

E-retail factors for customer activation and retention

Submitted by:

Shailendra Pratap Singh

**ACKNOWLEDGMENT**

Inside the Prediction of the dataset () I am guided by mine mentor (Sajid Choudhary), I also took some help from Kaggle, UCI Machine Learning Repository and YouTube in the completion of project.

**INTRODUCTION**

**1. 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. 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.

**2. Conceptual Background of the Domain Problem**

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.

**3. Review of Literature:**

(i) PREPROCESSING OF DATA  
(ii) DATA CLEANING  
(iii) DATA ANALYSIS

**4. Motivation for the Problem Undertaken**

Many companies think that marketing is all about selling of products and how the company will increase its profitability in the market share and not considering after sales service to be able to know whether the customers are satisfied with the products are able to use it.

The main problem is ‘what kind of strategies is expected from telecommunication industries to

be able to retain their customers?

The main problem is ‘what kind of strategies is expected from telecommunication industries to

be able to retain their customers?’

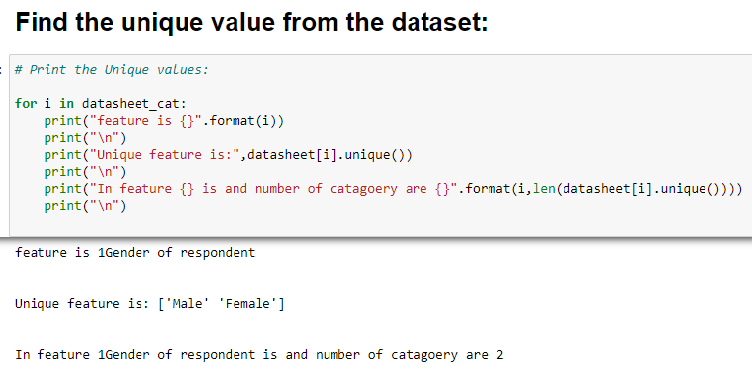
The main problem is ‘what kind of strategies is expected from telecommunication industries to

be able to retain their customers?’

**ANALYTICAL PROBLEM FRAMING**

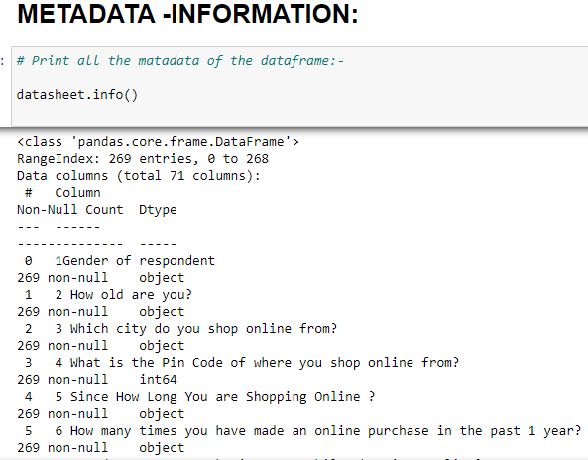
1. **Mathematical/ Analytical Modeling of the Problem**

In customer retention dataset all the feature inside any columns is in categories, so there are no Mathematical/Analytical steps are involved in the dataset analysis. We only count each unique feature



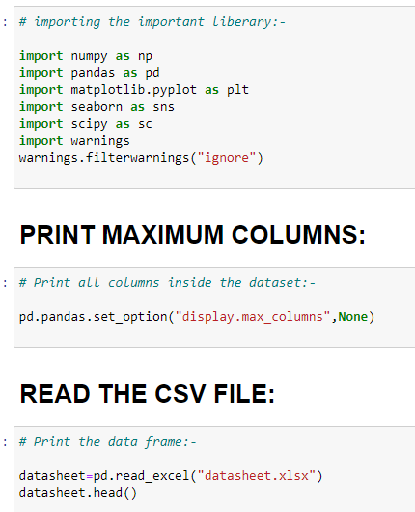
**2. Data Sources and their formats**

The source for this research was mainly primary source. The primary data gathered for this study was received through self-designed questionnaire administered to the management of different type of online retailers. The dataset is in excel format, here we got 2 kind of dataset one is customer retention(datasheet) and another one is customer retention (code sheet) both are compacted in single excel sheet we scrap one of them and made another excel sheet so that’s why both the dataset is separate and we could study anyone of them according to our convenience. Here I am working on the customer retention (datasheet). when we study about the dataset all the columns feature are in object datatype only pin code columns have int (64) datatype.

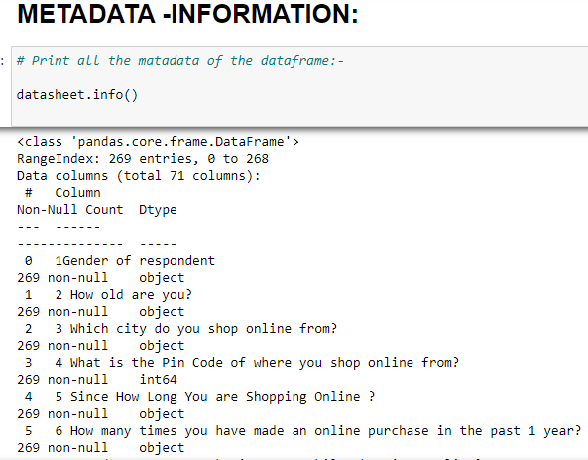


3. **Data Pre-processing Done**

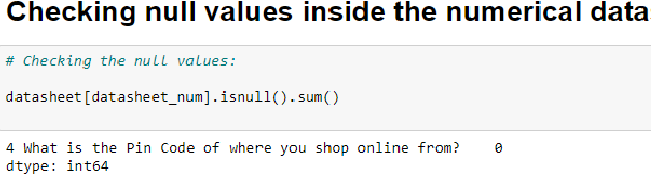
(i) Frist, we imported the important library and import the data inside the Jupiter notebook from the directory and give the read command for reading the dataset file.

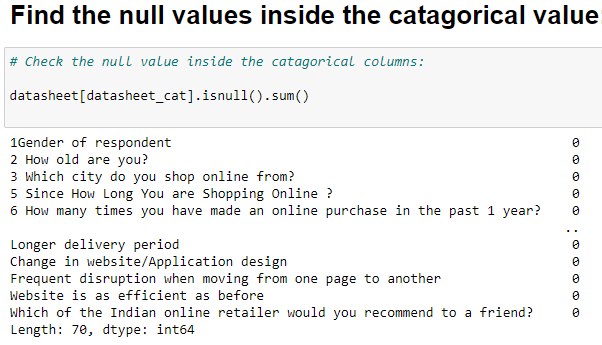


(ii) After reading the dataset we check the information of metadata from which we find all the initial information of dataset. Like our dataset have 291 row and 71 columns. All the dataset columns contain the object datatype value while pin code columns contain the int (64) datatype value.



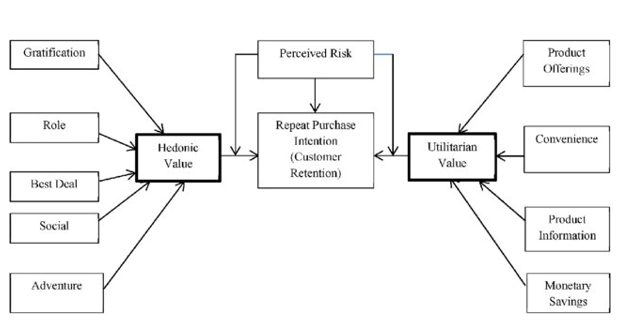
(iii) After checking the metadata, we extract the numerical variable and categorical variable separately and checking null value inside them so we get zero null value in both of the data.





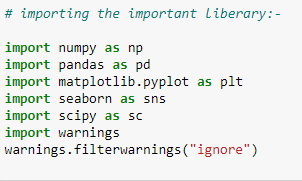
4. **Data Inputs- Logic- Output Relationships**

There is no input output is available inside the dataset here is some columns which contain hedonic value and some are contain utilitarian value both the values combined and define the quality and effectiveness of online shopping websites. **Hedonic value** is defined as that value a customer receives based on the subject experience of fun and playfulness and **Utilitarian value** is defined as that value that a customer receives based on a task-related and rational consumption behaviour.



6. **Hardware and Software Requirements and Tools Used**

The whole process is done on my laptop on jupitarnotebook, in the whole process we import different kind of library like NumPy, pandas, matplotlib, seaborn, we also import SciPy for statical analysis. For the machine model building we used skit learn.



**DATA ANALYSIS AND EVALUATION**

1. **Findings and Discussion of Data**

**TABLE 1: RESPONDENTS DISTRIBUTIONS BY GENDER**

|  |  |  |
| --- | --- | --- |
| **Gender** | **Frequency** | **Percent** |
| Male | 88 | 32.8 |
| Female | 181 | 67.2 |
| Total | 269 | 100 |

**Respondent profile:**

According toour data from 298 people 181 are female who is doing online shopping and 88 male person who is doing online shopping so from the percentage of overall ratio 67.2 percent is female and 32.8 percent is male so female is more shopped online in comparison male.

**TABLE 2:** **Age category of respondents.**

|  |  |  |
| --- | --- | --- |
| **Age** | **Frequency** | **Percentage** |
| 31-40 years | 81 | 30.1 |
| 21-30 years | 79 | 30.3 |
| 41-50 years | 70 | 26.6 |
| Less than 20 years | 20 | 7.4 |
| 51 years and above | 19 | 7.0 |

**Respondent Age:**

Inside our dataset age column there are 5 unique values in which 31-40 years (81,30.1) have highest frequency and percentage, after that 21-30 year (79,30.3) have heights frequency and percentage, after those 41-50 years (70,26.6) after than less than 20 years (20,7.4) and then 51 years and above (19,7.0) so the young people percentage is more than the elder once the percentage.

**TABEL 3: City category of respondents.**

|  |  |  |
| --- | --- | --- |
| **City** | **Frequency** | **Percentage** |
| Delhi | 58 | 21.56 |
| Greater Noida | 43 | 15.58 |
| Noida | 40 | 14.86 |
| Bangalore | 37 | 13.75 |
| Karnal | 27 | 10.03 |
| Solan | 18 | 06.69 |
| Ghaziabad | 18 | 06.69 |
| Gurgaon | 12 | 04.46 |
| Meerut | 9 | 03.34 |
| Moradabad | 5 | 01.85 |
| Bulandshahr | 2 | 00.74 |

**Respondent city:**

Delhi (58,21.56), Grater Noida (43,15.58), Noida (40,14.8) and Bangalore (37,13.75) is that place where people are more doing online shopping and Meerut (9,03.34), Moradabad (5,01.85) Bulandshahr is the place where people doing less online shopping (2,00.74).

**TABEL 4: How Long You are Shopping Online**

|  |  |  |
| --- | --- | --- |
| **Since How Long You are Shopping Online** | **Frequency** | **Percentage** |
| Above 4 years | 98 | 36.4 |
| 2-3 years | 65 | 24.1 |
| 3-4 years | 47 | 17.4 |
| Less than 1 year | 43 | 15.9 |
| 1-2 years | 16 | 05.9 |

**Respondent Since How Long You are Shopping Online:**

Science more people are doing online shopping. above 4 year (98,36.4).

**TABEL 5: Online Purchase In 1 Year.**

|  |  |  |
| --- | --- | --- |
| **Online purchase in 1 Year** | **Frequency** | **Percentage** |
| Less than 10 times | 114 | 42.37 |
| 31-40 times | 63 | 23.42 |
| 41 times and above | 47 | 17.47 |
| 11-20 times | 29 | 10.78 |
| 21-30 times | 10 | 03.71 |
| 42 times and above | 6 | 02.23 |

**Respondent Online purchase in 1 Year:**

Here we saw that most of the people are doing shopping less than 10 times (114,42.37) ,31-40 times (63,23.42),41 times and above (47,17.47),11-20 times (29,10.78) and very less people are doing shopping 21-30 times (10,03.71) and 42 times and above (6,02.23) in 1 year.

**TABEL 6:** **Easy to use website or application**

|  |  |  |
| --- | --- | --- |
| **Easy to use website or application** | **Frequency** | **Percentage** |
| Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com | 64 | 23.79 |
| Amazon.in, Flipkart.com, Myntra.com, Snapdeal.com | 44 | 16.35 |
| Amazon.in, Flipkart.com | 44 | 16.35 |
| Amazon.in | 29 | 10.78 |

|  |  |  |
| --- | --- | --- |
| Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com | 22 | 08.17 |
| Amazon.in, Paytm.com, Myntra.com | 20 | 07.43 |
| Amazon.in, Flipkart.com, Myntra.com | 19 | 07.06 |
| Paytm.com | 12 | 04.46 |
| Flipkart.com | 8 | 02.97 |
| Amazon.in, Paytm.com | 7 | 02.60 |

**Respondent Easy to use website or application:**

Most of the person like Amazon.in, Flipkart.com, Paytm.com, Myntra.com, Snapdeal.com (64,23.79).

**Data Analysis**:

The analysis of the data collected is done based on the objectives of this research. The first objective therefore, was to identify the various retention strategies adopted by many online companies.

**TABLE 7: To Identify the Retention Strategies Used by the Organization**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Question** | **Strongly**  **Agree** | | | **Agree** | | **Undecided** | | **Disagree** | | **Strongly**  **Disagree** | |
| **The**  **organization**  **retains its**  **customers by;** | **Frequency** | | **Percentage** | **Frequency** | **Percentage** | **Frequency** | **Percentage** | **Frequency** | **Percentage** | **Frequency** | **Percentage** |
| **The content on the website must be easy to read and understand;** | 164 | | 60.96 | 80 | 29.73 | 7 | 02.60 | NA | NA | 18 | 06.69 |
| **Information on similar product to the one highlighted is important for product comparison;** | 116 | | 43.12 | 92 | 34.20 | 43 | 15.98 | 18 | 06.69 | NA | NA |
| **User friendly Interface of the website;** | 189 | 70.26 | | 45 | 16.72 | 5 | 01.85 | 12 | 04.46 | 18 | 06.69 |
| **Being able to guarantee the privacy of the customer;** | 185 | 68.77 | | 58 | 21.56 | 26 | 09.66 | NA | NA | NA | NA |
| **Online shopping gives monetary benefit and discounts;** | 105 | 39.03 | | 85 | 31.59 | 50 | 18.58 | 18 | 06.69 | 11 | 04.08 |
| **Shopping online is convenient and flexible;** | 146 | 54.27 | | 78 | 28.99 | 33 | 12.26 | 12 | 04.46 | NA | NA |
| **User satisfaction cannot exist without trust;** | 122 | 45.35 | | 117 | 43.49 | 18 | 06.69 | 7 | 02.60 | 5 | 01.85 |
| **Getting value for money spent;** | 149 | 55.39 | | 82 | 30.48 | 38 | 14.12 | NA | NA | NA | NA |

**Respondent Easy to use website or application:**

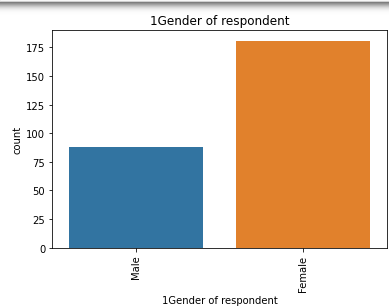
Inside our dataset there are several utilitarian values but I only select 10 feature which is responsible for making trust of customer here is some questioner which is based on customer satisfaction The content on the website must be easy to read and understand 164(60%) people is strongly agree with it and 18(6%) people is strongly disagree.Information on similar product to the one highlighted is important for product comparison 116 (43.12%) is strongly agree with it and 18 (6%) people is only disagreeUser friendly Interface of the website189 (70%) strongly agree and 18 (6%) disagree.

Being able to guarantee the privacy of the customer 185 (68.77%) and there is no people who shows the disagree NessOnline shopping gives monetary benefit and discounts.105() people is strongly agree and 11 (4%) people are disagree on that.Shopping online is convenient and flexible 146 (54.27%) people is strongly agreed and 12 people is showing disagree NessUser satisfaction cannot exist without trust 122 (46%) people is strongly agree and

5(1%) people is disagreed on that.Getting value for money spent 149 (56%) strongly agree and no people is disagreed on that.

**Visualizations**

For the visualization purpose we use several methods basically we use matplotlib and seaborn library for the visualization purpose here inside the dataset only categorical value are present so we only use count plot.



**CONCLUSION**

The study gives an insight about the opportunity as well as the challenges of retaining customers. It is noticed that companies cannot hold on to existing customers when the old strategies are being used, therefore to be able to retain customers new retention strategies should be used.

1. organizations’ total output greatly depends on existing customers

Introduce customer satisfaction rating to measure the performance of customers towards

the use of the organization’s product and services.

Introduce customer satisfaction rating to measure the performance of customers towards

the use of the organization’s product and services.

Introduce customer satisfaction rating to measure the performance of customers towards

the use of the organization’s product and services.

Introduce customer satisfaction rating to measure the performance of customers towards

the use of the organization’s product and services.

Introduce customer satisfaction rating to measure the performance of customers towards

the use of the organization’s product and services.

Introduce customer satisfaction rating to measure the performance of customers towards

the use of the organization’s product and services.

Introduce customer satisfaction rating to measure the performance of customers towards

the use of the organization’s product and services.

Introduce customer satisfaction rating to measure the performance of customers towards

the use of the organization’s product and services.

v Introduce customer satisfaction rating to measure the performance of customers towards

the use of the organization’s product and services.

Introduce customer satisfaction rating to measure the performance of customers towards

the use of the organization’s product and services.

Introduce customer satisfaction rating to measure the performance of customers towards

the use of the organiza