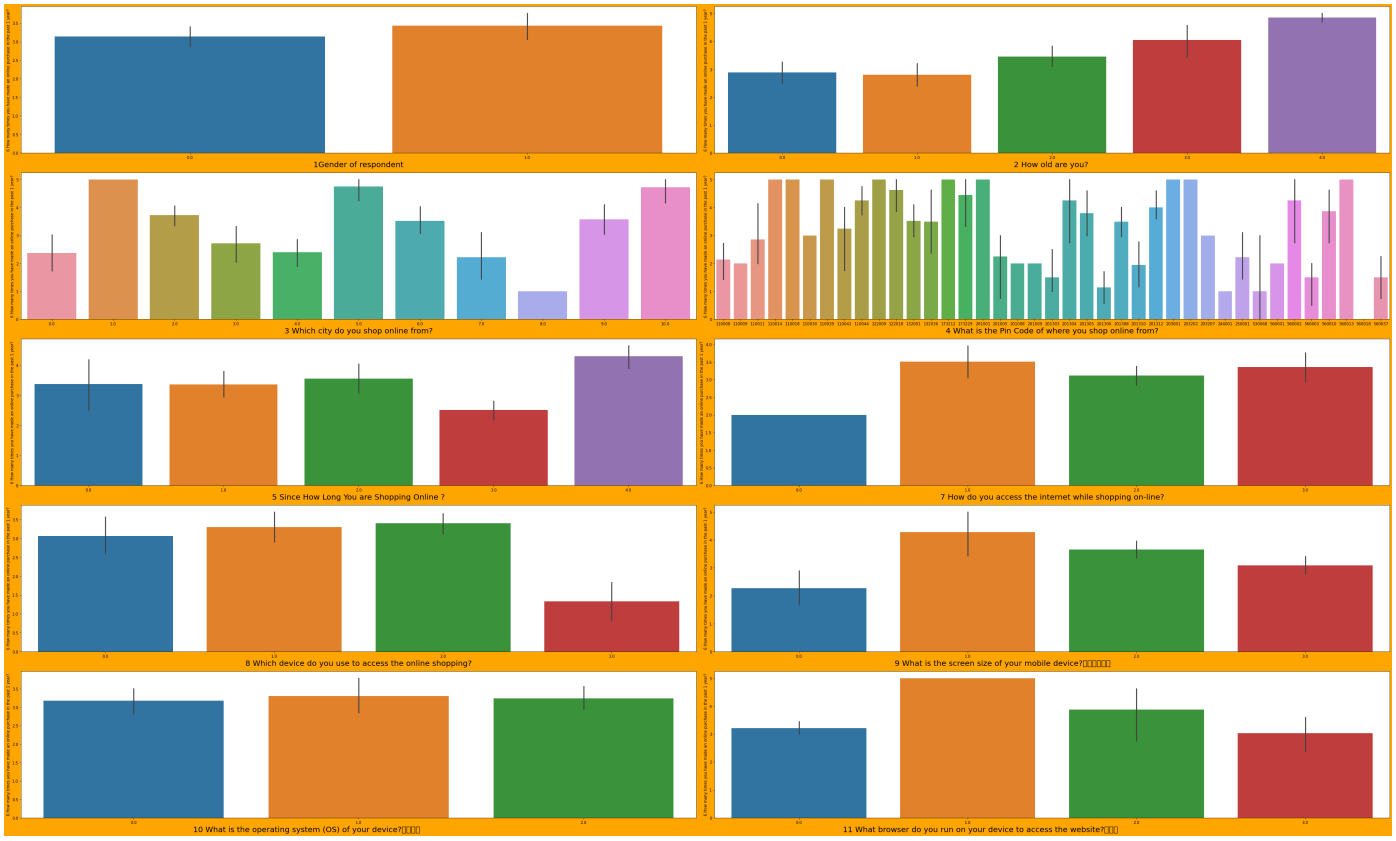
0.167532
3 Which city do you shop online from?

0.173871
12 Which channel did you follow to arrive at your favorite online store for the first ti
me?
0.215928
2 How old are you?

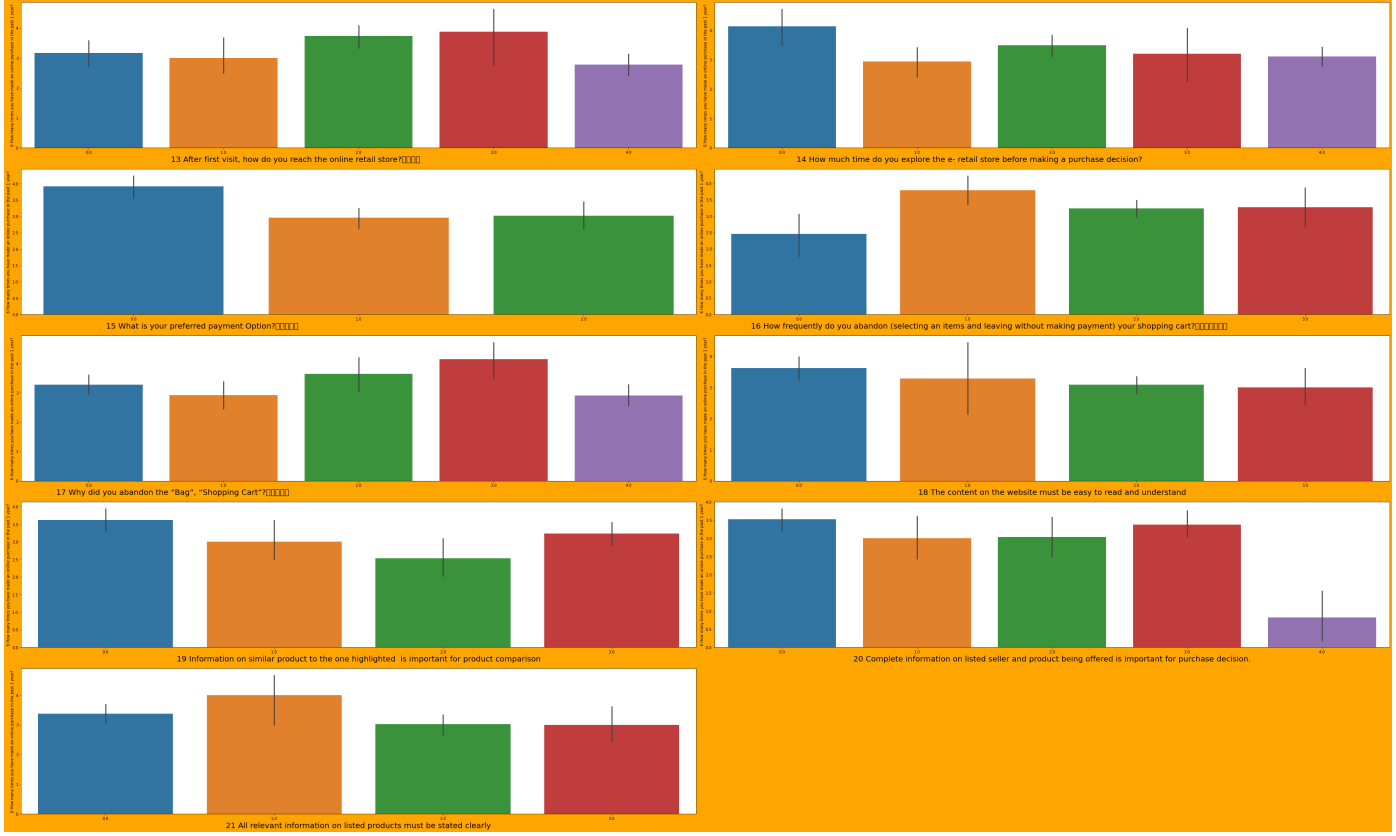
0.309575
6 How many times you have made an online purchase in the past 1 year?

```
1.000000
Name: 6 How many times you have made an online purchase in the past 1 year?, dtype: floa
t64
```

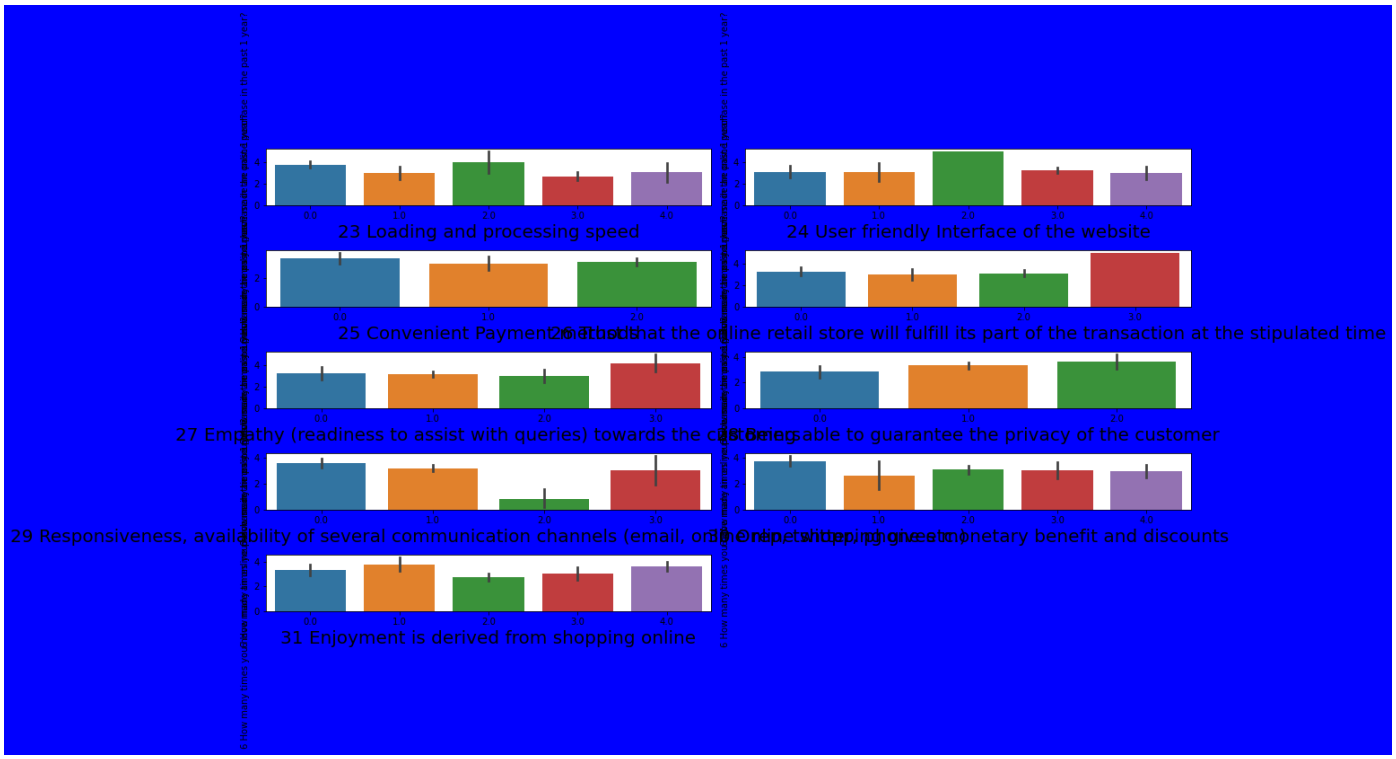
In [44]:
```
#Check for the relatinship between each features and Target
plt.figure(figsize=(50,30), facecolor='orange')#dimension and line color
plotnumber = 1 #like a form of counter which we will see its relevance in the 'for loop'

for column in features.iloc[:,0:10]:#triggering your for loop
    if plotnumber<=10:#since there are 3 features
        ax = plt.subplot(5,2,plotnumber)#3,4 means arrange the graphs in 3-rows and 3-co
        sns.barplot(x=features[column],y=df['6 How many times you have made an online pu
        plt.xlabel(column,fontsize=20)#labelling the x-axis with their respctive column
    plotnumber+=1#incrementals for plot number so as to allow ploting one by one until i
plt.tight_layout()#This is just to show you data in a compact form.
```



In [45]:
```
#Check for the relatinship between each features and Target
plt.figure(figsize=(50,30), facecolor='orange')#dimension and line color
plotnumber = 1 #like a form of counter which we will see its relevance in the 'for loop'

for column in features.iloc[:,11:20]:#triggering your for loop
    if plotnumber<=10:#since there are 3 features
        ax = plt.subplot(5,2,plotnumber)#3,4 means arrange the graphs in 3-rows and 3-co
        sns.barplot(x=features[column],y=df['6 How many times you have made an online pu
        plt.xlabel(column,fontsize=20)#labelling the x-axis with their respctive column
    plotnumber+=1#incrementals for plot number so as to allow ploting one by one until i
plt.tight_layout()#This is just to show you data in a compact form.
```

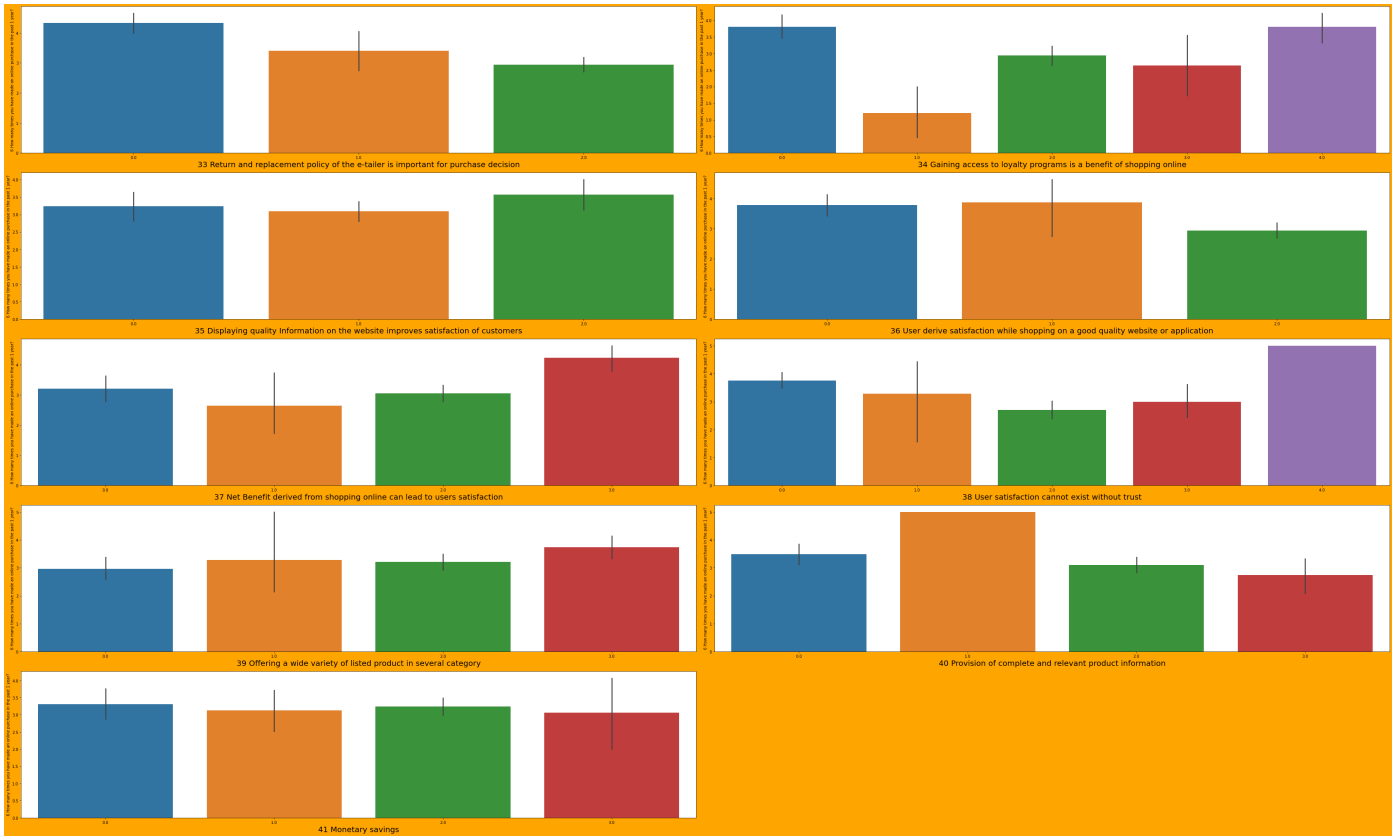


In [46]:
```
#Check for the relatinship between each features and Target
plt.figure(figsize=(15,8), facecolor='blue')#dimension and line color
plotnumber = 1 #like a form of counter which we will see its relevance in the 'for loop'

for column in features.iloc[:,21:30]:#triggering your for loop
    if plotnumber<=10:#since there are 3 features
        ax = plt.subplot(5,2,plotnumber)#3,4 means arrange the graphs in 3-rows and 3-co
        sns.barplot(x=features[column],y=df['6 How many times you have made an online pu
        plt.xlabel(column,fontsize=20)#labelling the x-axis with their respctive column
    plotnumber+=1#incrementals for plot number so as to allow ploting one by one until i
plt.tight_layout()#This is just to show you data in a compact form.
```



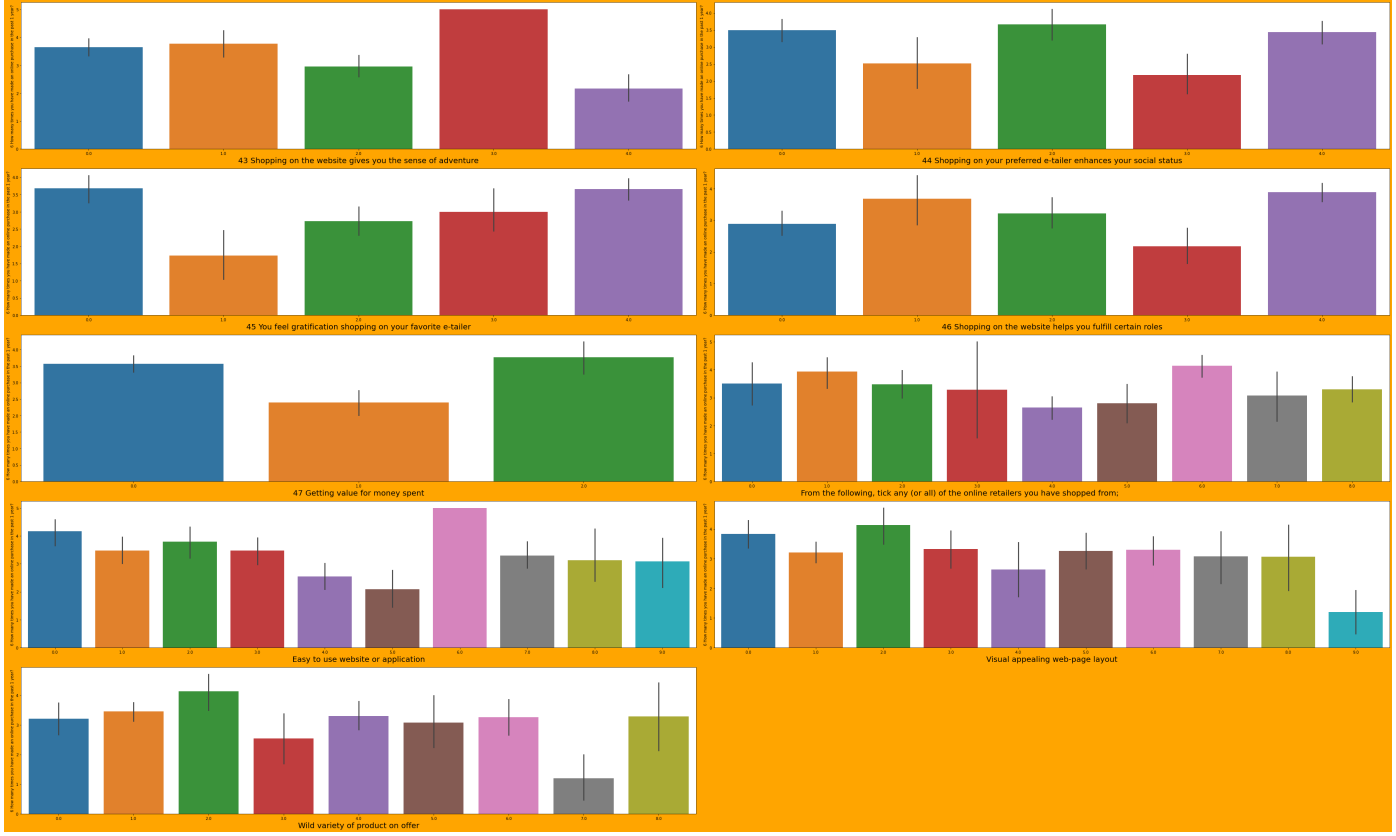
In [47]:
```
#Check for the relatinship between each features and Target
plt.figure(figsize=(50,30), facecolor='orange')#dimension and line color
```

```
plotnumber = 1 #like a form of counter which we will see its relevance in the 'for loop'

for column in features.iloc[:,31:40]:#triggering your for loop
    if plotnumber<=10:#since there are 3 features
        ax = plt.subplot(5,2,plotnumber)#3,4 means arrange the graphs in 3-rows and 3-co
        sns.barplot(x=features[column],y=df['6 How many times you have made an online pu
        plt.xlabel(column,fontsize=20)#labelling the x-axis with their respctive column
    plotnumber+=1#incrementals for plot number so as to allow ploting one by one until i
plt.tight_layout()#This is just to show you data in a compact form.
```



In [48]:

```
#Check for the relatinship between each features and Target
plt.figure(figsize=(50,30), facecolor='orange')#dimension and line color
plotnumber = 1 #like a form of counter which we will see its relevance in the 'for loop'

for column in features.iloc[:,41:50]:#triggering your for loop
    if plotnumber<=10:#since there are 3 features
        ax = plt.subplot(5,2,plotnumber)#3,4 means arrange the graphs in 3-rows and 3-co
        sns.barplot(x=features[column],y=df['6 How many times you have made an online pu
        plt.xlabel(column,fontsize=20)#labelling the x-axis with their respctive column
    plotnumber+=1#incrementals for plot number so as to allow ploting one by one until i
plt.tight_layout()#This is just to show you data in a compact form.
```



Observation based on the above visualisation

1. We have over 72 columns or features but we will select only 47 features an ignore the rest because it does not contribute to the target being customer retention

2.we will narrow down on the first 47 features and review its importance to customer retention or repurchase

Observation based on the features:

1. Gender of respondent': From the count plot we have more females than males making online purchases,see info from the count plot above
   ```
   Female    181
   Male       88
   ```
2. How old are you? ':As we can see from the above count plot most of the purchase came from people between the age of 31 to 50 years old.See the break down of the age limits as visualised from the count plot. 31-40 years 81 21-30 years 79 41-50 yaers 70 Less than 20 years 20 51 years and above 19

3. Which city do you shop online from?':As we can see we have more online purchases from delhi and the lowest figure in bulandshahr.

   ```
   Delhi            58
   Greater Noida    43
   Noida            40
   Bangalore        37
   ```

```
Karnal           27
Solan            18
Ghaziabad        18
Gurgaon          12
Merrut            9
Moradabad         5
Bulandshahr       2,
```

4. What is the Pin Code of where you shop online from?':The pin code is not really important,
5. Since How Long You are Shopping Online ?':As we can see we have more people who have shooped online for a longer period,

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

6. How many times you have made an online purchase in the past 1 year?:As we can see 114 people have made less purchase online in the last 1year

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

1. How do you access the internet while shopping on-line?:More people use their mobile phone to make online purchases.There is a need to pay attention to mobile devices

Mobile internet 142 Wi-Fi 76 Mobile Internet 47 Dial-up 4

1. Which device do you use to access the online shopping?:More people use their mobile phone to make online purchases.There is a need to pay attention to mobile devices Smartphone 141 Laptop 86 Desktop 30 Tablet 12

2. What is the screen size of your mobile device? Others 134 5.5 inches 99 4.7 inches 29 5 inches 7

3. What is the operating system (OS) of your device?\t\t\t\t ',

4. What browser do you run on your device to access the website?\t\t\t ',
5. Which channel did you follow to arrive at your favorite online store for the first time? ',
6. After first visit, how do you reach the online retail store?\t\t\t\t ',
7. How much time do you explore the e- retail store before making a purchase decision? ',
8. What is your preferred payment Option? ',
9. How frequently do you abandon (selecting an items and leaving without making payment) your shopping cart? ',
10. Why did you abandon the “Bag”, “Shopping Cart”? ',
11. The content on the website must be easy to read and understand',
12. Information on similar product to the one highlighted is important for product comparison',
13. Complete information on listed seller and product being offered is important for purchase decision.',
14. All relevant information on listed products must be stated clearly',
15. Ease of navigation in website', '23 Loading and processing speed',
16. User friendly Interface of the website',
17. Convenient Payment methods',
18. Trust that the online retail store will fulfill its part of the transaction at the stipulated time',
19. Empathy (readiness to assist with queries) towards the customers',
20. Being able to guarantee the privacy of the customer',

21. Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.)',
22. Online shopping gives monetary benefit and discounts',
23. Enjoyment is derived from shopping online',
24. Shopping online is convenient and flexible',
25. Return and replacement policy of the e-tailer is important for purchase decision',
26. Gaining access to loyalty programs is a benefit of shopping online',
27. Displaying quality Information on the website improves satisfaction of customers',
28. User derive satisfaction while shopping on a good quality website or application',
29. Net Benefit derived from shopping online can lead to users satisfaction',
30. User satisfaction cannot exist without trust',
31. Offering a wide variety of listed product in several category',
32. Provision of complete and relevant product information',
33. Monetary savings',
34. The Convenience of patronizing the online retailer',
35. Shopping on the website gives you the sense of adventure',
36. Shopping on your preferred e-tailer enhances your social status',
37. You feel gratification shopping on your favorite e-tailer',
38. Shopping on the website helps you fulfill certain roles',
39. Getting value for money spent',

Further observations:

As we can see from the above visualisations,features from 18 to 47 above contribute greatly to customer retention or repurchase as we can see from the count plot.More people agreed that these feature played a major factor to them coming back to the site to make more purchases online.

From the count plot visualisation we discovered that amazon was the people major choice based on the remaining factors from 47 to 72

Conclusion:

Based on the above analysis the features below contribute majorly to the customer retention or repurchase of any ecommerce site.

1. The content on the website must be easy to read and understand',
2. Information on similar product to the one highlighted is important for product comparison',
3. Complete information on listed seller and product being offered is important for purchase decision.',
4. All relevant information on listed products must be stated clearly',
5. Ease of navigation in website', '23 Loading and processing speed',
6. User friendly Interface of the website',

7. Convenient Payment methods',
8. Trust that the online retail store will fulfill its part of the transaction at the stipulated time',
9. Empathy (readiness to assist with queries) towards the customers',
10. Being able to guarantee the privacy of the customer',
11. Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.)',
12. Online shopping gives monetary benefit and discounts',
13. Enjoyment is derived from shopping online',
14. Shopping online is convenient and flexible',
15. Return and replacement policy of the e-tailer is important for purchase decision',
16. Gaining access to loyalty programs is a benefit of shopping online',
17. Displaying quality Information on the website improves satisfaction of customers',
18. User derive satisfaction while shopping on a good quality website or application',
19. Net Benefit derived from shopping online can lead to users satisfaction',
20. User satisfaction cannot exist without trust',
21. Offering a wide variety of listed product in several category',
22. Provision of complete and relevant product information',
23. Monetary savings',
24. The Convenience of patronizing the online retailer',
25. Shopping on the website gives you the sense of adventure',
26. Shopping on your preferred e-tailer enhances your social status',
27. You feel gratification shopping on your favorite e-tailer',
28. Shopping on the website helps you fulfill certain roles',
29. Getting value for money spent',