Exploratory Data Analysis

In [1]:

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

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

In [2]:

```
pip install openpyxl
```

```
Requirement already satisfied: openpyxl in c:\users\hema\anaconda3\lib\site-packages (3.0.7)
Requirement already satisfied: et-xmlfile in c:\users\hema\anaconda3\lib\site-packages (from openpyxl) (1.0.1)
Note: you may need to restart the kernel to use updated packages.
```

In [3]:

```
pip install nbconvert
```

```
Requirement already satisfied: nbconvert in c:\users\hema\anaconda3\lib\site-packages (6.0.7)
Requirement already satisfied: mistune<2,>=0.8.1 in c:\users\hema\anaconda3\lib\site-packages (from nbconvert) (0.8.
4)
Requirement already satisfied: nbformat>=4.4 in c:\users\hema\anaconda3\lib\site-packages (from nbconvert) (5.1.3)
Requirement already satisfied: pandocfilters>=1.4.1 in c:\users\hema\anaconda3\lib\site-packages (from nbconvert) (1
.4.3)
Requirement already satisfied: jinja2>=2.4 in c:\users\hema\anaconda3\lib\site-packages (from nbconvert) (2.11.3)
Requirement already satisfied: nbclient<0.6.0,>=0.5.0 in c:\users\hema\anaconda3\lib\site-packages (from nbconvert)
(0.5.3)
Requirement already satisfied: entrypoints>=0.2.2 in c:\users\hema\anaconda3\lib\site-packages (from nbconvert) (0.3
)
Requirement already satisfied: traitlets>=4.2 in c:\users\hema\anaconda3\lib\site-packages (from nbconvert) (5.0.5)
Requirement already satisfied: defusedxml in c:\users\hema\anaconda3\lib\site-packages (from nbconvert) (0.7.1)
Requirement already satisfied: jupyterlab-pygments in c:\users\hema\anaconda3\lib\site-packages (from nbconvert) (0.
1.2)
Requirement already satisfied: jupyter-core in c:\users\hema\anaconda3\lib\site-packages (from nbconvert) (4.7.1)
Requirement already satisfied: testpath in c:\users\hema\anaconda3\lib\site-packages (from nbconvert) (0.4.4)
Requirement already satisfied: pygments>=2.4.1 in c:\users\hema\anaconda3\lib\site-packages (from nbconvert) (2.8.1)
```

```
Requirement already satisfied: bleach in c:\users\hema\anaconda3\lib\site-packages (from nbconvert) (3.3.0)
Requirement already satisfied: MarkupSafe>=0.23 in c:\users\hema\anaconda3\lib\site-packages (from jinja2>=2.4->nbco
nvert) (1.1.1)
Requirement already satisfied: jupyter-client>=6.1.5 in c:\users\hema\anaconda3\lib\site-packages (from nbclient<0.6
.0,>=0.5.0->nbconvert) (6.1.12)
Requirement already satisfied: async-generator in c:\users\hema\anaconda3\lib\site-packages (from nbclient<0.6.0,>=0
.5.0->nbconvert) (1.10)
Requirement already satisfied: nest-asyncio in c:\users\hema\anaconda3\lib\site-packages (from nbclient<0.6.0,>=0.5.
0->nbconvert) (1.5.1)
Requirement already satisfied: pyzmq>=13 in c:\users\hema\anaconda3\lib\site-packages (from jupyter-client>=6.1.5->n
bclient<0.6.0,>=0.5.0->nbconvert) (20.0.0)
Requirement already satisfied: tornado>=4.1 in c:\users\hema\anaconda3\lib\site-packages (from jupyter-client>=6.1.5
->nbclient<0.6.0,>=0.5.0->nbconvert) (6.1)
Requirement already satisfied: python-dateutil>=2.1 in c:\users\hema\anaconda3\lib\site-packages (from jupyter-clien
t>=6.1.5->nbclient<0.6.0,>=0.5.0->nbconvert) (2.8.1)
Requirement already satisfied: pywin32>=1.0 in c:\users\hema\anaconda3\lib\site-packages (from jupyter-core->nbconve
rt) (227)
Requirement already satisfied: ipython-genutils in c:\users\hema\anaconda3\lib\site-packages (from nbformat>=4.4->nb
convert) (0.2.0)
Requirement already satisfied: jsonschema!=2.5.0,>=2.4 in c:\users\hema\anaconda3\lib\site-packages (from nbformat>=
4.4->nbconvert) (3.2.0)
Requirement already satisfied: attrs>=17.4.0 in c:\users\hema\anaconda3\lib\site-packages (from jsonschema!=2.5.0,>=
2.4->nbformat>=4.4->nbconvert) (20.3.0)
Requirement already satisfied: six>=1.11.0 in c:\users\hema\anaconda3\lib\site-packages (from jsonschema!=2.5.0,>=2.
4->nbformat>=4.4->nbconvert) (1.15.0)
Requirement already satisfied: setuptools in c:\users\hema\anaconda3\lib\site-packages (from jsonschema!=2.5.0,>=2.4
->nbformat>=4.4->nbconvert) (52.0.0.post20210125)
Requirement already satisfied: pyrsistent>=0.14.0 in c:\users\hema\anaconda3\lib\site-packages (from jsonschema!=2.5
.0,>=2.4->nbformat>=4.4->nbconvert) (0.17.3)
Requirement already satisfied: webencodings in c:\users\hema\anaconda3\lib\site-packages (from bleach->nbconvert) (0
.5.1)
Requirement already satisfied: packaging in c:\users\hema\anaconda3\lib\site-packages (from bleach->nbconvert) (20.9
)
Requirement already satisfied: pyparsing>=2.0.2 in c:\users\hema\anaconda3\lib\site-packages (from packaging->bleach
->nbconvert) (2.4.7)
Note: you may need to restart the kernel to use updated packages.
```

In [4]:

```
pip install nbconvert[webpdf]
```

```
Requirement already satisfied: nbconvert[webpdf] in c:\users\hema\anaconda3\lib\site-packages (6.0.7)
```

```
Requirement already satisfied: jupyterlab-pygments in c:\users\hema\anaconda3\lib\site-packages (from nbconvert[webp
df]) (0.1.2)
Requirement already satisfied: testpath in c:\users\hema\anaconda3\lib\site-packages (from nbconvert[webpdf]) (0.4.4
)
Requirement already satisfied: bleach in c:\users\hema\anaconda3\lib\site-packages (from nbconvert[webpdf]) (3.3.0)
Requirement already satisfied: mistune<2,>=0.8.1 in c:\users\hema\anaconda3\lib\site-packages (from nbconvert[webpdf
]) (0.8.4)
Requirement already satisfied: jinja2>=2.4 in c:\users\hema\anaconda3\lib\site-packages (from nbconvert[webpdf]) (2.
11.3)
Requirement already satisfied: traitlets>=4.2 in c:\users\hema\anaconda3\lib\site-packages (from nbconvert[webpdf])
(5.0.5)
Requirement already satisfied: pandocfilters>=1.4.1 in c:\users\hema\anaconda3\lib\site-packages (from nbconvert[web
pdf]) (1.4.3)
Requirement already satisfied: jupyter-core in c:\users\hema\anaconda3\lib\site-packages (from nbconvert[webpdf]) (4
.7.1)
Requirement already satisfied: nbformat>=4.4 in c:\users\hema\anaconda3\lib\site-packages (from nbconvert[webpdf]) (
5.1.3)
Requirement already satisfied: defusedxml in c:\users\hema\anaconda3\lib\site-packages (from nbconvert[webpdf]) (0.7
.1)
Requirement already satisfied: nbclient<0.6.0,>=0.5.0 in c:\users\hema\anaconda3\lib\site-packages (from nbconvert[w
ebpdf]) (0.5.3)
Requirement already satisfied: entrypoints>=0.2.2 in c:\users\hema\anaconda3\lib\site-packages (from nbconvert[webpd
f]) (0.3)
Requirement already satisfied: pygments>=2.4.1 in c:\users\hema\anaconda3\lib\site-packages (from nbconvert[webpdf])
(2.8.1)
Requirement already satisfied: pyppeteer==0.2.2 in c:\users\hema\anaconda3\lib\site-packages (from nbconvert[webpdf]
) (0.2.2)
Requirement already satisfied: urllib3<2.0.0,>=1.25.8 in c:\users\hema\anaconda3\lib\site-packages (from pyppeteer==
0.2.2->nbconvert[webpdf]) (1.26.4)
Requirement already satisfied: pyee<8.0.0,>=7.0.1 in c:\users\hema\anaconda3\lib\site-packages (from pyppeteer==0.2.
2->nbconvert[webpdf]) (7.0.4)
Requirement already satisfied: websockets<9.0,>=8.1 in c:\users\hema\anaconda3\lib\site-packages (from pyppeteer==0.
2.2->nbconvert[webpdf]) (8.1)
Requirement already satisfied: appdirs<2.0.0,>=1.4.3 in c:\users\hema\anaconda3\lib\site-packages (from pyppeteer==0
.2.2->nbconvert[webpdf]) (1.4.4)
Requirement already satisfied: tqdm<5.0.0,>=4.42.1 in c:\users\hema\anaconda3\lib\site-packages (from pyppeteer==0.2
.2->nbconvert[webpdf]) (4.59.0)
Requirement already satisfied: MarkupSafe>=0.23 in c:\users\hema\anaconda3\lib\site-packages (from jinja2>=2.4->nbco
nvert[webpdf]) (1.1.1)
```

```
Requirement already satisfied: jupyter-client>=6.1.5 in c:\users\hema\anaconda3\lib\site-packages (from nbclient<0.6
.0,>=0.5.0->nbconvert[webpdf]) (6.1.12)
Requirement already satisfied: nest-asyncio in c:\users\hema\anaconda3\lib\site-packages (from nbclient<0.6.0,>=0.5.
0->nbconvert[webpdf]) (1.5.1)
Requirement already satisfied: async-generator in c:\users\hema\anaconda3\lib\site-packages (from nbclient<0.6.0,>=0
.5.0->nbconvert[webpdf]) (1.10)
Requirement already satisfied: pyzmq>=13 in c:\users\hema\anaconda3\lib\site-packages (from jupyter-client>=6.1.5->n
bclient<0.6.0,>=0.5.0->nbconvert[webpdf]) (20.0.0)
Requirement already satisfied: tornado>=4.1 in c:\users\hema\anaconda3\lib\site-packages (from jupyter-client>=6.1.5
->nbclient<0.6.0,>=0.5.0->nbconvert[webpdf]) (6.1)
Requirement already satisfied: python-dateutil>=2.1 in c:\users\hema\anaconda3\lib\site-packages (from jupyter-clien
t>=6.1.5->nbclient<0.6.0,>=0.5.0->nbconvert[webpdf]) (2.8.1)
Requirement already satisfied: pywin32>=1.0 in c:\users\hema\anaconda3\lib\site-packages (from jupyter-core->nbconve
rt[webpdf]) (227)
Requirement already satisfied: jsonschema!=2.5.0,>=2.4 in c:\users\hema\anaconda3\lib\site-packages (from nbformat>=
4.4->nbconvert[webpdf]) (3.2.0)
Requirement already satisfied: ipython-genutils in c:\users\hema\anaconda3\lib\site-packages (from nbformat>=4.4->nb
convert[webpdf]) (0.2.0)
Requirement already satisfied: pyrsistent>=0.14.0 in c:\users\hema\anaconda3\lib\site-packages (from jsonschema!=2.5
.0,>=2.4->nbformat>=4.4->nbconvert[webpdf]) (0.17.3)
Requirement already satisfied: setuptools in c:\users\hema\anaconda3\lib\site-packages (from jsonschema!=2.5.0,>=2.4
->nbformat>=4.4->nbconvert[webpdf]) (52.0.0.post20210125)
Requirement already satisfied: six>=1.11.0 in c:\users\hema\anaconda3\lib\site-packages (from jsonschema!=2.5.0,>=2.
4->nbformat>=4.4->nbconvert[webpdf]) (1.15.0)
Requirement already satisfied: attrs>=17.4.0 in c:\users\hema\anaconda3\lib\site-packages (from jsonschema!=2.5.0,>=
2.4->nbformat>=4.4->nbconvert[webpdf]) (20.3.0)
Requirement already satisfied: packaging in c:\users\hema\anaconda3\lib\site-packages (from bleach->nbconvert[webpdf
]) (20.9)
Requirement already satisfied: webencodings in c:\users\hema\anaconda3\lib\site-packages (from bleach->nbconvert[web
pdf]) (0.5.1)
Requirement already satisfied: pyparsing>=2.0.2 in c:\users\hema\anaconda3\lib\site-packages (from packaging->bleach
->nbconvert[webpdf]) (2.4.7)
Note: you may need to restart the kernel to use updated packages.
```

In [5]:

```
file =('customer_retention_dataset.xlsx')
```

In [6]:

```
#extract the data from excel
```

In [7]:

```
newdata=pd.read_excel(file,sheet_name='datasheet')
```

In [8]:

```
newdata.head()
```

Out[8]:

	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	...	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?
0	Male	31-40 years	Delhi	110009	Above 4 years	31-40 times	Dial-up	Desktop	Others	Window/windows Mobile	...	Amazon.in	Amazon.in	Flipkart.com	Flipkart.com	Amazon.in	Paytm.com	Flipkart.com	Amazon.in	Amazon.in	Flipkart.com

	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	...	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?
1	Female	21-30 years	Delhi	110030	Above 4 years	41 times and above	Wi-Fi	Smartphone	4.7 inches	IOS/Mac	...	Amazon.in, Flipkart.com	Myntra.com	snapdeal.com	Snapdeal.com	Snapdeal.com	Snapdeal.com	Amazon.in	Myntra.com	Amazon.in, Flipkart.com	Amazon.in, Myntra.com
2	Female	21-30 ye	Greater Noida	201308	3-4 years	41 times and	Mobile Internet	Smartphone	5.5 inches	Android	...	Myntra.com	Myntra.com	Myntra.com	Myntra.com	Amazon.in	Paytm.com	Paytm.com	Paytm.com	Amazon.in	Amazon.in, Paytm.com,

	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	10 What is the operating system (OS) of your device? \t\t\t\t	...	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?
		ars				above															Myntra.com
3	Male	21-30 years	Karnal	132001	3-4 years	Less than 10 times	Mobile Internet	Smartphone	5.5 inches	IOS/Mac	...	Snapdeal.com	Myntra.com, Snapdeal.com	Myntra.com	Paytm.com	Paytm.com	Paytm.com	Amazon.in, Flipkart.com	Amazon.in, Flipkart.com	Amazon.in, Flipkart.com, Paytm.com	Amazon.in, Flipkart.com

	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	...	Longer time to get logged in (promotion, sales period)	Longer time in displaying graphics and photos (promotion, sales period)	Late declaration of price (promotion, sales period)	Longer page loading time (promotion, sales period)	Limited mode of payment on most products (promotion, sales period)	Longer delivery period	Change in website/Application design	Frequent disruption when moving from one page to another	Website is as efficient as before	Which of the Indian online retailer would you recommend to a friend?
4	Female	21-30 years	Bangalore	530068	2-3 years	11-20 times	Wi-Fi	Smartphone	4.7 inches	IOS/Mac	...	Flipkart.com, Paytm.com	Paytm.com	Paytm.com	Paytm.com	Snapdeal.com	Paytm.com	Amazon.in	Snapdeal.com	Paytm.com	Amazon.in, Myntra.com

5 rows × 71 columns

In [9]:

```
#Setting option to show max rows and max columns

pd.set_option('display.max_columns', None)
```

```
pd.set_option('display.max_rows', None)
```

Pre-processing the columns names

In [10]:

```
from string import digits
```

In [11]:

```
#Removing tab spaces
newdata.columns = newdata.columns.str.replace('/t','')
```

In [12]:

```
#Removing digits

remove_digits =str.maketrans('','',digits)
newdata.columns = newdata.columns.str.translate(remove_digits)
```

In [13]:

```
#Removing trailing and leading spaces

newdata.columns = newdata.columns.str.strip()
```

In [14]:

```
newdata.head()
```

Out[14]:

Gender of respondent	How old are you?	Which city do you shop online from?	What is the PIN Code of where you shop online from?	Since How Long are You Shopping Online?	How many times you have made an online purchase in t	How do you access the internet while shopping on-l	Which device do you use to access the online shoppi	What is the screen size of your mobile device?	What is the operating system (OS) of your device?	What browser do you run on your device to access the	Which channel did you follow to arrive at your favo	After first visit, how do you reach the online reta	How much time do you explore the e- retail store bef	What is your preferred payment Option?	How frequently do you abandon (selecting an items	Why did you abandon the “Bag”, “Shopping Cart”?	The content on the website must be easy to read and u	Information on similar product to the one highlig	Complete information on listed seller and produc	All relevant information on listed products must	Ease of navigation in website	Loading and processing speed	User friendly Interface of the website	Convenient Payment methods	Trust that the online retail store will fulfill it	Empathy (readiness to assist with queries) towar	Being able to guarantee the privacy of the custom	Responsiveness, availability of several commun	Online shopping gives monetary benefit and disco	Enjoyment is derived from shopping online	Shopping online is convenient and flexible	Return and replacement policy of the e-tailer is i	Gaining access to loyalty programs is a benefit of	Displaying quality Information on the website im	User derives satisfaction while shopping on a good	Net Benefit derived from shopping online can lead	User satisfaction cannot exist without trust	Offering a wide variety of listed products ever	Provision of complete and relevant product infor	Monetary savings	The Convenience of patronizing the online retail	Shopping on the website gives you the sense of adve	Shopping on your preferred e-tailer enhances you	You feel gratification shopping on your favorite	Shopping on the website helps you fulfill certain	Getting value for money spent	From the following, tick any (or all) of the online	Easy to use website or application	Visually appealing web-page layout	Wild variety of product on offer	Complete, relevant description information of p	Fast loading website speed of website and applica	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-cha	Longer time to get logged in (promotion, sales per

hepastyear?
ine?
ng?
website?
riteonlinestoreforthefirsttime?
ilstore?
oremakingapurchasedecision?
andleavingwithoutmakingpayment)yourshoppin
nderstand
htedisimportantforproductcomparison
tbeingofferedisimportantforpurchasedecisio
bestatedclearly
spartofthetransactionatthestipulatedtime
dsthecustomers
er
icationchannels(email,onlinerep,twitter,ph
unts
mportantforpurchasedecision
shoppingonline
provessatisfactionofcustomers
qualitywebsiteorapplication
tousersatisfaction
alcategory
mation
er
nture
rsocialstatus
e-tailer
roles
retailersyouhaveshoppedfrom;
roducts
tion
nnel
iod)Longertimeindisplayinggraphicsandphoto

gcart?

n.

oneetc.)

s(promotion,salesperiod)Latedeclarationofp rice(promotion,salesperiod)Longerpageloadi ngtime(promotion,salesperiod)Limitedmodeof payment on most products (promotion, sales perio

d) Longer delivery period Change in website/Appl icationdesignFrequentdisruptionwhenmovingf romonepagetoanotherWebsiteisasefficientasb eforeWhichoftheIndianonlineretailerwouldyo

																																																														urecommendtoafriend?
0	Male	31-40years	Delhi	110009	Above4years	31-40times	Dial-up	Desktop	Others	Window/windowsMobi	Googlechrome	SearchEngine	SearchEngine	6-10mins	E-wallets(Paytm,Free	Sometimes	Lackoftrust	Agree(4)	Indifferent(3)	Indifferent(3)	Agree(4)	Agree(4)	Stronglydisagree(	Dis-agree(2)	Dis-agree(2)	Disagree(2)	Stronglyagree(5)	Agree(4)	Agree(4)	indifferent(3)	Stronglydisagree(	Dis-agree(2)	Dis-agree(2)	Agree(4)	indifferent(3)	Stronglyagree(5)	Agree(4)	Stronglyagree(5)	indifferent(3)	indifferent(3)	Disagree(2)	indifferent(3)	Agree(4)	Agree(4)	Stronglyagree(5)	Agree(4)	Stronglyagree(5)	Amazon.in,Paytm.c	Paytm.com	Flipkart.com	Flipkart.com	Snapdeal.com	Snapdeal.com	Paytm.com	Paytm.com	Patym.com	Amazon.in	Amazon.in	Amazon.in	Flipkart.com	Paytm.com	Amazon.inAmazon.inFl

le	chargeetc.)	1)	1)	om	ipkart.comFlipkart.comAmazon.inPaytm.comFl

1	Female	21-30 years	Delhi	110030	Above 4 years	41 times and above	Wi-Fi	Smartphone	4.7 inches	IOS/Mac	Google chrome	Search Engine	Via application	more than 15 mins	Credit/Debit cards	Very frequently	Promo code not applicable	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	indifferent (3)	indifferent (3)	indifferent (3)	Strongly agree (5)	Strongly agree (5)	Strongly agree (5)	Amazon.in, Flipkart.com, Myntra.com, Snapdeal	Amazon.in, Flipkart.com, Myntra.com, Snapdeal	Amazon.in, Myntra.com	Flipkart.com, Myntra.com	Amazon.in, Flipkart.com, Myntra.com	Amazon.in, Flipkart.com, Myntra.com	Myntra.com	Amazon.com, Flipkart.com, Myntra.com	Amazon.in, Flipkart.com, Myntra.com	Amazon.in, Flipkart.com	Myntra.com	Myntra.com	Myntra.com	Amazon.in, Flipkart.com, Myntra.com	Amazon.in, Flipkart.com, Myntra.com, Snapdeal.c

on.inMyntra.comAmazon.in,Flipkart.comAmazo n.in, Myntra.com

2	Female	21-30 years	Greater Noida	201308	3-4 years	41 times and above	Mobile Internet	Smartphone	5.5 inches	Android	Google chrome	Search Engine	Via application	11-15 mins	E-wallets(Paytm, Freecharg	Sometimes	Promo code not applicable	Strongly agree(5)	Agree(4)	Agree(4)	Agree(4)	Agree(4)	Agree(4)	Strongly agree(5)	Strongly agree(5)	Strongly agree(5)	Strongly agree(5)	Strongly agree(5)	Strongly agree(5)	Strongly agree(5)	Strongly agree(5)	Strongly agree(5)	Strongly agree(5)	Strongly agree(5)	Strongly agree(5)	Strongly agree(5)	Strongly agree(5)	Agree(4)	Strongly agree(5)	Strongly agree(5)	Strongly agree(5)	Agree(4)	Agree(4)	Agree(4)	indifferent(3)	indifferent(3)	Agree(4)	Amazon.in, Paytm.com, Myntr	Amazon.in, Paytm.com, Myntr	Amazon.in, Paytm.com, Myntr	Amazon.in, Myntra.com	Amazon.in, Paytm.com, Myntr	Amazon.in, Paytm.com	Amazon.in, Paytm.com, Myntr	Amazon.com, Paytm.com, Mynt	Paytm.com, Myntra.com	Amazon.in	Amazon.in	Amazon.in, Paytm.com, Myntr	Amazon.in, Myntra.com	Myntra.com	Myntra.com Myntra.com Myntr

eetc.)

a.com

a.com

a.com

a.com

a.com

ra.com

a.com

a.comMyntra.comAmazon.inPaytm.comPaytm.com

Paytm.comAmazon.inAmazon.in,Paytm.com,Mynt

3	Male	21-30 years	Karnal	132001	3-4 years	Less than 10 times	Mobile Internet	Smartphone	5.5 inches	IOS/Mac	Safari	Search Engine	Search Engine	6-10 mins	Credit/Debit cards	Never	Better alternative offer	Agree (4)	Agree (4)	Indifferent (3)	Agree (4)	Strongly agree (5)	Agree (4)	Strongly agree (5)	Strongly agree (5)	Agree (4)	Strongly agree (5)	Strongly agree (5)	Agree (4)	Agree (4)	indifferent (3)	indifferent (3)	Agree (4)	indifferent (3)	indifferent (3)	Agree (4)	indifferent (3)	Agree (4)	Agree (4)	Agree (4)	Strongly agree (5)	Agree (4)	Agree (4)	Strongly agree (5)	Agree (4)	indifferent (3)	Agree (4)	Amazon.in, Flipkart.com, Paytm.com, M	Amazon.in, Flipkart.com, Paytm.com, M	Amazon.in, Flipkart.com, Paytm.com, M	Amazon.in, Flipkart.com	Amazon.in, Flipkart.com	Amazon.in, Flipkart.com, Snapdeal.c	Amazon.in, Flipkart.com, Paytm.com	Amazon.com, Flipkart.com, Paytm.c	Amazon.in, Flipkart.com, Myntra.c	Amazon.in, Flipkart.com, Snapdeal.c	Amazon.in, Flipkart.com, Myntra.c	Amazon.in, Flipkart.com, Snapdeal.c	Amazon.in, Flipkart.com, Snapdeal.c	Amazon.in, Flipkart.com, Myntra.com,	ra.com Snapdeal.comMyntra.com, Snapdeal.co

yntra.com...
yntra.com...
yntra.com...
om
om
om
om
om
om
om
Snapdeal
mMyntra.comPaytm.comPaytm.comPaytm.comAmaz

4	Female	21-30years	Bangalore	530068	2-3years	11-20times	Wi-Fi	Smartphone	4.7inches	IOS/Mac	Safari	ContentMarketing	Viaapplication	morethan15mins	Credit/Debitcards	Frequently	Betteralternativeoffer	Stronglyagree(5)	Indifferent(3)	Indifferent(3)	Agree(4)	Agree(4)	Stronglyagree(5)	Stronglyagree(5)	Stronglyagree(5)	Agree(4)	Stronglyagree(5)	Agree(4)	Stronglyagree(5)	Stronglyagree(5)	Stronglyagree(5)	Stronglyagree(5)	Stronglyagree(5)	Dis-agree(2)	Agree(4)	Stronglyagree(5)	Stronglyagree(5)	Stronglyagree(5)	Agree(4)	Agree(4)	Stronglyagree(5)	Agree(4)	indifferent(3)	Stronglydisagree(1)	Stronglyagree(5)	Stronglydisagree(1)	Stronglyagree(5)	Amazon.in,Flipkart.com,Paytm.c	Amazon.in,Flipkart.com,Paytm.c	Myntra.com	Myntra.com	Amazon.in,Flipkart.com,Paytm.c	Amazon.in	Amazon.in,Paytm.com,Myntra.c	Amazon.com,Flipkart.com,Paytm.	Amazon.in,Flipkart.com,Paytm.c	Amazon.in	Amazon.in,Paytm.com	Paytm.com	Amazon.in,Myntra.com	Amazon.in,Myntra.com	ipkart.com Flipkart.com,Paytm.comPaytm.co

om,Myntra.com...

om,Myntra.com...

om,Myntra.com...

om

com,Myntra.co...

om,Myntra.com...

mPaytm.comPaytm.comSnapdeal.comPaytm.comAm azon.inSnapdeal.comPaytm.comAmazon.in,Mynt

In [15]:

```
newdata.shape
```

Out[15]:

```
(269, 71)
```

In [16]:

```
#check each data type of columns and missing values
newdata.dtypes
```

Out[16]:

```
Gender of respondent                                                                                    object
How old are you?                                                                                        object
Which city do you shop online from?                                                                     object
What is the Pin Code of where you shop online from?                                                      int64
Since How Long You are Shopping Online ?                                                                object
How many times you have made an online purchase in the past  year?                                      object
How do you access the internet while shopping on-line?                                                  object
Which device do you use to access the online shopping?                                                  object
What is the screen size of your mobile device?                                                          object
What is the operating system (OS) of your device?                                                       object
What browser do you run on your device to access the website?                                           object
Which channel did you follow to arrive at your favorite online store for the first time?                object
After first visit, how do you reach the online retail store?                                            object
How much time do you explore the e- retail store before making a purchase decision?                     object
What is your preferred payment Option?                                                                  object
How frequently do you abandon (selecting an items and leaving without making payment) your shopping cart?    object
Why did you abandon the “Bag”, “Shopping Cart”?                                                         object
The content on the website must be easy to read and understand                                          object
Information on similar product to the one highlighted  is important for product comparison              object
Complete information on listed seller and product being offered is important for purchase decision.     object
All relevant information on listed products must be stated clearly                                      object
Ease of navigation in website                                                                           object
Loading and processing speed                                                                            object
User friendly Interface of the website                                                                  object
Convenient Payment methods                                                                              object
Trust that the online retail store will fulfill its part of the transaction at the stipulated time     object
Empathy (readiness to assist with queries) towards the customers                                        object
Being able to guarantee the privacy of the customer                                                     object
Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.)      object
```

Online shopping gives monetary benefit and discounts	object
Enjoyment is derived from shopping online	object
Shopping online is convenient and flexible	object
Return and replacement policy of the e-tailer is important for purchase decision	object
Gaining access to loyalty programs is a benefit of shopping online	object
Displaying quality Information on the website improves satisfaction of customers	object
User derive satisfaction while shopping on a good quality website or application	object
Net Benefit derived from shopping online can lead to users satisfaction	object
User satisfaction cannot exist without trust	object
Offering a wide variety of listed product in several category	object
Provision of complete and relevant product information	object
Monetary savings	object
The Convenience of patronizing the online retailer	object
Shopping on the website gives you the sense of adventure	object
Shopping on your preferred e-tailer enhances your social status	object
You feel gratification shopping on your favorite e-tailer	object
Shopping on the website helps you fulfill certain roles	object
Getting value for money spent	object
From the following, tick any (or all) of the online retailers you have shopped from;	object
Easy to use website or application	object
Visual appealing web-page layout	object
Wild variety of product on offer	object
Complete, relevant description information of products	object
Fast loading website speed of website and application	object
Reliability of the website or application	object
Quickness to complete purchase	object
Availability of several payment options	object
Speedy order delivery	object
Privacy of customers' information	object
Security of customer financial information	object
Perceived Trustworthiness	object
Presence of online assistance through multi-channel	object
Longer time to get logged in (promotion, sales period)	object
Longer time in displaying graphics and photos (promotion, sales period)	object
Late declaration of price (promotion, sales period)	object
Longer page loading time (promotion, sales period)	object
Limited mode of payment on most products (promotion, sales period)	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
dtype: object
```

In [17]:

```
newdata.isnull().sum()
```

Out[17]:

```
Gender of respondent                                                                                      0
How old are you?                                                                                          0
Which city do you shop online from?                                                                       0
What is the Pin Code of where you shop online from?                                                       0
Since How Long You are Shopping Online ?                                                                  0
How many times you have made an online purchase in the past  year?                                        0
How do you access the internet while shopping on-line?                                                    0
Which device do you use to access the online shopping?                                                    0
What is the screen size of your mobile device?                                                            0
What is the operating system (OS) of your device?                                                         0
What browser do you run on your device to access the website?                                             0
Which channel did you follow to arrive at your favorite online store for the first time?                  0
After first visit, how do you reach the online retail store?                                              0
How much time do you explore the e- retail store before making a purchase decision?                       0
What is your preferred payment Option?                                                                    0
How frequently do you abandon (selecting an items and leaving without making payment) your shopping cart?   0
Why did you abandon the “Bag”, “Shopping Cart”?                                                           0
The content on the website must be easy to read and understand                                            0
Information on similar product to the one highlighted  is important for product comparison                0
Complete information on listed seller and product being offered is important for purchase decision.       0
All relevant information on listed products must be stated clearly                                        0
Ease of navigation in website                                                                             0
Loading and processing speed                                                                              0
User friendly Interface of the website                                                                    0
Convenient Payment methods                                                                                0
Trust that the online retail store will fulfill its part of the transaction at the stipulated time        0
Empathy (readiness to assist with queries) towards the customers                                          0
Being able to guarantee the privacy of the customer                                                       0
Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.)   0
Online shopping gives monetary benefit and discounts                                                      0
Enjoyment is derived from shopping online                                                                 0
```

Shopping online is convenient and flexible	0
Return and replacement policy of the e-tailer is important for purchase decision	0
Gaining access to loyalty programs is a benefit of shopping online	0
Displaying quality Information on the website improves satisfaction of customers	0
User derive satisfaction while shopping on a good quality website or application	0
Net Benefit derived from shopping online can lead to users satisfaction	0
User satisfaction cannot exist without trust	0
Offering a wide variety of listed product in several category	0
Provision of complete and relevant product information	0
Monetary savings	0
The Convenience of patronizing the online retailer	0
Shopping on the website gives you the sense of adventure	0
Shopping on your preferred e-tailer enhances your social status	0
You feel gratification shopping on your favorite e-tailer	0
Shopping on the website helps you fulfill certain roles	0
Getting value for money spent	0
From the following, tick any (or all) of the online retailers you have shopped from;	0
Easy to use website or application	0
Visual appealing web-page layout	0
Wild variety of product on offer	0
Complete, relevant description information of products	0
Fast loading website speed of website and application	0
Reliability of the website or application	0
Quickness to complete purchase	0
Availability of several payment options	0
Speedy order delivery	0
Privacy of customers’ information	0
Security of customer financial information	0
Perceived Trustworthiness	0
Presence of online assistance through multi-channel	0
Longer time to get logged in (promotion, sales period)	0
Longer time in displaying graphics and photos (promotion, sales period)	0
Late declaration of price (promotion, sales period)	0
Longer page loading time (promotion, sales period)	0
Limited mode of payment on most products (promotion, sales period)	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
dtype: int64
```

In [18]:

```
newdata.nunique()
```

Out[18]:

```
Gender of respondent                                                                                       2
How old are you?                                                                                           5
Which city do you shop online from?                                                                       11
What is the Pin Code of where you shop online from?                                                       39
Since How Long You are Shopping Online ?                                                                   5
How many times you have made an online purchase in the past  year?                                         6
How do you access the internet while shopping on-line?                                                     4
Which device do you use to access the online shopping?                                                     4
What is the screen size of your mobile device?                                                             4
What is the operating system (OS) of your device?                                                          3
What browser do you run on your device to access the website?                                              4
Which channel did you follow to arrive at your favorite online store for the first time?                   3
After first visit, how do you reach the online retail store?                                               5
How much time do you explore the e- retail store before making a purchase decision?                        5
What is your preferred payment Option?                                                                     3
How frequently do you abandon (selecting an items and leaving without making payment) your shopping cart?  4
Why did you abandon the “Bag”, “Shopping Cart”?                                                            5
The content on the website must be easy to read and understand                                             4
Information on similar product to the one highlighted  is important for product comparison                 4
Complete information on listed seller and product being offered is important for purchase decision.        5
All relevant information on listed products must be stated clearly                                         4
Ease of navigation in website                                                                              4
Loading and processing speed                                                                               5
User friendly Interface of the website                                                                     5
Convenient Payment methods                                                                                 3
Trust that the online retail store will fulfill its part of the transaction at the stipulated time         4
Empathy (readiness to assist with queries) towards the customers                                           4
Being able to guarantee the privacy of the customer                                                        3
Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.)    4
Online shopping gives monetary benefit and discounts                                                       5
Enjoyment is derived from shopping online                                                                  5
Shopping online is convenient and flexible                                                                 4
Return and replacement policy of the e-tailer is important for purchase decision                           3
```

```
Gaining access to loyalty programs is a benefit of shopping online                                   5
Displaying quality Information on the website improves satisfaction of customers                     3
User derive satisfaction while shopping on a good quality website or application                     3
Net Benefit derived from shopping online can lead to users satisfaction                              4
User satisfaction cannot exist without trust                                                         5
Offering a wide variety of listed product in several category                                        4
Provision of complete and relevant product information                                               4
Monetary savings                                                                                     4
The Convenience of patronizing the online retailer                                                   3
Shopping on the website gives you the sense of adventure                                             5
Shopping on your preferred e-tailer enhances your social status                                      5
You feel gratification shopping on your favorite e-tailer                                            5
Shopping on the website helps you fulfill certain roles                                              5
Getting value for money spent                                                                        3
From the following, tick any (or all) of the online retailers you have shopped from;                 9
Easy to use website or application                                                                  10
Visual appealing web-page layout                                                                    10
Wild variety of product on offer                                                                     9
Complete, relevant description information of products                                              11
Fast loading website speed of website and application                                               10
Reliability of the website or application                                                           10
Quickness to complete purchase                                                                       9
Availability of several payment options                                                             11
Speedy order delivery                                                                                6
Privacy of customers’ information                                                                   11
Security of customer financial information                                                          11
Perceived Trustworthiness                                                                            9
Presence of online assistance through multi-channel                                                 10
Longer time to get logged in (promotion, sales period)                                              10
Longer time in displaying graphics and photos (promotion, sales period)                             10
Late declaration of price (promotion, sales period)                                                  8
Longer page loading time (promotion, sales period)                                                  11
Limited mode of payment on most products (promotion, sales period)                                   8
Longer delivery period                                                                               6
Change in website/Application design                                                                 7
Frequent disruption when moving from one page to another                                             8
Website is as efficient as before                                                                    8
Which of the Indian online retailer would you recommend to a friend?                                 8
dtype: int64
```

In [19]:

```
newdata.isnull().sum()/newdata.shape[0]*100
```

Out[19]:

```
Gender of respondent                                                                                      0.0
How old are you?                                                                                          0.0
Which city do you shop online from?                                                                       0.0
What is the Pin Code of where you shop online from?                                                       0.0
Since How Long You are Shopping Online ?                                                                  0.0
How many times you have made an online purchase in the past  year?                                        0.0
How do you access the internet while shopping on-line?                                                    0.0
Which device do you use to access the online shopping?                                                    0.0
What is the screen size of your mobile device?                                                            0.0
What is the operating system (OS) of your device?                                                         0.0
What browser do you run on your device to access the website?                                             0.0
Which channel did you follow to arrive at your favorite online store for the first time?                  0.0
After first visit, how do you reach the online retail store?                                              0.0
How much time do you explore the e- retail store before making a purchase decision?                       0.0
What is your preferred payment Option?                                                                    0.0
How frequently do you abandon (selecting an items and leaving without making payment) your shopping cart?   0.0
Why did you abandon the “Bag”, “Shopping Cart”?                                                           0.0
The content on the website must be easy to read and understand                                            0.0
Information on similar product to the one highlighted  is important for product comparison                0.0
Complete information on listed seller and product being offered is important for purchase decision.      0.0
All relevant information on listed products must be stated clearly                                        0.0
Ease of navigation in website                                                                             0.0
Loading and processing speed                                                                              0.0
User friendly Interface of the website                                                                    0.0
Convenient Payment methods                                                                                0.0
Trust that the online retail store will fulfill its part of the transaction at the stipulated time        0.0
Empathy (readiness to assist with queries) towards the customers                                          0.0
Being able to guarantee the privacy of the customer                                                       0.0
Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.)   0.0
Online shopping gives monetary benefit and discounts                                                      0.0
Enjoyment is derived from shopping online                                                                 0.0
Shopping online is convenient and flexible                                                                0.0
Return and replacement policy of the e-tailer is important for purchase decision                          0.0
Gaining access to loyalty programs is a benefit of shopping online                                        0.0
Displaying quality Information on the website improves satisfaction of customers                          0.0
```

```
User derive satisfaction while shopping on a good quality website or application                     0.0
Net Benefit derived from shopping online can lead to users satisfaction                              0.0
User satisfaction cannot exist without trust                                                         0.0
Offering a wide variety of listed product in several category                                        0.0
Provision of complete and relevant product information                                               0.0
Monetary savings                                                                                     0.0
The Convenience of patronizing the online retailer                                                   0.0
Shopping on the website gives you the sense of adventure                                             0.0
Shopping on your preferred e-tailer enhances your social status                                      0.0
You feel gratification shopping on your favorite e-tailer                                            0.0
Shopping on the website helps you fulfill certain roles                                              0.0
Getting value for money spent                                                                        0.0
From the following, tick any (or all) of the online retailers you have shopped from;                 0.0
Easy to use website or application                                                                   0.0
Visual appealing web-page layout                                                                     0.0
Wild variety of product on offer                                                                     0.0
Complete, relevant description information of products                                               0.0
Fast loading website speed of website and application                                                0.0
Reliability of the website or application                                                            0.0
Quickness to complete purchase                                                                       0.0
Availability of several payment options                                                              0.0
Speedy order delivery                                                                                0.0
Privacy of customers’ information                                                                    0.0
Security of customer financial information                                                           0.0
Perceived Trustworthiness                                                                            0.0
Presence of online assistance through multi-channel                                                  0.0
Longer time to get logged in (promotion, sales period)                                               0.0
Longer time in displaying graphics and photos (promotion, sales period)                              0.0
Late declaration of price (promotion, sales period)                                                  0.0
Longer page loading time (promotion, sales period)                                                   0.0
Limited mode of payment on most products (promotion, sales period)                                   0.0
Longer delivery period                                                                               0.0
Change in website/Application design                                                                 0.0
Frequent disruption when moving from one page to another                                             0.0
Website is as efficient as before                                                                    0.0
Which of the Indian online retailer would you recommend to a friend?                                 0.0
dtype: float64
```

Univariate Analysis

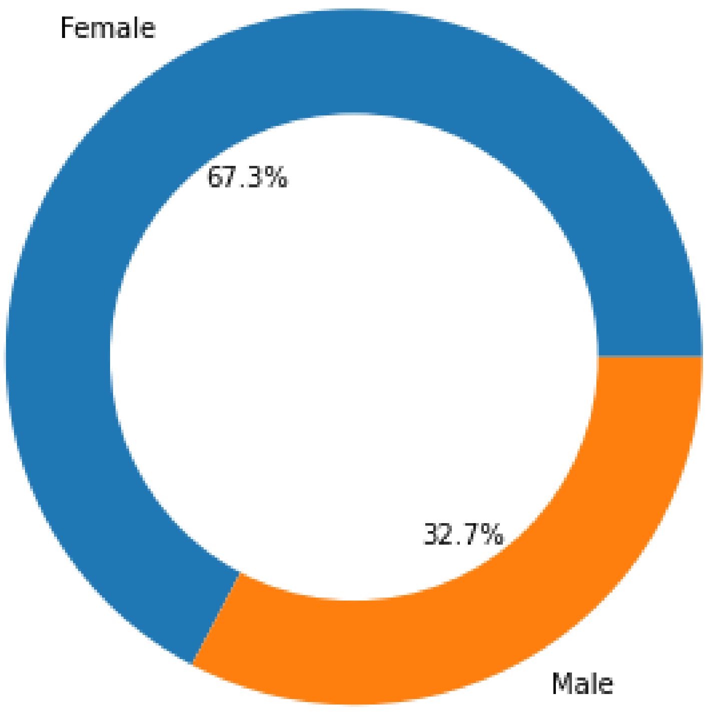
In [20]:

```
personal_info=['Gender of respondent','How old are you?','Which city do you shop online from?',
               'What is the Pin Code of where you shop online from?','Since How Long You are Shopping Online ?',
                   'How many times you have made an online purchase in the past  year?']
```

Personal Info

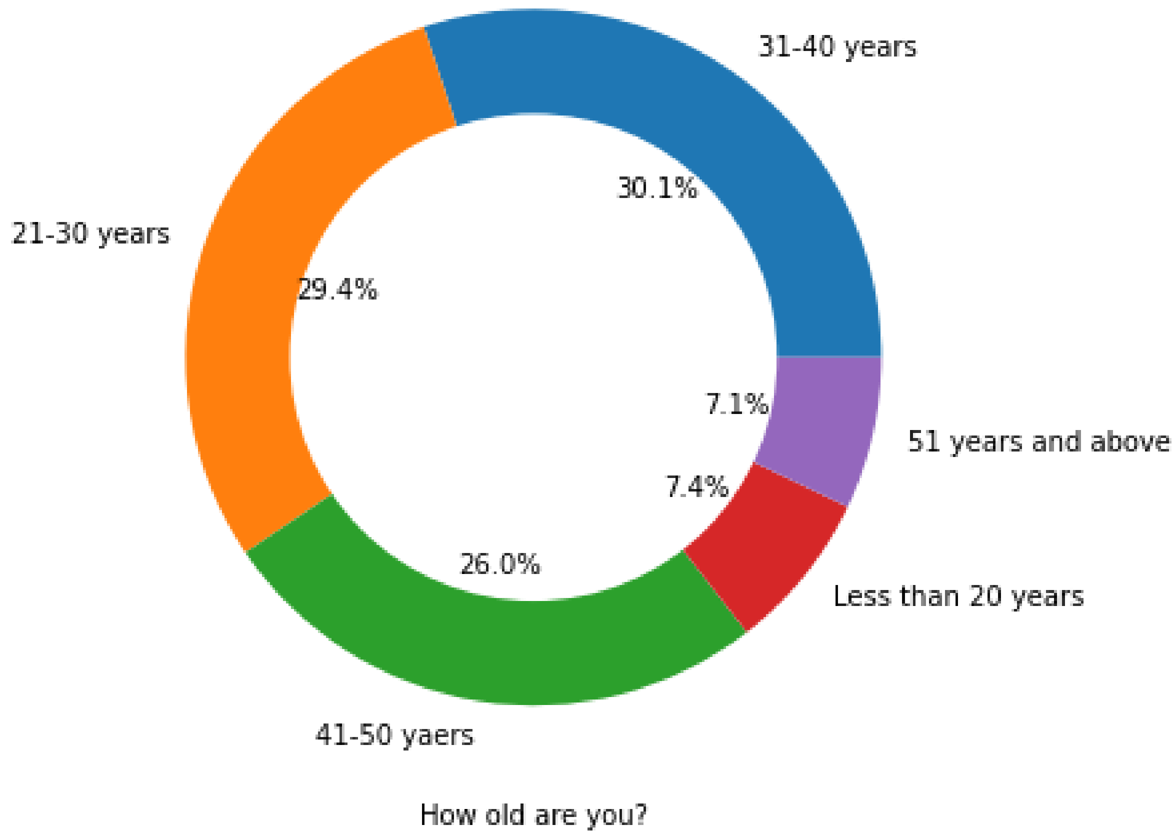
In [21]:

```
for i in personal_info:
    if i!='What is the Pin Code of where you shop online from?':
        plt.figure(figsize=(8,6))
        newdata[i].value_counts().plot.pie(autopct='%1.1f%%')
        centre=plt.Circle((0,0),0.7,fc='white')
        fig=plt.gcf()
        fig.gca().add_artist(centre)
        plt.xlabel(i)
        plt.ylabel('')
        plt.figure()
```

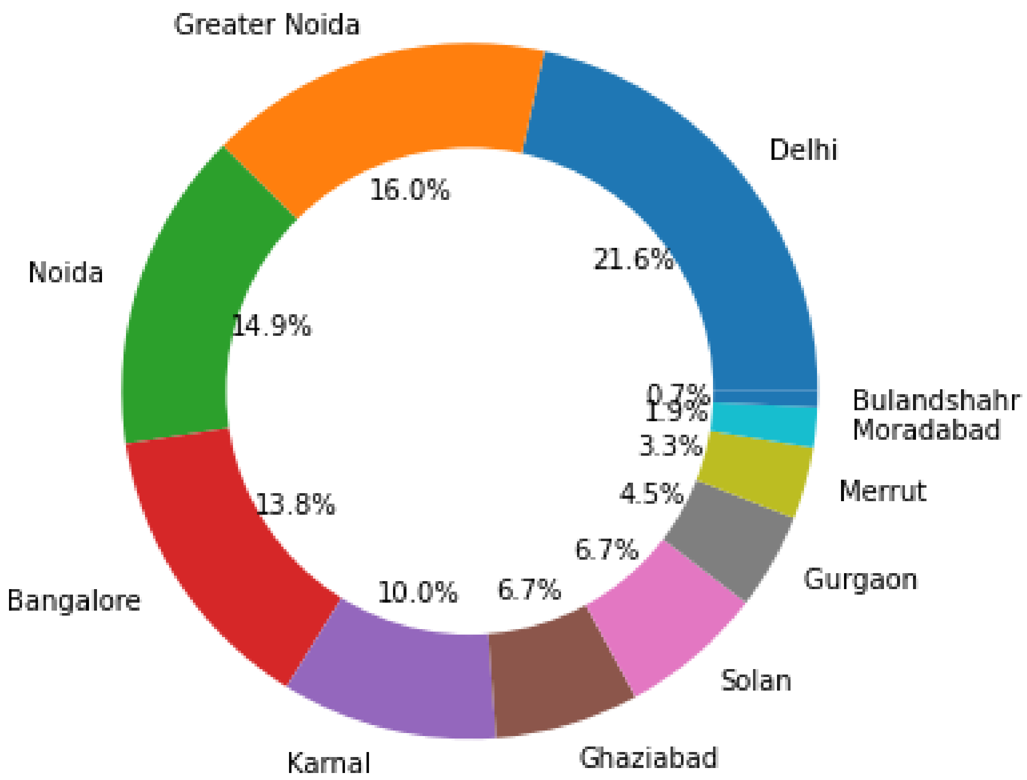


Gender of respondent

```
<Figure size 432x288 with 0 Axes>
```

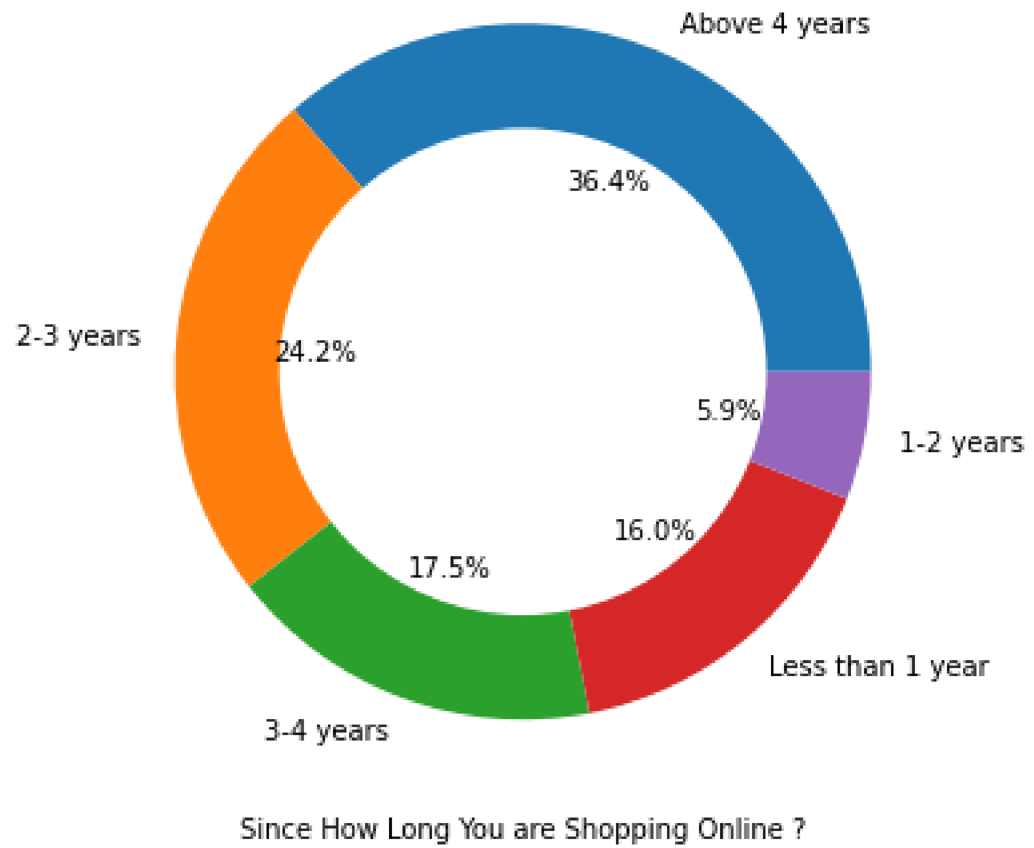


<Figure size 432x288 with 0 Axes>

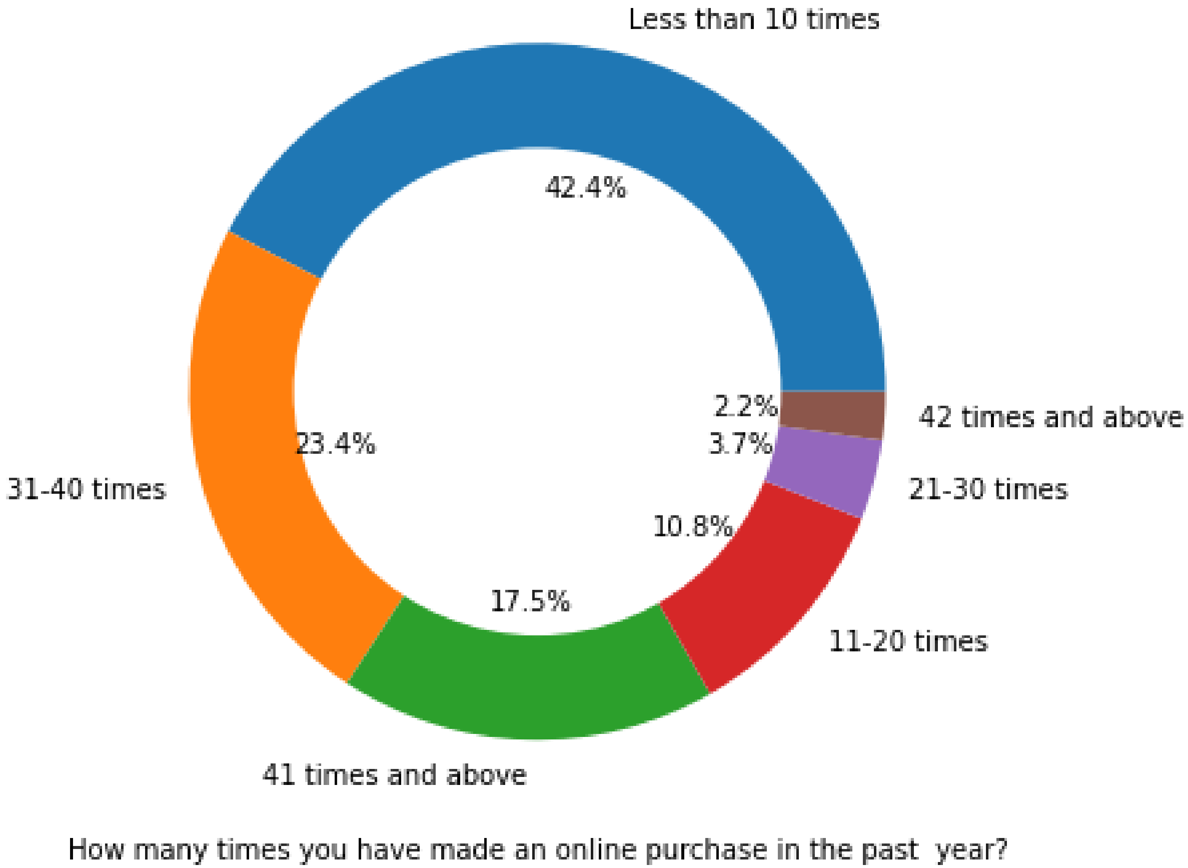


Which city do you shop online from?

<Figure size 432x288 with 0 Axes>



<Figure size 432x288 with 0 Axes>



<Figure size 432x288 with 0 Axes>

Analysis on the basis of various following factors

Intention of Repeat purchase

In [22]:

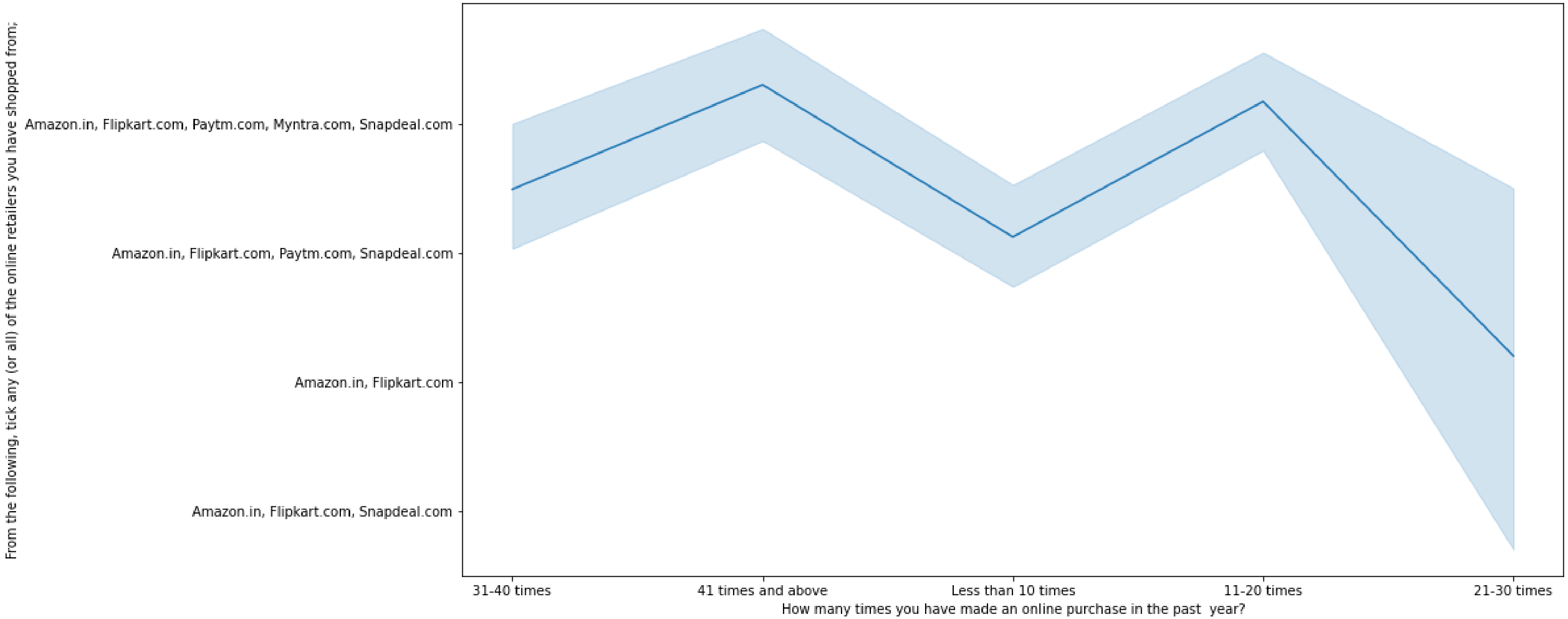
```
#Resolving ambiguity of column
#changing 42 times and above to 41 times and above

newdata['How many times you have made an online purchase in the past  year?'].replace('42 times and above','41 times
and above',
                                                                                inplace=True)
```

In [23]:

```
plt.figure(figsize=(15,8))
sns.lineplot(newdata['How many times you have made an online purchase in the past  year?'],
             newdata['From the following, tick any (or all) of the online retailers you have shopped from;'])
```

Out[23]:

```
<AxesSubplot:xlabel='How many times you have made an online purchase in the past  year?', ylabel='From the following
, tick any (or all) of the online retailers you have shopped from;'>
```



Converting years to numbers for better analysis

In [24]:

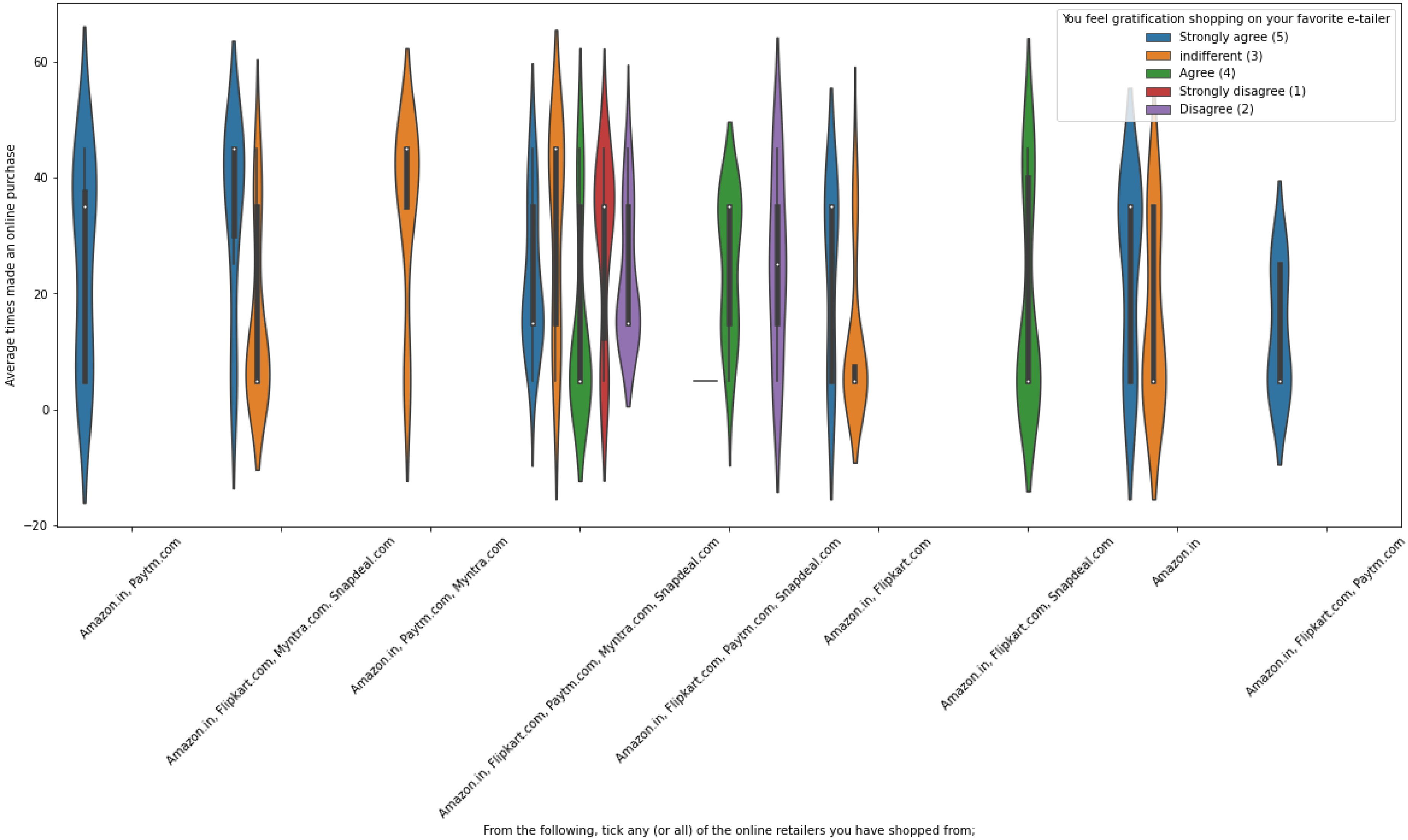
```
dict={'31-40 times':35,'41 times and above':45,'Less than 10 times':5,'11-20 times':15,'21-30 times':25}
newdata['Average times made an online purchase']=newdata['How many times you have made an online purchase in the past  year?'].replace(dict)
```

In [25]:

```
plt.figure(figsize=(20,8))
sns.violinplot(newdata['From the following, tick any (or all) of the online retailers you have shopped from;'],
               newdata['Average times made an online purchase'],hue=newdata['You feel gratification shopping on your favorite e-tailer'])
plt.xticks(rotation=45)
```

Out[25]:

```
(array([0, 1, 2, 3, 4, 5, 6, 7, 8]),
 [Text(0, 0, 'Amazon.in, Paytm.com'),
  Text(1, 0, 'Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com'),
  Text(2, 0, 'Amazon.in, Paytm.com, Myntra.com'),
  Text(3, 0, 'Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com'),
  Text(4, 0, 'Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com'),
  Text(5, 0, 'Amazon.in, Flipkart.com'),
  Text(6, 0, 'Amazon.in, Flipkart.com, Snapdeal.com'),
  Text(7, 0, 'Amazon.in'),
  Text(8, 0, 'Amazon.in, Flipkart.com, Paytm.com')])
```



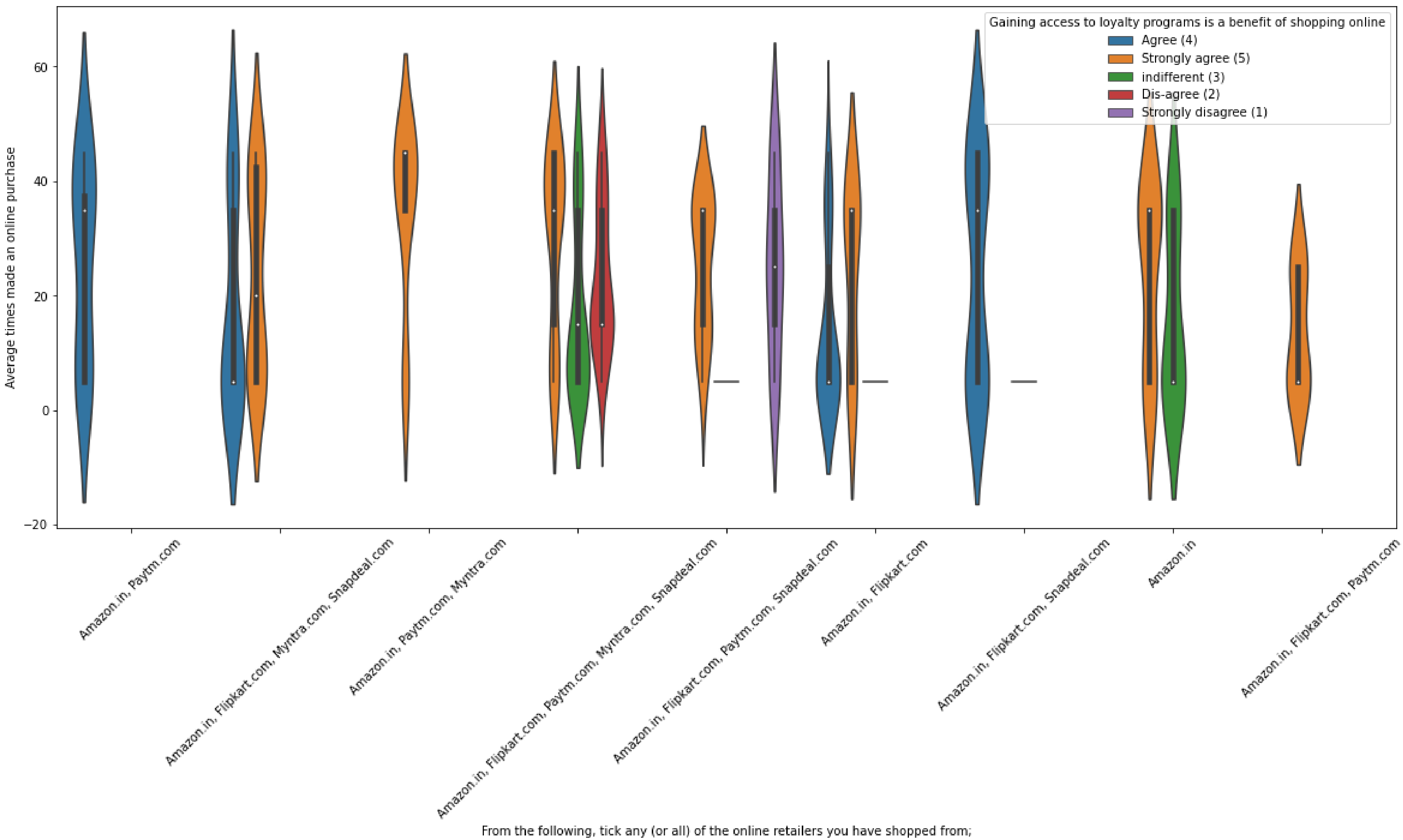
In [26]:

```
plt.figure(figsize=(20,8))
```

```python
sns.violinplot(newdata['From the following, tick any (or all) of the online retailers you have shopped from;'],
               newdata['Average times made an online purchase'],hue=newdata['Gaining access to loyalty programs is a benefit of shopping online'])
plt.xticks(rotation=45)
```

Out[26]:

```
(array([0, 1, 2, 3, 4, 5, 6, 7, 8]),
 [Text(0, 0, 'Amazon.in, Paytm.com'),
  Text(1, 0, 'Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com'),
  Text(2, 0, 'Amazon.in, Paytm.com, Myntra.com'),
  Text(3, 0, 'Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com'),
  Text(4, 0, 'Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com'),
  Text(5, 0, 'Amazon.in, Flipkart.com'),
  Text(6, 0, 'Amazon.in, Flipkart.com, Snapdeal.com'),
  Text(7, 0, 'Amazon.in'),
  Text(8, 0, 'Amazon.in, Flipkart.com, Paytm.com')])
```


Gaining access to loyalty programs is a benefit of shopping online
Agree (4)
Strongly agree (5)
indifferent (3)
Dis-agree (2)
Strongly disagree (1)
Average times made an online purchase
60
40
20
0
−20
Amazon.in, Paytm.com
Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com
Amazon.in, Paytm.com, Myntra.com
Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com
Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com
Amazon.in, Flipkart.com
Amazon.in, Flipkart.com, Snapdeal.com
Amazon.in
Amazon.in, Flipkart.com, Paytm.com
From the following, tick any (or all) of the online retailers you have shopped from;

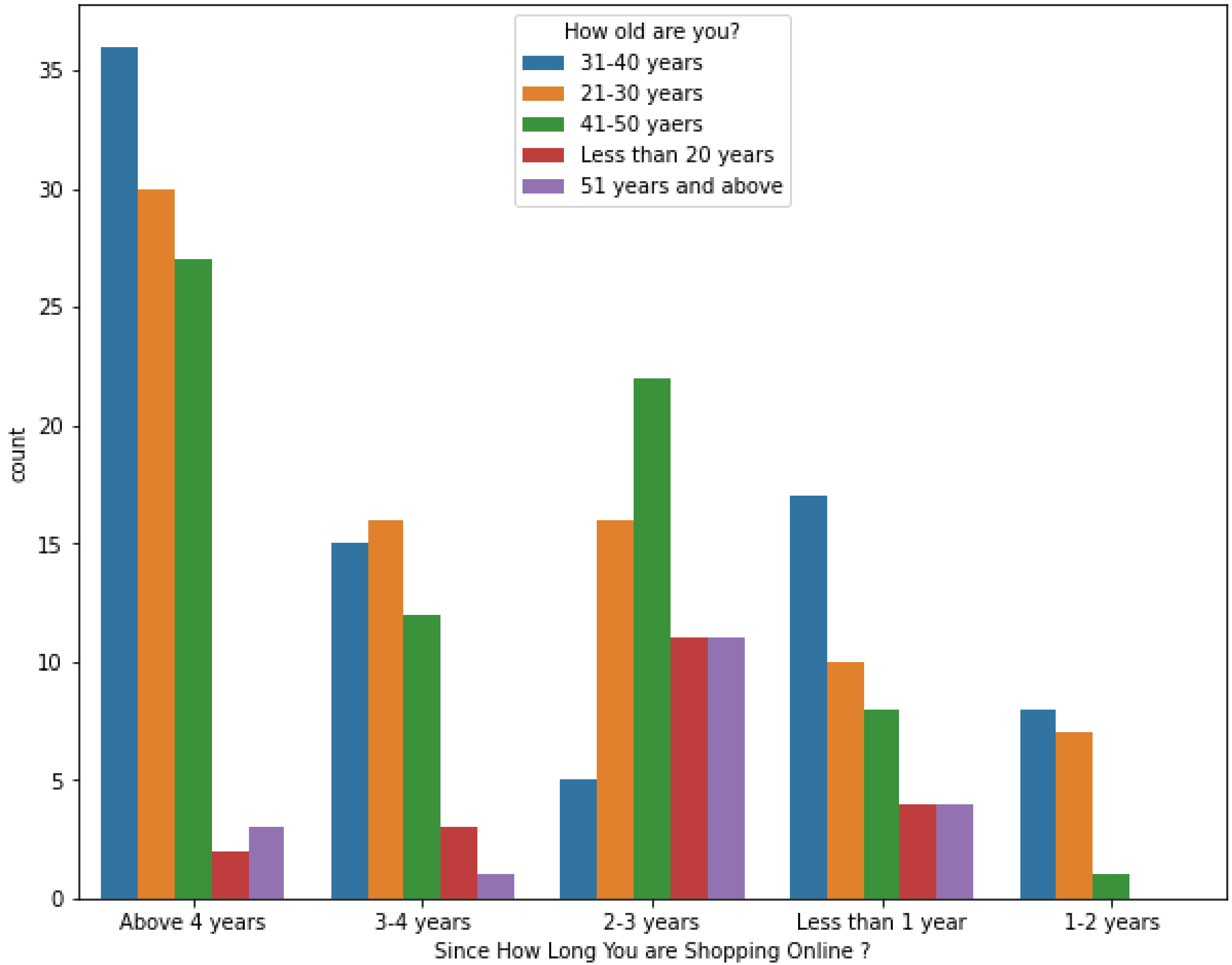
Online Retailing

In [27]:

```
plt.figure(figsize=(10,8))
sns.countplot(newdata['Since How Long You are Shopping Online ?'],hue=newdata['How old are you?'])
```

Out[27]:

```
<AxesSubplot:xlabel='Since How Long You are Shopping Online ?', ylabel='count'>
```


How old are you?
31-40 years
21-30 years
41-50 yaers
Less than 20 years
51 years and above
count
0
5
10
15
20
25
30
35
Above 4 years
3-4 years
2-3 years
Less than 1 year
1-2 years
Since How Long You are Shopping Online ?

Highest number of people have been shopping online for above 4 years except for the age group below 20 years and above 50 years. People who are shopping online for 1-2 years does not include teenagers and elder people.

Converting Years to numbers for better analysis

In [28]:

```
newdata['Since How Long You are Shopping Online ?'].unique()
```

Out[28]:

```
array(['Above 4 years', '3-4 years', '2-3 years', 'Less than 1 year',
       '1-2 years'], dtype=object)
```

In [29]:

```
dict={'Above 4 years':4.5,'3-4 years':3.5,'2-3 years':2.5,'1-2 years':1.5,'Less than 1 year':0.5}
newdata['Average years of shopping online']=newdata['Since How Long You are Shopping Online ?'].replace(dict)
```

In [30]:

```
newdata['Which city do you shop online from?'].unique()
```

Out[30]:

```
array(['Delhi', 'Greater Noida', 'Karnal ', 'Bangalore ', 'Noida',
       'Solan', 'Moradabad', 'Gurgaon ', 'Merrut', 'Ghaziabad',
       'Bulandshahr'], dtype=object)
```

In [31]:

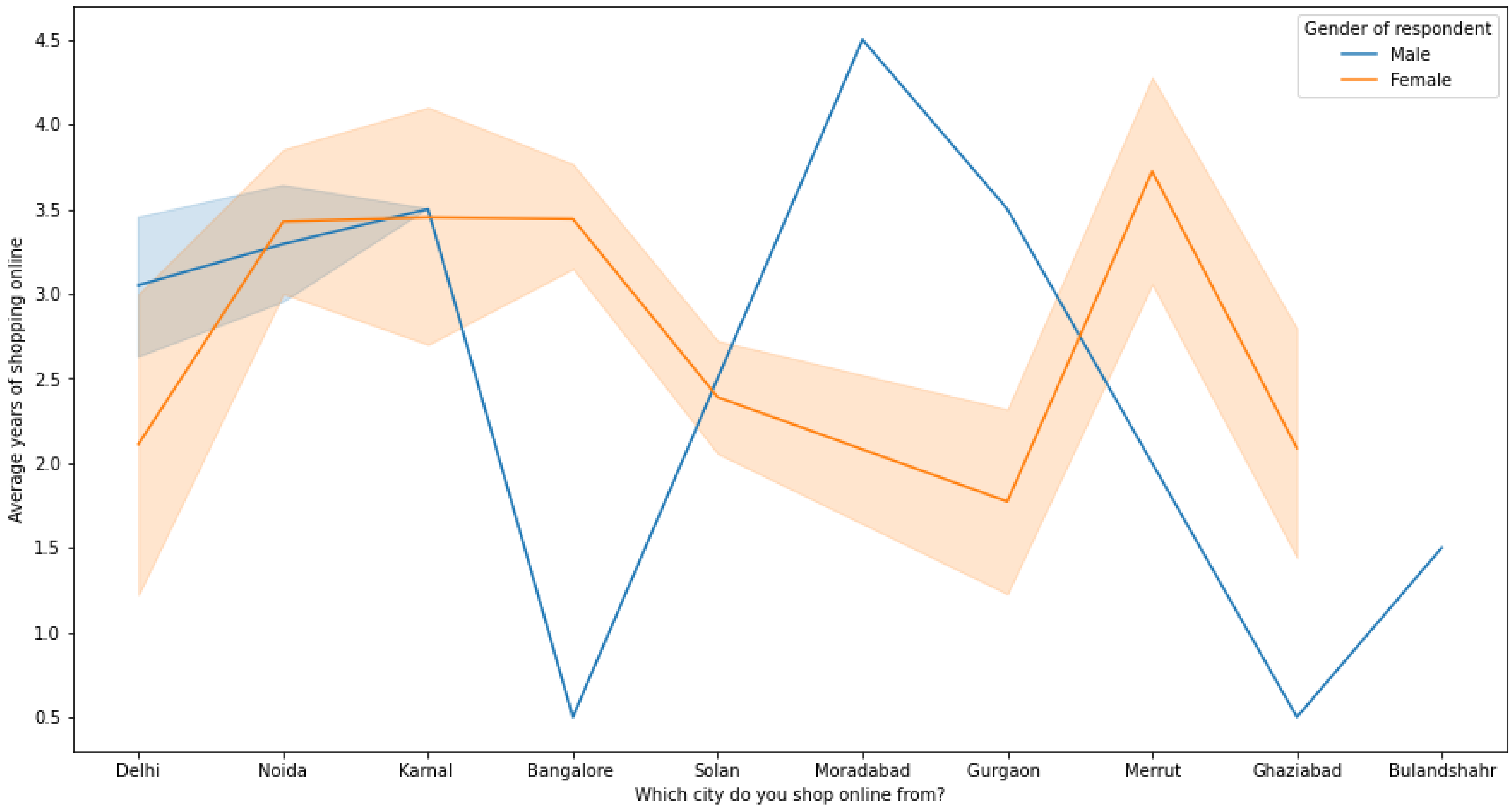
```
#Changing Greater noida to noida
newdata['Which city do you shop online from?'].replace({'Greater Noida':'Noida'},inplace=True)
```

In [32]:

```
plt.figure(figsize=(15,8))
sns.lineplot(newdata['Which city do you shop online from?'],newdata['Average years of shopping
online'],hue=newdata['Gender of respondent'])
```

Out[32]:

```
<AxesSubplot:xlabel='Which city do you shop online from?', ylabel='Average years of shopping online'>
```



In lines, we can see that density of female customers is more than male. Men living in banglore and ghaziabad shop have shopped online for less than 1 year. Highest number of men shopping online belong from delhi and noida, while men from moradabad have been shopping online for the longest. Women from meerut and noida have shopped the longest.

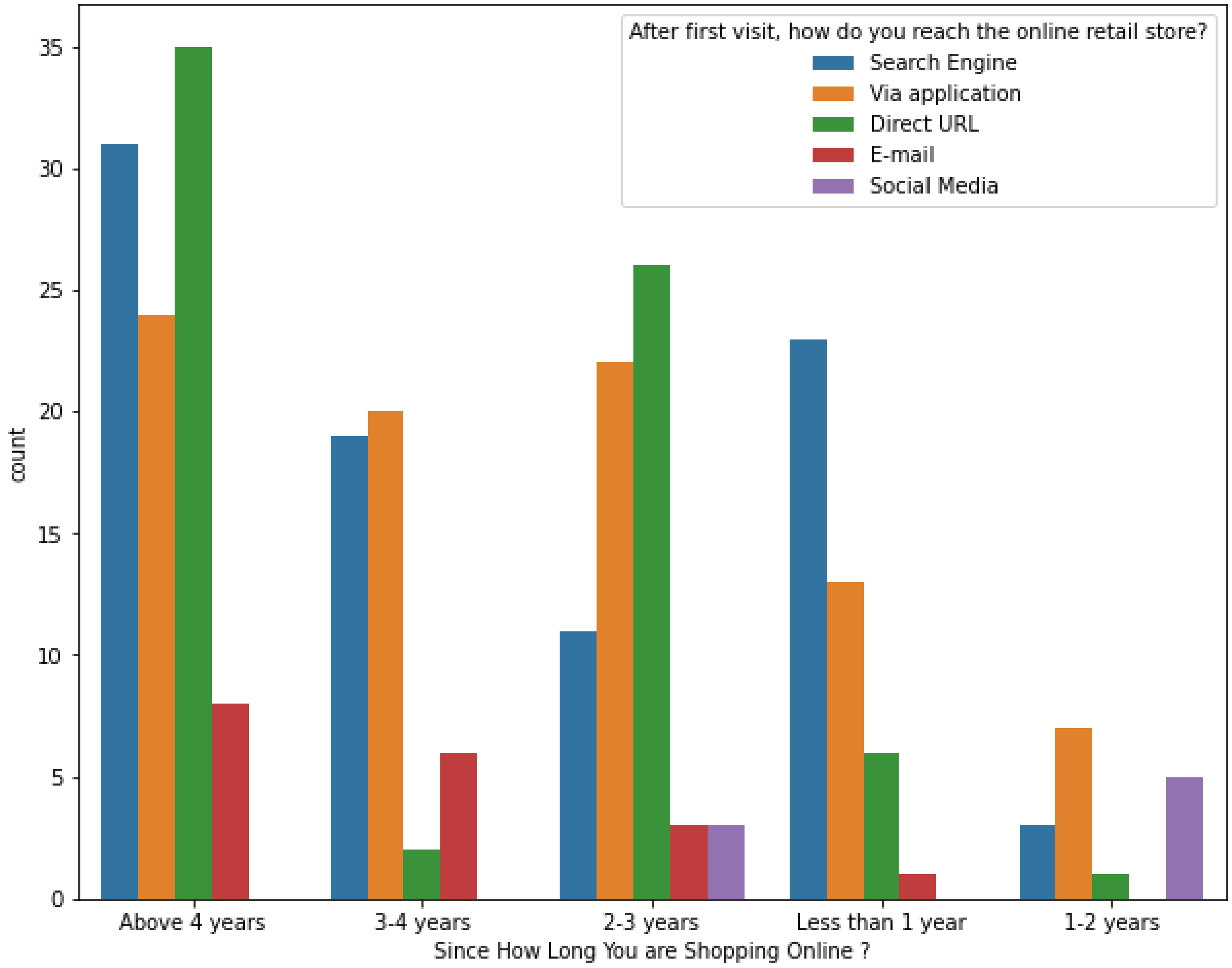
In [33]:

```
plt.figure(figsize=(10,8))
```

```
sns.countplot(newdata['Since How Long You are Shopping Online ?'],
              hue=newdata['After first visit, how do you reach the online retail store?'])
```

Out[33]:

```
<AxesSubplot:xlabel='Since How Long You are Shopping Online ?', ylabel='count'>
```


After first visit, how do you reach the online retail store?
Search Engine
Via application
Direct URL
E-mail
Social Media
count
35
30
25
20
15
10
5
0
Above 4 years
3-4 years
2-3 years
Less than 1 year
1-2 years
Since How Long You are Shopping Online ?

Even though people who are shopping online for more than 3 years donot use the application rather use search engine and direct url's in large number which indicates that online brands should update all their platforms rather than just application

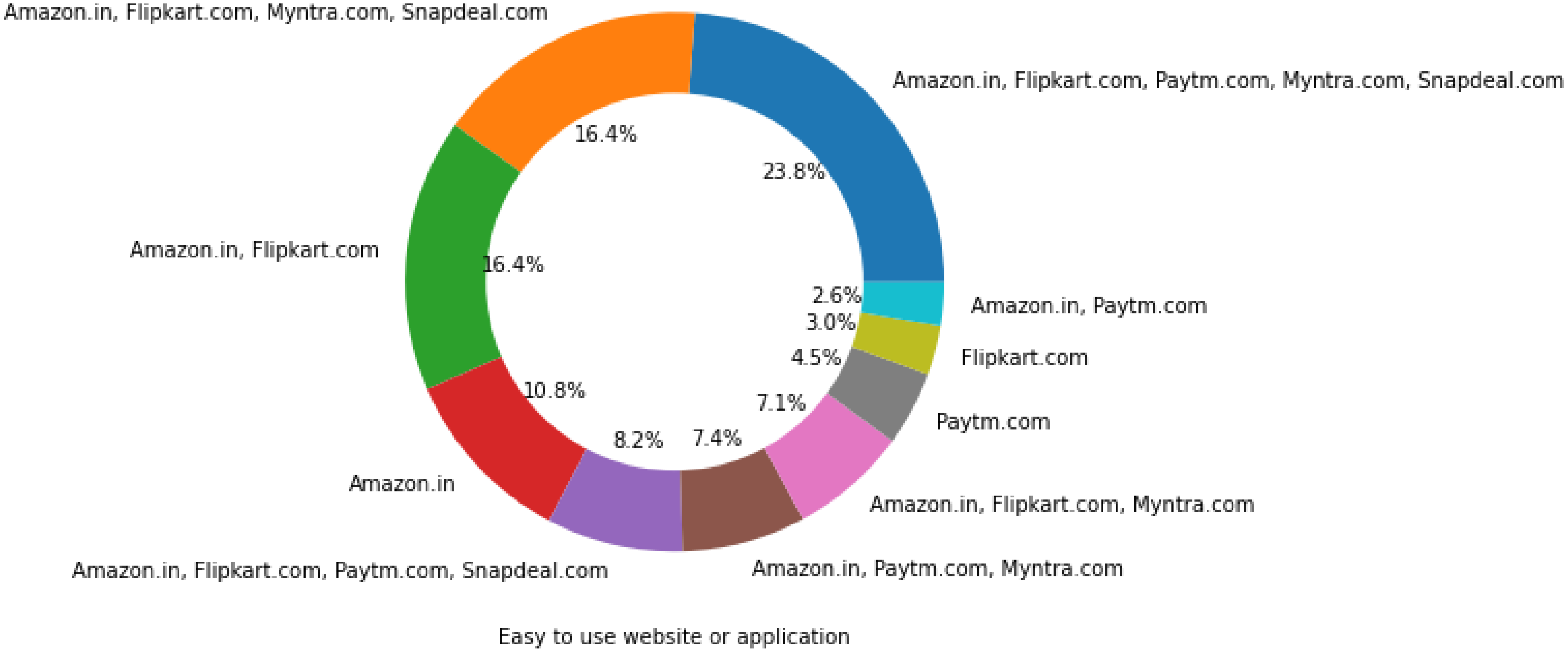
Brand Image

In [34]:

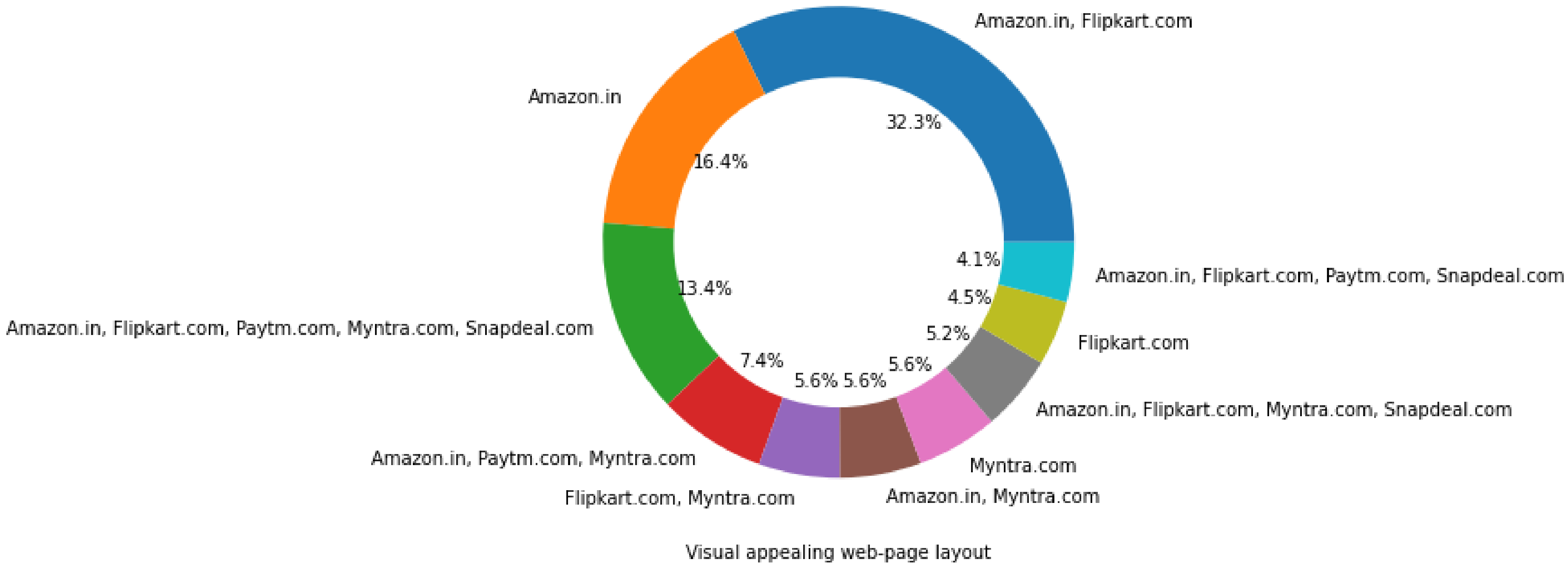
```
performance=['Easy to use website or application',
       'Visual appealing web-page layout', 'Wild variety of product on offer',
       'Complete, relevant description information of products',
       'Fast loading website speed of website and application',
       'Reliability of the website or 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']
```

In [35]:

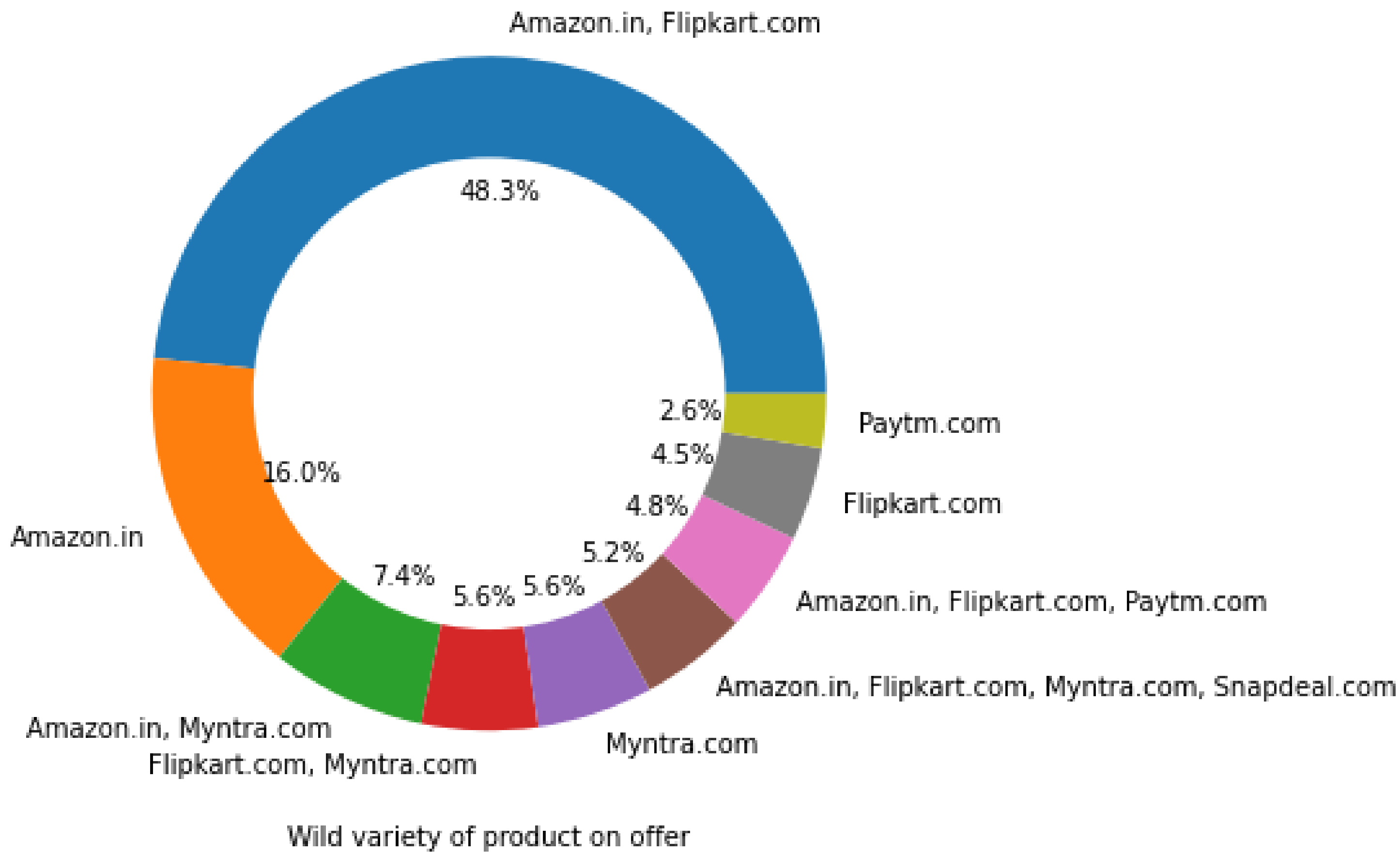
```
for i in performance:
        plt.figure(figsize=(8,6))
        newdata[i].value_counts().plot.pie(autopct='%1.1f%%')
        centre=plt.Circle((0,0),0.7,fc='white')
        fig=plt.gcf()
        fig.gca().add_artist(centre)
        plt.xlabel(i)
        plt.ylabel('')
        plt.figure()
```



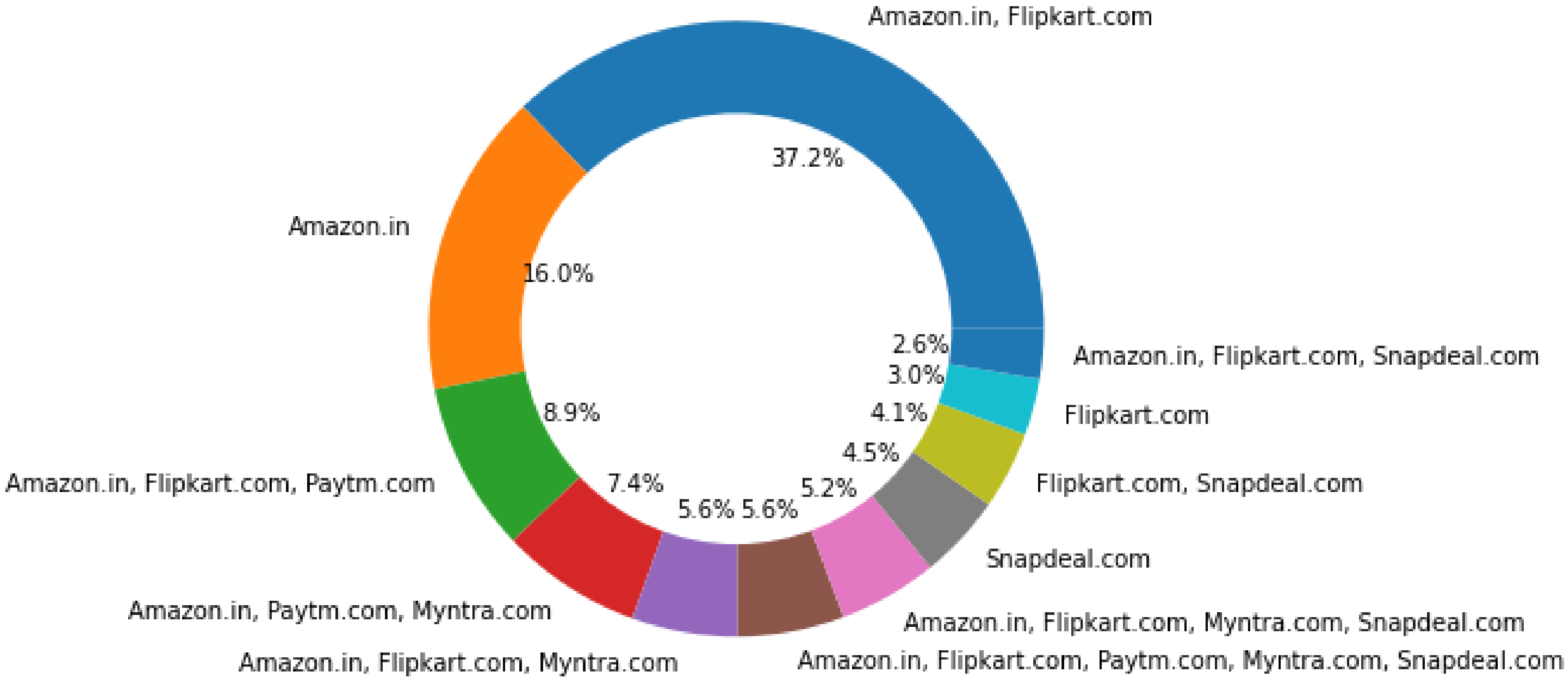
<Figure size 432x288 with 0 Axes>




```
<Figure size 432x288 with 0 Axes>
```

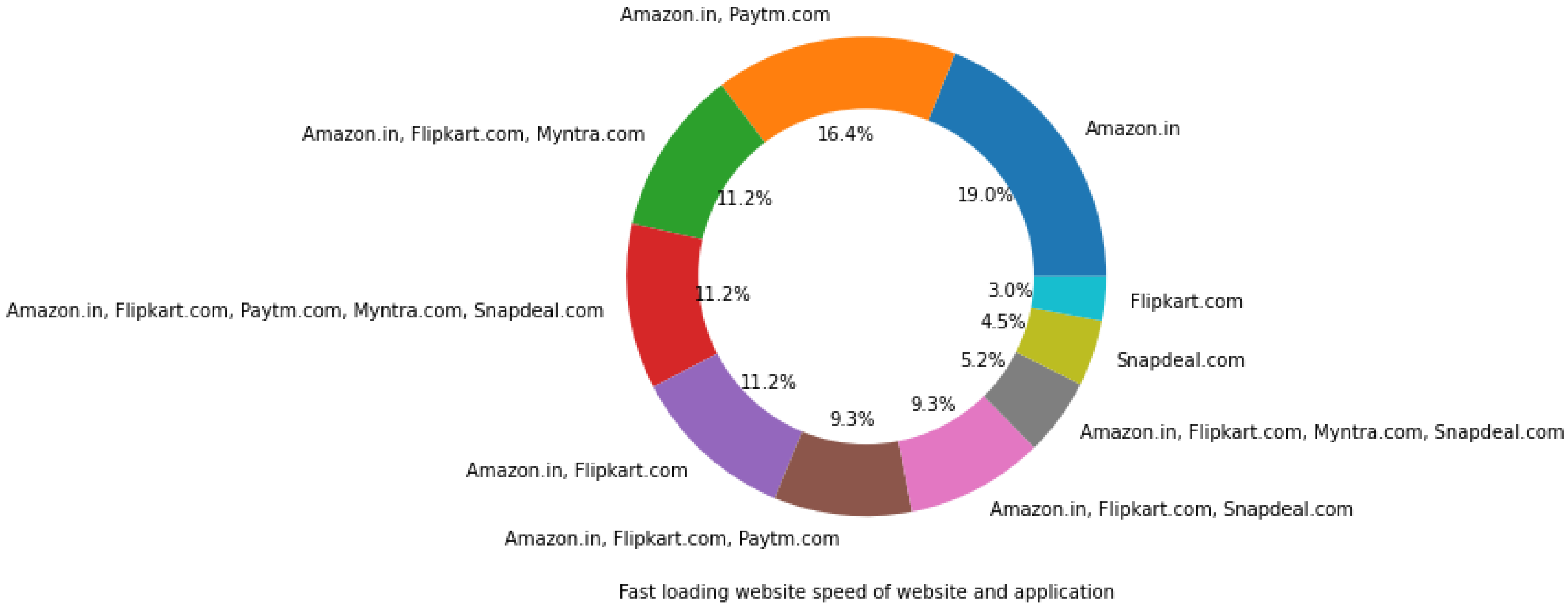



```
<Figure size 432x288 with 0 Axes>
```

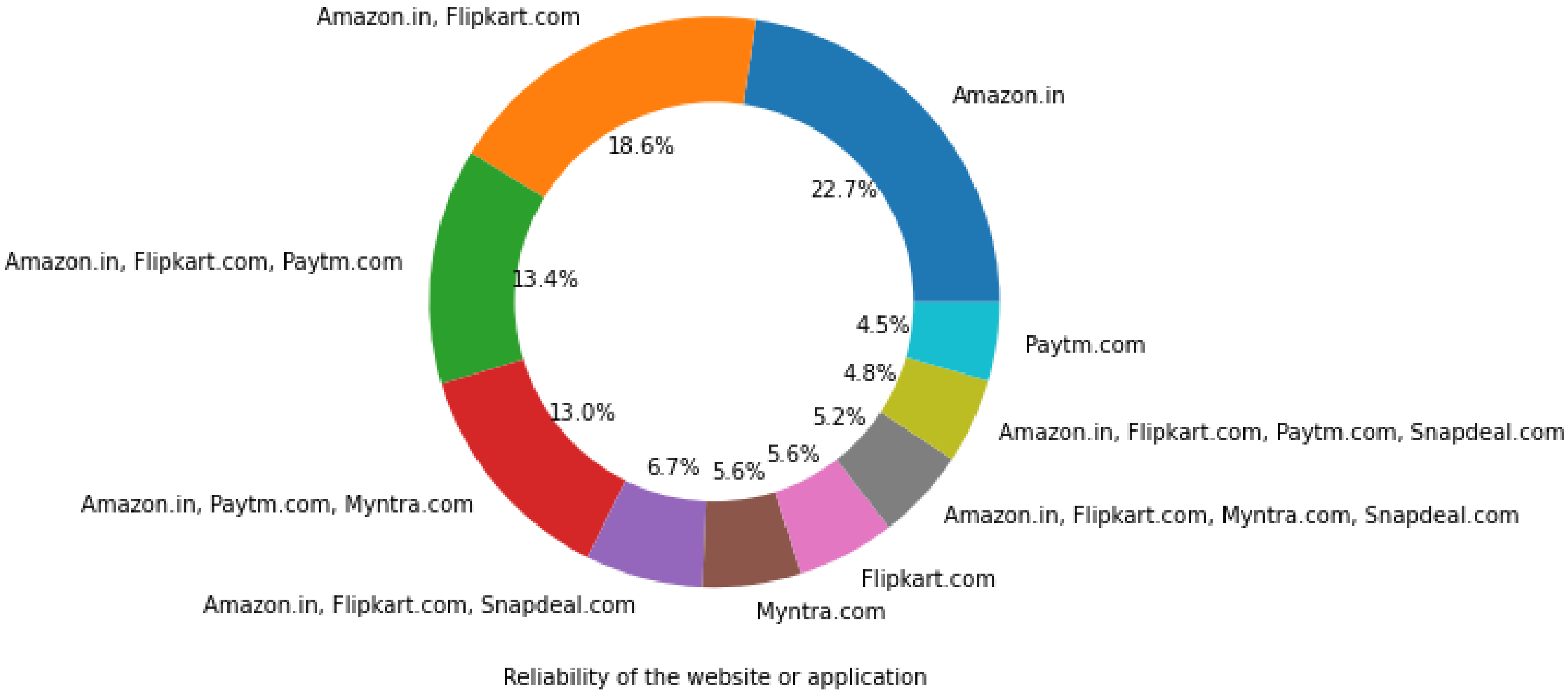


Complete, relevant description information of products

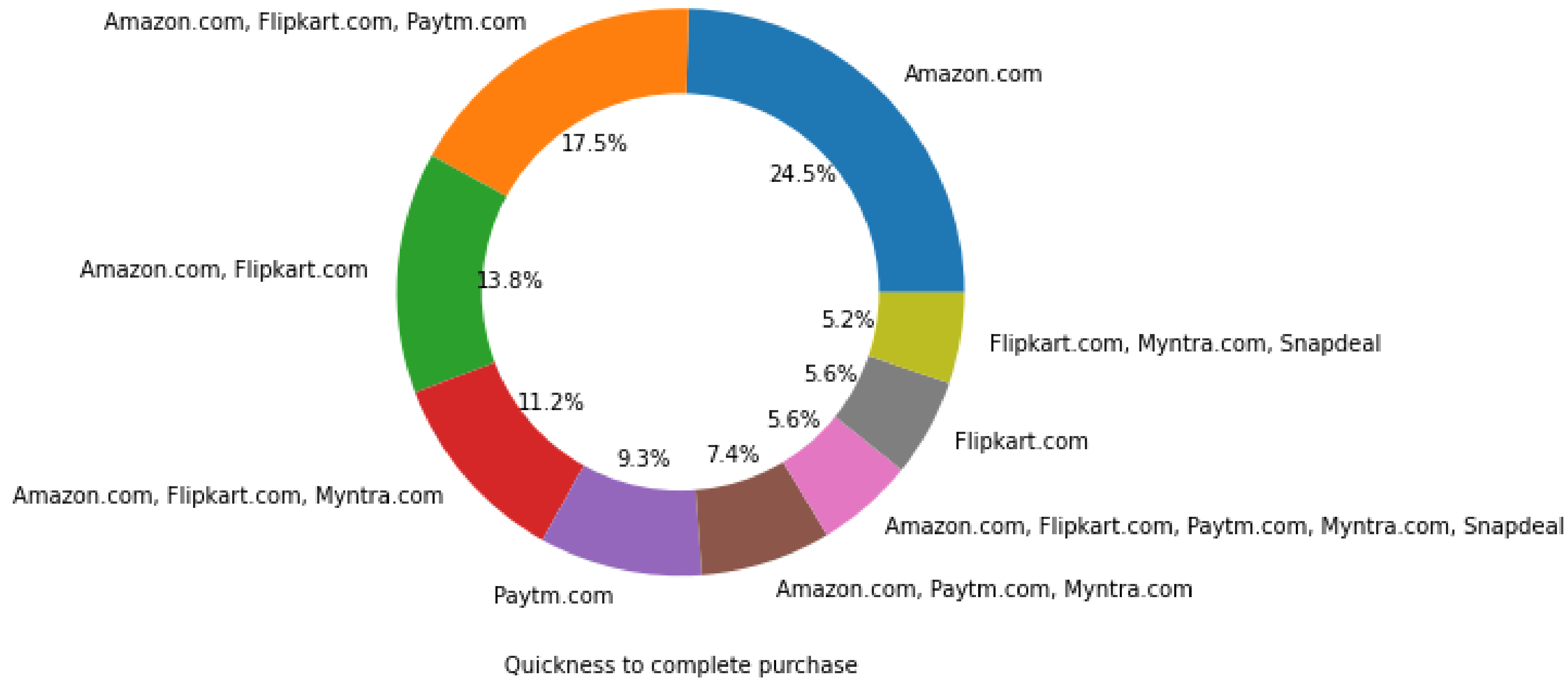
<Figure size 432x288 with 0 Axes>



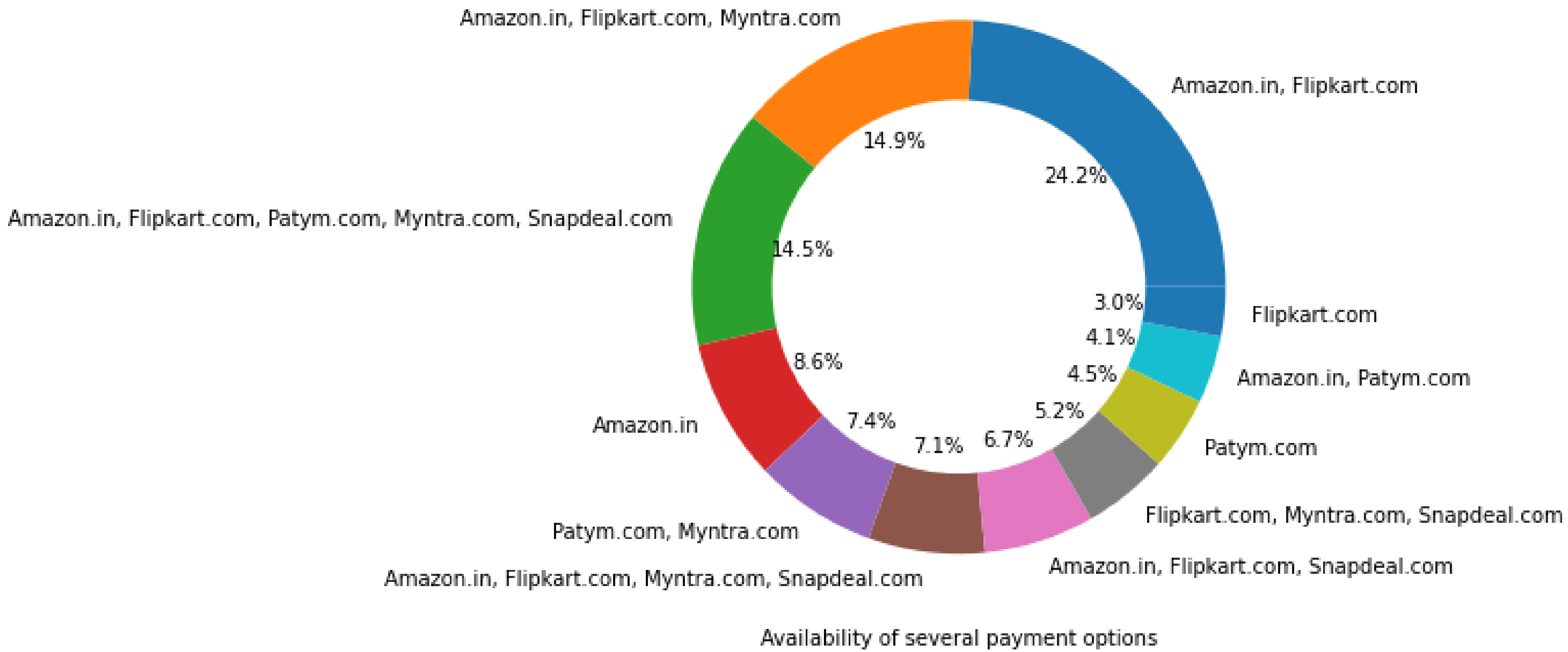
<Figure size 432x288 with 0 Axes>



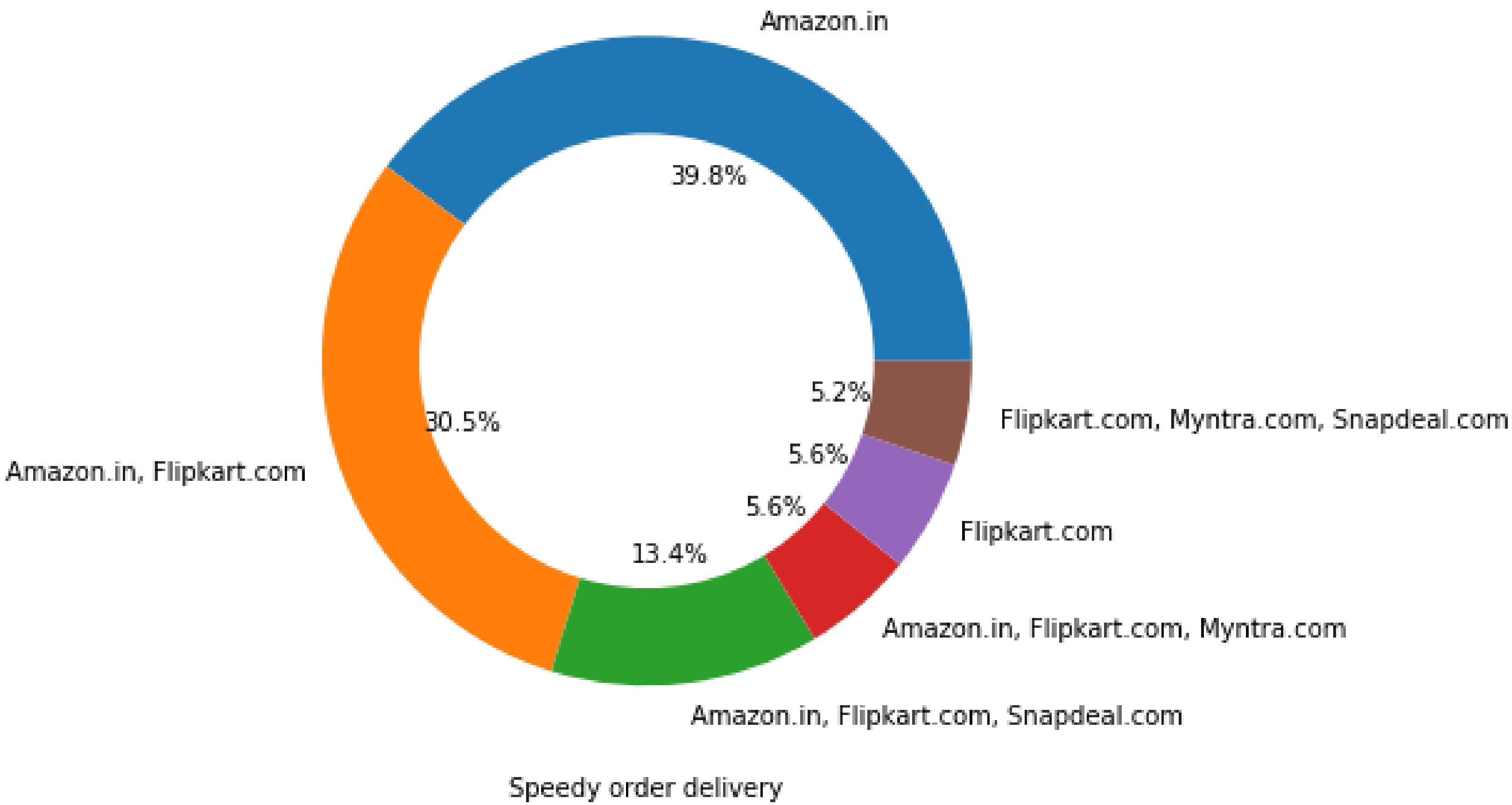

```
<Figure size 432x288 with 0 Axes>
```



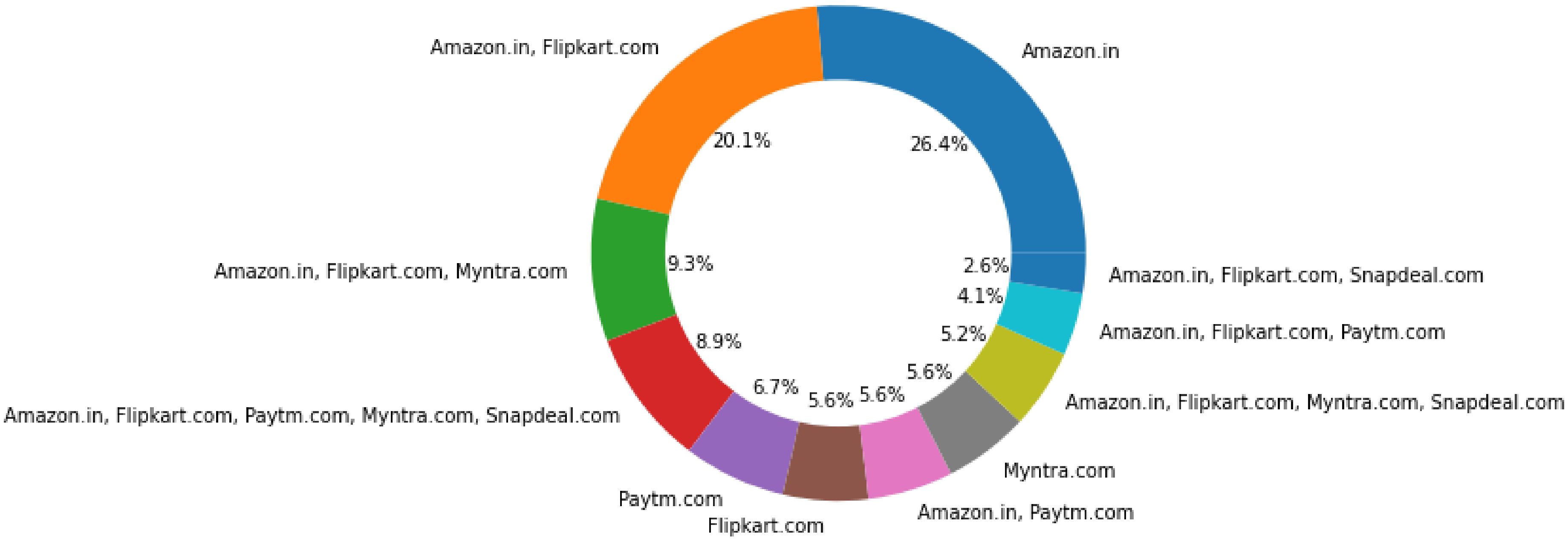

```
<Figure size 432x288 with 0 Axes>
```




```
<Figure size 432x288 with 0 Axes>
```

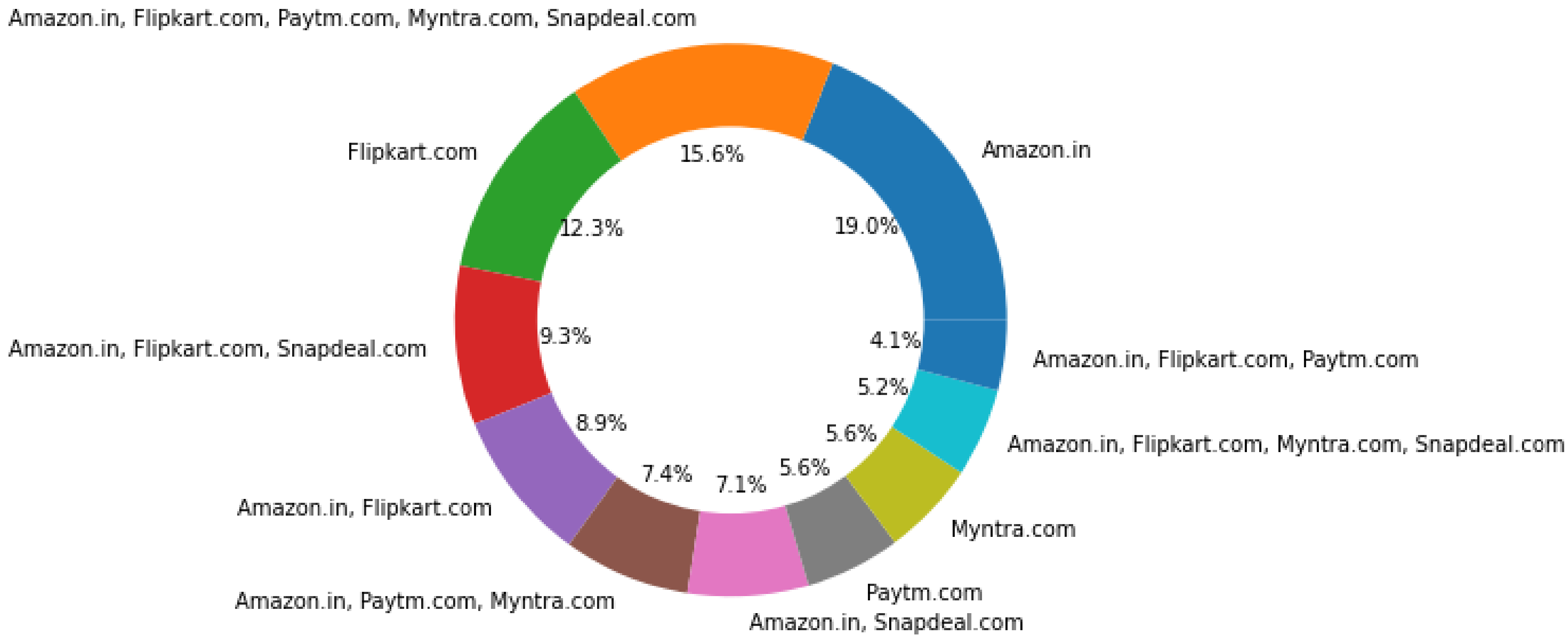


<Figure size 432x288 with 0 Axes>



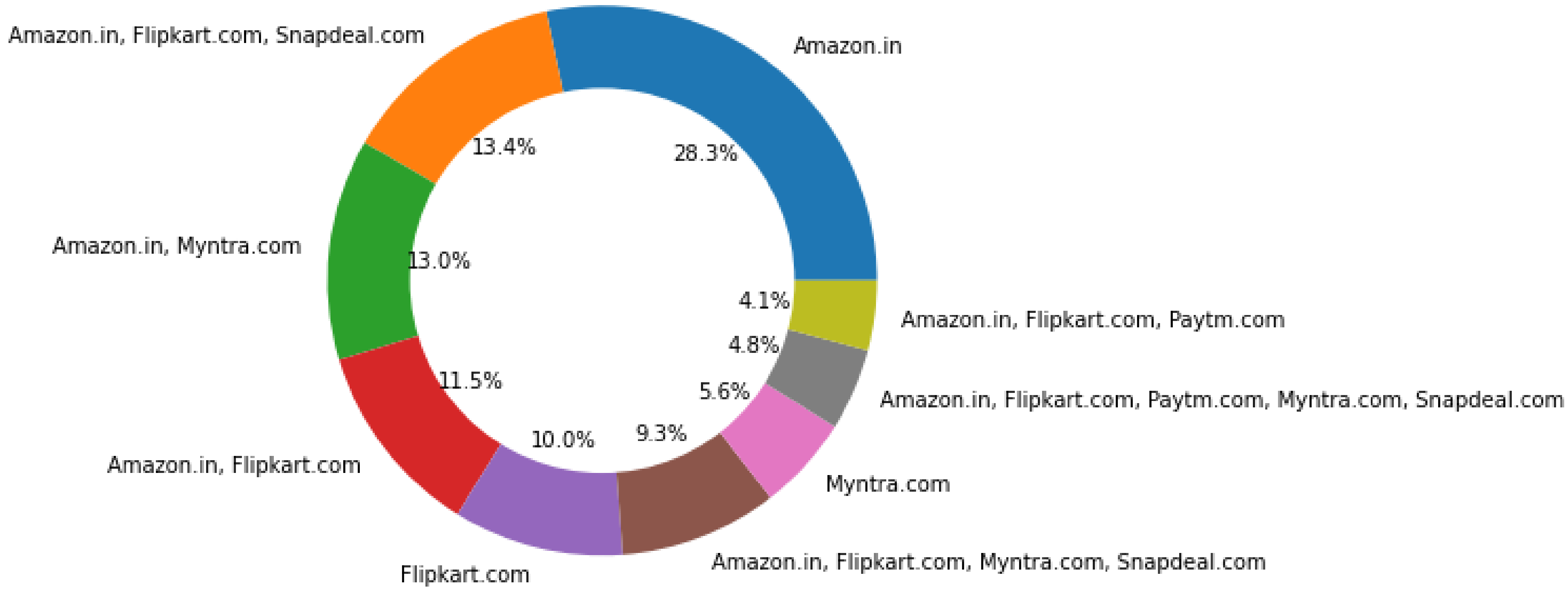
Privacy of customers' information

```
<Figure size 432x288 with 0 Axes>
```



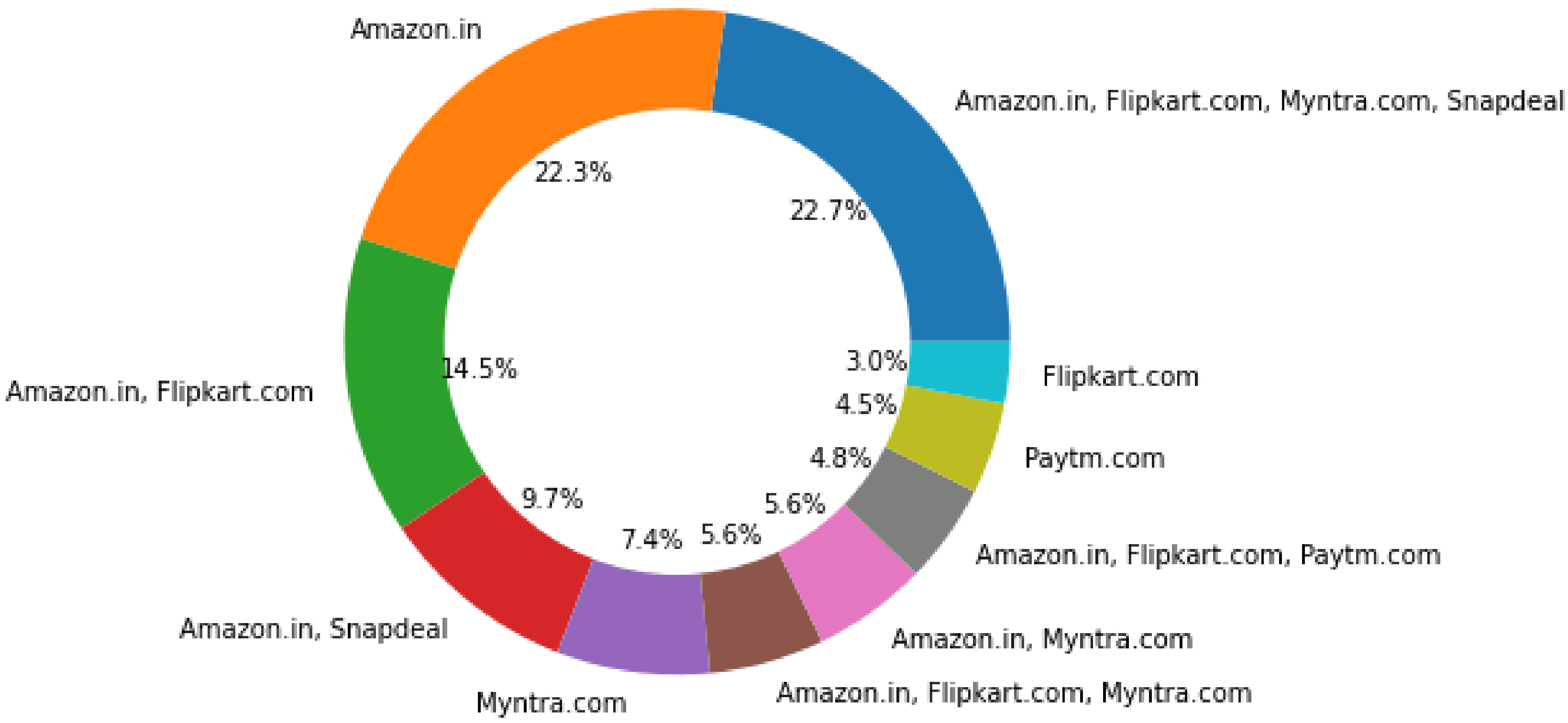
Security of customer financial information

<Figure size 432x288 with 0 Axes>



Perceived Trustworthiness

<Figure size 432x288 with 0 Axes>



Presence of online assistance through multi-channel

```
<Figure size 432x288 with 0 Axes>
```

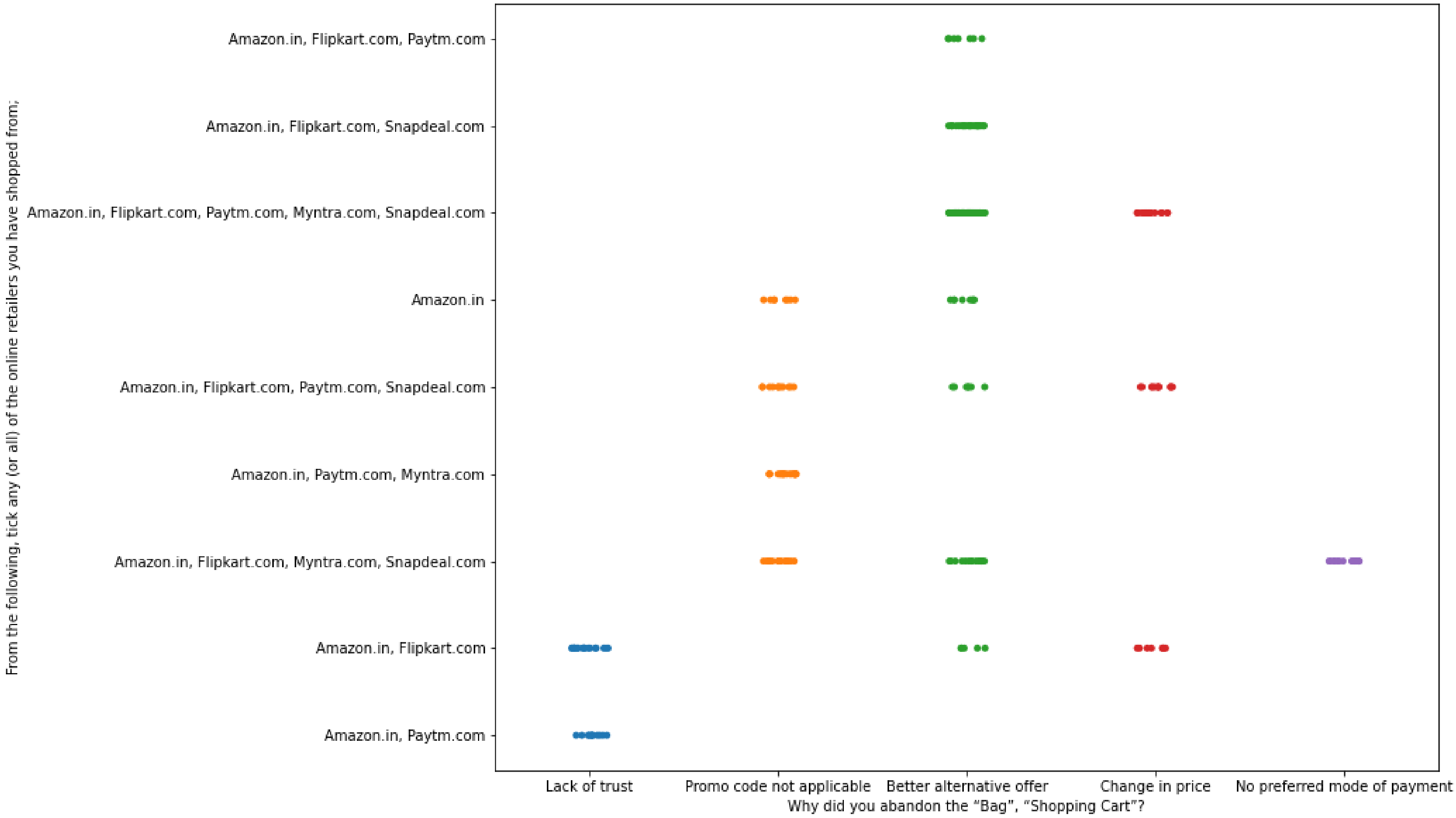
Amazon, Flipkart have been had the highest votes for having all the positive points and have maintained a very good brand image followed by paytm and the myntra.

In [36]:

```
plt.figure(figsize=(12,10))
sns.stripplot(newdata['Why did you abandon the "Bag", "Shopping Cart"?'],
              newdata['From the following, tick any (or all) of the online retailers you have shopped from;'])
```

Out[36]:

```
<AxesSubplot:xlabel='Why did you abandon the "Bag", "Shopping Cart"?', ylabel='From the following, tick any (or all) of the online retailers you have shopped from;'>
```



We can clearly see that most of the time people abandon the bag is beacuse they get a better alternative offer or promo code not applicable. There is also lack of trust seen in amazon, flipkart and paytm by some people.

Loyality

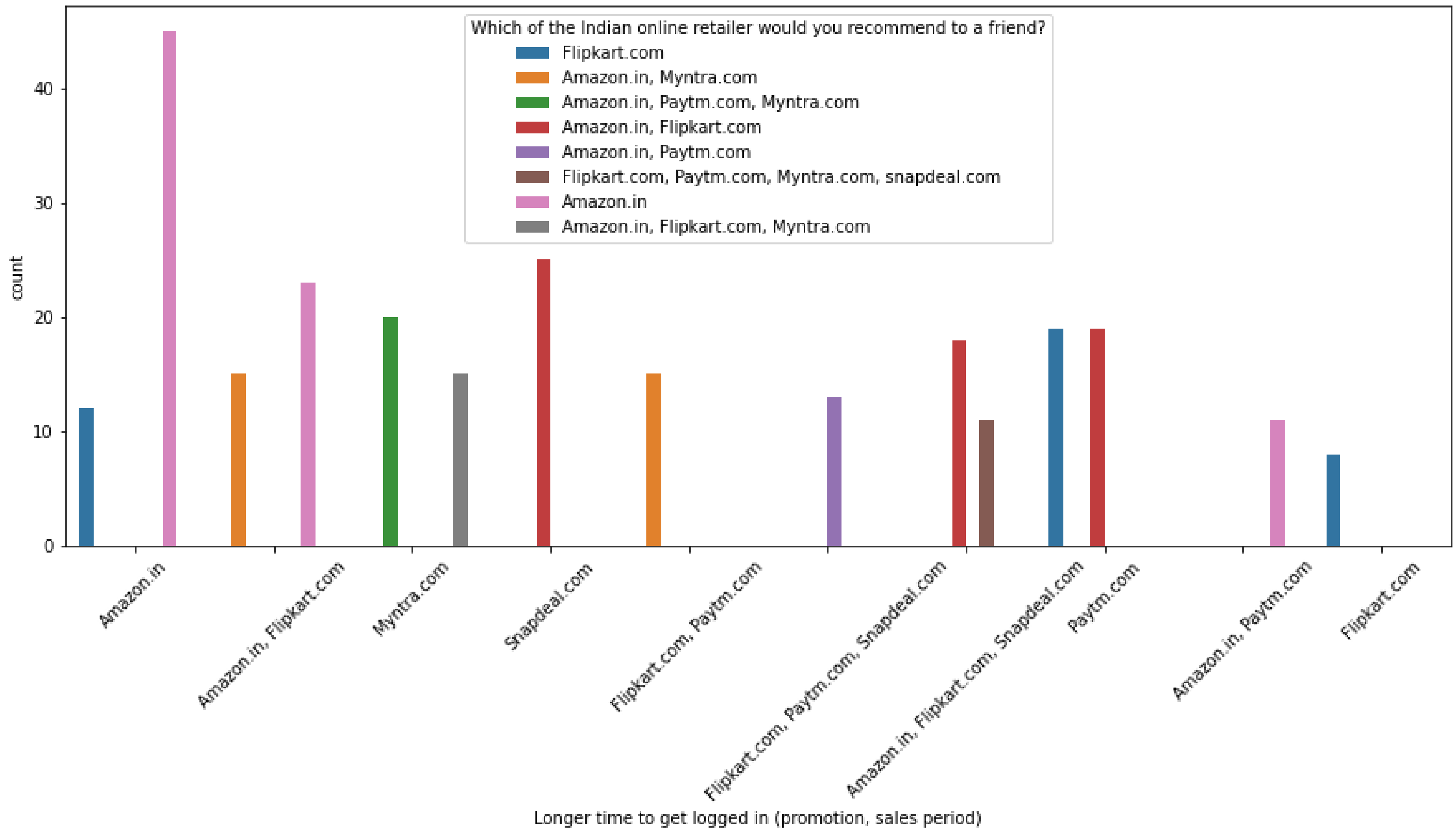
Loyal customers are those who keep using the same brand even if it is not good as other brands

In [37]:

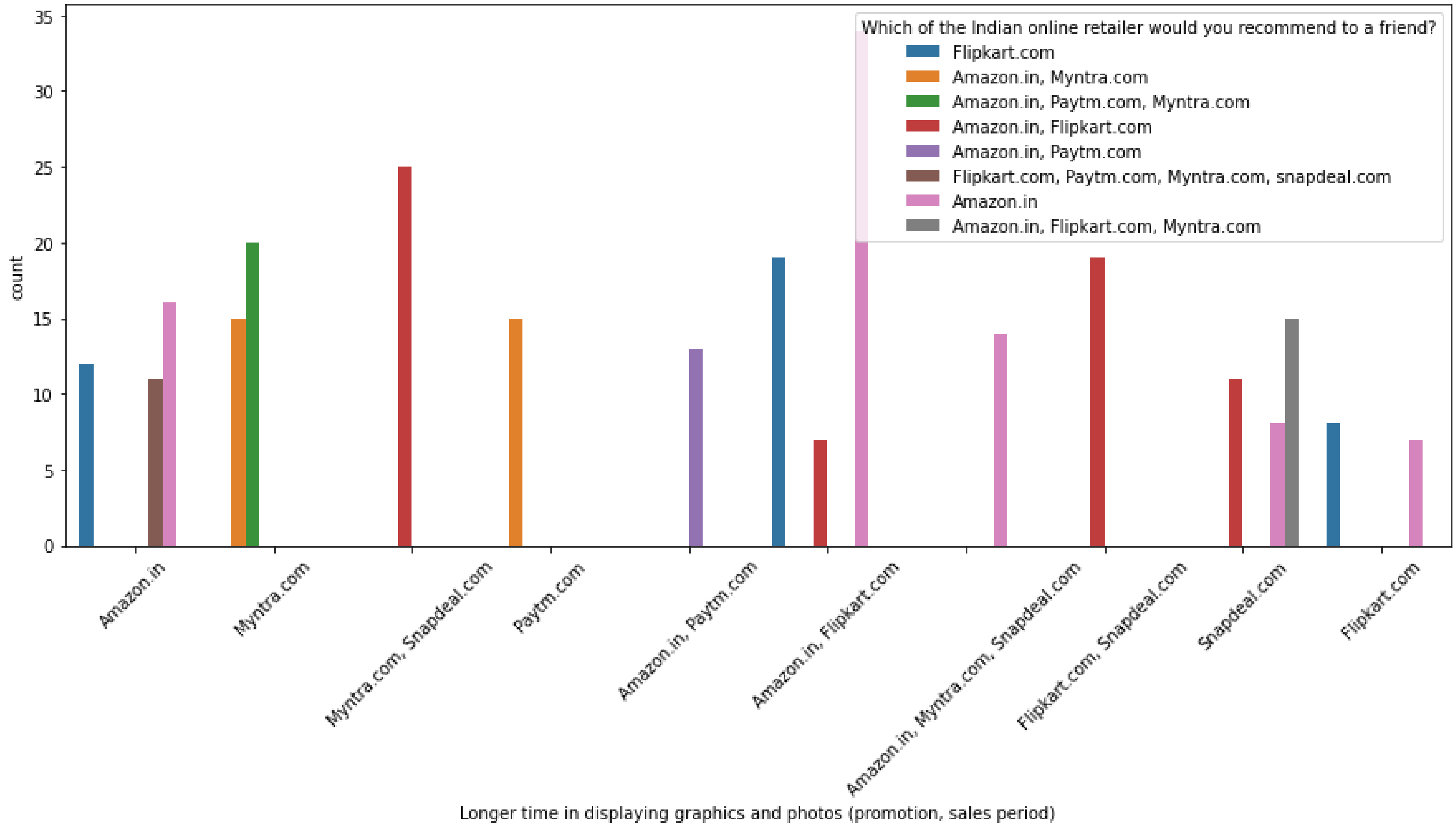
```
#Collecting all the negative remarks about a brand
bad=['Longer time to get logged in (promotion, sales period)',
       'Longer time in displaying graphics and photos (promotion, sales period)',
       'Late declaration of price (promotion, sales period)',
       'Longer page loading time (promotion, 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']
```

In [38]:

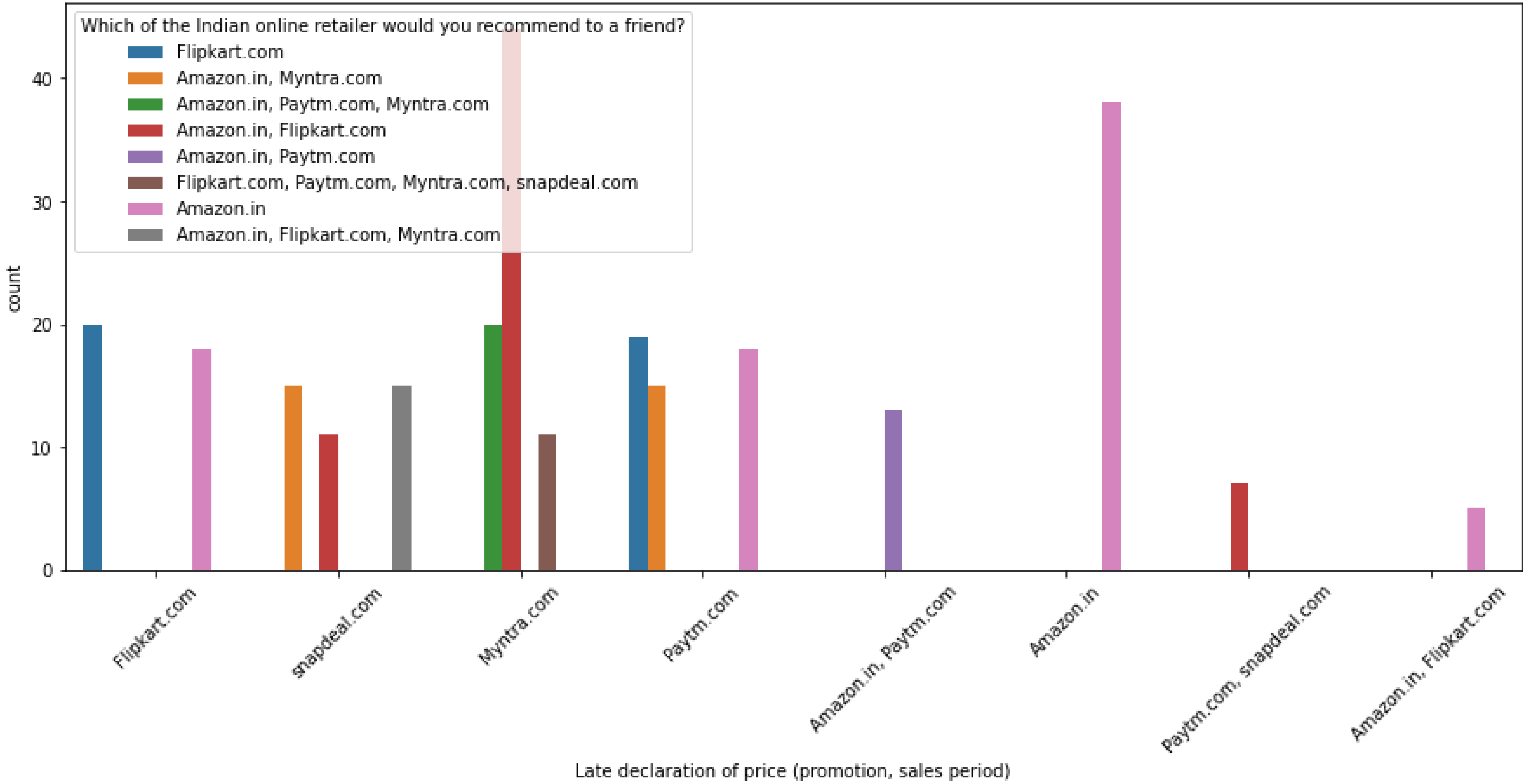
```
for i in bad:
        plt.figure(figsize=(15,6))
        sns.countplot(newdata[i],hue=newdata['Which of the Indian online retailer would you recommend to a 
friend?'])
        plt.xticks(rotation=45)
        plt.figure()
```



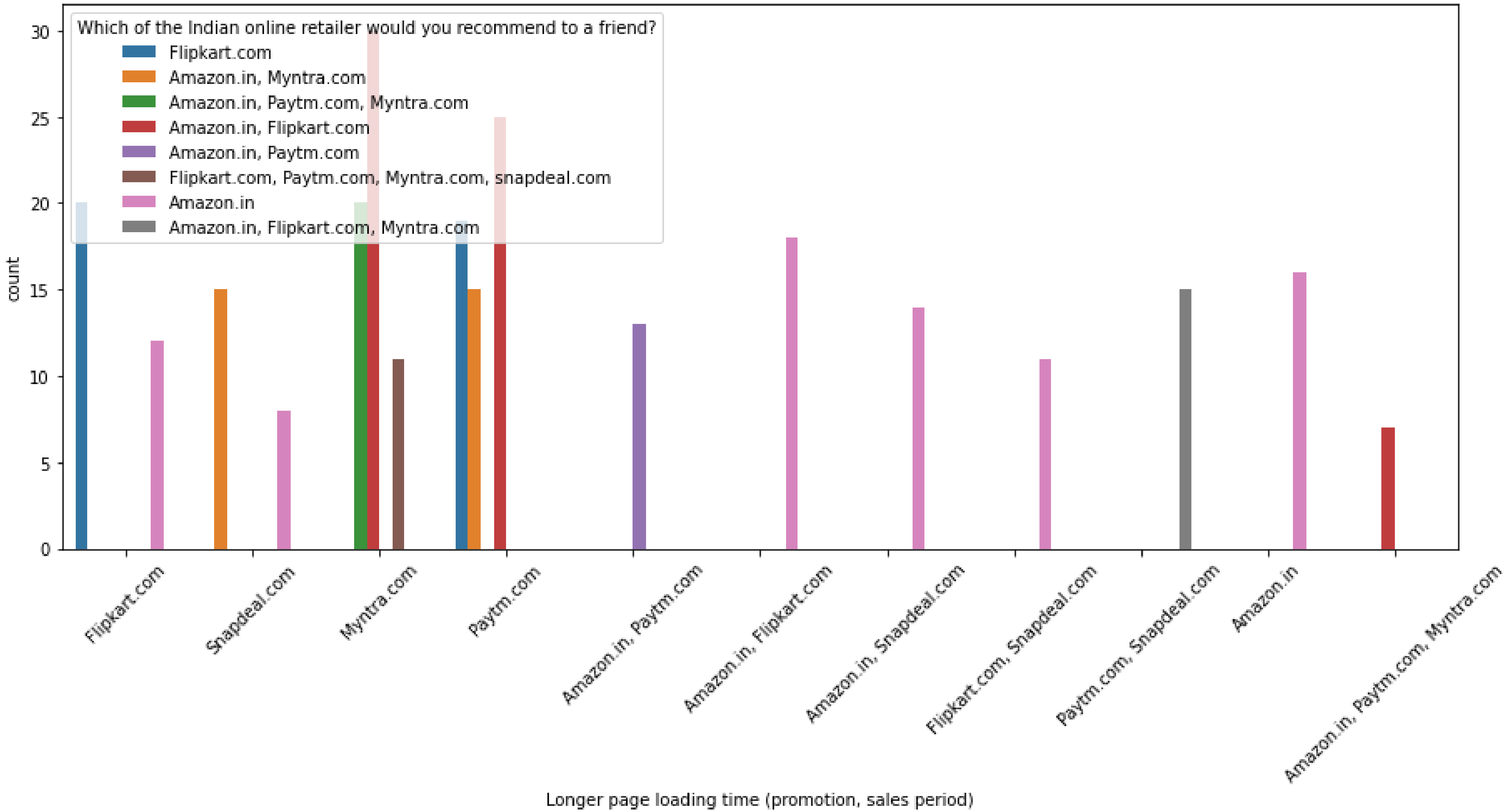

```
<Figure size 432x288 with 0 Axes>
```



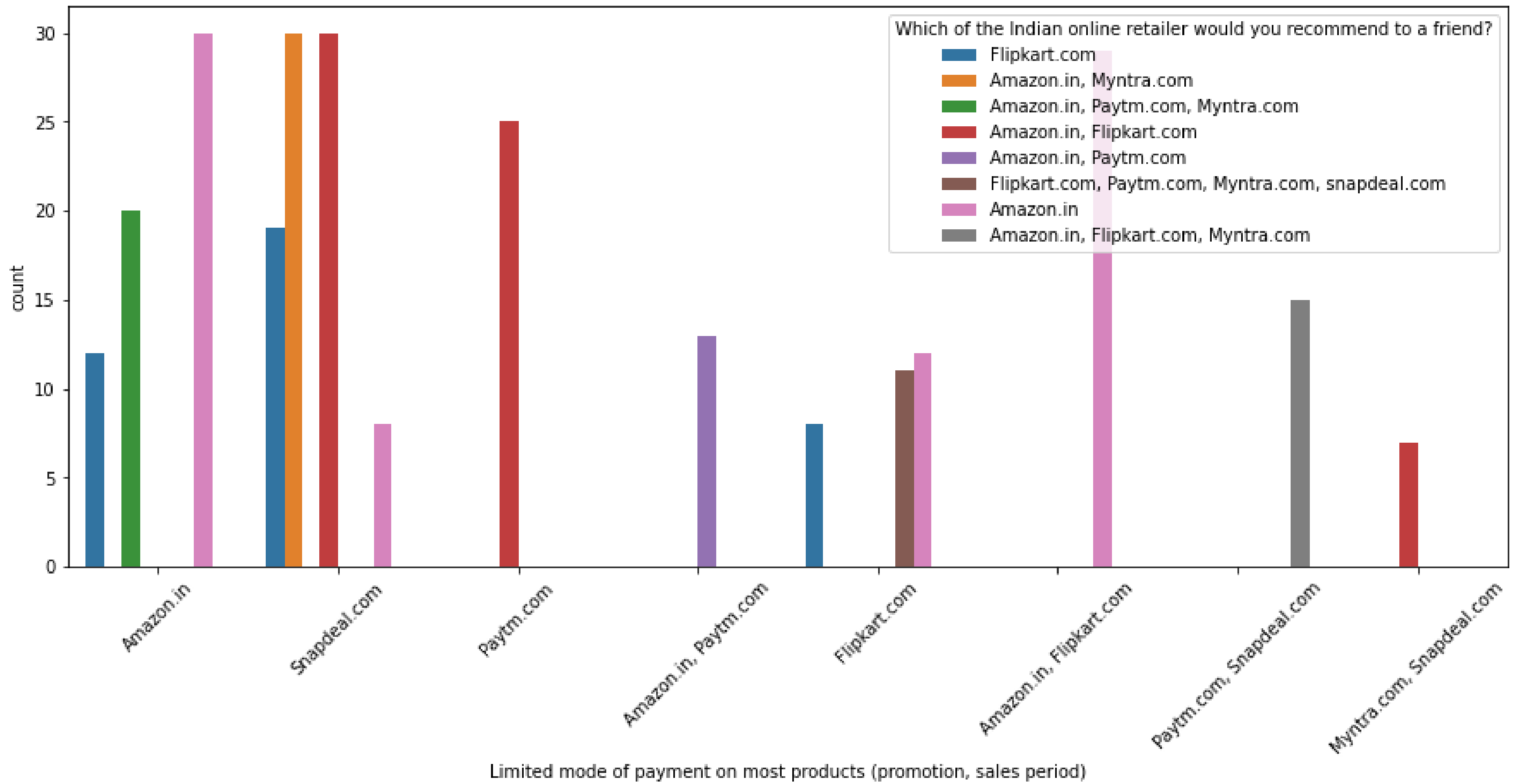

```
<Figure size 432x288 with 0 Axes>
```



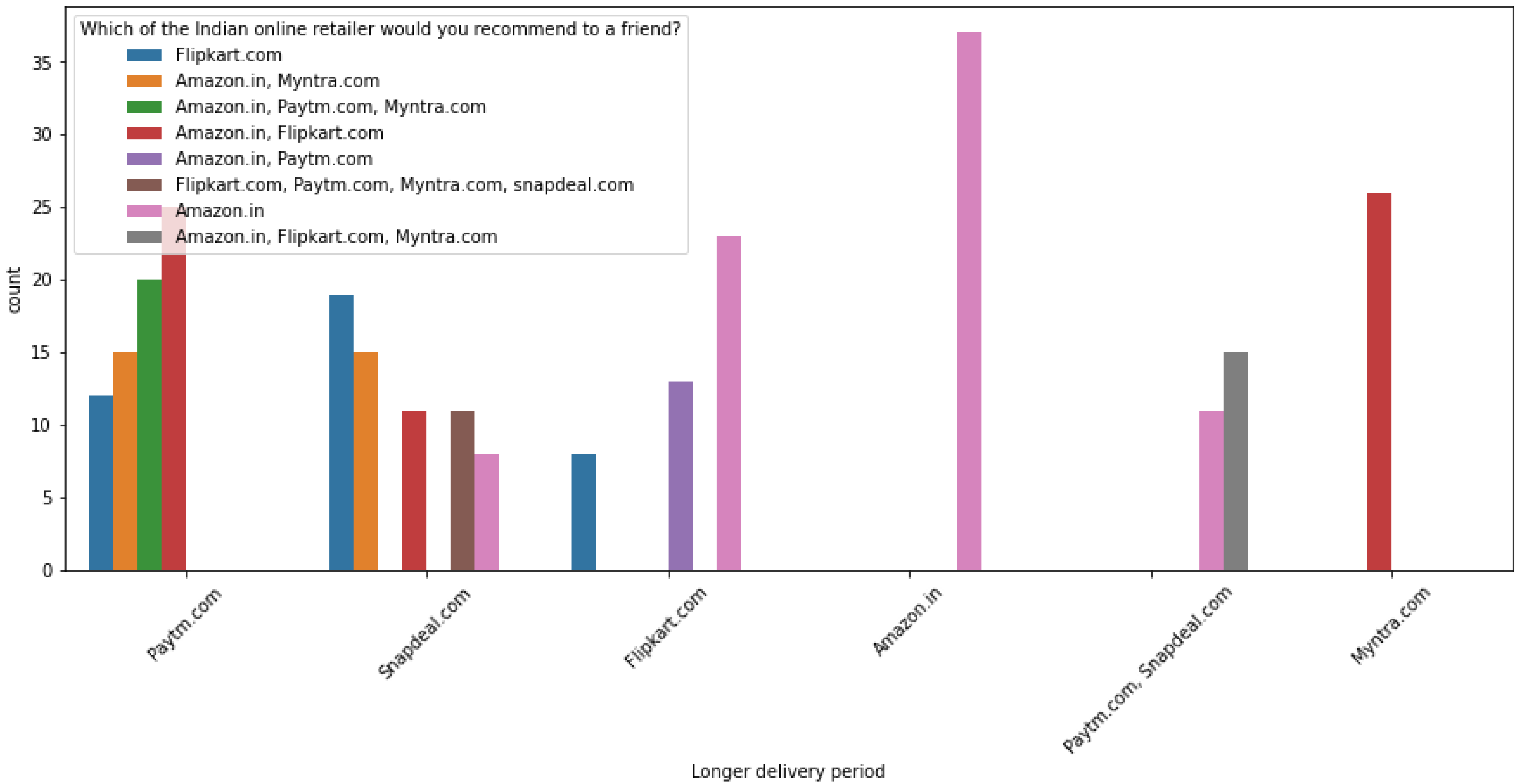

```
<Figure size 432x288 with 0 Axes>
```



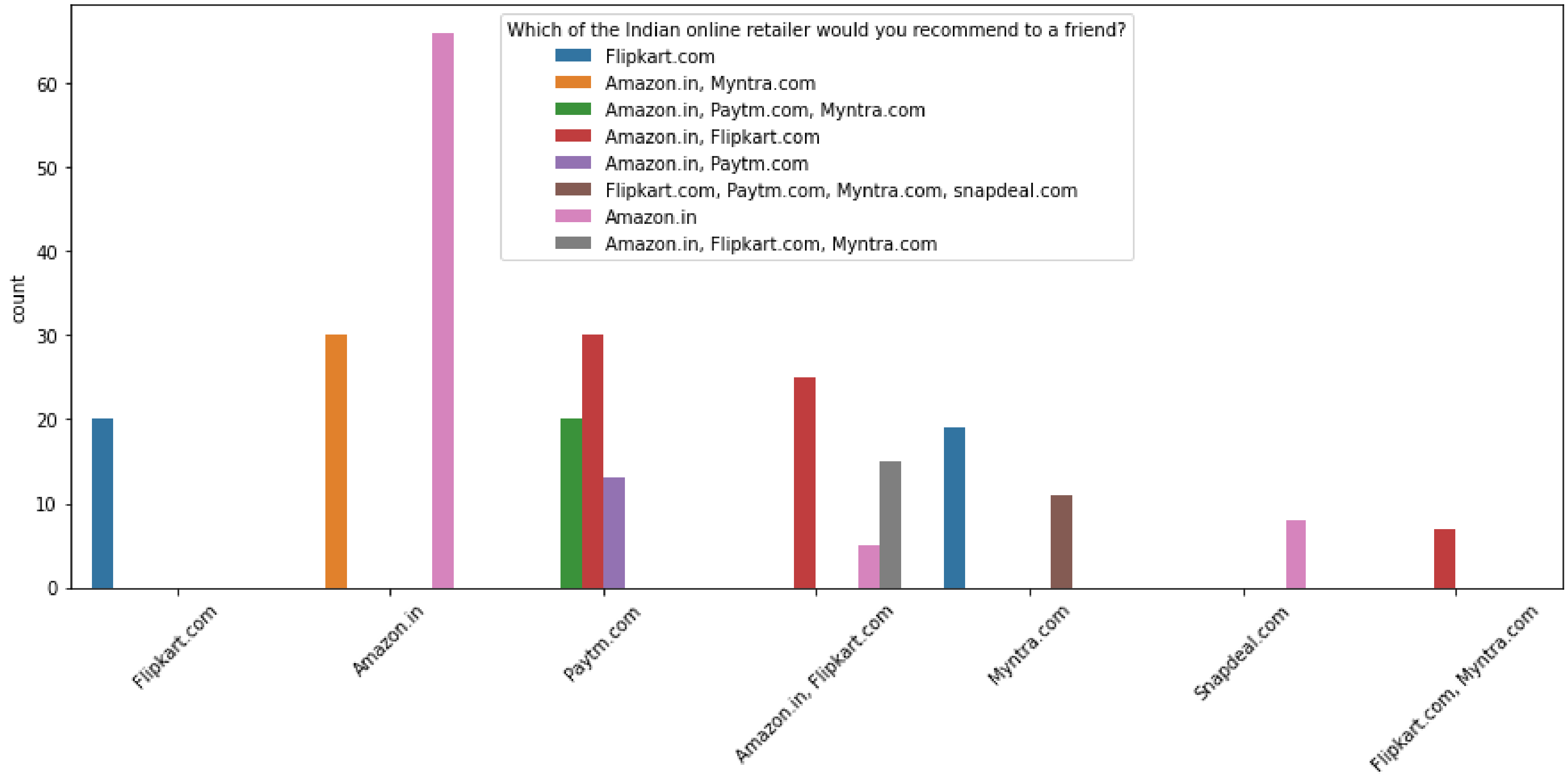

```
<Figure size 432x288 with 0 Axes>
```



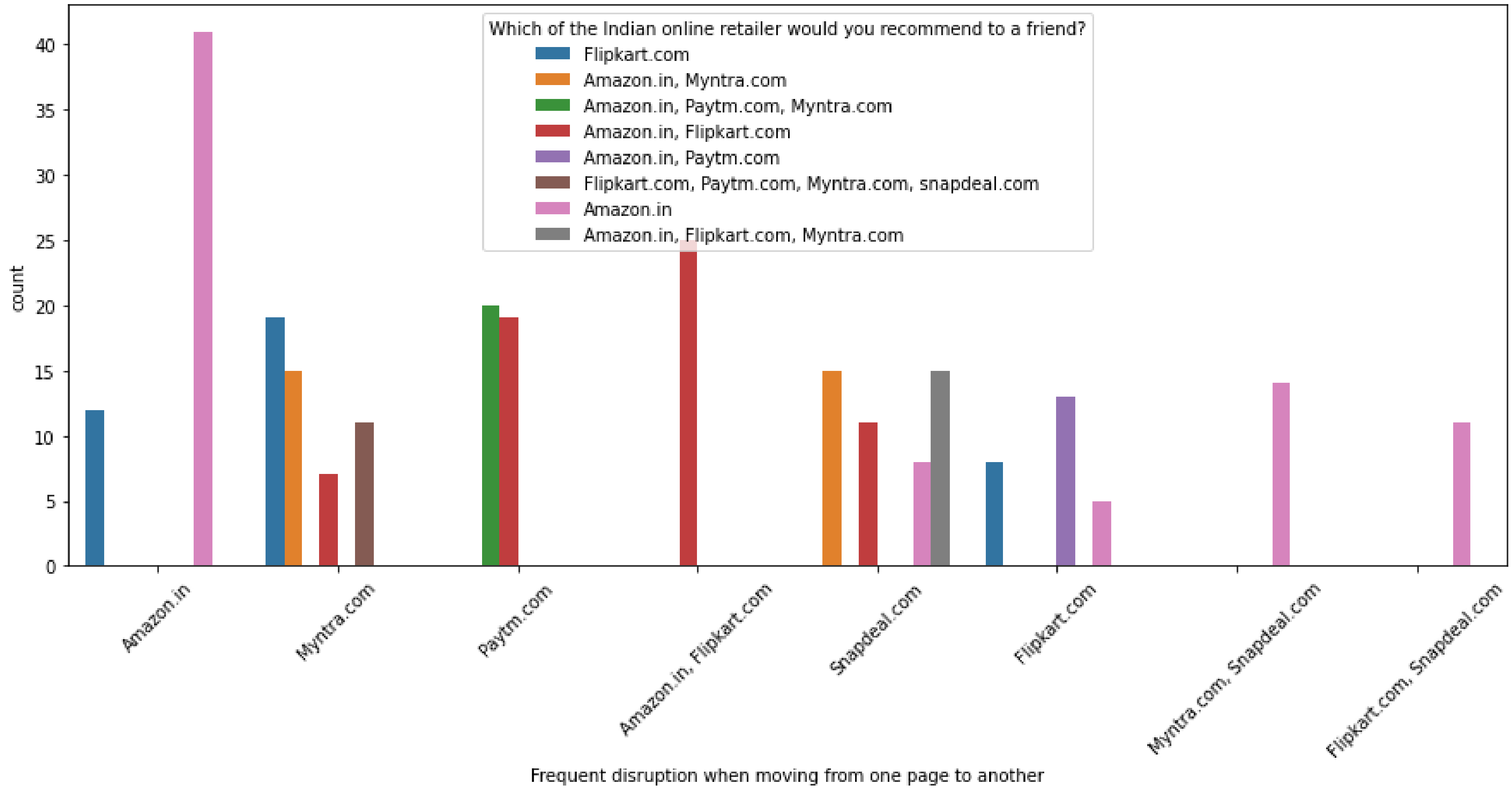

```
<Figure size 432x288 with 0 Axes>
```




```
<Figure size 432x288 with 0 Axes>
```




```
<Figure size 432x288 with 0 Axes>
```




```
<Figure size 432x288 with 0 Axes>
```

Customers seem to be more loyal to amazon, flipkart and paytm as even though many of them have given negative remarks about them still they would recommend these platforms to their friend

Processing the dataframe

Separating the label from rest of the features

In [39]:

```
x=newdata.copy()
x.drop('Which of the Indian online retailer would you recommend to a friend?',axis=1,inplace=True)
y=newdata['Which of the Indian online retailer would you recommend to a friend?']
```

In [40]:

```
#Encoding Categorical Features
cat=[i for i in x.columns if x[i].dtypes=='O']
```

In [41]:

```
from sklearn.preprocessing import OrdinalEncoder,LabelEncoder
encode=OrdinalEncoder()
labe=LabelEncoder()
```

In [42]:

```
#using ordinal encoder for independent features
for i in cat:
    x[i]=encode.fit_transform(x[i].values.reshape(-1,1))

#Using label encoder for Label Column
y=labe.fit_transform(y)
```

Scaling

In [43]:

```
from sklearn.preprocessing import MinMaxScaler
scaler=MinMaxScaler()
```

In [44]:

```
xd=scaler.fit_transform(x)
x=pd.DataFrame(xd,columns=x.columns)
```

Using various feature selection method to see which feature affects the most

Using Feature importance of random forrest

In [45]:

```
from sklearn.ensemble import RandomForestClassifier
m=RandomForestClassifier()
m.fit(x,y)
```

Out[45]:

```
RandomForestClassifier()
```

In [46]:

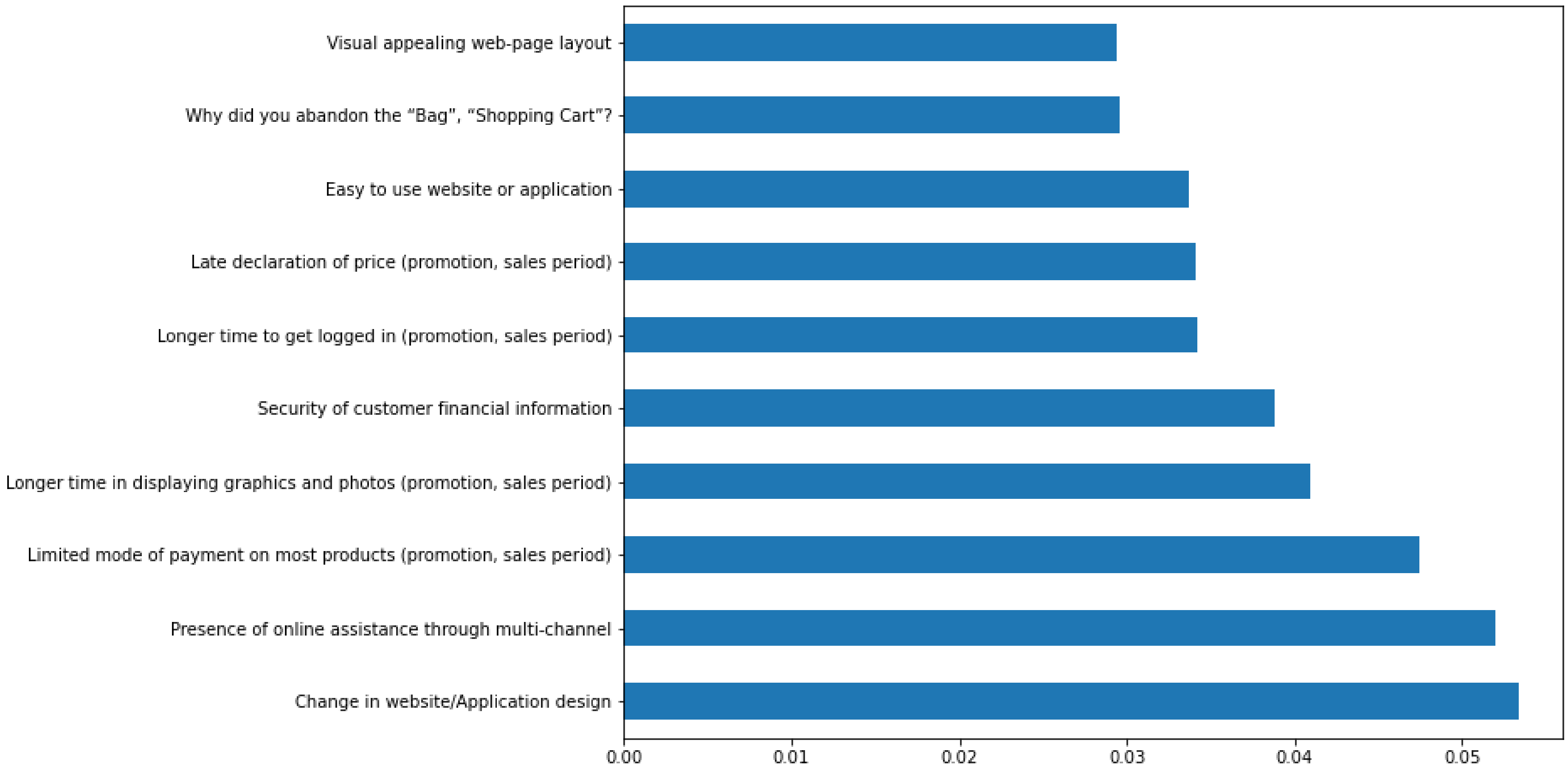
```
RandomForestClassifier()
```

Out[46]:

```
RandomForestClassifier()
```

In [47]:

```
#plot graph of feature importances for better visualization
feat_importances = pd.Series(m.feature_importances_, index=x.columns)
plt.figure(figsize=(10,8))
feat_importances.nlargest(10).plot(kind='barh')
plt.show()
```



Using chi2 test

In [48]:

```
from sklearn.feature_selection import SelectKBest
from sklearn.feature_selection import chi2
```

In [49]:

```
selection = SelectKBest(score_func=chi2)
fit = selection.fit(x,y)
```

In [50]:

```
dfscores = pd.DataFrame(fit.scores_)
dfcolumns = pd.DataFrame(x.columns)
featureScores = pd.concat([dfcolumns,dfscores],axis=1)
featureScores.columns = ['Features','Score']  #naming the dataframe columns
```

In [51]:

```
print(featureScores.nlargest(10,'Score'))  #print10 best features
feat=list(featureScores.nlargest(10,'Score')['Features'])
```

```
                                             Features      Score
16    Why did you abandon the “Bag”, “Shopping Cart”?  75.754028
22                       Loading and processing speed  59.810983
42  Shopping on the website gives you the sense of...  59.253569
10  What browser do you run on your device to acce...  57.171099
67               Change in website/Application design  55.301526
49                   Visual appealing web-page layout  54.245760
65  Limited mode of payment on most products (prom...  53.269266
61  Longer time to get logged in (promotion, sales...  48.222655
62  Longer time in displaying graphics and photos ...  48.130643
50                   Wild variety of product on offer  47.605973
```

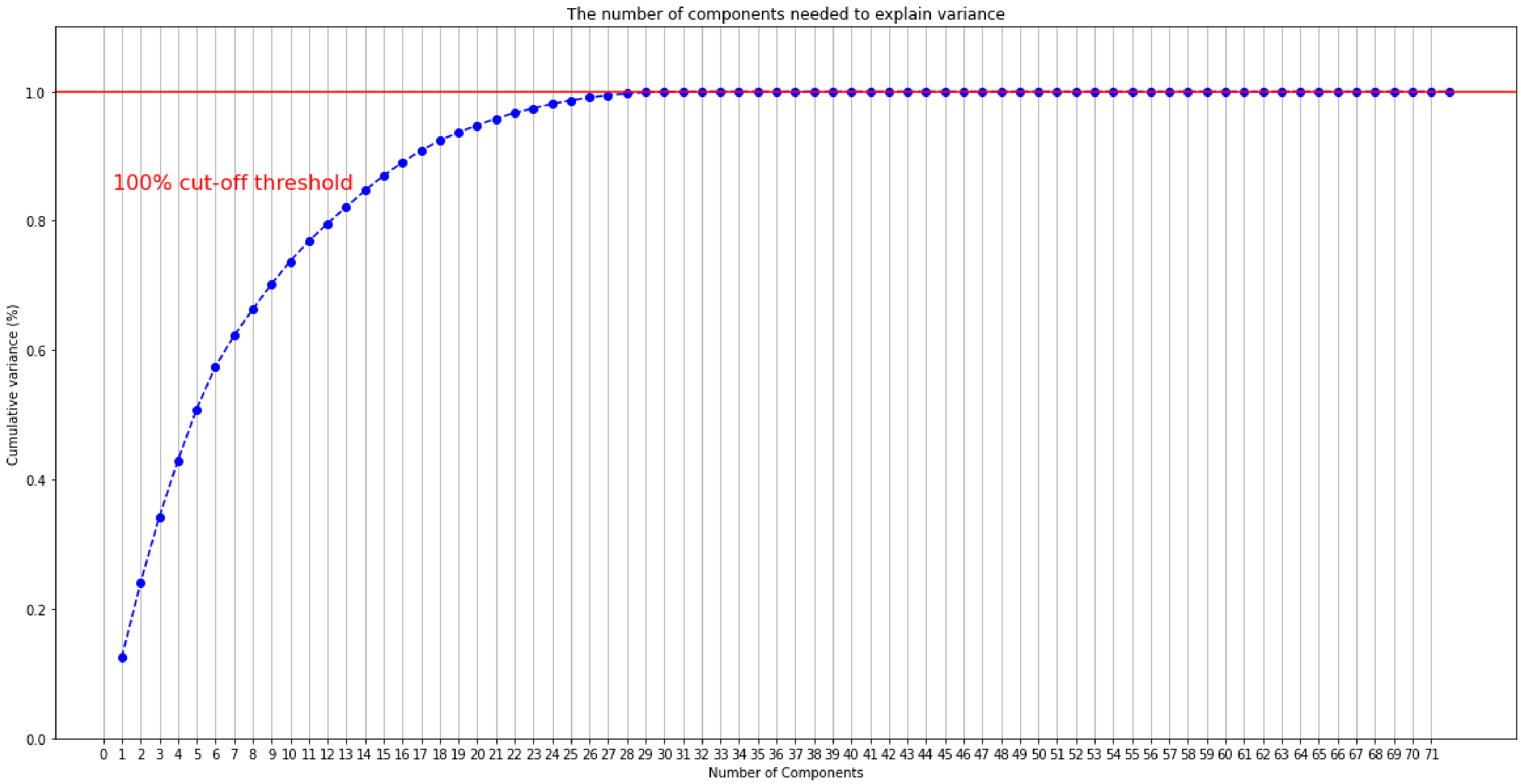
In [52]:

```
from sklearn.decomposition import PCA
pca = PCA().fit(x)
```

In [53]:

```
fig, ax = plt.subplots(figsize=(20,10))
xi = np.arange(1, 73, step=1)
yi = np.cumsum(pca.explained_variance_ratio_)

plt.ylim(0.0,1.1)
plt.plot(xi, yi, marker='o', linestyle='--', color='b')

plt.xlabel('Number of Components')
plt.xticks(np.arange(0, 72, step=1)) #change from 0-based array index to 1-based human-readable label
```

```
plt.ylabel('Cumulative variance (%)')
plt.title('The number of components needed to explain variance')

plt.axhline(y=1, color='r', linestyle='-')
plt.text(0.5, 0.85, '100% cut-off threshold', color = 'red', fontsize=16)

ax.grid(axis='x')
plt.show()
```



We can clearly see that with 29 features all the information can be retained

In [54]:

```
pca=PCA(n_components=29)
x=pca.fit_transform(x)
x=pd.DataFrame(x)
x.head()
```

Out[54]:

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
0	2.065419	-0.577759	-1.030081	-1.109784	0.652387	-1.137025	0.699876	-0.023177	-0.960103	-0.238855	-0.436650	-0.539191	0.130180	0.178306	-0.104644
1	0.048667	-1.490547	1.081348	0.641617	0.066388	-0.820495	0.072214	-0.644870	0.087754	-0.296247	-0.157354	0.881935	0.648067	0.345759	0.329984
2	1.671684	-0.120022	0.775570	-1.481374	0.128287	0.836151	-0.793600	0.102789	0.448813	-0.515949	-0.033307	-0.086125	0.368685	-0.447427	0.281074
3	-0.009522	2.146296	0.753236	-0.363176	-1.348954	-0.176575	0.567430	-0.548924	-0.142604	-0.084665	-0.341339	0.095133	0.089171	0.086230	-0.201624
4	0.051352	-0.187387	2.386865	0.914150	0.273219	-0.992250	-0.511792	0.701105	-0.225943	0.735107	0.138216	-0.814487	-0.204856	-0.294430	-0.127501

	15	16	17	18	19	20	21	22	23	24	25	26	27	28
0	-0.339286	0.379243	-0.431515	-0.272189	0.598931	0.068875	-0.266070	-0.009322	-0.128444	-0.024680	-0.225888	-0.104111	-0.203283	-0.207393
1	0.372285	-0.486708	-0.456116	-0.403685	-0.176390	-0.008384	0.155024	0.313679	0.079454	-0.162517	-0.101240	0.295586	-0.135537	0.122856
2	0.230154	-0.136122	0.074572	0.140000	0.038239	-0.068419	0.008284	0.215976	-0.037138	0.094985	-0.174314	0.156931	-0.006150	-0.166877
3	-0.132961	0.167764	0.160252	0.075425	0.025910	0.229481	-0.091051	0.190278	-0.069483	-0.059019	-0.140161	0.103730	0.035920	-0.116666
4	0.336204	0.214469	-0.030964	0.205835	-0.032298	0.130742	-0.195750	-0.163709	-0.071655	-0.186014	0.047217	0.185140	-0.035241	0.134780

Modelling Phase

In [55]:

```
from sklearn.model_selection import train_test_split,cross_val_score

from sklearn.ensemble import RandomForestClassifier
from xgboost import XGBClassifier

from sklearn.metrics import accuracy_score,confusion_matrix,classification_report,roc_auc_score,roc_curve
```

In [56]:

```
xtrain,xtest,ytrain,ytest=train_test_split(x,y,test_size=0.3,random_state=7)
```

Random Forest

In [57]:

```
model=RandomForestClassifier()
model.fit(xtrain,ytrain)
p=model.predict(xtest)
s=cross_val_score(model,x,y,cv=10)
```

In [58]:

```
print('Accuracy',np.round(accuracy_score(p,ytest),4))
print('-----------------------------------------------------------')
print('Mean of Cross Validation Score',np.round(s.mean(),4))
print('-----------------------------------------------------------')
print('Confusion Matrix')
print(confusion_matrix(p,ytest))
print('-----------------------------------------------------------')
print('Classification Report')
print(classification_report(p,ytest))
```

```
Accuracy 1.0
----------------------------------------------------------
Mean of Cross Validation Score 0.9926
```

```
----------------------------------------------------------
Confusion Matrix
[[26  0  0  0  0  0  0  0]
 [ 0 22  0  0  0  0  0  0]
 [ 0  0  4  0  0  0  0  0]
 [ 0  0  0  4  0  0  0  0]
 [ 0  0  0  0  5  0  0  0]
 [ 0  0  0  0  0  7  0  0]
 [ 0  0  0  0  0  0 11  0]
 [ 0  0  0  0  0  0  0  2]]
----------------------------------------------------------
Classification Report
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        26
           1       1.00      1.00      1.00        22
           2       1.00      1.00      1.00         4
           3       1.00      1.00      1.00         4
           4       1.00      1.00      1.00         5
           5       1.00      1.00      1.00         7
           6       1.00      1.00      1.00        11
           7       1.00      1.00      1.00         2

    accuracy                           1.00        81
   macro avg       1.00      1.00      1.00        81
weighted avg       1.00      1.00      1.00        81
```

Xgboost

In [59]:

```
model=XGBClassifier(verbosity=0)
model.fit(xtrain,ytrain)
p=model.predict(xtest)
s=cross_val_score(model,x,y,cv=10)
```

In [60]:

```
print('Accuracy',np.round(accuracy_score(p,ytest),4))
print('--------------------------------------------------------')
print('Mean of Cross Validation Score',np.round(s.mean(),4))
print('--------------------------------------------------------')
print('Confusion Matrix')
print(confusion_matrix(p,ytest))
print('--------------------------------------------------------')
print('Classification Report')
print(classification_report(p,ytest))
```

```
Accuracy 1.0
----------------------------------------------------------
Mean of Cross Validation Score 0.9926
----------------------------------------------------------
Confusion Matrix
[[26  0  0  0  0  0  0  0]
 [ 0 22  0  0  0  0  0  0]
 [ 0  0  4  0  0  0  0  0]
 [ 0  0  0  4  0  0  0  0]
 [ 0  0  0  0  5  0  0  0]
 [ 0  0  0  0  0  7  0  0]
 [ 0  0  0  0  0  0 11  0]
 [ 0  0  0  0  0  0  0  2]]
----------------------------------------------------------
Classification Report
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        26
           1       1.00      1.00      1.00        22
           2       1.00      1.00      1.00         4
           3       1.00      1.00      1.00         4
           4       1.00      1.00      1.00         5
           5       1.00      1.00      1.00         7
           6       1.00      1.00      1.00        11
           7       1.00      1.00      1.00         2

    accuracy                           1.00        81
   macro avg       1.00      1.00      1.00        81
weighted avg       1.00      1.00      1.00        81
```

Hyperparameter Tuning

In [61]:

```
from sklearn.model_selection import RandomizedSearchCV
```

Random Forest

In [62]:

```
params={'n_estimators':[100, 300, 500, 700],
        'min_samples_split':[1,2,3,4],
        'min_samples_leaf':[1,2,3,4],
        'max_depth':[None,1,2,3,4,5,6,7,8,9,10,15,20,25,30,35,40]}
```

In [63]:

```
g=RandomizedSearchCV(RandomForestClassifier(),params,cv=10)
```

In [64]:

```
g.fit(xtrain,ytrain)
```

Out[64]:

```
RandomizedSearchCV(cv=10, estimator=RandomForestClassifier(),
                   param_distributions={'max_depth': [None, 1, 2, 3, 4, 5, 6, 7,
                                                      8, 9, 10, 15, 20, 25, 30,
                                                      35, 40],
                                        'min_samples_leaf': [1, 2, 3, 4],
                                        'min_samples_split': [1, 2, 3, 4],
                                        'n_estimators': [100, 300, 500, 700]})
```

In [65]:

```
print(g.best_estimator_)
print(g.best_params_)
print(g.best_score_)
```

```
RandomForestClassifier(max_depth=30, min_samples_leaf=2, min_samples_split=4,
                       n_estimators=700)
{'n_estimators': 700, 'min_samples_split': 4, 'min_samples_leaf': 2, 'max_depth': 30}
0.9947368421052631
```

In [66]:

```
m=RandomForestClassifier(max_depth=20, min_samples_leaf=4, min_samples_split=4,n_estimators=700)
m.fit(xtrain,ytrain)
p=m.predict(xtest)
score=cross_val_score(m,x,y,cv=10)
```

In [67]:

```
print('Accuracy',np.round(accuracy_score(p,ytest),4))
print('------------------------------------------------------------')
print('Mean of Cross Validation Score',np.round(s.mean(),4))
print('------------------------------------------------------------')
print('Confusion Matrix')
print(confusion_matrix(p,ytest))
print('------------------------------------------------------------')
print('Classification Report')
print(classification_report(p,ytest))
```

```
Accuracy 1.0
------------------------------------------------------------
Mean of Cross Validation Score 0.9926
------------------------------------------------------------
Confusion Matrix
[[26  0  0  0  0  0  0  0]
 [ 0 22  0  0  0  0  0  0]
 [ 0  0  4  0  0  0  0  0]
 [ 0  0  0  4  0  0  0  0]
 [ 0  0  0  0  5  0  0  0]
 [ 0  0  0  0  0  7  0  0]
 [ 0  0  0  0  0  0 11  0]
 [ 0  0  0  0  0  0  0  2]]
------------------------------------------------------------
Classification Report
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        26
           1       1.00      1.00      1.00        22
           2       1.00      1.00      1.00         4
           3       1.00      1.00      1.00         4
           4       1.00      1.00      1.00         5
```

```
           5       1.00      1.00      1.00         7
           6       1.00      1.00      1.00        11
           7       1.00      1.00      1.00         2

    accuracy                           1.00        81
   macro avg       1.00      1.00      1.00        81
weighted avg       1.00      1.00      1.00        81
```

Xgboost

In [68]:

```
params={'n_estimators':[100,200,300,400,500],
      'learning_rate':[0.001,0.01,0.10,],
      'subsample':[0.5,1],
      'max_depth':[1,2,3,4,5,6,7,8,9,10]}
```

In [69]:

```
g=RandomizedSearchCV(XGBClassifier(),params,cv=10)
```

In [70]:

```
g.fit(xtrain,ytrain)
```

Out[70]:

```
RandomizedSearchCV(cv=10,
                   estimator=XGBClassifier(base_score=None, booster=None,
                                           callbacks=None,
                                           colsample_bylevel=None,
                                           colsample_bynode=None,
                                           colsample_bytree=None,
                                           early_stopping_rounds=None,
                                           enable_categorical=False,
                                           eval_metric=None, gamma=None,
                                           gpu_id=None, grow_policy=None,
                                           importance_type=None,
                                           interaction_constraints=None,
                                           learning_rate=None, max_bin=None...
                                           max_delta_step=None, max_depth=None,
                                           max_leaves=None,
                                           min_child_weight=None, missing=nan,
                                           monotone_constraints=None,
```

```
                                        n_estimators=100, n_jobs=None,
                                        num_parallel_tree=None,
                                        predictor=None, random_state=None,
                                        reg_alpha=None, reg_lambda=None, ...),
                   param_distributions={'learning_rate': [0.001, 0.01, 0.1],
                                        'max_depth': [1, 2, 3, 4, 5, 6, 7, 8, 9,
                                                      10],
                                        'n_estimators': [100, 200, 300, 400,
                                                         500],
                                        'subsample': [0.5, 1]})
```

In [71]:

```
print(g.best_estimator_)
print(g.best_params_)
print(g.best_score_)
```

```
XGBClassifier(base_score=0.5, booster='gbtree', callbacks=None,
              colsample_bylevel=1, colsample_bynode=1, colsample_bytree=1,
              early_stopping_rounds=None, enable_categorical=False,
              eval_metric=None, gamma=0, gpu_id=-1, grow_policy='depthwise',
              importance_type=None, interaction_constraints='',
              learning_rate=0.1, max_bin=256, max_cat_to_onehot=4,
              max_delta_step=0, max_depth=7, max_leaves=0, min_child_weight=1,
              missing=nan, monotone_constraints='()', n_estimators=100,
              n_jobs=0, num_parallel_tree=1, objective='multi:softprob',
              predictor='auto', random_state=0, reg_alpha=0, ...)
{'subsample': 0.5, 'n_estimators': 100, 'max_depth': 7, 'learning_rate': 0.1}
0.9894736842105264
```

In [72]:

```
m=XGBClassifier(max_depth=10,learning_rate=0.1,n_estimators=500,subsample=0.5)
m.fit(xtrain,ytrain)
p=m.predict(xtest)
score=cross_val_score(m,x,y,cv=10)
```

In [73]:

```
print('Accuracy',np.round(accuracy_score(p,ytest),4))
print('----------------------------------------------------------')
print('Mean of Cross Validation Score',np.round(s.mean(),4))
print('----------------------------------------------------------')
print('Confusion Matrix')
```

```
print(confusion_matrix(p,ytest))
print('-------------------------------------------------------------')
print('Classification Report')
print(classification_report(p,ytest))
```

```
Accuracy 1.0
-----------------------------------------------------------
Mean of Cross Validation Score 0.9926
-----------------------------------------------------------
Confusion Matrix
[[26  0  0  0  0  0  0  0]
 [ 0 22  0  0  0  0  0  0]
 [ 0  0  4  0  0  0  0  0]
 [ 0  0  0  4  0  0  0  0]
 [ 0  0  0  0  5  0  0  0]
 [ 0  0  0  0  0  7  0  0]
 [ 0  0  0  0  0  0 11  0]
 [ 0  0  0  0  0  0  0  2]]
-----------------------------------------------------------
Classification Report
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        26
           1       1.00      1.00      1.00        22
           2       1.00      1.00      1.00         4
           3       1.00      1.00      1.00         4
           4       1.00      1.00      1.00         5
           5       1.00      1.00      1.00         7
           6       1.00      1.00      1.00        11
           7       1.00      1.00      1.00         2

    accuracy                           1.00        81
   macro avg       1.00      1.00      1.00        81
weighted avg       1.00      1.00      1.00        81
```

Conclusion Both the models give accurate and equal results so we choose xgboost as or final model because of its quick speed.

Finalizing the best Model

In [74]:

```
model=XGBClassifier(max_depth=2,learning_rate=0.01,n_estimators=500,subsample=1)
model.fit(xtrain,ytrain)
p=model.predict(xtest)
score=cross_val_score(model,x,y,cv=10)
```

Evaluation Metrics

In [75]:

```
print('Accuracy',np.round(accuracy_score(p,ytest),4))
print('----------------------------------------------------------')
print('Mean of Cross Validation Score',np.round(s.mean(),4))
print('----------------------------------------------------------')
print('Confusion Matrix')
print(confusion_matrix(p,ytest))
print('----------------------------------------------------------')
print('Classification Report')
print(classification_report(p,ytest))
```

```
Accuracy 1.0
----------------------------------------------------------
Mean of Cross Validation Score 0.9926
----------------------------------------------------------
Confusion Matrix
[[26  0  0  0  0  0  0  0]
 [ 0 22  0  0  0  0  0  0]
 [ 0  0  4  0  0  0  0  0]
 [ 0  0  0  4  0  0  0  0]
 [ 0  0  0  0  5  0  0  0]
 [ 0  0  0  0  0  7  0  0]
 [ 0  0  0  0  0  0 11  0]
 [ 0  0  0  0  0  0  0  2]]
----------------------------------------------------------
Classification Report
              precision    recall  f1-score   support
```

```
           0       1.00      1.00      1.00        26
           1       1.00      1.00      1.00        22
           2       1.00      1.00      1.00         4
           3       1.00      1.00      1.00         4
           4       1.00      1.00      1.00         5
           5       1.00      1.00      1.00         7
           6       1.00      1.00      1.00        11
           7       1.00      1.00      1.00         2

    accuracy                           1.00        81
   macro avg       1.00      1.00      1.00        81
weighted avg       1.00      1.00      1.00        81
```

In [ ]: