E-retail factors for customer activation and retention

Submitted by:

Aaron D’souza

ACKNOWLEDGMENT

Link: https://www.youtube.com/watch?v=-o3AxdVcUtQ

Link: https://www.youtube.com/watch?v=dP170R-wO-8

Link: https://www.youtube.com/user/krishnaik06

Link: https://www.researchgate.net/publication/346412647_E-retail_factors_for_customer_activation_and_retention_An_empirical_study_from_Indian_e-commerce_customers

INTRODUCTION

- ## Business Problem Framing

  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.

  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.

  Large number of customers are getting attracted towards online retailing; this is because e-stores usually offer them a variety of services and products according to their preferences.

  The research furthermore investigated the factors that influence the online customers repeat purchase intention on the basis of the Means End Chain theory (MEC) and Prospect theory.

- ## Conceptual Background of the Domain Problem

Convenience, round the clock availability, flexible pricing, discounts as well as free door step delivery are some of the major benefits of shopping online.

Presently, more number of online retailers are beginning to experience increase in demand for products and services.

Indian online retail industry has been experiencing good times since the last six years; as a result of the constantly growing internet penetration, deployment of modern infrastructures, and a robust ecosystem for e-retail start-ups.

Several e-commerce start-ups have commenced operation with innovative strategies, which differs from what was pioneered by first generation e-commerce companies.

- ## Review of Literature

There were more than 1200 start-ups came up in 2018, including eight unicorns, taking the total number to 7200 in India.

India's B2C e-commerce revenue grew from €20 billion in 2017 to reach €25 billion in 2018 at a growth rate of 20% .

The number of internet users in India has been growing at a CAGR of 35% since the year 2007 according to a report by IAMAI-IMRB (2017).

Having grown from just over a 100 million internet users in 2010, India has since touched 500 million internet users in 2018.

Online retail businesses are positioning themselves to be able to take advantage of the massive internet user base by turning them into online shoppers.

- ## Motivation for the Problem Undertaken

Indian e-commerce market is poised to surpass that of the United States, making it the second largest in the world in less than twenty years from now (PwC India, 2018).

According to global payments firm Worldpay (The economist, 2016), this rapid growth is driven by a host of e-commerce players comprising of small, medium and large firms in terms of revenue and assets.

An amalgamation of factors contributed to India's rapid increase in number of online retail shoppers, which includes; digital friendly policies of the Govt.

Large investments made by the online retailers, growing internet and smartphone penetration, demographics distributions, emerging middle class and young population (IBEF, 2019; Agarwal and Dixit, 2020).

With western markets getting saturated (matured) and China becoming more restrictive, India is becoming the main battleground for the e-tailers.

Analytical Problem Framing

- ## Mathematical/ Analytical Modelling of the Problem

As an aspiring Data Scientist the goal is to apply my analytical skills to give findings and conclusions in detailed data analysis written in jupyter notebook.

This analysis will then be used by the E-commerce retailers to understand how exactly the variables are related to each other and what factors should be improved in order to attain high customer loyalty.

- ## Data Sources and their formats

The sample data is provided to us from our client database.

A total of two data types in the data set Int and Object.

The data provided was well organized structured data in .CSV (comma-separated values) format.

The data set has a total of 70 columns:

- **0 Gender of respondent : Gender**
- **1 How old are you? : Age**
- **2 Which city do you shop online from? : City**
- **3 What is the Pin Code of where you shop online from? : Pin**

- **4 Since How Long You are Shopping Online? : SHLSO**

- **5 How many times you have made an online purchase in the past 1 year? : NOPPY**

- **6 How do you access the internet while shopping on-line? : IA**

- **7 Which device do you use to access the online shopping? : Device_Usage**

- **8 What is the screen size of your mobile device? : Screen_size**

- **9 What is the operating system (OS) of your device? : OS**

- **10 What browser do you run on your device to access the website? : Browser**

- **11 Which channel did you follow to arrive at your favorite online store for the first time? : Channel**

- **12 After first visit, how do you reach the online retail store? : MRORS**

- **13 How much time do you explore the e- retail store before making a purchase decision? : Duration(minutes)**

- **14 What is your preferred payment Option? : Payment_pref**

- **15 How frequently do you abandon (selecting an items and leaving without making payment) your shopping cart? : FOIA**

- **16 Why did you abandon the "Bag", "Shopping Cart"? : RFA**

- **17 The content on the website must be easy to read and understand? : WCR**

- **18 Information on similar product to the one highlighted is important for product comparison? : IPCH**

- **19 Complete information on listed seller and product being offered is important for purchase decision? : IASAP**

- **20 All relevant information on listed products must be stated clearly? : RIAP**

- **21 Ease of navigation in website : EWN**

- **22 Loading and processing speed? : Speed(L&A)**

- **23 User friendly Interface of the website : UFI**

- **24 Convenient Payment methods : CPM**

- 25 Trust that the online retail store will fulfill its part of the transaction at the stipulated time : TOSCT

- 26 Empathy (readiness to assist with queries) towards the customers : Cust_Support_Empathy

- 27 Being able to guarantee the privacy of the customer : Privacy_Guarantee

- 28 Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.)? : RCC

- 29 Online shopping gives monetary benefit and discounts : Discounts

- 30 Enjoyment is derived from shopping online : EOS

- 31 Shopping online is convenient and flexible : C&F

- 32 Return and replacement policy of the e-tailer is important for purchase decision : R&R

- 33 Gaining access to loyalty programs is a benefit of shopping online : GALP

- 34 Displaying quality Information on the website improves satisfaction of customers : DQI

- **35 User derive satisfaction while shopping on a good quality website or application : SOA**

- **36 Net Benefit derived from shopping online can lead to users satisfaction : OSBUS**

- **37 User satisfaction cannot exist without trust : NSNT**

- **38 Offering a wide variety of listed product in several category : Wide_Variety**

- **39 Provision of complete and relevant product information : Complete_prod_info**

- **40 Monetary savings : MS**

- **41 The Convenience of patronizing the online retailer : CPR**

- **42 Shopping on the website gives you the sense of adventure : Adventure**

- **43 Shopping on your preferred e-tailer enhances your social status : Social_status**

- **44 You feel gratification shopping on your favorite e-tailer : Gratification**

- 45 Shopping on the website helps you fulfill certain roles : SFR
- 46 Getting value for money spent : VFM
- 47 From the following, tick any (or all) of the online retailers you have shopped from : C_E_R
- 48 Easy to use website or application : ETU
- 49 Visual appealing web-page layout : Best_Homepage
- 50 Wild variety of product on offer : VOP
- 51 Complete, relevant description information of products : RDOI
- 52 Fast loading website speed of website and application : Loading_speed
- 53 Reliability of the website or application : ROA
- 54 Quickness to complete purchase : Purchase_speed
- 55 Availability of several payment options : Payment_options
- 56 Speedy order delivery : Delivery_speed

- 57 Privacy of customers' information : Cust_privacy
- 58 Security of customer financial information : COFI
- 59 Perceived Trustworthiness : Trustworthiness
- 60 Presence of online assistance through multi-channel : Online_Ass
- 61 Longer time to get logged in (promotion, sales period) : Login(TD)
- 62 Longer time in displaying graphics and photos (promotion, sales period) : LTDG&P
- 63 Late declaration of price (promotion, sales period) : Late_declaration
- 64 Longer page loading time (promotion, sales period) : Page_load_time
- 65 Limited mode of payment on most products (promotion, sales period) : LTM_paymet_modes
- 66 Longer delivery period : Delayed_delivery
- 67 Change in website/Application design : CID
- 68 Frequent disruption when moving from one page to another : FDP

- **69 Website is as efficient as before: Website_eff**

- **70 Which of the Indian online retailer would you recommend to a friend? : Friend_Suggest**

```
1 df.head()
```

	Gender	Age	City	Pin	SHLSO	NOPPY	IA	Device_Usage	Screen_size	
0	Male	31-40 years	Delhi	110009	Above 4 years	31-40 times	Dial-up	Desktop	Others	Windo
1	Female	21-30 years	Delhi	110030	Above 4 years	41 times and above	Wi-Fi	Smartphone	4.7 inches	
2	Female	21-30 years	Greater Noida	201308	3-4 years	41 times and above	Mobile Internet	Smartphone	5.5 inches	

Fig.1 (Data frame Sample)

- # Hardware and Software Requirements and Tools Used

Hardware used:

- **OS: Windows 10 Home Single Language 64 bit**
- **Ram: 8 GB**
- **Processor: Intel I5**

Software used:

- **Jupyter Notebook**

Model/s Development and Evaluation

- Interpretation of the Results

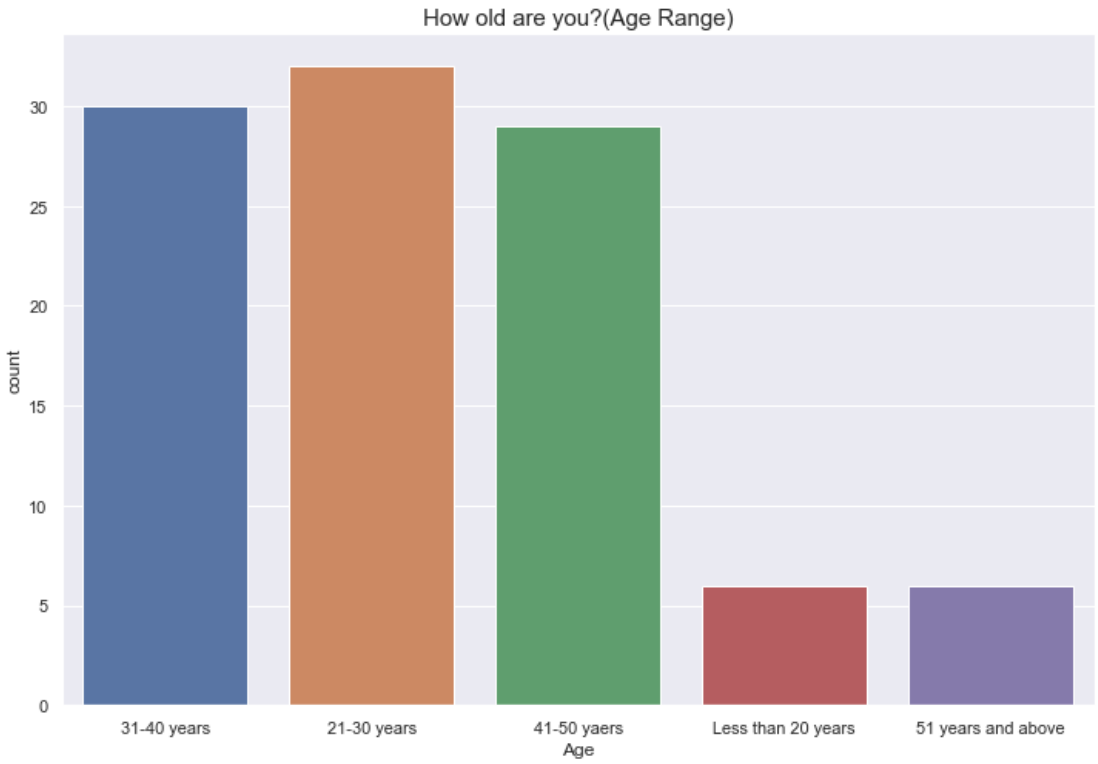

Fig.2 (Age)

The visualization in Fig.2 is a count plot from sea-born library it shows the distribution of Customer Age. Most of the customers are within the Age range (21 – 50) years. Less customers from age group below 20 years and above 51 years.

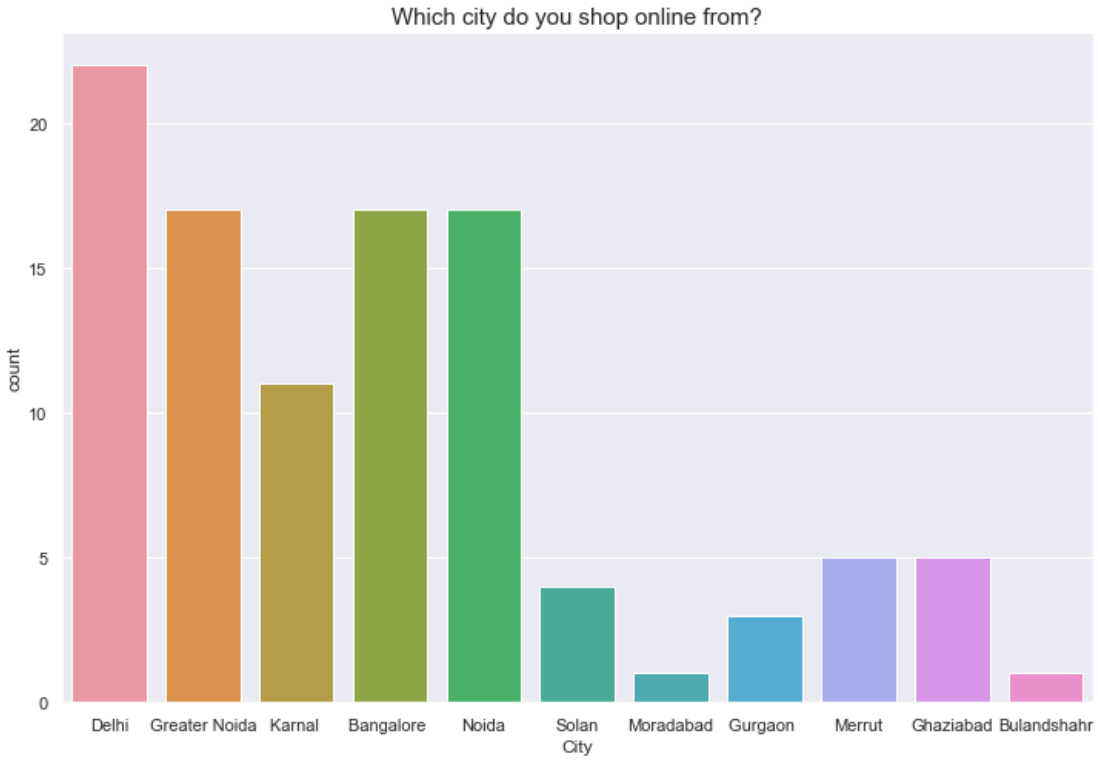

Fig.3 (City)

The visualization in Fig.3 is a count plot from sea-born library it shows the distribution of Cities the customers originate from. Delhi has the highest number of customers who are involved in online shopping (Important customers for the e-commerce company) followed by Greater Noida, Bangalore, and Noida.Moradabad, Bulandshahr have very few customers (the Ecom companies should investigate and find out the reasons for customers being less in those areas) and try to enhance sales in those areas.

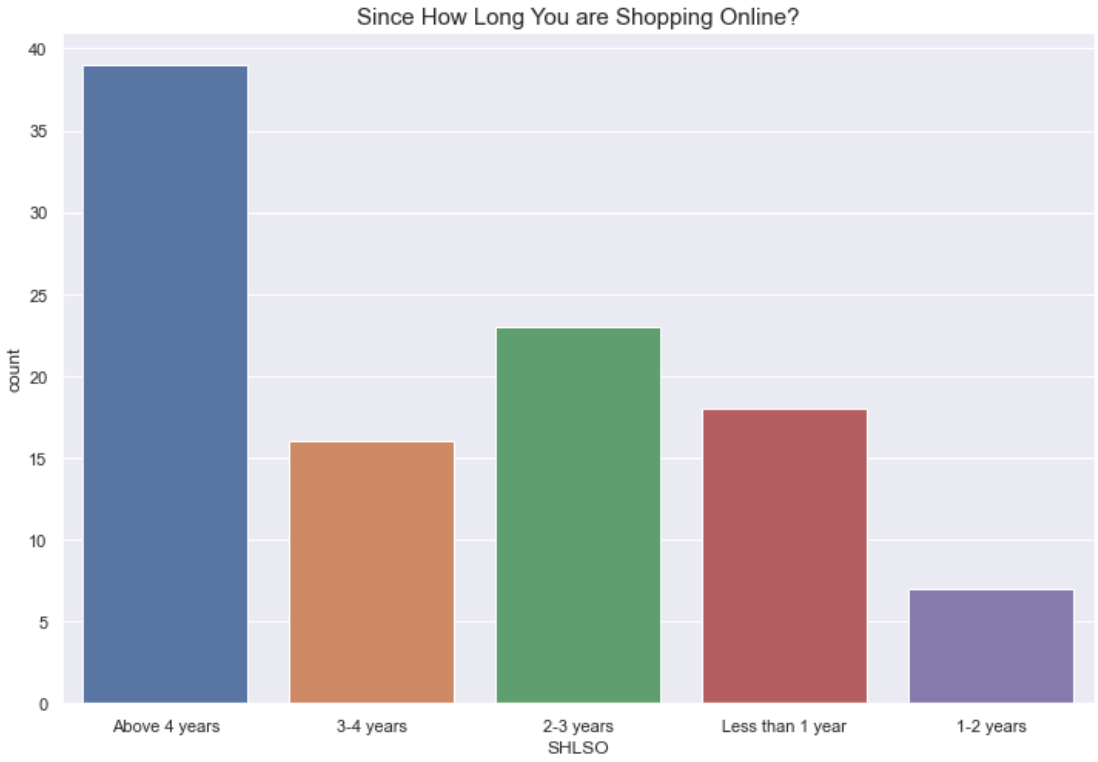

Fig.4 (SHLSO)

The visualization in Fig.4 is a count plot from sea-born library it shows the distribution for how long a customer is been shopping for. Most of the regular customers have been shopping for more than 4 years(Important customers, loyal customers). Need to provide best services to customers who have recently been shopping for less than 1 year (better the service more the loyalty).

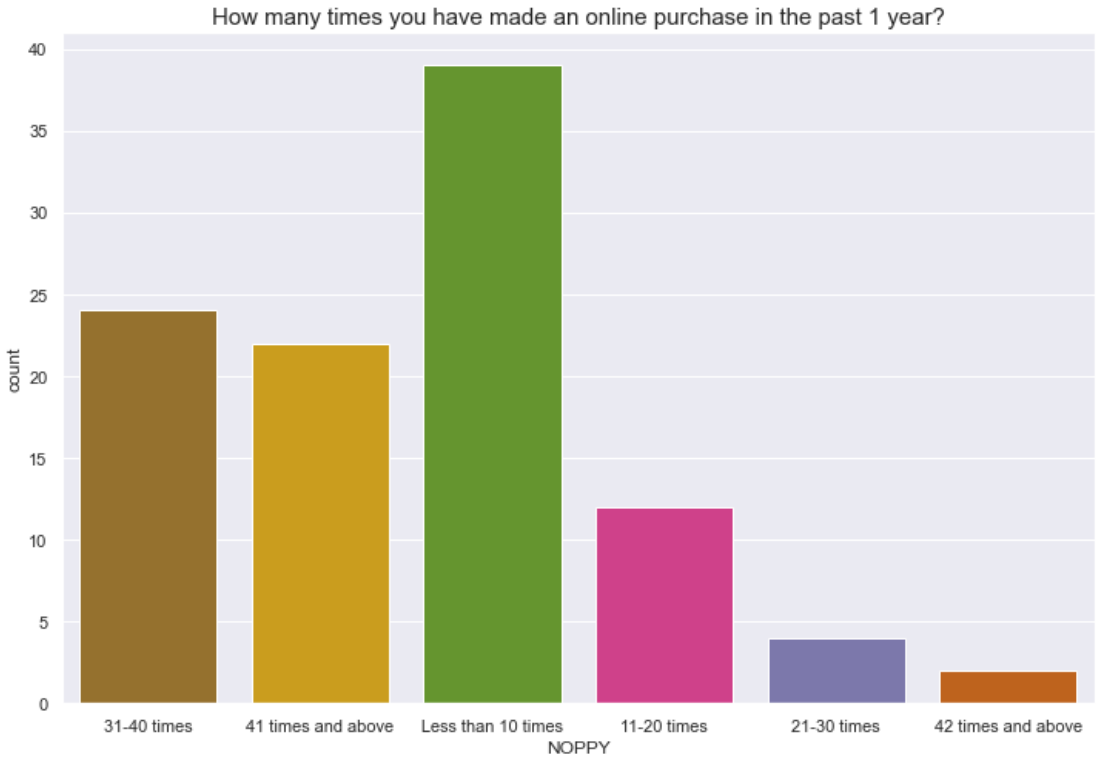

Fig.5 (NOPPY)

The visualization in Fig.5 is a count plot from sea-born library it shows the distribution for how many times a customer has made an online purchase in the past 1 year. Less than 10 times is the highest count for the customers who have purchased something online in the past 1 year(maybe average customers).41 times and 42 times and above are important customers to the e-commerce company (Customers who have a high engagement rate in online purchases). Mostly an average customer is buying ~ 10 items yearly.

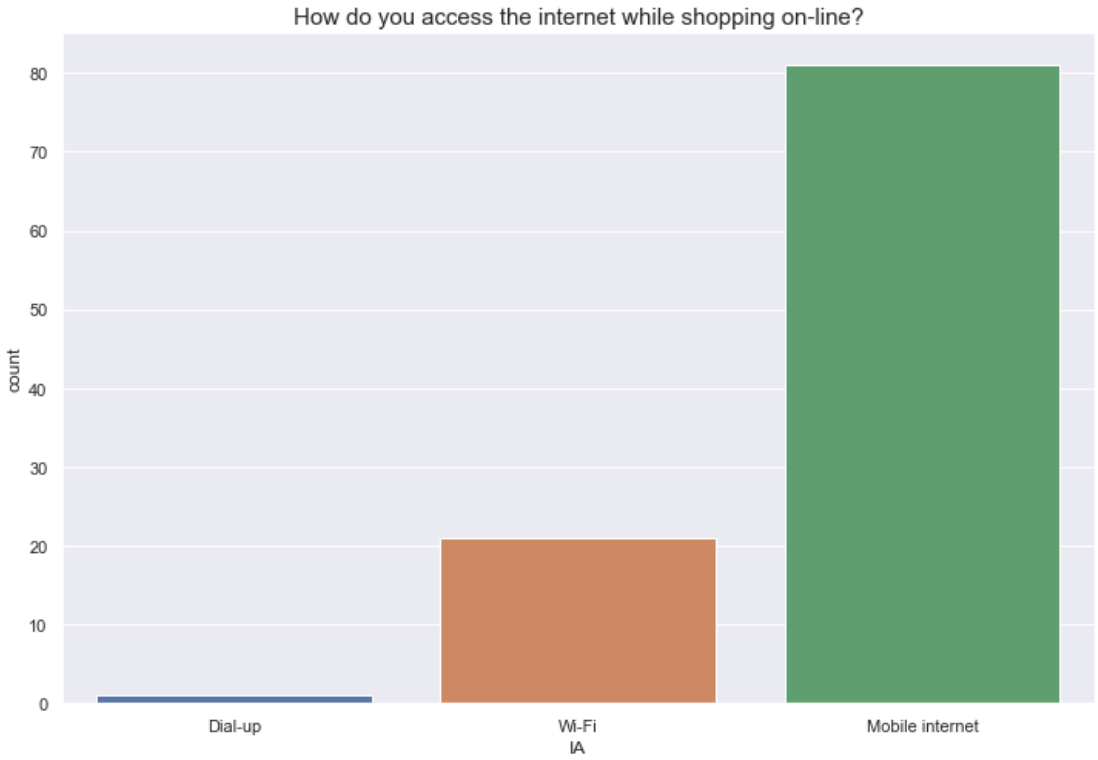

Fig.6 (IA)

The visualization in Fig.6 is a count plot from sea-born library it shows the distribution for how a customer access internet for online shopping. Most of the customers are using mobile data(Internet) while shopping online.

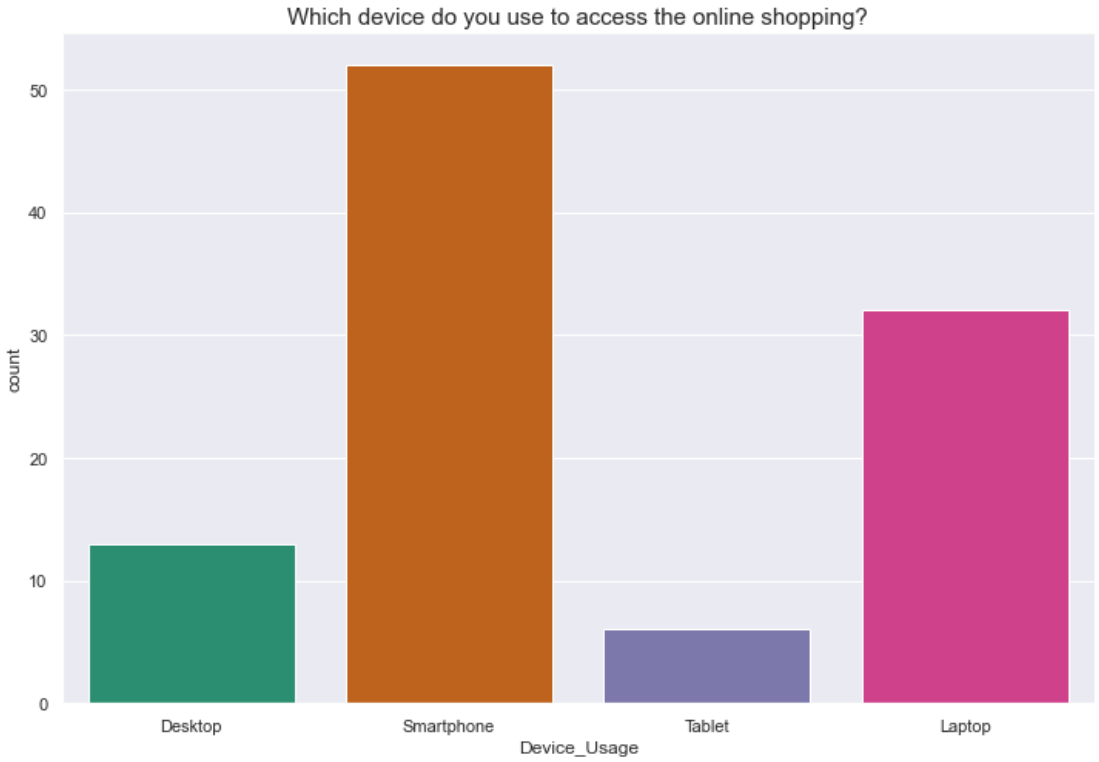

Fig.7 (Device_Usage)

The visualization in Fig.7 is a count plot from sea-born library it shows the distribution of device usage by customers for online shopping. Laptops and smartphones are the most used devices for online shopping. Tablet's are the least used devices.

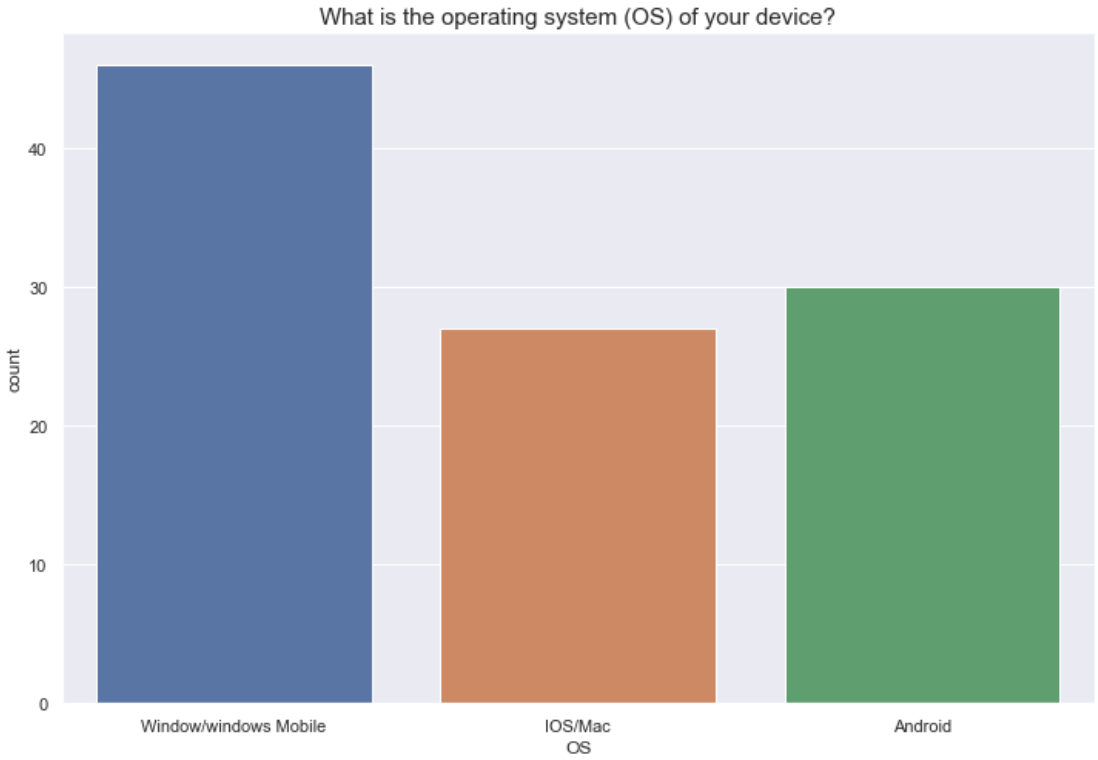

Fig.8 (OS)

The visualization in Fig.8 is a count plot from sea-born library it shows the distribution of which OS users prefer for online shopping. Most of the customers are using Window's operating system on their devices.

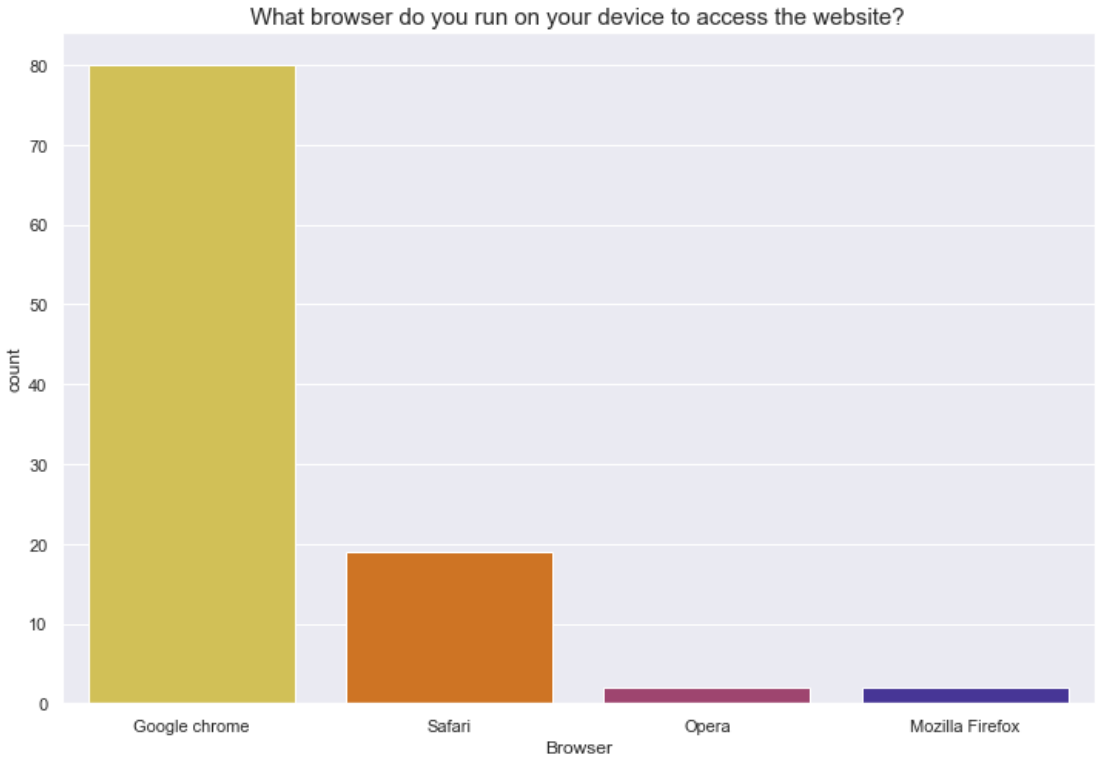

Fig.9 (Browser)

The visualization in Fig.9 is a count plot from sea-born library it shows the distribution of which OS users prefer for online shopping. Most of the customers prefer using Google Chrome as their default browser for online shopping.

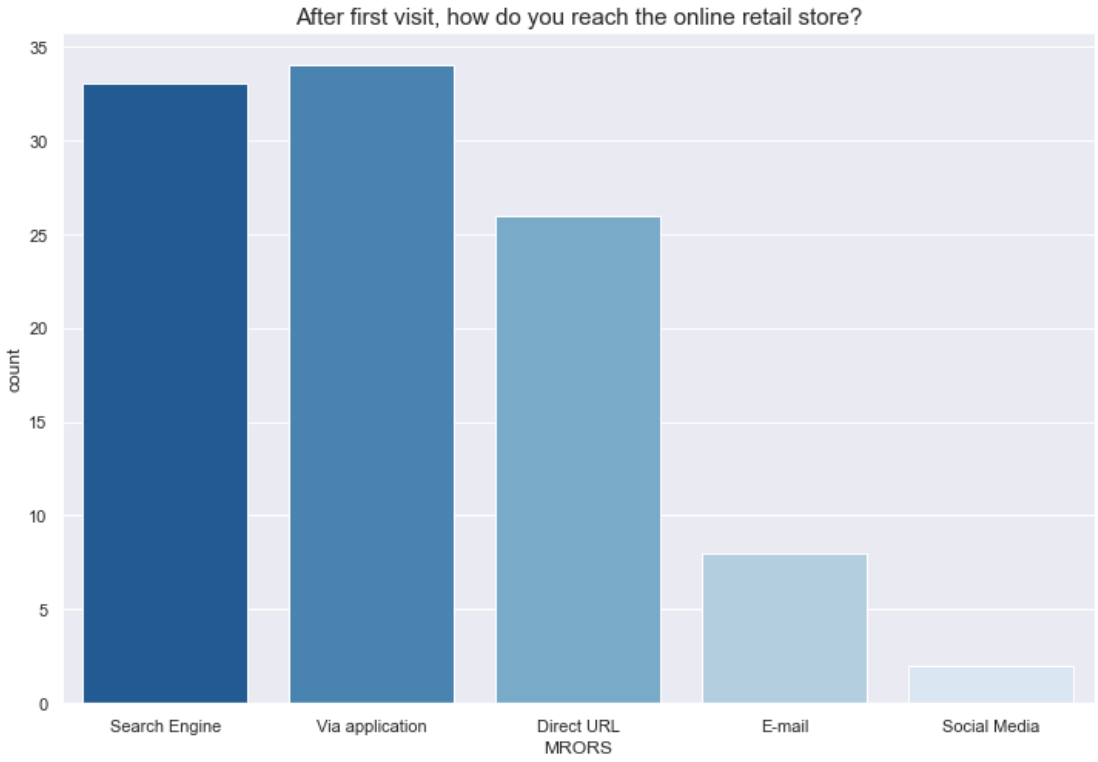

Fig.10 (MRORS)

The visualization in Fig.10 is a count plot from sea-born library it shows the distribution regarding how customers reached the online store after first visit. After the first visit, most of the individuals prefer to use the Application (mobile app) of the online store or used the Search Engine channel to reach the online store. Social media had very less influence (Companies can work on their social media marketing).

Fig.11 (Duration(minutes))

The visualization in Fig.11 is a count plot from the sea-born library it shows the distribution for how much time a customer explored the website before making any kind of purchase. On average an individual spent more than 15 minutes on the website before making a purchase decision.

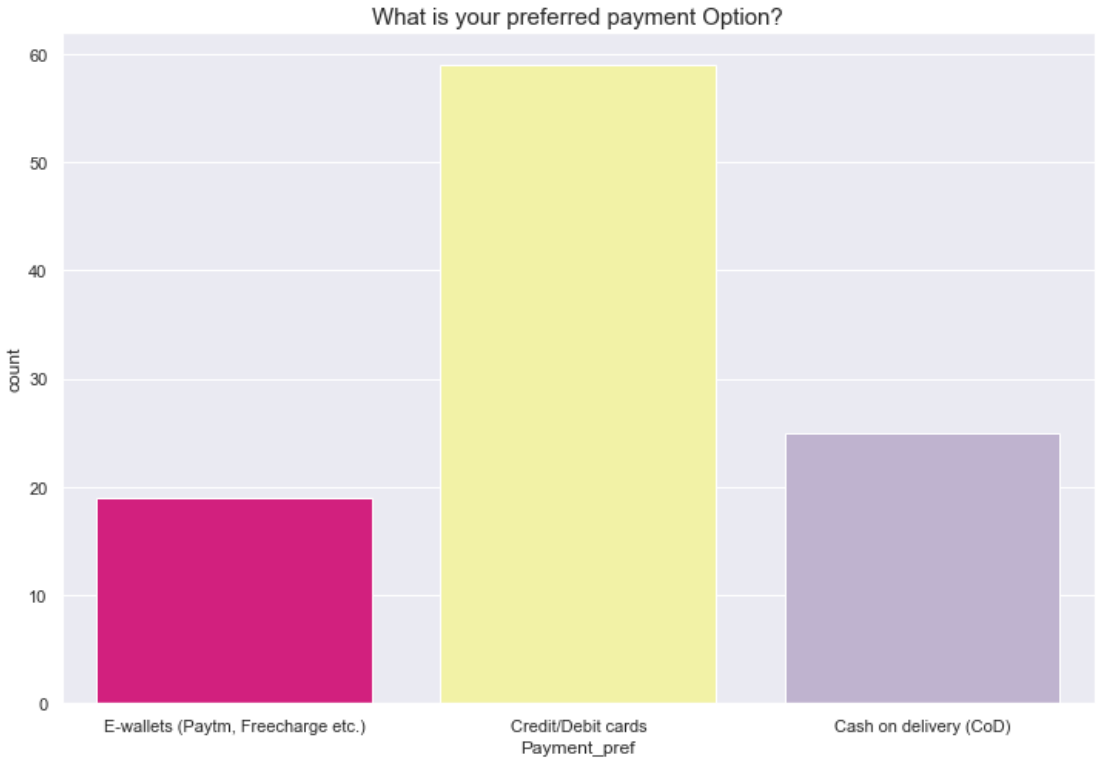

Fig.12 (Payment Preference)

The visualization in Fig.12 is a count plot from the sea-born library it shows the distribution of payment preference of customers. Most of the customers preferred a Credit/Debit cards as their payment method. Very few customers preferred E-wallets (Paytm, Freecharge, etc.)

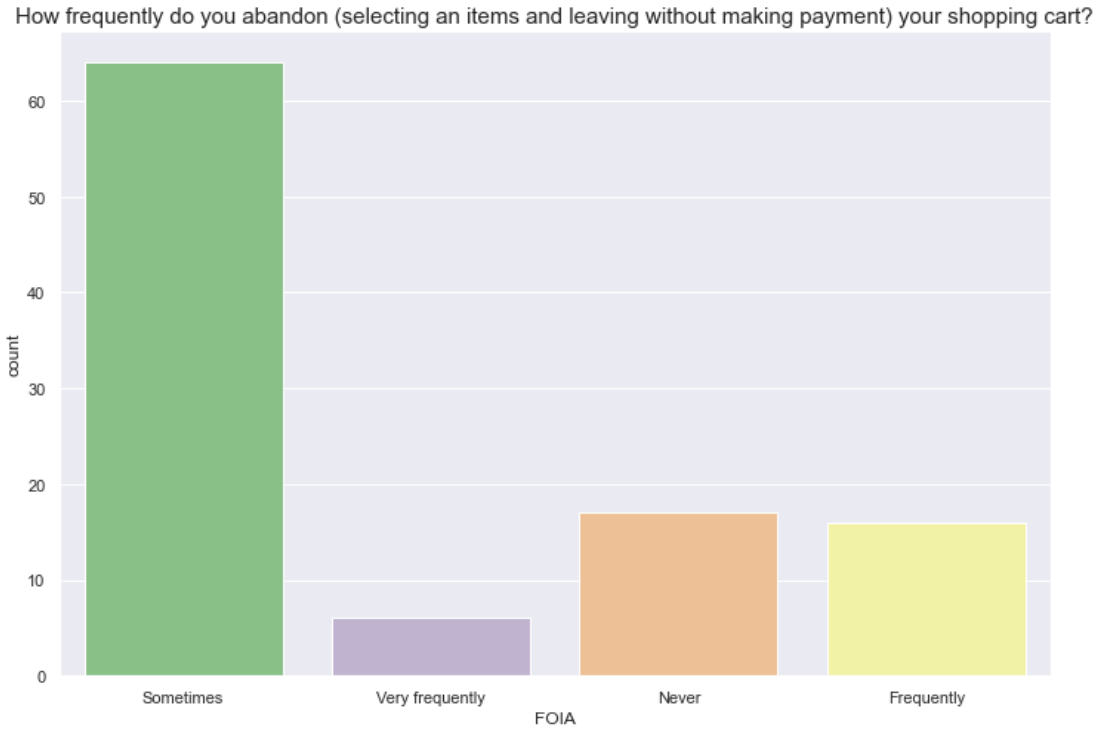

Fig.13 (FOIA)

The visualization in Fig.13 is a count plot from the sea-born library it shows the distribution of how frequently do customers abandon (selecting an items and leaving without making payment).When it comes to adding an item to the cart but not buying it, most of the customers tend to do it "sometimes" (Due to various personal reasons).

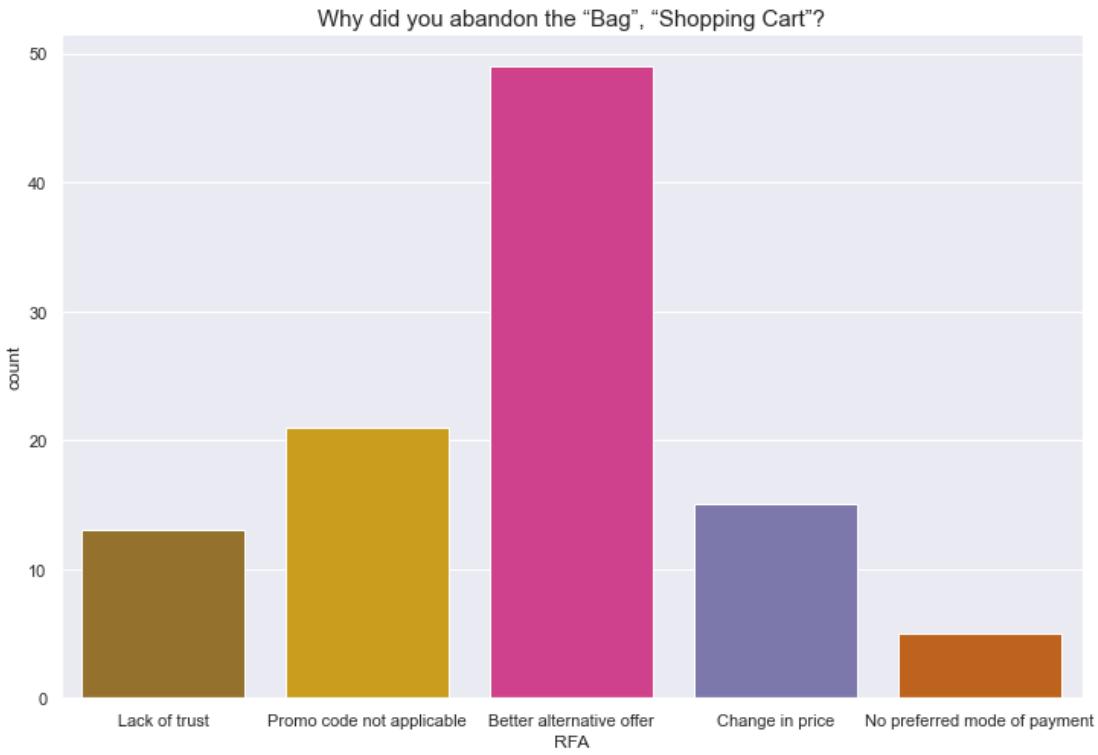

Fig.14 (RFA)

The visualization in Fig.14 is a count plot from the sea-born library it shows the distribution of why did a customer abandon the “Bag”, “Shopping Cart”.The most frequent reason for customers not buying the added item in the cart is having a better alternative offer on the same product.

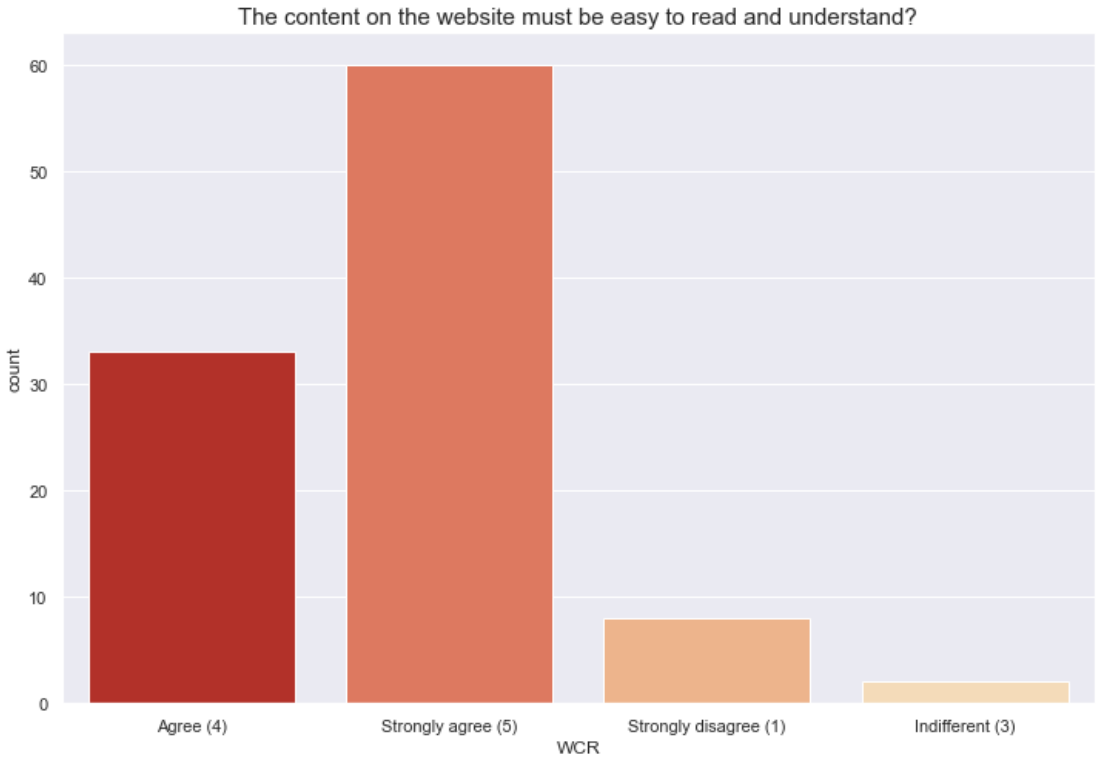

Fig.15 (WCR)

The visualization in Fig.15 is a count plot from the sea-born library it shows the distribution is regarding customer response on weather the content on the website must be easy to read and understand. It is very important that the content on the website must be easy to read and understand.

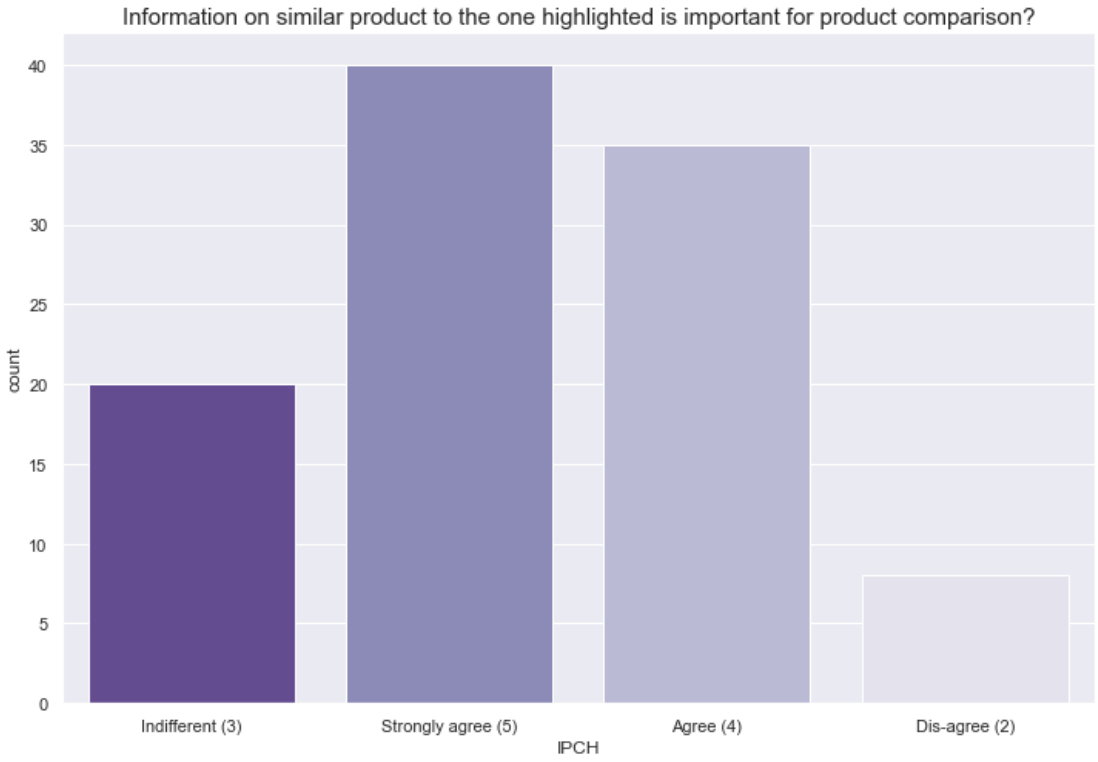

Fig.16 (IPCH)

The visualization in Fig.16 is a count plot from the sea-born library it clearly concludes that product comparison is a must and is an important factor as the customer will always try to buy what suits best for his needs. Suggesting similar products should be imminent.

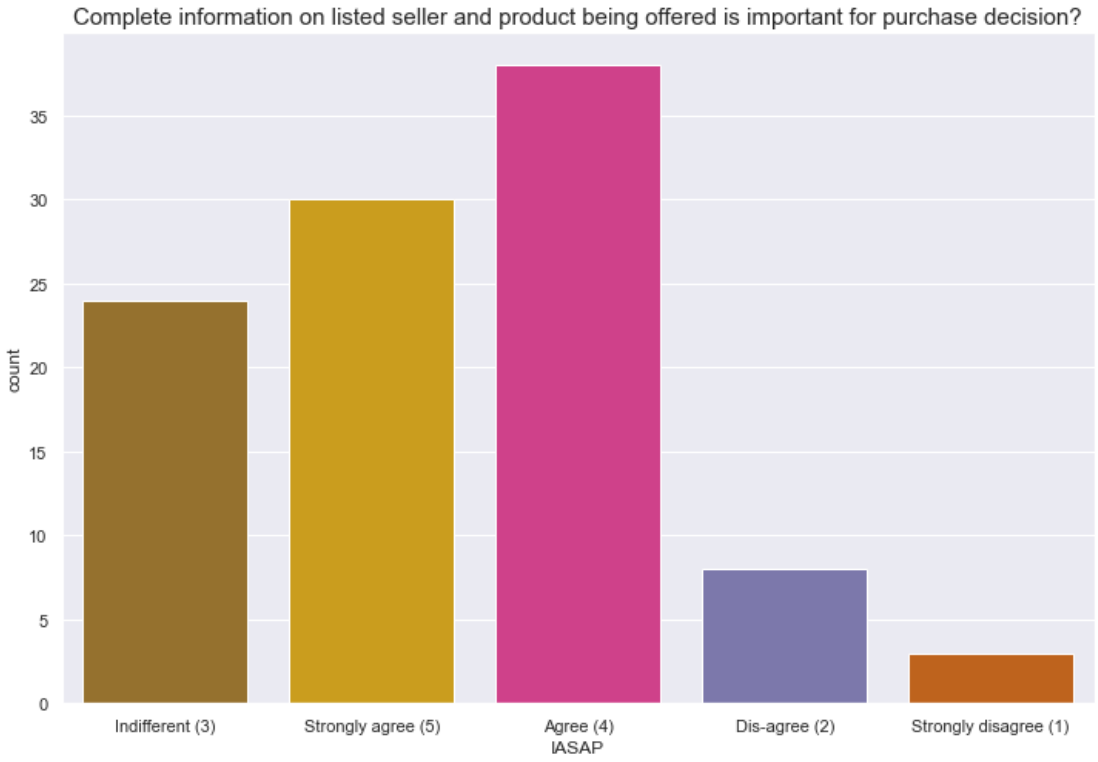

Fig.17 (IASAP)

The visualization in Fig.17 is a count plot from the sea-born library it clearly concludes that most of the customers agree with providing complete information on the listed sellers and products being offered as it gives them extra assurance in case something went wrong with the product.

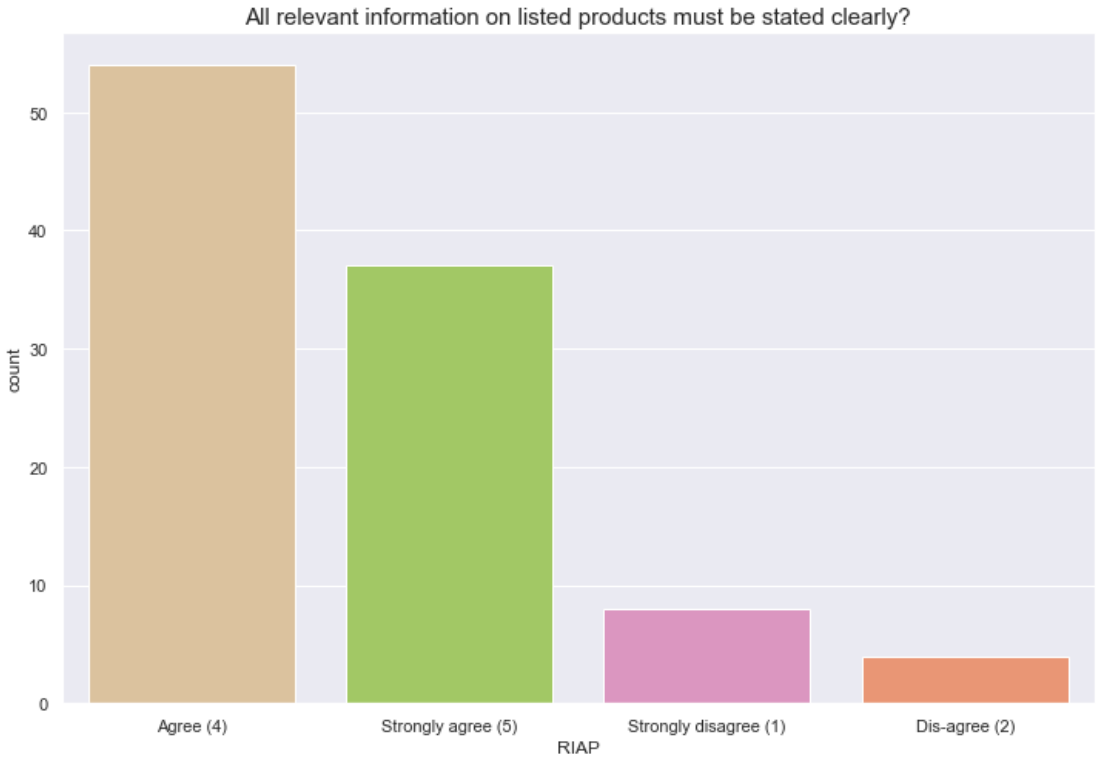

Fig.18 (RIAP)

The visualization in Fig.18 is a count plot from the sea-born library it clearly concludes that It is very important that the product information(specifications) should be accurate and authentic.

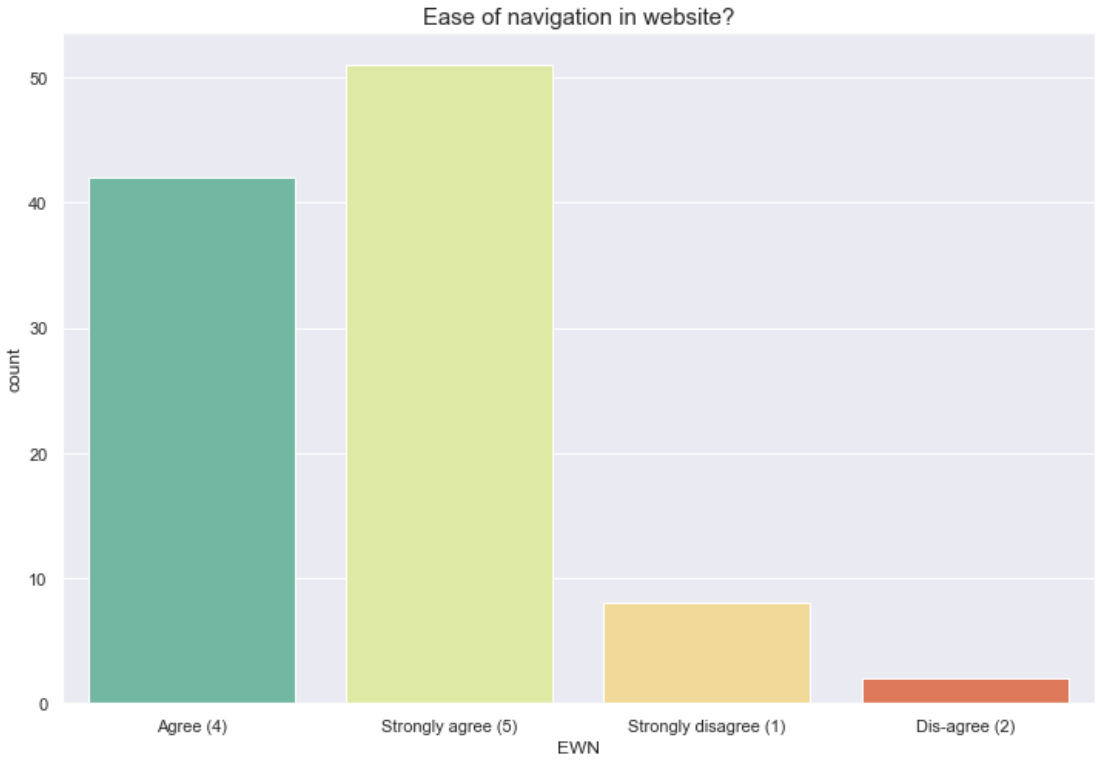

Fig.19 (EWN)

The visualization in Fig.19 is a count plot from the sea-born library it clearly concludes that The websites for online stores must be simple to navigate. New users must not have any difficulty while shopping.

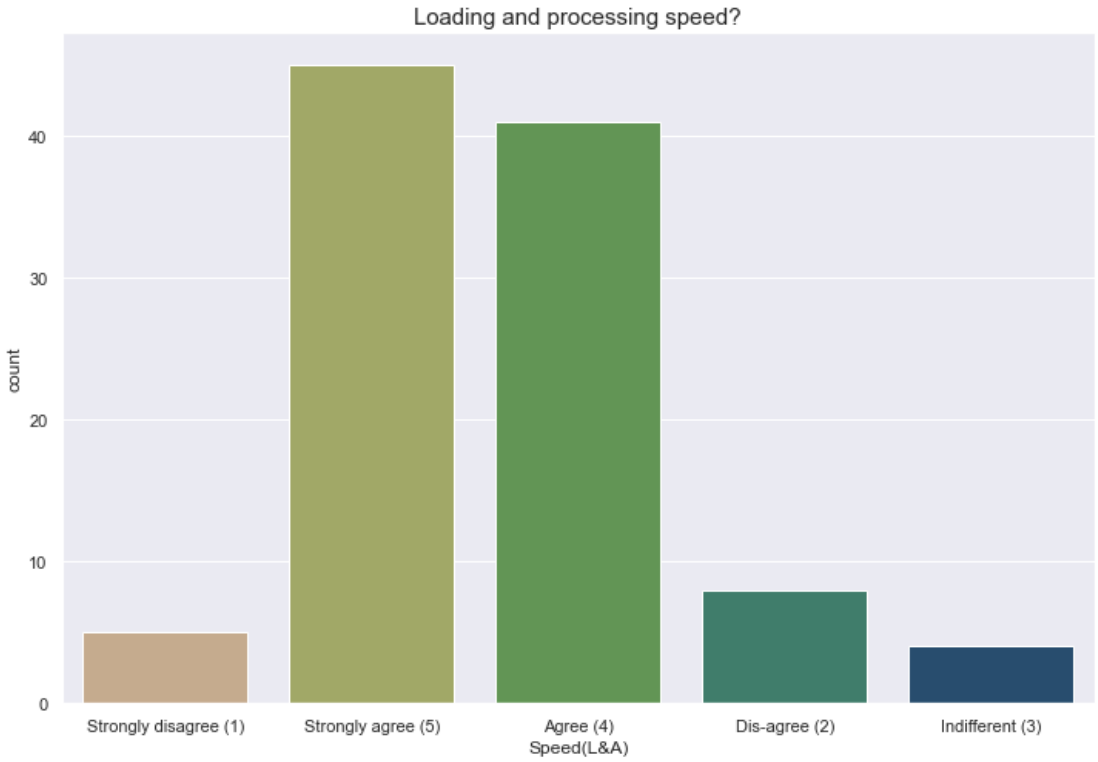

Fig.20 (L&A)

The visualization in Fig.20 is a count plot from the sea-born library it clearly concludes that It is very important for customers that the loading speed for their favorite online store should be adequate.

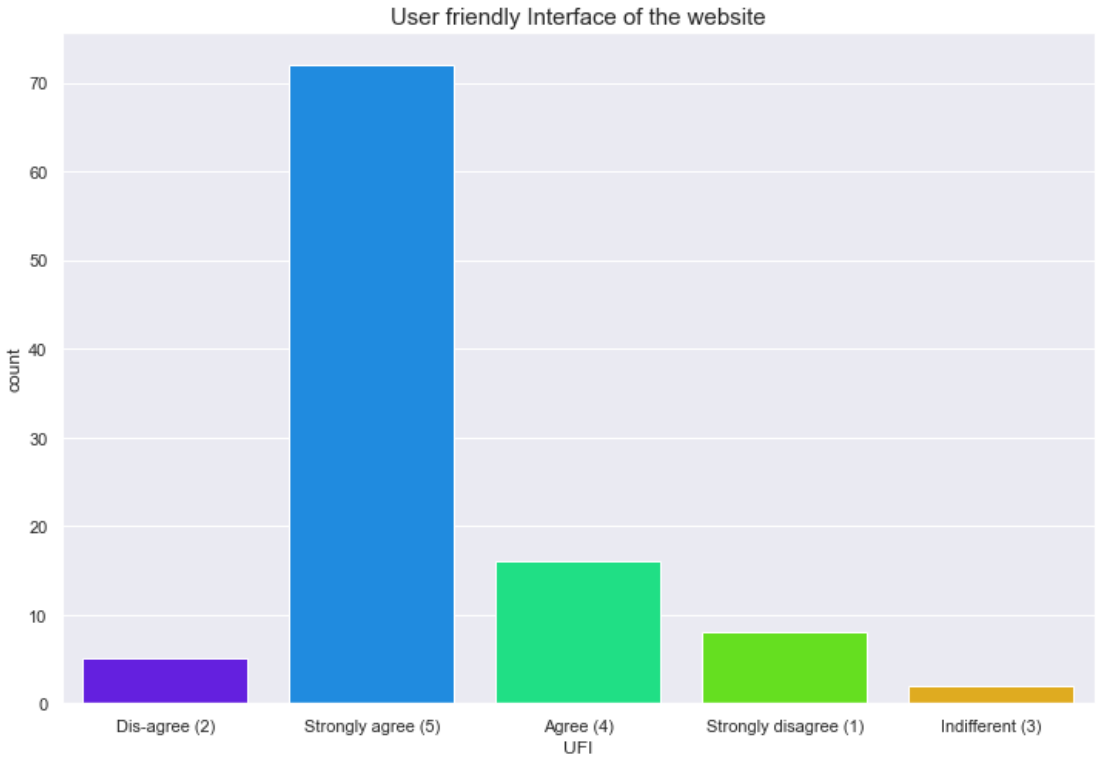

Fig.21 (UFI)

The visualization in Fig.21 is a count plot from the sea-born library it clearly concludes that from the user's perspective, a user-friendly UI(user interface) is very important.

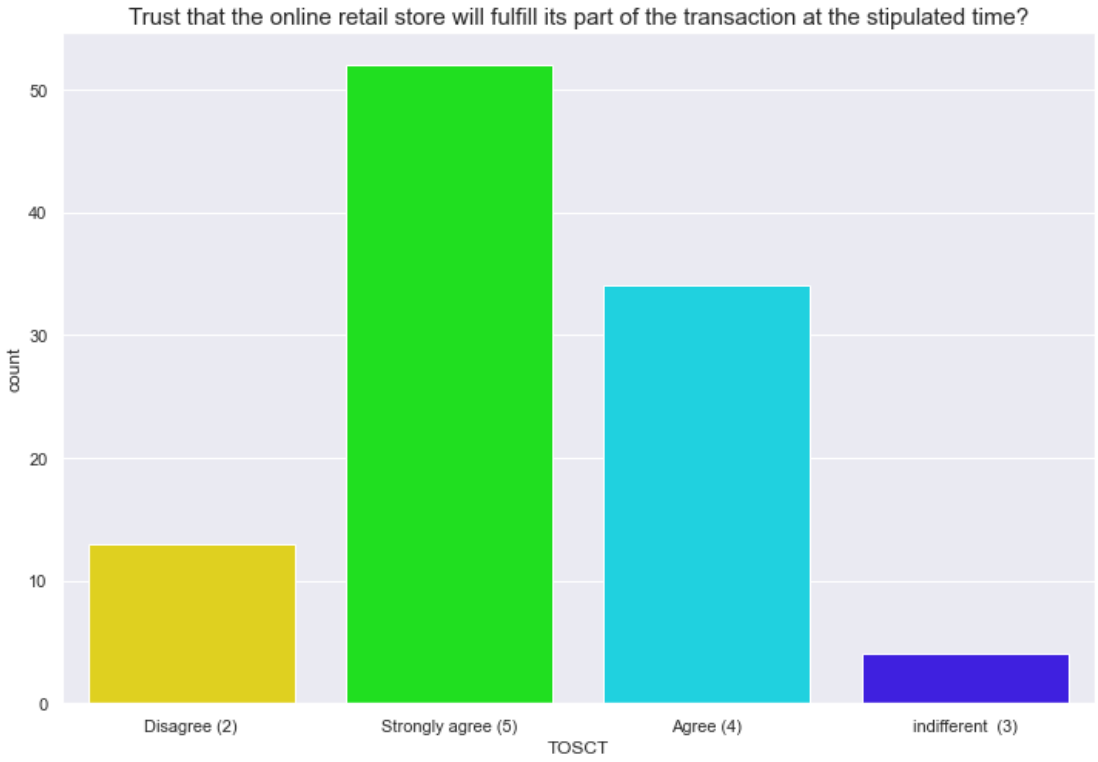

Fig.22 (TOSCT)

The visualization in Fig.22 is a count plot from the sea-born library it clearly concludes that, probably the most important factor for users in terms of payment. The payment should be completed in a time span as it will increase the trustworthiness of customers.

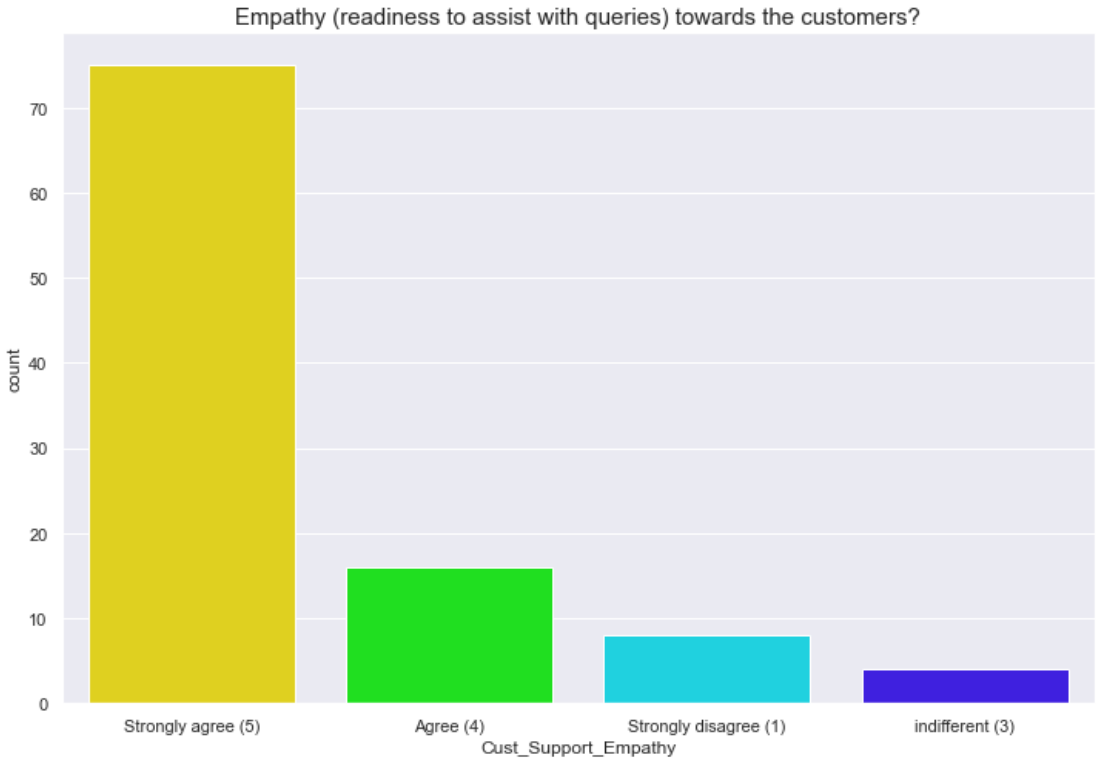

Fig.23 (Cust_Support_Empathy)

The visualization in Fig.23 is a count plot from the sea-born library it clearly concludes that, customer support is a crucial factor to improve customer loyalty and trust.

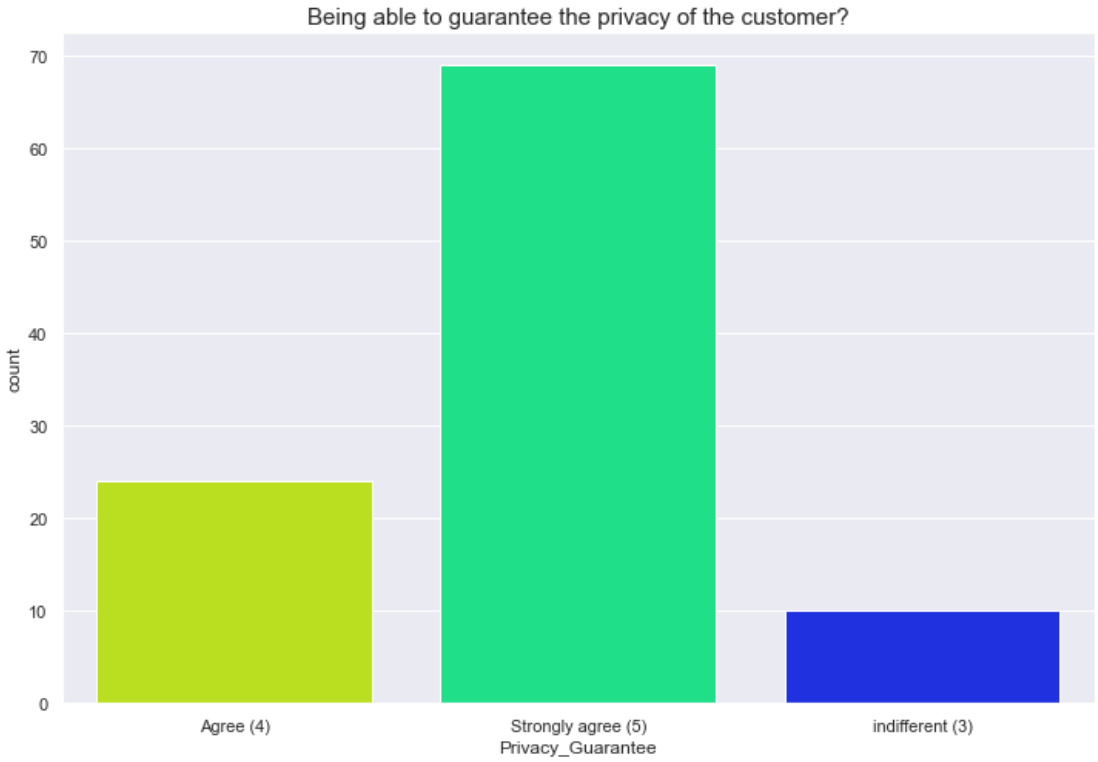

Fig.24 (Privacy_Guarantee)

The visualization in Fig.24 is a count plot from the sea-born library it clearly concludes that, customer privacy is a crucial factor. Customers should feel safe while shopping.

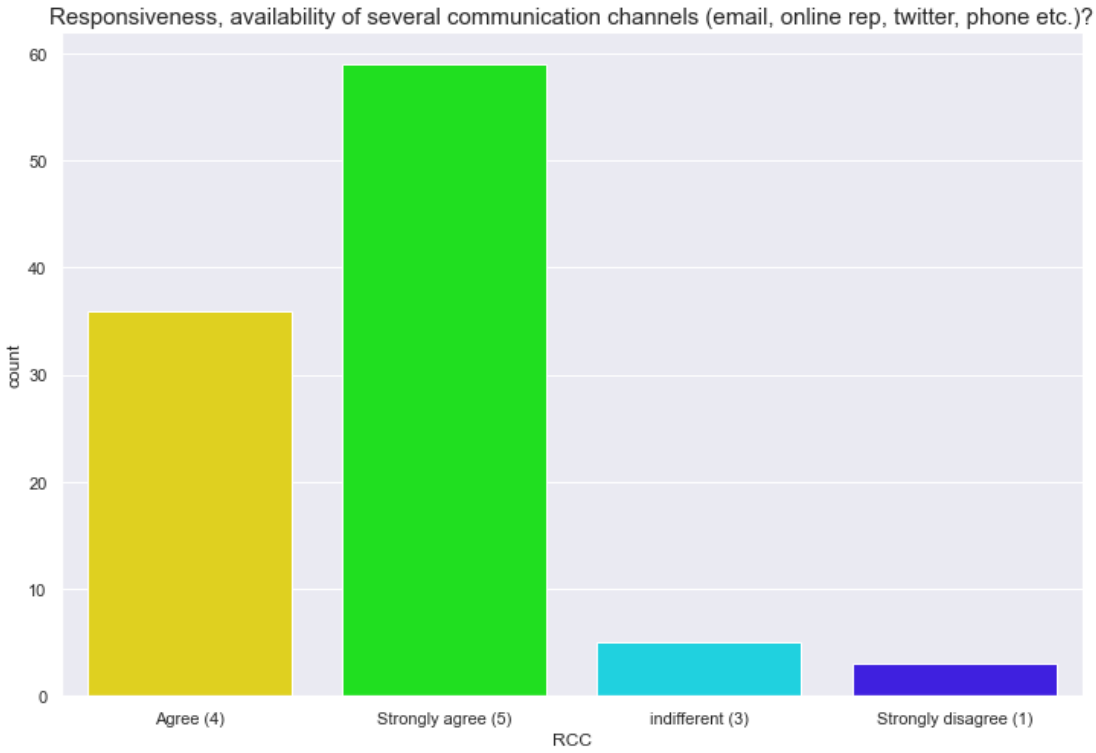

Fig.25 (RCC)

The visualization in Fig.25 is a count plot from the sea-born library it clearly concludes that, the website should be responsive as users will use different devices to access the website. Having various channels for handling user queries is crucial.

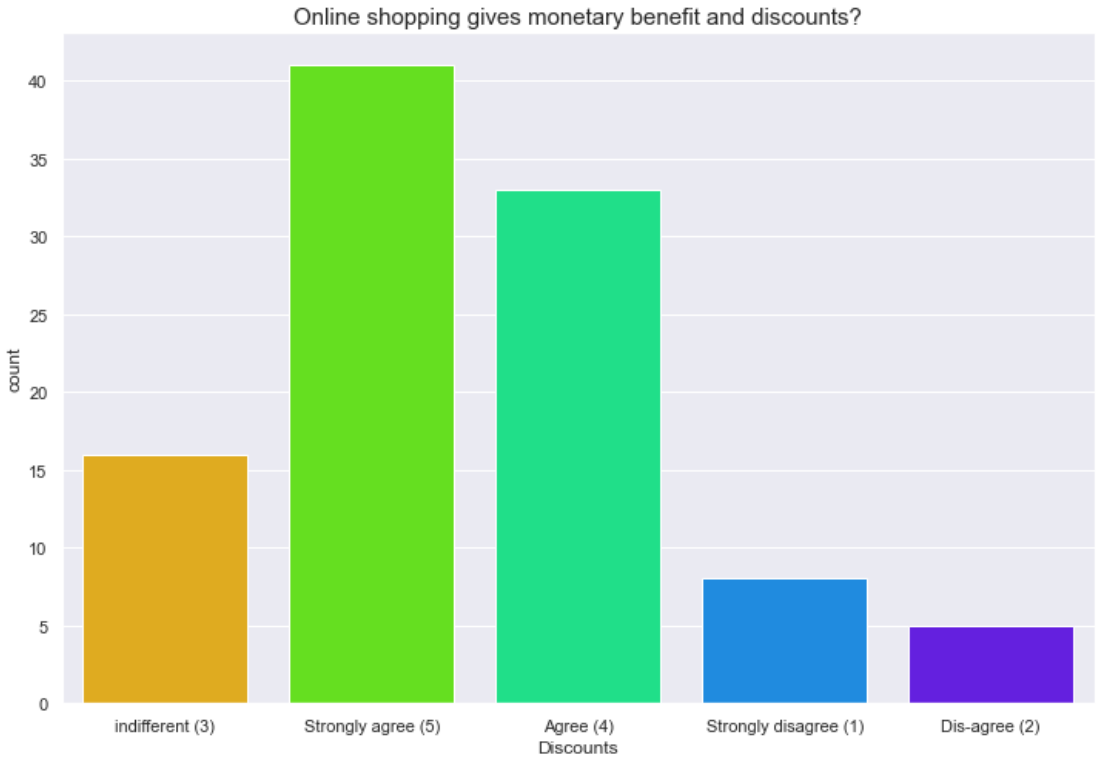

Fig.26 (Discounts)

The visualization in Fig.26 is a count plot from the sea-born library it clearly concludes that, monetary benefits and Discounts are influential to attract more customers in buying products online.

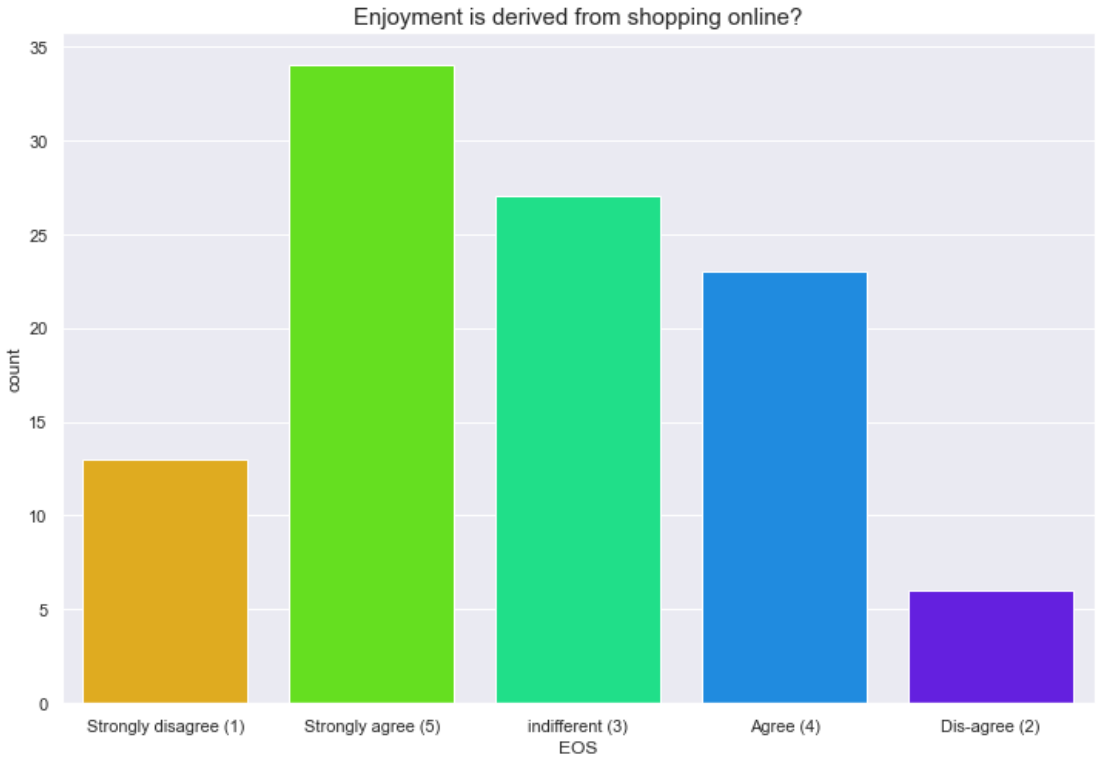

Fig.27 (EOS)

The visualization in Fig.27 is a count plot from the sea-born library it clearly concludes that, most of the users enjoy the experience of online shopping.

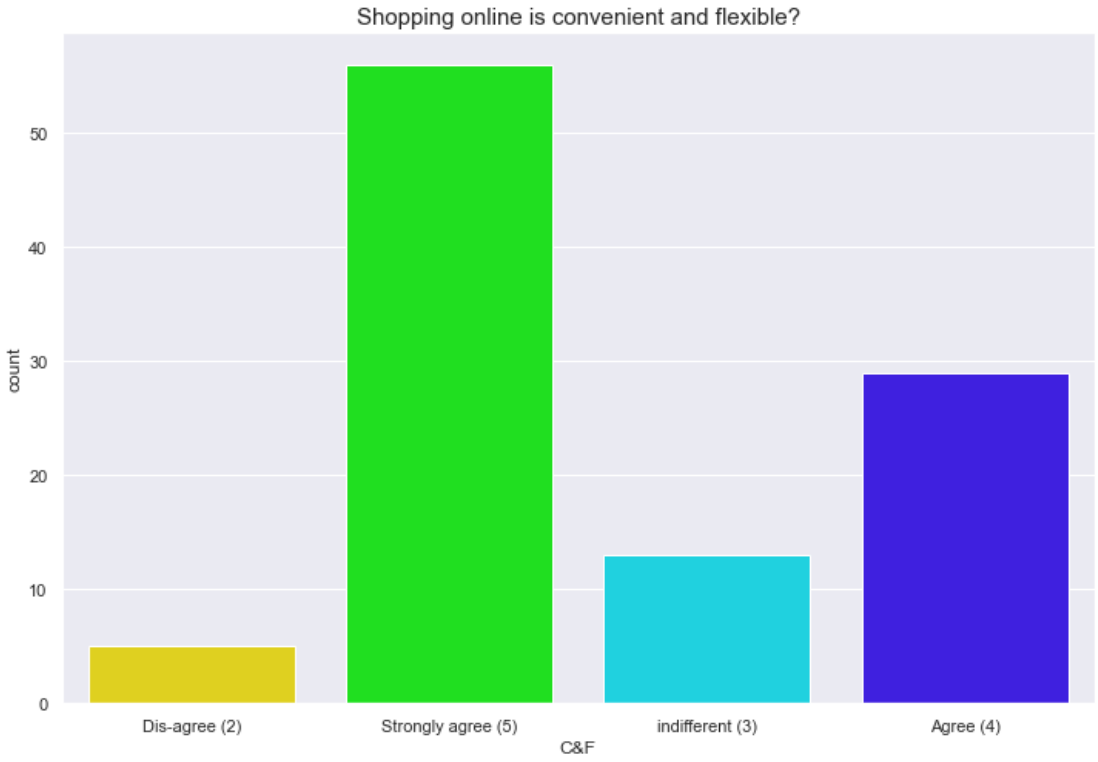

Fig.28 (C&F)

The visualization in Fig.28 is a count plot from the sea-born library it clearly concludes that, highest number of customers believe that online shopping is convenient and flexible.

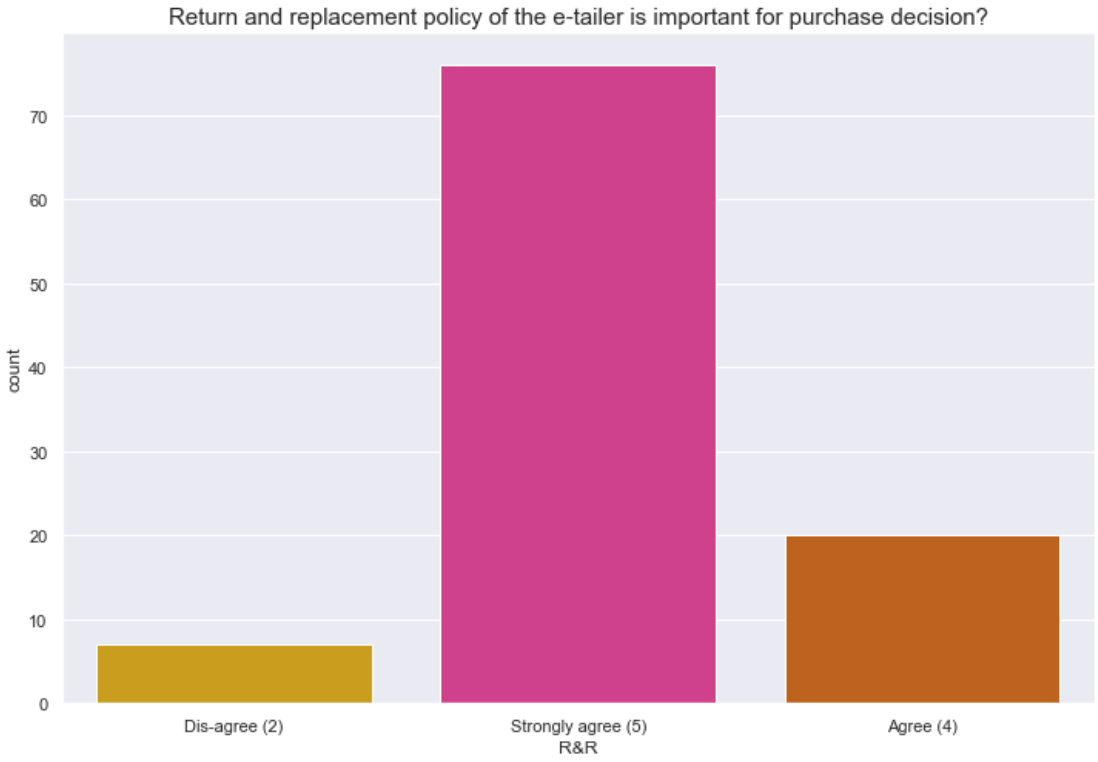

Fig.29 (R&R)

The visualization in Fig.29 is a count plot from the sea-born library it clearly concludes that, In case the delivered product turns out to be faulty than in such cases return and replacement policy is very crucial.

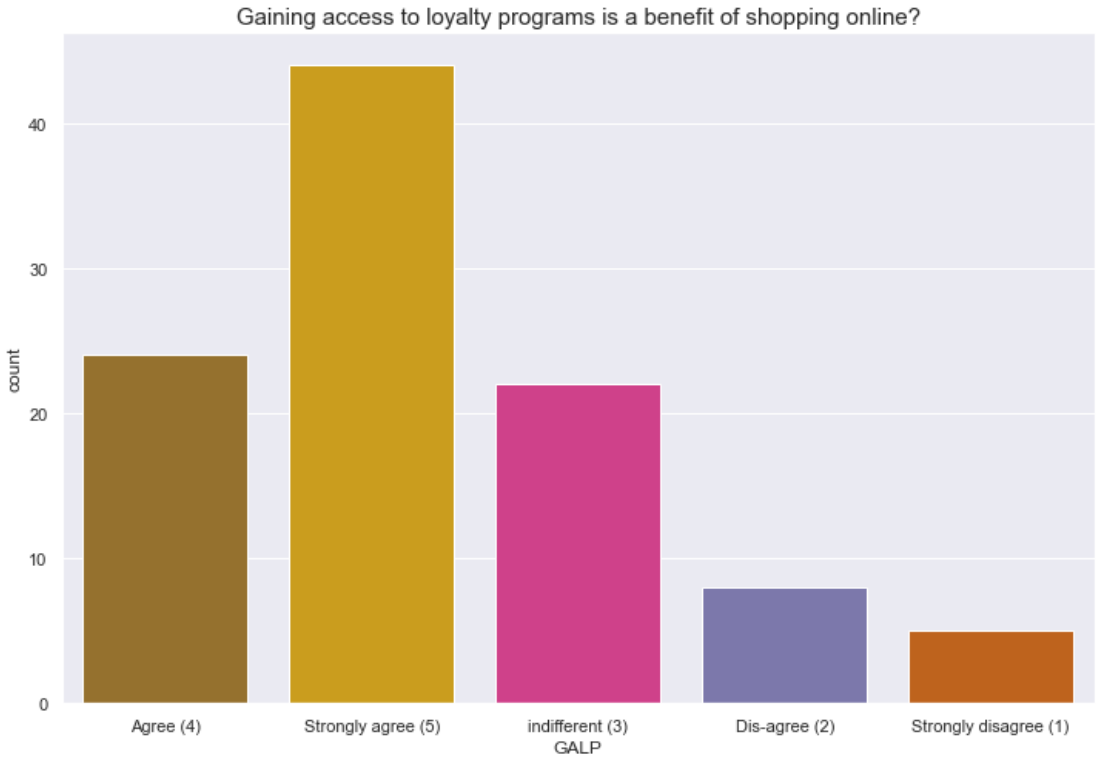

Fig.30 (GALP)

The visualization in Fig.30 is a count plot from the sea-born library it clearly concludes that, service like this can enhance the experience for loyal customers.

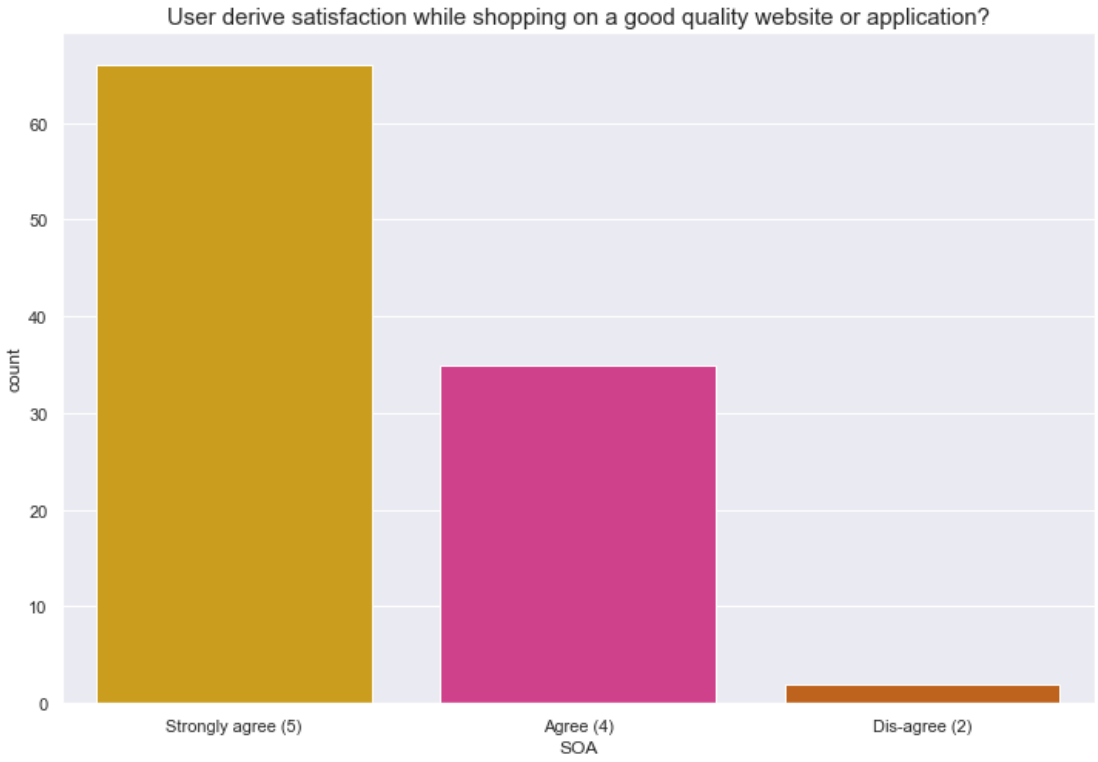

Fig.31 (SOA)

The visualization in Fig.31 is a count plot from the sea-born library it clearly concludes that the website should be attractive, easy to navigate, and should have good customer support for users to be satisfied.

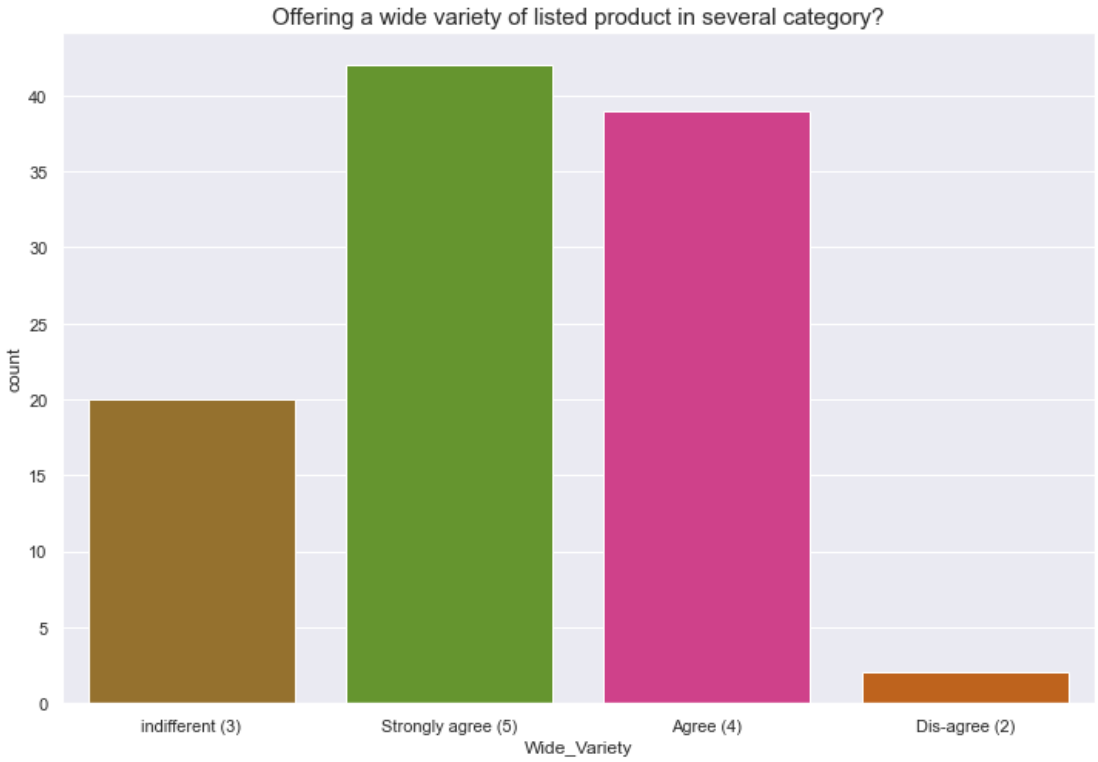

Fig.32 (Wide_Variety)

The visualization in Fig.32 is a count plot from the sea-born library it clearly concludes that, when it comes to shopping has a lot of alternatives for products is crucial as it helps the customer to choose the best option.

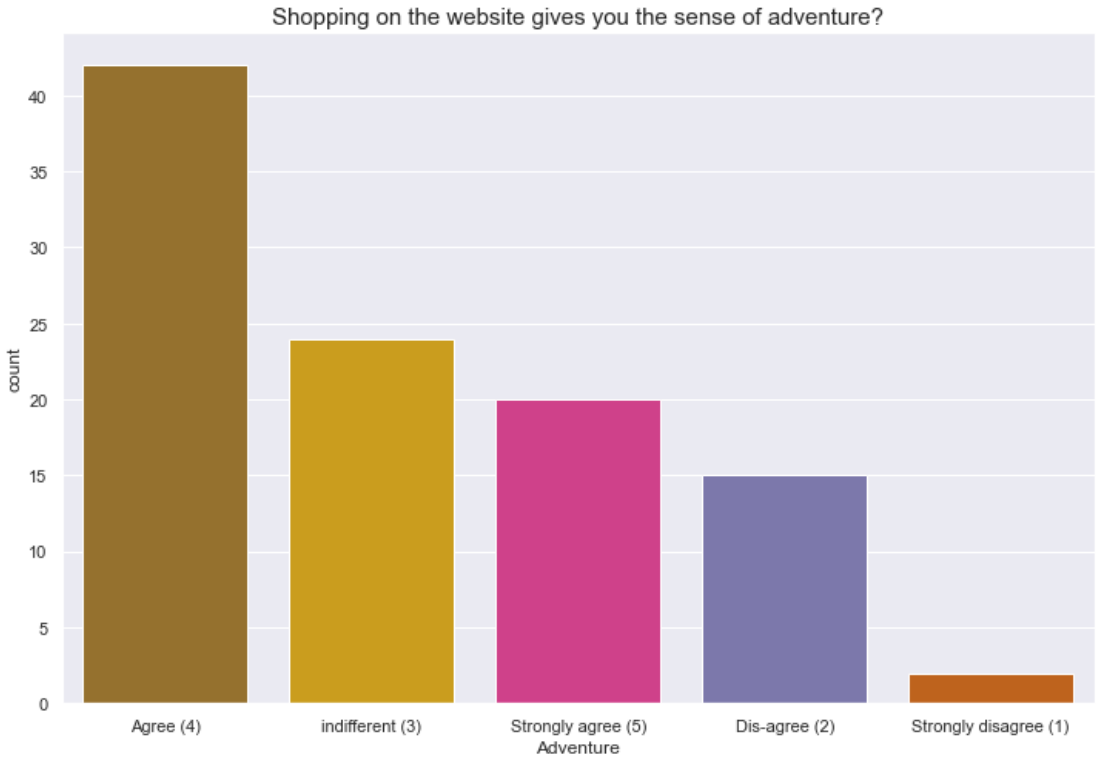

Fig.33 (Adventure)

The visualization in Fig.33 is a count plot from the sea-born library it clearly concludes that, most of the customers think that shopping online gives them a sense of adventure.

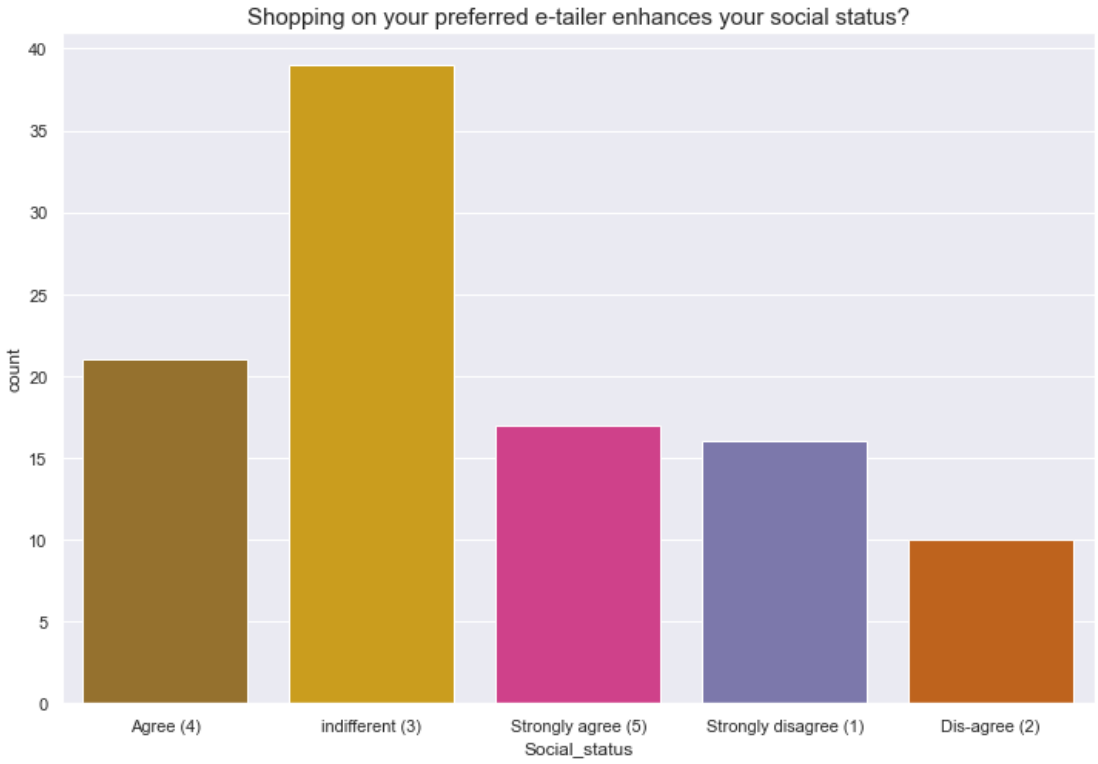

Fig.34 (Social_status)

The visualization in Fig.34 is a count plot from the sea-born library it clearly concludes that, according to most of the customers, there is no effect of online shopping on social status.

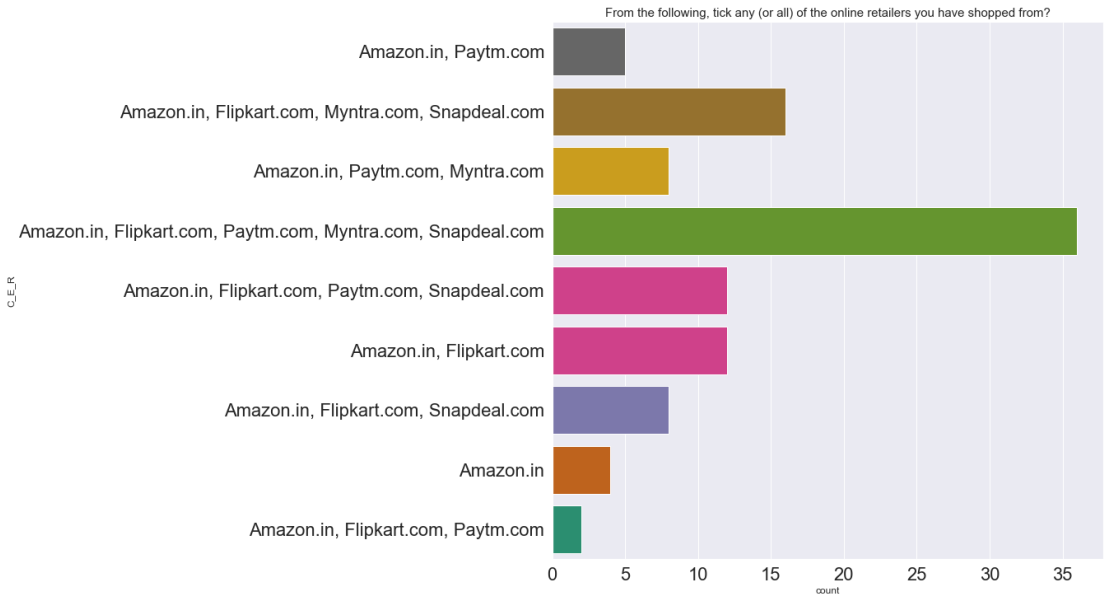

Fig.35(C_E_R)

The visualization in Fig.35 is a count plot from the sea-born library it clearly concludes that, Most of the customers prefer Amazon. in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com for shopping. we can say that these are the top 5 best e-commerce companies in India.

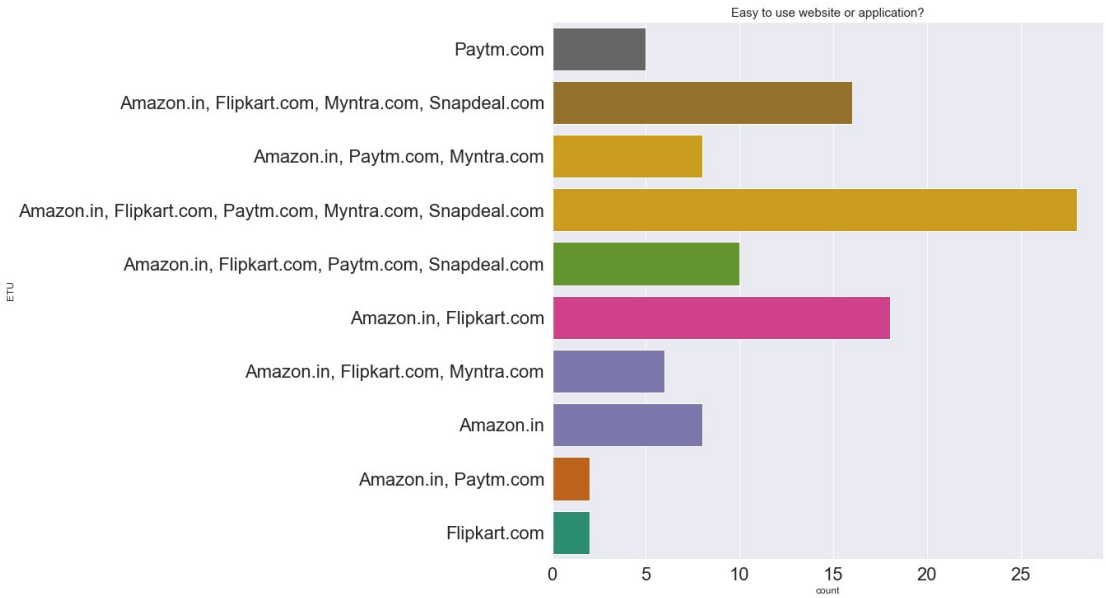

Fig.36 (ETU)

The visualization in Fig.36 is a count plot from the sea-born library it clearly concludes that, According to most of the customer's Amazon. in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com are easy-to-use websites for online shopping.

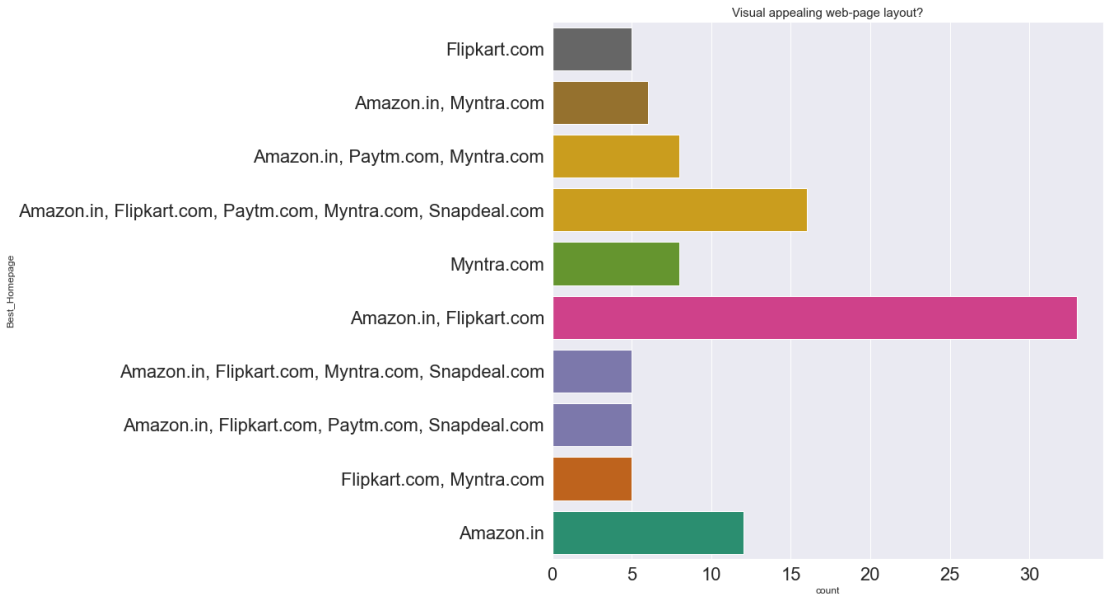

Fig.37 (Best_Homepage)

The visualization in Fig.37 is a count plot from the sea-born library it clearly concludes that, According to most of the customer's Amazon. in, Flipkart.com has the best home page in terms of design and user experience.

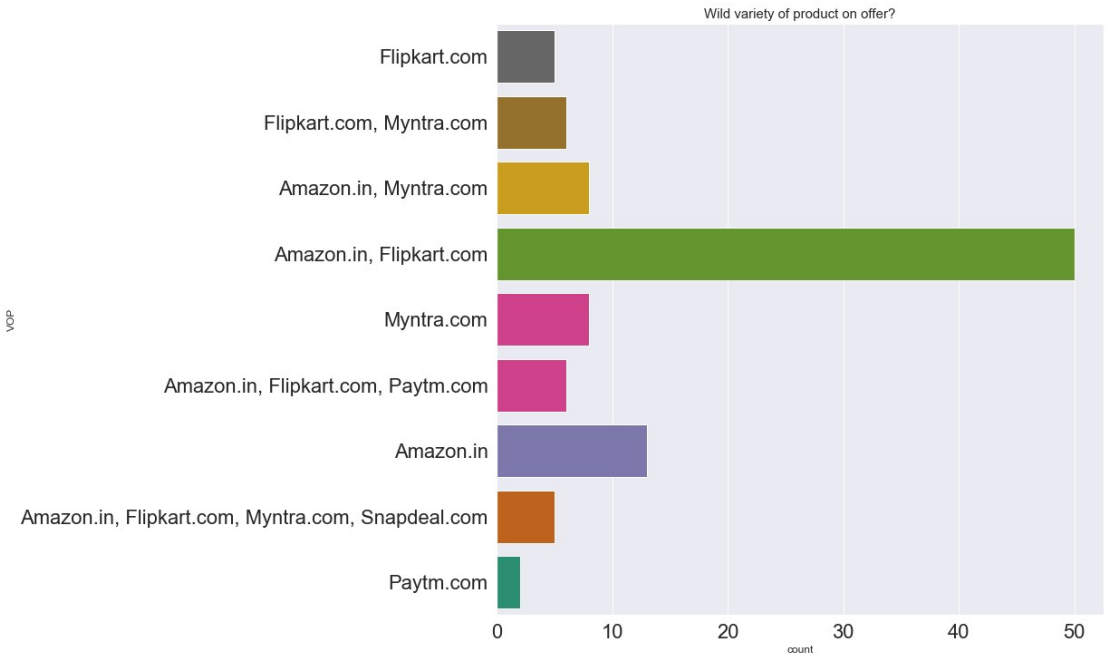

Fig.38 (VOP)

The visualization in Fig.38 is a count plot from the sea-born library it clearly concludes that, customers are voting for Amazon. in and Flipkart.com as the best e-commerce sites when it comes to the variety of products.

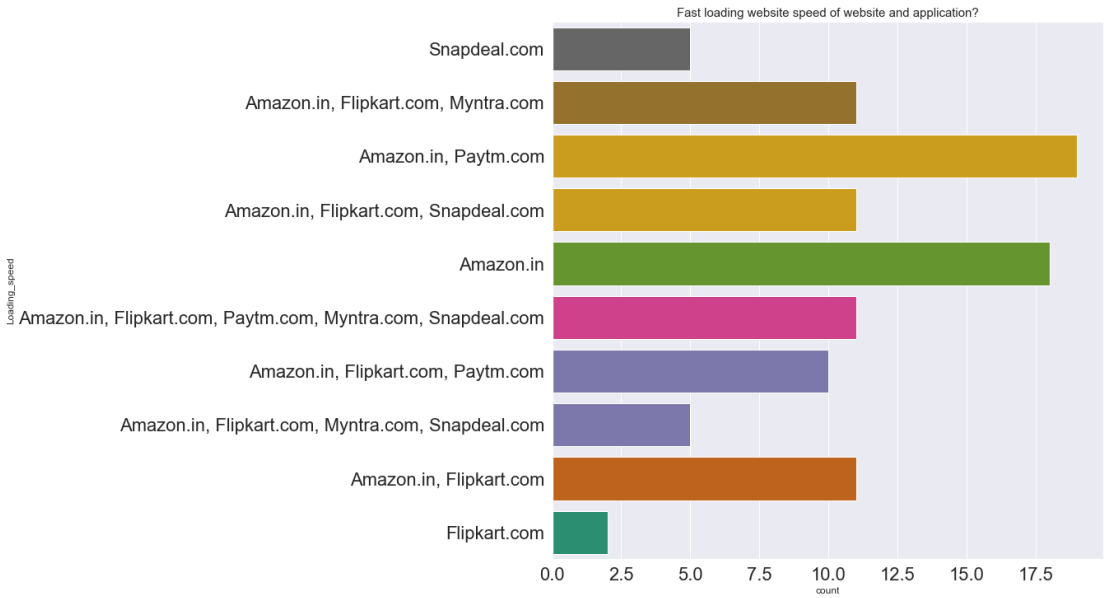

Fig.39 (Loading_speed)

The visualization in Fig.39 is a count plot from the sea-born library it clearly concludes that, When it comes to loading speed Amazon. in, Paytm.com we voted as the fastest loading websites. According to this survey, Flipkart.com has a slower loading speed.

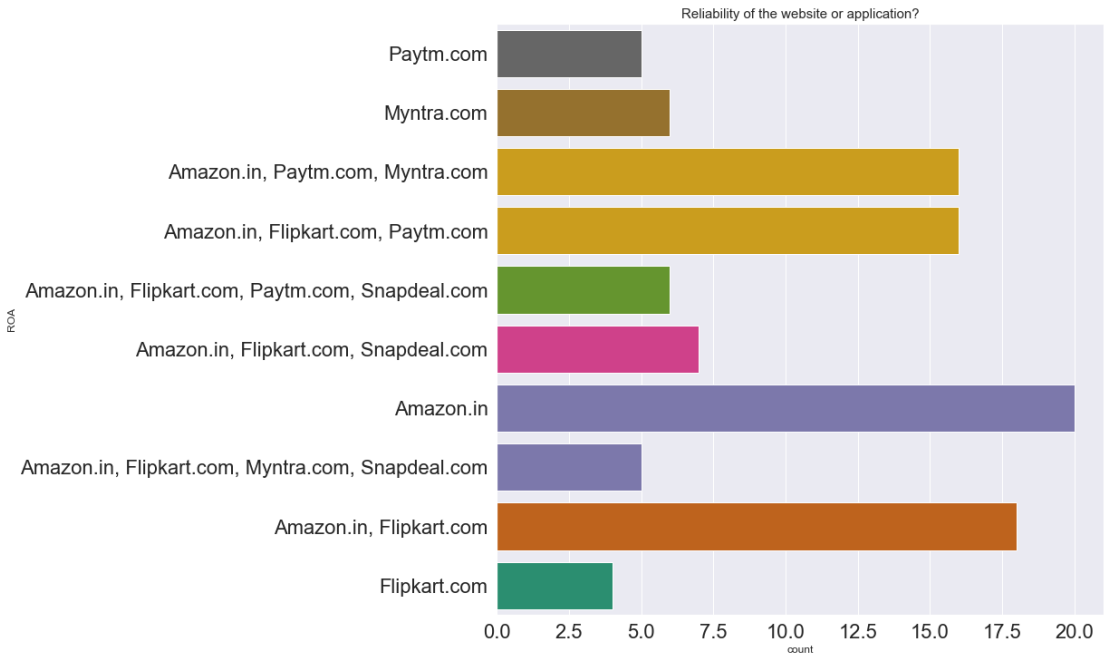

Fig.40 (ROA)

The visualization in Fig.40 is a count plot from the sea-born library it clearly concludes that, Amazon. in is the most reliable e Commerce website as compared to other online shopping websites.

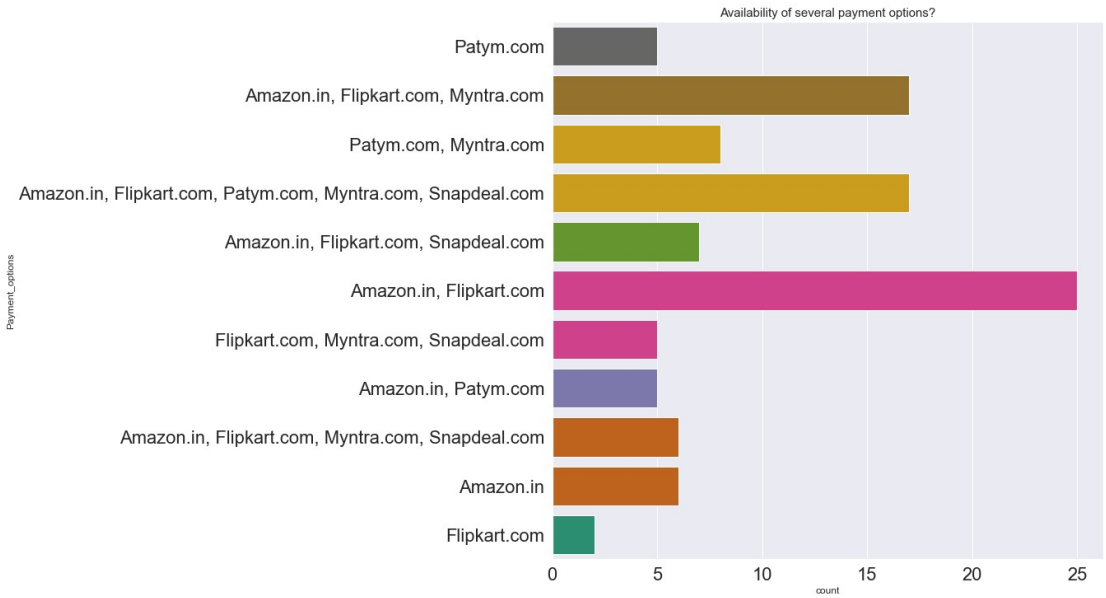

Fig.41 (Payment_options)

The visualization in Fig.41 is a count plot from the sea-born library it clearly concludes that, Amazon. in, Flipkart.com has the most number of payment alternatives.

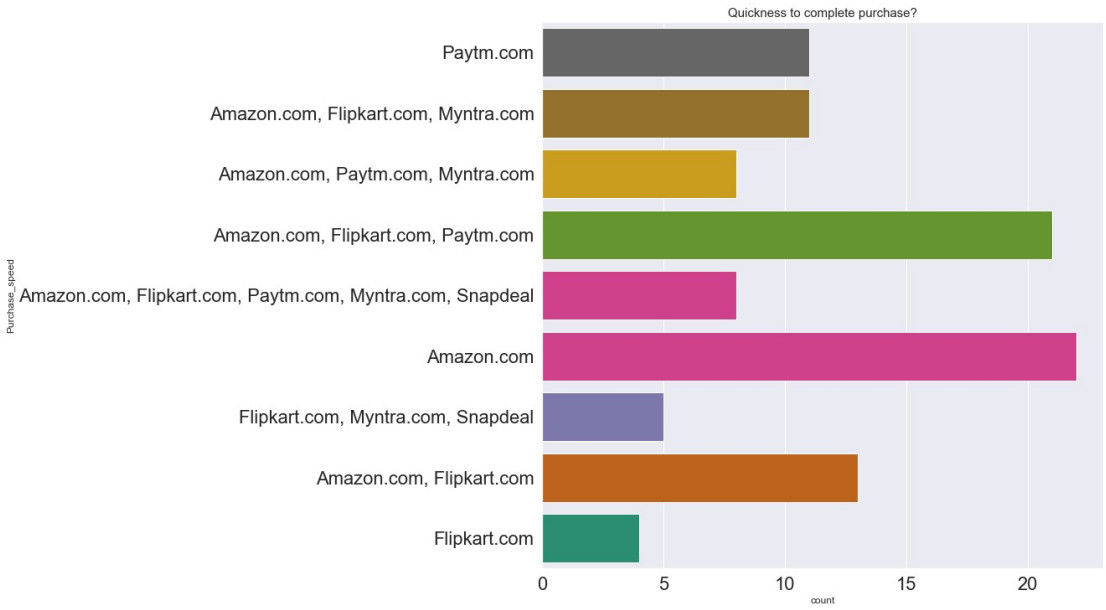

Fig.42 (Purchase_speed)

The visualization in Fig.42 is a count plot from the sea-born library it clearly concludes that, Amazon.in has the fastest purchase complete rate as compared to other websites.

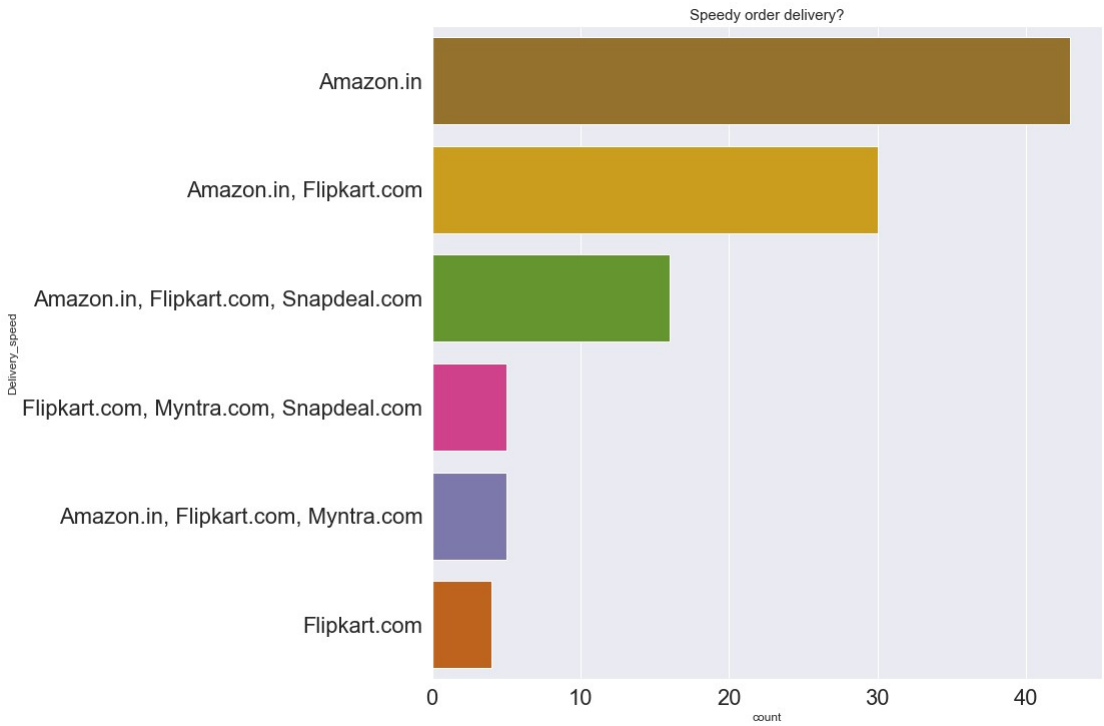

Fig.43 (Delivery_speed)

The visualization in Fig.43 is a count plot from the sea-born library it clearly concludes that, Amazon.com has the fastest delivery speed as compared to other e-commerce stores.

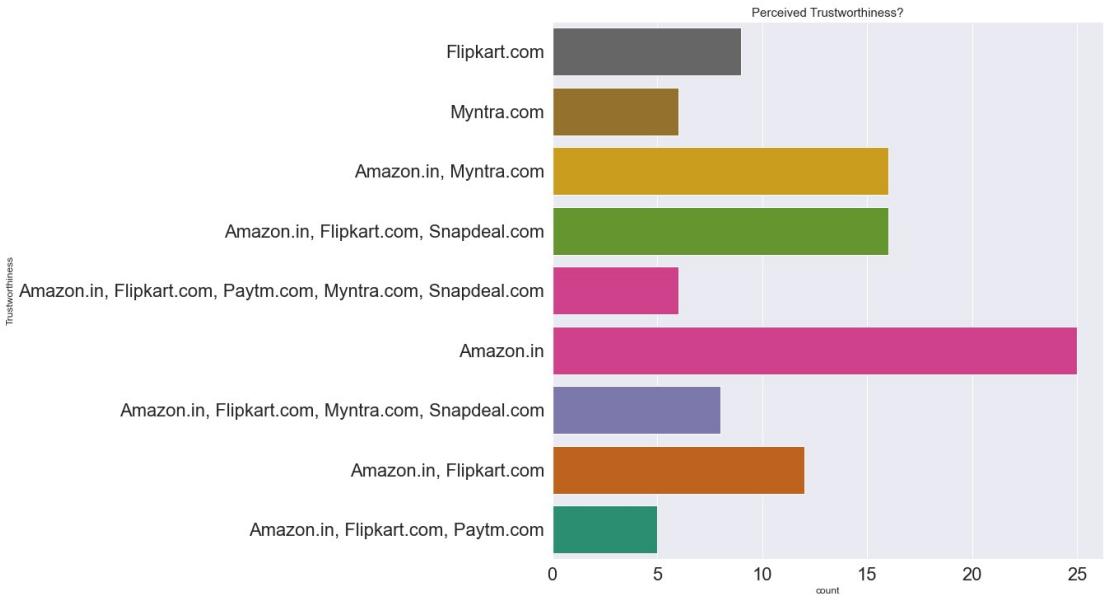

Fig.44 (Trustworthiness)

The visualization in Fig.44 is a count plot from the sea-born library it clearly concludes that, Amazon turns out to be the most trusted e-commerce website.

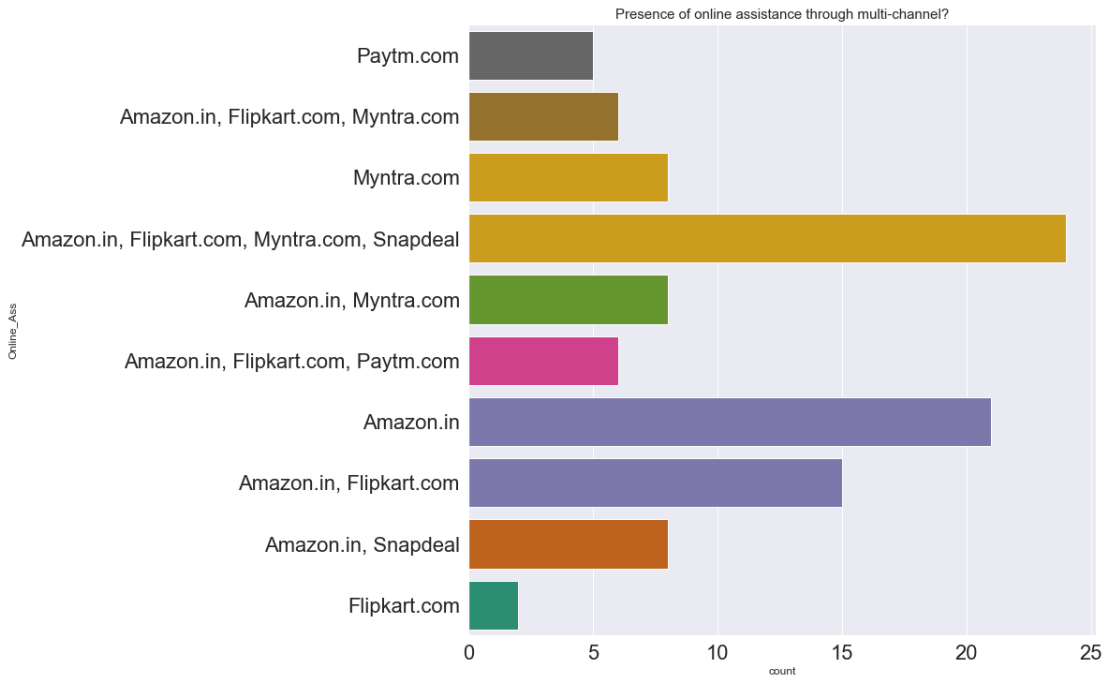

Fig.45 (Online_Ass)

The visualization in Fig.45 is a count plot from the sea-born library it clearly concludes that, Amazon.in, Flipkart.com, Myntra.com, Snapdeal have the highest customer votes for the Presence of online assistance through multi-channel.

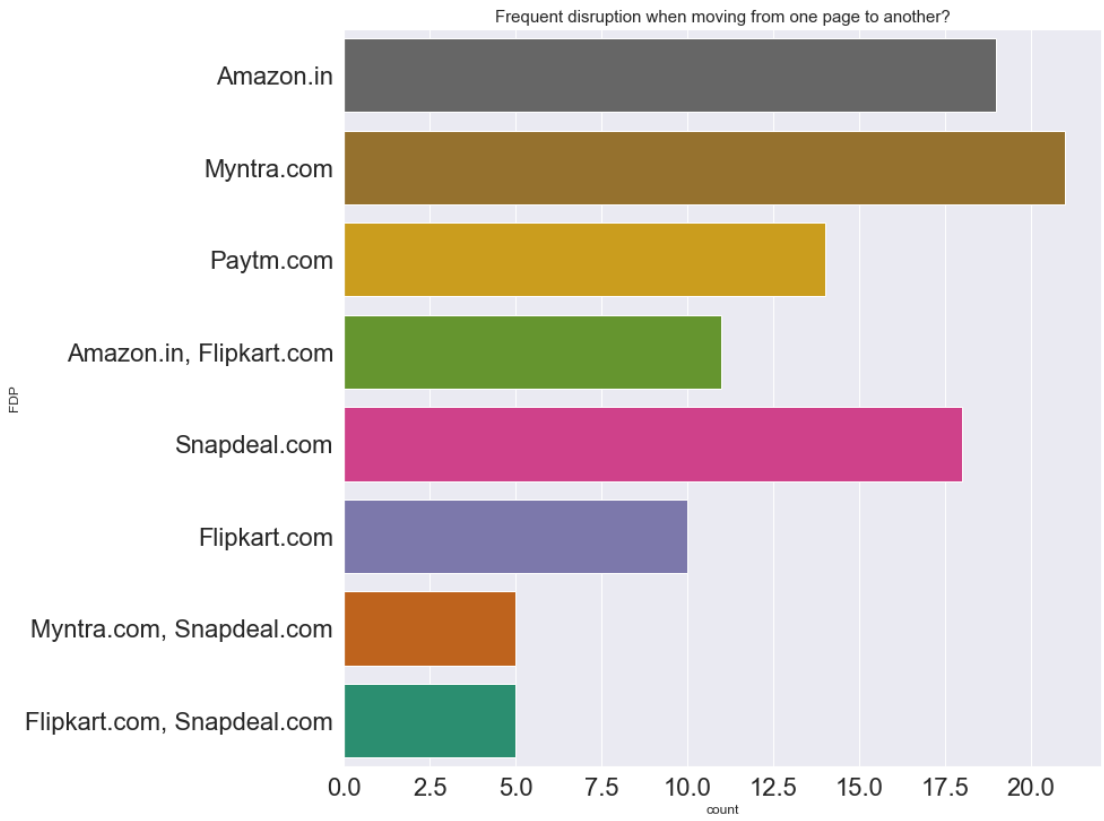

Fig.46 (FDP)

The visualization in Fig.46 is a count plot from the sea-born library it clearly concludes that, Myntra.com and Amazon.in have Frequent disruption when moving from one page to another.

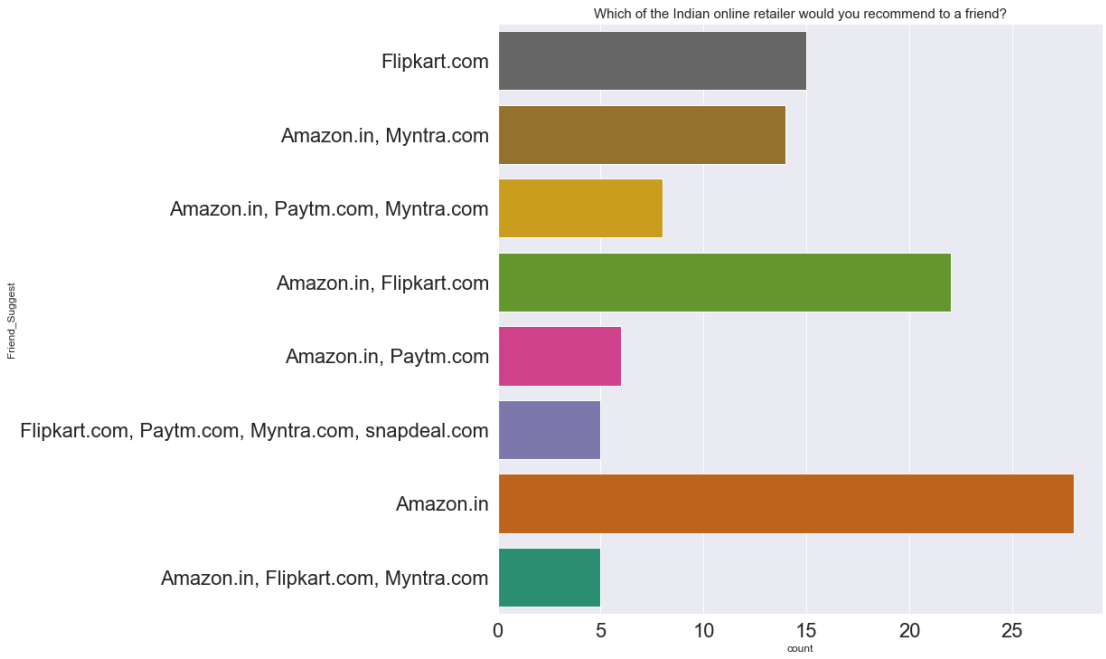

Fig.47 (Friend_Suggest)

The visualization in Fig.47 is a count plot from the sea-born library it clearly concludes that, Amazon, Flipkart.com have the highest votes when making recommendations to a friend about Indian online retailers.

CONCLUSION

- ## Key Findings and Conclusions of the Study

For an online business to be successful emphasis must be given to the factors necessary for user satisfaction. However, these factors may differ among the customers as per their preferences. From the above analysis we can conclude that Information, Trust and Service qualities prove to be the well-established factors necessary for online business users' satisfaction. Some other factors like quality of system, information, accessibility, security, and ease of use are needed to appeal to the consumer's utilitarian values. On the other hand, to appeal to the hedonistic values of the consumer, the e-store should incorporate elements that would offer value to the pleasure seeking behaviour of customers; for example: adventure, gratification, role shopping etc. The result further suggests that customers' perception of utilitarian and hedonistic values will inform their preference for a particular online store. Hence, attention to the customer specific content is very much important.

- ## Learning Outcomes of the Study in respect of Data Science

Learned a lot regarding, how various factors affect customer satisfaction and how an e-commerce company should focus on areas like service quality, Security to gain the trust of customers in order to improve customer loyalty.