



E-retail factors for customer activation and retention

Submitted by:

Aneesha B Soman

ACKNOWLEDGMENT

Acknowledgement The success and final outcome of this project required a lot of guidance and assistance from Sajid Choudhary Sir and I am Extremely fortunate to have got this all along the completion of my project work Whatever I have done is only due to such guidance and assistance and I would not forget to thank him.

I respect and thank Sajid Choudhary Sir, for giving me an opportunity to do the project work in Data Analytics and providing us all support and guidance which made me complete the project on time . I am extremely grateful to him for providing such a nice support and guidance though he had busy schedule managing the company affairs.

I have also referred to various articles in Towards Data Science and Kaggle to obtain codes on various visualisation methods.

INTRODUCTION

- Business Problem Framing

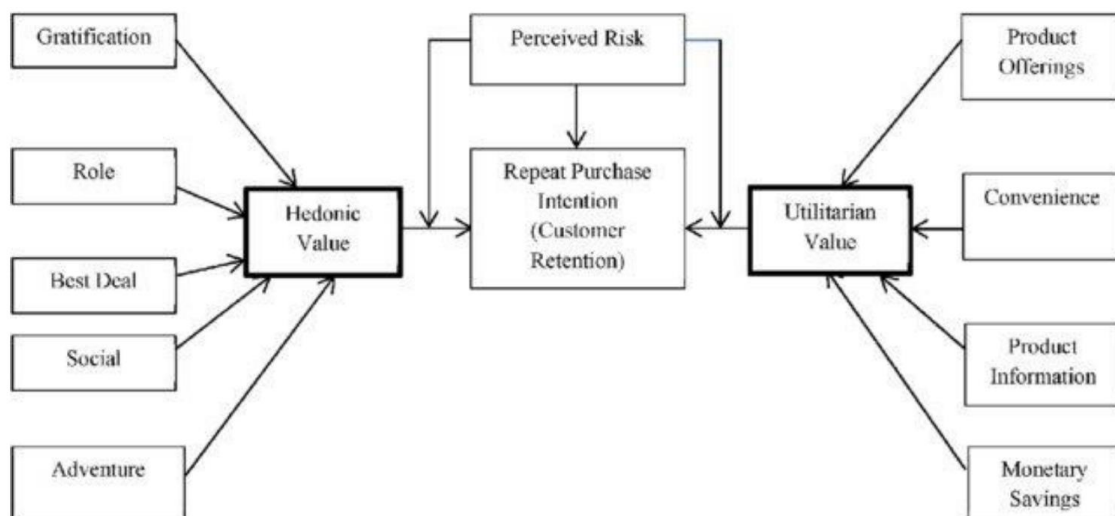
Customer satisfaction has emerged as one of the most important factors that guarantee the success of online store, hence analysing the extend to which the factors that leads to customer retention becomes significant

- Conceptual Background of the Domain Problem

Online shoppings customer retention becomes significant for the store and data needs to be taken from the sample pool of customers and analyse the factors that influenced them in choosing a website

- Review of Literature

Research has been done on a dataset which contains various parameters that can cause the success of a online store. Extensive research has been done on various parameters such as:



- Motivation for the Problem Undertaken

The objective of the projects was in understanding what all parameter can make an online store more efficient. It has been fascinating to understand the parameters like offers, as to how much less significance it can have. The domain is interesting as the current market is online. Hence through this analysis, any online store could be made better

Analytical Problem Framing

- Mathematical/ Analytical Modeling of the Problem

Visualisation of various factors have been done through matplotlib and seaborn

- Data Sources and their formats

Data set was in CSV format, it needs to be cleaned(columns mainly)

The columns contain the factors which can potentially affect the customer retention, where the factors contain how strongly the respondents agree to the factor in hand. Also some columns contains respondents analysis as to which website has performed the best in the factor

Gender	2 How old	3 Which c	4 What is	5 Since H	6 How ma	7 How do	8 Which d	9 What is	10 What i	11 What b	12 Which	13 After fi	14 How m	15 What i	16 How fr	17 Why di	18 The co	19 Inform	20
Vale	31-40 year	Delhi	110009	Above 4 yr	31-40 time	Dial-up	Desktop	Others	Window/w	Google ch	Search En	Search En	6-10 mins	E-wallets (Sometime	Lack of tru	Agree (4)	Indifferent	Inc	
Female	21-30 year	Delhi	110030	Above 4 yr	41 times a	Wi-Fi	Smartphor	4.7 inches	IOS/Mac	Google ch	Search En	Via applica	more than	Credit/Det	Very frequ	Promo coc	Strongly a	Strongly a	Sti
Female	21-30 year	Greater N	201308	3-4 years	41 times a	Mobile Int	Smartphor	5.5 inches	Android	Google ch	Search En	Via applica	11-15 mins	E-wallets (Sometime	Promo coc	Strongly a	Agree (4)	Ag	
Vale	21-30 year	Karnal	132001	3-4 years	Less than	Mobile Int	Smartphor	5.5 inches	IOS/Mac	Safari	Search En	Search En	6-10 mins	Credit/Det	Never	Better alte	Agree (4)	Agree (4)	Inc
Female	21-30 year	Bangalore	530068	2-3 years	11-20 time	Wi-Fi	Smartphor	4.7 inches	IOS/Mac	Safari	Content M	Via applica	more than	Credit/Det	Frequently	Better alte	Strongly a	Indifferent	Inc
Female	31-40 year	Noida	201308	Above 4 yr	41 times a	Wi-Fi	Smartphor	5.5 inches	Android	Google ch	Search En	Via applica	more than	E-wallets (Frequently	Better alte	Strongly a	Strongly a	Sti	
Vale	41-50 year	Delhi	110011	Above 4 yr	31-40 time	Wi-Fi	Tablet	Others	Android	Google ch	Display Ad	Direct URL	more than	Credit/Det	Sometime	Promo coc	Agree (4)	Agree (4)	Ag
Vale	41-50 year	Delhi	110018	3-4 years	Less than	Mobile Int	Desktop	Others	Window/w	Google ch	Search En	E-mail	6-10 mins	Credit/Det	Sometime	Change in	Strongly d	Dis-agree	Dis
Female	Less than	Solan	173229	2-3 years	Less than	Wi-Fi	Smartphor	5.5 inches	Android	Google ch	Search En	Direct URL	1-5 mins	Cash on d	Sometime	No prefer	Strongly a	Agree (4)	Sti
Female	31-40 year	Delhi	110039	Less than	Less than	Wi-Fi	Laptop	Others	Window/w	Google ch	Search En	Search En	more than	Cash on d	Sometime	Lack of tru	Agree (4)	Strongly a	Ag
Vale	31-40 year	Moradaba	244001	Above 4 yr	21-30 time	Mobile Int	Laptop	Others	Window/w	Google ch	Search En	Direct URL	6-10 mins	Credit/Det	Sometime	Change in	Strongly a	Indifferent	Ag
Female	41-50 year	Noida	201308	Above 4 yr	Less than	Wi-Fi	Laptop	Others	Window/w	Google ch	Search En	Direct URL	more than	Credit/Det	Sometime	Better alte	Strongly a	Strongly a	Ag
Female	51 years a	Gurgaon	122018	2-3 years	Less than	Wi-Fi	Laptop	Others	Window/w	Google ch	Search En	Direct URL	more than	Credit/Det	Sometime	Better alte	Strongly a	Strongly a	Sti
Female	31-40 year	Greater N	201310	Above 4 yr	11-20 time	Wi-Fi	Laptop	Others	Window/w	Google ch	Search En	Search En	11-15 mins	Credit/Det	Sometime	Better alte	Strongly a	Strongly a	Sti
Female	21-30 year	Karnal	132001	Less than	41 times a	Wi-Fi	Smartphor	5.5 inches	Android	Google ch	Search En	Via applica	11-15 mins	E-wallets (Sometime	Promo coc	Strongly a	Agree (4)	Ag	
Female	21-30 year	Bangalore	560037	2-3 years	41 times a	Mobile Int	Smartphor	5.5 inches	IOS/Mac	Safari	Search En	Search En	6-10 mins	Credit/Det	Never	Better alte	Agree (4)	Agree (4)	Inc
Vale	31-40 year	Noida	201308	1-2 years	21-30 time	Mobile Int	Laptop	Others	IOS/Mac	Safari	Content M	Via applica	more than	Credit/Det	Frequently	Better alte	Strongly a	Indifferent	Inc
Vale	41-50 year	Delhi	110011	Less than	Less than	Wi-Fi	Laptop	Others	Window/w	Google ch	Search En	Via applica	more than	E-wallets (Frequently	Better alte	Strongly a	Strongly a	Sti	
Female	41-50 year	Delhi	110018	1-2 years	Less than	Wi-Fi	Laptop	Others	Window/w	Google ch	Display Ad	Direct URL	more than	Credit/Det	Sometime	Promo coc	Agree (4)	Agree (4)	Ag

- Data Preprocessing Done

A.Encoding of the data done allowing the easy usage of the data

B.Column names we cleaned

- Data Inputs- Logic- Output Relationships

Data input and output are both categorical. The input is all the columns which are the factors and the output/label is the website that the respondent will refer a friend to

- the set of assumptions (if any) related to the problem under consideration

The assumption taken here are that the respondent has knowledge about Amazon, Flipkart, Snapdeal, Paytm and Myntra

- Hardware and Software Requirements and Tools Used

Hardware :PC Windows 10

Software: Jupyter notebook, Panda library, numpy library, Matplotlib library, Seaborn library

Model/s Development and Evaluation

Identification of possible problem-solving approaches (methods)

The factors need to be found which can impact the customer retention capacity. This can be done by analysing the various factors and the store the respondent prefers. This will be done by checking each of the factors impacts the respondents decision making.

Visualizations

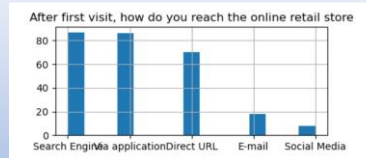
Customer activation

Methods used by respondent during customer activation



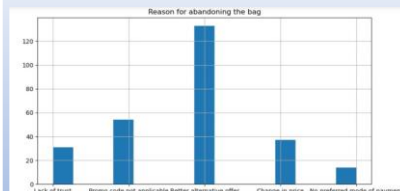
Search engine has been the primary usage during the first attempt to reach site

Method used by customer in 2nd time(Customer Retention)



In the second time, the method has search engines usage but also an equal usage of Via application

Factors hindering the customer activation

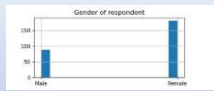


Maximum people and the highest frequency of dropping the bag is because of Better alternative

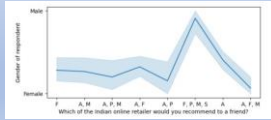


Respondents

Respondents



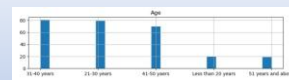
More number of female respondents. They consist of 3/4th



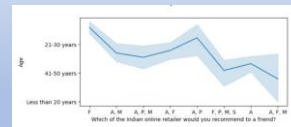
Male have chosen all flipkart, myntra, snapdeal and paytm but not amazon

Whereas women have chosen all ranges

Age



The respondents have been mostly between 21-30 years



21-30 years have not chosen snapdeal

41-50 years like all of them

Five major factors that contributed to the success of an e-commerce store have been identified as:

system quality

information quality

service quality

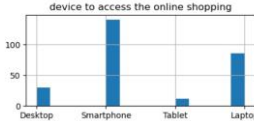
trust

Net benefit

System quality

The system quality

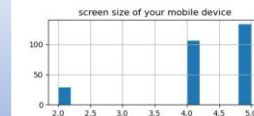
Smartphone 141
Laptop 85
Desktop 38
Tablet 12
Name: device to access the online shopping, dtype: int64



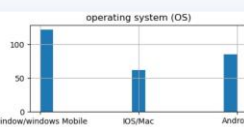
Largest number of online shopping done through Smartphone and least through tablet



Others 134
3-5 inches 99
4-7 inches 29
6-8 inches 7
Name: screen size of your mobile device, dtype: int64



Others is highest with 134 respondents for the device size



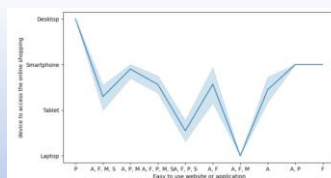
Maximum OS used are Windows



The respondent who would prefer	Device used	mobile device screen size	OS
amazon alone	smartphone	4.7 inches	windows/iOS
flipkart alone	Smartphone and tablet	not 4.7 inches/5.5 inches	windows
both amazon and flipkart	tablet	4.7 inches	iOS
both amazon and myntra	tablet	4.7 inches	iOS
amazon and paytm	smartphone	5.5 inches	android
amazon, myntra and flipkart	laptop	not 4.7 inches/5.5 inches	Windows
amazon, myntra and paytm	smartphone	5.5 inches	android

Easy to use websites and the device they used

This shows the device is most suited for the website



The respondent who would prefer	Device preferred
amazon alone	Smartphone and tablet
flipkart alone	smartphone
both amazon and flipkart	Smartphone and tablet
amazon and paytm	smartphone
amazon, myntra and flipkart	laptop
amazon, myntra and paytm	Smartphone
Paytm alone	Desktop
Amazon, myntra, flipkart and snap deal	tablet
Amazon, paytm, flipkart and snap deal	Tablet and laptop

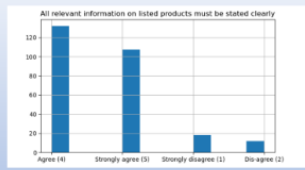
Information Quality

Content Readability

The readability of the product information is very significant as above 164 people have expressed how important it is to them. If not readable they are abandoning the bag



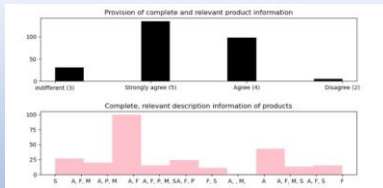
Clarity of content



The DETAILS of the product information is very significant as above 132 people have expressed how important it is to them. If details are not proper then they are abandoning the bag



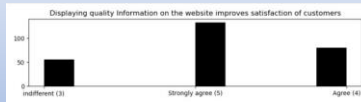
Complete, relevant description information of products websites/apps



The respondents strongly agree the significance of relevant product description. With complete description, the user friendly interface is being impacted. Flipkart gives the best description and then Amazon.



Relationship between product information and gratification

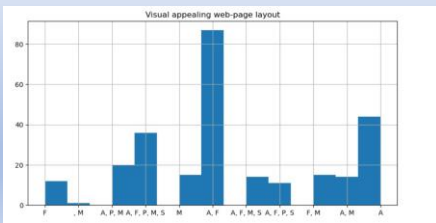


Displaying the quality of the product is very important for customer gratification

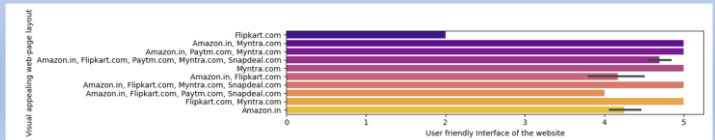


Visual appealing web-page layout

- Amazon and flipkart are having the best visual layout
- the visual appealing is impacting the User Friendly interace of the website
- it is not impacting much as a factor to be chosen again by the customer



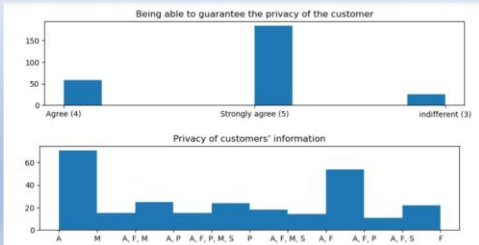
Visual appeal has less impact on recommendation to a friend



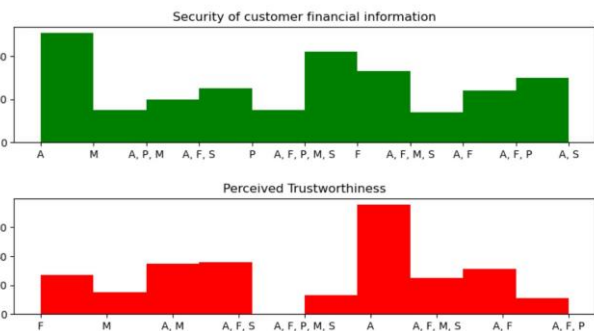
SERVICE QUALITY

PRIVACY

It is very important for an individual to have privacy and the highest privacy of customer information is seen in amazon

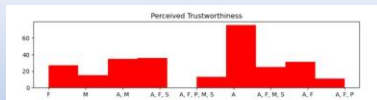


Highest privacy of customer financial information and perceived trustworthiness is also seen in amazon

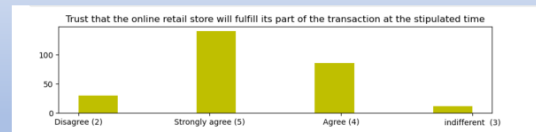
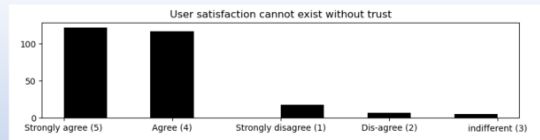
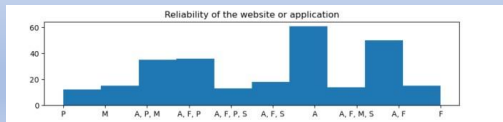


TRUST

TRUST



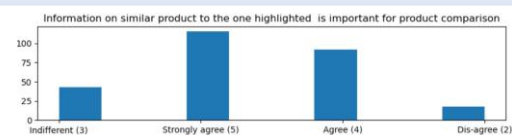
Amazon is the most trustworthy



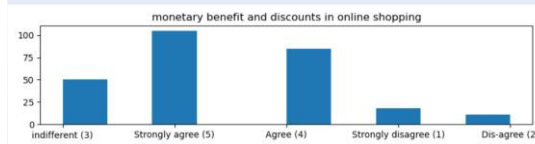
User satisfaction is directly proportional to trust that the product will be delivered in the proper amount of time

Monetary savings/Net benefit

Monetary savings/Net benefit



There is a wide proportion who does not need the product comparison and check net benefit



A large proportion of people is indifferent to the monetary benefits



Access to loyalty programs

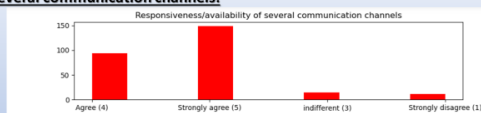
Access to loyalty programs



A large section is indifferent to loyalty programmes

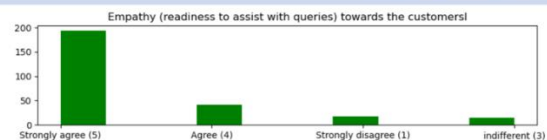
CUSTOMER CARE

Need for several communication channels:



There is a high need for several communication channels and Amazon is the best in online assistance through several communication channels

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



Interpretation from the visualisation done are:

1. Search engine has been the primary usage during the first attempt to reach site
2. In the second time, the method has search engines usage but also an equal usage of Via application
3. Maximum people and the highest frequency of dropping the bag is because of Better alternative
4. More number of female respondents.They consists 3/4th
5. Male have chosen all flipkart,myntra,snaodeal and paytm but not amazon
6. Whereas women have chosen all ranges
7. The respondents have been mostly been between 21-50years
8. 21-30years have not chosen snapdeal
9. 41-50years like all of them
10. Largest number of online shopping done through Smartphone and least through tablet
11. Others is highest with 134 respondents for the device size
12. Maximum OS used are Windows

The respondent who would prefer	Device used	mobile device screen size	OS
amazon alone	smartphone	4.7 inches	windows/IOS
flipkart alone	Smartphone and tablet	not 4.7 inches/5.5 inches	windows
both amazon and flipkart	tablet	4.7 inches	IOS
both amazon and myntra	tablet	4.7 inches	IOS
amazon and paytm	smartphone	5.5 inches	android
amazon,myntra and flipkart	laptop	not 4.7 inches/5.5 inches	Windows
amazon,myntra and paytm	smartphone	5.5 inches	android

13.

The respondent who would prefer	Device preferred
amazon alone	Smartphone and tablet
flipkart alone	smartphone
both amazon and flipkart	Smartphone and tablet
amazon and paytm	smartphone
amazon,myntra and flipkart	laptop
amazon,myntra and paytm	Smartphone
Paytm alone	Desktop
Amazon,myntra,flipkart and snap deal	tablet
Amazon,paytm,flipkart and snap deal	Tablet and laptop

14.

15. The readability of the product information is very significant as above 164 people have expressed how important it is to them. If not readable they are abandoning the bag

16. The DETAILS of the product information is very significant as above 132 people have expressed how important it is to them.If details are not proper then they are abandoning the bag
17. The respondents strongly agree the significance of relevant product description With complete description, the user friendly interface is being impacted
18. Flipkart gives the best description and then Amazon.
19. Displaying the quality of the product is very important for customer gratification
20. Amazon and flipkart are having the best visual layout
21. the visual appealing is impacting the User Friendly intercaace of the websiteit is not impacting much as a factor to be chosen again by the customer
22. It is very important for an individual to have privacy and the highest privacy of customer information is seen in amazon
23. Highest privacy of customer financial information and perceived trustworthiness is also seen in amazon
24. Amazon is the most trust worthy
25. User satisfaction is directly proportional to trust that the product will be delivered in the proper amount of time
26. There is a wide proportion who does not need the product comparison and check net benefit
27. A large proportion of people is indifferent to the monetary benefits
28. A large section is indifferent to loyalty programmes
29. There is a large section of respondents who are indifferent about shopping online and also many who strongly disagree that enjoyment is not present in shopping online
30. Return and replacement is very important for shopping online
31. Customers derive satisfaction only when the website quality is good
32. Displaying the quality of the product is very important for customer gratification
33. Mostly people dont get affected emotionally by which online store they buy from as they have no much attachment to it
34. Mostly people dont get affected emotionally by which online store they buy from as they have no much attachment to it
35. There is ease in patronizing a retailer
36. It is not an accepted idea among respondent,That online shopping increases the social status of the people are indifferent to it

37. Eventhough some respondents respond they come into the website for adventure in some small sense. But There is also a large number of respondent were indifferent to it
38. the largest section of respondent strongly agree or agree that user friendly websites impacts in taking decisions
39. Amazon ,flipkart,paytm and myntra all are easy to use websites
40. Amazon takes most amount of time to get logged in due to promotions and sales
41. Loading speed of amazon is fastest
42. Longest time in displaying graphics and photos is amazon and flipkart
43. Prices are declared late the most is myntra
44. Flexibility is very important factor for decision making
45. Amazon is the quickest for transactions
46. Delivery speed is highest for Amazon
47. It is very significant to show similar products descriptions as more than 70% have agreed to it
48. Eventhough a large group of people are drawn towards online shopping through discounts and loyalty shopping, a large group of people are indifferent to it

- **Interpretation of the Results**

The Customer retention depends on all the factors largely other than

a.Status increase through online shopping

b.discounts/offers and loyalty programmes

c.Adventure

CONCLUSION

- Key Findings and Conclusions of the Study

a.The customer retention improvement should be focussed on the hedonic and utilitarian factors(other than adventure) i.e .

Hedonic values	Utilitarian values
1.gratification 2.role 3.best deal	1.Product offering 2.convenience 3.product information

b.Also the success of the online store depends on:

- i. system quality
- ii. service quality
- iii. Net benefit
- iv. information quality
- v. trust

- Learning Outcomes of the Study in respect of Data Science

Matplotlib and Seaborn helped in making the vast data into more understandable and hence helping a person to analyse the data