OPEN ACCESS

e-ISSN: 2177-5184

https://doi.org/10.5585/2025.29842

Check for updates

Targeting profitable customers: a clv-driven resource allocation approach

Guilherme Augusto Rizzo [a1] **Guilherme Trez** [b2]

Universidade do Vale do Rio dos Sinos (UNISINOS), Programa de Pós-graduação / Escola: Programa de Pós-Graduação em Gestão e Negócios, Departamento: Unidade Acadêmica de Pesquisa e Pós-Graduação, Porto Alegre, RS, Brasil

Universidade do Vale do Rio dos Sinos (UNISINOS), Departamento: Unidade Acadêmica de Pesquisa e Pós-Graduação, Programa de Pós-graduação / Escola: Programa de Pós-Graduação em Gestão e Negócios, Porto Alegre, RS, Brasil

Editor-in-Chefe: André Torres Urdan
Scientific Editor: Leonardo Vils
Evaluation Process: Double Blind Review

Received: 02 Nov. 2023
Approved: 16 Oct. 2025

Authors' notes
Conflict of interest: The authors have not declared any potential conflicts of interest
Corresponding author: Guilherme Augusto Rizzo - guilhermerizzo.pr@gmail.com

Cite as – American Psychological Association (APA)

Augusto Rizzo, G., & Trez, G. Targeting profitable customers: a clv-driven resource allocation aproach. *Brazilian Journal of Marketing*, São Paulo, *24*(4), p. 1-28, e29842. https://doi.org/10.5585/2025.29842

Abstract

Purpose: This study aimed to develop an analytical framework for customer acquisition based on customer lifetime value (CLV). The framework seeks to identify the most profitable customer profiles and provide a replicable model for other companies, to guide strategic marketing resource allocation decisions.
Methodology/Approach: The model was developed by combining econometric analysis with predictive modeling, specifically linear regression and the Cox proportional hazards model. The proposed framework integrates well-established methods from the literature and is designed to be extensible, allowing the incorporation of new techniques and technologies.
Main results: The results confirmed the effectiveness of the CLV-based framework in identifying the most profitable profiles among new customers, highlighting significant differences in projected profitability across the segments analyzed. The framework proved to be a strategic tool for allocating marketing resources.
Theoretical/methodological contributions: This study contributes to the literature by integrating econometric and predictive models into a CLV-based customer-acquisition framework, a topic that remains relatively underexplored. The framework can be adapted by modifying CLV measurement methods, enabling its application in companies with varied operational configurations.
Relevance/originality: This work presents an original framework for customer acquisition, distinguishing itself from prior models focused on active customers. Based on the CLV, the proposal is applicable across industries, providing a practical systematic framework that combines predictive analysis with strategic allocation of resources, strengthening decision making in competitive markets.

Keywords: customer lifetime value, prospect lifetime value, customer segmentation, cox proportional hazards model, linear regression, customer selection.

[a] Master's degree in Management and Business from Universidade do Vale do Rio dos Sinos (UNISINOS, 2023) and Université de Poitiers, France (2023). Holds a specialization in Commercial Management from FGV (2017) and a Bachelor's degree in Chemical Engineering from UFRGS (2013). Professional with solid experience in strategic management positions in large companies, acting in the areas of administration, strategy, marketing, and commercial management. http://lattes.cnpq.br/8859531867086806 - guilhermerizzo.pr@gmail.com
[b] PhD in Business Administration from Universidade Federal do Rio Grande do Sul (UFRGS, 2009). Professor in the Graduate Program in Management and Business (PPGN, Professional Master's and Doctorate) and in the Graduate Program in Administration (PPGAdm, Academic Master's and Doctorate) at Universidade do Vale do Rio dos Sinos (UNISINOS). Academic and professional experience in marketing, marketing strategy, customer value management, market orientation, strategic management, and value-based healthcare management. http://lattes.cnpq.br/1546118902885558

Resumo

Objetivo do estudo: Este trabalho teve por objetivo desenvolver uma sistemática de análise para a prospecção de novos clientes baseada no *Customer Lifetime Value* (CLV). A proposta busca identificar os perfis de clientes mais rentáveis e oferecer um modelo replicável para outras empresas, contribuindo para decisões estratégicas de alocação de recursos de marketing.
Metodologia/abordagem: A sistemática foi desenvolvida utilizando uma abordagem econométrica combinada com técnicas preditivas, incluindo modelos de regressão linear e de riscos proporcionais de Cox. O *framework* proposto integra métodos consolidados na literatura e permite a futura incorporação de novas tecnologias e métodos.
Principais resultados: Os resultados confirmaram a eficácia da sistemática com base no CLV para identificar os perfis dos novos clientes mais rentáveis, destacando diferenças significativas na rentabilidade projetada entre os segmentos analisados. O *framework* demonstrou ser uma ferramenta estratégica para direcionar recursos de marketing.
Contribuições teóricas/metodológicas: Este estudo contribui para a literatura ao integrar diferentes modelos em um *framework* voltado para a prospecção de clientes com base no CLV, um tema que não foi extensivamente explorado na literatura. O *framework* pode ser adaptado com alterações nos métodos de mensuração do CLV, possibilitando sua aplicação em empresas com distintas configurações operacionais.
Relevância/originalidade: Este trabalho apresenta um *framework* original para a captação de novos clientes, diferenciando-se de modelos anteriores focados em clientes ativos. Com base no CLV, a proposta é aplicável a diferentes indústrias, fornecendo uma sistemática prática que combina análise preditiva com alocação estratégica de recursos, fortalecendo a tomada de decisão em mercados competitivos.

Palavras-chave: Customer Lifetime Value; valor vitalício do cliente; *prospect lifetime value*; segmentação de clientes; modelo de riscos proporcionais de cox; regressão linear; seleção de clientes.

1 Introduction

In recent years, customer lifetime value (CLV) has established itself as a central metric in organizational strategies, with recent studies highlighting its positive impact on companies' financial performance (Ali & Shabn, 2024). The search for greater precision in measuring CLV has led to the adoption of innovative methodologies, such as machine learning and artificial intelligence, which allow identifying more profitable customers and optimizing marketing resources (Venkatakrishna et al., 2021; Cowan et al., 2023). However, there remains a significant underexplored opportunity to broaden research into the use of CLV as a strategic metric for customer acquisition.

A customer-centric marketing strategy that seeks to increase customer equity is critical to firms’ long-term success (Rust et al., 2004). Investments in marketing intelligence, data, and specific marketing metrics such as customer lifetime value (CLV) can provide a strong competitive advantage, resulting in superior market performance over the competition (Verhoef & Lemon, 2013).

Studies have shown that not all customers are equally profitable for companies, and thus it may be desirable to allocate different resources to different groups of customers and discontinue some business relationships (Gupta et al., 2006; Rust et al., 2004). The CLV presents itself as a metric capable of identifying the most profitable customers, in such a way that it directs the allocation of financial resources, seeking a greater return on invested capital (Gupta et al., 2006; Kumar & Reinartz, 2006).

This work aimed to develop a CLV-based framework for allocating marketing resources in customer acquisition. It seeks to identify the most profitable customer profiles and offer a replicable model for other companies, contributing to strategic marketing resource allocation decisions. The company analyzed in this research operates in the Brazilian energy market, in a contractual and continuous configuration with its customers. The "RStudio" software was used for the development of the models and the statistical treatment of the data.

To develop the investment allocation analysis system, the cost of acquiring new customers was considered, and the prospect lifetime value (PLV) metric was used (Jain & Singh, 2002). The results showed that there is a significant difference between the projected profitability of the different segments in which the analyzed company operates, allowing to direct marketing resources and to seek greater profitability.

In this work, a framework that contributes to the understanding of the use of CLV is proposed, guiding decisions for the effective allocation of investments, reinforcing the importance of a portfolio view between the different customer profiles, balancing the risk and return of the different segments. This research reinforces the idea that the analysis and selection of the different customer profiles and segments in which a company operates directly impacts the customer lifetime value, as well as the customer equity, the return on invested marketing resources and, consequently, the value of the entire company.

While the existing literature broadly addresses customer retention and segmentation based on CLV (Gupta et al., 2006; Mosaddegh et al., 2021), there is a significant gap in the application of CLV to guide potential customer acquisition decisions in the context of companies that operate with ongoing contractual relationships. This study is based on the logic of the theory-in-use (TIU) (Zeithaml et al, 2020), as a practical approach that seeks to offer contributions to the advancement of knowledge through perspectives based on the relationship between context, actions and results. The framework uses econometric approaches that can be combined with new methodologies, such as machine learning and big data, expanding the possibilities of using CLV for strategic customer acquisition decisions (Venkatakrishna et al., 2021; Cowan et al., 2023).

In addition to this introduction, this article is organized into five main sections. Section 2 presents a literature review on customer lifetime value (CLV), highlighting its applications and limitations. In section 3, we detail the methodology adopted, including the description of the data and models used to develop the proposed framework. Section 4 presents the logic of the developed framework. Section 5 discusses the results obtained from the data analysis. Finally, Section 6 shows the conclusions of the study, its main theoretical and practical contributions as well as the managerial implications, limitations, and suggestions for future research.

2 Customer Lifetime Value

Customer lifetime value can be defined as the present value of all expected future profits generated by a customer over the duration of the relationship. This definition accounts for the time value of money by discounting cash flows at an appropriate, firm- and country-specific rate, similar to the discounted cash flow approach in finance. It also incorporates customer lifetime by reflecting the probability of churn or switching to competitors over time (Gupta et al., 2006).

The fundamental model for calculating the CLV presented by Gupta et al. (2006) evaluates this metric at an individual level for each costumer and considers the customer acquisition cost, allowing the calculation of the CLV for current and future customers of the company. The CLV in this case can be calculated using the following equation (Gupta et al., 2004; Gupta et al., 2006; Reinartz & Kumar, 2003):

$$CLV = \sum_{t=0}^{T} \frac{(p_t - c_t) r_t}{(1+i)^t} - AC \quad (1)$$

Where: pt = price paid by a consumer at time t; ct = direct cost of serving the customer at time t; i = interest rate or cost of capital to the company; rt = probability of the customer repeating the purchase or being “alive” at the time t; AC = acquisition cost (to measure a customer's residual CLV, this value must be disregarded); and T = time horizon used to estimate the CLV.

Among the most widespread models in literature, CLV econometric models use economic theory to explain consumer behavior based on empirical data and evidence, such as *logit*, *probit* and *hazard* models (Kumar & Shah, 2015). Generally, these studies model three components: customer acquisition, customer retention, and margin expansion. After this measurement, these components are combined to estimate the CLV (Gupta et al., 2006).

Since not all CLV models explicitly incorporate customer acquisition cost, Singh and Jain (2013) proposed prospect lifetime value (PLV) to value prospective customers, a metric that incorporates all components of CLV while net of customer acquisition cost(CAC) . In this way, CLV models that omit acquisition cost can still be applied to new customers.

Venkatesan e Kumar (2004) evaluated the application of CLV as a metric for customer selection and marketing resource allocation. They demonstrated that customers who are selected based on CLV provide higher profits in future periods than those selected by other traditional customer metrics. Malthouse e Blattberg (2005) studied the predictive capacity of CLV models in different industries with different applications. Linear regression and neural network models were used, and one of the linear regression models tested was superior in the comparative analysis.

Donkers et al. (2007) studied the application of CLV econometric models in the insurance industry, a typically contractual and lost-for-good setting. To determine CLV, the authors separated the modeling into two steps: the first with the objective of estimating customer retention and the second with the objective of determining revenues and margin expansion resulting from purchasing behavior. Simple models performed well when compared to existing more complex models. The profit regression models, average retention model, retention based on monetary values and the “tobit II” model had the highest accuracy for predicting CLV.

Recent machine learning and big data integrations have expanded CLV prediction possibilities (Venkatakrishna et al., 2021; Ali & Shabn, 2024). These approaches offer valuable insights that complement traditional models, enabling greater accuracy in identifying profitable customers.

3 Methodology

To develop a resource allocation framework to target the most profitable customer segments in a contractual setting, we first assessed which CLV models from the literature were suitable for this purpose. Further investigations of the market and of the focus company were carried out to understand the variables that can influence the profitability and behavior of its customers, while seeking, in the literature, compatible and adherent models to predict the future profits of customers in this commercial relationship configuration.

The framework is grounded in the Theories-in-Use (TIU) approach, aimed at developing theories relevant and applicable to marketing. According to Zeithaml et al. (2020), TIU is a practical approach that seeks to develop new and organic theories from the perspectives of stakeholders close to the problem. The work was developed based on data from an organization and its structuring allows the identification of new ways of using CLV to solve strategic problems for marketing managers. This context reinforces the contribution of the study by proposing a model that addresses a gap very little explored until now: the use of CLV for prospecting new customers.

Once the models for measuring the CLV were defined, the characteristics of the market and the company were understood and the prerequisites of the database were established, we selected a software environment to implement the models. We used the RStudio IDE for R, an open integrated development environment designed for the R programming language. The objective was to facilitate the statistical analysis, modeling, calculations and graphical manipulations (RStudio, 2022). After applying the models in the database, it was possible to evaluate model fit and predictive performance, as well as to assess the application of the models to direct investments in the different customer profiles, enabling the evaluation of the company's prospecting strategy.

The study was conducted in collaboration with a Brazilian energy company operating under ongoing contractual and continuous relationships. The company consented to participate and provided supplementary data under a confidentiality and data-sharing agreement; all customer-level data were anonymized and handled in compliance with the LGPD, and commercially sensitive information was safeguarded given the market’s highly competitive nature.

In this work, data from the company's commercial transactions at an individual customer level were accessed and used, which allowed a bottom-up model of CLV to be built (Kumar, 2008). The data made available referred to one of the company's regional branches and were organized on a monthly basis, from 2016 to 2021. The databases were made available in spreadsheets organized by annual period and consolidated for use in the models developed in RStudio. For each of the components of the CLV model, a database was prepared. The variables selected to compose the database used for the models were defined considering the literature (Donkers et al., 2007; Gupta et al., 2006; Malthouse & Blattberg, 2005; Venkatesan & Kumar, 2004).

The database used contained the following costumer information at an individual level: revenue, sales price, sales volume, expected volume when the customer was acquired, raw-material cost, gross margin, expected gross margin at the time when customer was acquired, industry (based on the customer's economic activity), consumption start date, consumption end date, sale period (month and year) and investments made at acquisition. In addition to the customer level information, internal operating cost data (fixed and variable) were provided, as well as marketing and sales expenses included in the customer acquisition cost. The WACC (weighted average cost of capital) was also provided.

To understand the requirements of the model, it is essential to understand the business dynamics of the sector. The development of a client presupposes an initial investment in equipment for the client, made by the company, signing a continuous commercialization contract and establishing minimum consumption parameters, exclusively and for a determined period. This contract aims to guarantee the return on investment made by the company. This commercial relationship between the company and the customer is configured as lost-for-good (the customer is either active or is lost forever) (Jackson, 1985), since they cannot consume from another supplier. It was also possible to classify this operation as contractual and continuous, as sales happen only when the customer consumes, which occurs at different intervals throughout the contract (Kumar & Shah, 2015).

Additionally, within the company's business model, customers are segmented by economic activity and sales volume. This information characterizes the form of relationship with customers, as well as the analysis of financial viability, based on volume and price projections, which guides the decision on investments in new customers, evaluating the Net Present Value (NPV), internal rate of return and payback period of the project, without differences for the different segments of customers served.

Differences in customer profile and behaviors are also characteristic of the industry. Accounting for this when proposing the model can yield important contributions, bringing theory closer to practice and encouraging the adoption of CLV to support strategic decisions. Some segments tend to have a longer customer lifetimes and lower propensity to churn, while others show higher churn probability and bankruptcy risk. Therefore, the expected cash flows from these customers do not always materialize as projected in the feasibility analysis carried out for investments in new customers.

Despite the contractually defined time horizon, there were cases in which customers terminated the contract in the first period, before the investments were returned to the company. Regarding to margins, the contracts stipulate adjustments that affect customer prices. Therefore, margin is not constant throughout the commercial relationship. Furthermore, an econometric approach was selected for this research, modeling the retention and margin expansion components, as presented by Donkers et al. (2007), Gupta et al. (2006), Malthouse and Blattberg (2005) as well as Venkatesan and Kumar (2004). A bottom-up model, which evaluates customers individual level, was also chosen (Kumar, 2008), since the objective of this study is to analyze investment in new customers based on the CLV.

To calculate the retention component (i.e., the probability that the customer remains active), the proportional hazards (PH) model was selected. Gupta et al. (2006) point out that proportional hazards models are a widely used class of duration models. To this end, initially, a survival analysis was elaborated using the Kaplan-Meier estimator and, subsequently, the COX proportional hazards model was used in order to assess the impact of other variables on the probability of the client being active or lost in future periods. The choice of these models is also due to the fact that the database provided by the company contains type I censorship on the right (Colosimo & Giolo, 2006). In other words, at the end of the analyzed period, not all clients in the study had been lost.

As Rich et al. (2010) suggest, the Kaplan-Meier estimator can be used to compare the survival curves of different groups of interest over time. In the study in question, the curves of the different segments of economic activity were compared. The Kaplan-Meier estimator can be defined by:

$$\hat{S}(t) = \prod_{j\,:\,t_j < t} \left(\frac{n_j - d_j}{n_j} \right) = \prod_{j\,:\,t_j < t} \left(1 - \frac{d_j}{n_j} \right) \quad (2)$$

Where: S(t) is the survival function; nj the number of individuals at risk in tj, that is, the individuals who did not fail and were not censored until the instant immediately prior to tj; dj the number of failures in tj , j = 1, ... , k; t1 < t2 ... < tk, the distinct and ordered k times of failure (Kaplan & Meier, 1992).

The proportional hazards model or Cox regression allows modeling of the covariate-adjusted survival function using censored data (Cox, 1972). The formula can be described by H(t,X1,X2, …,Xk) as in the equation below (Kleinbaum & Klein, 2012):

$$H(t,X) = H_o(t)^{\sum_{i=o}^{p} \beta_i X_i} + \varepsilon \quad (3)$$

Where Ho (t) is the base risk function, which does not involve the explanatory variables X and which is configured as the parametric component of the model; t is the time or period; β are the parameters associated with the covariates; X are the explanatory variables, independent of time and ε the associated error. In this research, the explanatory variables considered were the estimated customer consumption at the closing of the contract, the unit margin at the closing of the contract, the investment required for each customer and the different segments of economic activity. To determine the explanatory variables of the final proportional hazards model, the p-value of each of them was evaluated (Ferreira & Patino, 2015).

To calculate the profit regression model, or margin growth over time, the linear regression model was used, via the ordinary least squares (OLS) method, using as a dependent variable the future margin and as independent variables the margins of the previous period, as well as the volume (quantity acquired), the segment of economic activity, and how long the customer has been with the company. Thus, the interaction between these variables was evaluated.

The models were adjusted considering the data available between 2016 and 2020 and, subsequently, their accuracy was measured in the 2021 data, which allowed us to compare the best model for the application. For each model, a Receiver Operator Characteristic (ROC) curve was created, according to which the percentage accuracy (percentage of error in the validation sample) was analyzed for different cutoff points (Lopes et al., 2014). The percentage accuracy of each observation i in the validation sample (2021) was calculated as follows:

$$ACC = \frac{|Observed(i) - Predicted(i)|}{Observed\ (i)} x100 \ (4)$$

Where Observed(i) represents the value observed in the validation sample and Predicted(i) represents the value predicted by the built model.

The mean percentage accuracy was also calculated from the root mean squared percentage error (RMSPE), described in equation 5 (Göçken et al., 2016), following the same model comparison procedure used by Donkers et al. (2007), hereafter referred to below as AccPercentual.

$$AccPercentual = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(100\ x\frac{|Observed(i) - Predicted(i)|}{Observed(i)}\right)^2} \ (5)$$

Where Observed(i) represents the value observed in the validation sample and Predicted(i) represents the value predicted by the built model. Therefore, the lower the RMSPE (AccPercentual), the better the model approximates the predicted data to the data observed in the validation sample.

Having defined the margin expansion model and the model for the probability of a customer being active, we combined them to calculate CLV. Because the margin expansion model produced a unit margin basis, it was multiplied by each customer's expected purchase volume to obtain the expected profit in Brazilian reais. To obtain each customer's net profit, this expected margin (MarginT) was subtracted by the unit cost to serve, multiplied by the expected volume, as described in Equation 6.

$$CLVi = \sum_{t=o}^{T}\frac{Vol\ x\ (MarginT - Costs)\ x\ Prob\ Active\%(t)}{(1+d)^t} \ (6).$$

Finally, MarginT was obtained through the profit expansion model. The Prob Active%(t) (Active Probability %) was estimated using Cox proportional hazards model. For Vol (volume), we used the expected consumption volume for the new customer. Costs refer to the expenses to serve a new customer and were provided by the company, calculated at the individual level of each customer. The discount rate d is the weighted average cost of capital (WACC), also provided by the company. The period T considered was five years, the standard contract length signed with new customers, and time t was measured in years.

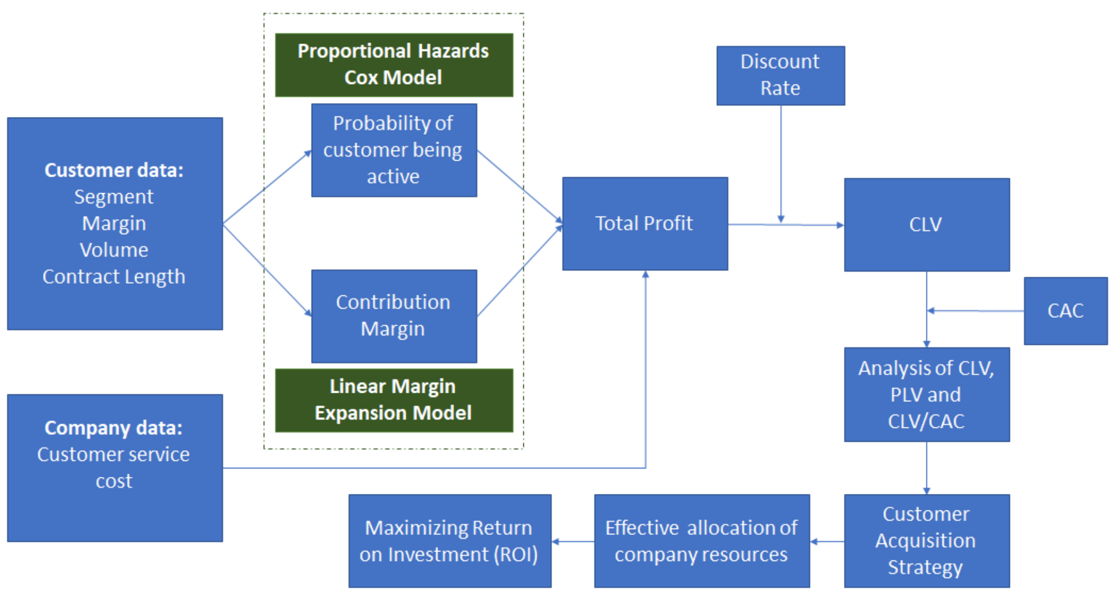
As described by Singh and Jain (2013), customer acquisition cost (CAC) was considered for investment allocation decision. Two metrics were analyzed, PLV, defined as the CLV minus CAC and the CLV-to-CAC ratio (CLV/CAC) wich compares expected CLV to realized acquisition cost.

4 Implementation Framework

Figure 1 presents an implementation framework that summarizes the systematic analysis of the company's resources allocation to new customers in order to maximize the return on capital invested in customer acquisition, focusing on the most profitable profiles and segments, using CLV as the central metric.

Figure 1

Applied Framework for Customer Acquisition Based on CLV



Source: prepared by the authors.

Based on customer and company data, the Cox proportional hazards models, used to determine the probability of the customer being active, was combined with a linear regression with interaction terms, selected as the best performing model to estimate future customer margin in this application. Thus, it was possible to estimate the expected total profit of a new customer which, when discounted at the company's discount rate, results in the CLV. Based on the customer acquisition cost, the allocation of investment across different profiles can be compared, directing the company's prospecting strategy and resulting in an effective allocation of its resources and, finally, maximizing the return on invested capital.

While Venkatesan and Kumar (2004) presented a resource allocation model focused on already active customers, the framework proposed in this work differs by concentrating efforts on prospecting new customers. This approach is new in the literature because it integrates different models to assess the CLV of potential customers, offering a practical metric to guide strategic acquisition decisions.

The proposed framework allows integration with advanced technologies to predict the probability of customer retention and margin expansion. Recent studies (Venkatakrishna et al., 2021) demonstrate how machine learning algorithms can improve CLV accuracy, suggesting paths for future extensions of the proposed model.

5 Analysis and Discussion of the Results

With the previously defined models and the processed dataset, implementation was carried out in RStudio Software following these steps: 1 Kaplan-Meier estimator; 2 Cox proportional hazards model; 3 Margin expansion model; and 4 Calculation and analysis of CLV, PLV and CLV/CAC.

5.1 Kaplan-Meier Estimator

The Kaplan-Meier estimator was used to separate the different segments of economic activity of the customers, and the survival curves were generated based on data from 2016 to 2021. Customers who remained active after this period were considered censored, since the number of years that they remained active was not observed. Data from survival curves can be seen in Table 1.

Table 1

Survival Over Time by Kaplan-Meier Estimator

Segment	*1st year*	*2nd year*	*3rd year*	*4th year*	*5th year*
Segment A	86%	81%	76%	69%	58%
Segment B	91%	77%	67%	58%	48%
Segment C	95%	82%	74%	69%	65%
Segment D	92%	80%	69%	61%	54%
Segment E	99%	93%	88%	83%	82%

Source: Prepared by the authors.

Using the Kaplan-Meier estimator, it was possible to identify significant differences across segments in their survival curves. Segments E and C were most likely to be active both at the end of the first year and at the end of the last period analyzed, respectively. Segment A had the lowest probability of being active at the end of the first year, but segment B had the lowest probability of being active at the end of the analyzed period. Segment D had the second highest propensity to churn at the end of the fifth year.

5.2 Cox Proportional Hazards Model

In the Cox proportional hazards model, in addition to the segment dependent covariate, new covariates were included to analyze their effect on the customer survival curve. The dependent covariates were: segment, customer's initial margin (mg_orcada_rton), expected consumption volume (vol_orcado_ton) and investment in the customer (investment). The data are presented in Table 2.

Table 1

Initial Cox Proportional Hazards Model

Parameter	Estimate	exp(Estimate)	Standard error	Z-value	P-value
Segment B	0.27	1.31	0.25	1.10	0.27
Segment C	-0.18	0.83	0.27	-0.70	0.49
Segment D	0.16	1.17	0.27	0.57	0.57
Segment E	-0.98	0.37	0.27	-3.69	<0.01
vol_orçado_ton	-0.01	0.99	0.02	-0.38	0.70
mg_orçada_rton	0.00	1.00	0.00	3.35	<0.01
investment	0.00	1.00	0.00	0.44	0.66

Source: Prepared by the authors.

Based on the p-values, we observed that only the covariates mg_orcada_rton and segment were statistically significant, (p-value<0.05). Therefore, the null hypothesis for these covariates were rejected, and these variables were retained for a new model that is presented in Table 3.

Table 2

Final Cox Proportional Hazards Model

Parameter	Estimate	exp(Estimate)	Standard error	Z-value	P-value
Segment B	0.263	1.301	0.248	1.063	0.28
Segment C	-0.196	0.822	0.264	-0.740	0.45
Segment D	0.151	1.163	0.272	0.553	0.57
Segment E	-0.995	0.370	0.265	-3.752	<0.01
mg_orçada_rton	0.000	1.000	0.000	3.391	<0.01

Source: Prepared by the authors.

From the final model, the baseline cumulative hazard function Ho(t) was obtained, which represents the baseline cumulative risk of loss up to each period (year) and is presented in Table 4. The survival probability for a client with given covariates (segment and mg_orcada_rton) is calculated from this function.

Table 3

Ho(t) of the Final Cox Proportional Hazards Model

t	H0(t)
1	0.0721855813
2	0.2023166245
3	0.3205721060
4	0.4330755921
5	0.5427747408

Source: Prepared by the authors.

Through the Cox proportional hazards model and the baseline cumulative hazard function, Ho(t), the equations for a customer's survival curve at the end of year t are presented with an initial margin m for each of the segments, which will be called S(t|Segment,m), as follows:

$$S(t|A,m) = H_o(t)^{(0{,}0002602325\ x\ m)} + \varepsilon \ (7)$$

$$S(t|B,m) = H_o(t)^{(0{,}2633256967 +\ 0{,}0002602325\ x\ m)} + \varepsilon \ (8)$$

$$S(t|C,m) = H_o(t)^{(-0{,}1957358913 +\ 0{,}0002602325\ x\ m)} + \varepsilon \ (9)$$

$$S(t|D,m) = H_o(t)^{(0{,}1508499033 +\ 0{,}0002602325\ x\ m)} + \varepsilon \ (10)$$

$$S(t|E,m) = H_o(t)^{(-0{,}9949215867 +\ 0{,}0002602325\ x\ m)} + \varepsilon \ (11)$$

Using the Cox proportional hazards model, we confirmed that, in addition to the economic activity segment, the margin of a new client also affects its survival curve, with an inverse effect, that is, the higher the margin, or the price charged to this client at the initial moment of the commercial relationship, the higher the probability of customer loss with each passing year, demonstrating a relationship between risk and expected return.

Segment B had the highest propensity to *churn*. Segments E and C had the lowest loss propensities. An effect of margin on the probability of customer loss was not expected, and, for this company, there was no prior view regarding the churn propensity of segments A and D.

5.3 Margin Expansion Model

The previously defined models were implemented in R Software to evaluate their performance, using linear regression, with and without interactions, via the ordinary least squares (OLS) method. The models were trained using data from 2016 to 2020, and 2021 data were used to assess model performance.

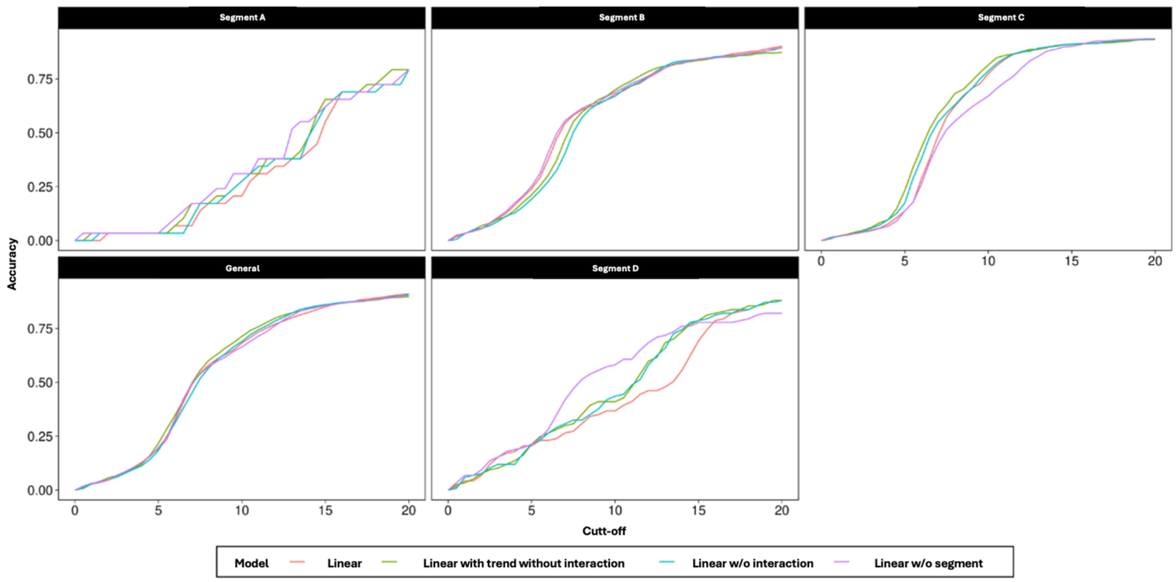
Four models were selected to analyze and compare the results, as shown below:

- Linear Model without Segment: In this model, the dependent variable was the future margin, and the independent variable was the margin of the previous year.
- Linear Model without Interaction: The dependent variable was the future margin, and the independent variables were the margin of the previous year and the segment.
- Linear Model with Trend without Interaction: The dependent variable was the margin of the following year, and the independent variables were the margin of the previous year, the segment and how long the customer has been with the company.
- Linear Model (with Interaction): The dependent variable was the future margin, and the independent variables were the margin of the previous year and the segment, including the interaction between the independent variables margin of the previous year and the segment.

The ROC curves were drawn for each of the models, evaluating the criterion of percentage accuracy, defined as ACC and presented in equation 4, for cut-offs that ranged from 5% to 20%, which are presented in Figure 2. By analyzing the ROC curves of the models, it was possible to verify that the models had a different performance across the analyzed segments. However, this criterion alone was not sufficient to determine the best model for use, since, in the overall analysis of the data, for all segments, all models had similar accuracy for the analyzed ranges.

Figure 2

ROC Curves for Percentage Accuracy



Source: Prepared by the authors.

The root mean squared percentage error (RMSPE) was then calculated and is presented in equation 5 in the section related to the method used (3 - Methodology) for each model, allowing the comparison of the models and the identification of the one with the lowest RMSPE, i.e, the smallest difference between observed and predicted values. The RMSPE and percentage accuracy (ACC) results, presented alongside the ROC curves, are detailed in Table 5 for the data of all segments combined, since the objective is to predict margin expansion for all segments of interest.

Table 4

RMSPE and Percentage Accuracy (ACC) of Margin Expansion Models at Each Cut-off

SEGMENT	MODEL	CUT-OFF	RMSPE	ACC
General	Linear	5.00	19.19	0.22
General	Linear	10.00	19.19	0.68
General	Linear	15.00	19.19	0.85
General	Linear	20.00	19.19	0.91
General	Linear w/o interaction	5.00	19.22	0.22
General	Linear w/o interaction	10.00	19.22	0.68
General	Linear w/o interaction	15.00	19.22	0.86
General	Linear w/o interaction	20.00	19.22	0.91
General	Linear w/o segment	5.00	19.32	0.20
General	Linear w/o segment	10.00	19.32	0.66
General	Linear w/o segment	15.00	19.32	0.86
General	Linear w/o segment	20.00	19.32	0.90
General	Linear with trend without interaction	5.00	19.39	0.24
General	Linear with trend without interaction	10.00	19.39	0.70
General	Linear with trend without interaction	15.00	19.39	0.86
General	Linear with trend without interaction	20.00	19.39	0.90

Source: Prepared by the authors.

The Linear model (with Interaction) was the model that presented the lowest RMSPE. The percentage accuracies, as shown in the ROC curves, presented similar performance. Thus, the Linear Interaction model was chosen as the margin expansion model to compose the CLV calculation. This model is presented in equation 12:

$$MarginT+1_i = \begin{cases} \beta_0 + \beta_1 MarginT_i \quad + \varepsilon_t \text{ ,} if\ segment\ A \\ (\beta_0+\beta_2)+(\beta_1+\beta_3)MarginT_i \quad + \varepsilon_t,\ if\ B \\ (\beta_0+\beta_4)+(\beta_1+\beta_5)MarginT_i \quad + \varepsilon_t,\ if\ \ C \\ (\beta_0+\beta_6)+(\beta_1+\beta_7)MarginT_i \quad + \varepsilon_t,\ if\ \ D \\ (\beta_0+\beta_8)+(\beta_1+\beta_9)MarginT_i \quad + \varepsilon_t,\ if\ \ E \end{cases} \quad (12)$$

Table 6 shows the coefficients of this model:

Table 6

Coefficients of the Linear Model (with Interaction)

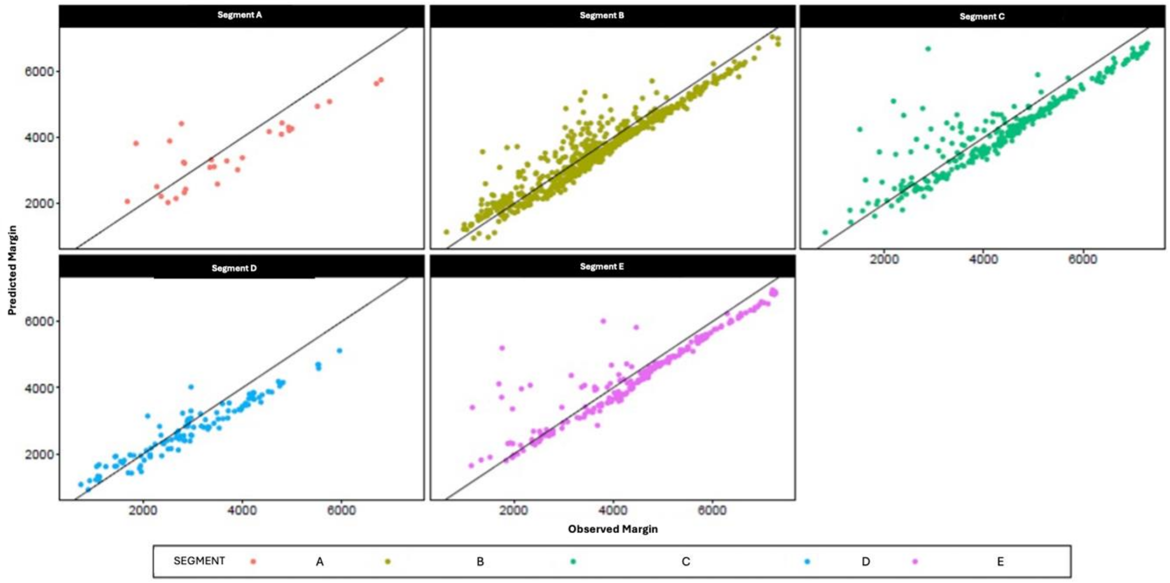
Coefficient	Estimate
(Intercept)	785.56
MarginT	0.94
SegmentB	-118.04
SegmentC	59.35
SegmentD	-152.27
SegmentE	158.26
MarginT:SegmentC	0.01
MarginT:SegmentD	-0.03
MarginT:SegmentE	0.01

Source: Prepared by the authors.

Figure 3 shows the data predicted by the model compared to the observed data. When evaluating the coefficients of the Linear model (with Interaction), it was possible to verify that, with their different intercepts and curve slopes, the segments will have a greater growth in their margins in the following order: E, C, B, A and D.

Figure 3

Observed x Predicted Margin by the Linear Model (with Interaction)



Source: Prepared by the authors.

5.4 Calculation and Analysis of Customer Lifetime Value, Prospect Lifetime Value and Customer Lifetime Value/Cost of Customer Acquisition

To demonstrate how the company can use the proposed model, a scenario was generated to analyze how customers with the same volume, the same initial price and the same acquisition cost across each of the different segments would behave and what future profits would be the obtained by the company. The CLV was calculated for this scenario within the previously established models.

As previously defined, CLV was calculated by equation 6. The input data considered were: estimated volume of 12 tons per year, an initial margin of R$1,450.00 per ton, a customer acquisition cost (CAC) of R$15,000.00, a discount rate (WACC) of 11% per year, unit costs of R$500 per ton, an average inflation for cost correction of 8% per year, and a five year analyzed period. The Prob Active(%) variable is obtained from the Cox proportional hazards model and the MarginT variable is obtained by linear regression with interaction. The WACC discount rate was provided by the company, but its composition was not detailed due to the confidentiality requested, since the market is highly concentrated.

For the customer acquisition cost, the value refers to the average investment made in customers of this size by the company, covering marketing and sales expenses. The values were obtained based on the investments made during 2021. The PLV was calculated by subtracting the CAC from the CLV, and the CLV/CAC was calculated by dividing the CLV by the customer acquisition cost. Table 7 shows the projected results of CLV, PLV and CLV/CAC for the comparison of the different segments.

Table 7

CLV, PLV and CLV/CAC Projected for the Different Segments

	CLV	PLV	CLV/CAC
Segment A	R$55,591.92	R$40,591.92	3.7
Segment B	R$48,500.01	R$33,500.01	3.2
Segment C	R$63,615.25	R$48,615.25	4.2
Segment D	R$43,199.72	R$28,199.72	2.9
Segment E	R$85,590.34	R$70,590.34	5.7

Source: Prepared by the authors.

Considering that we used the same volume, the same initial price and the same customer acquisition cost, the main difference between the calculations presented in Table 7 is related to the probability of the customer being active and the evolution of margins over time. Thus, it was possible to identify that there are two special segments that bring superior returns to the company. Segments C and E, which are most likely to be active over time, according to the Cox proportional hazards model, and show greater margin growth over time according to the linear regression model with interaction. Segment D, despite not having the highest propensity to churn, has the lowest growth in margins over time. Thus, it endes up presenting the lowest CLV, PLV and CLV/CAC.

The results corroborate the view of Mosaddegh et al. (2021), that segmentation based on behavioral dynamics allows identifying high-potential segments and design strategies more aligned with the customer life cycle. Through this model, the company can direct its prospecting strategy, seeking to reach customers in the most profitable segments and maximize its CLV, PLV and CLV/CAC. PLV and CLV/CAC allow an analysis of investment allocation in new customers across different scenarios, different volumes, margins and investments. This model may enhance the return on capital invested by the company.

In any case, as proposed by Malthouse and Blattberg (2005), the CLV, PLV and CLV/CAC projections should serve as guidelines for the company, recognizing that there is inherent error linked to the model and that customers considered less profitable may become highly profitable in the future, as well as customers considered highly profitable may not meet the company's expectation. Therefore, this evaluation provides greater information and support in decision making, as it incorporates variables of customers' purchasing behavior. In this way, the company will be able to allocate its capital with a balanced and effective mix, aiming at a portfolio of clients from different segments, but with a greater proportion in the highest profitability segments.

The results obtained reinforce the validity of the framework in a practical context, but its strength lies in the system developed to choose profitable customers, which is replicable in other sectors. The framework provides a structure that can be adapted for different markets by adjusting the methods for calculating the likelihood of the customer being active and the evolution of margins.

6 Conclusions

The model and analysis system developed in this work fulfills its role as a driver for selecting new customers, along with their respective profiles and segments for the company. There is a significant difference in expected profitability across the segments in which the company operates. Thus, managers can balance and allocate investments in new clients to optimize the risk–return trade-off, focusing efforts on the highest-profitability segments while maintaining diversification across segments, since there is error inherent to the models, as in any econometric or probabilistic approach.

This work contributes to the literature by combining different models to develop a framework aimed at prospecting new customers, an area still underexplored in research on customer lifetime value (CLV). While most studies focus on evaluating CLV for active customers (Ali & Shabn, 2024; Mosaddegh et al., 2021; Donkers et al., 2007; Gupta et al., 2006; Venkatesan & Kumar, 2004), this study proposes a system directly replicable in companies operating under contractual and continuous relationships. The framework converges with the literature by adopting CLV as a central metric (Mosaddegh et al., 2021; Venkatesan & Kumar, 2004) but differs by expanding its application to strategic decisions aimed at customer acquisition. In addition, while works such as Ali and Shabn (2024) highlight the use of CLV in organizational strategies, this study advances by proposing an adaptable model for different business configurations, with a specific focus on prospecting new customers.

As suggested by Mosaddegh et al. (2021), the analysis of segment dynamics is crucial to predict CLV and, consequently, to support strategic decisions related to customer acquisition and retention. The comparison of PLV and CLV/CAC across segments demonstrates the impact of customer behavior on expected future profitability. Currently, the analyzed company, as well as many others, evaluates all customers in the same way at the time of acquisition. However, using a predictive metric can increase the acquisition of more profitable customers, thereby increasing the value of the customer base and, consequently, the value of the company.

From a practical point of view, this paper proposes an implementation framework that companies and practitioners can use to structure a model and to allocate marketing resources for acquiring new customers based on CLV. The paper discusses how the company can manage its resources by envisioning a portfolio of new customers, seeking effective allocation and higher profitability.

As a practical implication, beyond the decision of how to allocate the company's investments, other strategies can be developed to maximize CLV in customer segments considered less profitable. New relationship strategies for approaching customers can be designed to reduce churn propensity throughout the business relationship, acting on the probability component of the customer being active. New opportunities can be generated by offering new products and services, cross-selling, or increasing the amount sold (up-selling), directly influencing the margin expansion component proposed in the model. In addition, actions can be taken regarding other items addressed in the framework, such as reducing the cost to serve and the cost of acquiring customers.

Among the limitations of this research, we note the use of data from a regional branch of the analyzed company, since customers from different geographic regions may exhibit different behaviors. Moreover, to apply this model across the whole company, in other companies, or in other locations, it would be necessary to validate it under those new configurations. The analysis in large economic-activity segment groups is broad, and different subgroups within these segments may have distinct behaviors, which may alter the survival curves and margin-expansion curves constructed. Future studies can evaluate the application of this model in different regions, industries, and scenarios. Segmentation to capture customer behavior could be evaluated from other perspectives, beyond the economic activity segment used here, and new variables, not considered in this work, may be included in future evaluations.

The available period of company data was five years, which the company also reported as the standard contract term it signs with clients. However, the company also mentions that many customers renew their contracts for new periods and, therefore, the CLV projection may be underestimated—an effect cited by Gupta and Lehmann (2003). Other studies could analyze a longer horizon to project CLV over longer periods.

Regarding model limitations, as noted above, the models have calculated accuracy, allowing a margin of error in the analysis of a new customer. As Gujarati and Porter (2011) affirm, regression analysis based on time-series data assumes that the underlying series are stationary. Therefore, stationarity may affect the developed model. The model used was based on approaches already studied in the literature, but future CLV models can employ economic forecasting techniques for time-series data—such as panel data, exponential smoothing methods, single-equation regression models, simultaneous-equation regression models, autoregressive integrated moving average processes, and vector autoregressions (Gujarati & Porter, 2011).

Although the study used data from a specific company, the framework can be applied directly to any organization that operates under contractual and ongoing relationships, such as energy, fuel, telecommunications, internet, or subscription businesses. For distinct operational configurations, predictive models can be adapted to the particularities of the sector or target market.

As described by Malthouse and Blattberg (2005), projecting CLV based on historical data is not an easy task, and customers expected to be the most profitable may not remain so. The models contain associated error, and thus the company must evaluate its strategies seeking a balance between risk and return. This also applies to the model developed in this work..

Future research could verify how the projections materialize over time, comparing them with new models to evaluate which have lower associated error. Given the company's configuration, business relationships, and the structure of the data series—which in this case contained censoring—this was the chosen approach; however, under different settings, other models could be tested to assess predictive capacity.

Finally, based on the applied framework presented and the literature, future research could evaluate the impacts on the company's expected profitability, aiming at effective allocation of marketing resources through different actions to maximize CLV, PLV, and CLV/CAC by adjusting the proposed components: customer acquisition cost, cost to serve, survival curve, and margin evolution over time.

Authors' contribution

Contribution	Rizzo, G. A	Trez, G
Conceptualization	X	X
Methodology	X	X
Software	X	X
Validation	X	X
Formal analysis	X	X
Investigation	X	X
Resources	X	X
Data Curation	X	-------
Writing - Original Draft	X	X
Writing - Review & Editing	X	X
Visualization	X	X
Supervision	X	X
Project administration	X	-------

References

Ali, N., & Shabn, O. S. (2024). Customer lifetime value (CLV) insights for strategic marketing success and its impact on organizational financial performance. Cogent Business & Management, 11(1). https://doi.org/10.1080/23311975.2024.2361321

Blattberg, R. C., & Deighton, J. (1996). Manage Marketing by the Customer Equity Test. *Harvard Business Review*, *74*, *4*, 136–144.

Colosimo, E. A., & Giolo, S. R. (2006). *Análise de Sobrevivência Aplicada* (1º). Edgard Blücher.

Cowan, G., Mercuri, S., & Khraishi, R. (2023). Modelling customer lifetime-value in the retail banking industry. http://arxiv.org/abs/2304.03038

Cox, D. R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, *34*(2), 187–202. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x

B., Verhoef, P. C., & de Jong, M. G. (2007). Modeling CLV: A test of competing models in the insurance industry. *Quantitative Marketing and Economics*, *5*(2), 163–190. https://doi.org/10.1007/s11129-006-9016-y

Ferreira, J. C., & Patino, C. M. (2015). What does the p value really mean? *Jornal Brasileiro de Pneumologia*, *41*(5), 485–485. https://doi.org/10.1590/S1806-37132015000000215

Göçken, M., Özçalıcı, M., Boru, A., & Dosdoğru, A. T. (2016). Integrating metaheuristics and Artificial Neural Networks for improved stock price prediction. *Expert Systems with Applications*, *44*, 320–331. https://doi.org/10.1016/j.eswa.2015.09.029

GUJARATI, D. N., & Porter, D. C. (2011). *Econometria Básica* (5º). Grupo A. https://integrada.minhabiblioteca.com.br/#/books/9788580550511/

Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N., Ravishanker, N., & Sriram, S. (2006). Modeling Customer Lifetime Value. *Journal of Service Research*, *9*(2), 139–155. https://doi.org/10.1177/1094670506293810

Gupta, S., & Lehmann, D. R. (2003). Customers as assets. *Journal of Interactive Marketing*, *17*(1), 9–24. https://doi.org/10.1002/dir.10045

Gupta, S., Lehmann, D. R., & Stuart, J. A. (2004). Valuing Customers. *Journal of Marketing Research*, *41*(1), 7–18. https://doi.org/10.1509/jmkr.41.1.7.25084

Jackson, B. B. (1985). *Winning and Keeping Industrial Customers: The Dynamics of Customer Relationships*. Lexington Books.

Jain, D., & Singh, S. S. (2002). Customer lifetime value research in marketing: A review and future directions. *Journal of Interactive Marketing*, *16*(2), 34–46. https://doi.org/10.1002/dir.10032

Kaplan, E. L., & Meier, P. (1992). *Nonparametric Estimation from Incomplete Observations*. *53*(282), 319–337. https://doi.org/10.1007/978-1-4612-4380-9_25

Kleinbaum, D. G., & Klein, M. (2012). *Survival Analysis* (Third, Vol. 39, Issue 2). Springer New York. https://doi.org/10.1007/978-1-4419-6646-9

Kumar, V. (2008). Customer Lifetime Value – The Path to Profitability. *Foundations and Trends® in Marketing*, *2*(1), 1–96. https://doi.org/10.1561/1700000004

Kumar, V., & Reinartz, W. J. (2006). *Customer relationship management: a databased approach* (N. J. : J. W. & S. Hoboken (ed.)).

Kumar, V., & Shah, D. (2015). Handbook of Research on Customer Equity in Marketing. In *Handbook of Research on Identity Theory in Marketing*. Edward Elgar Publishing. https://doi.org/10.4337/9781781004982

Lopes, B., Ramos, I. C. de O., Ribeiro, G., Correa, R., Valbon, B. de F., Luz, A. C. da, Salomão, M., Lyra, J. M., & Ambrósio Junior, R. (2014). Biostatistics: fundamental concepts and practical applications. *Revista Brasileira de Oftalmologia*, *73*(1). https://doi.org/10.5935/0034-7280.20140004

Malthouse, E. C., & Blattberg, R. C. (2005). Can we predict customer lifetime value? *Journal of Interactive Marketing*, *19*(1), 2–16. https://doi.org/10.1002/dir.20027

Mosaddegh, A., Albadvi, A., Sepehri, M. M., & Teimourpour, B. (2021). Dynamics of customer segments: A predictor of customer lifetime value. Expert Systems with Applications, 172, 114606. https://doi.org/10.1016/j.eswa.2021.114606

Reinartz, W. J., & Kumar, V. (2003). The impact of customer relationship characteristics on profitable lifetime duration. *Journal of Marketing*, *67*(1), 77–99. https://doi.org/10.1509/jmkg.67.1.77.18589

Rich, J. T., Neely, J. G., Paniello, R. C., Voelker, C. C. J., Nussenbaum, B., & Wang, E. W. (2010). A practical guide to understanding Kaplan-Meier curves. *Otolaryngology–Head and Neck Surgery*, *143*(3), 331–336. https://doi.org/10.1016/j.otohns.2010.05.007

RStudio. (2022). *Products*. https://www.rstudio.com/products/rstudio/

Rust, R. T., Lemon, K. N., & Zeithaml, V. A. (2004). Return on Marketing: Using Customer Equity to Focus Marketing Strategy. *Journal of Marketing*, *68*(1), 109–127. https://doi.org/10.1509/jmkg.68.1.109.24030

Singh, S. S., & Jain, D. C. (2013). Measuring Customer Lifetime Value. In *Faculty & Research*. https://doi.org/10.1108/S1548-6435(2009)0000006006

Venkatakrishna, M. R., Mishra, M. P., Sneha, M., & Tiwari, P. (2021). Customer Lifetime Value Prediction and Segmentation using Machine Learning. International Journal of Research in Engineering and Science (IJRES) ISSN, 9, 36–48. www.ijres.org

Venkatesan, R., & Kumar, V. (2004). Framework for Customer Selection. *Journal of Marketing*, *68*(October), 106–125.

Verhoef, P. C., & Lemon, K. N. (2013). Successful customer value management: Key lessons and emerging trends. *European Management Journal*, *31*(1), 1–15. https://doi.org/10.1016/j.emj.2012.08.001

Zeithaml, V. A., Jaworski, B. J., Kohli, A. K., Tuli, K. R., Ulaga, W., & Zaltman, G. (2020). A Theories-in-Use Approach to Building Marketing Theory. Journal of Marketing, 84(1), 32-51. https://doi.org/10.1177/0022242919888477