

# Estimating Customer Lifetime Value: A Metric Alternative for Customer Segmentation applied to the Retail Sector

Melissa Montero, Econ. <sup>1</sup>, Diana Suarez, Econ. <sup>2</sup>, Milton Ismael Paredes Aguirre, Msc. <sup>3</sup> and Ronald Enrique Campoverde Aguirre, Ph.D. <sup>4</sup>

<sup>1,2,3,4</sup> Escuela Superior Politécnica del Litoral, ESPOL, Facultad de Ciencias Sociales y Humanísticas - FCSH, Campus Gustavo Galindo Km. 30.5 Vía Perimetral, P.O. Box 09-01-5863, Guayaquil, Ecuador, mgmonter@espol.edu.ec<sup>1</sup>, djsuarez@espol.edu.ec<sup>2</sup>, isparede@espol.edu.ec<sup>3</sup>, recampov@espol.edu.ec<sup>4</sup>

*Abstract— The success of a company depends on its ability to build and maintain long-term relationships with its most valuable buyers. Ecuadorian companies face the challenge of identifying, differentiating and segmenting their clients. In response to this need, this study seeks to identify an unconventional form of segmentation based on the expected monetary value of customer billing in the retail sector. To this end, a probabilistic model was implemented that is derived from the Pareto/NBD methodology with a Bayesian hierarchy approach to estimate the customer value (CLV), based on a RFM data structure that considers the recency, frequency and monetary value of customers. After the estimate, the respective segmentation of customers was made where 24% of them were found to generate 68% of sales, therefore, three-quarters of customers are considered unprofitable. This research proposes a more efficient segmentation alternative based on customer value in the context of a retail company.*

*Keywords—Customer lifetime value, Segmentation, Pareto NBD, MCMC.*

Digital Object Identifier (DOI):  
<http://dx.doi.org/10.18687/LACCEI2020.1.1.327>  
ISBN: 978-958-52071-4-1 ISSN: 2414-6390

# Estimando el Customer Lifetime Value: Una Métrica Alternativa para la Segmentación de Clientes aplicado al Sector Retail

Melissa Montero, Econ.<sup>1</sup>, Diana Suarez, Econ.<sup>2</sup>, Milton Ismael Paredes Aguirre, Msc.<sup>3</sup> and Ronald Enrique Campoverde Aguirre, Ph.D.<sup>4</sup>

<sup>1,2,3,4</sup> *Escuela Superior Politécnica del Litoral, ESPOL, Facultad de Ciencias Sociales y Humanísticas - FCSH, Campus Gustavo Galindo Km. 30.5 Vía Perimetral, P.O. Box 09-01-5863, Guayaquil, Ecuador, mgmonter@espoledu.ec<sup>1</sup>, djsuarez@espol.edu.ec<sup>2</sup>, isparede@espol.edu.ec<sup>3</sup>, recampov@espol.edu.ec<sup>4</sup>*

**Resumen--** *El éxito que logre alcanzar una empresa depende de su capacidad para construir y mantener relaciones a largo plazo con sus compradores más valiosos. Las empresas ecuatorianas se enfrentan al reto de poder identificar, diferenciar y segmentar a sus clientes. En respuesta a esta necesidad, el presente estudio busca identificar una forma no convencional de segmentación basado en el valor monetario esperado de la facturación de clientes en el sector retail. Para ello, se implementó un modelo probabilístico que se desprende de la metodología de Pareto/NBD con un enfoque de jerarquía bayesiana para estimar el valor del cliente (CLV), basados en una estructura de datos RFM que considera la recencia, frecuencia y monto de compra de los clientes. Después de realizar la estimación, se realizó la respectiva segmentación de clientes donde se halló que el 33% de ellos generan el 68% de las ventas, por lo tanto, las tres cuartas partes de clientes son considerados menos rentables. Esta investigación propone una alternativa de segmentación más eficiente en función del valor del cliente en el contexto de una empresa de retail.*

**Palabras Clave--** *Valor de vida del cliente, Segmentación, Pareto NBD, MCMC.*

**Abstract--** *The success of a company depends on its ability to build and maintain long-term relationships with its most valuable buyers. Ecuadorian companies face the challenge of identifying, differentiating and segmenting their clients. In response to this need, this study seeks to identify an unconventional form of segmentation based on the expected monetary value of customer billing in the retail sector. To this end, a probabilistic model was implemented that is derived from the Pareto/NBD methodology with a Bayesian hierarchy approach to estimate the customer value (CLV), based on a RFM data structure that considers the recency, frequency and monetary value of customers. After the estimate, the respective segmentation of customers was made where 24% of them were found to generate 68% of sales, therefore, three-quarters of customers are considered unprofitable. This research proposes a more efficient segmentation alternative based on customer value in the context of a retail company.*

**Keywords--** *Customer lifetime value, Segmentation, Pareto NBD, MCMC.*

## I. INTRODUCCIÓN

En la actualidad, el mundo de los negocios se está digitalizando y cada día las empresas enfrentan un mayor nivel de competencia, haciendo más difícil la

supervivencia [1]. Las economías persisten en explorar nuevas estrategias para mejorar su competitividad, a paso lento, las empresas buscan mejorar sus modelos de negocios, cambiando la forma tradicional de tomar decisiones [2]. El progreso de las tecnologías de información ha hecho posible que las compañías conozcan información de las preferencias de sus clientes. Sin embargo, estas constantemente fallan en el uso de esta información y en la definición eficiente de criterios de segmentación, debido a que no van acorde a sus objetivos.

Por ello, contar con información sobre las preferencias de los clientes es importante mantener una relación a largo plazo [2]. En el contexto del CRM, la relación de una empresa con sus compradores es el foco principal. Por ello, los gestores deben usar herramientas que le permitan identificar los clientes que tienen mayor probabilidad de permanecer con la empresa y reconocer los patrones de comportamientos futuros. De esta manera, las empresas contarán con información necesaria para poder gestionar las decisiones de marketing y valorar su base de clientes con el fin de mejorar la planificación estratégica [3].

El CLV es una métrica clave para identificar a los clientes con mayor valor en términos monetarios [4]. Una segmentación a base de esta métrica es esencial para gestionar adecuadamente la relación con diferentes tipos de clientes [5]. Empresas como Harrah's, IBM, Capital One, LL Bean, Netflix, entre otras, usan CLV de manera rutinaria como una herramienta para administrar y medir el éxito de sus negocios.

Este estudio centra su análisis en empresas de comercio minorista. La importancia de este sector radica en su acelerado crecimiento, gracias a la evolución tecnológica y la mejora en los procesos de distribución, brindando una mayor disponibilidad de productos en diferentes tipos de canales [6], además de ser un sector generador de empleo [7]. Una característica importante de esta industria es que se desarrolla en entornos no contractuales, debido a que sus clientes son principalmente consumidores finales quienes no declaran en ningún momento su inactividad, sino que simplemente dejan de concurrir [8].

Este tipo de comercio cuenta con una gran variedad de clientes cuyo comportamiento de compra es heterogéneo, haciendo que se convierta en un verdadero desafío identificar a los clientes más valiosos. Si se parte del principio de Pareto, donde se establece que el 20% de los clientes son capaces de

Digital Object Identifier (DOI):

<http://dx.doi.org/10.18687/LACCEI2020.1.1.329>

ISBN: 978-958-52071-4-1 ISSN: 2414-6390

generar el 80% de las ventas ¿Las empresas ecuatorianas pueden identificar ese 20% de clientes más valiosos? ¿Están en condiciones de estimar cuánto les van a comprar sus clientes los próximos meses?

Según datos del INEC, para el 2016 el comercio en el Ecuador representó el 51,16% del total de empresas, participación importante debido a que este mercado engloba todas las empresas especializadas en la comercialización masiva de productos y/o servicios. Específicamente, en el comercio del retail las empresas con mayor representatividad son supermercados. En el año 2018, las 5 cadenas de supermercados más grandes generaron una cifra superior a los 4.2 mil millones de dólares, el equivalente al 4% del PIB ecuatoriano.

El éxito de los supermercados radica en la posibilidad de adquirir múltiples productos y servicios bajo un mismo techo. La oferta de productos en este sector es muy variada desde productos básicos de consumo diario hasta productos de compra ocasional como electrodomésticos, ropa, zapatos, entre otros. A pesar de su gran alcance, estas empresas se enfrentan a un conjunto de consumidores con comportamiento y características heterogéneas. Por lo tanto, es un gran desafío para ellos lograr identificar los grupos de clientes que visitan sus tiendas, entender sus necesidades y poder ofrecer servicios que logren satisfacerlos.

Un criterio eficiente de segmentación se convierte en una alternativa esencial para manejar una gran variedad de clientes con preferencias variadas [9]. Hoy en día, las empresas no deben limitarse a tener un establecimiento atractivo con productos accesibles, sino ir un paso más allá generando estrategias a corto y largo plazo que contribuyan a la fidelización de los clientes a fin de garantizar su continuidad en un entorno cada vez más competitivo.

La segmentación en función al CLV es una alternativa más eficiente para modelar el comportamiento de compra. Reconocer a los clientes como activos intangibles que aumentan el rendimiento empresarial, permitirá administrar mejor la creación de valor para la empresa y el proceso de gestión de flujo de efectivo, de esta manera mejorar su crecimiento y su sostenibilidad en el mercado. Adicionalmente, la información que otorga sobre el cliente permitirá establecer mecanismos para aumentar su nivel de satisfacción y la retención del comprador. Por las razones expuestas, la presente investigación emprende el análisis del comportamiento de compra de los clientes a través de sus actividades transaccionales, considerando su dinamismo durante un determinado período de tiempo y así estimar el valor que representa uno o un segmento de clientes para la empresa.

## II. MARCO TEÓRICO

### A. Customer Relationship Management

El *Customer Relationship Management* (CRM) es un modelo de negocios y una estrategia cuyo eje central es el cliente; se define como un proceso dedicado a establecer, mantener y mejorar las relaciones con clientes de alto valor a

largo plazo [10], [11]. Para lograr una gestión eficaz, el CRM debe tener en cuenta que no todos los clientes son iguales, por ende, deben ser tratados de manera distinta [12]. En virtud de ello, es necesario determinar una asignación de recursos según los diferentes niveles de clientes. Una métrica clave para la gestión efectiva de los clientes es el Customer Lifetime Value (CLV), el cual permite identificar a los clientes de alto valor para la empresa [13], [14].

A medida que la competencia en el mercado se vuelve más severa, mantener a clientes valiosos se vuelve un desafío para las empresas. Blattberg, Gary y Thomas [15] resaltan la importancia de identificar los tipos de cliente ya que estos no generan la misma rentabilidad. De acuerdo con el principio 80/20, también llamada Ley de Pareto, el 20% de los clientes aportan el 80% del beneficio a la empresa [16]. Por lo tanto, identificar y retener ese 20% de los clientes es fundamental para una planificación estratégica de las compañías [17].

### B. Customer Lifetime Value

El CLV es definido como el valor presente de los beneficios futuros obtenidos de un cliente durante su vida en relación con la empresa [18]. La versatilidad de esta métrica ha permitido su aplicación a diferentes niveles como: (a) el cálculo y la predicción del CLV [19], (b) los efectos de estrategias de mercadotecnia basados en CLV [20] y (c) la optimización de la asignación de recursos de marketing a los clientes basados en CLV [21]. De manera general, la estimación del CLV ayuda a las empresas a gestionar efectivamente las relaciones con sus clientes en términos de cómo iniciarlas, mantenerlas y mejorarlas [22].

El desafío clave en la medición del CLV es el desarrollo de un modelo confiable para pronosticar el flujo futuro de ganancias que cada cliente proporcionará a la empresa. Los académicos han desarrollado una variedad de modelos que se aplican a contextos específicos. Estos generalmente se dividen en contextos de configuraciones contractuales y no contractuales [23], [24], tal como lo muestra la tabla 1. En una configuración contractual, la empresa observa la deserción del cliente, mientras que en una configuración no contractual, la deserción no se observa [22].

TABLA 1  
Técnicas de modelado

	Ajustes No Contractuales	Ajustes Contractuales
Compras Continuas	Modelos Probabilísticos Pareto/NBD [25] BG/NBD [26] Pareto GGG [27] Modelo Gamma Gompertz GG/NBD [28] Bayesiano Jerárquico [29] Modelos Machine Learnings CART + Regresión Logit/Lineal [30] Modelos de Markov Modelo oculto de Markov [22]	Modelo Gamma Exponencial Modelo Gamma-Weibull [31]
Compras Discretas	BB/G Modelo de Estructura Básica de CLV	Modelo Jerárquico Bayesiano [32]

Estos modelos de estimación del CLV generalmente se enfocan en la frecuencia, recencia y el valor monetario promedio en cada compra. Varios enfoques relacionan estas variables con factores gestionables desde el punto de vista administrativo que permiten analizar el impacto de ellas en la rentabilidad esperada [21]. Otros enfoques modelan estas variables utilizando supuestos del comportamiento de compra, sin embargo, omiten covariables explicativas que podrían mejorar las estimaciones [33]. La ventaja principal de los modelos estocásticos mencionados es que se basan en un simple conjunto de datos sobre el comportamiento del cliente, esta estructura se adapta a casi todas las empresas y permite una comprensión más profunda de cómo se relaciona dicho comportamiento con el valor del cliente.

Investigaciones anteriores han demostrado que estos modelos estocásticos son adecuados para abordar una serie de problemas de gestión. Por ejemplo, la selección de clientes para la asignación de recursos y la medición de CLV durante períodos cortos y largos de tiempo. Uno de los primeros modelos en realizar este tipo de análisis es el modelo Pareto / NBD propuesto por Schmittlein, Morrison y Colombo [25]. A partir de su trabajo pionero, otros estudios han presentado extensiones mejoradas del modelo original [34], [35], [36], [33], [29], [37].

Sin embargo, han existido críticas en cuanto a la finalidad del CLV. Kumar y George [38] y Homburg, Droll y Totzek [39] han argumentado que esta estrategia puede tener efectos negativos en las relaciones de una empresa con sus clientes de bajo valor; esto debido a que los coloca en un nivel de baja prioridad, por lo que, se puede caer en el error de dejarlos insatisfechos [40]. Esta insatisfacción podría desencadenar una negativa publicidad boca a boca, dando lugar a una reducción en ventas y los beneficios a largo plazo [41]. Paralelamente, la teoría apunta un efecto positivo a largo plazo debido a la retención de clientes de alto valor.

Un hallazgo común es que los clientes de alto valor no reciben suficiente atención y que algunas empresas gastan demasiado en clientes de bajo valor [42]. La orientación de CLV debe mejorar la satisfacción de los clientes de alto valor a través de un tratamiento preferencial respecto al producto ofrecido, el precio, la atención y la comunicación con el cliente [39]. Por lo tanto, las estrategias basadas en el CLV provocan reacciones positivas en los clientes de alto valor que mejoran los ingresos y los beneficios a través de un incremento en los volúmenes de ventas.

### C. Estructura de datos RFM

Varios modelos probabilísticos parten de una estructura de datos con variables de RFM, *Recency*, *Frequency* and *Monetary Value*. El análisis bajo esta estructura es una técnica robusta y ampliamente utilizada en el marketing de bases de datos, su aplicación se ajusta a varios escenarios que implican un gran número de clientes como las compras en línea, ventas minoristas, entre otras [43]. Así mismo, es utilizado para

clasificar a los clientes según su historial de compra en base a tres dimensiones:

#### *Recencia*

Se refiere al tiempo transcurrido desde la última transacción del cliente. Un menor valor de recencia implica que el cliente visitará la empresa en un corto período de tiempo. Por otro lado, un mayor valor supone que es menos probable que el cliente visite la empresa en breve.

#### *Frecuencia*

Es el número transacciones que un cliente realizó durante un período específico. Cuanto mayor sea el valor de la frecuencia, los clientes presentan ser más fieles a la empresa.

#### *Valor Monetario*

Se define como la cantidad de dinero gastada por el cliente durante un período determinado. Cuanto mayor sea la cantidad de dinero que el cliente gasta, esto genera más ingresos a la empresa.

La estructura RFM tiene la capacidad de predecir el comportamiento de los clientes en el futuro, por lo tanto, existe una relación intrínseca en la estimación del CLV [33], [38]. Sin embargo, la estructura como tal no provee información sobre la posibilidad de que los clientes realicen futuras compras, tampoco proporciona alguna estimación económica que refleje el valor de los clientes. En virtud de ello, es necesario un modelo que complemente el poder predictivo de las variables RFM.

### D. Modelo de Pareto/NBD

Uno de los primeros modelos en abordar la probabilidad de que un cliente seguirá siendo activo en el futuro y sea capaz de describir su comportamiento de compra es el modelo de Pareto/NBD. Este modelo fue desarrollado por Schmittlein, Morrison y Colombo [25] el cual describe el flujo de transacciones de los clientes en entornos no contractuales.

Basado en este modelo inicial, Fader, Hardie y Lee [33] construyeron el modelo BG / NBD-GG aplicable en un entorno amplio de industrias. Sin embargo, este modelo se basa en el supuesto de deserción permanente del cliente y no toma en consideración la existencia de heterogeneidad en el comportamiento de los compradores. Estas limitaciones pueden conducir a predicciones inexactas debido a que estas suposiciones no se cumplen en la mayoría de los entornos [22].

Por otro lado, estos modelos generalmente son estimados bajo máxima verosimilitud (MLE) presentando soluciones únicas, dichas soluciones singulares generan restricciones al momento de intentar relajar algunos supuestos subyacentes de comportamiento de compra de los clientes. Para resolver estas desventajas un método de estimación alternativo aplicado en modelos probabilísticos son las simulaciones de Markov-Chain-Monte-Carlo (MCMC). Esta metodología realiza cálculos más complejos lo que deriva a costos altos de implementación. Sin embargo, este método tiene ventajas superiores dado que permite realizar estimaciones de

parámetros a nivel individual y general, además genera mayor flexibilidad en los supuestos restrictivos de manera que proporciona estimaciones más realistas.

Bajo esta metodología de estimación se han desarrollado varias extensiones del modelo original propuesto por Schmittlein y Colombo [25]. Ma y Liu [44] han desarrollado una variante jerárquica de Bayes, esta propuesta mejora los resultados de estimación, pero aún no aprovecha al máximo la flexibilidad obtenida, debido a que mantiene los supuestos restrictivos del modelo original. Mientras que, la propuesta de Abe [29] de Pareto / NBD consigue aprovechar las ventajas de estimación relajando el supuesto de independencia en del proceso de compra y la tasa de abandono, además considera la heterogeneidad existente en la cohorte de clientes. Por último, Platzer y Reutterer [27] proponen Pareto / GGG, el cual permite un grado variable de regularidad en los tiempos de transacción, sin embargo, este modelo mantiene el supuesto de independencia entre la frecuencia de compra y la tasa abandono, lo que no es realista en la mayoría de los casos.

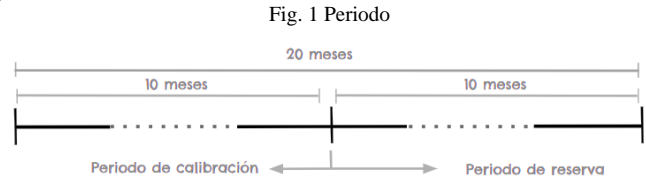
### III. METODOLOGÍA

En este estudio se aplicó la extensión del modelo de Pareto/NBD propuesto por Abe [29] que utiliza un enfoque jerárquico bayesiano para predecir el comportamiento de compra de los clientes en contextos no contractuales. El método de estimación utilizado está basado en simulaciones de Markov-Chain-Monte-Carlo (MCMC), con un algoritmo que realiza la estimación de los parámetros individuales y generales del modelo probabilístico, en base a registros de transacciones históricas del cliente. Finalmente, una evaluación de métricas de error permite explorar la idoneidad del modelo. Se usó el software estadístico R versión 3.6.1 para el cálculo del CLV.

Se realizó un análisis a nivel individual con datos transaccionales de múltiples clientes de una de las principales tiendas de supermercados en el país. En específico, se seleccionó una sucursal que opera en la ciudad de Guayaquil. Tomando como referencia artículos anteriores en el cálculo de CLV [34], [21], [33], se aplicó un proceso de dos fases para la estimación que consisten en: (a) pronosticar el número de transacciones futuras de cada individuo y su probabilidad de mantener una relación activa con la empresa, (b) Calcular el beneficio promedio individual por transacción. De esta manera, con el producto de la cantidad futura de transacciones, el beneficio por transacción, la probabilidad de que se realice compras en el futuro y la tasa de descuento se obtuvo una aproximación del CLV para cada cliente.

Para realizar las estimaciones se dividió el conjunto de datos en un período de calibración y un período de reserva para poder validar las predicciones del modelo [45]. Los datos a nivel de cliente se transformaron en una estructura RFM, *Recency, Frequency and Monetary Value*. El conjunto de datos preparado del período de calibración se utilizó para ajustar los parámetros de Pareto/NBD Abe. Las predicciones a nivel individual sobre el número futuro de compras, la probabilidad de estar activo y el valor monetario se utilizan para calcular el

CLV y posteriormente se descuentan mediante una tasa. Finalmente, se estiman dos valores de CLV: el CLV predicho a través del modelo de Pareto/NBD Abe y el CLV real en el período de reserva.



#### A. Modelando la Frecuencia de Compra

El modelo de Pareto/NBD Abe analiza la relación entre la frecuencia y recencia, de esta manera provee la probabilidad de que un cliente individual esté activo en un momento determinado en el tiempo. En este modelo, el comportamiento del consumidor se caracteriza por: (a) distribución de compra de Poisson con el parámetro  $\lambda$  que describe la tasa de compra, y además (b) el tiempo de vida modelado como una función exponencial con  $\mu$  como parámetro para la tasa de deserción [29]. El presente modelo captura la heterogeneidad de los clientes a través de la estimación de parámetros específicos individuales con un enfoque jerárquico bayesiano. En particular, la propuesta mantiene los supuestos del comportamiento propuestos por Schmittlein, Morrison y Colombo [25], pero tiene dos diferencias características: (1) complementa la parte analítica de la distribución global, considerando una distribución individual que en cuenta la heterogeneidad de los clientes con un método de simulación [29].

Los datos de frecuencia, recencia y tiempo están representados con las siguientes variables  $(x, t_x, T)$ . La vida comienza a el tiempo 0 (cuando se produce la primera transacción) y las transacciones de los clientes se monitorean hasta el tiempo  $T$ .  $x$  es el número de transacciones repetidas observadas en el período de tiempo  $(0, T]$ , con la última compra ( $x$ th repeticiones) ocurriendo en  $t_x$ . Por lo tanto, la recencia se define como  $T - t_x$ .

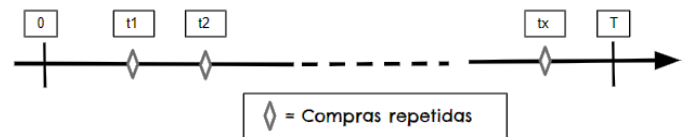


Fig. 2 Línea de tiempo transaccional

Usando estas notaciones matemáticas, los supuestos de este modelo se pueden expresar de la siguiente manera:

*Proceso de compra de Poisson:*

$$P[x|\lambda] = \begin{cases} \frac{(\lambda T)^x}{x!} e^{-\lambda T} & \text{if } \tau > T \\ \frac{(\lambda \tau)^x}{x!} e^{-\lambda \tau} & \text{if } \tau > T \end{cases} \quad x = 0, 1, 2, \dots, \quad (1)$$

Donde  $\tau$  es la vida del cliente, variable no observada.

Tiempo de vida con una distribución exponencial:

$$f(\tau) = \mu e^{-\mu\tau} \quad \tau \geq 0, \quad (2)$$

Incorporando Covarianzas

$$\begin{bmatrix} \log(\lambda) \\ \log(\mu) \end{bmatrix} \sim MVN \left( \theta_0 = \begin{bmatrix} \theta_\lambda \\ \theta_\mu \end{bmatrix}, \Gamma_0 = \begin{bmatrix} \sigma_\lambda^2 & \sigma_{\mu\lambda} \\ \sigma_{\mu\lambda} & \sigma_\mu^2 \end{bmatrix} \right) \quad (3)$$

donde MVN denota una distribución normal multivariada.

Introduciendo variables latentes

Para estimar  $\lambda$  y  $\mu$  de forma individual dadas las funciones a priori, se incorporan las variables latentes para obtener las funciones posteriori. Es decir, que no se puede aprovechar la simple actualización bayesiana si se desconocen las variables no observadas (si el cliente  $i$  está activo en  $T_i$  y, si no, el tiempo de deserción  $y_i < T_i$ ). Esta es la verdadera razón del complejo proceso de estimación asociado con el modelo Pareto / NBD. Por ello, se incorporan variables no observables como variables latentes en el presente modelo. Para reducir la notación, el subíndice  $i$  se elimina.  $z$  se define como 1 si un cliente está activo en el momento  $T$  y 0 de lo contrario. Otra variable latente es el tiempo de deserción  $y$  que aparece cuando  $z = 0$  (el cliente se encuentra inactivo). Si se conocen  $z$  e  $y$ , implica que la función de probabilidad para los datos de frecuencia y recencia  $(x, t_x, T)$  se convierte en una expresión simple para  $x > 0$ .

Caso  $z = 0$  (el cliente está activo en el tiempo  $T$ )

$P(x$ th compra en  $t_x$  y está activo hasta  $T$  y no compra entre  $[t_x, T])$

$$\begin{aligned} &= \frac{\lambda^x t_x^{x-1}}{\Upsilon(x)} e^{-\lambda t_x} \times e^{-\lambda(T-t_x)} \\ &= \frac{\lambda^x t_x^{x-1}}{\Upsilon(x)} e^{-(\lambda+\mu)T} \end{aligned}$$

Caso  $z = 1$  (el cliente está inactivo en el tiempo  $T$ )

$P(x$ th compra en  $t_x$  y no compra entre  $[t_x, T]$  y abandona en  $y \leq T$ )

$$\begin{aligned} &= \frac{\lambda^x t_x^{x-1}}{\Upsilon(x)} e^{-\lambda t_x} \times e^{-\lambda(y-t_x)} \times \mu e^{-\mu y} \\ &= \frac{\lambda^x t_x^{x-1}}{\Upsilon(x)} \mu e^{-(\lambda+\mu)y} (t_x \leq y \leq T) \end{aligned}$$

Combinando estos dos casos, se obtuvo una notación más compacta para la función de probabilidad:

$$L(x, t_x, T | \lambda, \mu, z, y) = \frac{\lambda^x t_x^{x-1}}{\Upsilon(x)} \mu^{1-z} e^{-(\lambda+\mu)\{zT+(1-z)y\}} \quad (4)$$

Para  $x = 0$ , no hay compras repetidas y  $t_x = 0$ . Por lo tanto,  $\Gamma(x = 0)$  y  $t_x^{x-1}$  no están definidas. La función de probabilidad apropiadas es  $e^{-(\lambda+\mu)T}$  para  $z = 1$ , y  $\mu e^{-(\lambda+\mu)y}$  para  $z = 0$ . Por lo tanto, la ecuación (4) se convierte en

$$L(x = 0, t_x, T | \lambda, \mu, z, y) = \mu^{1-z} e^{-(\lambda+\mu)\{zT+(1-z)y\}} \quad (5)$$

Sin embargo, como no se observa  $z$  ni  $y$ , son considerados como datos faltantes, por ello se aplica una técnica de aumento de datos [34]. Para simular  $z$  en nuestro procedimiento de estimación de MCMC, se usa la siguiente expresión para la probabilidad de que un cliente esté activo en  $T$ , o de manera equivalente  $z = 1$ .

$$\begin{aligned} &P[\tau > T | \lambda, \mu, T, t_x] \\ &= P[z = 1 | \lambda, \mu, T, t_x] \\ &= \frac{1}{1 + (\mu/(\lambda + \mu)) [e^{(\lambda+\mu)(T-t_x)} - 1]} \quad (6) \end{aligned}$$

## B. Especificaciones previas

Para estimar los parámetros específicos por cliente denotamos  $\theta_i = [\log \lambda_i, \log \mu_i]$  el cual se distribuye normalmente con media  $\beta' d_i$  y la matriz de varianza-covarianza  $\Gamma_0$ , donde el subíndice  $i$  indica los clientes individuales. El objetivo será estimar los parámetros  $\{\theta_i, y_i, z_i, \forall i; \beta, \Gamma_0\}$  en base a las variables observadas de frecuencia y recencia  $\{x_i, t_{x(i)}, T_i; \forall i\}$ .

Los parámetros  $\lambda_i$  y  $\mu_i$  se eligen como lognormal con la función prior especificada en la ecuación (3). A su vez los hiperparámetros  $\beta$  y  $\Gamma_0$  se estiman respectivamente de forma bayesiana con un prior multivariado normal y un prior Wishart inverso.

$$\beta \sim MVN(\beta_0, \Sigma_0), \quad \Gamma_0 \sim IW(v_{00}, \Gamma_{00})$$

## C. Modelando MCMC

Dadas las especificaciones previas se puede estimar los parámetros  $\{\theta_i, y_i, z_i, \forall i; \beta, \Gamma_0\}$  utilizando el método MCMC. Para estimar la densidad conjunta, se genera secuencialmente cada parámetro, desde su distribución condicional hasta lograr la convergencia. El procedimiento se describe a continuación:

- 1) Establecer el valor inicial para  $\theta_i^{(0)} \forall i$ .
- 2) Para cada cliente  $i$ ,
  - a. Generar  $\{z_i | \theta_i\}$  de acuerdo con la ecuación (6).
  - b. Si  $z_i = 0$ , se genera  $\{y_i | z_i, \theta_i\}$  usando una distribución exponencial truncada.
  - c. Obtener  $\{\theta_i | z_i, y_i\}$  utilizando la ecuación (4).
- 3) Generar  $\{\beta, \Gamma_0 | \theta_i, \forall i\}$  usando una actualización de regresión normal multivariada estándar.
- 4) Iterar los pasos (2) y (3) hasta que se logre la convergencia.

#### D. Validación de resultados Ajuste y Error

##### Errores de Predicción a Nivel Individual

Para evaluar el desempeño de los modelos a nivel individual se aplicaron varias métricas de error que sugiere la literatura: Error Absoluto Medio, MAE; Error Cuadrático Medio, Media RMSE; Error Cuadrático de la Mediana, Mediana RMSE y las medidas de sensibilidad fueron calculadas y analizadas en este estudio. Las métricas se definen a continuación en la tabla 2.

TABLA 2  
Errores de Predicción a Nivel Individual

Nombre	Evaluación de los autores	Ejemplos de estudios
Error Absoluto Medio (MAE)	Es una medida sencilla utilizada por múltiples autores	Jasek, Vrana, Sperkova, Smutny y Kobulsky [8]; Platzer y Reutterer [27]; Schwartz, Bradlow and Fader; Platzer [46]
Error Cuadrático Medio (RMSE Medio)	Es sensible a los datos atípicos	Platzer [46]; Wübben y Wangenheim [47]
Error Cuadrático de la Mediana (Mediana RMSE)	Presenta menor sensibilidad a los datos atípicos	Wübben y Wangenheim [47]
Sensibilidad	El porcentaje de clientes que están correctamente clasificados para pertenecer al 10% superior	Jasek et al. [8]; Malthouse y Blattberg [48]; Wübben y Wangenheim [47]
Correlación	Proposed by Fader, Hardie, & Lee [33] y cuestionado por Platzer [46].	Fader, Hardie y Lee [33]; Platzer [46]

##### Errores de Predicción a Nivel Agregado

Con el objetivo de evaluar el desempeño del modelo probabilístico aplicado, se compararon los resultados de las métricas de error con estudios similares realizados por varios autores en el pasado. A nivel agregado, la literatura sugiere dos métricas de error para comparar los CLV: Forecast vs Actual (FvA) y el Error porcentual absoluto medio (MAPE). El MAPE es una de las medidas de error de mayor uso, sin embargo, presenta limitaciones cuando la data contiene valores extremos o ceros. Por esta razón, no es una medida aplicable en este estudio debido a la existencia de clientes desertores que generan CLV de cero. En virtud de ello, la evaluación del FvA será suficiente para comprender la precisión del modelo a nivel agregado [8].

#### IV. ANÁLISIS DE RESULTADOS

Para el presente estudio, se analizaron clientes de uno las tres cadenas de supermercados más representativos en Ecuador, dirigido principalmente a sectores de ingresos medios y bajos. En específico, se seleccionaron a los clientes con tarjeta de fidelización que hayan realizado transacciones, por lo menos una vez durante el mes de septiembre del 2017. El conjunto de datos suministrado por la cadena de supermercado comprende

la facturación correspondiente a la compra de los clientes para un total de 20 meses, desde el 1 de septiembre del 2017 hasta el 30 de abril del 2019. Tomando como referencia los estudios de Glady, Baesens, & Croux [28] se divide el conjunto de datos en dos partes iguales, los primeros 10 meses corresponden al período de calibración y los subsiguientes meses al período de reserva. Se evaluó un período poco extenso, considerando la necesidad gerencial es proporcionar información que permita una asignación de marketing basada en el valor de los clientes en un futuro próximo, dado que los flujos futuros de períodos más amplios se descuentan considerablemente.

En cuanto a los productos, solo se incluyeron los ingresos generados por la venta de 10 secciones de la tienda debido a que éstas guardan la naturaleza de un supermercado, los cuales son: artículos de limpieza, bebidas, carnes, comestibles, congelados, embutidos, frutas y verduras, golosinas, lácteos y perfumería. Adicionalmente, como restricción se excluyeron productos cuyo PVP fuera mayor a USD \$100, para evitar la inclusión de datos atípicos lo cual generaría una sobreestimación en los montos de compra.

##### A. Evaluación y desempeño del Modelo

Para evaluar la competencia del modelo de Pareto/NBD Abe [29], se analizan las métricas del error tanto a nivel individual y como a nivel agregado. A nivel individual se estimó el MAE, la media y la mediana del RMSE, así también se analizaron los parámetros de sensibilidad. Con el objetivo de evaluar el desempeño del modelo probabilístico aplicado, se compararon los resultados de las métricas de error con estudios similares realizados por varios autores en el pasado. El MAE del modelo Abe Pareto/NBD es de 40% del CLV promedio, un valor más bajo comparando con el estudio de Bernant, Koning y Fok [49] con un MAE de 85,6% en una empresa de retail. La media del RMSE es aproximadamente el 72% del promedio del CLV y la mediana RMSE un 19%, estos resultados son comparables con el estudio de Do Ruibin y Vintilescu Borglöv [50] aplicado en un contexto similar que presentan un 120% y 34% para la media y mediana RMSE respectivamente, mostrando la superioridad del modelo de Pareto/NBD [29] frente al modelo original Pareto/NBD [25].

Respecto a la sensibilidad, el modelo también presenta una mejor clasificación respecto al original. En su estudio, Wübben y Wangenheim [47] encontraron que el modelo de Pareto/NBD clasificaba correctamente cerca del 62% del 10% de los clientes reales y al nivel del 20% clasificaba adecuadamente un 67%. Mientras que, el modelo de Pareto/NBD (Abe) presenta una mejor clasificación tanto al 10% como al 20%, tal como lo muestra la tabla 3.

A nivel agregado, se realiza la evaluación en base al Forecast vs Actual (FvA), que es de 108,5% para el período de calibración, lo cual indica un error pequeño del modelo Pareto/NBD (Abe) en comparación con el Pareto/NBD de 128% [38], ya que sobreestima solo un 8,5% el CLV promedio.

TABLA 3  
Métricas de error a nivel individual

		Pareto/ NB (Abe)
Nivel Individual	MAE	39,83
	Media RMSE	72,17
	Mediana RMSE	18,7
	Sensibilidad 10%	70
	Sensibilidad 20%	78
Nivel Agregado	FvA	108,5

### B. Resultados de la estimación del modelo

Se predice el monto en un período de un año iniciando en el mes de mayo del 2019. En base a la frecuencia de compra se realizó un muestreo aleatorio estratificado, donde se tomó una muestra representativa de 500 clientes para realizar el análisis. La tasa de descuento utilizada para estimar el valor actual de los flujos futuros de los clientes fue de 9%. Esta tasa se determinó con la colaboración del departamento del área de finanzas de la empresa que nos proporcionó la información.

La fórmula matemática para descontar los flujos del CLV se define de la siguiente forma:

$$CLV_i = \sum_{t=0}^T \frac{V_{it}}{(1+r)^t} \quad (3.1)$$

Donde:

$V_{it}$  = Compra futura del cliente  $i$  en la semana  $t$ .

$r$  = tasa de descuento semanal.

$t$  = indicativo del período en semanas.

El resultado de  $CLV_i$  corresponde al valor presente de todos los montos que gastará el cliente en el futuro. Los resultados a nivel general son:

TABLA 4  
CLV general

	Promedio	Mínimo	Máximo	Suma Total
CLV	\$ 338,88	\$ -	\$ 3.026,78	\$ 169.438,25

El promedio del CLV es aproximadamente \$339 el cual no es tan bajo considerando un periodo corto de pronóstico de un año y que el supermercado está dirigido a sectores populares. Además, solo se incluyeron los ingresos generados por 10 secciones de la tienda, aquellos guardan la naturaleza de un supermercado, y se excluyeron productos cuyo PVP fuera su superior a USD \$100.

TABLA 5  
Descripción CLV por Segmento de clientes

Segmento	Promedio	Mínimo	Máximo	Suma Total	%	% Según Contribución
Platino	\$ 1.931,63	\$ 1.503,27	\$ 3.026,78	\$ 25.111,13	2,6%	14,82%
Oro	\$ 1.193,29	\$ 1.000,41	\$ 1.468,67	\$ 38.185,28	6,4%	22,54%
Plata	\$ 708,13	\$ 500,94	\$ 988,58	\$ 52.401,41	14,8%	30,93%
Bronce	\$ 141,05	\$ -	\$ 496,16	\$ 53.740,44	76,2%	31,72%
Total	\$ 338,88	\$ -	\$ 3.026,78	\$ 169.438,25	100,0%	100,0%

Dado que se obtuvieron las estimaciones del CLV de todos los clientes de la muestra representativa, se procede a la segmentación de los consumidores en base al valor económico que le generan a la empresa. Al segmentar a los clientes, la empresa puede comprender mejor la importancia de cada cliente para el beneficio total de la empresa. La segmentación basada en los aportes del cliente, generará estrategias de marketing más rentables.

Se clasifican los clientes en cuatro segmentos descritos como: a) los clientes platino, son aquellos clientes con mayor valor para la compañía con un valor estimado de compra superior a \$1500 durante un año; b) oro se refiere a aquellos que generan entre \$1000 y \$1500; c) plata quienes mantiene compras entre \$500 y \$1000 y por último d) los clientes bronce para aquellos con compras menores a \$500 en el año. La tabla 5 resume un breve análisis descriptivo por grupo.

Los clientes Platino, los más valiosos para la compañía, cuentan con un valor promedio de \$1.931,63, estos consumidores representan apenas el 2,6% del total de clientes, un incremento de clientes en este segmento sería una gran oportunidad para la empresa, debido a que les permitiría tener mayor confianza frente al futuro incierto, así mismo obtener mayores niveles de ingresos. Los clientes Oro también tienen poca representatividad un 6,4%, a continuación, los clientes Plata con un CLV promedio de \$703,87 una participación 14,8% y por último los clientes Bronce tienen una mayor participación de 76,2% con un escaso valor económico promedio de \$141,05.

Por otro lado, la contribución económica que generan los grupos, muestra la verdadera rentabilidad que genera cada segmento a la compañía. A pesar de que la representatividad física de los clientes platino (2,6%), oro (6,4%) y plata (14,8%) es apenas un 33%, estos grupos generan una contribución económica superior al 60%, es decir un tercio de los clientes aporta más del 68% de las ventas.

Adicionalmente, es posible caracterizar los segmentos de clientes considerando los datos recopilados por las empresas. En función a los datos disponibles, se pudo conocer que el segmento más valioso, los clientes platino, está conformado principalmente por hombres (61.5%), casados (69%), entre 35 a 50 años (61.5%) y son compradores frecuentes que acuden al menos dos veces por semana (61.5%). En contraste, los menos valiosos, los clientes bronce, son principalmente mujeres (57.8%), solteras (39.7%) entre 35 a 50 años (43.3%) y se categorizan como compradores ocasionales con visitas mínimas de 2 veces al mes (46.7%)



De forma similar, dado que el modelo se estima a nivel individual es posible identificar patrones específicos de compra. Por ejemplo, de los clientes analizados se pudo identificar que el cliente más rentable del establecimiento fue hombre casado de 38 años con un valor promedio de compra de \$19.48 que frecuenta la tienda alrededor de 3 o 4 veces a la semana, su primera sección de preferencia son los comestibles y su compra anual pronosticado es de USD \$3.026,87. Obtener resultados niveles específicos, genera una gran utilidad para la compañía, en virtud de que se puede analizar las características demográficas y el comportamiento de compra de los clientes más rentables para estar en la capacidad de una estratégica asignación de recursos basándose en el valor del cliente para la empresa

## V. CONCLUSIONES

En la actualidad, las compañías se benefician constantemente de los avances en tecnológicos. Con la evolución de las TICs y las estrategias de marketing uno a uno, CRM y marketing de bases de datos, las empresas han generado mejores tácticas para crear relaciones estables con sus clientes. Sin embargo, las empresas frecuentemente fallan en identificar los receptores adecuados para sus estrategias, debido a que segmentan sus clientes en base a criterios poco tangibles como sus características demográficas, geográficas, psicográficas, actitudinales, de uso e inclusive aquellos que incursionan en la segmentación en base al nivel de facturación lo realizan con datos pasados o actuales. En respuesta a ello, este estudio propone una metodología alternativa predictiva que permite calcular la probabilidad de permanencia, frecuencia de compra y el flujo promedio esperado de efectivo de los clientes que en conjunto permiten la estimación del CLV cuyo valor constituye el criterio base para la segmentación.

El modelo probabilístico usado en la estimación fue la extensión Pareto/NBD propuesta por Abe [29] que captura la heterogeneidad de los clientes y muestra mayor flexibilidad en los supuestos. Este modelo tiene cualidades superiores al modelo original Pareto NBD [25] en las estimaciones tanto a nivel global como a nivel individual, ya que posee menores tasas de error respecto a otros estudios realizados en campos similares.

Con base a la estimación del modelo, se realizó la segmentación de clientes con diferentes niveles de CLV en 4 categorías: (a) platino; (b) oro; (c) plata y (d) bronce. Los clientes platino y oro fueron muy leales y altamente rentables. Mientras que, los clientes plata y bronce fueron considerados como clientes de bajo valor. Los resultados confirman que los consumidores tienen un comportamiento heterogéneo, cuyo detalle es útil para realizar un seguimiento de los cambios en el comportamiento a lo largo del tiempo.

La comprensión informada de los clientes clave, ayudaría a los gerentes a formular y cambiar el enfoque estratégico hacia la cartera de clientes que se ajuste a los objetivos de la organización. No obstante, los gerentes y tomadores de decisión deben determinar qué nivel de detalle es conveniente analizar en función a sus objetivos. Por ejemplo, el análisis del nivel de

cliente individual es de útil para determinar diferentes estrategias como la selección de campañas publicitarias o las preferencias de atención al cliente. La aplicación varía según los objetivos comerciales de la empresa en un determinado momento, la cual decidirá incluir o excluir a un grupo determinado de clientes. Por otro lado, el análisis por segmento de clientes permite analizar las características comunes que comparten los consumidores por segmentos de valor para así poder identificar los factores promotores de crecimiento del segmento de cliente. Por último, a nivel general de clientes, las estimaciones de CLV son útiles para una mejor perspectiva en la planificación de negocios y una adecuada gestión estratégica.

En resumen, el CLV es uno de los indicadores más importantes para las empresas en general. A pesar de su subestimación, esta métrica es de uso transversal puesto que contribuye a mejorar diversas actividades comerciales como: marketing, servicio al cliente, administración, finanzas, entre otras. El uso de esta métrica incentiva la competencia en análisis de datos a nivel de negocios, a través del aprovechamiento de la información que ya dispone la empresa. De esta forma, es posible analizar el comportamiento del cliente, ahorrando recursos y esfuerzos, haciendo uso de la información disponible en lugar de realizar inversiones en nuevas investigaciones de mercado.

Para un adecuado uso del CLV se recomienda efectuar una fase previa de experimentación que permita conocer la factibilidad de una campaña segmentada o individualizadas para aprovechar el alto valor esperado de los clientes. Por otro lado, debido a la existencia de múltiples metodologías para la estimación del CLV se sugiere otros estudios que puedan comparar diferentes modelos, otros más recientes y menos conocidos, en búsqueda de un rendimiento superior.

Dada la capacidad predictiva en la estimación, el CLV puede considerarse una medida de rendimiento efectiva para las ventas o ganancias. Por lo tanto, esta métrica podría ser de gran utilidad para las empresas al momento de evaluar la efectividad de las promociones, programas de fidelidad y otros tipos de esfuerzos de marketing. Sin embargo, la metodología empleada en el cálculo no permite determinar los posibles factores causales para el CLV. Por ello, se recomienda que las futuras investigaciones puedan analizar la relación entre la vida del cliente y su rentabilidad.

## REFERENCIAS

- [1] F. Cahen y F. M. Borini, «International digital competence,» *Journal of International Management*, vol. 26, n° 1, p. 100691, 2020. doi: 10.1016/j.intman.2019.100691
- [2] A. Gilchrist, *Industry 4.0: the industrial internet of things*, New York : Apress, 2016. doi: 10.1007/978-1-4842-2047-4

- [3] R. T. Rust y T. S. Chung, «Marketing models of service and relationships,» *Marketing science*, vol. 25, n° 6, pp. 560-580, 2006. doi: 10.1287/mksc.1050.0139
- [4] V. Kumar y A. Pansari, «National culture, economy, and customer lifetime value: Assessing the relative impact of the drivers of customer lifetime value for a global retailer,» *Journal of International Marketing*, vol. 24, n° 1, pp. 1-21, 2016. doi: 10.1509/jim.15.0112
- [5] M. M. Yoo, B. Bai y A. Singh, «Loyalty program effectiveness: Investigating the evolution of behavioral loyalty and Customer Lifetime Value over time,» de *In 8th advances in hospitality and tourism marketing and management (ahtmm) conference*, 2018.
- [6] V. Kumar y D. Shah, Edits., *Handbook of research on customer equity in marketing*, Edward Elgar Publishing, 2015. doi: 10.4337/9781781004982
- [7] Statista, «Statista,» 2017. [En línea]. Available: <https://www.statista.com/markets/423/retail-trade/>.
- [8] P. Jasek, L. Vrana, L. Sperkova, Z. Smutny y M. Kobulsky, «Modeling and Application of Customer Lifetime Value in online Retail,» pp. 2-22, 2018. doi: 10.3390/informatics5010002
- [9] T. Hong y E. Kim, «Segmenting customers in online stores based on factors that affect the customer's intention to purchase,» *Expert Systems with Applications*, vol. 39, n° 2, pp. 2127-2131, 2012. doi: 10.1016/j.eswa.2011.07.114
- [10] A. Parvatiyar y N. Sheth, «Conceptual framework of customer relationship management,» *Emerging concepts tools and applications*, pp. 3-25, 2001. doi: 978-0-07-43504-9
- [11] A. Payne y P. Frow, «A strategic framework for customer relationship management,» *Journal of Marketing*, p. 167-176, 2005. doi: 10.1509/jmkg.2005.69.4.167
- [12] W. Reinartz, M. Krafft y W. Hoyer, «The customer relationship management process: Its measurement and impact on performance,» *Journal of Marketing Research*, p. 293-305, 2004. doi: 10.1509/jmkr.41.3.293.35991
- [13] R. Blattberg, P. Kim y S. Neslin, «Database marketing: analyzing and managing customers,» *New York, NY: Springer*, 2008. doi: 10.1007/978-0-387-72579-6\_6
- [14] R. Thakur y L. Workman, «Customer Portfolio Management (CPM) for Improved Customer Relationship Management (CRM): Are your Customer Platinum, Gold, Silver, or Bronze?,» *Journal of Business Research*, pp. 4095-5102, 2016. doi: 10.1016/j.jbusres.2016.03.042
- [15] R. Blattberg, G. Gary y J. S. Thomas, «Customer Equity: Building and Managing Relationships as Valuable Assets,» *Harvard Business School Press*, 2001.
- [16] L. Berry, «Relationship marketing of services growing interest, emerging perspective,» *Journal of the Academy of Marketing Science*, pp. 236-245, 1995. doi: 10.1177/009207039502300402
- [17] S.-I. Wu y P.-C. Li, «The relationship between CRM, RQ, and CLV based on different hotel preferences,» *International Journal of Hospitality Management*, 2011. doi: 10.1016/j.ijhm.2010.09.011
- [18] A. Avinash, P. Sahu y A. Pahari, «Big Data Analytics for Customer Lifetime Value Prediction,» *Telecom Business Review*, vol. 12, n° 1, p. 46, 2019.
- [19] R. Rust, V. Kumar y R. Venkatesan, «Will the frog change into a prince?: Predicting future customer profitability,» *International Journal of Research in Marketing*, pp. 94-281, 2011. doi: 10.1016/j.ijresmar.2011.05.003
- [20] R. T. Rust, K. N. Lemon y V. A. Zeithaml, «Return on Marketing: Using Customer Equity to Focus Marketing Strategy,» *Journal of Marketing*, pp. 27-109, 2004. doi: 10.1509/jmkg.68.1.109.24030
- [21] R. Venkatesan y V. Kumar, «A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy,» *Journal of Marketing*, pp. 25-106, 2004. doi: 10.1509/jmkg.68.4.106.42728
- [22] J. Romero, R. Lans y B. Wierenga, «A Partially Markov Model of Customer Dynamics for CLV Measurement,» *Journal of Interactive Marketing*, 2013. doi: 10.1016/j.intmar.2013.04.003
- [23] P. Fader y B. Hardie, «Probability Models for Customer-Base,» *Journal of Interactive Marketing*, pp. 61-9, 2009. doi: 10.1016/j.intmar.2008.11.003
- [24] W. J. Reinartz y V. Kumar, «On the Profitability of Long-Life Customers in a Noncontractual Setting: An Empirical Investigation and Implications for Marketing,» *Journal of Marketing*, p. 17-35., 2000. doi: 10.1509/jmkg.64.4.17.18077
- [25] D. C. D. G. M. Schmittlein y R. Colombo, «Counting your Customers: Who Are They and What Will They Do Next?,» *Management Science*, pp. 1-24, 1987. doi: 10.1287/mnsc.33.1.1
- [26] P. Fader, B. Hardie y K. Lee, «Counting your Customers the Easy Way: An Alternative to the Pareto/NBD Model,» *Marketing Science*, pp. 84-275, 2005. doi: 10.1287/mksc.1040.0098
- [27] M. Platzer y T. Reutterer, «Ticking Away the Moments: Timing Regularity Helps to Better Predict Customer Activity,» *Marketing Science*, vol. 35, pp. 693-829, 2016. doi: 10.1287/mksc.2015.0963
- [28] N. Gladys, B. Baesens y C. Croux, «A modified Pareto/NBD approach for predicting customer lifetime value,» *Expert Systems with Applications*, pp. 2062-2071, 2009. doi: 10.1016/j.eswa.2007.12.049

- [29] M. Abe, «“Counting Your Customers” One by One: A Hierarchical Bayes Extension to the Pareto/NBD Model,» *Marketing Science*, p. 541–553, 2009. doi: 10.1287/mksc.1090.0502
- [30] Z. Jamal y A. Zhang, «DMEF Customer Lifetime Value Modeling,» *Journal of Interactive Marketing*, pp. 279–283, 2009. doi: 10.1016/j.intmar.2009.04.004
- [31] D. G. Morrison y D. C. Schmittlein, «Jobs, strikes, and wars: Probability models for duration,» *Organizational Behavior and Human Performance*, pp. 224–251, 1980. doi: 10.1016/0030-5073(80)90065-3
- [32] S. S. S. Borle y J. Dipak, «Customer Lifetime Value Measurement,» *Management Science*, 2007. doi: 10.1287/mnsc.1070.0746
- [33] P. Fader, B. Hardie y K. Lee, «RFM and CLV: Using Iso Value Curves for Customer Base Analysis,» *Journal of Marketing Research*, pp. 30–415, 2005. doi: 10.1509/jmkr.2005.42.4.415
- [34] D. Schmittlein y R. Peterson, «Customer Base Analysis: An Industrial Purchase Process Application,» *Marketing Science*, pp. 41–67, 1994. doi: 10.1287/mksc.13.1.41
- [35] R. Colombo y W. Jiang, «A Stochastic RFM Model,» *Journal of Interactive Marketing*, pp. 2–12, 1999. doi: 10.1002/(SICI)1520-6653(199922)13:3<2::AID-DIR1>3.0.CO;2-H
- [36] W. Reinartz y V. Kumar, «The impact of customer relationship characteristics on profitable lifetime duration,» *Journal of Marketing*, pp. 77–99, 2003. doi: 10.1509/jmkg.67.1.77.18589
- [37] P. Fader, B. Hardie y J. Shang, «Customer Base Analysis in a Discrete Time Noncontractual Setting,» *Marketing Science*, pp. 108–1086, 2010. doi: 10.1287/mksc.1100.0580
- [38] V. Kumar y M. George, «Measuring and maximizing customer equity: A critical analysis,» *Journal of the Academy of Marketing Science*, pp. 157–171, 2007. doi: 10.1007/s11747-007-0028-2
- [39] C. Homburg, M. Droll y D. Totzek, «Customer prioritization: Does it pay off, and how should it be implemented?,» *Journal of Marketing*, pp. 110–130, 2008. doi: 10.1509/jmkg.72.5.110
- [40] E. Gerstner y B. Libai, «Why does poor service prevail?,» *Marketing Science*, pp. 601–603, 2006.
- [41] J. Hogan, K. Lemon y B. Libai, «What is the true value of a lost customer?,» *Journal of Service Research*, p. 196–208, 2003. doi: 10.1177/1094670502238915
- [42] V. R. R. & L. K. Zeithaml, «The customer pyramid creating and serving profitable customers,» *California Management Review*, pp. 118–142, 2001. doi: 10.2307/41166104
- [43] J. Christy, L. Umamakeswari y A. Priyatharsini, «RFM-Ranking And effective approach to customer segmentation,» *Journal of King Saud University - Computer and Informacion Science*, 2018. doi: 10.1016/j.jksuci.2018.09.004
- [44] S.-H. Ma y J.-L. Liu, «The MCMC Approach for Solving the Pareto/NBD Model and Possible Extensions,» de *Third International Conference on Natural Computation (ICNC 2007)*, Haikou, China, 2007. doi: 10.1109/ICNC.2007.728
- [45] B. Nikkhahan, A. H. Badrabad y M. J. Tarokh, «Customer lifetime value model in an online toy store,» *Journal of Industrial Engineering International*, pp. 19–31, 2011.
- [46] M. Platzer, *Stochastic Models of Noncontractual Consumer Relationships*, Vienna University of Economics and Business Administration, 2008.
- [47] M. Wübben y F. Wangenheim, «Instant Customer Base Analysis: Managerial Heuristics Often “Get It Right”,» *Journal of Marketing*, vol. 72, n° 3, p. 82–93, 2008. doi: 10.1509/jmkg.72.3.082
- [48] E. C. Malthouse y R. C. Blattberg, «Can we predict customer lifetime value?,» *Journal of interactive marketing*, vol. 1, n° 2–16, p. 19, 2005. doi: 10.1002/dir.20027
- [49] J. R. Bernat, A. J. Koning y D. Fok, «Modelling Customer Lifetime Value in a Continuous, Non-Contractual Time Setting,» *Business Analytics & Quantitative Marketing*, 2018. Disponible en: <https://pdfs.semanticscholar.org/d713/c994f3141ac0e11734b38c46ce6f7c43f847.pdf>
- [50] K. Do Ruibin y T. Vintilescu Borglöv, «Predicting Customer Lifetime Value: Understanding its accuracy and drivers from a frequent flyer program perspective.,» 2018. Disponible en: <http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1234459&dsid=8609>
- [51] P. Fader, Hardie, B. & Lee y K.L., «RFM and CLV: Using Iso-Value Curves for Customer Base Analysis,» *Journal of Marketing Research*, pp. 415–430, 2005. doi: 10.1509/jmkr.2005.42.4.415
- [52] A. Tanner y W. H. Wong, «The claculation of pesoterior distributions by data augmentation(theory and methods),» 1987, pp. 528–540. doi: 10.1080/01621459.1987.10478458
- [53] M. Wübben y F. Wangenheim, «Instant Customer Base Analysis: Managerial Heuristics Often “Get It Right”,» *Journal of Marketing*, p. 82–93, 2008. doi: 10.1509/jmkg.72.3.082
- [54] R. Hyndman y A. Koehler, «Another look at measures of forecast accuracy,» *International Journal of Forecasting*, p. 679–688, 2006.