EDA on Customer Retention dataset.

ACKNOWLEDGMENT

I have taken efforts in this project. However, it would not have been possible without the kind support and help of many individuals and organizations. I would like to extend my sincere thanks to all of them. I am highly indebted to Flip Robo Technologies Bangalore for their guidance and constant supervision as well as for providing necessary information regarding the project & also for their support in completing the project. I want to thank my SME Ms. Shristhi Maan for providing the Dataset and helping us to solve the problem and addressing out our Query in right time. I would like to express my gratitude towards my parents & members of Flip Robo for their kind co-operation and encouragement which help me in completion of this project. I would like to express my special gratitude and thanks to industry persons for giving me such attention and time.

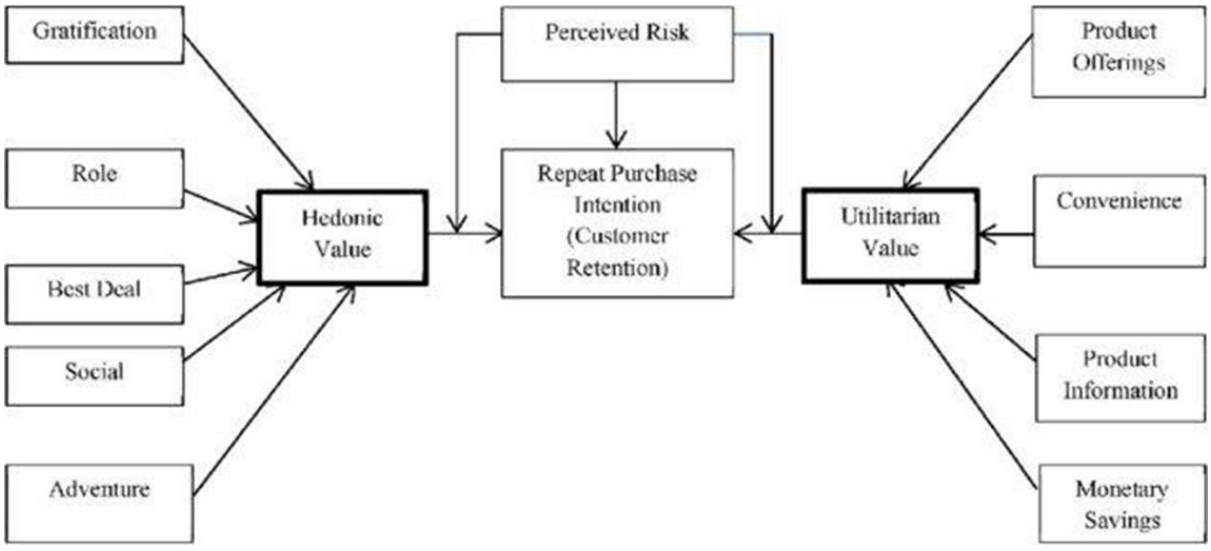
INTRODUCTION

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 ecommerce 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. Conceptual Background of the Domain Problem The data is collected from the Indian online shoppers. Results indicate the e-retail success factors, which are very much critical for customer satisfaction. Five major factors that contributed to the success of an ecommerce 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. Motivation for the Problem Undertaken Our main objective of doing this project is to analyse whether the users are shopping products from ecommerce websites, how did they give feedbacks to these websites on the basis of several

positive and negative factors and also the details of the users on basis of factors like age, gender, etc.

Diagrammatic Representation of Customer Retention



The Hedonic value consists of factors like Gratification, Role, Best Deal, Social and Adventure. The Utilitarian

value consists of factors like Product Offerings, Convenience, Product Information and Monetary Savings. Customer Retention is based on 3 factors, according to the above diagram. They are: Perceived Risk, Hedonic value and Utilitarian value . Data Sources and their formats The data is been given by a highly-confidential company and they gave it to us in an excel file. They also had provided the problem statement by explaining what they need from us and also the required criteria to be satisfied. Let's check the data now. Below I have attached the snapshot below to give an overview.

```
data=pd.read_excel("D:/fliprobo/project/P1/Customer_retention_dataset_1/Customer_retention_dataset/Customer_retention_datase
data.head().T
```

	0	1	2	3	4
1Gender of respondent	Male	Female	Female	Male	Female
2 How old are you?	31-40 years	21-30 years	21-30 years	21-30 years	21-30 years
3 Which city do you shop online from?	Delhi	Delhi	Greater Noida	Karnal	Bangalore
4 What is the Pin Code of where you shop online from?	110009	110030	201308	132001	530068
5 Since How Long You are Shopping Online ?	Above 4 years	Above 4 years	3-4 years	3-4 years	2-3 years
...	...	...	...	...	...
Longer delivery period	Paytm.com	Snapdeal.com	Paytm.com	Paytm.com	Paytm.com
Change in website/Application design	Flipkart.com	Amazon.in	Paytm.com	Amazon.in, Flipkart.com	Amazon.in
Frequent disruption when moving from one page to another	Amazon.in	Myntra.com	Paytm.com	Amazon.in, Flipkart.com	Snapdeal.com
Website is as efficient as before	Amazon.in	Amazon.in, Flipkart.com	Amazon.in	Amazon.in, Flipkart.com, Paytm.com	Paytm.com
Which of the Indian online retailer would you recommend to a friend?	Flipkart.com	Amazon.in, Myntra.com	Amazon.in, Paytm.com, Myntra.com	Amazon.in, Flipkart.com	Amazon.in, Myntra.com

71 rows × 5 columns

There are totally 269 rows and 71 columns in this dataset.

Our objective is to find the insights of the data and to do thorough data analysis.

Hardware and Software Requirements and Tools Used For doing this project, the hardware used is a laptop with high end specification and a stable internet connection. While coming to software part, I had used anaconda navigator and launched Jupyter notebook to do my EDA. For using an excel file, Microsoft excel is needed. In Jupyter notebook, I had used lots of python libraries to carry out this project and I have mentioned below with proper justification:

1. Pandas-: A library which is used to read the data, visualisation and analysis of data.

2. NumPy-: Used for working with array and various mathematical techniques.

3. Seaborn-: Visualization tool for plotting different types of plot.

4. Matplotlib-: It provides an object-oriented API for embedding plots into applications.

Data Analysis::

```
data.columns.tolist()
```

```
['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 decision?
',
```

Firstly, we clear the columns names. As we can see there are spaces in the columns name. So, we strip these columns.

```
In [5]: data.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-n
ull    object
 1   2 How old are you?                                         269 non-n
ull    object
```

Showing information of each columns that is the datatypes of all columns with non-null values.

```
In [6]: for col in data:
            print(f"{col}::\n----------\nUnique values::{data[col].unique()}\n----------\n")
```

```
1Gender of respondent::
----------
Unique values::['Male' 'Female']
----------

2 How old are you?::
----------
Unique values::['31-40 years' '21-30 years' '41-50 yaers' 'Less than 20 years'
 '51 years and above']
----------

3 Which city do you shop online from?::
```

```
In [6]:
for col in data:
    print(f"{col}::\n----------\nUnique values::{data[col].unique()}\n----------\n")
```

```
1Gender of respondent::
----------
Unique values::['Male' 'Female']
----------

2 How old are you?::
----------
Unique values::['31-40 years' '21-30 years' '41-50 yaers' 'Less than 20 years'
 '51 years and above']
----------

3 Which city do you shop online from?::
```

Showing Unique values of all columns. And analyse that the internet access has values like Mobile internet and Mobile Internet. So, we replace to get one value like replaced Mobile internet with Mobile Internet.

```
data.shape
```

```
(269, 71)
```

There are now 269 no. of records.

```
data.isnull().sum()
```

```
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
```

Showing shape that has 269 no. of rows and 71 columns.

And no null values in any columns.

In [14]:
```
for col in data:
    print(col)
    print(data[col].value_counts())
    print("-"*40)
    print("-"*40)
```

```
1Gender of respondent
Female    181
Male       88
Name: 1Gender of respondent, dtype: int64
----------------------------------------
----------------------------------------
2 How old are you?
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
----------------------------------------
```

No. of each value present in their respective columns.

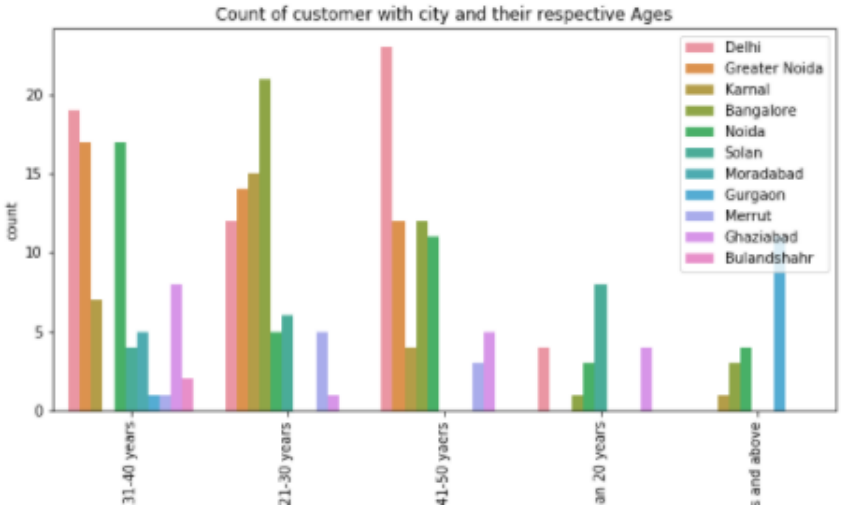
In [15]:
```
customerDetails=data.iloc[:,:7]        # sepearating data for taking details of Customer
deviceDeatils=data.iloc[:,7:11]        # Taking Deatails of device seperately
firstVisit=data.iloc[:,11:15]          # FirstVisit of customer through which platform
abondon=data.iloc[:,15:17]             # customer abondon the products reasons
services=data.iloc[:,17:47]            # Sevices of customer
onlineReatilers=data.iloc[:,47:]       # Online retailers feedback by customer details.
```

In [22]:
```
fig=plt.figure(figsize=(10,5))
data[['1Gender of respondent',""'2 How old are you?']].value_counts().plot(kind='bar')
plt.title("Count of customer with genderrespondent and their respective Ages")
plt.show()
fig.savefig('Count of customer with genderrespondent and their respective Ages.jpg')
```

Then split data into such manner that split data that customer details, Device details, First Visit of Customer, abandoning products by customer and how much, Services provided to customer and the feedbacks that was made by customers.

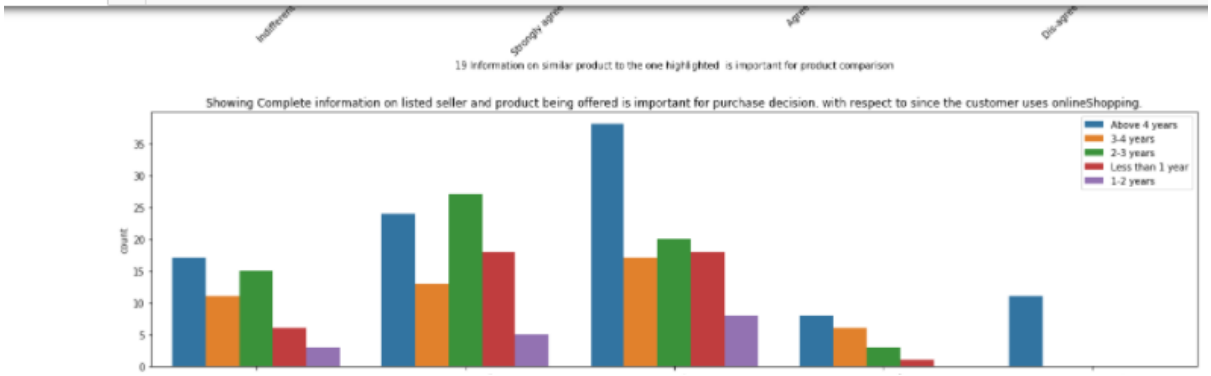
Then plot to know the no. of customer with their respective ages and their gender.

In [25]:
```
fig=plt.figure(figsize=(10,5))
sns.countplot(data['2 How old are you?'],hue=data["3 Which city do you shop online from?"])
plt.xticks(rotation=90)
plt.title("Count of customer with city and their respective Ages")
plt.legend(loc=1)
plt.show()
```



Plotting graph for the customers ages with respect to their city and analyse there are most customer from Delhi.

In [297]:
```
for col in services.iloc[:,:-1]:
    fig=plt.figure(figsize=(20,5))
    sns.countplot(data[col],hue=data["5 Since How Long You are Shopping Online ?"])
    plt.xticks(rotation=45)
    plt.title("Showing "+col[3:]+" with respect to since the customer uses onlineShopping.")
    plt.legend(loc=1)
    plt.show()
```



Plotted graphs since how long the customer using the online Shopping and what they got their services and analyses that :

- Mostly customers who are using websites more than 4 years are not satisfied with information listed seller and product being offered to customer from websites.

- Mostly customer who are using websites more than 3 years are not satisfied by the enjoyment service from online shopping. while others are happy/satisfied with that.
- only new customer who are customers for less than 1 year are convenient of patronizing the online retailer.
- Mostly Customers who are using less than 2 years doesn't feels adventure while using websites.
- Mostly customer who are using websites less than 1 year or 2-3 years of using doesn,t feels gratification while shopping.
- Mostly customer who are using websites less than 1 year doesn't helps to fulfill their certain roles.

Analysis of website feedbacks obtained :

We can see that after column 47, there are both positive and negative feedbacks of the websites, which are given by the correspondents. We will analyse those data by using data analysis process.

In [33]:
```
#A separate dataframe for displaying the positive feedback
posFeed=onlineReatilers.drop(["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", "Frequent disruption when moving from one page to another"], axis=1)
posFeed.head(2).T
```

Out[33]:

	0	1
From the following, tick any (or all) of the online retailers you have shopped from;	Amazon.in, Paytm.com	Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com
Easy to use website or application	Paytm.com	Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com
Visual appealing web-page layout	Flipkart.com	Amazon.in, Myntra.com
Wild variety of product on offer	Flipkart.com	Flipkart.com, Myntra.com
Complete, relevant description information of products	Snapdeal.com	Amazon.in, Flipkart.com, Myntra.com
Fast loading website speed of website and application	Snapdeal.com	Amazon.in, Flipkart.com, Myntra.com
Reliability of the website or application	Paytm.com	Myntra.com
Quickness to complete purchase	Paytm.com	Amazon.com, Flipkart.com, Myntra.com
Availability of several payment options	Patym.com	Amazon.in, Flipkart.com, Myntra.com
Speedy order delivery	Amazon.in	Amazon.in, Flipkart.com
Privacy of customers' information	Amazon.in	Myntra.com
Security of customer financial information	Amazon.in	Myntra.com
Perceived Trustworthiness	Flipkart.com	Myntra.com
Presence of online assistance through multi-channel	Paytm.com	Amazon.in, Flipkart.com, Myntra.com
Change in website/Application design	Flipkart.com	Amazon.in
Website is as efficient as before	Amazon.in	Amazon.in, Flipkart.com
Which of the Indian online retailer would you recommend to a friend?	Flipkart.com	Amazon.in, Myntra.com

First, we will extract only the feedbacks data and then save it in a new data frame, which will be used for further process.

A separate dataframe for displaying the positive feedback posFeed=onlineReatailer.drop(["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", "Frequent disruption when moving from one page to another"], axis=1) # A separate dataframe for displaying the negative feedback negFeed=onlineReatailer [["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", "Frequent disruption when moving from one page to another"]] Now, we will analyse the negative feedbacks first by checking the count of websites and the type of feedbacks given to each website. Then, we will save the obtained data in a new data frame and rename the column names.

.

In [44]:
```
negFeed=onlineReatilers[["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", "Frequent disruption when moving from one page to another"]]
negFeed.head(2).T
```

Out[44]:

	0	1
Longer time to get logged in (promotion, sales period)	Amazon.in	Amazon.in, Flipkart.com
Longer time in displaying graphics and photos (promotion, sales period)	Amazon.in	Myntra.com
Late declaration of price (promotion, sales period)	Flipkart.com	snapdeal.com
Longer page loading time (promotion, sales period)	Flipkart.com	Snapdeal.com
Limited mode of payment on most products (promotion, sales period)	Amazon.in	Snapdeal.com
Longer delivery period	Paytm.com	Snapdeal.com
Frequent disruption when moving from one page to another	Amazon.in	Myntra.com

Now, we will analyse the negative feedbacks by checking the count of websites and the type of feedbacks given to each website. Then, we will save the obtained data in a new data frame and rename the column names.

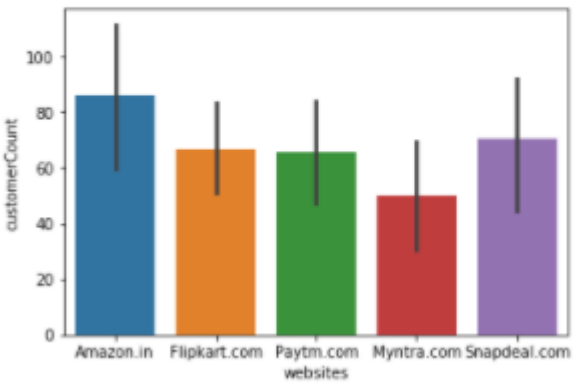
In [35]:
```
#Analysing the negative feedbacks separately
website_list=['Amazon.in','Flipkart.com','Paytm.com','Myntra.com','Snapdeal.com']  #Website list
col_names=[]   #Empty list for column names
websites=[]    #Empty list for websites
count=[]       #Empty list for checking the count of no of times the websites are mentioned

for col in negFeed.columns:
    for i in website_list:
        present=len(negFeed[negFeed[col].str.contains(i)])  #Checking if websites in dataframe are available in the website
        col_names.append(col)   #Appending the column names
        websites.append(i)    #Appending the website names
        count.append(present) #Appending the count ofcustomer at website present in feedback
```

In [36]:
```
#Creating negative feedback dataframe
negative_df=pd.DataFrame([col_names,websites,count],index=["columnsName","websites","customerCount"]).T
negative_df
```

Out[36]:

	columnsName	websites	customerCount
0	Longer time to get logged in (promotion, sales...	Amazon.in	135
1	Longer time to get logged in (promotion, sales...	Flipkart.com	103
2	Longer time to get logged in (promotion, sales...	Paytm.com	77
3	Longer time to get logged in (promotion, sales...	Myntra.com	35

In [37]:
```
#Plotting barplot for people_count vs website relationship
sns.barplot(x='websites',y='customerCount',data=negative_df)
plt.show()
```



Plotted graph showing the no. of customer using particular websites for negative feedbacks.

Extracting feature wise comparison from negative feedback dataframe:

In [38]:
```
#Extracting feature wise comparison from negative feedback dataframe
b=range(0,81,5)
try:
    for i in b:
        features=negative_df.iloc[i:i+5:]
        sns.barplot(x='websites',y='customerCount',data=features)
        plt.title(features['columnsName'][i])
        plt.show()
        print("\n")
except ValueError:   #As the value has some error while iterating, we are passing it by using except
    pass
```

Observations ::

1.Amazon takes longer time to get logged in during promotion, followed by flipkart, paytm and snapdeal.

2.Amazon takes longer time in displaying graphics and photos followed by flipkart and snapdeal.com.

3.Myntra and Paytm makes late declaration of price during promotion.

4.Paytm takes longer time to load the page during promotion.

5.Snapdeal and Amazon have limited mode of payment on most of products during promotion.

6.Paytm and Snapdeal take a longer delivery period, whereas Myntra and Amazon takes lesser delivery period.

7.Amazon, Snapdeal and Myntra have frequent discrepencies, when moving from one page to another.

Now, we will analyse the positive feedbacks by checking the count of websites and the type of feedbacks given to each website. Then, we will save the obtained data in a new data frame and rename the column names.

In [39]:
```
#Analysing the positive feedbacks separately
website_list=['Amazon.in','Flipkart.com','Paytm.com','Myntra.com','Snapdeal.com']  #Website list
col_names=[]   #Empty list for column names
websites=[]    #Empty list for websites
count=[]       #Empty list for checking the count of customer present to websites are mentioned

for col in posFeed.columns:
    for i in website_list:
        present=len(posFeed[posFeed[col].str.contains(i)])  #Checking if websites in dataframe are available in the websit
        col_names.append(col)   #Appending the column names
        websites.append(i)    #Appending the website names
        count.append(present) #Appending the count of customer at website present in feedback

#Creating positive feedback dataframe
positive_df=pd.DataFrame([col_names,websites,count],index=["columnsName","websites","customerCount"]).T
positive_df
```

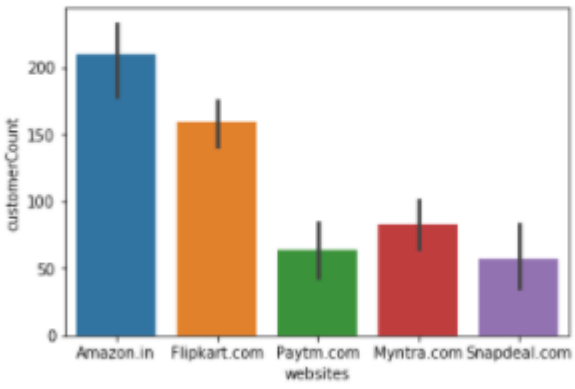
Out[39]:

	columnsName	websites	customerCount
0	From the following, tick any (or all) of the o...	Amazon.in	269
1	From the following, tick any (or all) of the o...	Flipkart.com	221
2	From the following, tick any (or all) of the o...	Paytm.com	150
3	From the following, tick any (or all) of the o...	Myntra.com	146
4	From the following, tick any (or all) of the o...	Snapdeal.com	182

Plotted graph showing the no. of customer using particular websites for negative feedbacks.

In [40]:
```
#Plotting barplot for people_count vs website relationship
sns.barplot(x='websites',y='customerCount',data=positive_df)
```

Out[40]: <matplotlib.axes._subplots.AxesSubplot at 0x1e84e82ba20>



We can observe that Amazon has received overall good feedback from the customers, followed by Flipkart, and snapdeal has received less number of feedbacks.

Percentage of customer who gives the feedbacks for positive dataframe.

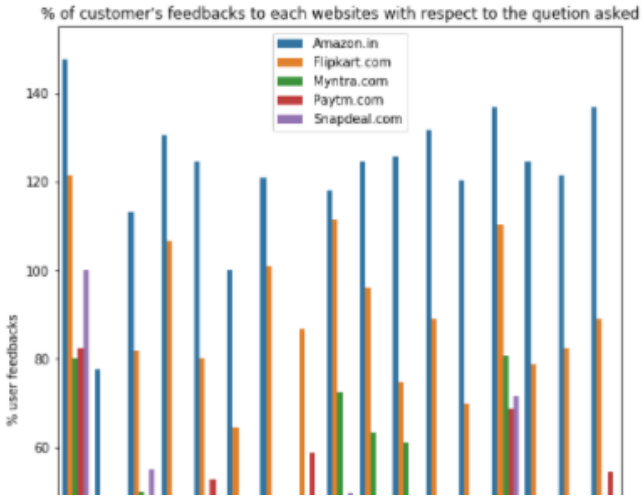
In [49]:
```
#Calculating the percentage for all data
positive_df['% user feedbacks']=(positive_df['customerCount']*100)/269 # Amazon users
positive_df['% user feedbacks']=(positive_df['customerCount']*100)/221 # Flipkart users
positive_df['% user feedbacks']=(positive_df['customerCount']*100)/147 # Myntra users
positive_df['% user feedbacks']=(positive_df['customerCount']*100)/150 # Paytm users
positive_df['% user feedbacks']=(positive_df['customerCount']*100)/182 # Snapdeal users
```

In [50]:
```
positive_df.reset_index(drop=True)
```

Out[50]:

	columnsName	websites	customerCount	% user feedbacks
0	From the following, tick any (or all) of the o...	Amazon.in	269	147.802198
1	Change in website/Application design	Amazon.in	141	77.472527
2	Security of customer financial information	Amazon.in	206	113.186813
3	Complete, relevant description information of ...	Amazon.in	238	130.769231
4	Reliability of the website or application	Amazon.in	227	124.725275
...	...	...	...	...

In [96]:
```
fig=plt.figure(figsize=(8,10))
sns.barplot(positive_df['columnsName'],positive_df['% user feedbacks'],hue=positive_df['websites'])
plt.title("% of customer's feedbacks to each websites with respect to the quetion asked")
plt.xticks(rotation=90)
plt.legend(loc=0)
plt.show()
fig.savefig("% of customer's feedbacks to each websites with respect to the quetion asked.png")
```



Plotted graph and we observe that ::

- Mostly customers are using Amazon, because:
    - this websites has change in website/application,
    - Security of information for customer financial information,
    - Complete, relevant description information of the products,
    - amazon websites is as efficient before,
    - has wild variety of product offers,
    - has Visual appealing web-page layout,
    - as presence of online assistance through multi-channel,
    - has speedy order delivery,
    - customers recommend amazon to their friends because they are easy to use,
    - and are perceived trustworthiness,
    - keeps privacy of customer information and is fast loading website/application.

- flipkart has highest quickness to complete purchase.
- amazon and flipkart has highest availiability of several payments options.

Observations from positive feedback:

1.Flipkart ranks about 90% in satisfying customers, followed by myntra.

2.The maximum percentage paytm and snapdeal could score here is 83 and 71 respectively.

3.No one is willing to refer snapdeal to their contacts as it has the less percentage among all websites.

4.On an average, snapdeal and paytm scores are less when compared to amazon, flipkart and myntra.

Percentage of customer who gives the feedbacks for positive dataframe.

In [67]:
```
#Calculating the percentage for all data
negative_df['% user feedbacks']=(negative_df['customerCount']*100)/135 # Amazon users
negative_df['% user feedbacks']=(negative_df['customerCount']*100)/103 # Flipkart users
negative_df['% user feedbacks']=(negative_df['customerCount']*100)/75 # Myntra users
negative_df['% user feedbacks']=(negative_df['customerCount']*100)/98 # Paytm users
negative_df['% user feedbacks']=(negative_df['customerCount']*100)/109 # Snapdeal users
```

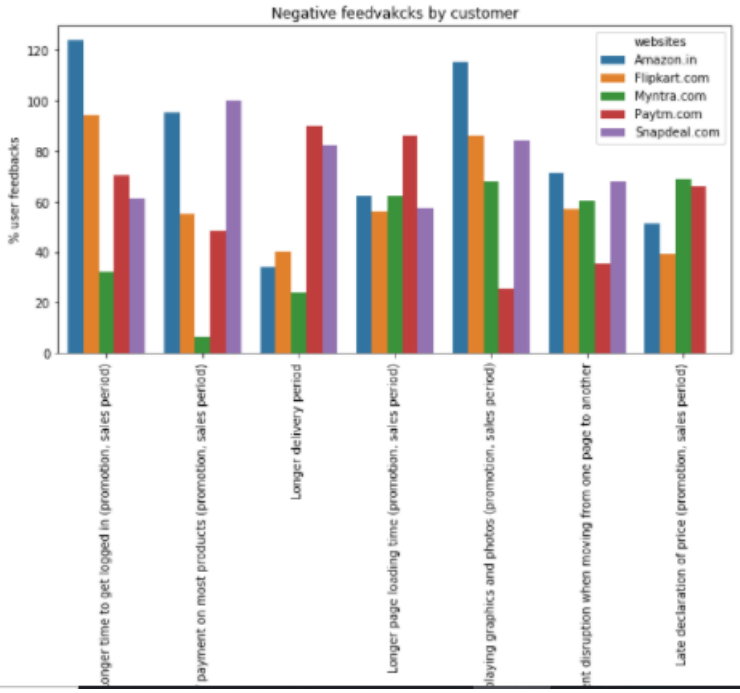
In [68]:
```
negative_df.sort_values(by=['% user feedbacks','websites'],ascending=False)
```

Out[68]:

	columnsName	websites	customerCount	% user feedbacks	Percentage
0	Longer time to get logged in (promotion, sales...	Amazon.in	135	123.853211	43.13099
5	Longer time in displaying graphics and photos ...	Amazon.in	126	115.59633	40.255591
24	Limited mode of payment on most products (prom...	Snapdeal.com	109	100.0	34.824281
20	Limited mode of payment on most products (prom...	Amazon.in	104	95.412844	33.226837
1	Longer time to get logged in (promotion, sales...	Flipkart.com	103	94.495413	32.907348
27	Longer delivery period	Paytm.com	98	89.908257	31.309904
17	Longer page loading time (promotion, sales per...	Paytm.com	94	86.238532	30.031949

In [75]:

```
fig=plt.figure(figsize=(10,5))
sns.barplot(negative_df['columnsName'],negative_df['% user feedbacks'],hue=negative_df['websites'])
plt.xticks(rotation=90)
plt.title("Negative feedvakcks by customer")
plt.show()
fig.savefig("Negative feedbacks by customer.png")
```



Plotted graph and we observe that:

Observations:

Mostly amazon customer got longer time to logged in and amazon takes longer time in displaying graphics and photos.

snapdeal has most limited mode for products transaction.

Paytm Customer got long term delivery period and long page loading.

amazon and snapdeal has taken frequent disruption when moves from one page to another.

Myntra and paytm has late declaration of product Price.

Conclusion Individual recommendations and feedbacks to the websites

Amazon.com To be improved: During promotions, try to give a disturbance free shopping experience to customers. Give more payment options to customers. Try to give price early during promotion. Reduce the delivery time of the products. Positive feedback summary: Convenient to use and also a good website for shopping. Fast delivery of products. Availability of complete information of the products. Presence of online assistance through multi-channels. Reliable website or app, perceived trustworthiness.

Flipkart.com To be improved: During promotions, try to give a disturbance free shopping experience to customers. Give more payment options to customers. Try to give the price early during promotion. Reduce the delivery time of the products. Flipkart and Amazon almost share the same feedbacks with varying percentages as the only difference. Positive feedback summary: Convenient to use and also a good website for shopping. Fast delivery of products. Availability of complete information of the products. Presence of online assistance through multi-channels. Reliable website or app, perceived trustworthiness. Wild variety of products to offer.

Myntra.com To be improved: During promotions, try to give a disturbance free shopping experience to customers. Try to give the price early during promotions. Reduce the delivery time of the products during promotions. Positive feedback summary: Convenient to use and also a good website. Availability of several payment options. Faster products delivery. Complete information of products available. Reliable website or app, perceived trustworthiness. Wild variety of product to offer

Paytm.com To be improved: Reduce the delivery time of the products during promotions. Try to give the price early during promotion. During promotions, try to give a disturbance free shopping experience to customers. Late declaration of price and discounts. Frequent disturbance is occurring while moving

from one page to another. Positive feedback summary Convenient to use and a good website. Quickness to complete a purchase. About 64% of the customers feel that either web or app is reliable. Around 20% of the customers believe that paytm has a wild variety of products on offer.

Snapdeal.com To be improved: Reduce the delivery time of the products during promotions. Try to give the price early during promotion. During promotions, try to give a disturbance free shopping experience to customers. Late declaration of price and discounts. No one has expressed to recommend snapdeal to a contact as it has the most negative feedbacks among all other websites. Positive feedback summary: Convenient to use. 54% of the customers are happy about the availability of financial information security.

General suggestions and recommendations to all the e-commerce websites

Improve the experience of shopping for customers, as there is a lot of scope in enhancing the shopping experience to the customers using AI. Continue giving more financial benefits like coupons, cashbacks, etc. as customers are very much attracted to it. Trustworthiness and approachability through various channels is still highly rated by customers. Majority of the customers are working class women and their age is between 20-40. Always bring variety of products targeting them. Provide more customer friendly approach like fast delivery, complaint resolution, etc.

Therefore, we had analysed the given dataset by using various data analysis process and also, we had concluded the analysis by observing the positive and negative feedbacks obtained. We recommended some suggestions for the websites to improve further in the future.