



# Customer Churn Analysis

---

**A Project Report by:**

**Archit Khanduja**

## Acknowledgement

The project would not have been built without the constant support from **DataTrained** and **Fliprobo** teams.

Following are the research papers, discussions and articles that helped me in completing the project:

<https://www.optimove.com/resources/learning-center/customer-churn-prediction-and-prevention>

<https://seaborn.pydata.org/>

## **Introduction**

### **Business problem Framing**

Customer churn refers to when a customer ceases his or her relationship with a company. Online businesses typically treat a customer as churned once a particular amount of time has elapsed since the customer's last interaction with the site or service. The full cost of churn includes both lost revenue and the marketing costs involved with replacing those customers with new ones. Reducing churn is a key business goal of every online business.

### **Conceptual Background of the Domain Problem**

In order to succeed at retaining customers who would otherwise abandon the business, marketers and retention experts must be able to (a) predict in advance which customers are going to churn through churn analysis and (b) know which marketing actions will have the greatest retention impact on each particular customer. Armed with this knowledge, a large proportion of customer churn can be eliminated. Given the type of dataset we have, there are only 269 samples and any Machine Learning model will not be able to generalize itself and will be prone to overfitting.

### **Motivation for the Problem Undertaken**

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

## Analytical Problem Framing

### Mathematical/ Analytical Modelling of the Problem

- The dataset contains 71 features and 269 rows.
- The dataset contains no NULL values.
- All the attributes are of 'object' type except the PIN code.

Not much data cleaning is required as the dataset already seems very clean. So we will proceed to our Analytical modelling.

### Exploratory Data Analysis

Let us see which city people shop online from first:

```
[10]: dataset['3 Which city do you shop online from?'].value_counts()

[10]: Delhi          58
      Greater Noida  43
      Noida          40
      Bangalore     37
      Karnal        27
      Ghaziabad     18
      Solan         18
      Gurgaon       12
      Merrut        9
      Moradabad     5
      Bulandshahr   2
      Name: 3 Which city do you shop online from?, dtype: int64

[11]: 58+43+40+37+18+12

[11]: 208

[12]: 208/269

[12]: 0.7732342007434945

77% of the customers belong to metropolitan cities
```

77% of the customers belong to metropolitan cities. The company should pay more emphasis on catering to the needs of these customers in order to retain them.

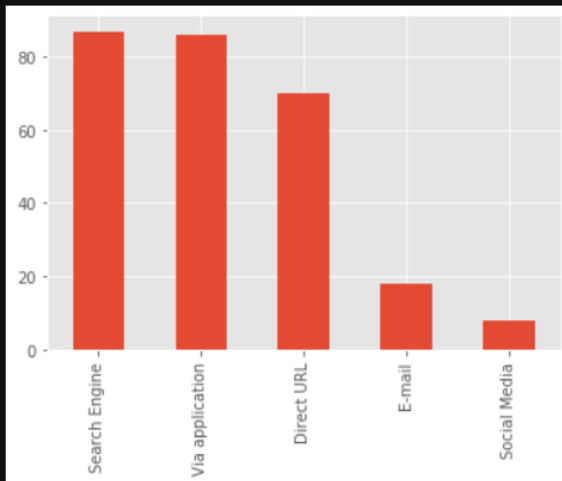
Now, there is another important aspect which needs to be explored. After the first experience of online shopping, how did the customer reach the online store. My theory is that if the customer installed the application, the customer is there with you for the long term.

```
[13]: dataset["13 After first visit, how do you reach the online retail store?"].value_counts()

[13]: Search Engine      87
      Via application    86
      Direct URL        70
      E-mail            18
      Social Media       8
      Name: 13 After first visit, how do you reach the online retail store?, dtype: int64

[14]: dataset["13 After first visit, how do you reach the online retail store?"].value_counts().plot.bar()

[14]: <matplotlib.axes._subplots.AxesSubplot at 0x23373dc4188>
```



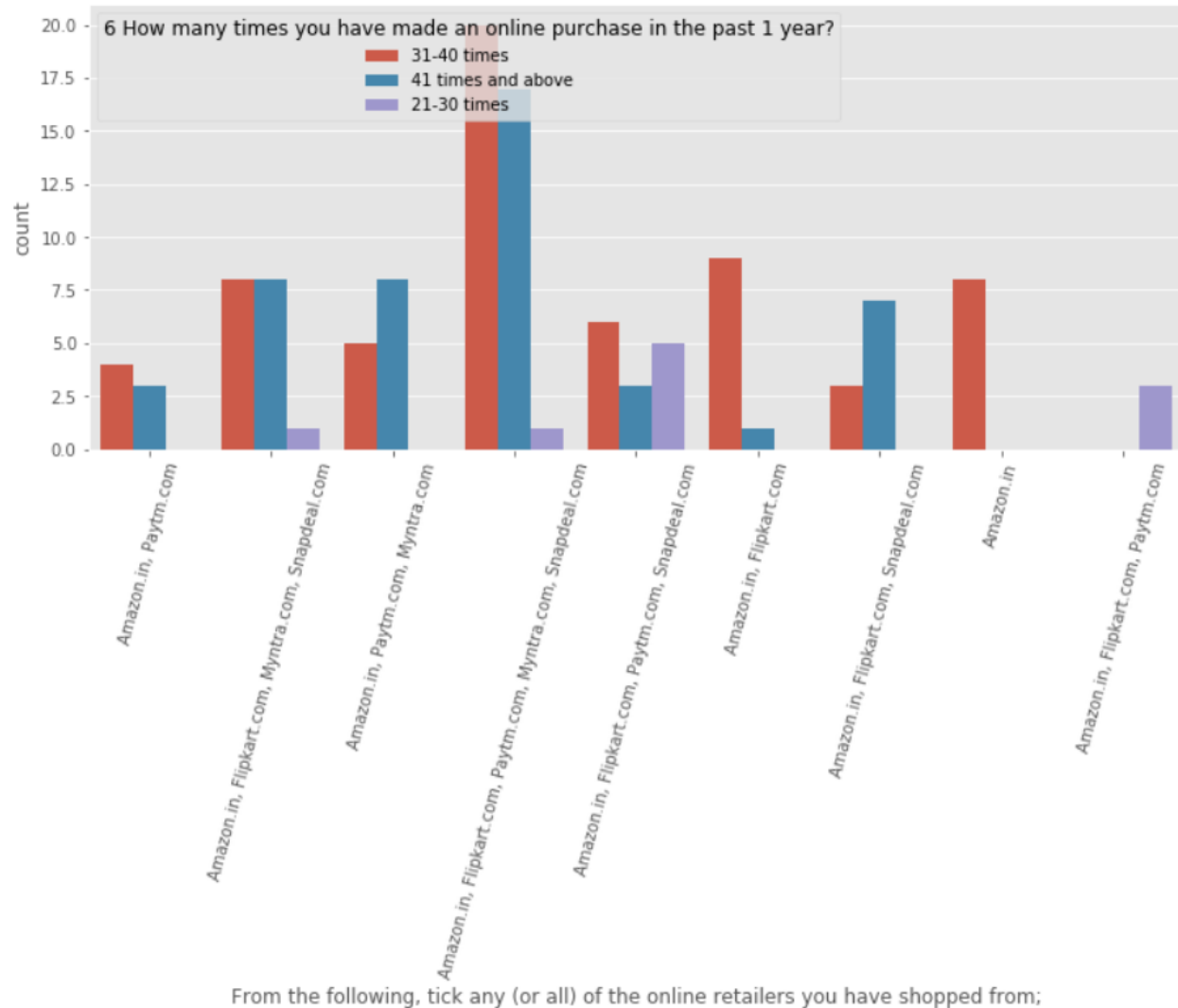
86/269 customers have downloaded the app after their first usage of the online retail store. We can safely assume that these customers have been retained by the E commerce website as they have downloaded the application on their phone and have a fair chance of shopping again. Also, there are a few customers who are using Desktop/Laptop, for which, there are no applications for online shopping and you have to stick with the web browser.

**If we are building a machine learning model, we can use this column as our target variable.**

The only challenge with this will be the users who are using desktop and cannot have an application installed in their browser and we need to create another model to predict the churn for those using the desktop and feed the output to our new model.

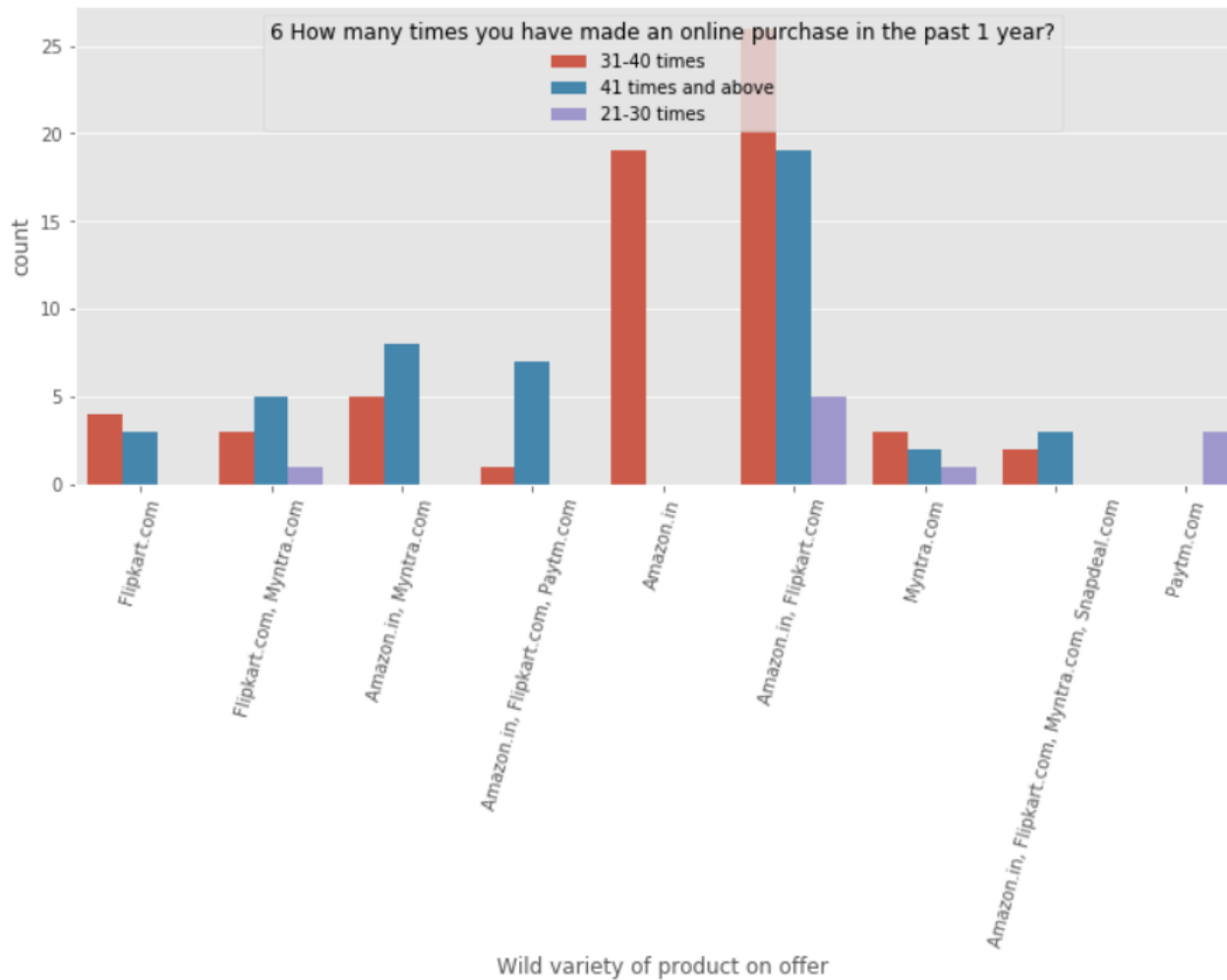
### Bivariate Analysis on the best Online platform

We will analyze our data for “How many times have you made an online purchase in the past 1 year” against other features and will analyze only for those customers who have made a purchase more than 20 times in an year.

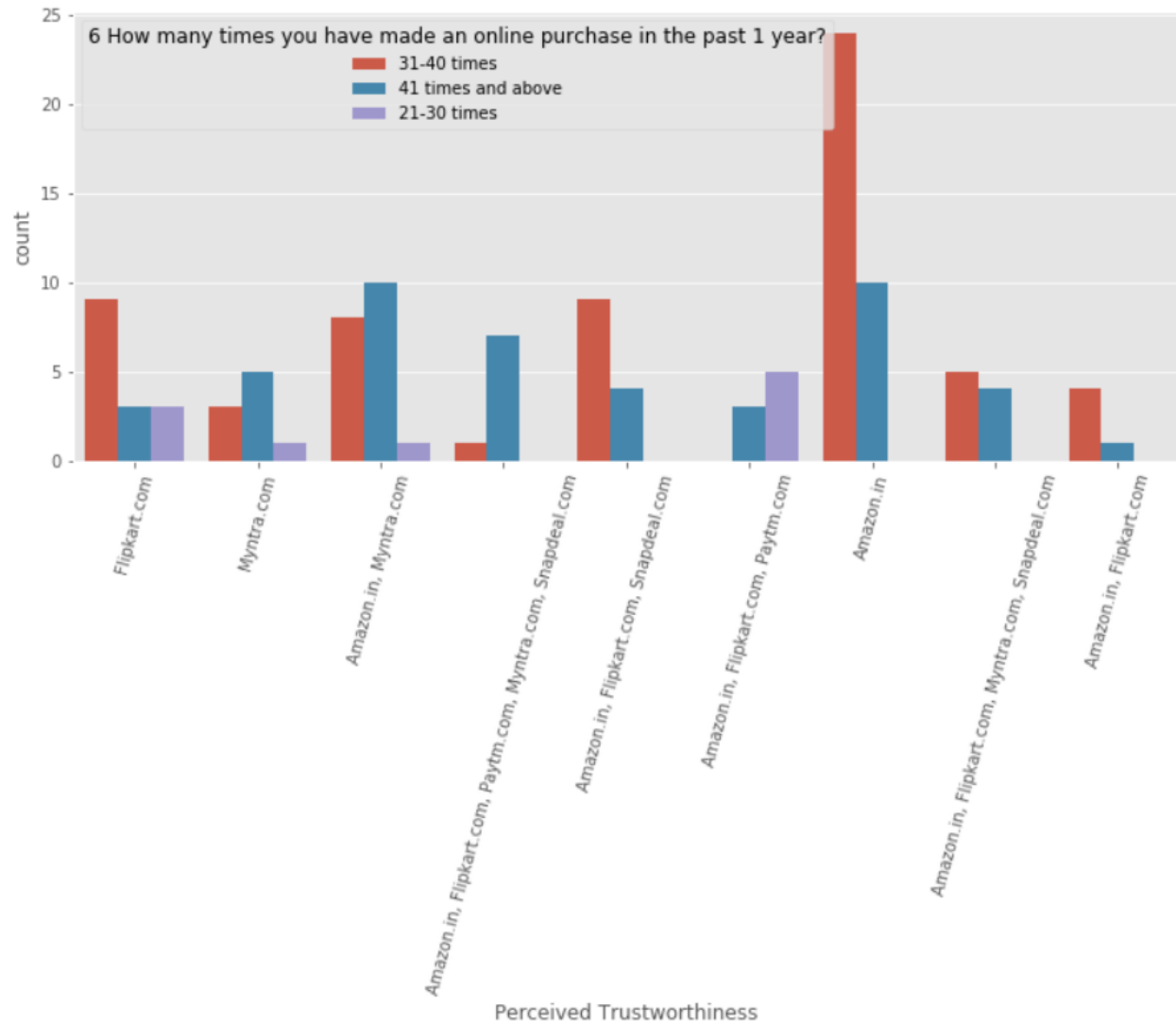


Highest number of customers have shopped from Amazon.com, Paytm and Myntra. There is no customer who has not shopped from Amazon, making Amazon the market leader in this domain.

Also, Amazon has the wildest variety of products on offer as depicted by the following graph:



Also, when it comes to the trustworthiness, Amazon again has earned a lot of brownie points and have outperformed the other companies as shown below.



### Conclusion from Bivariate Analysis

We can safely conclude that Amazon has outperformed the other portals in all the aspects of the survey and has clearly maintained its image. It has been able to retain its customers successfully as compared to the other players in the market.



## Exploratory Data Analysis on “Strongly Agree”

In our survey, we have a lot of questions asked upon Agreement on a few questions. Here, we are going to analyze which factors are the most important for the customers for choosing their perfect online retailer by counting the percentage of strongly agree and sorting the values.

```
[46]: df_strongly_agree=df_value_counts['Strongly agree (5)']

[47]: df_strongly_agree.sort_values()

[47]: 46 Shopping on the website helps you fulfill certain roles      0.141264
44 Shopping on your preferred e-tailer enhances your social status 0.178439
43 Shopping on the website gives you the sense of adventure      0.200743
42 The Convenience of patronizing the online retailer            0.200743
45 You feel gratification shopping on your favorite e-tailer     0.241636
47 Getting value for money spent                                  0.304833
31 Enjoyment is derived from shopping online                     0.319703
20 Complete information on listed seller and product being offered is important for purchase decision. 0.323420
30 Online shopping gives monetary benefit and discounts          0.390335
21 All relevant information on listed products must be stated clearly 0.397770
39 Offering a wide variety of listed product in several category 0.412639
23 Loading and processing speed                                  0.427509
34 Gaining access to loyalty programs is a benefit of shopping online 0.427509
19 Information on similar product to the one highlighted is important for product comparison 0.431227
38 User satisfaction cannot exist without trust                  0.453532
35 Displaying quality Information on the website improves satisfaction of customers 0.494424
40 Provision of complete and relevant product information        0.501859
26 Trust that the online retail store will fulfill its part of the transaction at the stipulated time 0.524164
22 Ease of navigation in website                                 0.524164
32 Shopping online is convenient and flexible                    0.542751
41 Monetary savings                                              0.550186
29 Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.) 0.553903
25 Convenient Payment methods                                    0.591078
37 Net Benefit derived from shopping online can lead to users satisfaction 0.609665
36 User derive satisfaction while shopping on a good quality website or application 0.650558
28 Being able to guarantee the privacy of the customer          0.687732
24 User friendly Interface of the website                       0.702602
27 Empathy (readiness to assist with queries) towards the customers 0.721190
33 Return and replacement policy of the e-tailer is important for purchase decision 0.736059
Name: Strongly agree (5), dtype: float64
```

Here, we can see the percentage of strongly agree in various attributes of online shopping. We have sorted this in the descending order.

A few insights from this analysis are as follows:

1. Shopping on your preferred e-tailer enhances your social status: This attribute has only 17% of the customers who strongly agree. This is because most of the customers will go to the shopping malls for the social status as ordering online doesn't usually help people in showing off their shopping bags publically.
2. The Convenience of patronizing the online retailer : This attribute has only 20% of the customers who strongly agree. This means that the customers are not much concerned about patronizing with the retailer.

Coming to some of the important attributes

3. Return and replacement policy of the e-tailer is important for purchase decision: It has whopping 73% of the customers who strongly agree to it. It is evident from the fact that people cannot actually try, touch and feel the products that they are purchasing before they reach home and they would want to return or replace in case of dissatisfaction. Online shopping websites should make strategies around easy return and replacement policy if they want to retain their customers.

4. Empathy (readiness to assist with queries) towards the customers: This also has a high percentage of 72. The reason is insecurity amongst the customers. If they have a provision of a helpline number associated with the website, that sense of insecurity vanishes and the customer can trust the website more. A 24X7 support is a must for these online retail stores.

5. User friendly Interface of the website: Just like the navigation in a physical store should be easy, same is the case with the online retail store so that the customers do not have to work around much and the overall shopping experience is smooth. The retail stores should invest heavily in creating user friendly apps and websites.

6. Being able to guarantee the privacy of the customer: This also got 68% strongly agree. Everyone is concerned of their privacy. If there is a website that posts about your recent purchase on its social media and tags you without your consent can be a breach of privacy. Imagine buying underwears and being tagged by amazon on facebook for doing so.

7. Convenient Payment methods: This also had 59% people strongly agreeing. Cash on delivery, online cards and upi, all the methods should be available for the customers to make the transaction.

8. Responsiveness, availability of several communication channels (email, online rep, twitter, phone etc.): In case one channel is not available, customers can reach out to multiple channels which again is an important factor. We have recently seen customers reaching out to Amazon and Flipkart's Twitter handles and raising issues over there which are readily resolved.

