

Analyze the Effect of Customer Retention Rate on Customer Lifetime Value(CLV) - Based on the Regression Model

Yueru Jin
emmayueru.jin02@gmail.com

University of California, Berkeley, California, United States, 94709

Abstract—As companies become more profit-driven, they often ignore the importance of boosting or recovering from lost customer retention rate, which potentially leads to greater customer lifetime value and greater success. This paper first provides graphical analyses of the relationship between customer retention rate and customer lifetime value (CLV), and compares the trend with the relationship between profit per customer and CLV. In this theoretical data analysis, the author finds an exponential relationship between customer retention rate and CLV and a linear relationship between profit per customer and CLV. To test the real-world relevance of the result, linear regression analyses are performed on the real-world dataset on telecom customers. The general trend is similar, as indicated in the multivariate regression analysis: increase in customer retention rate impacts CLV more than profit per customer does. However, from the actual data, no significant relationship between profit per customer and CLV is found, but a strong linear relationship between customer retention rate and CLV exists.

Keywords-Customer retention rate, customer lifetime value, profit, linear regression, correlation

1 INTRODUCTION

Nowadays, companies are focusing more on driving customer acquisition than on retention. According to online statistics, 44% of companies have a greater focus on customer acquisition, while only 18% focus more on customer retention [1]. In addition, popular marketing methods such as user generated content advertising are veered towards customer acquisition than towards customer retention. According to research by You and Joshi, user generated content advertising is more effective for acquisition while traditional media gains an edge in retention. [2] However, this lack of focus on customer retention rate may miss potential payoff. According to research by Bain & Company, a 5% increase in customer retention rate produces more than a 25% increase in profit. [3] Therefore, there is an inefficient investment of resources for many companies in driving long-term value and success. Even worse, companies are losing their customer base at a high rate.

The goal of this paper is to model the effect of change in customer retention rate on Customer Lifetime Value and compare it with the effect of change in profit per customer. The hypothetical data will be analyzed graphically first, and then compared with the linear regression analyses on the real-world dataset. Lastly, a conclusion is reached that customer

retention rate has a greater impact on CLV than profit per customer, and that companies should focus more on business strategies that boost retention rate.

2 METHODOLOGY

This paper focuses on Customer Lifetime Value (CLV) as an indicator of a company's success. CLV shows a "relationship-oriented marketing perspective", demonstrating how "changes in customer behavior (e.g. increased purchase, retention) can influence future profitability". CLV can be used to evaluate the return on customer acquisition, retention, and the effectiveness of certain marketing strategies. The four main variables that influence CLV are profit per customer, discount rate, retention rate, and horizon of the customer relationship. For simplicity, this article assumes an infinite horizon of the customer relationship and uses the formula of $\text{profit} / (1 + \text{discount rate} - \text{retention rate})$ to calculate CLV. [4]

2.1 Hypothetical Data Modeling

The author is interested in studying and comparing the effect of profit and retention rate on CLV. To eliminate the effect of confounding variables, the two relationships are first modeled with hypothetical data.

The hypothetical data used are profit = 10, discount rate = 0.1, and retention rate = 0.8. To ease the calculating process of floating numbers, the CLV is rounded to the hundredth (Table 1).

Table 1 Original Hypothetical Data

Profit	10
Discount rate	0.1
Retention rate	0.8

2.2 The relationship between profit and CLV

To discover the effect of retention rate on CLV, the author assumes that profit and discount rate stays constant (profit = 10, discount rate = 0.1), and plots the retention rate of 100 data with an interval of [0%,99%] against the corresponding calculated CLV (as shown in Table 2 and figure 1).

Table 2 Hypothetical Data with Retention Rate from 0%-99%

Profit	10
Discount rate	0.1
Retention rate	0%~99%

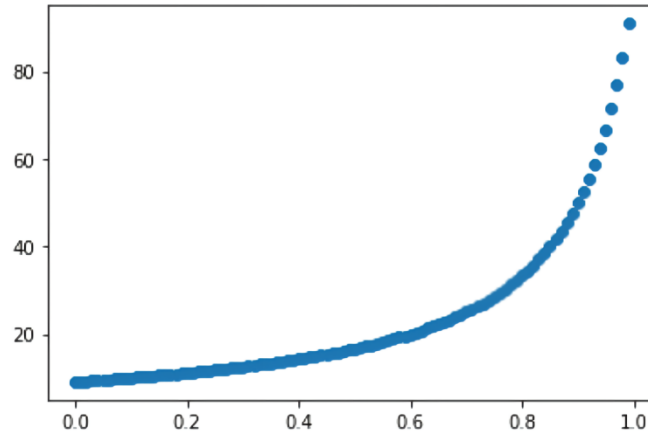


Figure 1 Graph of Hypothetical Data with Retention Rate from 0%-99%

Then a graph of exponential growth is generated, which increases at a lower rate from 0%-60% and more swiftly above 60%. Above a retention rate of 80%, a small increase in retention rate would result in a huge increase in CLV. For instance, an increase in retention rate from 80% to 85% would lead to a 20% increase in CLV (from 33.33 to 40) (Table 3).

Table 3 Hypothetical Data with Profit from 1 to 100

Profit	1~100
Discount rate	0.1
Retention rate	0.8

2.3 The relationship between retention rate and CLV

Next, to discuss the effect of profit on CLV, the author plots a scatterplot of profit from 1 to 100 on the x-axis and CLV on the y-axis. The author gets a linear graph this time, meaning that a unit increase in retention rate always results in the same amount of increase in CLV. For example, an 100% increase in profit from 10 to 20 would result in an equally 100% increase in CLV from 33.33 to 66.67 (Figure 2).

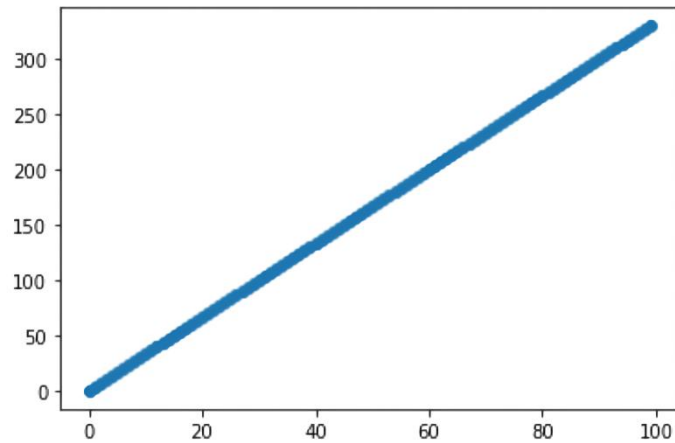


Figure 2 Graph of Hypothetical Data with Profit from 1 to 100

2.4 Comparison and Implication

From the two graphs and analyses, it can be seen that the customer retention rate has higher growth potential in boosting CLV than profit, especially when companies have reached the critical point of 60%. Therefore, investing in business strategies that drive retention rates may bring more value to the companies, especially if they already have a retention rate above 60%.

In reality, it is common to see a drop in the retention rate for companies. If the retention rate drops, how much of an increase in profit would companies need to make up for the loss in CLV? The following table study four such scenarios (Table 4).

Table 4 Percentage Increase in Profit needed for a percentage decrease in Retention Rate (r_rate stands for customer retention rate)

Old r_rate	75%	80%	85%	90%
New r_rate	70%	75%	80%	85%
Old profit	10	10	10	10
New profit	11.43	11.67	12.00	12.5
%increase of profit	14.28%	16.66%	20%	25%
%decrease of r_rate	6.67%	6.25%	5.88%	5.56%

From the table, it can be seen that the greater the old retention rate, the greater percentage increase of profit is needed to offset the decrease in retention rate. When the old retention rate is 75%, the percentage decrease in retention rate was nearly 50% of the percentage increase of profit. As the old retention rate goes up to 90%, the percentage decrease in retention rate went down to nearly 25% of percentage increase of profit. This demonstrates two things: first, a small decrease in retention rate is usually made up by a large increase in profit. Second, the higher the original retention rate, the bigger the gap between percentage decrease of retention rate and percentage increase of profit is.

In fact, the boosting retention rate can not only boost CLV, also provide companies with a stable source of income. Unlike newly acquired customers who are more cautious of their purchase, “loyal customers increase their spending at an increasing rate, purchase at a full margin rather than at discount prices, and create operating efficiencies”.[5] In addition, it is often less costly to retain than acquire customers. It is generally inefficient for companies to spend a lot on massive-scale marketing research and market to people who are not familiar with their products, because the conversion rate may be very low. Moreover, the startup costs incurred through customer acquisition could only be recovered by several years of relationship building with them. However, if companies could make use of the data obtained from their existing users, they would enhance their purchasing experience effectively at a lower cost.

3 REAL-WORLD DATASET MODELING

To prove the effectiveness of the model, the author uses the actual dataset to test it. The dataset used in this paper is from the website Kaggle, which originally consists of 100 variables and approximately 100,000 records of telecom customer data.[6] To compare the trend across states, the author groups the record of each individual customer by their state, and calculates the average of the variables that are studied for each state. This article is mainly interested in two variables in the dataset: profit per customer and retention rate. In this dataset, the average total charge per customer is estimated as the average profit per customer, since telecom companies often have a large amount of fixed costs but very little variable cost. The author also gets retention rate by calculating average churn rate and subtracting it from 1. For simplicity, the author assumes the discount rate in this CLV model to be a constant 10%. After getting those three variables, the CLV formula can be applied and be used to calculate the average CLV value for each state.

3.1 Linear Regression Analysis

To investigate the relationship between each of the two variables and CLV and see if they align with the result of the hypothetical data modeling, the author conducts a linear regression analysis. In the first regression model, the average profit per customer is plotted as the independent variable and average CLV as the dependent variable.

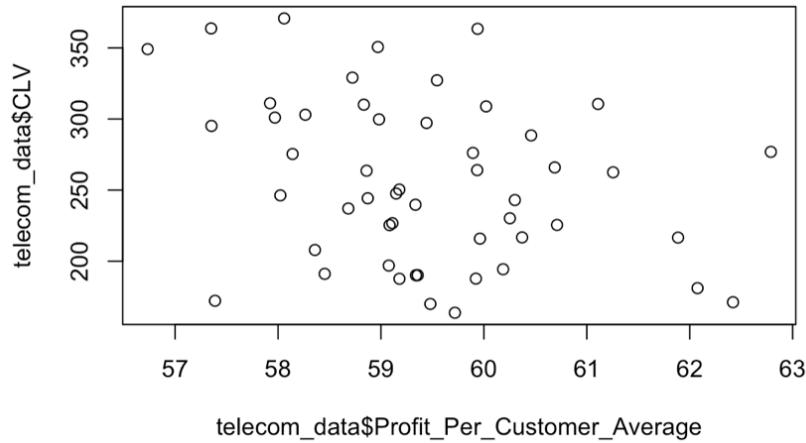


Figure 3 The Scatterplot of profit per customer against average CLV

From the scatterplot and linear regression model analysis, it can be concluded that there is no significant linear relationship between average profit per customer and CLV (Figure 3). The t-values of the independent and dependent variables are 3.005 and -2.288, separately, and the p-values are not as significant. Besides, the multiple r-squared value is only 0.0965, which means that linear regression does not fit the data well. This runs contrary to the previous hypothetical data graphing, which shows a positive linear relationship and between profit and CLV. This could be caused by confounding factors or because the independent variable data are not as spread out, with a range of only 57 to 63.

In the second regression model, the goal of the study was to plot the average retention rate as the independent variable and CLV as the dependent variable.

From the graph, we can see a positive linear relationship between average retention rate and average CLV. The relationship is very significant, with t-values of -20.75 and 30.04 for the intercept and average retention rate, separately. The r-squared value is as high as 0.9485. It is interesting to notice that the exponential relationship is very weak, unlike what is predicted in the hypothetical model, in which the curve is concave up and the second derivative of the graph is positive, especially above a retention rate of 60% (Figure 4). This indicates that the real effect of retention rate on CLV might be positive but relatively constant.

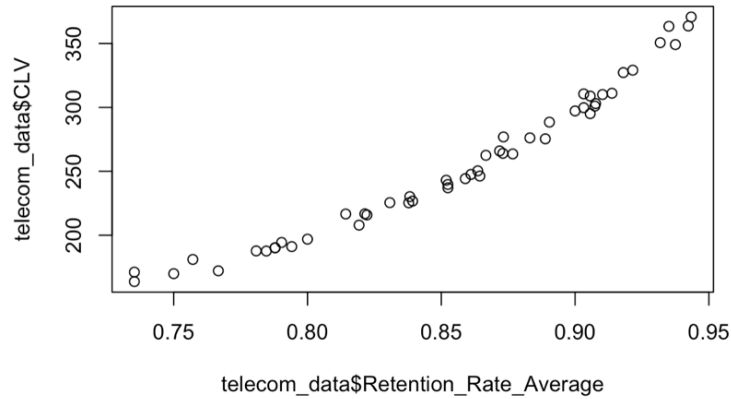


Figure 4 The Scatterplot of average customer retention rate against average CLV

3.2 Multivariate Regression Analysis

Next, to investigate the extent of the effect that the targeted two variables have on CLV, a multivariate regression analysis was performed. The summary shows that a unit increase in retention rate drives up CLV by 9.87 units, while a unit increase in profit per customer only increases CLV by nearly 1.99 units. Thus, it can be concluded that retention rate has a greater impact than profit per customer on boosting the value of CLV. The relationship is also very significant as indicated by a r-squared of 0.9503 (Figure 5).

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-706.7570	104.6236	-6.755	1.75e-08 ***
Retention_Rate_Average	9.8716	0.3439	28.706	< 2e-16 ***
Profit_Per_Customer_Average	1.9888	1.5207	1.308	0.197

Figure 5 The summary of multivariate regression of Retention Rate and Profit per Customer Against CLV

4 CONCLUSION

To compare the effect of retention rate and profit per customer on Customer Lifetime Value, this article first graphs the relationship using a set of hypothetical data, which has a span of 1%-100% for customer retention rate and 1-100 for profit per customer. The hypothetical data is direct in showing the relationship without the distraction of confounding variables. From the graphs, it can be seen that there is a linear relationship between profit per customer and customer lifetime value, and an exponential relationship between customer retention rate and customer lifetime value, which indicates that customer retention rate has a higher growth potential than profit per customer. It can also be seen from the scenario analysis that a small decrease in customer retention rate is usually made up by a large increase in profit, and the gap is larger if the original retention rate is higher.

To test the conclusions better, the next step is to perform linear regression analyses on a real-world telecom customer dataset. Unlike the hypothetical data, the author failed to get a relationship between profit per customer and CLV. However, the author got a strong linear relationship instead of an exponential relationship between customer retention rate and CLV. This could be due to the limited range of data or confounding variables. The multivariate regression analysis confirms the result from hypothetical data, though, that customer retention rate has a greater effect on CLV than profit per customer. Thus, it can be concluded that companies should be focusing more on business strategies that boost customer retention rates.

Acknowledgement. I want to show my sincere appreciation to Prof. Zhang from MIT for the marketing models and business analytics methods she taught me, which inspires me to start this research.

REFERENCES

- [1] Gagan, (2021), Customer Acquisition Statistics and Trends, Visually, <https://visual.ly/community/Infographics/business/customer-acquisition-statistics-and-trends>
- [2] Ya You & Amit M. Joshi (2020) The Impact of User-Generated Content and Traditional Media on Customer Acquisition and Retention, *Journal of Advertising*, 49:3, 213-233, DOI: 10.1080/00913367.2020.1740631
- [3] Reichheld, F. (2001). Prescription for cutting costs. Harvard Business School Publishing. https://media.bain.com/Images/BB_Prescription_cutting_costs.pdf
- [4] Fripp G, Customer Lifetime Value, The Complete Guide to CLV, <https://www.clv-calculator.com/existing-customer-clv-formula/>
- [5] Nasir, S. (2017). Customer Retention Strategies and Customer Loyalty. In *Advertising and Branding: Concepts, Methodologies, Tools, and Applications* (pp. 1178-1201). IGI Global. https://www.researchgate.net/publication/317001664_Customer_Retention_Strategies_and_Customer_Loyalty
- [6] Abhinav, 2017, Telecom Customer Churn Prediction, <https://www.kaggle.com/abhinav89/telecom-customer>