

Retention Analysis Project Description

Problem: What drives customer retention and who are its most valuable customers?

Retention analytics is at the core of all digital B2C businesses. In many cases, it is crucial to ensure the economic feasibility of a product/service. This is especially the case for e-commerce businesses, who want the customers that they have acquired to make as many orders as possible to increase their lifetime value and drive up the ROI on their acquisition, this is the volume effect. Long-term customer retention allows for another value leverage, the price effect, as loyal customers tend to be less price-sensitive (Kanghyun & Thanh, 2011).

Because of these factors, it is interesting and important for digital marketplaces to analyze their customer retention data and identify opportunities for minimizing churn and improving their unit economics. It also allows the firm to perform a more granular customer segmentation and identify its most valuable clients. We aim to perform this analysis and answer these questions for a British e-commerce retailer.

Dataset: E-Commerce Public Dataset

<https://archive.ics.uci.edu/ml/datasets/online+retail>

This dataset, available through the UC Irvine Machine Learning website, provides the details of every order placed on an anonymous UK-based online retailer between January 2010 and September 2011. In a comma-separated value .csv table, it lists factors including price, quantity, and approximate location; additionally, unlike some other online retailer datasets, it clearly includes the customer ID on each order.

General Plan of Attack:

First, we will perform initial exploratory analysis of the dataset and classify customers as either retained or lost between 2016 and 2018. Then we will predict Customer Lifetime Value (CLTV) for Olist customers, experimenting with different models including a beta-geometric/negative binomial (BG/NBD) distribution (Fader et al., 2005). By the end of this analysis, we will have identified which customers should be retained to maximize the ROI on acquisition cost. This will position us to recommend interventions such as targeted advertisement for customers with high potential CLTV.

Group Members: Amaury Laurent apl2163; Daiyang Li dl3393; Charlotte Liao el3073; Jonathan Jasper jfj2121; Nicholas Etz npe2105