

E-retail factors for customer activation and retention: A case study from Indian e-commerce customers.

PREPARED BY: - ABHINAV KUMAR

Customers have an abundance of options when it comes to products for purchase. This excess of options, however, increases the risk of poor customer retention. Since acquiring new customers costs much more than keeping current customers, a higher retention rate is always better.

Customer retention represents the number of customers who continue purchasing from a company after their first purchase. This is usually measured as the **customer retention rate**, which is the percentage of customers your company has retained over a certain time period. The opposite of retention rate is **churn rate**, which represents the percentage of customers a company has lost over a given time period.

Customer retention analytics can be done through machine learning, allowing companies to base their product and marketing strategies on predictive customer analytics rather than less reliable predictions made manually.

But here we have focused majorly on the data analysis that how customers from different regions are affected with different factors.

Here we have analyzed more than 50 columns and the polls for different questions were answered by people.

Customer satisfaction has emerged as one of the most important factors that guarantee the success of online store; it has been posited as a key stimulant of purchase, repurchase intentions and customer loyalty.

A comprehensive review of the literature, theories and models have been carried out to propose the models for customer activation and customer retention.

Five major factors that contributed to the success of an e-commerce store have been identified as: service quality, system quality, information quality, trust and net

benefit. The research furthermore investigated the factors that influence the online customers repeat purchase intention.

The combination of both utilitarian value and hedonistic values are needed to affect the repeat purchase intention (loyalty) positively. The data is collected from the Indian online shoppers. Results indicate the e-retail success factors, which are very much critical for customer satisfaction.

What Is Customer Retention Analysis?

Customer retention analysis is the application of statistics in order to understand how long customers are retained before churning out and to identify trends in customer retention. This type of analysis discerns how long customers usually stick around, whether or not seasonality affects customer retention, and discovers behaviors and factors that differentiate retained customers from churned customers.

Why Is Customer Retention Analysis Important For Your Company?

Customer retention analysis is important for your company (here e-commerce websites) because it helps you understand which personas have higher retention rates and discern which features impact retention. This provides actionable insights that can help you make more effective product and marketing decisions.

It can be difficult for a product or sales team to know how well a product is actually performing with the target audience. They may think that features and messaging is on brand and clear because acquisition numbers are growing. However, just because

new customers are purchasing a product does not necessarily mean customers like the product or service enough to stick around.

That is where customer retention analysis comes in. Every company needs data in order to make effective business and marketing decisions. Machine learning makes this easier than it has ever been before, which is great news for companies that wish to leverage this data.

In the past, e-commerce companies spent lots of resources when trying to make demand and supply predictions. After the invention of artificial intelligence, these companies are now relying on machines to make the projections. The advantage of using machine learning algorithms in this context is that they can offer unique insights on certain data sets.

Using the demand and supply predictions, online retailers can also rely on ML algorithms to set analytical goals. These retailers have the opportunity to make data-backed decisions that will help them improve their services and products. They can only achieve these goals by using the accurate data analyses that machine learning provides.

How Do You Analyze Customer Retention?

Machine learning for customer retention analytics uses past customer data to predict future customer behavior. In today's data-driven world, companies can track hundreds of data points about thousands of customers. Therefore, the input data for the customer retention model could be any combination of the following:

- **How many times you have made an online purchase in the past 1 year?**
 - **User friendly Interface of the website**
 - **Convenient Payment methods**
 - **Trust that the online retail store will fulfill its part of the transaction at the stipulated time**
 - **Empathy (readiness to assist with queries) towards the customers**
 - **Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.)**
 - **Enjoyment is derived from shopping online**
 - **Shopping online is convenient and flexible**
 - **Net Benefit derived from shopping online can lead to users satisfaction**
 - **Enjoyment is derived from shopping online**
- Etc.....**

So, the input variables shown above are just for examples but here we have collected the results for more than 50 questions for which a survey is conducted to collect the polls in different parts of the country.

During the model training process, this data will be used to find correlations and patterns to create the final trained model to predict customer retention. Not only does this tell you the overall churn risk of your customer base, but it can determine churn risk down to the individual customer level. You could use this data to proactively market to those customers with higher churn risk or find ways to improve your product, customer service, messaging, etc. in order to lower your overall churn rate.

Factors affecting the churning out of customers: -

Optimize Prices

Pricing is crucial to the success of any online business. Online retailers should note that for them to drive sales, they shouldn't rely on local market prices or set markup rates. These days, retailers can easily check and compare the prices that their competitors set for certain products or services. Furthermore, online buyers are always on the lookout for better deals.

Provide Quality Customer Self-service and Support

It can be quite challenging to provide customers with quality self-service or support when running an e-commerce business. The good news for managers of e-commerce enterprises is that ML technology offers intelligent chatbots to handle this daunting task. The chatbots rely on natural language to initiate conversations aimed at identifying and solving issues that customers have. When e-commerce enterprises automate their customers self service and support initiatives, it is easier for them and their clients to experience satisfaction. These automated processes also help them to take their service delivery mechanisms a notch higher.

Detect and Prevent Fraud

Just like other businesses, e-commerce enterprises are prone to fraud. At times, online enterprises experience chargebacks when fraudsters gain unauthorized access to financial systems. They are also at risk of having a tarnished reputation

after being a victim of fraud. It is close to impossible detecting and preventing fraud.

Offer Personalized Product Suggestions

E-commerce giants such as Amazon can attest to the fact that personalized product suggestions work. To be precise, Amazon relies on a product suggestion tool known as the Recommendation Engine to drive 35 percent of its sales. It took the engineers of Recommendation Engine lots of effort to come up with a tool that can help in finding the right patterns in shopping behavior and product sales.

So above discussed are some of the key parameters that are helpful in retaining the customers.

Now we will go for the **data analysis** that is done for the given dataset: -

So, it all starts with importing the required libraries i.e pandas with the help of which I have imported my excel file, after this I have checked the shape for the dataset which contains around 269 rows and 70 columns.

Then I have went to check if my dataset contains any null values and here we are lucky enough we don't have any null values in the dataset.

The dataset is in encoded form i.e all the object data type is converted into int data type.

For the few columns that are not encoded for that with the help of label encoder technique I have converted the remaining in int dtype.

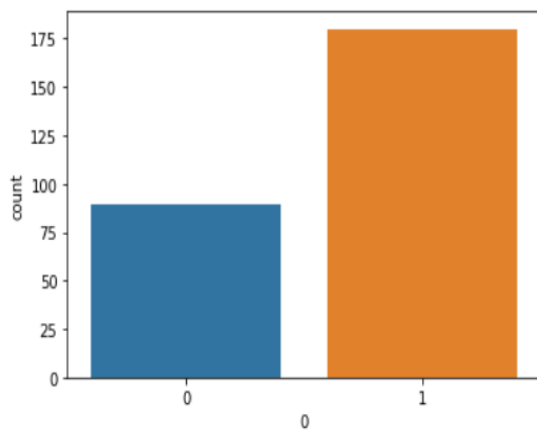
Since the dataset is large, we have used label encoding technique instead of one hot encoder in order to avoid the more consumption of memory.

Now comes the important part that is more useful in analyzing the dataset more efficiently is data visualization.

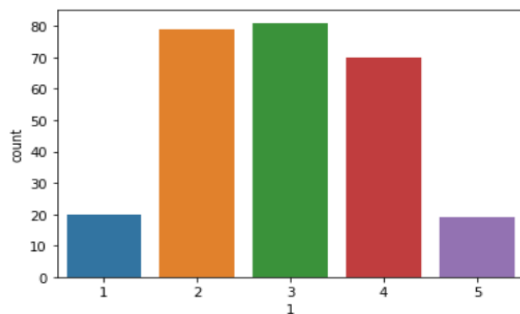
Data visualization helps in analyzing different columns that is useful in achieving our goal and building an efficient machine learning model.

For that we will import our required library i.e. matplotlib and seaborn with the help of which we have plotted different graphs. The instance of which is shown below.

```
: sns.countplot(df[0]) #male=0, female=1
: <matplotlib.axes._subplots.AxesSubplot at 0x2c350a415c8>
```



```
] : #How old are you?
sns.countplot(df[1])
]: <matplotlib.axes._subplots.AxesSubplot at 0x2c34f3936c8>
```

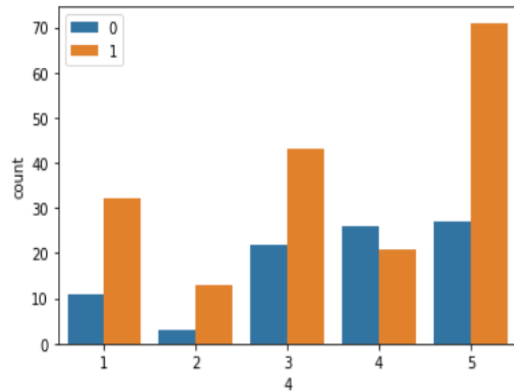


Conclusion:- most of the people are from age 21-50 that prefer online shopping from e commerce website


```
#Since How Long You are Shopping Online ?
```

```
sns.countplot(df[4],hue=df[0]) #less than 1 is 1, bet 1-2 is 2...
```

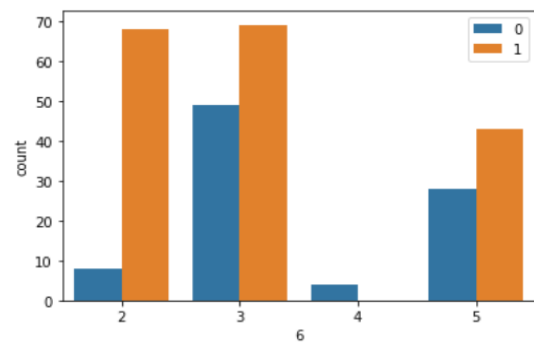
```
<matplotlib.axes._subplots.AxesSubplot at 0x2c3541887c8>
```



```
# How do you access the internet while shopping on-line?
```

```
sns.countplot(df[6],hue=df[0]) #wifi, mobile internet, dialup
```

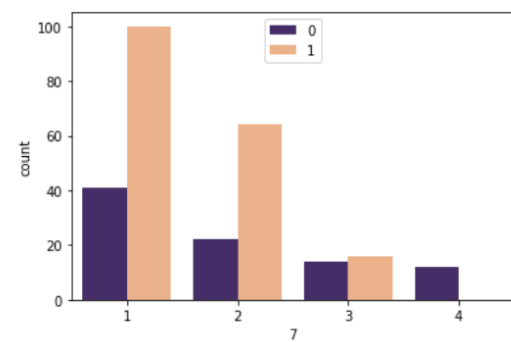
```
<matplotlib.axes._subplots.AxesSubplot at 0x2c3543c1548>
```



```
: sns.countplot(df[7],hue=df[0],palette=['#432371',"#FAAE7B"]) # Which device do you use to access the online shopping?
```

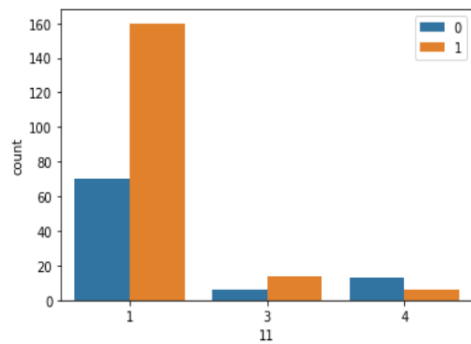
```
# 1---Smartphone, 4---Tablet, 3--Desktop, 2--Laptop
```

```
: <matplotlib.axes._subplots.AxesSubplot at 0x2c350a65948>
```



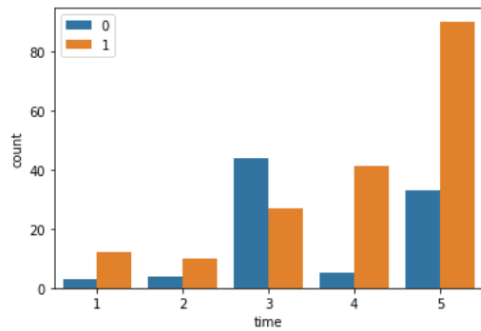
```
: # Which channel did you follow to arrive at your favorite online store for the first time?
sns.countplot(df[11],hue=df[0]) #3--Content Marketing, 1--Search Engine, 4--Display Adverts
```

```
: <matplotlib.axes._subplots.AxesSubplot at 0x2c34f168b08>
```



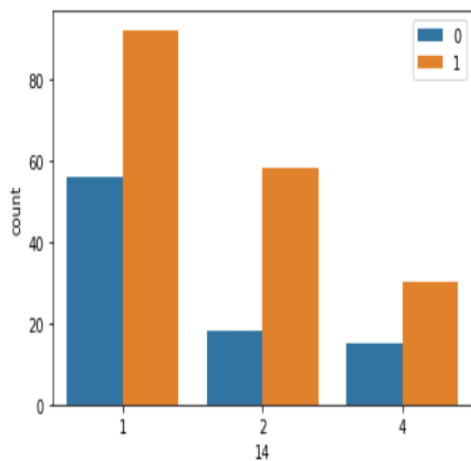
```
: sns.countplot(df[13],hue=df[0]) #How much time do you explore the e- retail store before making a purchase decision?
plt.xlabel('time')
plt.ylabel('count') #starting from less than 1 min---1, 1-5--2,6-10---3,11-15---4,>15---5
```

```
: Text(0, 0.5, 'count')
```



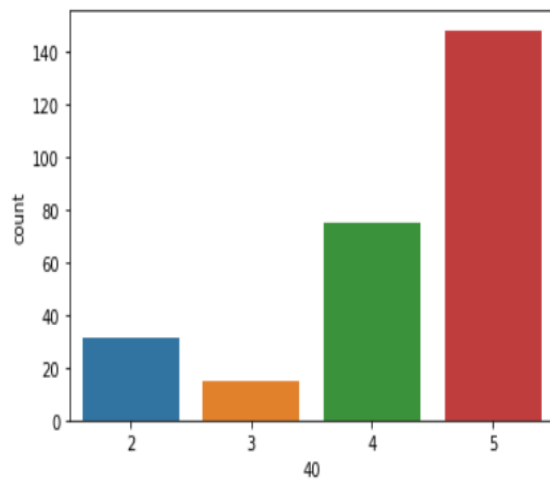
```
: sns.countplot(df[14],hue=df[0])#preferred payment Option---1---debit/credit, 2-e wallet,4---cod
```

```
: <matplotlib.axes._subplots.AxesSubplot at 0x2c351c025c8>
```



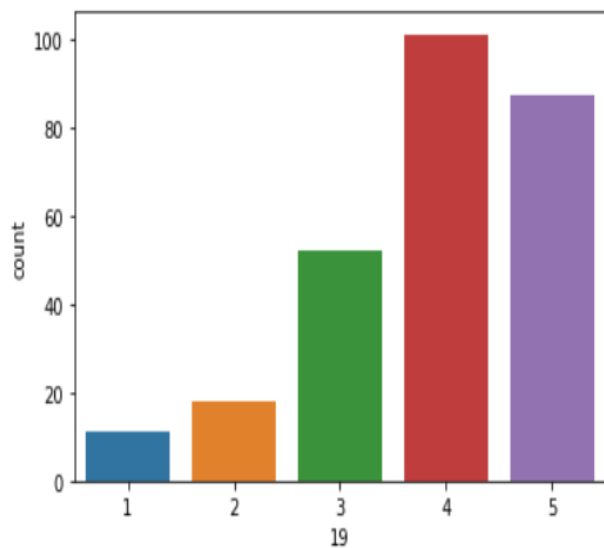
```
: #Monetary savings  
sns.countplot(df[40]) #Disagree (2) Strongly agree (5) indifferent (3) Agree (4)
```

```
: <matplotlib.axes._subplots.AxesSubplot at 0x2c351c43108>
```



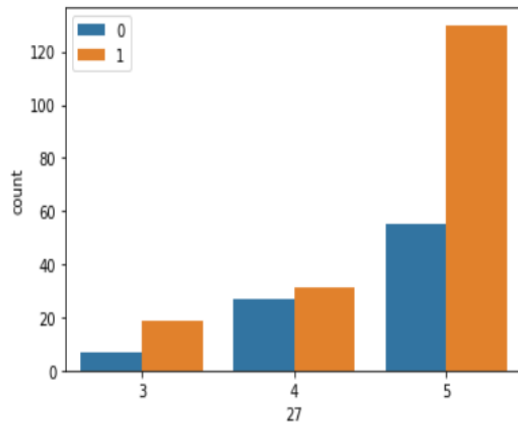
```
#Complete information on listed seller and product being offered is important for purchase decision.  
sns.countplot(df[19])  
#Indifferent (3) Strongly agree (5) Agree (4) Dis-agree (2)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x2c35424f6c8>
```



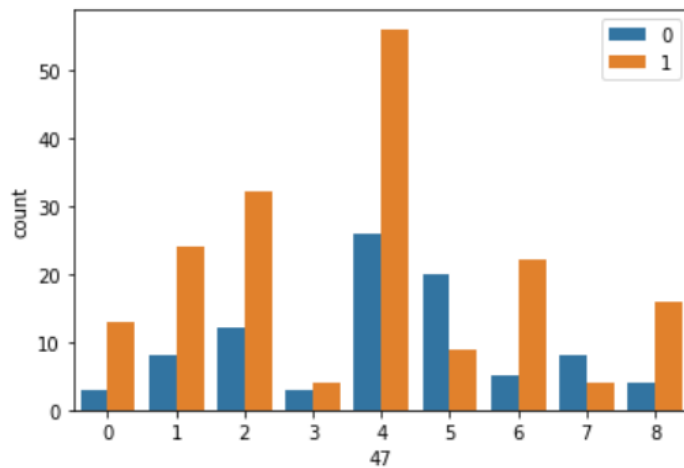
```
: #Being able to guarantee the privacy of the customer
sns.countplot(df[27],hue=df[0])
#Strongly agree (5) Agree (4) indifferent (3)

: <matplotlib.axes._subplots.AxesSubplot at 0x2c35433bec8>
```



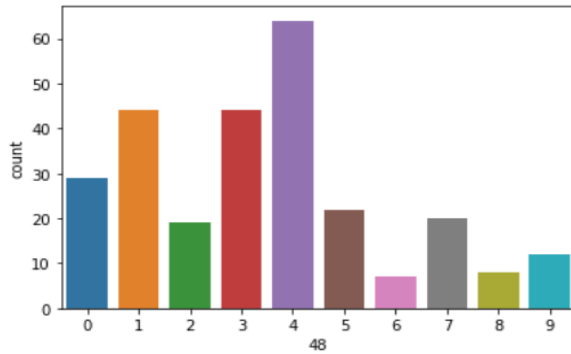
```
: #From the following, tick any (or all) of the online retailers you have shopped from
sns.countplot(df[47],hue=df[0])
#5---Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com
#0---Amazon.in
#1---Amazon.in, Flipkart.com
#3---Amazon.in, Flipkart.com, Paytm.com
#4---Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com
#6---Amazon.in, Flipkart.com, Snapdeal.com
#2---Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com
#7---Amazon.in, Paytm.com
#8---Amazon.in, Paytm.com, Myntra.com

: <matplotlib.axes._subplots.AxesSubplot at 0x2c3516bbd48>
```



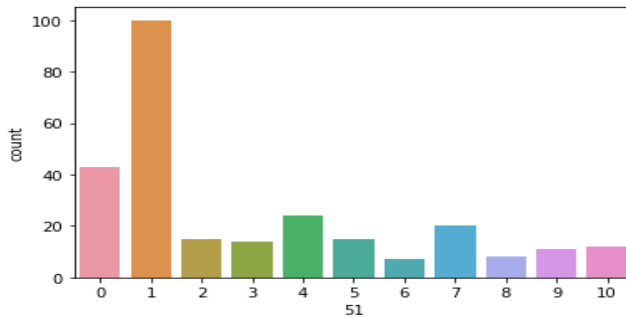
```
#Easy to use website or application
sns.countplot(df[48])
#9--Paytm.com
#3--Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com
#7--Amazon.in, Paytm.com, Myntra.com
#4--Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com
#1--Amazon.in, Flipkart.com
#5--Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com
#2--Amazon.in, Flipkart.com, Myntra.com
#6--Amazon.in, Paytm.com
#8---Flipkart.com
```

<matplotlib.axes._subplots.AxesSubplot at 0x2c356cd7148>



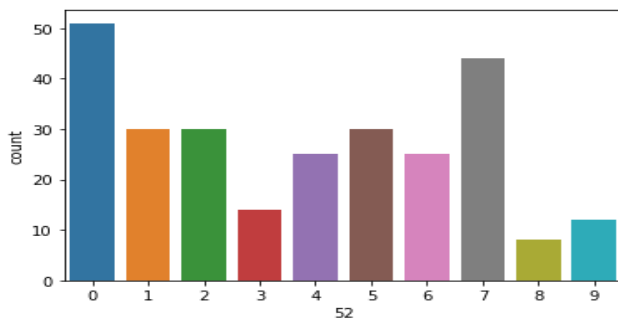
```
] : #Complete, relevant description information of products
sns.countplot(df[51]) #1---Amazon.in, Flipkart.com
```

<matplotlib.axes._subplots.AxesSubplot at 0x2c35467fb88>



```
] : #Fast Loading website speed of website and application
sns.countplot(df[52]) #0--Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com
#7--- Amazon.in, Paytm.com
```

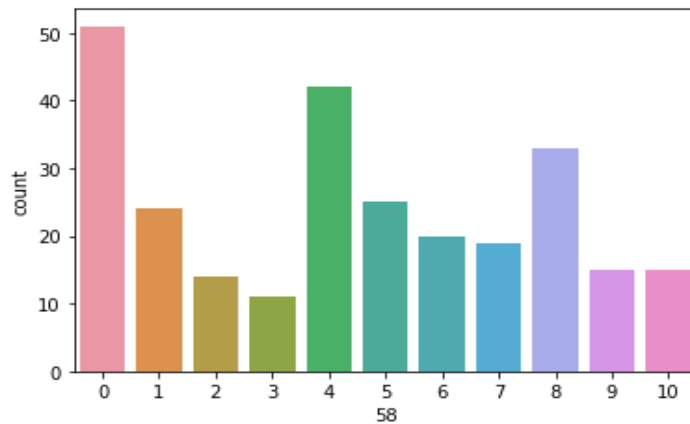
<matplotlib.axes._subplots.AxesSubplot at 0x2c356d51608>



```
In [70]: #Security of customer financial information
#0---Amazon.in
#4---Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com

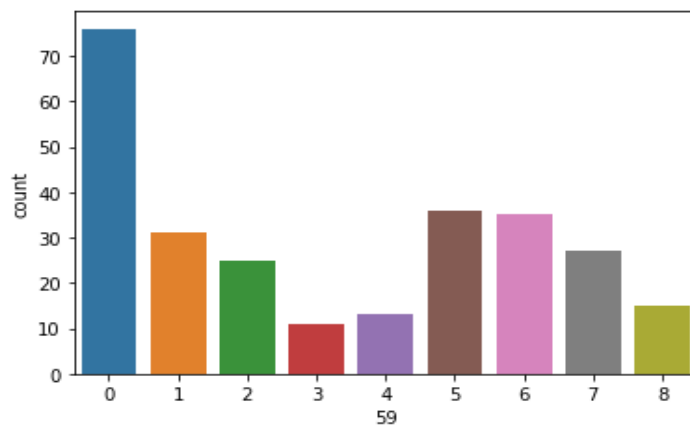
sns.countplot(df[58])
```

Out[70]: <matplotlib.axes._subplots.AxesSubplot at 0x2c356ee5048>



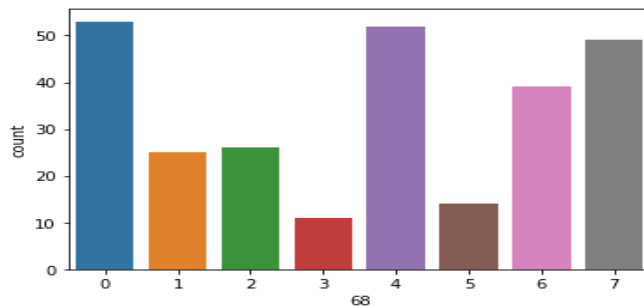
```
In [71]: #Perceived Trustworthiness
#0--Amazon.in
sns.countplot(df[59])
```

Out[71]: <matplotlib.axes._subplots.AxesSubplot at 0x2c356f6c788>



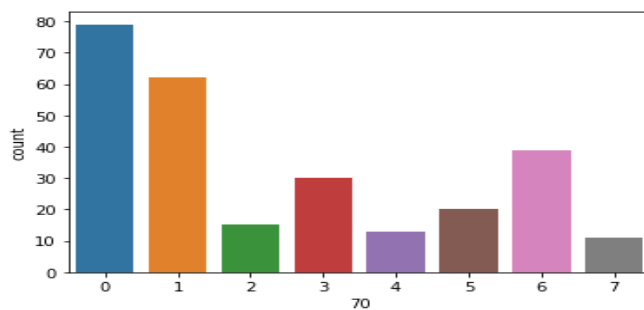
```
In [80]: #Website is as efficient as before
#0--Amazon.in 4---Myntra.com 6---Paytm.com 7--Snapdeal.com
sns.countplot(df[68])
```

Out[80]: <matplotlib.axes._subplots.AxesSubplot at 0x2c3573d7848>



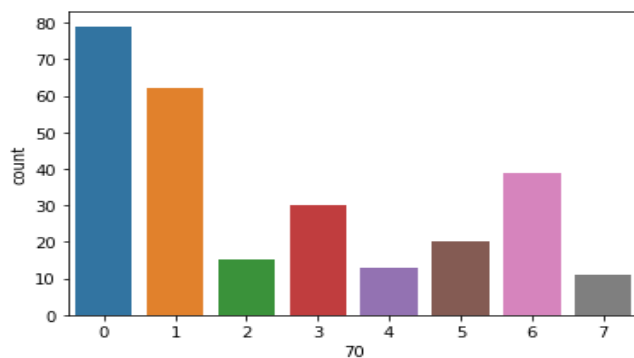
```
In [82]: #Which of the Indian online retailer would you recommend to a friend?
#Amazon.in
sns.countplot(df[70])
```

Out[82]: <matplotlib.axes._subplots.AxesSubplot at 0x2c3574d3b08>



```
In [82]: #Which of the Indian online retailer would you recommend to a friend?
#Amazon.in
sns.countplot(df[70])
```

Out[82]: <matplotlib.axes._subplots.AxesSubplot at 0x2c3574d3b08>



So, these are some of the visuals plotted from the dataset from which we can conclude that the e-commerce website 'AMAZON.IN' is the best website in comparison with other e-commerce website and since each coin has two faces similarly this website also lags in few aspects from others, but over and above most people prefer this website for online shopping.