



UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

# **DESARROLLO DE UN MÉTODO DE VALORACIÓN DE CLIENTES EN UNA EMPRESA DEL SECTOR AUTOMOTRIZ**

**MEMORIA PARA OPTAR AL TÍTULO DE  
INGENIERO CIVIL INDUSTRIAL**

**ALEX ALFREDO OSSES GODOY**

PROFESOR GUÍA:

RICARDO SAN MARTÍN ZURITA

MIEMBROS DE LA COMISIÓN:

ALEJANDRA PUENTE CHANDÍA

JORGE LARA BACCIGALUPPI

SANTIAGO DE CHILE

2015

## **DESARROLLO DE UN MÉTODO DE VALORACIÓN DE CLIENTES EN UNA EMPRESA DEL SECTOR AUTOMOTRIZ**

En el mundo competitivo de hoy, el conocimiento de los clientes puede transformarse en una ventaja para las empresas. Lo anterior se debe a que cada individuo es diferente y pueden presentar distintos comportamientos, existiendo algunos que son más rentables que otros, por lo que detectarlos y enfocarse en los clientes correctos representa potenciales ganancias a futuro. En base a lo anterior, en este proyecto se propone el desarrollo de una metodología de valoración de clientes de una importante empresa del sector automotriz, específicamente en el subsector de camiones medianos de una de las marcas que representa.

El método propuesto tiene como objetivo apoyar a las áreas comerciales a focalizar la retención y fidelización de los clientes más valiosos mediante la generación de recomendaciones de distintos tipos de acciones de marketing a enfocar en diferentes grupos de clientes para así mejorar su gestión. Lo anterior implica la estimación del valor futuro de los clientes, para lo cual se utiliza la métrica de CLV (Customer Lifetime Value). Para estimarla, se proponen 2 modelos ampliamente utilizados en distintas industrias para la estimación del número de transacciones: Uno probabilístico (BG-NBD) y uno econométrico (Logit-Poisson-Markov). Por otro lado, para la estimación de los montos se utiliza el modelo probabilístico Gamma-Gamma.

Las métricas de ajuste utilizadas para validar los modelos indican que el modelo econométrico es el que presenta el mejor desempeño para la estimación del número de transacciones, el cual se utiliza en conjunto con el modelo probabilístico que estima los montos para así proyectar el valor futuro de los clientes.

Utilizando el valor histórico generado por los clientes y su valor proyectado para el año actual y para un plazo de 5 años (utilizando el modelo desarrollado) se propone una agrupación de clientes en base a estas 3 variables. A partir de la agrupación propuesta se realizan recomendaciones de acciones de marketing a los distintos grupos generados, logrando así apoyar a las áreas comerciales a la focalización de los esfuerzos a realizar con la cartera de clientes.

Los clientes más valiosos para el plazo de 5 años utilizados resultan ser los clientes tipo "Empresa" que se desempeñan en el rubro "R2" y que además de tener vehículos medianos poseen vehículos livianos y/o pesados dentro de la misma marca. Estos representan un valor promedio de \$ 172.735.837 en el plazo señalado.

El modelo de valoración y su utilización propuesta significan el punto de partida para la empresa en la búsqueda de un enfoque relacional, que busca comprender, mejorar y aumentar las relaciones que posee actualmente con su cartera de clientes.

## **AGRADECIMIENTOS**

Agradezco enormemente a mi familia por todo el apoyo incondicional que me brindaron durante esta hermosa etapa vivida. Los sacrificios realizados para facilitarme el proceso fueron enormes y sin ellos no puedo imaginar haber podido pasar satisfactoriamente este gran desafío.

A mis amigos que siempre estuvieron en los buenos y malos momentos vividos a largo de estos años. En particular a los amigos formados dentro de la Universidad que siempre fueron un apoyo en los momentos de estrés que frecuentemente viví durante la vida universitaria.

A todos los compañeros con los que alguna vez me tocó trabajar y quizás no tuve la oportunidad de conocer más profundamente, gracias por el esfuerzo entregado en cada trabajo.

A la empresa que me abrió las puertas para realizar mi trabajo de título. A mis jefes y compañeros de trabajo por su apoyo, sus consejos y los buenos momentos vividos a lo largo del proceso.

Finalmente, debo agradecer a los profesores integrantes de mi comisión, quienes aportaron distintas visiones al trabajo y lograron ayudarme enormemente, en especial en las etapas iniciales del trabajo. Gracias por sus críticas constructivas, por su buena disposición y la motivación y consejos que me brindaron a lo largo de este año.

# Tabla de Contenido

<b>I.-Introducción</b> .....	<b>1</b>
<b>I.1.-Antecedentes Generales</b> .....	<b>1</b>
<b>I.2.-Descripción del proyecto y justificación</b> .....	<b>5</b>
<b>I.3.-Objetivos</b> .....	<b>9</b>
<b>I.4.-Alcances</b> .....	<b>9</b>
<b>I.5.-Resultados esperados</b> .....	<b>10</b>
<b>II.-Marco Conceptual</b> .....	<b>11</b>
<b>II.1.-Cadenas de Markov</b> .....	<b>11</b>
<b>II.2.- Regresión Logística</b> .....	<b>11</b>
<b>II.3.-Regresión de Poisson Truncada en Cero.</b> ....	<b>12</b>
<b>II.4.-Modelos de CLV</b> .....	<b>13</b>
II.4.1.-Horizonte de tiempo (T): .....	14
II.4.2.-Número de Transacciones ( $Y_{jt}$ ): .....	14
II.4.2.1-Modelos Probabilísticos:.....	15
II.4.2.1.1.-Modelo BG/NBD:.....	15
II.4.2.2-Cadenas de Markov: .....	16
II.4.2.2.1.-Probabilidad de compra ( $P_{jt}$ ) .....	16
II.4.2.2.2.-Número esperado de transacciones condicionado ( $n_{jt}$ ).....	18
II.4.3.-Montos por transacción ( $m_{jt}$ ).....	19
II.4.3.1.-Modelo Gamma-Gamma .....	19
<b>II.5.-Validación de Modelos</b> .....	<b>20</b>
<b>II.6.-Significancia de variables</b> .....	<b>22</b>
II.6.1.-Test Anova (Análisis de varianza) .....	22
II.6.2.-Test de Wald .....	22
<b>III.-Descripción de la Metodología</b> .....	<b>23</b>
<b>III.1.-Descripción de la Data Disponible</b> .....	<b>23</b>
<b>III.2.-Análisis del Comportamiento Histórico</b> .....	<b>23</b>
<b>III.3.-Selección de Variables Relevantes.</b> .....	<b>23</b>
<b>III.4.-Aplicación y Validación de los Modelos.</b> .....	<b>24</b>
<b>III.5.-Propuesta de uso del Modelo</b> .....	<b>24</b>
<b>IV.-Desarrollo Metodológico</b> .....	<b>25</b>
<b>IV.1.-Descripción de la data disponible</b> .....	<b>25</b>
IV.1.1.-Data transaccional .....	25
IV.1.2.-Data de perfilamiento .....	26
IV.1.3.-Data Macroeconómica .....	27
IV.1.4.Limpieza, Pre-Procesamiento y Transformación de datos. ....	27
IV.1.4.1.-Variables de Perfilamiento .....	28
IV.1.5.-Descomposición de la data transaccional en observaciones por cliente .....	30
<b>IV.2.-Análisis del comportamiento histórico.</b> .....	<b>32</b>
IV.2.1.-Determinación del valor histórico de los clientes de la empresa. ....	32

IV.2.2.-Relación entre las variables de comportamiento y las variables de perfilamiento.....	33
IV.2.2.1.-Número de Transacciones ( $x^*$ ), Montos por Transacción ( $m^*$ ) y variables de perfilamiento.....	33
IV.2.2.2.-Decisión de compra ( $Ct$ ), decisión del número de transacciones condicionada ( $xt^{***}$ ) y su relación con variables de comportamiento. ....	41
IV.2.2.3.-El efecto del PIB .....	45
<b>IV.3.-Selección de variables relevantes.....</b>	<b>46</b>
IV.3.1.- Test ANOVA .....	47
IV.3.2.- Test Chi-Cuadrado .....	48
<b>IV.4.-Aplicación y validación de los modelos.....</b>	<b>49</b>
IV.4.1.-Periodo Calibración-Validación .....	49
IV.4.2.-Estimación del número de transacciones.....	49
IV.4.2.1.-Modelo BG/NBD .....	49
IV.4.2.2.-Modelo Logit-Poisson-Markov (L-P-M).....	52
IV.4.2.4.-Selección del modelo para la estimación del número de transacciones .....	56
IV.4.3.-Estimación de los montos por transacción.....	57
IV.4.3.1.-Modelo Gamma-Gamma .....	57
IV.4.4.-Estimación del valor de los clientes.....	58
IV.4.4.1.-Validación de la capacidad de detección de valor del modelo .....	59
IV.4.5.-Aplicación del modelo a toda la cartera de clientes.....	60
<b>IV.5.-Propuesta de uso del modelo.....</b>	<b>62</b>
IV.5.1.-Recomendaciones de acciones al área de Marketing .....	62
IV.5.1.1.-Acciones a corto plazo .....	62
IV.5.1.2.-Acciones a mediano-largo plazo .....	62
IV.5.2.-Agrupación de Clientes.....	63
IV.5.3.-Recomendación de acciones a los distintos grupos de clientes .....	65
IV.5.4.-Evaluación del impacto de la propuesta .....	66
<b>V.-Conclusiones Finales.....</b>	<b>69</b>
V.1.-Conclusión del trabajo .....	69
V.2.-Trabajos Futuros.....	70
<b>VI.-Bibliografía .....</b>	<b>71</b>
<b>VII.-Anexos.....</b>	<b>73</b>
Anexo A: Lista de rubros disponibles y su descripción .....	73
Anexo B: Frecuencia de clientes según la región a la que pertenecen .....	74
Anexo C: Frecuencia de clientes según su Rubro.....	74
Anexo D: Frecuencia de observaciones del número de transacciones realizadas por los clientes en los años que deciden comprar.....	75
Anexo E: Frecuencia de observaciones de acuerdo a decisión de compra y recencia .....	75
Anexo F: Resultados Anova para la variable Monto por Transacción.....	76
Anexo G: Cambios en las matrices de transición entre un año y otro .....	76
Anexo H: Predicción Acumulada a 2018 considerando los perfiles formados por las variables “Naturaleza”, “Rubro” y “Tiene Vehículo Otro Segmento” .....	77

## Índice de Ilustraciones

<i>Ilustración 1: Ciclo de vida de un cliente.....</i>	<i>2</i>
<i>Ilustración 2: Evolución de las ventas de camiones en Chile .....</i>	<i>3</i>
<i>Ilustración 3: Evolución de las ventas de buses en Chile.....</i>	<i>3</i>
<i>Ilustración 4: Organigrama de la división de camiones y buses. ....</i>	<i>5</i>
<i>Ilustración 5: Comportamiento de clientes adquiridos de la Marca A en la línea de camiones medianos.....</i>	<i>7</i>
<i>Ilustración 6: Promedio de “Monto gastado/Años distintos de compra” de Clientes marca A según su Naturaleza....</i>	<i>8</i>
<i>Ilustración 7: Promedio de “Monto gastado/Años distintos de compra” de Clientes marca A según su Tamaño de Flota.....</i>	<i>8</i>
<i>Ilustración 8: Matriz de Agrupación de clientes para la recomendación de acciones.....</i>	<i>9</i>
<i>Ilustración 9: Representación de estados y su conexión.....</i>	<i>17</i>
<i>Ilustración 10: Matriz de probabilidades de transición entre estados para la estimación de la probabilidad de compra .....</i>	<i>18</i>
<i>Ilustración 11: Matriz de probabilidades de transición entre estados para la estimación del número de transacciones .....</i>	<i>19</i>
<i>Ilustración 12: Representación de curva ROC.....</i>	<i>21</i>
<i>Ilustración 13: Evolución de las decisiones de compra en el tiempo .....</i>	<i>31</i>
<i>Ilustración 14: Evolución del promedio de la decisión del número de transacciones .....</i>	<i>32</i>
<i>Ilustración 15: Evolución del promedio de Transacciones según Naturaleza y Antigüedad.....</i>	<i>34</i>
<i>Ilustración 16: Evolución del Monto promedio por Transacción según Naturaleza y Antigüedad .....</i>	<i>35</i>
<i>Ilustración 17: Evolución del promedio de Transacciones según Rubro y Antigüedad del cliente.....</i>	<i>36</i>
<i>Ilustración 18: Evolución del Monto promedio por Transacción según Rubro y Antigüedad del cliente .....</i>	<i>36</i>
<i>Ilustración 19: Evolución del promedio de Transacciones según Zona y Antigüedad del cliente .....</i>	<i>37</i>
<i>Ilustración 20: Evolución del Monto promedio por Transacción según Zona y Antigüedad del cliente.....</i>	<i>38</i>
<i>Ilustración 21: Evolución del promedio de Transacciones según Tamaño de flota y Antigüedad del cliente .....</i>	<i>39</i>
<i>Ilustración 22: Evolución del Monto promedio por Transacción según Tamaño de flota y Antigüedad del cliente .....</i>	<i>39</i>
<i>Ilustración 23: Evolución del promedio de Transacciones según “Vehículos de otro segmento” y Antigüedad.....</i>	<i>40</i>
<i>Ilustración 24: Evolución del Monto promedio por Transacción según Si tiene vehículos de otro segmento y Antigüedad del cliente .....</i>	<i>41</i>
<i>Ilustración 25: Evolución de las Tasas de recompra según Naturaleza de los clientes.....</i>	<i>42</i>
<i>Ilustración 26: Evolución de las Tasas de recompra según el Rubro de los clientes .....</i>	<i>43</i>
<i>Ilustración 27: Evolución de las Tasas de recompra según si el cliente tiene o no vehículos de otro segmento .....</i>	<i>44</i>
<i>Ilustración 28: Relación entre el incremento del PIB y la decisión de compra.....</i>	<i>46</i>
<i>Ilustración 29: Relación entre el Incremento del PIB Y el Número promedio de transacciones .....</i>	<i>46</i>
<i>Ilustración 30: Curva ROC probabilidades de compra 2012 modelo L3.....</i>	<i>54</i>
<i>Ilustración 31: Curva ROC probabilidades de compra 2013 modelo L3.....</i>	<i>54</i>
<i>Ilustración 32: Boxplot del CLV .....</i>	<i>61</i>
<i>Ilustración 33: Recomendación de acciones por grupos generados en base a las 3 variables de valor .....</i>	<i>66</i>
<i>Ilustración 34: Matriz de Transición para el año 2012 de clientes Tipo Empresa, Rubro R7 y que poseen vehículos de otros segmentos. ....</i>	<i>76</i>
<i>Ilustración 35: Matriz de Transición para el año 2013 de clientes Tipo Empresa, Rubro R7 y que poseen vehículos de otros segmentos. ....</i>	<i>77</i>

## Índice de tablas

Tabla 1: Clasificación de camiones según su peso.....	2
Tabla 2: Resumen de datos transaccionales.....	25
Tabla 3: Variables transaccionales por cliente.....	26
Tabla 4: Frecuencia según Naturaleza del cliente.....	28
Tabla 5: Frecuencia según Tamaño de flota.....	28
Tabla 6: Frecuencia según Zona.....	29
Tabla 7: Frecuencia según nuevo rubro.....	29
Tabla 8: Frecuencia según vehículos de otro segmento del cliente.....	30
Tabla 9: Resumen decisiones de compra.....	30
Tabla 10: Resumen decisión del número de transacciones.....	31
Tabla 11: Resultados Valor Histórico.....	33
Tabla 12: Transacciones y Montos por Transacción según Naturaleza del cliente.....	34
Tabla 13: Transacciones y Montos por Transacción según Rubro del cliente.....	35
Tabla 14: Transacciones y Montos por Transacción según Zona Geográfica del cliente.....	37
Tabla 15: Transacciones y Montos por Transacción según Categoría del cliente.....	38
Tabla 16: Transacciones y Montos por Transacción según “vehículos de otro segmento”.....	40
Tabla 17: Resumen de la decisión de compra según la Naturaleza del cliente en la observación.....	41
Tabla 18: Resumen de la decisión de compra según el Rubro del cliente en la observación.....	42
Tabla 19: Resumen de la decisión de compra según si el cliente tiene Vehículos de otro segmento.....	43
Tabla 20: Resumen decisión de compra según la recencia del cliente en la observación.....	44
Tabla 21: Resumen de decisión de compra según la cantidad de años distintos en que el cliente ha comprado al momento de la observación.....	45
Tabla 22: Resumen de decisión del número de transacciones a realizar según la decisión tomada por el cliente el año anterior.....	45
Tabla 23: Resultados ANOVA variable Número de Interacciones.....	47
Tabla 24: Resultados chi-cuadrado para la variable decisión de compra.....	48
Tabla 25: Resultados Chi-cuadrado evolución transacciones.....	49
Tabla 26: Parámetros obtenidos en la estimación del modelo BG/NBD.....	50
Tabla 27: Resultados y errores BG/NBD proyección año 2012.....	51
Tabla 28: Resultados y errores BG/NBD proyección año 2013.....	51
Tabla 29: Resultados y errores BG/NBD proyección acumulada años 2012-2013.....	51
Tabla 30: Coeficientes modelo logístico L 3.....	53
Tabla 31: Resultados Test de Wald para la variable Rubro.....	53
Tabla 32: Resultados y errores L-P-M proyección año 2012.....	55
Tabla 33: Resultados y errores L-P-M proyección año 2013.....	55
Tabla 34: Resultados y errores L-P-M proyección acumulada años 2012-2013.....	55
Tabla 35: Comparación BG/NBD 3 y LPM 3.....	56
Tabla 36: Comparación Predicción Acumulada por perfiles BG/NBD 3 y L-P-M 3.....	57
Tabla 37: Métricas de ajuste para la estimación de los montos por transacción.....	58
Tabla 38: Métricas de ajuste para la predicción del valor generado durante el periodo de validación.....	58
Tabla 39: Predicción acumulada en el periodo de validación según perfil “Naturaleza”-“Rubro”.....	59
Tabla 40: Comparación Predicción Valor 2012.....	60
Tabla 41: Comparación Predicción Valor 2013.....	60
Tabla 42: Predicción CLV acumulado a 2018 a nivel de Perfil “Naturaleza”-“Rubro”.....	61
Tabla 43: Valor Promedio para grupos generados con el Valor Histórico hasta el 2014.....	64
Tabla 44: Valor Promedio para grupos generados con el Valor Proyectado para el año 2015.....	64
Tabla 45: Valor Promedio para grupos generados con el Valor acumulado proyectado entre el 2015 y 2018.....	64
Tabla 46: Movimientos de clientes entre grupos Considerando el Valor histórico y el Proyectado acumulado al 2018.....	65
Tabla 47: Recomendación de acciones por grupos según Valor histórico y valor proyectado al 2018.....	65
Tabla 48: Movimiento de clientes entre grupos Considerando las 3 variables de valor.....	66
Tabla 49: Lista de rubros y su descripción.....	73

<i>Tabla 50: Frecuencia de clientes según su región.....</i>	<i>74</i>
<i>Tabla 51: Frecuencia de clientes según su rubro .....</i>	<i>74</i>
<i>Tabla 52: Frecuencia de observaciones del número de transacciones realizadas por los clientes que deciden comprar .....</i>	<i>75</i>
<i>Tabla 53 Frecuencia de observaciones según decisión de compra y recencia real.....</i>	<i>75</i>
<i>Tabla 54: Resumen Predicción CLV acumulado a 2018 a nivel de Perfil “Naturaleza”-“Rubro”-“Vehículo de otro segmento” .....</i>	<i>78</i>

# I.-Introducción

## I.1.-Antecedentes Generales

### Gestión de Clientes:

La alta competencia que experimentan las empresas hoy en día ha generado que estas comiencen a implementar estrategias con un enfoque cada vez más relacional. Estas estrategias se engloban bajo el concepto de *Customer Relationship Management* (CRM), entendido como una estrategia de negocio en la cual los clientes se transforman en el foco de la compañía, transformándose en recursos claves de la misma. [1]

El desarrollo de la tecnología ha facilitado la aplicación de este tipo de estrategias en las empresas, a través de distintas herramientas que tienen diferentes objetivos y funcionalidades. En base a lo anterior se pueden encontrar 4 tipos de aplicaciones de estrategias CRM: [2]

- CRM estratégico: Se enfoca en la creación de una cultura de negocio centrada en el cliente con el fin de invertir de mejor forma los recursos de la empresa para entregar valor.
- CRM operacional: Busca la automatización de los procesos en los cuáles se interactúa con los clientes, donde se incluyen tareas de marketing, ventas y servicios.
- CRM analítico: Se construye sobre el CRM operacional y analizan la data generada por los clientes con el fin de crear información para la segmentación de clientes, análisis de comportamiento y del valor que estos representan para la compañía.
- CRM colaborativo: Tiene como objetivo la integración de los clientes utilizando distintos canales de interacción.

El presente trabajo se enfoca en el CRM analítico, buscando una mejor comprensión del comportamiento y valor de los clientes de una empresa del sector automotriz. Por esta razón es importante comprender el ciclo de vida de la relación cliente-empresa, el que se pueden resumir en: [3]

- Adquisición del cliente: Se basa principalmente en la identificación de potenciales clientes, sus necesidades y la forma de llegar a ellos.
- Desarrollo de la relación con clientes: Se busca ofrecer los productos y/o servicios adecuados a las necesidades de los clientes.
- Retención del cliente: El objetivo es fidelizar al cliente para que éste continúe su relación con la empresa.

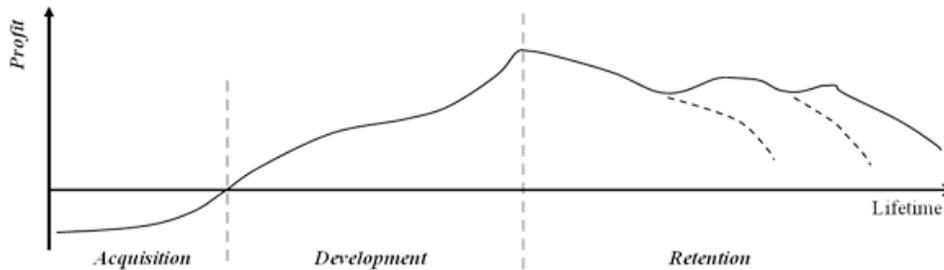


Ilustración 1: Ciclo de vida de un cliente<sup>1</sup>

Dentro de los objetivos de un CRM se encuentran la maximización de las utilidades generadas por los clientes mediante la disminución de los costos de adquisición, el aumento de los ingresos a lo largo de la relación cliente-empresa, y la extensión de la misma.

Para lograr lo anterior, el CRM analítico cumple un rol fundamental. Debido a que no todos los clientes se comportan de la misma forma, estos representan distinto valor para la empresa, por lo tanto, dependiendo del tipo de cliente y su comportamiento histórico y potencial, las empresas pueden enfocar de mejor forma sus esfuerzos en las 3 etapas mencionadas para el ciclo de vida de la relación cliente-empresa. [4]

En este contexto, analizar el valor tanto de los clientes como de los prospectos es una tarea fundamental para la cual existen distintos métodos que, utilizando la información generada por los clientes, buscan determinar el valor histórico y el valor potencial de los mismos. Ejemplos de lo anterior son el RFM (Recency, Frequency, Monetary) y CLV (Customer Lifetime Value) que se han utilizado en diversas industrias.

### **Sector Automotriz en Chile – Vehículos Pesados:**

Como ya se mencionó, el trabajo se desarrolla en una empresa del sector automotriz, específicamente en el subsector de vehículos pesados donde se concentra la representación y distribución de distintas marcas de camiones de diferentes tipos (Ver “Tabla 1”) y buses que se comercializan en Chile.

	PBV (kgs)	
	Desde	Hasta
Liviano	3.860	6.350
Mediano	6.351	14.968
Pesado	14.969	->

Tabla 1: Clasificación de camiones según su peso

El mercado de camiones en Chile vivió su “peak” durante los años 2011 y 2012 donde se registraron ventas de 18505 y 18309 unidades respectivamente. Estas ventas

<sup>1</sup> Fuente: <http://www.ieseg-databasemarketing.com/about-database-marketing/database-marketing-customer-lifecycle/>

históricas se vieron impulsadas principalmente por el dinamismo que experimentaron distintos sectores económicos, principalmente la minería, la construcción y el comercio exterior, los que fortalecieron la demanda de este tipo de vehículos durante el periodo en cuestión. Durante el 2013 se experimentó una importante caída en la cantidad de vehículos vendidos (14753 unidades) la que se debió a factores como la desaceleración de la inversión privada en maquinarias, a la postergación de proyectos de la minería, la escases de choferes lo que limita la incorporación de nuevas unidades a las flotas, los grandes proyectos de inversión que se han pospuesto y la variación de los precios al incorporarse tecnologías de mayor costo. Estos mismos factores hacen que para este año 2014 se esperen resultados a la baja en comparación al 2013, proyectándose ventas de 14.500 unidades en el sector. [5]

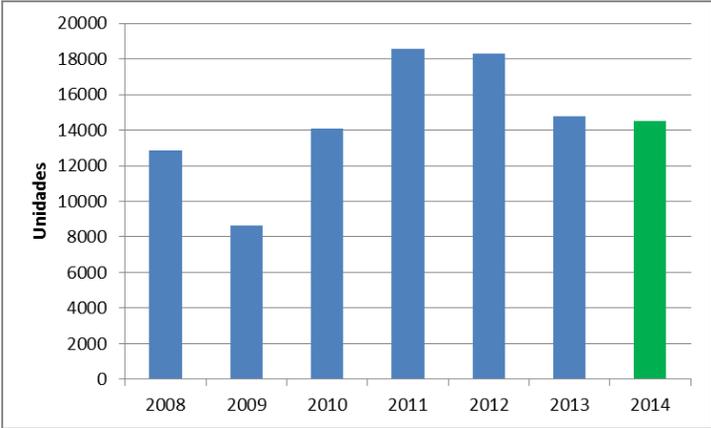


Ilustración 2: Evolución de las ventas de camiones en Chile<sup>2</sup>

Respecto al mercado de buses, este ha mostrado un comportamiento similar al de camiones, experimentando su “peak” en el año 2012 cuando se vendieron 5196 unidades. Este gran nivel de ventas fue influenciado fuertemente por la licitación de buses de Transantiago que se realizó ese año.

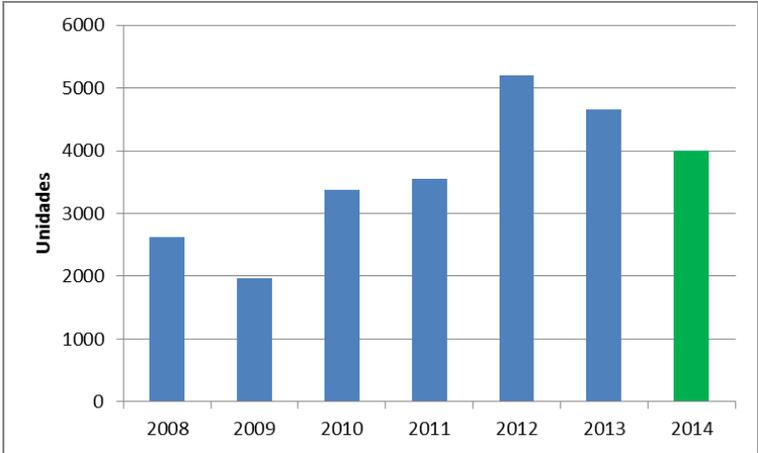


Ilustración 3: Evolución de las ventas de buses en Chile<sup>3</sup>

<sup>2</sup> Fuente: ANAC

<sup>3</sup> Fuente: ANAC

## **Gestión de clientes en el sector automotriz:**

Al igual que en muchas industrias, el concepto de CRM tiene mucha relevancia en la industria automotriz. Muchas empresas de este rubro utilizan estrategias de este tipo con el objetivo de atraer nuevos clientes, aumentar la lealtad a la marca, reducir costos, incrementar eficiencia y lograr mantener una ventaja competitiva. [6]

Respecto al concepto del valor del cliente en esta industria, se han realizado estudios en donde se aplican métodos de RFM y CLV, especialmente en lo que respecta a la industria de vehículos livianos. Un ejemplo de lo anterior es el estudio realizado por Chan el 2008 [3] en el cual se propone una combinación de ambos métodos para decidir en qué clientes enfocar determinados tipos de campaña utilizando data de la marca Nissan<sup>4</sup>. Otro ejemplo dentro de la industria de vehículos livianos se encuentra en el trabajo realizado por C.-J Chenga, S.W. Chiub, C.-B. Chengc y J.-Y. Wuc el 2011 [7], en donde utilizan distintas técnicas de data mining para modelar y predecir el comportamiento de los clientes de un servicio técnico de la marca Nissan.

Dentro del subsector de vehículos pesados, la cantidad de estudios es menor. Sin embargo, destaca el caso de Caterpillar [8] quienes, utilizando su base de datos tanto de clientes como de prospectos, aplicaron métodos de CLV y CHAID (CHi-squared Automatic Interaction Detection) para construir un ranking de clientes y prospectos que permitieron a las áreas de Venta y Marketing enfocar mejor sus esfuerzos. Esto aumentó un 5% el Market Share de la marca en una época en que el mercado iba a la baja.

---

<sup>4</sup> El problema de este estudio, es que no especifican los modelos utilizados para pronosticar el comportamiento futuro de los clientes ni el impacto que tienen las campañas de Marketing en ellos.

## I.2.-Descripción del proyecto y justificación

### Descripción del problema

Como se ha mencionado, el proyecto se desarrolla en la división de camiones y buses de una empresa del sector automotriz. Específicamente, el trabajo a desarrollarse afecta al área de gestión de clientes de la división, la cual es un área transversal a las 5 marcas que esta representa.

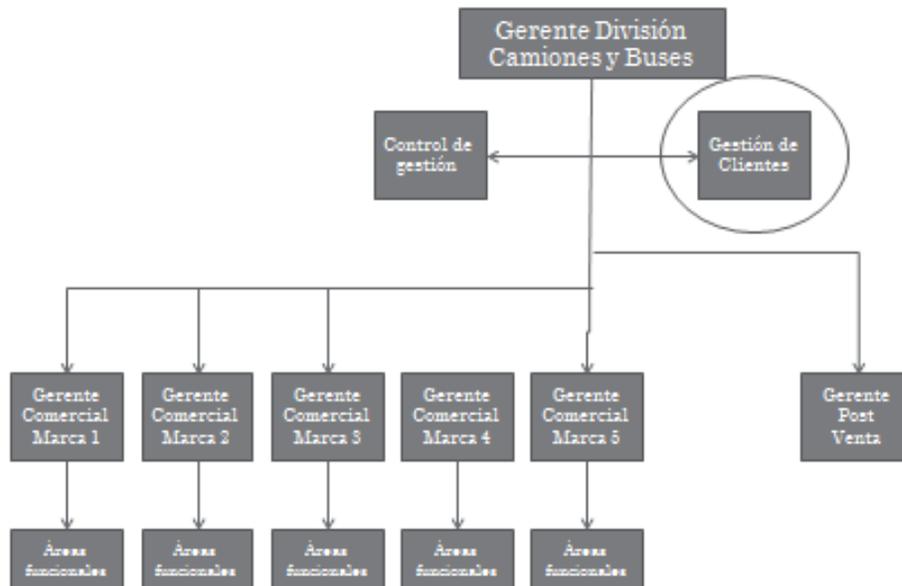


Ilustración 4: Organigrama de la división de camiones y buses<sup>5</sup>.

El objetivo central del área es impartir una cultura CRM dentro de la división, para lo cual tiene una serie de funciones dentro de las cuáles las que se intentan mejorar en este trabajo corresponden a la “Gestión de interacciones entre cliente-empresa” y el “Apoyo a las Áreas de Marketing de las distintas marcas”.

Para comprender el problema en el cual se enfoca el trabajo es importante mencionar una serie de hitos que han afectado a su generación:

- El Área de gestión de clientes se crea a principios del 2010 debido a que en los años anteriores el mercado venía a la baja y la compañía deseaba comenzar a cuidar más a sus clientes. Previo a esto, no se le daba mucha importancia al tema de gestión de clientes.
- Durante los años 2010 y 2012 el mercado se recupera y vive sus mejores años<sup>6</sup>. Como el área de Gestión de Clientes se estaba recién formando, las marcas siguieron sin darle importancia al tema, avalados por los buenos resultados que

<sup>5</sup> Fuente: Elaboración propia.

<sup>6</sup> Ver Ilustraciones 2 y 3.

se estaban experimentando. Lo anterior indica que este éxito en las ventas generó un problema para la penetración de esta nueva área en la empresa.

- Durante el año 2013 el mercado vuelve a experimentar una baja en sus ventas y como los pronósticos para el 2014 no eran auspiciosos, la empresa comienza a preocuparse nuevamente por la gestión de su cartera de clientes.

Lo anterior genera que la empresa se plantee las siguientes preguntas:

- ¿En qué clientes debemos enfocarnos?
- ¿Cuáles han sido nuestros clientes más rentables?
- ¿Cuáles están más propensos a interactuar próximamente con la empresa?
- ¿Cuáles representan el mayor valor de cara al futuro?
- ¿Tienen algún perfil particular?

La incapacidad por parte de la empresa de responder las preguntas anteriores da origen al problema, el cual se puede resumir en:

“Poco conocimiento de la cartera de clientes impide enfocar correctamente los esfuerzos de las áreas de Marketing y Venta”.

### **Justificación del problema:**

Actualmente las marcas que conforman la división de camiones y buses se enfocan en el Marketing masivo, el cual tiene un mayor impacto o está más enfocado en la adquisición de clientes que en la retención [9]. Lo anterior ha llevado a tasas de retención por debajo del promedio<sup>7</sup>, tal como se observa en la “Ilustración 5”, donde se muestra el comportamiento futuro de los clientes adquiridos año a año de los clientes de una de las marcas de la compañía (De ahora en adelante “**Marca A**”) en la línea de negocio de camiones medianos<sup>8</sup>.

Considerando que los clientes deberían generar una nueva compra en un plazo máximo de 5-6 años<sup>9</sup>, se tiene que aquellos que han comprado por primera vez en los años 2007 y 2008 ya han experimentado este plazo. Como se observa en la “Ilustración 5”, en los años 2007 y 2008 se adquirieron 145 y 101 nuevos clientes respectivamente, de los cuales cerca de un 21% ha vuelto a comprar posteriormente. Se tiene además que los clientes del 2007 que volvieron a comprar generaron, en términos de facturación, un promedio de \$ 69.072.311 entre los años 2008 y 2013, mientras que los clientes del 2008 generaron un promedio de \$109.113.576 entre los años 2009 y 2013. Con esto, se estima que si la tasa de retención de estos clientes hubiese sido, por ejemplo, un 5%

<sup>7</sup> En promedio, esta tasa en la industria es superior al 50%. Fuente: (2012) <http://www.etrucker.com/2004/09/14/truck-brand-loyalty-rising>  
<http://automotivedigest.com/2013/09/j-d-power-study-truck-owners-are-brand-loyal>

<sup>8</sup> Posteriormente, en la sección de alcances, se justificará la elección de esta marca y de esta línea de negocio.

<sup>9</sup> Según expertos de la industria, este debería ser el plazo de renovación promedio de los vehículos de este tipo. Es posible que se produzcan compras antes, debido a ampliaciones de flota o renovaciones de vehículos de otras marcas, sin embargo los 5-6 años son un “límite máximo a considerar”

mayor (7 clientes para el año 2007 y 5 clientes para el año 2008), se hubiesen obtenido \$ 1.051.797.812 durante los años 2008 y 2013.

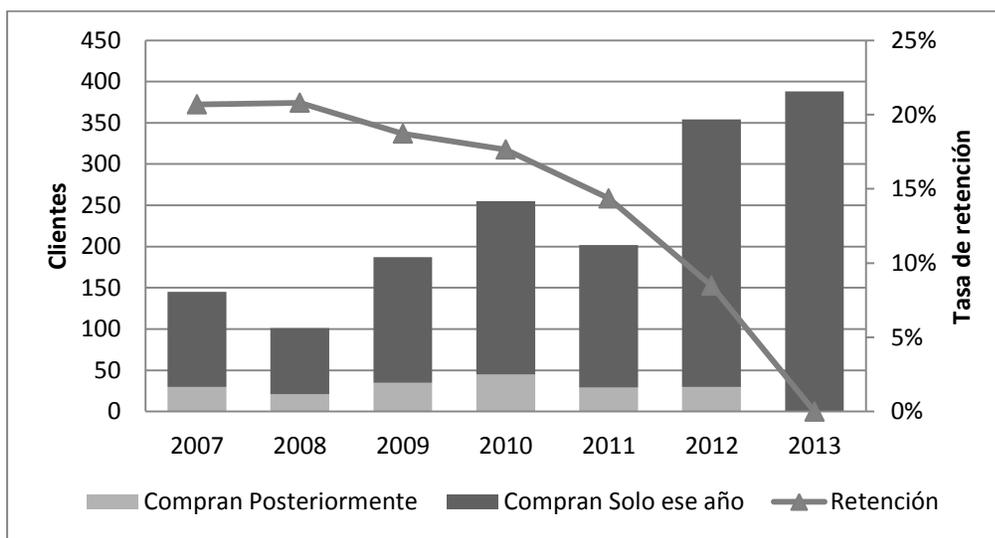


Ilustración 5: Comportamiento de clientes adquiridos de la Marca A en la línea de camiones medianos<sup>10</sup>

La buena relación calidad/precio que posee la marca en el segmento de camiones medianos<sup>11</sup>, junto con las aceptables tasas de Satisfacción de Clientes en el proceso de Venta y Servicio<sup>12</sup>, permiten inferir que las bajas tasas de retención se deben posiblemente a la falta de gestión de la cartera de clientes. Por lo tanto es necesario un cambio de enfoque en lo que respecta al marketing, con el fin de lograr orientar de mejor forma sus esfuerzos para obtener mejores resultados.

Para lograr lo anterior se propone comprender de mejor forma el comportamiento de los clientes y el valor que estos representan para la empresa. Sin embargo.... ¿El comportamiento de los clientes es heterogéneo? Si la respuesta a esta pregunta fuese negativa, significaría que todos los clientes serían igual de valiosos, por lo tanto deberían ser tratados de manera homogénea.

Como se puede apreciar en la “Ilustración 6” y la “Ilustración 7”, analizando por separado 2 variables de comportamiento (Naturaleza del cliente<sup>13</sup>, y Tamaño de Flota<sup>14</sup>) y una de comportamiento (Promedio de “Monto Gastado/Años Distintos de Compra”) se aprecia que existe heterogeneidad en el comportamiento de distintos grupos de clientes.

<sup>10</sup> Fuente: Elaboración propia.

<sup>11</sup> Estudio de Mercado realizado por una empresa externa el 2013 que muestra que la marca A tiene la segunda mejor relación calidad precio del segmento.

<sup>12</sup> 9.1 y 8.3 de un máximo de 10 respectivamente, como promedio anual para el año 2013

<sup>13</sup> Se refiere a si el cliente es una persona natural o una empresa

<sup>14</sup> Existen 4 categorías: Particular (hasta 2 vehículos), Flotista Pequeño (3-5 vehículos), Flotista Mediano (6-15 vehículos) y Flotista Grande( 16 o más vehículos)

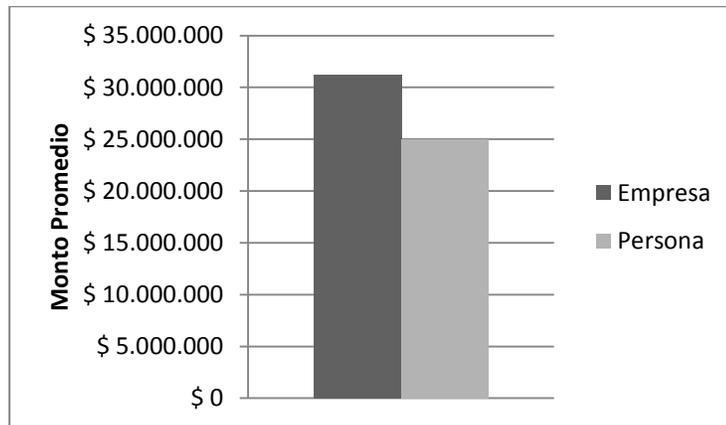


Ilustración 6: Promedio de “Monto gastado/Años distintos de compra” de Clientes marca A según su Naturaleza

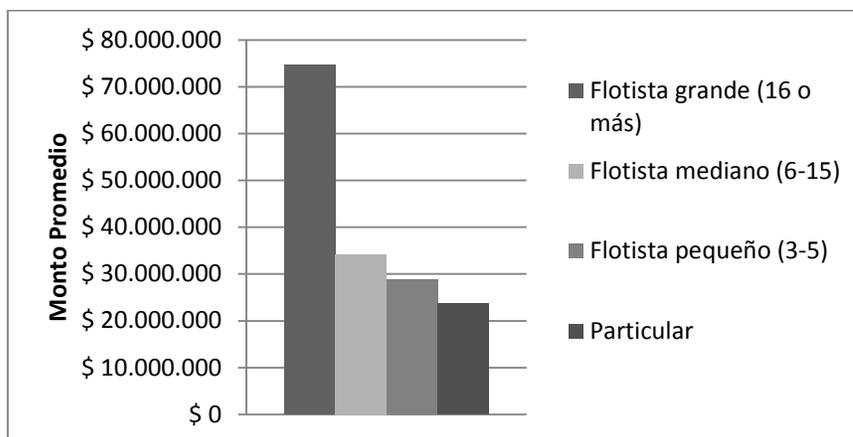


Ilustración 7: Promedio de “Monto gastado/Años distintos de compra” de Clientes marca A según su Tamaño de Flota

Lo anterior indica que al parecer existen clientes que han generado mayor valor en comparación a otros, por lo tanto se justifica la aplicación de métodos que permitan determinar quiénes serán los clientes más valiosos para la compañía en base a la comprensión de la evolución de su comportamiento histórico con el fin de poder determinar su comportamiento potencial.

**Solución propuesta:**

Se propone trabajar en el desarrollo e implementación de un método de valoración de clientes que permita principalmente generar un ranking de los clientes que actualmente posee la compañía según su valor histórico y potencial.

Se desea lograr una agrupación como la que se muestra en la “Ilustración 8”, definiendo, en conjunto con las jefaturas de la Marca, los límites de cada segmento y los posibles grupos de acciones de Marketing a focalizar en cada uno de estos.

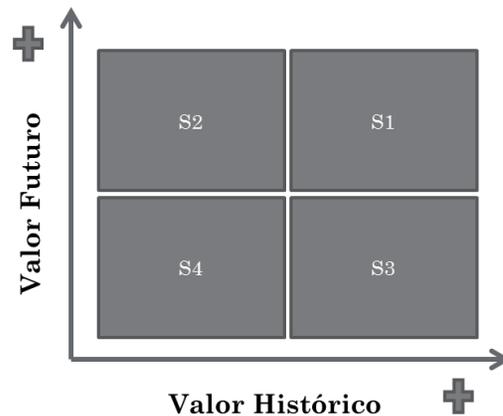


Ilustración 8: Matriz de Agrupación de clientes para la recomendación de acciones

### I.3.-Objetivos

#### Objetivo General

Determinar el valor de clientes con el fin de ayudar a las áreas comerciales a focalizar la retención y fidelización de los clientes más valiosos.

#### Objetivos Específicos

- Consolidar la información transaccional generada en las distintas líneas de negocio.
- Comprender el comportamiento que ha presentado la cartera de clientes de la compañía a lo largo del tiempo.
- Determinar las variables de perfilamiento que influyen en el comportamiento de los clientes.
- Generar ranking de clientes en base al valor que estos representan para la compañía.
- Generar recomendaciones del uso del modelo a las áreas de marketing para orientar el enfoque de acciones a la cartera de clientes.

### I.4.-Alcances

- Para el desarrollo del modelo se utilizará información transaccional y de los clientes pertenecientes solo a la "Marca A".
- Las conclusiones del modelo serán válidas solo para el caso de estudio, pero la metodología será replicable para las otras marcas dentro de la división, ya que se desea generar un procedimiento que en un futuro sea transversal a todas.
- No se realizarán propuestas de campañas de Marketing, sin embargo se generarán recomendaciones del uso del modelo para la selección de clientes a los cuales enfocar distintos tipos de acciones.

### **I.5.-Resultados esperados**

- Mejor comprensión del comportamiento transaccional de los clientes de la compañía.
- Una metodología para la valoración de la cartera de clientes de la compañía en base a variables transaccionales y de perfilamiento.
- Una metodología de selección de clientes a participar en determinadas campañas en base al objetivo de estas.
- Resultados de fácil uso y acceso para el área de Marketing y Ventas de la compañía.

## II.-Marco Conceptual

### II.1.-Cadenas de Markov

Tal como se indica en [10], una cadena de Markov es un proceso estocástico, es decir, una familia arbitraria de variables aleatorias  $\{X_t\}_{t \in T}$ , en donde cada  $X_t$  es una función del espacio muestral en algún conjunto  $E$  (espacio de estados). Estas variables están indexadas por un conjunto ordenado denominado “*dominio de tiempo*” ( $T$ ), donde en cada instante  $t \in T$  se lleva a cabo un experimento donde cuyo resultado queda determinado en el valor de  $X_t$ . El objetivo principal, es conocer con qué probabilidad cada  $X_t$  toma valores dentro del conjunto  $E$  y también saber si esta probabilidad depende de la historia del proceso. Dentro de los procesos estocásticos se distinguen distintos casos según si  $T$  es un conjunto continuo o discreto y según la cardinalidad de  $E$ , entre otros aspectos. En este estudio interesa el caso donde  $T = \mathbb{N} = \{0,1,2,3,4, \dots\}$  y  $E$  es un conjunto numerable.

En particular, las cadenas de Markov son procesos estocásticos que cumplen con la “*propiedad Markoviana*”, la que indica que para todo  $n \geq 0$  y todo  $j, i_n, \dots, i_0 \in E$  se cumple que:

$$P(X_{n+1} = j | X_n = i_n, \dots, X_0 = i_0) = P(X_{n+1} = j | X_n = i_n)$$

Esto indica que el valor de la variable en el siguiente periodo solo depende del valor que esta obtuvo en el periodo anterior y no del resto de la historia.

Es importante notar que la “*Probabilidad de evolución*”  $P(X_{n+1} = j | X_n = i)$ , además de depender de los valores de  $i$  y  $j$ , podría depender del instante de tiempo  $n$ . En base a lo anterior una cadena de Markov se denominará homogénea si se cumple que:

$$P(X_{n+1} = j | X_n = i) = P(X_1 = j | X_0 = i) = p_{ij}, \quad \forall n \geq 0, \quad \forall i, j \in E (*)$$

De esta forma, a la matriz  $P = (p_{ij})_{i,j \in E}$  se le denomina “*Matriz de Transición de la cadena*”, la que no depende del tiempo.

Cuando no se cumple la condición descrita en (\*), se habla de cadenas no homogéneas, donde se tendrá

$$P(X_{n+1} = j | X_n = i) = p_{ij}(n), \quad \forall n \geq 0, \quad \forall i, j \in E$$

Con lo anterior se tiene que la “*Matriz e transición de la cadena*”  $P(n) = (p_{ij}(n))_{i,j \in E}$  va cambiando en el tiempo.

### II.2.- Regresión Logística

Busca explicar el comportamiento de una variable dependiente dicotómica (1 o 0)  $Y$  en función de una o más variables explicativas cualitativas o cuantitativas. Por ejemplo,

podría utilizarse para describir la decisión de compra de un cliente en un determinado momento y analizar las variables que en el influyen.

El objetivo es describir el efecto de los cambios de las variables explicativas sobre la probabilidad de que  $Y$  valga 1 (probabilidad de éxito).

Sea:

$p = \text{Probabilidad } (Y = 1)$   
 $x = \text{variable explicativa}$

El modelo de regresión logística simple plantea:

$$\text{logit}(p) = \ln \frac{p}{1-p} = \beta_0 + \beta_1 x$$

Despejando  $p$  se obtiene:

$$p = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}$$

Para determinar los parámetros que dan vida al modelo se utiliza el enfoque máxima verosimilitud.

### II.3.-Regresión de Poisson Truncada en Cero.

Esta regresión se utiliza para modelar procesos de conteo discretos, donde el valor de la variable en estudio excluye el 0 como opción posible [11].

Corresponde a la distribución de probabilidad condicional de una Regresión de Poisson tradicional, por lo que su distribución de probabilidad se puede derivar de esta última de la siguiente manera:

- Considerando que la función de probabilidad de una distribución de Poisson viene dada por:

$$f(k; \lambda) = P(x_i = k | \lambda) = \frac{\lambda^k \cdot e^{-\lambda}}{k!}$$

- Se condiciona la probabilidad considerando solo valores positivos:

$$P(y_i = k | \lambda, k > 0) = \frac{f(k; \lambda)}{1 - f(0; \lambda)} = \frac{\lambda^k e^{-\lambda}}{k! (1 - e^{-\lambda})} = \frac{\lambda^k}{(e^\lambda - 1) \cdot k!}$$

- Sea  $Y$  la variable dependiente que se busca explicar a través del vector  $x \in \mathbb{R}^n$  de variables independientes. La regresión de Poisson asume que  $Y_x \sim \text{Poisson}(\exp(\beta' \cdot x))$ , por lo que para la regresión truncada se tendrá:

$$P(y_i = k | \beta', x, k > 0) = \frac{f(k; \lambda)}{1 - f(0; \lambda)} = \frac{\lambda^k e^{-\lambda}}{k! (1 - e^{-\lambda})} = \frac{e^{k \cdot \beta' \cdot x}}{(e^{e^{\beta' \cdot x}} - 1) \cdot k!}$$

Los coeficientes de las variables explicativas en la regresión son obtenidos con el método de máxima verosimilitud.

#### II.4.-Modelos de CLV

El CLV se define como el valor presente de todas las utilidades generadas y las esperadas por parte del cliente a lo largo de su relación con la compañía.

Dentro de los beneficios que promete este método se encuentran [12]:

- Diferenciar a los clientes más rentables para poder enfocarse en ellos de mejor forma
- Formular e implementar estrategias específicas de los clientes buscando aumentar las ganancias que se obtienen a lo largo del ciclo de vida de la relación cliente-empresa, y buscando a la vez aumentar el largo de esta.
- Saber cuánto se puede invertir en conservar al cliente con el fin de obtener retorno sobre la inversión.

El CLV se puede dividir en 2 componentes:

$$CLV_j = VH_j + VF_j$$

Donde **VH** representa el Valor Histórico del cliente *j* y se puede escribir como:

$$VH_j = \sum_{t=0}^{Hoy} (m_{jt} - c_{jt}) \cdot (1 + i)^{Hoy-t} - AC_j, \quad (1)$$

Dónde:

- $m_{jt}$  = precio pagado por el cliente *j* en el periodo *t*.
- $c_{jt}$  = costo de servir al cliente *j* en el periodo *t*.
- $i$  = tasa de descuento o costo de interés del capital de la empresa.
- $AC_j$  = Costo de adquirir al cliente *j*.
- **Hoy** = Periodo en que se realiza el análisis

Y **VF<sub>j</sub>** representa el Valor Futuro del cliente y corresponde a:

$$VF_j = \sum_{t=Hoy}^T \frac{E(m_{jt} - c_t) \cdot E(Y_{jt})}{(1 + i)^t}, \quad (2)$$

Dónde, asumiendo independencia entre el número de transacciones y los montos por transacción:

- $E(Y_{jt}) =$  Esperanza del número de transacciones que el cliente  $j$  realizará en el periodo  $t$
- $E(p_j - c_t) =$  Esperanza de los márgenes por transacción que el cliente  $j$  generará en el periodo  $t$ .
- $T =$  Horizonte de tiempo para la estimación del CLV.

El VH es sencillo de calcular y simplemente se debe determinar la tasa correcta para descontar los flujos.

Por otro lado, el VF representa mayor complejidad debido a que se debe tener cuidado en los supuestos a considerar para realizar las proyecciones correspondientes. Para esto, existen distintas alternativas al cálculo de este modelo. Estas se diferencian en la forma de considerar el horizonte de tiempo ( $T$ ) y en la forma de proyectar a futuro las distintas variables, específicamente las tasas de retención (o probabilidades de compra) y los montos generados por los clientes.

#### **II.4.1.-Horizonte de tiempo ( $T$ ):**

Respecto a  $T$ , algunos estudios sugieren considerar un tiempo arbitrario fijo [13], mientras que en otros casos se propone que el tiempo sea considerado infinito pues argumentan que el horizonte de tiempo limitado sobre-estima el CLV. [14].

En el presente trabajo se utiliza un plazo fijo para la evaluación del valor del cliente, el cual corresponde a 4 años.

#### **II.4.2.-Número de Transacciones ( $Y_{jt}$ ):**

Para el cálculo del número esperado transacciones esperadas existen distintos métodos propuestos por diferentes autores para su aplicación en diversas industrias. A modo de ejemplo, se pueden encontrar tesis que aplican diferentes metodologías, como en [15] y [16] que utiliza modelos probabilísticos para proyectar el número de transacciones y compara su desempeño contrastándolo con modelos simples; en [17] se utilizan cadenas de Markov para modelar los cambios de comportamiento de los clientes relacionados a su tasa de retención; en [18] se utiliza el modelo Logit para el cálculo de las tasas de retención; en [19] se utiliza un modelo de clases latentes; en [20] se propone el uso de modelos jerárquicos bayesianos para la estimación del CLV de los clientes de una cadena de supermercados.

Cada uno de los modelos tiene sus ventajas y desventajas, y muchas veces sus capacidades predictivas varían dependiendo del tipo de problema a utilizar.

A continuación se profundiza en 2 de ellos, que han sido utilizados en diversas industrias de forma exitosa y que serán utilizados en el presente trabajo:

### II.4.2.1-Modelos Probabilísticos:

Estos modelos asumen que el comportamiento de los clientes sigue ciertas distribuciones de probabilidad, buscando determinar los parámetros de las distribuciones que mejor se ajusten a la data para utilizarlos como predictores. Ejemplos de estos modelos son el Pareto/NBD, BG/NBD y BG/BB que deben su nombre a las distribuciones de probabilidad que utilizan para representar el comportamiento de los clientes. Las ventajas de estos modelos es que son relativamente sencillos de calcular, sin embargo los supuestos que consideran pueden no ser aplicables a determinadas industrias.

A continuación se explica uno de estos modelos.

#### II.4.2.1.1.-Modelo BG/NBD:

Este modelo considera los siguientes supuestos: [21]

- Mientras está activo el número de transacciones que realiza en un periodo de largo  $t$ , sigue una distribución de Poisson de tasa  $\lambda$ .
- Heterogeneidad en  $\lambda$  se explica mediante una distribución gamma de parámetros  $r, \alpha$ .
- Después de cualquier transacción, un cliente se vuelve inactivo con probabilidad  $p$ . (i.e, La fuga esta descrita a través de una distribución geométrica de parámetro  $p$ ).
- Heterogeneidad en  $p$  sigue una distribución Beta de tasas  $a, b$ .

El modelo, además de los parámetros de las distribuciones, utiliza como input las siguientes variables transaccionales:

- $x$  = número de transacciones realizadas por el cliente a la fecha.
- $t_x$  = tiempo transcurrido entre la primera y última transacción del cliente.
- $T$  = tiempo transcurrido entre la primera transacción del cliente y el final del periodo de observación.

Los parámetros del modelo se estiman utilizando el método de máxima verosimilitud, la que a nivel individual se expresa como:

$$L(r, \alpha, a, b | X = x, t_x, T) = \frac{B(a, b + x)}{B(a, b)} \cdot \frac{\Gamma(r + x)\alpha^r}{\Gamma(r)(\alpha + T)^{r+x}} + \delta_{x>0} \frac{B(\alpha + 1, b + x - 1)}{B(a, b)} \cdot \frac{\Gamma(r + x)\alpha^r}{\Gamma(r)(\alpha + t_x)^{r+x}}$$

Una vez estimados los parámetros es posible derivar una expresión para calcular el número de transacciones esperadas en un periodo futuro de largo  $t$ :

$$E(Y(t)|X = x, t_x, T, r, \alpha, a, b) = \frac{a + b + x - 1}{a - 1} \left[ 1 - \left( \frac{\alpha + T}{\alpha + T + t} \right)^{r+x} F_1(r + x, b + x; a + b + x - 1; \frac{t}{\alpha + T + t}) \right] \\ = \frac{1 + \delta_{x>0} \frac{a}{b + x - 1} \left( \frac{\alpha + T}{\alpha + t_x} \right)^{r+x}}{1 + \delta_{x>0} \frac{a}{b + x - 1} \left( \frac{\alpha + T}{\alpha + t_x} \right)^{r+x}}$$

Donde  $F_1$  corresponde a la función hipergeométrica gaussiana.

#### II.4.2.2-Cadenas de Markov:

También conocidos como “Modelos de Migración”, debido a que permiten a los clientes moverse entre distintos estados. La forma más común de definir estos estados es en términos de la recencia de la última compra realizada por los clientes [22], aunque también hay trabajos en donde se consideran como la cantidad de transacciones realizadas por el cliente en el último periodo [7] e incluso otros que consideran ambas variables [23].

Bajo este enfoque, en este trabajo se considera que el número de transacciones ( $Y_{jt}$ ) a realizar por el cliente  $j$  en el año  $t$  se puede descomponer en:

$$Y_{jt} = P_{jt} \cdot n_{jt}$$

Dónde:

- $P_{jt}$  = Probabilidad de que el cliente  $i$  compre en e periodo  $t$ .
- $n_{jt}$  = Número esperado de transacciones a realizar por el cliente  $j$  en el periodo  $t$  condicionado a que se que me va a comprar.

El objetivo de trabajar con estos 2 componentes es que la decisión de comprar o no comprar y la decisión de cuánto comprar pueden ser independientes, y podría ser conveniente analizarlas de forma aislada, ya que algunas estrategias pueden ir orientadas a aumentar su probabilidad de compra y otras a aumentar el número de transacciones.

De esta forma, cada uno de estos componentes será modelado como una cadena de Markov, identificando durante el transcurso del trabajo las variables de estado y de perfilamiento que afectan las probabilidades de transición de las variables para así poder definir las.

La ventaja es que permiten representar de forma intuitiva el comportamiento del cliente y las variables que lo afectan, sin embargo, la condición de homogeneidad y el orden (primer orden, en la mayoría de los casos) pueden limitar su capacidad predictiva.

##### II.4.2.2.1.-Probabilidad de compra ( $P_{jt}$ )

Para esta variable los estados se definirán como una combinación  $(r, ad)^{15}$ , donde  $r$  indica la recencia (en años) del cliente y  $ad$  indica la cantidad de periodos (años) distintos en que el cliente ha realizado transacciones, por lo que la evolución entre estados indicará la probabilidad de que un cliente compre o no compre. A modo de ejemplo, la “Ilustración 9” indica una representación de los estados y sus conexiones, donde la lógica indica que de un estado en particular se tienen 2 posibles salidas que representan la decisión de compra del cliente en un determinado periodo, por ejemplo, si un cliente está en el estado (1,1) (compró en el periodo anterior, y ha comprado solo en un periodo; i.e, la compra del periodo anterior fue su primera transacción) y decide comprar nuevamente en el siguiente periodo pasará al estado (1,2) (habrá comprado en el periodo anterior, pero ahora ya ha comprado en 2 periodos distintos), mientras que si decide no comprar pasará al estado (2,1) (no compra desde hace 2 periodos y ha comprado solo en 1).



Ilustración 9: Representación de estados y su conexión

En el estudio en cuestión, se consideran 6 estados de recencia donde el último es absorbente lo que indica que si un cliente no ha comprado en 7 periodos seguidos, se considerará como perdido para la empresa y su probabilidad de compra será 0. Además se consideran 3 estados para la variable "ad" donde el tercero indica que el cliente ha comprado en 3 periodos distintos o más. Debido a esto se tendrán 21 posibles estados, lo que dará origen a una Matriz de probabilidades de transición como la que se presente en la “Ilustración 10”.

<sup>15</sup> La justificación se realiza durante el trabajo

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	(1;1)	(1;2)	(1;3)	(2;1)	(2;2)	(2;3)	(3;1)	(3;2)	(3;3)	(4;1)	(4;2)	(4;3)	(5;1)	(5;2)	(5;3)	(6;1)	(6;2)	(6;3)
1 (1;1)	0	p(1,2)	0	p(1,4)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2 (1;2)	0	0	p(2,3)	0	p(2,5)	0	0	0	0	0	0	0	0	0	0	0	0	0
3 (1;3)	0	0	p(3,3)	0	0	p(3,6)	0	0	0	0	0	0	0	0	0	0	0	0
4 (2;1)	0	p(4,2)	0	0	0	0	p(4,7)	0	0	0	0	0	0	0	0	0	0	0
5 (2;2)	0	0	p(5,3)	0	0	0	0	p(5,8)	0	0	0	0	0	0	0	0	0	0
6 (2;3)	0	0	p(6,3)	0	0	0	0	0	p(6,9)	0	0	0	0	0	0	0	0	0
7 (3;1)	0	p(7,2)	0	0	0	0	0	0	0	p(7,10)	0	0	0	0	0	0	0	0
8 (3;2)	0	0	p(8,3)	0	0	0	0	0	0	0	p(8,11)	0	0	0	0	0	0	0
9 (3;3)	0	0	p(9,3)	0	0	0	0	0	0	0	0	p(9,12)	0	0	0	0	0	0
10 (4;1)	0	p(10,2)	0	0	0	0	0	0	0	0	0	0	p(10,13)	0	0	0	0	0
11 (4;2)	0	0	p(11,3)	0	0	0	0	0	0	0	0	0	0	p(11,14)	0	0	0	0
12 (4;3)	0	0	p(12,3)	0	0	0	0	0	0	0	0	0	0	0	p(12,15)	0	0	0
13 (5;1)	0	p(13,2)	0	0	0	0	0	0	0	0	0	0	0	0	0	p(13,16)	0	0
14 (5;2)	0	0	p(14,3)	0	0	0	0	0	0	0	0	0	0	0	0	0	p(14,17)	0
15 (5;3)	0	0	p(15,3)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	p(15,18)
16 (6;1)	0	p(16,2)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17 (6;2)	0	0	p(17,3)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	p(17,17)
18 (6;3)	0	0	p(18,3)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	p(18,18)

Ilustración 10: Matriz de probabilidades de transición entre estados para la estimación de la probabilidad de compra

Para las probabilidades de compra, y así las probabilidades de transición de la matriz, se utilizarán regresiones logísticas.

Cuando se han determinado las probabilidades de transición, el cálculo de la variable quedará definido por:

$$P_{jt} = I_j \cdot P_t^t \cdot a$$

Dónde:

- $I_j$  = Vector que representa el estado inicial del cliente.
- $P$  = Matriz de probabilidades de transición del sistema.
- $a$  = Vector auxiliar que permitirá rescatar los valores correspondientes de la matriz de probabilidad.
- $t$  = Periodo para el cual se quiere determinar la probabilidad.

#### II.4.2.2.2.-Número esperado de transacciones condicionado ( $n_{jt}$ )

Respecto a esta componente, se utilizará un enfoque similar al propuesto en [7], donde los estados de la cadena de Markov están definidos por el número de transacciones realizadas por el cliente en el periodo anterior<sup>16</sup>. En particular, se tendrán 5 estados (0, 1, 2, 3, 4), lo que genera una matriz de transición mucho más sencilla que la utilizada para la evolución de las probabilidades de compra. Sin embargo, a diferencia de esa matriz, en esta cada estado tendrá 4 salidas que representan la probabilidad de pasar de realizar en un periodo  $t$  una cantidad de transacciones igual a  $r \in \{0, 1, 2, 3, 4\}$  a realizar en  $t + 1$  una cantidad de transacciones  $s \in \{1, 2, 3, 4\}$ . Notar que  $s$  no puede ser 0 debido a que estamos condicionando en que el cliente comprará. La matriz descrita se puede apreciar en la “Ilustración 11”.

<sup>16</sup> Esta selección de estados se justificará durante el trabajo.

		1	2	3	4	5
		(0)	(1)	(2)	(3)	(4)
1	(0)	0	p(1,1)	p(1,2)	p(1,3)	p(1,4)
2	(1)	0	p(2,1)	p(2,2)	p(2,3)	p(2,4)
3	(2)	0	p(3,1)	p(3,2)	p(3,3)	p(3,4)
4	(3)	0	p(4,1)	p(4,2)	p(4,3)	p(4,4)
5	(4)	0	p(5,1)	p(5,2)	p(5,3)	p(5,4)

Ilustración 11: Matriz de probabilidades de transición entre estados para la estimación del número de transacciones

Para determinar las probabilidades de transición se utilizan Regresiones de Poisson Truncadas en Cero. Esta es una diferencia importante a lo realizado en [7], donde utilizan regresiones de Poisson debido a que no realizan la descomposición de la probabilidad de compra.

Una vez determinadas las probabilidades de transición entre estados, el cálculo de la variable en cuestión es similar al realizado para las probabilidades de compra, es decir:

$$n_{jt} = I_j \cdot P^t \cdot a$$

Dónde:

- $I_j$  = Vector que representa el estado inicial del cliente.
- $P$  = Matriz de probabilidades de transición del sistema.
- $a$  = Vector auxiliar que permitirá rescatar los valores correspondientes de la matriz de probabilidad.
- $t$  = Periodo para el cual se quiere determinar la probabilidad.

### II.4.3.-Montos por transacción ( $m_{jt}$ )

#### II.4.3.1.-Modelo Gamma-Gamma

Este método se basa en los siguientes supuestos: [24]

- El valor monetario de las transacciones realizadas por un cliente varía aleatoriamente en torno a su monto promedio por transacción.
- El monto promedio por transacción varía entre clientes, pero no varía en el tiempo para un mismo cliente.
- La distribución del monto promedio por transacción a nivel de cliente es independiente del proceso que determina el número de transacciones que este realizará.
- Sea  $z_i$  el valor generado por un determinado cliente en su transacción  $i$ -ésima, se asume que  $z_i$  sigue una distribución gamma de parámetros  $p, v$ .
- Heterogeneidad en el parámetro  $v$  sigue una distribución gamma con parámetros  $q, y$ .

Con lo anterior, se tiene que el monto estimado por transacción para un cliente que ha realizado  $x$  transacciones a la fecha se expresa como:

$$E(m|p, q, y, \bar{z}, x) = \frac{p(y + x\bar{z})}{px + q - 1}$$

Dónde:

- $\bar{z}$  = monto promedio de las  $x$  transacciones realizadas por el cliente a la fecha.

La estimación de los parámetros del modelo se realiza utilizando el método de máxima verosimilitud.

## II.5.-Validación de Modelos

Con el fin de evaluar el desempeño del modelo que se decida utilizar, se propone el análisis de diferentes métricas de ajuste como:

**MAE (Error absoluto medio):** Busca reflejar en cuánto se equivoca el modelo. Matemáticamente corresponde a:

$$MAE = \frac{1}{n} \sum_i |y_i - \hat{y}_i|$$

Dónde:

- $y_i$  es la variable dependiente del registro.
- $\hat{y}_i$  es la estimación de la variable.
- $\bar{y}$  es la media de la variable.
- $n$  es el número total de registros.

**MAPE (Error porcentual absoluto medio):** Entrega una medida sobre cuánto se equivoca porcentualmente el modelo. Matemáticamente corresponde a :

$$MAPE = \frac{1}{n} \sum_i \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

**MPE (Error porcentual medio):** Busca reflejar si el modelo tiene tendencia a sobreestimar o subestimar el valor real de la variable.

$$MPE = \frac{1}{n} \sum_i \left( \frac{\hat{y}_i - y_i}{y_i} \right)$$

**WMAPE (Error porcentual absoluto ponderado):** Es una modificación del MAPE. Trata de salvar los altos valores de desviación causados por valores de predicción pequeños. La idea es entregar la importancia adecuada a cada valor a estimar.

$$WAPE = \frac{\sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot y_i}{\sum_{i=1}^n y_i}$$

Es importante notar que en cada una de las medidas, si el valor real  $y_i$  de la variable es igual a 0, los indicadores se indeterminan. Debido a lo anterior y tal como lo hacen en [15], no se considerarán estos casos para el análisis de los errores. Es decir, los errores entregados corresponden a los cometidos en los clientes que realmente compran en el periodo de validación, por lo que además se necesitarán métodos adicionales para evaluar la capacidad predictiva global del modelo.

**Curva ROC (Receiver Operating Characteristic):** Permite evaluar la capacidad de discriminación de un modelo que clasifica observaciones en 2 grupos (1 o 0) como lo es el caso de la regresión logística.

Para realizar esta evaluación se analizan distintos puntos de corte ( $\pi_0$ ) para la probabilidad sobre la cual la observación sería calificada como 1 por el modelo. En base a estos puntos de corte se analiza la sensibilidad y especificidad del modelo, donde:

- Sensibilidad:  $P(\hat{y}_i = 1 | y_i = 1)$
- Especificidad:  $P(\hat{y}_i = 0 | y_i = 0)$

La curva ROC representa en un gráfico la “sensibilidad” (o TPR) en función de “1-especificidad” (o FPR). De esta forma se evalúa la capacidad predictiva del modelo comparándose con una clasificación aleatoria, la que se representa con una recta en el gráfico. Un ejemplo de este tipo de curvas se muestra en la “Ilustración 12”

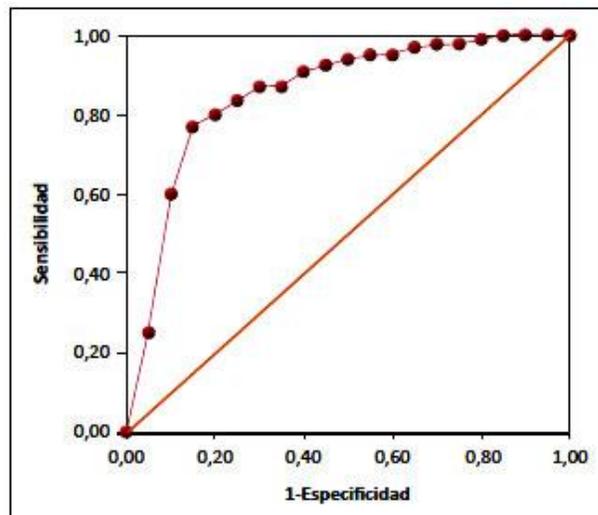


Ilustración 12: Representación de curva ROC

Mientras mayor sea el área bajo la curva ROC, mayor será la capacidad predictiva del modelo.

## II.6.-Significancia de variables

### II.6.1.-Test Anova (Análisis de varianza)

Se utiliza para determinar si una variable independiente categórica es estadísticamente significativa respecto a una variable dependiente cuantitativa que se quiere estudiar.

El test consiste en probar si los distintos grupos formados por la variable categórica poseen la misma media en relación a la variable dependiente. La hipótesis que se plantea es

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_n$$
$$H_1 = \text{No todas las medias son iguales}$$

Donde  $\mu_i, i = 1, \dots, n$  corresponde a las media de cada uno de los grupos formados por la variable categórica.

La regla de decisión consiste en rechazar  $H_0$  si el nivel de significancia arrojado por el test es menor a cierto número, el que generalmente se fija en 0.05 lo que equivale a asumir un nivel de confianza de un 95%.

### II.6.2.-Test de Wald

Para evaluar si los parámetros obtenidos para las variables (principalmente variables categóricas con más de 2 opciones posibles) en una regresión logística son significativos se utiliza el Test de Wald.

Este permite establecer qué variables son importantes para explicar la probabilidad del suceso ( $Y = 1$ ), mediante la prueba de las siguientes hipótesis:

$$H_0: \beta_i = 0$$
$$H_1: \beta_i \neq 0$$

Donde  $\beta_i$  corresponde al coeficiente de la variable independiente que se quiere validar. Es decir, la Hipótesis nula plantea que, al ser 0 el coeficiente de la variable en explicativa, esta no afecta en el comportamiento de la variable en estudio.

Para llevar a cabo el test de Wald se construye el estadístico de Wald ( $W_i$ ) de la siguiente manera:

$$W_i = \frac{\beta_i}{S_{\beta_i}} \rightsquigarrow N(0,1) \Leftrightarrow \left( \frac{\beta_i}{S_{\beta_i}} \right)^2 \rightsquigarrow \chi_1^2$$

Es decir, el estadístico se compara con una distribución chi-cuadrada, donde el criterio consiste en rechazar la Hipótesis nula cuando el nivel de significancia es menor a 0,05.

### **III.-Descripción de la Metodología**

Para cumplir con los objetivos propuestos, el desarrollo del trabajo contempla una serie de etapas, las que se describen a continuación:

#### **III.1.-Descripción de la Data Disponible**

Con el objetivo de comprender la data de la que se dispone y las distintas variables que la componen, el desarrollo del trabajo comienza con la descripción de cada una de las variables que podrían ser útiles para la descripción y análisis del comportamiento de los clientes.

Esta etapa involucra la descripción de la data transaccional disponible, que corresponde a las ventas de camiones realizadas por la “Marca A” desde el año 1999 al año 2013. Además, se poseen variables relacionada a los clientes que han realizado estas transacciones, las que también son descritas en esta etapa. Por otro lado, se agrega la descripción de ciertas variables macroeconómicas que podrían afectar el comportamiento presentado por los clientes.

Además, se presenta una descomposición de la data transaccional en observaciones por cliente, con lo que se busca evitar perder información relacionada a su comportamiento que podría ser útil en los modelos que se desarrollarán de forma posterior. Para esto, por cada año a partir del cual el cliente compra por primera vez en la empresa, se crean variables anuales que indican si el cliente compra o no compra y la cantidad de transacciones realizadas (en el caso de que compre).

#### **III.2.-Análisis del Comportamiento Histórico**

Una vez descritas las variables disponibles, se procede a analizar la relación entre las distintas variables disponibles y las variables de comportamiento que se buscan proyectar (número de transacciones y montos por transacción), con el fin de comprender de mejor manera la forma que han afectado a las decisiones tomadas periodo a periodo por los clientes. Para esto, se realiza un análisis de la evolución del comportamiento de los clientes a lo largo de los años y de cómo se diferencia este comportamiento según las variables de perfilamiento disponibles.

Además, en esta etapa se determina el valor histórico generado por los clientes a lo largo de su relación con la empresa, almacenando este dato en una nueva variable para que de esta forma pueda ser utilizada en las etapas futuras del presente trabajo.

#### **III.3.-Selección de Variables Relevantes.**

Con el fin de determinar si las relaciones, analizadas en la etapa anterior, entre las variables de comportamiento a proyectar y el resto de las variables disponibles son realmente significativas, se procede a analizar la independencia entre las variables

mediante la aplicación del Test ANOVA y el Test Chi-Cuadrado dependiendo de la naturaleza de las variables a estudiar, utilizando el primero de ellos para el caso en que la variable dependiente es numérica y la independiente es categórica, y el segundo para el caso en que ambas variables son categóricas.

La importancia de esta etapa radica en lograr identificar las variables relevantes a incluir en los modelos de proyección de comportamiento que se proponen posteriormente, para de esta forma lograr obtener los mejores resultados posibles en cuanto a la capacidad predictiva que estos modelos entregan.

#### **III.4.-Aplicación y Validación de los Modelos.**

Una vez seleccionadas las variables a utilizar para la predicción del número de transacciones y los montos por transacción, se procede a aplicar los modelos propuestos dividiendo la data entre los años 1999-2011 y 2012-2013 para los periodos de calibración y validación respectivamente.

En una primera instancia se estima el número de transacciones anuales a realizar por los clientes utilizando el modelo probabilístico BG/NBD al cual se le realizan 3 variaciones, donde en cada una de ellas se van agregando variables explicativas al modelo inicial, el que utiliza como base las variables recencia y frecuencia del cliente al momento del análisis.

Luego se utiliza el Modelo Logit-Poisson-Markov (L-P-M) para la estimación del número de transacciones, utilizando, al igual que en primer modelo, 3 variaciones en donde se van incorporando variables explicativas.

Para la estimación de los montos por transacción a gastar por los clientes se utiliza el modelo probabilístico Gamma-Gamma.

Con el fin de validar los modelos se utilizan distintas métricas de ajuste como el MAPE y MAE, además de una estimación del error porcentual agregado de cada uno de los modelos y sus variaciones, seleccionando el que posea los errores más bajos durante el periodo de validación.

Finalmente se aplica el modelo, seleccionado en base al criterio anterior, a toda la cartera de clientes y se proyecta el valor anual de estos hasta el año 2018.

#### **III.5.-Propuesta de uso del Modelo.**

Para finalizar, se propone el uso del modelo para la creación de grupos de clientes de acuerdo a su valor histórico, su valor proyectado para el año actual, y su valor proyectado de cara al año 2018. Para los grupos generados a partir del modelo, se proponen distintas acciones de marketing que buscan potenciar la relación con los clientes, en especial con los más valiosos.

Con lo anterior se busca dar recomendaciones a las áreas de Marketing para que estas tengan más claridad a la hora de enfocar sus esfuerzos en la cartera de clientes, con el fin de maximizar la rentabilidad que estas acciones generan para la empresa.

## IV.-Desarrollo Metodológico

### IV.1.-Descripción de la data disponible

Para comprender la data de la que se dispone, y las distintas variables que la componen se realizan las siguientes etapas:

#### IV.1.1.-Data transaccional

El trabajo se desarrolla utilizando la información histórica de la “Marca A” en su línea de negocio de venta de camiones medianos. Se posee información transaccional desde el año 1999 hasta el 2013. A continuación se muestra un resumen general de esta data:

<b>Fecha Primera transacción</b>	12-02-1999
<b>Fecha Última transacción</b>	31-12-2013
<b>Número de clientes</b>	1872
<b>Número de Vehículos vendidos</b>	3292
<b>Número de transacciones</b>	2616
<b>Monto total</b>	MM\$ 85.896

Tabla 2: Resumen de datos transaccionales

Donde una transacción equivale a una fecha única de compra por parte de un cliente, en la cual puede comprar uno o más vehículos y el monto total está expresado en términos de la facturación realizada, esto último debido a que no se maneja la información de los costos asociados a cada una de las transacciones históricas. Además, los valores de las facturaciones están actualizados a valor presente en términos de la inflación con el objetivo de hacer comparables los montos entre años lejanos. Este cálculo se realiza utilizando la siguiente fórmula:

$$Valor\ actualizado = Valor\ factura \cdot \frac{Valor\ UF\ Hoy}{Valor\ UF\ Fecha\ de\ compra}$$

Es importante mencionar que este número de transacciones del que se dispone para su utilización corresponde al 93% del total histórico, debido a que existen transacciones que tienen asociadas como clientes a una institución financiera<sup>17</sup>, por lo que estas quedaron fuera del análisis debido a la imposibilidad de asociarles un cliente final<sup>18</sup>.

<sup>17</sup> Esto se debe a que históricamente, cuando un vehículo era financiado a través de algún crédito o leasing, quedaba la institución financiera asociada a la transacción y no el cliente final, y muchas veces no almacenaban la información del “cliente real” o “usuario” en ningún sistema.

<sup>18</sup> Es posible que se en un plazo de 2 semanas se pueda depurar esta data y así aumentar el % de data disponible a trabajar, así que los valores aquí mostrados en este análisis parcial podrían cambiar.

En base a la data disponible, las variables transaccionales que se poseen para cada cliente son:

Abreviación	Fórmula	Descripción
$x$		Número de vehículos comprados hasta el año $t$
$x^*$		Número de transacciones <sup>19</sup> realizadas hasta el año $t$
$M$		Monto total gastado
$Hoy$	'01-01-2014'	Fecha "simulada" en que se realiza el análisis
$t_1$		Fecha de la primera transacción
$t_f$		Fecha de la última transacción
$T_1$	$Año(t_1)$	Año en que el cliente realiza la primera transacción
$T_f$	$Año(t_f)$	Año en que el cliente realiza la última transacción
$T$	$\frac{Hoy - t_1}{365}$	Antigüedad, en años, del cliente
$t_x$	$\frac{t_n - t_1}{365}$	Tiempo transcurrido, en años, entre la primera y la última transacción
$A$		Años distintos en que el cliente ha comprado
$x^{**}$	$\frac{x^*}{A}$	Número de interacciones por año en que ha comprado
$m^*$	$\frac{M}{x^*}$	Monto promedio por transacción
$R$	$Año(Hoy) - T_f$	Recencia, en años, de un cliente

Tabla 3: Variables transaccionales por cliente

Además de estas variables que resumen el comportamiento histórico del cliente, se realizan cortes anuales de la data, por lo que para todo cliente se tiene la información acumulada anual de cada variable. A modo de ejemplo se posee la variable:

- $x_t^* = \text{Número de transacciones históricas realizadas por el cliente hasta finales del año } t$

#### IV.1.2.-Data de perfilamiento

Las variables de perfilamiento con las que se dispone son:

- Naturaleza del Cliente: Indica si el cliente es una persona natural o una empresa.

<sup>19</sup> Donde las transacciones se definen como compras con distintas fechas.

- Tamaño de flota: Variable categórica que entrega una referencia de la cantidad de vehículos<sup>20</sup> que posee el cliente. Posee 4 categorías:
  - Particular: Máximo 2 vehículos.
  - Flotista pequeño (3-5)
  - Flotista mediano (6-15)
  - Flotista grande (16 o más)
- Región
- Zona: Indica la zona geográfica a la cual pertenece el cliente según su región.
- Rubro: Indica la industria para la cual el cliente utiliza sus vehículos. Actualmente existen 26 opciones para este atributo, las que se describen en el “Anexo A”
- Vehículos de otro segmento: Indica si el cliente tiene, a la fecha, algún vehículo de otro segmento dentro de la marca. Es decir, si el cliente posee algún “Vehículo liviano” o “Vehículo pesado” además de los vehículos que posee dentro del segmento de vehículos medianos.

#### ***IV.1.3.-Data Macroeconómica***

Con el fin de analizar el impacto que tiene el movimiento de variables macroeconómicas en el comportamiento de la cartera de clientes, se decide incluir variables de este tipo, sin embargo para poder incluirlas en el estudio, se debe conocer la evolución histórica de la variable (desde 1999 a la fecha) y las proyecciones futuras de ella (hasta el 2018). Debido a lo anterior, se decide comenzar a analizar el efecto de estas variables tomando como ejemplo una que pueda ser significativa y presente valores conocidos para el rango de tiempo mencionado. Esta variable corresponde a:

- Crecimiento del PIB (% anual): Tasa de crecimiento anual porcentual del PIB a precios de mercado en moneda local.

#### ***IV.1.4.Limpieza, Pre-Procesamiento y Transformación de datos.***

A continuación se muestran la cantidad de clientes que poseen una determinada característica, y las respectivas limpiezas y tratamiento de datos en caso de ser necesarias.

---

<sup>20</sup> Independiente de la marca y el tipo que estos sean.

#### ***IV.1.4.1.-Variables de Perfilamiento***

##### Naturaleza del cliente:

Todos los clientes poseen este dato y en la “Tabla 4” se aprecia que, considerando todos los clientes que han comprado hasta finales del año 2013, la cantidad de empresas es superior a la cantidad de personas.

<b>Naturaleza del Cliente</b>	<b>Frecuencia</b>	<b>%</b>
<b>Empresa</b>	1078	57,55%
<b>Persona</b>	794	42,45%
<b>Total general</b>	<b>1872</b>	<b>100,00%</b>

Tabla 4: Frecuencia según Naturaleza del cliente

##### Tamaño de flota:

En la “Tabla 5” se aprecia que la mayoría de los clientes que posee la empresa hasta el periodo de estudio corresponden a “Particulares” y “Flotistas pequeños”, lo que a priori parece indicar un “bajo potencial” de compra para la cartera de clientes en general<sup>21</sup> debido a que la lógica indica que a mayor cantidad de flota que manejan los clientes, mayor será la frecuencia de renovación y la cantidad de vehículos a comprar por estos.

<b>Tamaño de flota</b>	<b>Frecuencia</b>	<b>%</b>
<b>Flotista grande (16 o más)</b>	65	3,47%
<b>Flotista mediano (6-15)</b>	242	12,93%
<b>Flotista pequeño (3-5)</b>	598	31,94%
<b>Particular</b>	967	51,66%
<b>Total general</b>	<b>1872</b>	<b>100,00%</b>

Tabla 5: Frecuencia según Tamaño de flota

##### Región:

En el “Anexo B” se presenta la distribución de clientes por región, donde se aprecia que la mayoría de los clientes pertenece a la región metropolitana, seguidos de lejos por los de la Octava Región. Esto se debe a que el fuerte de las ventas se concentra en la región metropolitana a través de las sucursales propias que posee la empresa.

Con el objetivo de hacer esta variable más manejable, se decide agrupar a los clientes en zonas geográficas de acuerdo a la región a la que pertenecen. En base a esta agrupación se define la siguiente variable:

##### Zona Geográfica:

---

<sup>21</sup> Sin embargo, más adelante se indicará que esta hipótesis no es correcta.

De los resultados de la variable anterior se desprende que la mayoría de los clientes están ubicados en la zona centro, tal como se observa en la “Tabla 6”, seguidos por los ubicados en el norte.

Zona	Frecuencia	%
Centro	1493	79,75%
Norte	271	14,48%
Sur	108	5,77%
<b>Total general</b>	<b>1872</b>	<b>100,00%</b>

Tabla 6: Frecuencia según Zona

### Rubro:

Debido a la existencia de muchas opciones para esta variable, se propone una agrupación en relación a su frecuencia de repetición, manteniendo así los 7 rubros que más se repiten y agrupando el resto en la categoría otros. Esto, con el fin de incluir de forma más sencilla esta variable en los modelos que posteriormente se utilizarán y porque además existen muchos rubros que, de acuerdo a la descripción que poseen, son muy similares en cuanto a la utilidad que los clientes le dan a los vehículos.

En el “Anexo C” se observa la distribución de rubros actual, y en base a esta información y a la regla de agrupación mencionadas anteriormente, las 8 etiquetas<sup>22</sup> que finalmente definirán esta variable quedan:

Rubro	Frecuencia	%
R1	185	9,88%
R2	259	13,84%
R3	77	4,11%
R4	271	14,48%
R5	78	4,17%
R6	438	23,40%
R7	97	5,18%
R8	467	24,95%
<b>Total general</b>	<b>1872</b>	<b>100,00%</b>

Tabla 7: Frecuencia según nuevo rubro

### Vehículos de otro segmento:

La mayoría de los clientes, dentro del segmento de vehículos medianos, no tienen vehículos de otros segmentos dentro de la misma marca, siendo solo un poco menos del 14% los que poseen además algún vehículo liviano o pesado. La lógica indicaría que estos clientes, al poseer otro tipo de vehículos dentro de la marca, representan un mayor grado de fidelidad, por lo que debería ser más probable vuelvan a comprar en el

<sup>22</sup> Donde se esconde el nombre de los rubros, con el fin de proteger información que la empresa considera relevante.

futuro en comparación de los que solo poseen vehículos dentro del segmento de medianos.

Tiene vehículos otro segmento	Frecuencia	%
NO	1611	86,06%
SI	261	13,94%
<b>Total general</b>	<b>1872</b>	<b>100,00%</b>

Tabla 8: Frecuencia según vehículos de otro segmento del cliente

#### **IV.1.5.-Descomposición de la data transaccional en observaciones por cliente**

Debido a que los volúmenes transaccionales generados en la industria son pequeños, se descompone la data en observaciones por cliente con el fin de evitar perder información relacionada a su comportamiento.

Para esta descomposición se considera que a partir del año siguiente al año en que el cliente realiza su primera transacción con la empresa (es decir, su primera transacción no se considera como una observación), el cliente puede tomar 2 decisiones en cada año. Estas decisiones son:

- Comprar o no comprar ( $C_t$ ).
- Si decide comprar, debe tomar la decisión del Número de transacciones a realizar ( $x_t^{***}$ ).

De esta forma, a lo largo de la historia se tienen 6092 observaciones de decisión de compra por parte de la cartera de clientes, las cuáles se descomponen en:

Decisión	Frecuencia	%
Compra	456	7,49%
No compra	5636	92,51%
<b>Total general</b>	<b>6092</b>	<b>100,00%</b>

Tabla 9: Resumen decisiones de compra.

Estas 456 observaciones en las cuales la decisión tomada por los clientes es “Compra”, corresponden a las transacciones realizadas año a año por los clientes antiguos, es decir, clientes que han comprado previamente en la empresa. En la “Ilustración 13” se observa la evolución que ha tenido esta decisión en el tiempo. Aquí se aprecia que año a año son muy pocos los clientes que vuelven a realizar una nueva transacción, lo que vuelve a justificar el problema planteado inicialmente de que las tasas de recompra de la cartera de clientes son bajas.

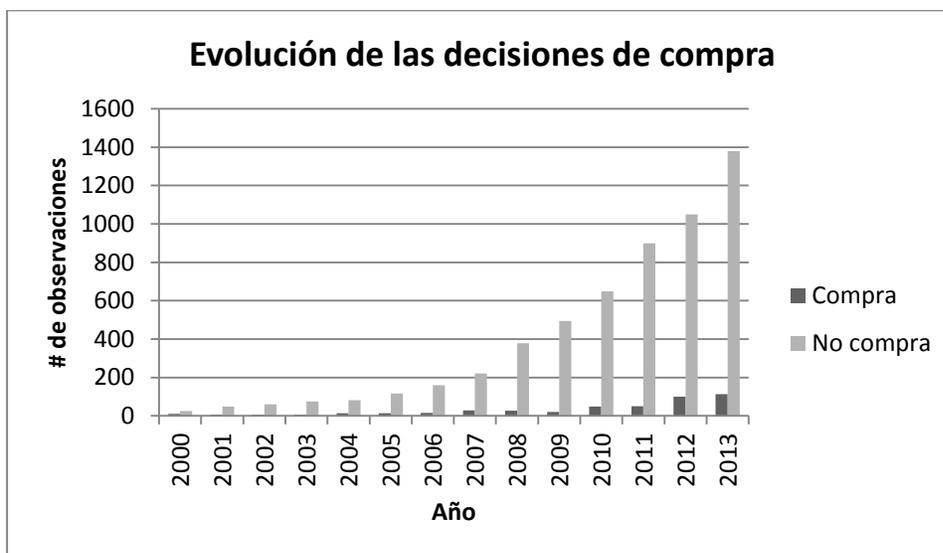


Ilustración 13: Evolución de las decisiones de compra en el tiempo

Respecto a la decisión del número de transacciones a realizar, dado que el cliente decide comprar ( $x_t^{**}$ ) se tienen 456 observaciones (las decisiones de compra positiva del punto anterior). La distribución de esta variable se muestra en el “Anexo D” donde se aprecia que son pocas las observaciones en donde los clientes realizan más de 4 transacciones. Por esta razón se decide categorizar esta variable de la forma en que se muestra en la “Tabla 10”.

Número de transacciones	Frecuencia	%
1	345	75,66%
2	75	16,45%
3	21	4,61%
4 o +	15	3,29%
<b>Total general</b>	<b>456</b>	<b>100,00%</b>

Tabla 10: Resumen decisión del número de transacciones

Se observa que la mayoría de los clientes que vuelven a comprar realizan solo una transacción con la empresa en el año de observación, lo que parece ser la lógica debido a que este tipo de compras se realizan de forma planificada y por lo general los clientes deberían comprar una vez al año (en donde podrían comprar más de un vehículo). Sin embargo, en cerca de un 25% de las observaciones en donde los clientes deciden comprar, estos realizan transacciones en fechas distintas dentro del mismo año, por lo que esta decisión debe ser considerada al momento de estudiar el comportamiento y valor de los clientes.

En la “Ilustración 14” se presenta la evolución que ha tenido en el tiempo el promedio de la variable  $x_t^{**}$ . Se observa que no hay grandes variaciones anuales, a excepción del periodo transcurrido entre los años 2007 y 2009, y que en general los promedios se acercan a 1,5 transacciones anuales.

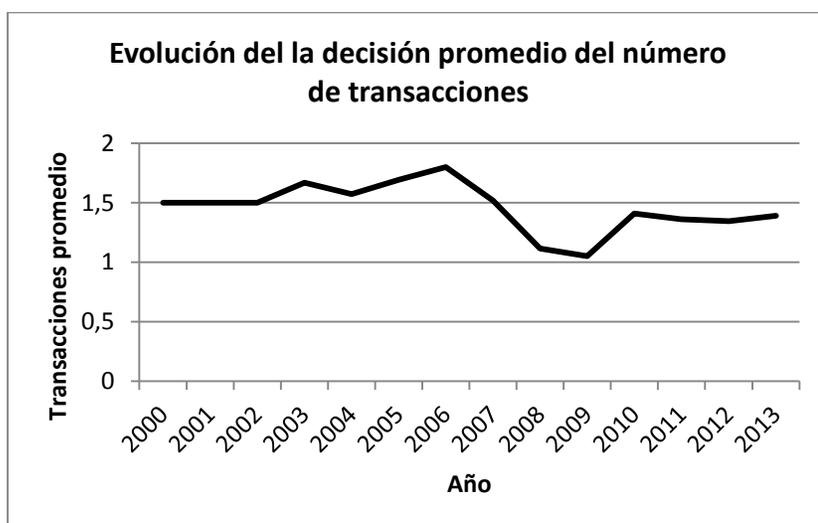


Ilustración 14: Evolución del promedio de la decisión del número de transacciones

Es importante mencionar que para cada observación, tanto de decisión de compra como de la cantidad de transacciones a realizar, se poseen las variables de comportamiento (descritas en la sección “III.1.1”) que el cliente posee al momento de la toma de decisiones.

## IV.2.-Análisis del comportamiento histórico.

Con el fin de comprender el comportamiento histórico de los clientes se busca analizar la relación entre las distintas variables disponibles y las acciones que ha tomado la cartera de clientes a través de los siguientes pasos:

### IV.2.1.-Determinación del valor histórico de los clientes de la empresa.

Para calcular el valor histórico de cada cliente se ajustan por inflación los montos pagados por los clientes en el pasado. De esta forma, el valor histórico de cada cliente queda determinado por:

$$Valor\ Histórico_i = \sum_{t=12-02-1999}^{HOY} \pi(t) \cdot \frac{Valor\ UF_{HOY}}{Valor\ UF_t}$$

La “Tabla 11” muestra los resultados a nivel general, del Valor Histórico de los clientes. Se aprecia que un 20 % de los clientes ha generado un 56% de los ingresos en términos de facturación, lo que está lejos de la regla de Pareto. Además se aprecia que la contribución de los grupos 3, 4 y 5 es casi idéntica.

Segmento	# Clientes Acum	% Clientes Acum	Monto Acum	% Monto Acum
1	374	20%	MM\$ 48.670	56%
2	748	40%	MM\$ 62.460	72%
3	1122	60%	MM\$ 72.166	83%
4	1496	80%	MM\$ 80.115	92%
5	1872	100%	MM\$ 86.888	100%

Tabla 11: Resultados Valor Histórico

Sin embargo, no basta solo con tener esta mirada histórica e identificar y perfilar, por ejemplo, al 20% superior que aparece en la tabla.

Esto, debido a que hay muchos clientes que han experimentado pocas oportunidades de compra, por lo que no se debe determinar el perfil, sino que los patrones que han determinado la evolución del comportamiento de la cartera de clientes.

#### ***IV.2.2.-Relación entre las variables de comportamiento y las variables de perfilamiento.***

Las variables de comportamiento relevantes para la determinación del valor del cliente corresponden a la cantidad de interacciones realizadas ( $x^*$ ) y a los montos gastados por interacción ( $m_i$ ). Además, resulta importante comprender los factores que afectan a la decisión de comprar o no comprar ( $C_t$ ) tomada periodo a periodo por los clientes, y la decisión del número de transacciones a realizar dado que decidió comprar ( $x_t^{**}$ ). Por esta razón, a continuación se realiza una descripción del comportamiento histórico de estas variables y su relación con las distintas variables de perfilamiento descritas:

##### ***IV.2.2.1.-Número de Transacciones ( $x^*$ ), Montos por Transacción ( $m^*$ ) y variables de perfilamiento***

A continuación se describe el comportamiento de las variables “Número de Transacciones” y “Montos por Transacción” según los distintos valores de cada variable de perfilamiento. La descripción se realiza en base a resúmenes históricos y a la evolución que año a año han presentado:

##### Naturaleza del cliente:

En la “Tabla 12” se observa que los clientes tipo empresa han interactuado, en promedio, más veces que los clientes tipo persona, siendo también la desviación estándar de sus interacciones más alta lo que indicara que además existe una mayor variabilidad en su comportamiento. Lo mismo ocurre con el monto promedio por interacción. Lo anterior parece indicar cierta dependencia entre ambas variables de comportamiento y la variable de perfilamiento en estudio.

Naturaleza del cliente	Promedio de $x^*$	Desviación estándar de $x^*$	Promedio de $m^*$	Desviación estándar de $m^*$
Empresa	1,5816	1,8212	\$ 31.120.132	\$ 24.788.122
Persona	1,1474	0,5119	\$ 26.235.150	\$ 10.712.765
<b>Total general</b>	<b>1,3974</b>	<b>1,4375</b>	<b>\$ 29.048.190</b>	<b>\$ 20.203.162</b>

Tabla 12: Transacciones y Montos por Transacción según Naturaleza del cliente

Observando tanto la “Ilustración 15” como la “Ilustración 16” se desprende que año a año el comportamiento de los clientes antiguos<sup>23</sup> es mucho mejor, en promedio, al comportamiento de los clientes nuevos<sup>24</sup>, presentando mayores promedios tanto para  $x^*$  como para  $m^*$ . Además, se tiene que para los clientes que se adquieren en cada periodo, ambas variables de comportamiento son bastante similares, siendo solo un poco mayor para los clientes tipo empresa en ambos casos.

Por otro lado, considerando la evolución del comportamiento de los clientes antiguos, se observa que para ambas variables de comportamiento, los promedios son mucho mayores para los clientes tipo empresa lo que parece confirmar lo que se desprendía del resumen histórico donde se observó una posible dependencia entre las ambas variables de comportamiento y la variable “Naturaleza” del cliente

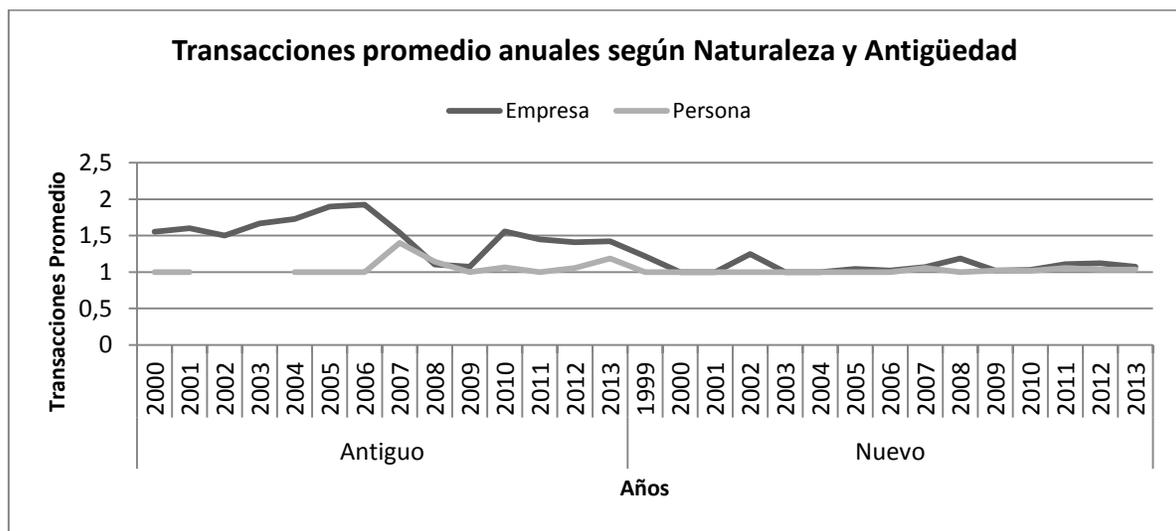


Ilustración 15: Evolución del promedio de Transacciones según Naturaleza y Antigüedad.

Es importante mencionar que las disminuciones que se aprecian anualmente de los montos promedios por transacción en la “Ilustración 16” no se deben necesariamente a que los clientes hayan disminuido drásticamente la cantidad de vehículos que compran por transacción, sino más bien a una disminución de precios que, sumado al factor de actualización por inflación (que aumenta aún más los precios pagados en años anteriores) genera el “efecto a la baja” que se aprecia. Lo importante es observar año a año la diferencia entre los distintos tipos de clientes según su perfil. Este detalle se debe considerar para el resto de los gráficos que se presentan en esta sección.

<sup>23</sup> Clientes que han comprado en algún año anterior al año al cuál se está analizando

<sup>24</sup> Clientes que compran por primera vez en el año que se está analizando

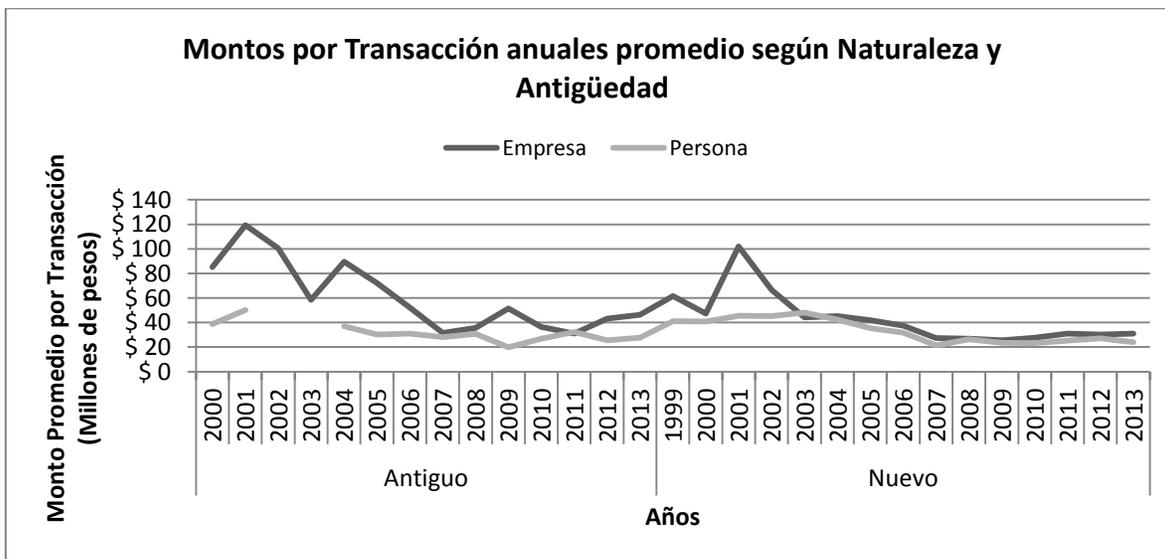


Ilustración 16: Evolución del Monto promedio por Transacción según Naturaleza y Antigüedad

Rubro:

Los clientes que utilizan sus camiones en el rubro “R3” y “R5” son los que han realizado, en promedio, una mayor cantidad de transacciones con la empresa considerando el resumen histórico del comportamiento de la cartera de clientes representado en la “Tabla 13”. Además, para estos mismos rubros, la desviación estándar del promedio de interacciones es mayor lo que parece indicar una mayor heterogeneidad en su comportamiento.

Por otro lado, los clientes del rubro “R2” son los que poseen un mayor monto promedio por transacción y además la mayor desviación estándar para esta variable.

Rubro	Promedio de $x^*$	Desviación estándar de $x^*$	Promedio de $m^*$	Desviación estándar de $m^*$
R1	1,162	0,517	\$ 28.344.637	\$ 15.677.618
R2	1,490	1,227	\$ 33.464.344	\$ 38.937.120
R3	1,779	2,210	\$ 32.468.137	\$ 24.220.466
R4	1,240	0,749	\$ 28.033.619	\$ 15.788.463
R5	1,744	2,360	\$ 33.059.081	\$ 16.992.235
R6	1,326	1,675	\$ 27.368.991	\$ 13.575.724
R7	1,515	1,251	\$ 28.260.792	\$ 14.025.154
R8	1,452	1,505	\$ 27.971.110	\$ 13.453.577
<b>Total general</b>	<b>1,397</b>	<b>1,438</b>	<b>\$ 29.048.190</b>	<b>\$ 20.203.162</b>

Tabla 13: Transacciones y Montos por Transacción según Rubro del cliente

Al analizar la evolución en el tiempo de ambas variables de comportamiento en relación al Rubro de los clientes, se aprecia que para los clientes nuevos, las transacciones promedio y sus montos son prácticamente iguales independiente de su rubro. Sin embargo, al observar lo ocurrido para los clientes antiguos, se aprecia que existen

mayores diferencias en el comportamiento en relación al rubro, siendo algunas de estas mantenidas a lo largo de los años como es el caso del promedio de la variable  $m^*$ , donde se observa que los clientes del rubro “R2” han mantenido elevados valores para esta variable a lo largo de los años en comparación al resto de los rubros. Lo anterior se desprende de la “Ilustración 17” y la “Ilustración 18”

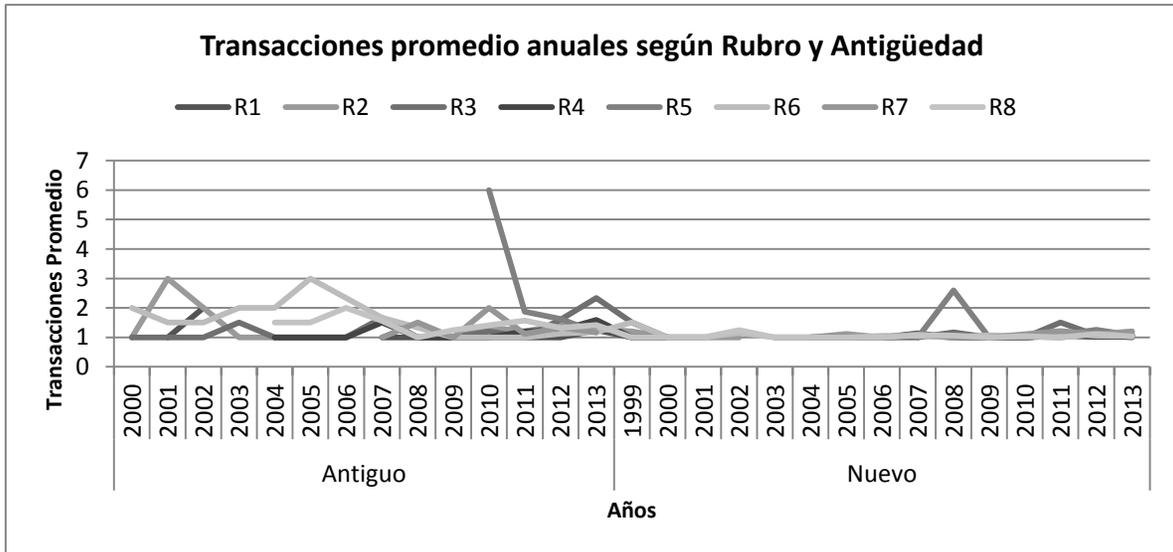


Ilustración 17: Evolución del promedio de Transacciones según Rubro y Antigüedad del cliente

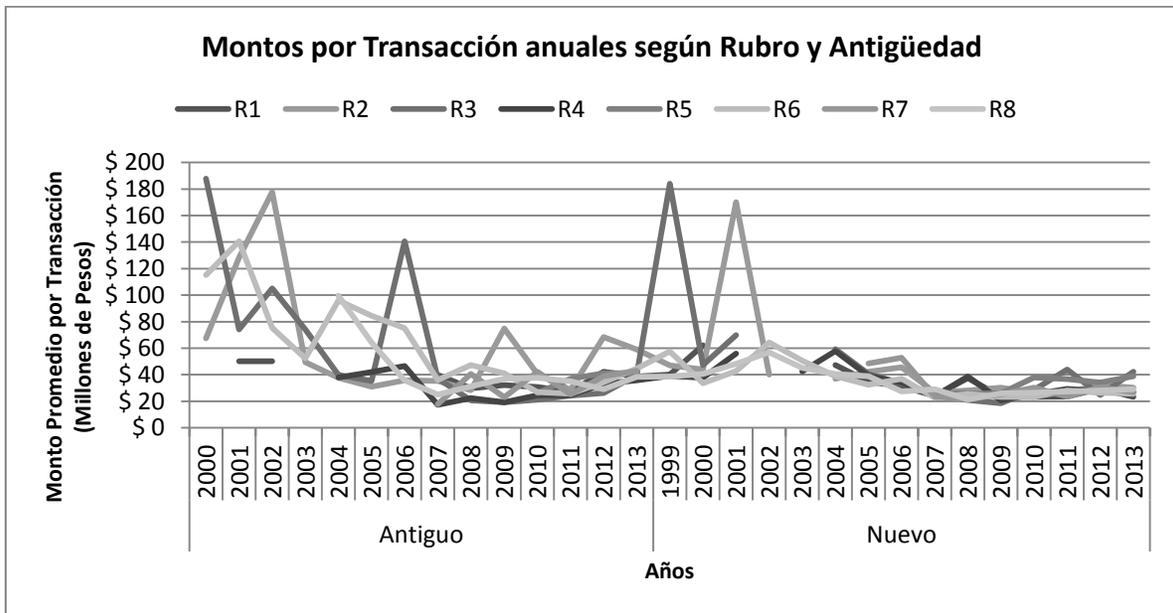


Ilustración 18: Evolución del Monto promedio por Transacción según Rubro y Antigüedad del cliente

Zona:

En base al resumen histórico para esta variable, se observa en la “Tabla 14” que los clientes de la zona central son los que poseen los mayores promedios para ambas variables, por lo que parece existir cierta dependencia entre la zona y las variables de comportamiento en estudio. Además, para el resto de las zonas, las interacciones y sus

montos además de ser más bajos poseen una desviación estándar mucho menor, lo que indica un comportamiento más similar para estos clientes.

Zona	Promedio de $x^*$	Desviación estándar de $x^*$	Promedio de $m^*$	Desviación estándar de $m^*$
Centro	1,436	1,575	\$ 29.614.954	\$ 22.084.591
Norte	1,262	0,673	\$ 27.081.774	\$ 8.989.316
Sur	1,204	0,543	\$ 26.147.447	\$ 10.472.765
<b>Total general</b>	<b>1,397</b>	<b>1,438</b>	<b>\$ 29.048.190</b>	<b>\$ 20.203.162</b>

Tabla 14: Transacciones y Montos por Transacción según Zona Geográfica del cliente

Por otro lado, al observar la evolución año a año de estas variables en la “Ilustración 19” y la “Ilustración 20”, junto con la variable “Antigüedad” de los clientes, se aprecia que para todas las “Zonas” el comportamiento de los clientes nuevos es similar para ambas variables de comportamiento en estudio. Una situación similar se tiene para el caso de los clientes antiguos, donde las interacciones promedios y sus montos se han movido de forma similar a lo largo del tiempo. Lo anterior indica que, a pesar de que en el resumen histórico se aprecian diferencias en el comportamiento de los clientes según su zona, al analizar la evolución en el tiempo estas diferencias no parecen ser tan relevantes en lo que se refiere al valor de los clientes.

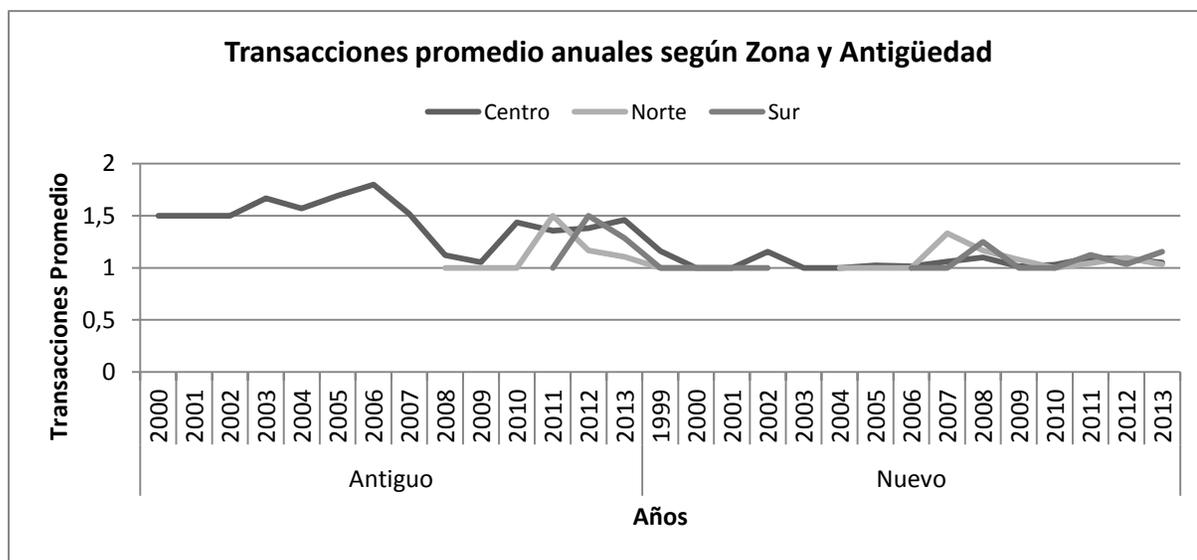


Ilustración 19: Evolución del promedio de Transacciones según Zona y Antigüedad del cliente

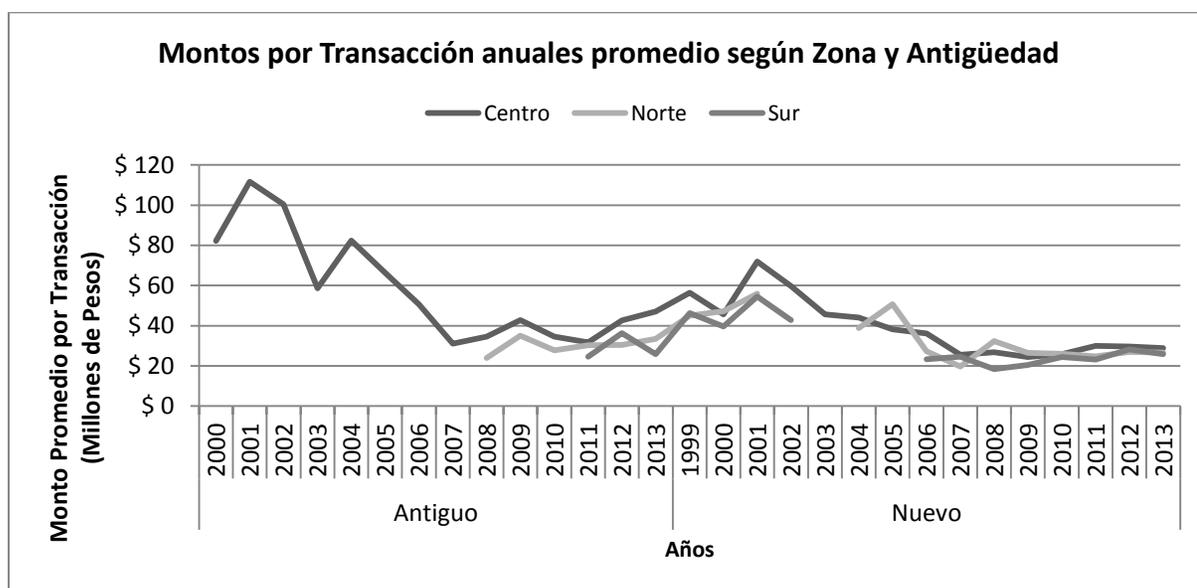


Ilustración 20: Evolución del Monto promedio por Transacción según Zona y Antigüedad del cliente

### Tamaño de flota:

En esta variable se producen las mayores diferencias en relación a las 2 variables en estudio, ya que se observa, de la “Tabla 15” que a mayor cantidad de flota general que manejan los clientes<sup>25</sup> mayores son los promedios a nivel resumen histórico tanto de transacciones realizadas como de los montos por transacción. Además se observa que los clientes tipo “Particular” poseen los promedios y desviaciones estándar más bajos para ambas variables, lo que indica que este tipo de clientes se comporta de forma similar.

Tamaño de flota	Promedio de $x^*$	Desviación estándar de $x^*$	Promedio de $m^*$	Desviación estándar de $m^*$
<b>Flotista grande (16 o más)</b>	4,246	5,640	\$ 61.509.868	\$ 74.204.968
<b>Flotista mediano (6-15)</b>	1,884	1,732	\$ 31.543.354	\$ 17.893.724
<b>Flotista pequeño (3-5)</b>	1,376	0,739	\$ 29.791.399	\$ 16.866.783
<b>Particular</b>	1,097	0,336	\$ 25.782.133	\$ 9.304.099
<b>Total general</b>	<b>1,397</b>	<b>1,438</b>	<b>\$ 29.048.190</b>	<b>\$ 20.203.162</b>

Tabla 15: Transacciones y Montos por Transacción según Categoría del cliente

Observando la “Ilustración 21” y la “Ilustración 22” se desprende que para los clientes antiguos, los promedios para  $x^*$  son mucho mayores a lo largo del tiempo para los clientes que poseen “Flotas Grandes” seguidos por los “Flotistas Medianos” lo que confirma lo analizado previamente en el resumen histórico. En relación a los promedios de la variable  $m^*$  se observa que las diferencias no son tan grandes como en la variable anterior.

<sup>25</sup> Recordar que esta variable indica la cantidad de camiones que el cliente posee independiente de la marca o el tipo que estos sean.

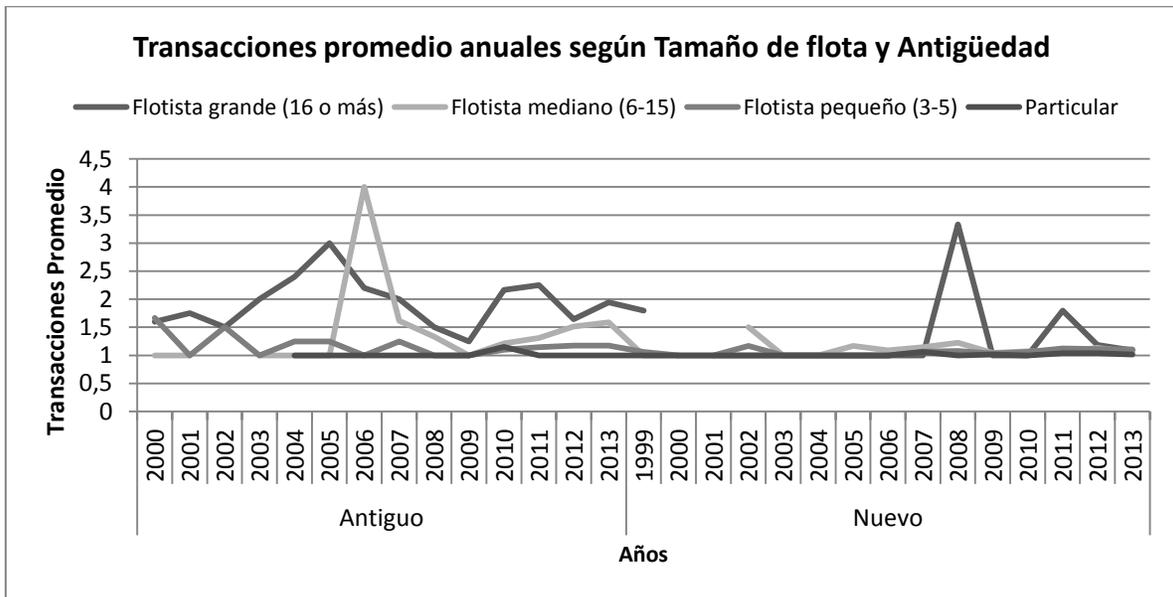


Ilustración 21: Evolución del promedio de Transacciones según Tamaño de flota y Antigüedad del cliente

Respecto a los clientes nuevos que año a año va adquiriendo la empresa, se aprecia que en los promedios de ambas variables de comportamiento las diferencias año a año no parecen ser muy grandes para clientes con distinto tamaño de flota, aunque los “*Flotistas grandes*” siguen presentando los mejores comportamientos.

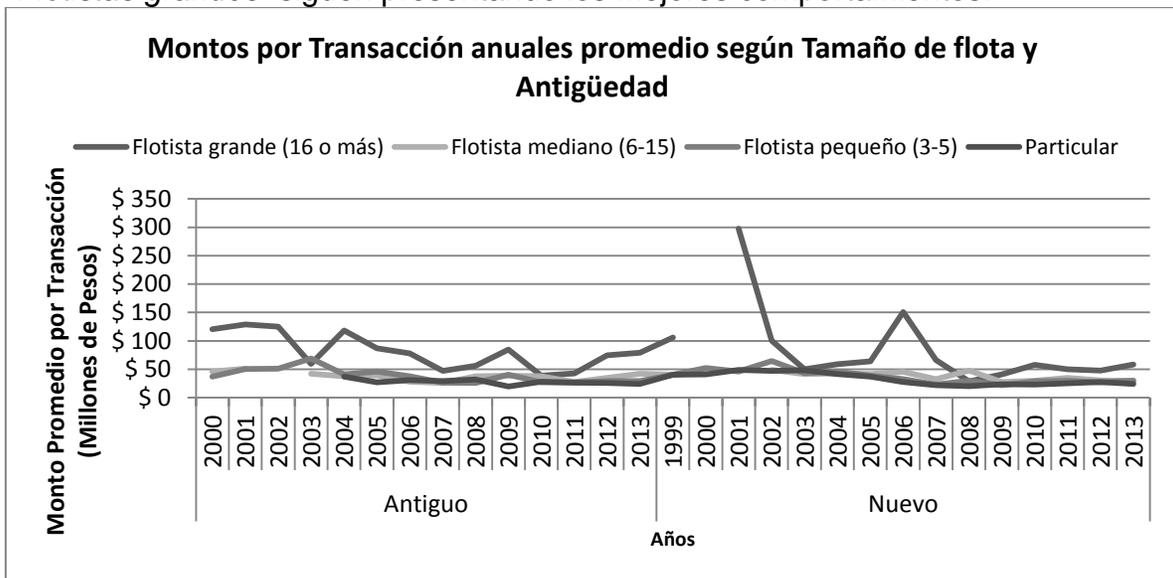


Ilustración 22: Evolución del Monto promedio por Transacción según Tamaño de flota y Antigüedad del cliente

### Vehículos de otro segmento:

Los clientes que poseen, además de vehículos medianos, vehículos de otros segmentos dentro de la marca, han realizado en promedio más interacciones y sus montos promedios por interacción son más altos que los clientes que solo poseen vehículos medianos. Estas diferencias notadas a nivel histórico y que se muestra en la “Tabla 16” parecen indicar que esta variable afecta el comportamiento transaccional de los clientes.

Tiene vehículos de otro segmento	Promedio de $x^*$	Desviación estándar de $x^*$	Promedio de $m^*$	Desviación estándar de $m^*$
NO	1,2731	1,0491	\$ 28.182.190	\$ 19.975.891
SI	2,1648	2,7146	\$ 34.393.501	\$ 20.807.546
<b>Total general</b>	<b>1,397</b>	<b>1,438</b>	<b>\$ 29.048.190</b>	<b>\$ 20.203.162</b>

Tabla 16: Transacciones y Montos por Transacción según "vehículos de otro segmento"

Como se aprecia en la "Ilustración 23" y en la "Ilustración 24" la evolución anual de las variables de comportamiento en estudio indica que para los clientes antiguos se diferencian según esta variable de perfilamiento, siendo los clientes que si poseen vehículos de otros segmentos los que año a año realizan un promedio de interacciones levemente mayor que lo clientes que no poseen otros tipos de vehículos dentro de la marca. Esta diferencia no es tan notoria al analizar la evolución de la variable  $m^*$  donde no se observan importantes diferencias año a año que hacen difícil establecer una relación entre esta variable de comportamiento y la variable de perfilamiento en estudio.

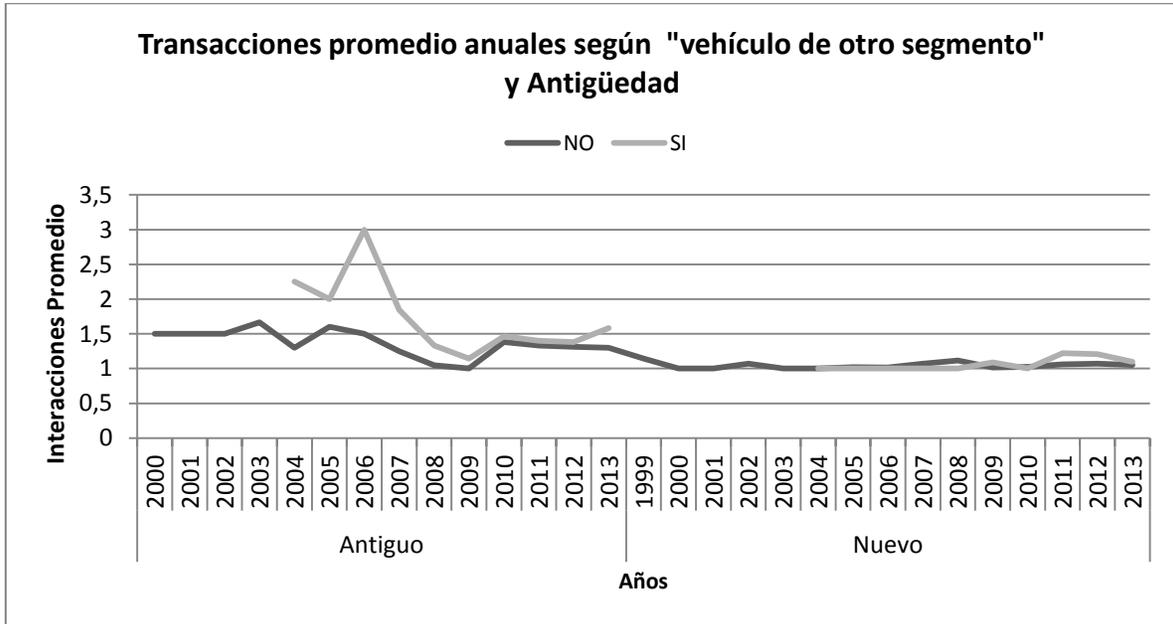


Ilustración 23: Evolución del promedio de Transacciones según "Vehículos de otro segmento" y Antigüedad

Por otro lado, observando la evolución anual del comportamiento de los clientes nuevos, se tiene que para ambas variables son los clientes que si poseen vehículos de otros segmentos, además de medianos, los que poseen los valores más altos, sin embargo, estas diferencias parecen ser pequeñas.

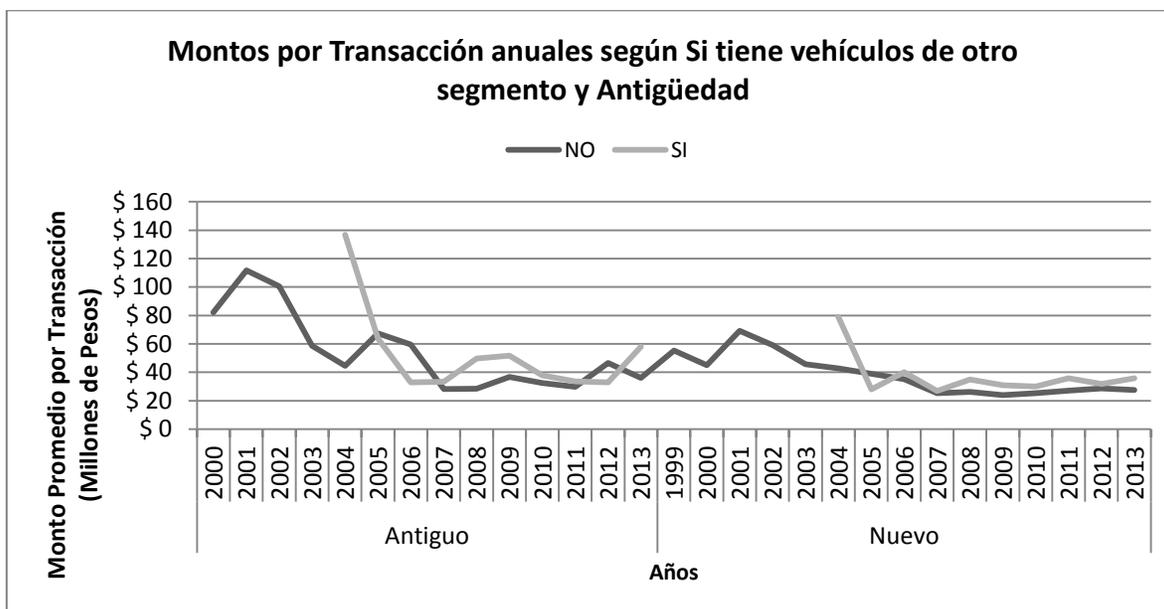


Ilustración 24: Evolución del Monto promedio por Transacción según Si tiene vehículos de otro segmento y Antigüedad del cliente

**IV.2.2.2.-Decisión de compra ( $C_t$ ), decisión del número de transacciones condicionada ( $x_t^{**}$ ) y su relación con variables de comportamiento.**

Naturaleza:

En la “Tabla 17” Se muestra la relación entre la decisión de compra del cliente y su naturaleza para la totalidad de observaciones. Se aprecia que son los clientes tipo empresa los que en general han vuelto a comprar en el futuro.

Naturaleza	Decisión			
	# Observaciones	%	# Observaciones	%
<b>Empresa</b>	369	10,50%	3146	89,50%
<b>Persona</b>	87	3,38%	2490	96,62%

Tabla 17: Resumen de la decisión de compra según la Naturaleza del cliente en la observación

Analizando la evolución de las decisiones de compra en el tiempo presentada en la “Ilustración 25”, se observa que año a año la proporción de clientes tipo empresa que compra nuevamente es mucho mayor a la proporción de personas que lo hace. Además se observa que el porcentaje de recompra (compras realizadas por clientes antiguos, es decir, que han comprado en un periodo previo al actual) de clientes tipo persona es bastante parejo año a año, mientras que el de empresas tiende a variar un poco más en el tiempo. Este resultado es relevante, ya que deja en evidencia una dependencia entre la decisión de compra y la Naturaleza del cliente.

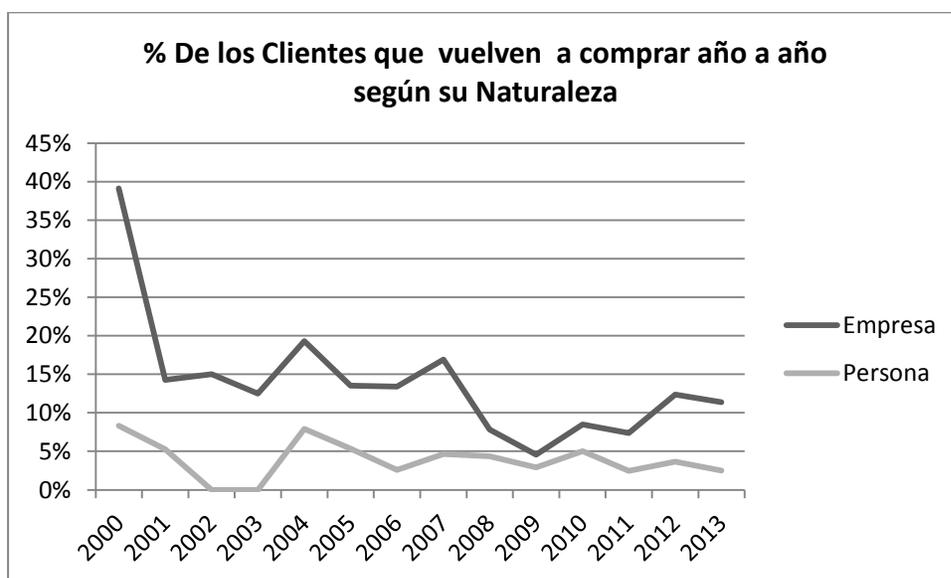


Ilustración 25: Evolución de las Tasas de recompra según Naturaleza de los clientes

### Rubro:

Considerando todas las observaciones hasta el 2013 cuyos resultados se presentan en la “Tabla 18”, se tiene que las tasas de recompra a nivel de rubros es bastante pareja, siendo los rubros “R3” y “R5” los que poseen tasas levemente mayores. Si se consideran solo las decisiones de compra positiva, se tiene que la mayoría de ellas corresponden al rubro “R8”, sin embargo esto se debe a la alta cantidad de clientes pertenecientes a este rubro y no a que estos representen un mejor comportamiento.

Rubro	Decisión			
	Compra		No compra	
	# Observaciones	%	# Observaciones	%
R1	21	3,18%	639	96,82%
R2	84	9,58%	793	90,42%
R3	38	13,82%	237	86,18%
R4	45	6,31%	668	93,69%
R5	24	13,64%	152	86,36%
R6	81	5,50%	1391	94,50%
R7	30	10,91%	245	89,09%
R8	133	8,09%	1511	91,91%

Tabla 18: Resumen de la decisión de compra según el Rubro del cliente en la observación

Respecto a la evolución de las tasas de recompra según el Rubro de los clientes se aprecia en la “Ilustración 26” que estas presentan una alta variabilidad, sin embargo los clientes del rubro “R3”, en general, han mantenido sus tasas altas en comparación al resto.

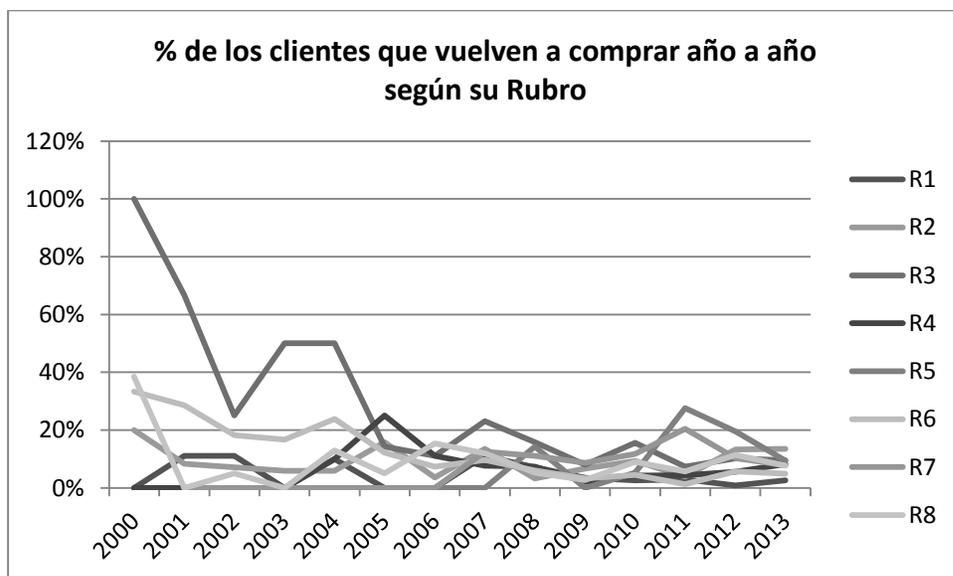


Ilustración 26: Evolución de las Tasas de recompra según el Rubro de los clientes

#### Tiene Vehículo de Otro segmento:

Los clientes que poseen vehículos de otros segmentos dentro de la marca, han presentado un mejor comportamiento en cuanto a las compras futuras, a pesar de que la mayoría de las recompras las han realizado los clientes que solo poseen vehículos medianos. Esto último se debe a que estos clientes representan mayoría en cuanto al total de clientes dentro de la base, y por esta misma razón se posee una mayor cantidad de observaciones por parte de ellos.

Tiene Vehículos de otro segmento	Etiquetas de columna			
	Compra		No compra	
	# Observaciones	%	# Observaciones	%
<b>NO</b>	307	5,83%	4963	94,17%
<b>SI</b>	149	18,13%	673	81,87%

Tabla 19: Resumen de la decisión de compra según si el cliente tiene Vehículos de otro segmento

Al analizar la evolución en el tiempo de las tasas de recompra de los clientes según si tienen vehículos de otros segmentos, en la "Ilustración 27" se confirma lo observado a nivel histórico, es decir, que los clientes que si poseen otros tipos de vehículos dentro de la marca parecen estar más dispuestos a volver a comprar nuevamente.

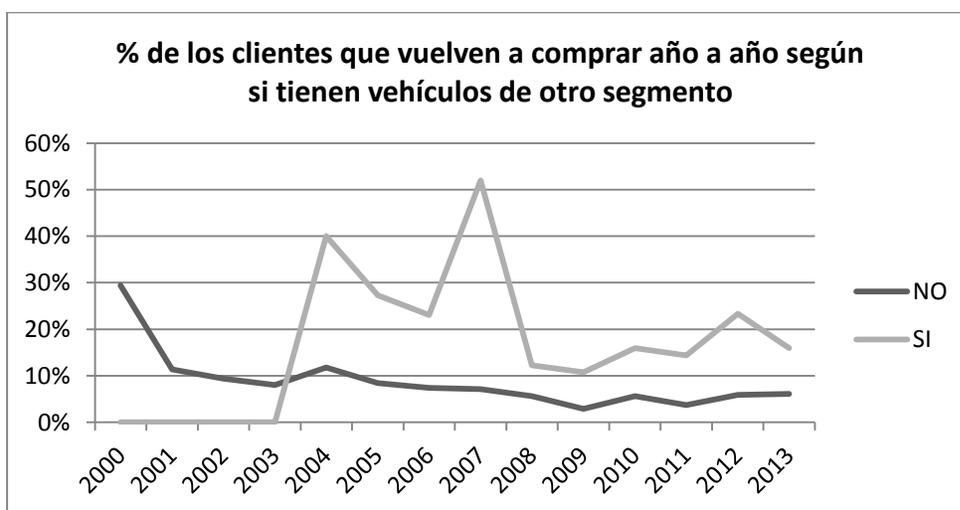


Ilustración 27: Evolución de las Tasas de recompra según si el cliente tiene o no vehículos de otro segmento

### Decisión de compra y Recencia:

Como se aprecia en el “Anexo E” son pocas las observaciones en donde el cliente decide comprar teniendo una recencia alta (mayor a 6). Por esta razón se propone la transformación de esta variable numérica en una categórica, de aquí en adelante “*Recencia Anual agrupada*” ( $R^*$ ), que mantiene los 5 estados de recencia iniciales y los resultados correspondientes al resto los agrupa bajo la etiqueta “6 o más”.

Recencia	Decisión			
	Compra		No compra	
	# Observaciones	%	# Observaciones	%
1	232	12,64%	1604	87,36%
2	90	7,38%	1130	92,62%
3	55	5,88%	881	94,12%
4	21	3,20%	636	96,80%
5	18	3,79%	457	96,21%
6 o +	40	4,13%	928	95,87%

Tabla 20: Resumen decisión de compra según la recencia del cliente en la observación

En la “Tabla 20” se aprecia que a medida que aumenta la recencia del cliente al momento de la observación, la proporción de los clientes que compra va disminuyendo hasta llegar a la Recencia “6 o +” donde esta vuelve a aumentar levemente.

### Decisión de compra y “Años distintos en que el cliente ha comprado” ( $A$ ) al momento de la observación:

Por razones similares a la agrupación realizada con la variable frecuencia, se agrupa la variable a bajo 3 etiquetas, dando origen a la variable  $A^*$ . La relación entre esta nueva variable y la decisión de compra tomada por los clientes a nivel de observaciones, se presenta en la “Tabla 21”. Aquí se aprecia que a mayor valor de  $A^*$  aumenta

considerablemente la proporción de clientes que decide comprar, por lo tanto esta parece ser una variable importante a la hora de describir la decisión de compra.

Decisión	Decisión			
	Compra		No compra	
	# Observaciones	%	# Observaciones	%
<b>1</b>	290	5,05%	5458	94,95%
<b>2</b>	91	43,13%	120	56,87%
<b>3 o +</b>	75	56,39%	58	43,61%

Tabla 21: Resumen de decisión de compra según la cantidad de años distintos en que el cliente ha comprado al momento de la observación

Número de interacciones a realizar dado que el cliente va a comprar ( $x_t^{***}$ ), y número de interacciones realizadas por el cliente en el periodo anterior al momento de la observación ( $x_{t-1}$ )

Utilizando la agrupación realizada previamente para esta variable, se aprecia en la “Tabla 22” que la mayoría de los clientes decide realizar solo una transacción, a diferencia de los clientes que previamente han realizado “3 o +” transacciones, quienes parecen preferir volver a realizar un número mayor de transacciones.

Número de transacciones en el periodo anterior	Número de transacciones periodo actual				Total general
	1	2	3	4 o +	
<b>0</b>	188	27	6	3	<b>224</b>
<b>1</b>	128	30	8	4	<b>170</b>
<b>2</b>	22	13	3	2	<b>40</b>
<b>3</b>	5	3	2	2	<b>12</b>
<b>4 o +</b>	2	2	2	4	<b>10</b>
<b>Total general</b>	<b>345</b>	<b>75</b>	<b>21</b>	<b>15</b>	<b>456</b>

Tabla 22: Resumen de decisión del número de transacciones a realizar según la decisión tomada por el cliente el año anterior.

#### IV.2.2.3.-El efecto del PIB

##### PIB y Decisión de compra ( $C_t$ ):

En la “Ilustración 28” se muestra la evolución del crecimiento porcentual del PIB y la evolución de la proporción de clientes antiguos que deciden año a año comprar. Se aprecia que existe una relación entre ambas variables, es decir, cuando el PIB aumenta, parece incrementar el porcentaje de clientes que decide comprar.

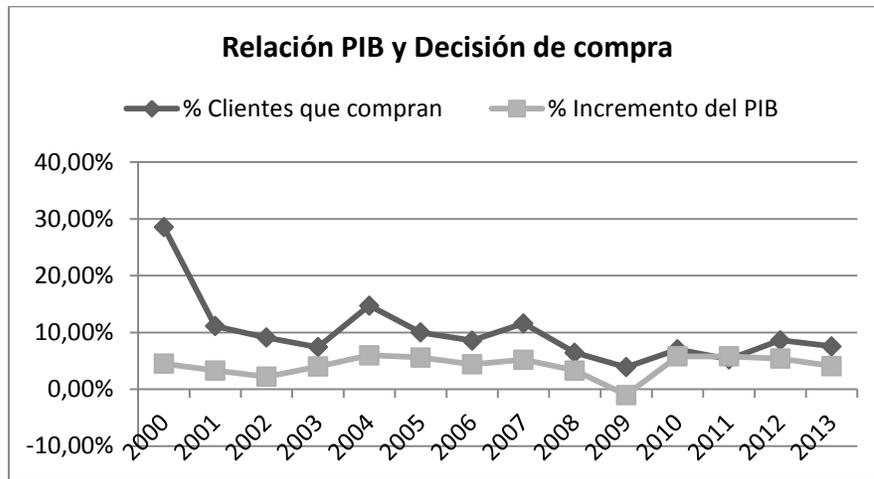


Ilustración 28: Relación entre el incremento del PIB y la decisión de compra

PIB y decisión del número de transacciones ( $x_t^{**}$ ):

Como se aprecia en la “Ilustración 29”, y al igual que en el caso de la decisión de compra, el número de transacciones promedio realizadas por los clientes dado que estos deciden comprar, parece estar fuertemente relacionado a los movimientos del que sufre el PIB. Del gráfico se desprende la posible existencia de una correlación positiva entre ambas variables.

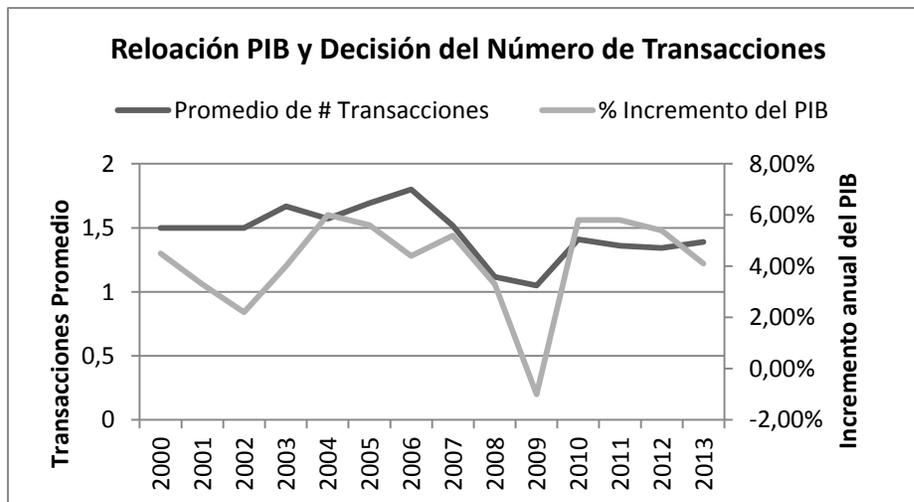


Ilustración 29: Relación entre el Incremento del PIB Y el Número promedio de transacciones

**IV.3.-Selección de variables relevantes.**

En la sección anterior se describió la relación entre las variables de comportamiento relevantes y las variables de perfilamiento, sin embargo, no se sabe si las diferencias que se observaron para los distintos valores de una variable de perfilamientos son relevantes e indican dependencia con las variables de comportamiento en estudio. Por esta razón, a continuación se analiza la dependencia entre variables de comportamiento y de perfilamiento a través de 2 métodos:

### IV.3.1.- Test ANOVA

En este caso las variables de comportamiento relevantes que se buscan relacionar son el “número de transacciones” ( $x^*$ ) y el “monto por transacción” ( $m^*$ ) de los clientes.

Las variables con las que se busca relación, corresponden a las variables de perfilamiento descritas: Tipo de cliente, rubro, naturaleza, tamaño de flota y si tiene o no vehículos de otro segmento.

El test se aplica a la totalidad de la data donde para cada cliente que compone la cartera de tiene el valor resumen de sus variables de comportamiento, es decir, su número de interacciones históricas y su monto por interacción y los resultados se muestran a continuación. En particular, es importante notar, que si se trabaja directamente con la variable  $x^*$  se estaría pasando por alto que hay clientes que tienen distinta antigüedad y que por esta razón, han tenido la oportunidad de realizar más transacciones. Por esta razón, la variable resumen que se utilizará corresponde al número de transacciones por año en que el cliente ha comprado

Variable Número de Transacciones por año en que el cliente ha comprado ( $x^{**}$ )

Se observa, de la columna final de la “Tabla 23”, que con un 95% de confianza las variables “Tipo de Cliente”, “Rubro”, “Tamaño de flota” y “Tiene vehículo de otro segmento” influyen en la variable de comportamiento en estudio, por lo tanto son candidatas a ser consideradas en los modelos.

Variable		DF	Sum sq	Mean sq	F value	pr(>F)
<b>Tipo de cliente</b>	Inter-grupos	1	2,47	2,4655	27,58	1,68e-07
	Intra-grupos	1870	167,15	0,0894		
	Total	1871	169,62			
<b>Rubro</b>	Inter-grupos	7	2,08	0,29754	3,31	0,00166
	Intra-grupos	1864	167,53	0,08988		
	Total	1871	169,61			
<b>Zona</b>	Inter-grupos	2	0,05	0,02418	0,267	0,766
	Intra-grupos	1869	169,57	0,09073		
	Total	1871	169,62			
<b>Tamaño de Flota</b>	Inter-grupos	3	11,25	3,750	44,23	<2e-16
	Intra-grupos	1868	158,37	0,085		
	Total	1871	169,62			
<b>Tiene vehículo de otro tipo</b>	Inter-grupos	1	1,91	1,9077	21,27	4,26e-06
	Intra-grupos	1870	167,71	0,0897		
	Total	1871	169,62			

Tabla 23: Resultados ANOVA variable Número de Interacciones

### Variable Monto por Transacción ( $m^*$ )

De forma similar a lo realizado para la variable anterior, se realiza un Test Anova para analizar la dependencia de esta variable con distintos atributos del cliente. Los resultados en este caso se encuentran en el “Anexo F”, donde se desprende que, al igual que en el caso anterior, todas las variables son estadísticamente significativas a un 95% de confianza a excepción de la variable “Zona”. Estas variables se deben utilizar en el modelo que proyecta los montos por transacción de los clientes.

Es importante considerar que tanto la variable “*Tamaño de flota*” como la variable “*Tiene vehículo de otro segmento*” no son estáticas y podrían haber cambiado en el transcurso del tiempo y/o lo pueden hacer de cara al futuro. Es decir, un cliente que hoy en día si no posee vehículos de otro segmento es posible que en el futuro si posea, o un cliente que comenzó a comprar hace 5 años y que al principio no tenía vehículos de otro segmento puede haber adquirido otro un par de años más tarde. Respecto a esta variable, se poseen las variaciones que año a año han experimentado los clientes, por lo que solo se debería tener en consideración una posible evolución en el futuro. Sin embargo, para el caso de la variable “Tamaño de flota” no se posee a nivel de clientes un historial que registre los movimientos de esta variable, por lo que hoy en día esta variable puede estar desactualizada, es decir, un cliente puede haber pasado de ser “Flotista pequeño” a “Flotista grande” y no se sabrá lo anterior hasta que se realice una actualización de la información del cliente. Debido a que estas no están actualizadas, se decide no considerarla en los modelos que se proponen en este trabajo.

### **IV.3.2.- Test Chi-Cuadrado**

Para evaluar las variables que resultan significativas en la decisión de compra de los clientes ( $C_t$ ) que se desprenden de las observaciones generadas de la data, se realizan test chi-cuadrado, cuyos resultados se resumen en la “Tabla 24”

Variable	X-squared	df	p-value
Zona	4,6281	2	0,09886
Naturaleza	108,9053	1	< 2,2e-16
Rubro	64,0376	7	2,347E-08
R*	116,4008	5	< 2,2e-16
A*	895,8665	2	< 2,2e-16

Tabla 24: Resultados chi-cuadrado para la variable decisión de compra

Aquí se observa que, con un 95% de confianza, se puede decir que existe dependencia entre la variable en estudio y el resto de las variables presentadas con excepción de la variable zona, cuyo p-valor es mayor al 5%, por lo que debe ser rechazada la dependencia entre ambas.

Para la evolución del número de transacciones dado que los clientes deciden comprar ( $x_t^{***}$ ), se testea el efecto de otras variables de comportamiento en su decisión (Recordar que esta se transformó en una variable categórica).

Variable	X-squared	df	p-value
Zona	5,0846	6	0,533
Naturaleza	16,0611	3	0,001102
Rubro	17,9331	12	0,1177
R*	22,1405	15	0,1042
A*	29,7986	6	4,29E-02
$x_{t-1}^{***}$	82,2369	12	1,54E-09

Tabla 25: Resultados Chi-cuadrado evolución transacciones

Es importante mencionar que la relación entre las variables de decisión de compra y del número de transacciones se testeará en el mismo modelo, incluyéndola como variable explicativa y observando si los coeficientes entregados por el modelo para dicha variable resultan ser significativos.

#### **IV.4.-Aplicación y validación de los modelos.**

Para la estimación del valor potencial de los clientes se proponen 2 modelos: El modelo BG/NBD y un modelo de Markov-Logit. La aplicación y validación de los modelos se realiza en las siguientes etapas:

##### **IV.4.1.-Periodo Calibración-Validación**

La data total se divide de la siguiente manera:

- Periodo de calibración: Información transaccional hasta el 31 de Diciembre del 2011. Hasta esta fecha se tienen:
  - 1148 clientes.
  - 1551 transacciones.
  - 3451 observaciones del comportamiento anual de los clientes.
- Periodo de validación: Información transaccional desde el 1 de Enero del 2012 hasta el 31 de Diciembre del 2013

Lo que se busca es intentar predecir el comportamiento en los próximos 2 años de los clientes que han comprado por primera vez antes del 31 de Diciembre del 2011.

##### **IV.4.2.-Estimación del número de transacciones**

###### **IV.4.2.1.-Modelo BG/NBD**

Este modelo ocupa variables que resumen el comportamiento transaccional histórico del cliente a la fecha, y en base a estas variables intenta proyectar el comportamiento futuro.

Se realizaron 3 variaciones de este modelo:

- BG/NBD 1: Estimación utilizando solo variables de comportamiento, en particular las variables utilizadas corresponden a:
  - Número de transacciones que los clientes han realizado hasta finales del año 2011 ( $x_{2011}^*$ ).
  - Antigüedad, en años, del cliente hasta finales del año 2011 ( $T_{2011}$ ).
  - Tiempo transcurrido, en años, entre la primera y la última transacción hasta finales del año 2011 ( $t_{x_{2011}}$ ).
- BG/NBD 2: Estimación utilizando, además de las variables anteriores, la variable “*Tipo de Cliente*” para establecer diferencias en los parámetros calibrados por el modelo.
- BG/NBD 3: Estimación utilizando, además de las variables anteriores, la variable “*Rubro*” y la variable “*Tiene vehículos de otro segmento*” para establecer diferencias en los parámetros a calibrar por el modelo.

Es importante recordar que en este modelo se busca determinar los parámetros  $r, \alpha$  que maximizan la verosimilitud de la de la distribución Binomial Negativa que representa el número de transacciones realizadas por los clientes, además de los parámetros  $a, b$  que maximizan la verosimilitud de la distribución Beta Geométrica que representa la fuga de los clientes. Lo anterior se realiza utilizando la siguiente ecuación que representa la verosimilitud a nivel individual:

$$L(r, \alpha, a, b | X = x, t_x, T) = \frac{B(a, b + x)}{B(a + b)} \cdot \frac{\Gamma(r + x)\alpha^r}{\Gamma(r)(\alpha + T)^{r+x}} + \delta_{x>0} \frac{B(\alpha + 1, b + x - 1)}{B(a, b)} \cdot \frac{\Gamma(r + x)\alpha^r}{\Gamma(r)(\alpha + t_x)^{r+x}}$$

En la “*Tabla 26*” se muestran los parámetros obtenidos al maximizar la verosimilitud para las variaciones 1 y 2 del modelo.

Parámetro	Variación		
	BG/NBD1	BG/NBD 2	
		Empresa	Persona
<b>r</b>	0,166	0,200	0,215
<b>α</b>	1,245	1,068	4,036
<b>a</b>	0,577	0,724	0,192
<b>b</b>	1,169	1,984	0,191
<b>LL</b>	-1054,55	-760,75	-274,01

Tabla 26: Parámetros obtenidos en la estimación del modelo BG/NBD

Con los parámetros obtenidos para cada una de las variaciones del modelo, se procede a proyectar el número de transacciones a realizar por los clientes durante el periodo de validación.

A continuación se presentan los resultados de las distintas variaciones del modelo para el periodo de validación. Además se presentan distintas métricas de ajuste que buscan cuantificar el error de los resultados entregados por cada uno de ellos.

Como se aprecia en la “Tabla 27”, la segunda variación del modelo predice de mejor forma a nivel agregado el número de transacciones a realizar por la cartera de clientes de cara al año 2012, lo que se observa en el menor Error Porcentual Agregado que posee en comparación a los otros 2. Sin embargo, a nivel individual se observa que la tercera variación del modelo posee los mejores resultados, lo que se observa en el menor valor tanto para el MAPE como para el WMAPE obtenidos.

Modelo	Proyección 2012	Real 2012	Error porcentual agregado	MAE	MPE	MAPE	WMAPE
BG/NBD 1	89,18	133	32,95%	0,160	-81,62%	87,19%	87,58%
BG/NBD 2	89,40	133	32,78%	0,158	-80,64%	86,36%	86,67%
BG/NBD 3	88,07	133	33,77%	0,156	-81,57%	84,35%	85,10%

Tabla 27: Resultados y errores BG/NBD proyección año 2012

En la “Tabla 28” se observa que para el año 2013, la segunda variación del modelo vuelve a entregar la mejor estimación a nivel agregado, sin embargo nuevamente es la tercera variación del modelo la que entrega los mejores resultados a nivel individual.

Modelo	Proyección 2013	Real 2013	Error porcentual agregado	MAE	MPE	MAPE	WMAPE
BG/NBD 1	80,11	111	27,83%	0,145	-88,29%	89,14%	89,12%
BG/NBD 2	80,77	111	27,24%	0,144	-87,54%	88,35%	88,38%
BG/NBD 3	78,25	111	29,51%	0,142	-87,77%	87,77%	88,22%

Tabla 28: Resultados y errores BG/NBD proyección año 2013

Finalmente, la “Tabla 29” muestra que, considerando la proyección acumulada para los años 2012 y 2013, se vuelve a repetir la situación ya mencionada en los casos anteriores, es decir, la segunda variación entrega mejores resultados a nivel agregado y la tercera ofrece los mejores resultados a nivel individual.

Modelo	Proyección acum 2012-2013	Real acum 2012-2013	Error porcentual agregado	MAE	MPE	MAPE	WMAPE
BG/NBD 1	169,29	244	30,62%	0,283	-76,40%	84,61%	84,57%
BG/NBD 2	170,17	244	30,26%	0,278	-74,93%	83,25%	83,33%
BG/NBD 3	166,32	244	31,84%	0,273	-76,16%	81,04%	81,83%

Tabla 29: Resultados y errores BG/NBD proyección acumulada años 2012-2013

En conclusión, se puede apreciar que al aumentar la cantidad de variables utilizadas para estimar los parámetros del modelo, aumenta la precisión a nivel individual, por lo que la variación “BG/NBD 3” ofrece los mejores resultados para la estimación del número de transacciones según este modelo.

A modo de referencia, los resultados presentados a nivel individual son mejores que los obtenidos en [15], donde se aplica un modelo similar tanto para estimar el valor de los clientes como en los supuestos utilizados para calcular los errores asociados. A diferencia de lo realizado en ese trabajo, aquí no se aplican constantes para tratar de reducir el error, debido a que se corre el riesgo de sobreajuste.

#### ***IV.4.2.2.-Modelo Logit-Poisson-Markov (L-P-M)***

En este modelo, a diferencia del anterior, no se utilizan variables de resumen del comportamiento histórico, sino que utiliza todas las observaciones obtenidas de la cartera de clientes para calibrar los parámetros usados para la estimación del número de transacciones.

Como se explicó en la sección “II.5.2.2”, este modelo entrega la esperanza del número de transacciones que un cliente realizará en cada periodo. Para esto estima las probabilidades de compra en un periodo utilizando regresiones logísticas, y el número de transacciones dado que el cliente compra utilizando una Regresión de Poisson truncada en 0. Multiplicando los valores de salida de cada modelo se obtiene la esperanza del número de interacciones para el siguiente periodo. Posteriormente, se utilizan Cadenas de Markov para representar la evolución en el tiempo tanto de las probabilidades de compra, como de las interacciones a realizar año a año (y así, con estas 2 variables, se representa la evolución del número de transacciones esperadas).

Respecto a la Probabilidad de compra y su evolución en el tiempo, se realizan 3 variaciones del modelo Logit que las estima:

- L1: Estimación utilizando solo variables de comportamiento, las que son usadas como estados de la cadena de Markov que representa la evolución del comportamiento de los clientes, en particular las variables utilizadas para definir los estados corresponden a:
  - Años distintos en que el cliente ha comprado hasta la fecha ( $A_t^*$ ).
  - Hace cuántos años que el cliente compró por última vez ( $R_t^*$ ).
- L2: Estimación utilizando, además de las variables anteriores, la variable “*Tipo de Cliente*” para establecer diferencias en los parámetros calibrados por el modelo.
- L3: Estimación utilizando, además de las variables anteriores, las variables “*Rubro*”, “*Tiene vehículo de otro segmento*” y “*PIB*” para establecer diferencias en los parámetros a calibrar por el modelo.

En relación al Modelo Logit, la “Tabla 30” muestra los parámetros obtenidos y su significancia para la tercera variación del modelo. Respecto a la interpretación de los resultados entregados por el modelo se aprecia, por ejemplo, que los clientes que han comprado en el periodo anterior son los que tienen la mayor probabilidad de volver a comprar en el periodo actual. Además, el hecho de haber comprado en más de un periodo previo aumenta enormemente la probabilidad de que el cliente vuelva a comprar. Por otro lado, los clientes tipo empresa son más propensos a comprar que los

tipo persona. Tener vehículos de otro segmento afecta positivamente la probabilidad de compra. Respecto a los rubros, los clientes del Rubro “R5” parecen ser los más propensos a interactuar nuevamente. Además, tal como se intuía, el PIB tiene un efecto positivo en la probabilidad de compra.

Variable		Beta	Significancia	Exp (Beta)
Intercepto		-3,23045	<2e-16	0,039
R*	[2]	-0,41394	0,02948	0,661
	[3]	-1	0,000154	0,388
	[4]	-1,06534	0,000319	0,344
	[5]	-1,97307	0,001014	0,139
	[6 o +]	-0,54649	0,029403	0,578
A*	[2]	2,32395	<2e-16	10,215
	[3]	2,50485	<2e-16	12,241
Naturaleza	[Persona]	-0,55122	0,002117	0,576
Rubro	[R2]	0,4535	0,18029	1,573
	[R3]	0,76883	0,043965	2,157
	[R4]	0,0747	0,841205	1,077
	[R5]	1,1909	0,009521	3,290
	[R6]	0,23607	0,468015	1,266
	[R7]	0,96237	0,016345	2,617
	[R8]	0,38031	0,226873	1,462
Otro vehiculo	[SI]	0,78563	0,00000673	2,193
PIB		0,095	0,004464	1,099

Tabla 30: Coeficientes modelo logístico L 3

Respecto al PIB, es importante notar que este genera que las matrices de transición de probabilidad obtenidas a través del modelo variarán año a año, por lo que genera “heterogeneidad en el tiempo”. Como muestra de lo anterior, el “Anexo G” muestra la matriz de transición de cara al 2012 y de cara al 2013 para un determinado perfil.

Es importante notar que dentro de la variable “Rubro”, existen opciones que por sí solas no parecen ser significativas (R2, R4, R6, R8), sin embargo, a la aplicar el “Test de Wald” y evaluar la significancia global de la variable, se tiene que efectivamente es significativa a un 95% de confianza. Lo anterior se observa de los resultados presentados en la “Tabla 31”.

X2	df	P(>X2)
15,8	7	0,027

Tabla 31: Resultados Test de Wald para la variable Rubro

Para evaluar la capacidad predictiva de las probabilidades entregadas por el modelo Logit, se presentan las curvas ROC de la variación “L3”<sup>26</sup>. En la “Ilustración 30” se presenta la curva de las probabilidades obtenidas para el año 2012, la que posee un

<sup>26</sup> Que, como se verá posteriormente, resulta ser la mejor variación del modelo

área bajo la curva de 0,74 lo que indica que el modelo gana 0,24 sobre la probabilidad de lanzar una moneda aleatoriamente. A modo de referencia, este resultado obtenido es mejor a los presentados en [18]

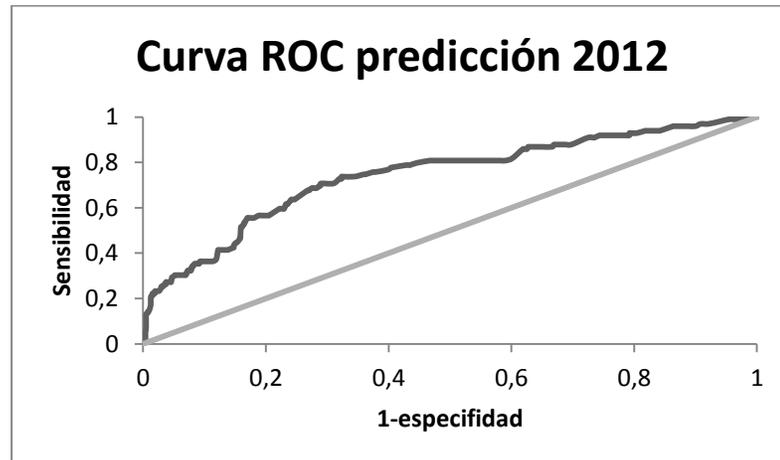


Ilustración 30: Curva ROC probabilidades de compra 2012 modelo L3

Respecto a las predicciones obtenidas de cara al año 2013, se presenta la curva ROC asociada en la “Ilustración 31”. El área bajo la curva corresponde a 0,72, la que es un poco menor a la predicción obtenida para el año 2012. Lo anterior indica que la calidad de predicción disminuye al aumentar la cantidad de periodos en que se quiere predecir, sin embargo, esta disminución no es demasiado grande, por lo menos entre un año y otro.

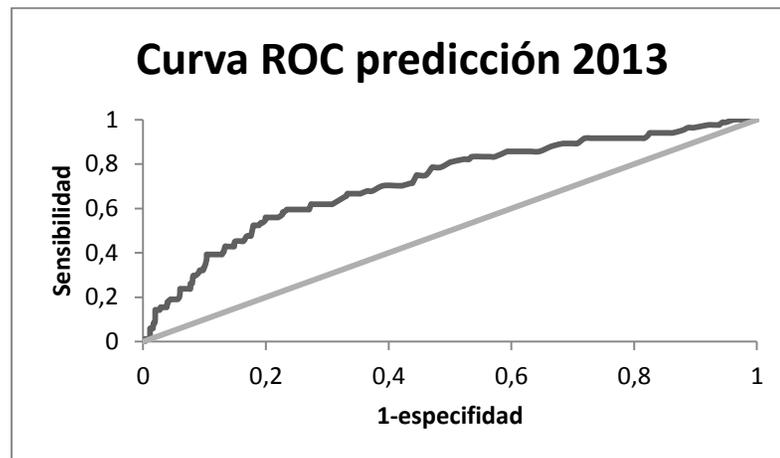


Ilustración 31: Curva ROC probabilidades de compra 2013 modelo L3

Respecto al Número de Transacciones dado que el cliente compra ( $x_t^{***}$ ), se utilizaron como variables explicativas:

- El número de transacciones realizadas por el cliente el año anterior ( $x_{t-1}^{***}$ ).
- La “Naturaleza” del cliente.
- El “PIB”.

De esta forma, uniendo los resultados obtenidos de la Regresión Logística y de la Regresión de Poisson y manteniendo la lógica de las variaciones presentadas, de obtienen tres variaciones del modelo para el cálculo de la esperanza del número de interacciones.

Como se aprecia en la “Tabla 32”, la tercera variación del modelo predice de mejor forma a nivel agregado el número de interacciones a realizar por la cartera de clientes de cara al año 2012, lo que se observa en el menor Error Porcentual Agregado que posee en comparación a los otros 2. Además, este mismo modelo es el que predice mejor a nivel individual, lo que se observa en el menor valor tanto para el MAPE como para el RMAPE obtenidos.

Modelo	Proyección 2012	Real 2012	Error porcentual agregado	MAE	MPE	MAPE	WMAPE
L-P-M 1	108,14	133	18,69%	0,172	-81,00%	83,51%	84,45%
L-P-M 2	104,32	133	21,56%	0,168	-80,73%	83,34%	84,20%
L-P-M 3	134,42	133	1,07%	0,171	-68,27%	75,51%	75,99%

Tabla 32: Resultados y errores L-P-M proyección año 2012

En la “Tabla 33” se observa que para el año 2013 la segunda variación del modelo predice de mejor forma el número de interacciones a realizar por la cartera de clientes a nivel agregado. Sin embargo, al considerar los indicadores MAPE y RMAPE se desprende que, a nivel individual, es la tercera variación del modelo la que posee los mejores resultados.

Modelo	Proyección 2013	Real 2013	Error porcentual agregado	MAE	MPE	MAPE	WMAPE
L-P-M 1	130,19	111	17,29%	0,179	-81,81%	82,37%	83,99%
L-P-M 2	124,02	111	11,73%	0,173	-81,34%	81,89%	83,57%
L-P-M 3	151,92	111	36,78%	0,183	-73,04%	75,58%	77,33%

Tabla 33: Resultados y errores L-P-M proyección año 2013

Finalmente, la “Tabla 34” muestra que considerando la proyección acumulada para los años 2012 y 2013 es la primera variación del modelo la que entrega el mejor resultado a nivel agregado, mientras que la tercera variación sigue siendo la que entrega los mejores resultados a nivel individual.

Modelo	Proyección acum 2012-2013	Real acum 2012-2013	Error porcentual agregado	MAE	MPE	MAPE	WMAPE
L-P-M 1	238,34	244	2,32%	0,316	-69,49%	75,41%	77,53%
L-P-M 2	228,35	244	6,42%	0,306	-68,94%	75,17%	77,16%
L-P-M 3	286,25	244	17,34%	0,314	-55,63%	70,92%	70,29%

Tabla 34: Resultados y errores L-P-M proyección acumulada años 2012-2013

En conclusión, se puede apreciar que, al igual que para el modelo BG/NBD, al aumentar la cantidad de variables utilizadas por el modelo para realizar la predicción, aumenta la precisión a nivel individual, por lo que la variación “L-P-M 3” ofrece los mejores resultados para la estimación del número de interacciones según este modelo.

#### IV.4.2.4.-Selección del modelo para la estimación del número de transacciones

Como se aprecia en las 2 secciones anteriores, para los modelos propuestos y sus variaciones se obtienen mejores resultados al agregar una mayor cantidad de variables de perfilamiento y de comportamiento como variables explicativas del número de transacciones a realizar por los clientes.

	Modelo	Proyección	Real	Error porcentual agregado	MAE	MPE	MAPE	WMAPE
2012	BG/NBD 3	88,07	133	33,77%	0,156	-81,57%	84,35%	85,10%
	LPM 3	134,42	133	1,07%	0,171	-68,27%	75,51%	75,99%
2013	BG/NBD 3	78,25	111	29,51%	0,142	-87,77%	87,77%	88,22%
	LPM3	151,92	111	36,78%	0,183	-73,04%	75,58%	77,33%
Acumulado	BG/NBD 3	166,32	244	31,84%	0,273	-76,16%	81,04%	81,83%
	LPM 3	286,25	244	17,34%	0,314	-55,63%	70,92%	70,29%

Tabla 35: Comparación BG/NBD 3 y LPM 3

Sin embargo, al observar la “Tabla 35” que compara los 2 mejores variaciones de cada modelo, se tiene que es el modelo “L-P-M 3” el que entrega los mejores resultados tanto a nivel individual como a nivel agregado. Además, sus resultados tanto a nivel anual como a nivel acumulado son mejores, a excepción del año 2013, en dónde es el modelo “BG/NBD 3” el que proporciona el menor error a nivel agregado.

Además de observar los resultados a nivel individual y a nivel agregado, resulta interesante observar cómo se comportan estos modelos en predicciones a nivel de perfil de los clientes. Esto debido a que los errores en las predicciones a nivel individual siguen siendo muy altos, por lo que es importante conocer cuál de los 2 modelos trabaja mejor a un nivel de perfil de clientes, pensando en la posible implementación de acciones a este nivel. En base a lo anterior, la “Tabla 36” muestra que en general el modelo “L-P-M 3” entrega los menores errores a nivel agregado dentro de la combinación “Naturaleza”-“Rubro” considerando las predicciones acumuladas entre los años 2012 y 2013. Además considerando los errores a nivel individual, estos son relativamente más bajos que en el modelo “BG/NBD 3”.

Naturaleza	Rubro	Real Acumulado	L-P-M 3			BG/NBD 3		
			Prediccion	Error Absoluto	MAPE	Prediccion	Error Absoluto	MAPE
Empresa	R1	1	0,67	32,82%	98,80%	3,27	227,22%	83,34%
	R2	52	61,43	18,13%	55,85%	22,12	57,45%	77,39%
	R3	18	15,16	15,76%	50,20%	9,63	46,49%	68,57%
	R4	22	18,88	14,19%	81,12%	11,73	46,68%	85,92%
	R5	18	25,33	40,71%	96,71%	18,70	3,92%	86,69%
	R6	18	19,68	9,32%	80,32%	11,85	34,15%	90,22%
	R7	14	14,58	4,15%	53,88%	20,54	46,72%	42,12%
	R8	67	89,46	33,53%	63,94%	35,65	46,79%	80,64%
Persona	R1	3	2,30	23,26%	97,78%	5,64	87,86%	95,52%
	R2	5	4,73	5,49%	95,13%	4,01	19,78%	94,56%
	R3	3	2,05	31,68%	58,26%	4,45	48,45%	68,16%
	R4	3	2,08	30,75%	92,40%	3,88	29,41%	91,26%
	R5	1	2,30	130,20%	50,82%	0,57	42,61%	85,67%
	R6	12	18,42	53,53%	88,20%	8,84	26,30%	94,60%
	R7	1	2,17	116,65%	92,14%	1,92	91,74%	97,92%
	R8	6	7,01	16,85%	92,81%	3,50	41,73%	95,46%

Tabla 36: Comparación Predicción Acumulada por perfiles BG/NBD 3 y L-P-M 3

En base a lo anterior, se decide utilizar el modelo “L-P-M 3” para proyectar el número de transacciones a realizar por los clientes de cara al futuro.

#### **IV.4.3.-Estimación de los montos por transacción**

##### **IV.4.3.1.-Modelo Gamma-Gamma**

Para la aplicación de este modelo se utilizan las siguientes variables explicativas en la obtención de los parámetros:

- $x_{2011}^*$  = Número de transacciones realizadas hasta finales del año 2011
- $m_{2011}^*$  = Monto promedio por transacción hasta finales del año 2011
- Rubro
- Naturaleza
- Tiene vehículo de otro segmento

Las métricas de ajuste para la estimación de esta variable se presentan en la “Tabla 37” donde se aprecia que, a diferencia de lo ocurrido en la estimación del número esperado de transacciones a realizar donde en general el modelo subestimaba el comportamiento de los clientes a nivel individual (Valores de MPE negativos), para la estimación de esta variable se sobreestima un poco su valor. Además, los errores a nivel individual tanto anuales como el acumulado para el periodo 2012-2013 son considerablemente menores a los obtenidos en la estimación del número de transacciones.

Modelo	MPE	MAPE	WMAPE
<b>2012</b>	16,83%	38,99%	51,86%
<b>2013</b>	12,72%	40,20%	47,65%
<b>Acumulado</b>	12,19%	36,00%	49,05%

Tabla 37: Métricas de ajuste para la estimación de los montos por transacción

#### **IV.4.4.-Estimación del valor de los clientes**

Una vez seleccionado el modelo a utilizar para el cálculo del número de transacciones, y teniendo los resultados del modelo que estima los montos por transacción, la multiplicación de ambas variables para cada cliente permite obtener el valor proyectado para la cartera en el periodo de validación.

Observando la “Tabla 38” se desprende que los errores de la predicción a nivel agregado son relativamente pequeños, mientras que los errores a nivel individual para los clientes que realmente compran son similares a los obtenidos para la estimación del número de transacciones.

Modelo	Proyección	Real	Error porcentual agregado	MPE	MAPE	WMAPE
<b>2012</b>	\$ 4.681.427.751	\$ 5.398.538.516	23,28%	-67,69%	78,50%	82,08%
<b>2013</b>	\$ 5.283.675.568	\$ 5.173.431.337	2,13%	-73,31%	76,95%	83,26%
<b>Acumulado</b>	\$ 9.965.103.319	\$ 10.571.969.853	5,74%	-57,00%	71,97%	77,24%

Tabla 38: Métricas de ajuste para la predicción del valor generado durante el periodo de validación

Además, en la tabla se aprecia que a nivel acumulado, es decir, considerando lo ocurrido en ambos años, el error es mucho menor. Lo anterior indica que si se piensa utilizar el modelo para desarrollar acciones enfocadas en el mediano-largo plazo, como se verá más adelante, el modelo parece aumentar la certeza de sus predicciones y parece ser más confiable.

Por otro lado, resulta interesante observar cómo se comportan las predicciones a nivel de perfil considerando las combinaciones “Rubro” y “Naturaleza”. En la “Tabla 39” se aprecia que, considerando estas variables de perfilamiento, los errores a nivel agregado son relativamente pequeños a excepción de las combinaciones “Persona-R7” y “Persona-R5”, donde el modelo tiene a sobreestimar en demasía el valor de los clientes de estos perfiles. Por otro lado, los errores a nivel individual siguen siendo altos, aunque hay perfiles a los que se les predice con muy poco error a este nivel como lo es el caso de los clientes Tipo “Persona” y Rubro “R5”.

Naturaleza	Rubro	# Clientes	Monto promedio modelo	Monto Promedio	Error Absoluto	MAPE
Empresa	R1	38	\$ 655.115	\$ 1.186.749	45%	99,0%
	R2	83	\$ 30.913.062	\$ 42.483.783	27%	61,2%
	R3	33	\$ 16.233.984	\$ 19.891.995	18%	68,4%
	R4	106	\$ 5.065.418	\$ 8.372.003	39%	82,8%
	R5	36	\$ 24.605.480	\$ 21.525.263	14%	85,6%
	R6	121	\$ 6.500.800	\$ 8.179.085	21%	79,8%
	R7	43	\$ 9.788.462	\$ 13.690.955	29%	65,3%
	R8	194	\$ 15.558.908	\$ 11.412.987	36%	62,6%
Persona	R1	81	\$ 862.279	\$ 1.161.356	26%	97,8%
	R2	75	\$ 1.846.594	\$ 2.112.821	13%	95,5%
	R3	16	\$ 2.920.942	\$ 4.936.937	41%	51,7%
	R4	38	\$ 1.210.821	\$ 1.567.196	23%	91,4%
	R5	5	\$ 21.780.472	\$ 4.411.418	394%	5,5%
	R6	160	\$ 3.160.341	\$ 1.806.353	75%	86,4%
	R7	160	\$ 4.856.273	\$ 1.500.092	224%	88,3%
	R8	104	\$ 1.938.043	\$ 1.584.531	22%	92,0%

Tabla 39: Predicción acumulada en el periodo de validación según perfil “Naturaleza”-“Rubro”

#### **IV.4.4.1.-Validación de la capacidad de detección de valor del modelo**

Debido a que son pocos los clientes que compran en el periodo de validación, y a que las métricas de ajuste presentadas (MPE, MAPE, WMAPE) se enfocan en lo que ocurre en ellos, dejando de lado las predicciones realizadas por el modelo para los clientes que no compran, es posible que para este grupo de clientes (los que no compran) el modelo proyecte mayores valores que para los que realmente compran.

Lo anterior podría afectar de forma importante la posterior utilidad del modelo a un nivel más desagregado, por lo que resulta necesario intentar validar su capacidad de discriminación en comparación a alguna alternativa simple que actualmente podría utilizar la empresa para identificar a sus clientes más valiosos.

El método de comparación realizado consistió en seleccionar al 20% de los clientes que han resultado más valiosos históricamente hasta finales del año 2011 y observar que porcentaje generan del valor real en los años de validación del modelo. Esto se contrasta con el 20% más valioso predicho para cada año utilizando el modelo. El hecho de contrastar el modelo desarrollado con una agrupación simple en base al valor histórico, permite obtener una noción de si realmente se gana conocimiento de los clientes con el modelo propuesto en comparación a algún método sencillo que podría utilizar la compañía para valorizar a sus clientes.

Los resultados de la comparación señalada anteriormente se muestran en la “Tabla 40” y en la “Tabla 41”

	20% Mayor según el modelo	20% Mayor según historia
<b>Monto real capturado</b>	\$ 4.091.752.752	\$ 3.228.495.206
<b>% capturado del real total capturado</b>	75,79%	59,80%

Tabla 40: Comparación Predicción Valor 2012

	20% Mayor según el modelo	20% Mayor según historia
<b>Monto real capturado</b>	\$ \$ 3.575.019.782	\$ 3.116.485.224
<b>% capturado del real total capturado</b>	69,10%	60,24%

Tabla 41: Comparación Predicción Valor 2013

En las tablas anteriores se aprecia que el quintil más valioso predicho por el modelo de cara al 2012 genera más del 70%, mientras que el seleccionado en base al valor histórico generó solo un 60%. Una situación similar se da para el año 2013, sin embargo, la ganancia que se genera con el modelo resulta ser un poco menor.

Es importante mencionar que, debido a que el modelo propuesto trabaja en base a esperanzas y no con un método de clasificación (compra o no compra), más que la exactitud a nivel individual, se busca que el modelo tenga la capacidad de distinguir entre los que son más o menos valiosos. De esta forma, aunque los valores predichos no sean exactos lo que importa es poder detectar a los que poseen una esperanza mayor y que estos realmente sean los más valiosos para la compañía. Por esta razón, la “validación” que se acaba de realizar indica que vale la pena utilizar el modelo con el fin de comenzar a enfocar acciones o esfuerzos diferenciados a la cartera de clientes.

#### ***IV.4.5.-Aplicación del modelo a toda la cartera de clientes***

Con los parámetros obtenidos de la calibración y validación realizadas previamente, se procede a aplicar el modelo a toda la data disponible (Clientes que han comprado por primera vez hasta el 31-12-2013, y se proyecta su valor en un plazo de 5 años, es decir, desde el 2014 hasta el año 2018. Este plazo escogido se debe al hecho de que anualmente la precisión del modelo va disminuyendo, y a pesar de que esta disminución no fue muy grande para el plazo de validación de 2 años, el comportamiento de la cartera podría cambiar drásticamente dentro de los próximos años, por lo que intentar un plazo mayor de predicción resultaría muy arriesgado. Además, este plazo es suficiente para, como se plantea más posteriormente, utilizar los resultados del modelo en la recomendación de acciones de Marketing. Es importante mencionar que los montos futuros proyectados para los clientes se descuentan a una tasa del 18%<sup>27</sup> anual.

La “Ilustración 32” presenta el Box-plot de las predicciones de CLV hasta el año 2018, donde se aprecia que en general los ingresos predichos son pequeños están en el orden de los \$10.000.000 y \$30.000.000, lo que se traduce en la compra de, a lo más, un vehículo. Sin embargo existen clientes que parecen representar un valor muy grande en comparación al resto.

<sup>27</sup> Tasa de referencia dada por la empresa

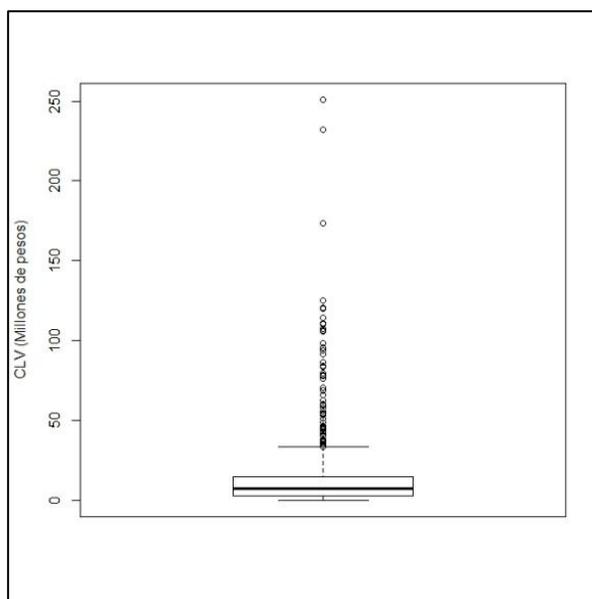


Ilustración 32: Boxplot del CLV

En la “Tabla 42” se muestra las predicciones acumuladas realizadas por el modelo hasta el año 2018 a nivel de “Naturaleza” y “Rubro”, donde se aprecia que en promedio, los clientes más rentables serán los de tipo “Empresa” que se dedican al rubro “R2”, mientras que los menos rentables corresponden a los de tipo “Persona” y pertenecientes al rubro “R1”. Las predicciones realizadas por el modelo a los perfiles formados además por la variable “Tiene vehículo de otro segmento” se encuentran en el “Anexo H”.

Naturaleza	Rubro	# Clientes	Predicción a 2018	Promedio	Desviación Estándar	Valor Máximo	Valor Mínimo
Empresa	R1	56	\$ 131.953.217	\$ 2.356.307	\$ 1.757.934	\$ 8.214.253	\$ 383.011
	R2	137	\$ 13.524.269.043	\$ 98.717.292	\$ 215.345.451	\$ 2.281.823.273	\$ 7.142.588
	R3	53	\$ 1.855.748.621	\$ 35.014.125	\$ 62.718.101	\$ 287.754.259	\$ 4.217.799
	R4	202	\$ 2.817.608.162	\$ 13.948.555	\$ 10.483.364	\$ 72.075.967	\$ 2.016.988
	R5	61	\$ 1.382.443.150	\$ 22.663.002	\$ 16.796.804	\$ 94.432.887	\$ 5.052.974
	R6	187	\$ 2.369.969.548	\$ 12.673.634	\$ 11.771.277	\$ 123.773.775	\$ 2.013.686
	R7	74	\$ 899.984.558	\$ 12.161.953	\$ 14.598.741	\$ 75.021.139	\$ 2.942.657
	R8	308	\$ 6.872.995.217	\$ 22.314.920	\$ 18.641.789	\$ 148.346.665	\$ 2.743.087
Persona	R1	129	\$ 208.907.466	\$ 1.619.438	\$ 565.933	\$ 4.544.441	\$ 90.767
	R2	122	\$ 542.497.034	\$ 4.446.697	\$ 2.348.889	\$ 14.410.554	\$ 1.364.045
	R3	24	\$ 254.052.610	\$ 10.585.525	\$ 3.706.356	\$ 17.222.218	\$ 3.564.286
	R4	69	\$ 187.483.881	\$ 2.717.158	\$ 1.310.377	\$ 8.879.409	\$ 122.393
	R5	17	\$ 477.202.226	\$ 28.070.719	\$ 12.558.885	\$ 41.214.426	\$ 1.600.391
	R6	251	\$ 784.489.837	\$ 3.125.458	\$ 2.312.988	\$ 18.925.926	\$ 516.876
	R7	23	\$ 71.310.309	\$ 3.100.448	\$ 1.063.392	\$ 6.594.581	\$ 1.725.694
	R8	159	\$ 281.396.422	\$ 1.769.789	\$ 669.787	\$ 5.548.079	\$ 275.497
<b>Total general</b>		<b>1872</b>	<b>\$ 32.662.311.302</b>	<b>\$ 17.447.816</b>	<b>\$ 64.656.511</b>	<b>\$ 2.281.823.273</b>	<b>\$ 90.767</b>

Tabla 42: Predicción CLV acumulado a 2018 a nivel de Perfil “Naturaleza”-“Rubro”

## **IV.5.-Propuesta de uso del modelo.**

A continuación se presenta una propuesta de utilización del modelo para el enfoque de acciones de Marketing a la cartera de clientes que posee actualmente (hasta finales de Diciembre del 2013) la empresa. Es importante mencionar que históricamente la empresa nunca ha desarrollado un plan de trabajo enfocado en su cartera de clientes, debido a que todos los esfuerzos se han enfocado en el Marketing Masivo el cual genera un impacto difícil de medir en el comportamiento de la cartera de clientes y además posee un impacto mayor en la adquisición de cuentas más que en el cuidado de las mismas.

### ***IV.5.1.-Recomendaciones de acciones al área de Marketing***

A grandes rasgos, y según la opinión de los expertos en el área de Marketing dentro de la compañía, se pueden distinguir 2 posibles grupos de acciones a enfocar en la cartera de clientes, de acuerdo al plazo de tiempo en el que buscan generar resultados:

#### ***IV.5.1.1.-Acciones a corto plazo***

Tienen como objetivo generar una respuesta del cliente en el periodo actual, donde periodo se define como el “año actual”. Dentro de este tipo de este grupo de acciones se pueden distinguir:

- **Promoción**: Si se detecta que un cliente está propenso a realizar una interacción en el periodo actual, se debe intentar asegurar que la transacción la realice con la empresa. Este tipo de acciones no involucra necesariamente algún descuento, sino que el objetivo es generar contacto con el cliente para ser su primera opción.
- **Recuperación de clientes**: La competencia genera que muchas veces los clientes abandonen la empresa y decidan comenzar a comprar en otra marca. Si se detecta lo anterior y el cliente en cuestión posee un considerable valor histórico, valdría la pena realizar alguna acción para recuperarlo. Una herramienta útil para detectar lo anterior es el valor proyectado por el cliente a futuro, ya que si este comienza a disminuir se puede atribuir a una “posible fuga”. Esta es una decisión que se toma en el corto plazo, pensando en los resultados que puede traer una posible recuperación del cliente en el mediano-largo plazo. Dentro de este grupo se encuentran acciones como:
  - Contactar a los clientes para conocer las razones de su posible fuga.
  - En base a estas razones, analizar si vale la pena realizar alguna oferta para recuperar al cliente.

#### ***IV.5.1.2.-Acciones a mediano-largo plazo***

Las acciones de este tipo no buscan generar una respuesta o acción inmediata por parte del cliente, sino que buscan asegurar o mejorar el posible comportamiento de algún grupo en particular. Dentro de estas acciones se encuentran:

- Fidelización: A los clientes que han generado importantes ingresos históricamente a la empresa, y que además tienen un gran potencial se seguir realizando en el futuro, se les debe enfocar acciones de este tipo, las cuales tiene como objetivo que el cliente se sienta identificado con la marca para así ser su primera opción en una futura decisión de compra. Dentro de este grupo se encuentran acciones como:
  - Invitación a eventos realizados por la marca.
  - Invitación a presentación de nuevos modelos de vehículos.
  - Premios y regalos relacionados a la marca.
  - Comunicación constante para conocer su satisfacción con el producto.
  - Descuentos en el servicio técnico de los vehículos.
  - Descuentos en vehículos de otros segmentos.
  - Alta prioridad a la hora se resolver problemas.
  
- Desarrollo o crecimiento: Estos esfuerzos se enfocan en tratar de elevar el valor potencial del cliente, a través del aumento de los ingresos generados por transacción o de la frecuencia en que realiza las compras. Posibles acciones a tomar corresponden a:
  - Incentivar la renovación de los vehículos.
  - Incentivar o facilitar el aumento de flota de los clientes, haciendo descuentos, por ejemplo, en base a la cantidad de vehículos que el cliente compre por transacción.
  - Descuentos en servicio técnico de los vehículos.
  
- Sin esfuerzos adicionales: Más que una acción o esfuerzo, esta es una decisión tomada pensando en la relación al largo plazo entre el cliente y la empresa. Todas las compañías poseen clientes que les representan poco valor (y a veces incluso pérdidas), ya sea porque compran poco frecuentemente, sus montos por transacción son muy bajos, etc. A este grupo de clientes, muchas veces no conviene destinar recursos debido a la alta probabilidad de que la inversión no sea recuperada en el tiempo.

#### ***IV.5.2.-Agrupación de Clientes***

Utilizando el valor histórico de los clientes hasta el año 2013 y los valores obtenidos por el modelo para las proyecciones año a año de cara al 2018, se plantea la siguiente agrupación de clientes:

Al igual como se realiza en [25], se utiliza la regla 20, 80, 30 para realizar la agrupación de clientes. Esta regla indica que el 20% de los mejores clientes genera el 80% de los beneficios para la empresa, mientras que el 30% peor genera pérdidas. A diferencia de lo realizado en el citado trabajo donde solo utilizan el valor futuro proyectado, aquí se propone la utilización de 3 variables: el valor histórico, el valor de cara al año actual (2015), y el valor proyectado al 2018.

De acuerdo a lo anterior, para cada una de estas 3 variables se crean 3 grupos, los que corresponden a:

- 1 = 20% de los clientes con mayor valor.
- 2 = 50% de los clientes que siguen en cuanto a valor.
- 3 = 30% final de los clientes.

Debido a que se tiene data transaccional hasta el año 2013, el comportamiento de la cartera para el año 2014 se predice utilizando el modelo, y se suma esta predicción al valor histórico al año 2013. Con lo anterior se obtienen los grupos y los promedios de valor por cliente que se aprecian en la “Tabla 43”.

Grupo	Promedio de Valor hasta 2014
1	\$ 141.275.884
2	\$ 31.662.785
3	\$ 19.900.583
<b>Total general</b>	<b>\$ 50.030.809</b>

Tabla 43: Valor Promedio para grupos generados con el Valor Histórico hasta el 2014

Para el valor proyectado al año 2015 se obtienen los resultados mostrados en la “Tabla 44”.

Grupo	Promedio de Valor proyectado en 2015
1	\$ 13.590.454
2	\$ 1.808.552
3	\$ 369.973
<b>Total general</b>	<b>\$ 3.730.534</b>

Tabla 44: Valor Promedio para grupos generados con el Valor Proyectado para el año 2015

Mientras que para el valor proyectado acumulado entre los años 2015 y 2018 para la cartera de clientes se obtienen los grupos mostrados en la “Tabla 45”:

Grupo	Promedio de Valor proyectado entre 2015-2018
1	\$ 48.506.756
2	\$ 7.401.660
3	\$ 1.465.247
<b>Total general</b>	<b>\$ 13.831.704</b>

Tabla 45: Valor Promedio para grupos generados con el Valor acumulado proyectado entre el 2015 y 2018

### IV.5.3.-Recomendación de acciones a los distintos grupos de clientes

Considerando las variables “Valor Histórico” y “Valor proyectado entre 2015 y 2018” y combinándolos, se obtiene la siguiente matriz de movimiento de clientes entre los grupos generados:

Valor Histórico	Valor Potencial 2018		
	1	2	3
1	207	121	46
2	148	513	275
3	19	302	241

Tabla 46: Movimientos de clientes entre grupos Considerando el Valor histórico y el Proyectado acumulado al 2018

De las acciones de Marketing descritas anteriormente, y considerando las combinaciones (Valor Histórico, Valor Potencial a 2018) se propone la siguiente aplicación de esfuerzos, ordenados de mayor a menor de acuerdo al nivel de esfuerzo económico que requieren, o que la empresa debería estar dispuesta a realizar:

- Fidelización: Se recomienda aplicar este tipo de acciones a los clientes con mayor valor de cara al futuro. Estos clientes serían los grupos formados por las combinaciones (1,1), (2,1) y (3,1).
- Desarrollo o crecimiento: A los clientes con mediano valor potencial, pero cuyo valor histórico parece indicar que podrían ser rentables de cara al futuro. Estos clientes son los pertenecientes a los grupos (1,2) y (2,2).
- Recuperación: Se recomienda aplicar este tipo de esfuerzos a los clientes que hayan representado históricamente un valor considerable, pero cuyo valor potencial es bajo. Estos son los clientes pertenecientes a los grupos (1,3).
- Sin esfuerzos adicionales: Clientes que no hayan representado un valor importante históricamente y que tampoco parece que lo vayan a generar de cara al futuro. Estos clientes son los pertenecientes a los grupos (2,3), (3,2) y (3,3).

En base a lo anterior se obtiene la siguiente matriz.

Valor Histórico	Valor Potencial 2018			
	1	2	3	
1	207	121	46	Fidelizar
2	148	513	275	Desarrollar
3	19	302	241	Recuperar
				Sin esfuerzos adicionales

Tabla 47: Recomendación de acciones por grupos según Valor histórico y valor proyectado al 2018.

Ahora, al agregar la variable “Valor en 2015” y agregando a las recomendaciones de acciones ya definidas para cada uno de los grupos, se obtiene la siguiente matriz:

		Grupo proyección año 2018								
		1			2			3		
		Grupo Proyección a 2015			Grupo Proyección a 2015			Grupo Proyección a 2015		
Grupo Histór	1	2	3	1	2	3	1	2	3	
1	201	6	0	6	113	2	0	4	42	
2	133	15	0	20	463	30	0	38	237	
3	14	5	0	0	282	20	0	10	231	

Tabla 48: Movimiento de clientes entre grupos Considerando las 3 variables de valor

Así, considerando las combinaciones (Valor Histórico, Valor Potencial a 2018, Valor Projectado en 2015), se propone realizar acciones de promoción en el periodo actual a todas las combinaciones (-,-,1), debido a que se debe asegurar que el grupo de clientes con mayor valor para el año actual tenga presente el nombre de la marca en el caso de que efectivamente esté pensando en realizar alguna transacción en el periodo en cuestión.

De esta forma, se generan las combinaciones de acciones indicadas en la “Ilustración 33”.

		Grupo proyección año 2018								
		1			2			3		
		Grupo Proyección a 2015			Grupo Proyección a 2015			Grupo Proyección a 2015		
Grupo Histór	1	2	3	1	2	3	1	2	3	
1	201	6	0	6	113	2	0	4	42	
2	133	15	0	20	463	30	0	38	237	
3	14	5	0	0	282	20	0	10	231	

Fidelizar y promocion	
Fidelizar	
Desarrollar y promocion	
Desarrollo	
Recuperación y Promoción	
Recuperación	
Sin esfuerzos adicionales y promoción	
Sin esfuerzos adicionales	

Ilustración 33: Recomendación de acciones por grupos generados en base a las 3 variables de valor

#### IV.5.4.-Evaluación del impacto de la propuesta

Debido a que la empresa históricamente no ha desarrollado ninguna metodología para trabajar con la cartera de clientes, el cambio de enfoque que se plantea es imposible de cuantificar debido a que no se tienen experiencias previas que midan el impacto que generan este tipo de acciones en los distintos grupos de clientes. Lo anterior impide optimizar las decisiones de inversión en cada grupo en esta etapa inicial, por lo que en conjunto con la empresa se plantea la siguiente puesta en marcha:

### Cantidad a invertir de cara al 2015:

De cara al próximo año, la empresa planea comenzar a invertir parte del presupuesto de su área de Marketing en acciones enfocadas en su cartera de clientes. El presupuesto anual del que dispone es de aproximadamente \$300.000.000, y se desea comenzar destinando un pequeño porcentaje a modo de ir testeando el impacto que genera la inversión en la cartera de clientes. Este porcentaje aún no está definido por la gerencia, pero se estima que la cota máxima será un 10% lo que equivale a \$30.000.000.

### Monto Máximo a invertir por grupo de cliente:

Utilizando el valor futuro proyectado para cada cliente y aplicando un margen de 20% a los ingresos proyectados<sup>28</sup>, se puede estimar la cantidad de dinero máximo a invertir anualmente en las distintas recomendaciones de acciones para los distintos grupos generados.

### Grupos en los cuáles comenzar a testear:

Los deseos de la empresa son destinar parte de este presupuesto a las siguientes combinaciones (Valor Histórico, Valor Potencial a 2018, Valor Proyectado a 2015):

- (1,1,1), Se desea analizar el impacto que generan posibles acciones en este grupo.
- (2,2,2), Se desea analizar la dificultad que significa aumentar el valor de los clientes de este grupo, que representan el mayor en cuanto a cantidad de clientes.
- (1,3,3), Se desean analizar la capacidad de recuperación de clientes que en algún momento fueron valiosos.

### Impacto mínimo a generar:

El margen del 20% mencionado anteriormente implica que por cada peso invertido en el cliente, este debe generar un mínimo de 5 pesos para que el esfuerzo haya valido la pena. De acuerdo a lo mencionado, anteriormente, esta inversión máxima de \$30.000.000 debe generar un aumento de \$150.000.000 en el valor que se predice a la cartera de clientes.

### Medición del impacto:

Se deben registrar todas las inversiones realizadas en los clientes seleccionados y las decisiones que este va tomando en relación a los esfuerzos realizados. Posteriormente, por cada grupo se debe comparar el retorno promedio generado por los clientes en los cuáles se invirtió con el retorno promedio generado por los clientes del mismo grupo en

---

<sup>28</sup> Recordar que el modelo predice ingresos y no márgenes. Este margen de 20% fue dado como una referencia del promedio que marginan los vehículos del segmento.

los cuáles no se invirtió. De esta forma es posible estimar el impacto que van generando determinados esfuerzos en los distintos grupos de clientes.

Es importante mencionar, que las acciones específicas a realizar como parte de las recomendaciones definidas a cada grupo son tareas que definirá el área de Marketing de la marca, siendo el área de Gestión de clientes, desde la cual nace el trabajo, la encargada de cuantificar el impacto real que estas generan.

## **V.-Conclusiones Finales**

### **V.1.-Conclusión del trabajo**

El modelo de valoración y su utilización propuesta significan el punto de partida para la empresa en la búsqueda de un enfoque relacional, que busca comprender, mejorar y aumentar las relaciones que posee actualmente con su cartera de clientes. Este cambio de enfoque surge de la necesidad de elevar las bajas tasas de retención que posee la compañía. Para esto se desarrolló un modelo que busca identificar los clientes que parecen ser más valiosos de cara al futuro.

El problema de las bajas tasas que de retención dificultó el desarrollo del modelo debido a que, dada la naturaleza de la industria, se necesitan varios años de data disponible para poder identificar los patrones que afectan la recompra por parte de los clientes. Además, el mal manejo de la información presentado históricamente por la empresa, generó un arduo trabajo de recopilación de información y limpieza de datos con el que se logró componer la historia desde el año 1999 hasta el año 2013, con una completitud cercana al 90% de la realidad según los expertos dentro de la empresa.

Respecto al modelo propuesto, a pesar de que los errores que entrega son elevados a nivel individual, cuando se trabaja con grupos de clientes las predicciones no se alejan tanto de la realidad, lo que permite a la compañía tener una noción del valor actual que representa su cartera de clientes como un “todo”, y a nivel de distintas combinaciones de perfil.

Por otro lado, como el modelo propuesto trabaja con la esperanza del ingreso generado por los clientes y, como ya se mencionó, las tasas de recompra históricamente han sido bajas, era esperable que las predicciones a nivel individual estuvieran alejadas de la realidad. Sin embargo, se testeó la capacidad de identificación de clientes valiosos por parte del modelo, y los resultados indican que aunque las agrupaciones generadas por el modelo no predicen el valor exacto de estos (al trabajar con esperanzas), si se logra identificar cuáles son los más valiosos en relación al resto.

Dentro de los patrones de compra detectados, resulta llamativo el hecho de que los clientes que han comprado en más de un periodo previo poseen una probabilidad de compra mucho mayor que los que han comprado solo una vez. Por otro lado, los clientes tipo empresa son más propensos a comprar que los tipo persona. Tener vehículos de otro segmento afecta positivamente la probabilidad de compra. Respecto a los rubros, los clientes del Rubro “R5” parecen ser los más propensos a interactuar nuevamente. Además, tal como se intuía, el PIB tiene un efecto positivo en la probabilidad de compra.

Además, la metodología propuesta utilizando cadenas de Markov junto a regresiones logísticas y de Poisson resulta interesante debido a la facilidad que tiene para modelar la evolución del comportamiento futuro. Dentro del modelo, las variables macroeconómicas como el PIB se encargan de dar cierta heterogeneidad en el tiempo a

las probabilidades de transición entre estados, lo que mejora un posible supuesto de homogeneidad si se hubiesen ocupado matrices de transición constantes en el tiempo.

Respecto a la utilización del modelo, se logra agrupar a los clientes de forma sencilla en base a su valor histórico y el valor futuro proyectado, permitiendo de esta forma generar recomendaciones de distintos tipos de acciones de marketing para los distintos grupos de clientes obtenidos. Lo anterior permite a las áreas comerciales tener una mejor noción de “qué hacer” con su cartera de clientes para así enfocar de mejor forma sus esfuerzos.

## **V.2.-Trabajos Futuros**

De cara al futuro se propone analizar el comportamiento y valor de los clientes en la línea de negocio de servicio y en la venta de vehículos livianos y pesados, e intentar relacionar el comportamiento de los clientes en todas líneas de negocio de la empresa, para de esta forma tener una visión global del valor del cliente.

Como se mencionó previamente, la medición del impacto que genera la recomendación de acciones realizada utilizando el modelo es una tarea fundamental de cara al futuro. Una vez que se logren cuantificar el impacto de las distintas acciones se puede pensar en el desarrollo de modelos de optimización de la inversión a realizar en la cartera de clientes

Determinar otras variables macroeconómicas que afecten el comportamiento de los clientes y para las cuales existan proyecciones futuras de sus valores. Estas variables son importantes ya que agregan heterogeneidad en el tiempo a la matriz de transición de probabilidades, lo que debería aumentar la capacidad predictiva del modelo.

Determinar variables del entorno competitivo que podrían ser relevantes en la explicación del comportamiento de los clientes.

Evaluar la aplicación de otro tipo de metodologías y modelos que permitan llegar a mejores resultados respecto a la estimación del CLV a nivel individual.

## VI.-Bibliografía

- [1] Aurora Garrido-Moreno, Antonio Padilla-Meléndez, Ana Rosa Del Águila-Obra, «Exploring the Importance of Knowledge Management for CRM Success,» *World Academy of Science, Engineering and Technology*, vol. 4, pp. 79-80, 2010.
- [2] Khalid Rababah, Haslina Mohd, and Huda Ibrahim, «Customer Relationship Management (CRM) Processes from Theory to Practice: The Pre-implementation Plan of CRM System,» *International Journal of e-Education, e-Business, e-Management and e-Learning*, vol. 1, pp. 22-23, 2011.
- [3] C. C. H. Chan, «Intelligent value-based customer segmentation method for campaign management: A case study of automobile retailer,» *Expert Systems with Applications*, p. 2758, 2008.
- [4] SAP White paper, ANALYTICAL CRM, 2001.
- [5] ANAC, «Mercado automotor Diciembre 2013,» Santiago, 2013.
- [6] SAP Solutions, «Customer Relationship Management in the automotive industry,» 2005.
- [7] C.-J Chenga, S.W. Chiub, C.-B. Chengc, J.-Y. Wuc, «Customer lifetime value prediction by a Markov chain based data mining model: Application to an auto repair and maintenance company in Taiwan,» *Scientia Iranica*, 2011.
- [8] A. Hughes, «dbmarketing,» Marzo 2013. [En línea]. Available: <http://www.dbmarketing.com/2010/03/building-caterpillar-market-share-with-a-database/>. [Último acceso: 12 Octubre 2014].
- [9] F. B. Rizal Ahmad, «Customer retention: a potentially potent marketing management strategy,» *Journal of strategic Marketing*, p. 9, 2001.
- [10] René A. Caldentey, Susana V. Mondschein, «Cadenas de Markov,» de *Modelos de Decisión en Ambientes Inciertos*, 1999, pp. 55-75.
- [11] Yang Wang, Michael Ong, Honghu Liu, «Compare Predicted Counts between Groups of Zero Truncated Poisson Regression based on Recycled Predictions Method,» *Statistics in Epidemiology*, pp. 2478-2487, 2011.
- [12] V. Kumar, «Customer Lifetime Value».
- [13] Reinartz, Werner J., and V. Kumar, «The impact of customer relationship characteristics on profitable lifetime duration,» *Journal of Marketing*, pp. 77-99, 2003.
- [14] Gupta y Lehmann, «Managing Customers as Investments,» 2005.
- [15] Ha Nui Baek, «Estimación de customer lifetime value a nivel de clientes usando variables socio-demográficas y transaccionales; Memoria para optar al título de ingeniero civil industrial,» Santiago, Chile, 2006.
- [16] P. A. M. Vicuña, «Estimación de lifetime value basada en comportamiento transaccional; Memoria para optar al título de Ingeniero Civil Industrial,» Santiago, Chile, 2005.
- [17] D. Gropellier, «Modelo para estimar el valor de vida de los clientes de un banco; Memoria para optar al título de Ingeniero Civil Industrial,» 2002.
- [18] Ignacio Domingo Trujillo Silva, «Valor esperado del cliente en el mercado de seguros patrimoniales; Memoria para optar al título de Ingeniero Civil Industrial,» 2013.

- [19] M. C. G. Rosas, «Estimación del valor de clientes de un club de fidelización de una tienda de retail; Memoria para optar al título de Ingeniero Civil Industrial,» 2009.
- [20] F. I. R. Tello, «Estimación del valor de clientes fidelizados en una cadena de Supermercados Utilizando Modelos Jerárquicos Bayesianos; Memoria para optar al título de Ingeniero Civil Industrial,» 2011.
- [21] Peter S. Fader, Bruce G. S. Hardie, Ka Lok Lee, «“Counting Your Customers” the Easy Way:»,» *Marketing Science*, vol. 24, nº 2, pp. 275-284, 2005.
- [22] Robert C. Blattberg, Byung-Do Kim, Scott A. Neslin, «Database Marketing: Analyzing and Managing Customers,» *International Series in Quantitative Marketing*, pp. 114-121, 2008.
- [23] Phillip E. Pfeifer and Robert L. Carraway, «Modeling Customer Relationships,» *Journal of Interactive Marketing*, pp. 43-55, 2000.
- [24] Peter S. Fader, Bruce G. S. Hardie, «The Gamma-Gamma Model of Monetary,» 2013.
- [25] Yanhong Guo, Hui Wang and Wei Liu, «Improved Pareto/NBD Model and Its Applications in Customer Segmentation based on Personal Information Combination,» *International Journal of Database Theory and Application*, vol. 6, nº 5, pp. 175-186, 2013.
- [26] D. Birant, «Data Mining Using RFM Analysis,» *Intech*, pp. 91-92, 2011.
- [27] Bruce A. Craig, Peter P. Sendi, «Estimation of the transition matrix of a discrete-time,» *HEALTH ECONOMICS*, pp. 33-42, 2002.
- [29] Sunil Gupta, Dominique Hanssens, Bruce Hardie, William Kahn, V. Kumar, Nathaniel Lin, Nalini Ravishanker S. Sriram, «Modeling Customer Lifetime Value,» *Journal of Service Research*, vol. 9, pp. 142-143, 2006.

## VII.-Anexos

### Anexo A: Lista de rubros disponibles y su descripción

Etiqueta	RUBROS	DESCRIPCIÓN
1	Agrícola	Frutas y verduras
2	Alimentación general	Productos alimenticios
3	Refrigerados	Productos que requieren refrigeración
4	Áridos	Arena, ripio, gravilla
5	Carga delicada	Vidrios
6	Cargas peligrosas	Explosivos, productos químicos
7	Combustible	Bencina, parafina, petróleo
8	Construcción	Movimiento de tierra, obras civiles, montaje, obras viales
9	Courrier	Paquetería, documentos
10	Eventos	Arriendo de carpas, toldos, muebles, etc
11	Ferretería	
12	Forestal	Vehículo que trabaje o preste servicios al área
13	Ganadería	Transporte de animales
14	Gas	Cilindros y granel
15	Lubricantes	Aceite, riles
16	Minería	Camión para trabajo en minera
17	Muebles	
18	Mudanza	
19	Otro	
20	Rentacar	
21	Retail	Grandes tiendas
22	Servicios especiales	Lubricador, aljibe, compactador basura, capacho
23	Transporte de carga general	Clientes que adaptan su flota al transporte de distintas cargas sin tener un foco principal en algún tipo
24	Transporte de maquinaria	
25	Transporte de pasajeros	
26	Transporte de valores	

Tabla 49: Lista de rubros y su descripción.

**Anexo B: Frecuencia de clientes según la región a la que pertenecen.**

Región	Frecuencia	%
I Región	16	0,85%
II Región	117	6,25%
III Región	69	3,69%
IV Región	68	3,63%
V Región	83	4,43%
Región Metropolitana	1126	60,15%
VI Región	73	3,90%
VII Región	48	2,56%
VIII Región	163	8,71%
IX Región	52	2,78%
X Región	40	2,14%
XI Región	3	0,16%
XII Región	9	0,48%
XIV Región	4	0,21%
XV Región	1	0,05%
<b>Total general</b>	<b>1872</b>	<b>100,00%</b>

Tabla 50: Frecuencia de clientes según su región

**Anexo C: Frecuencia de clientes según su Rubro**

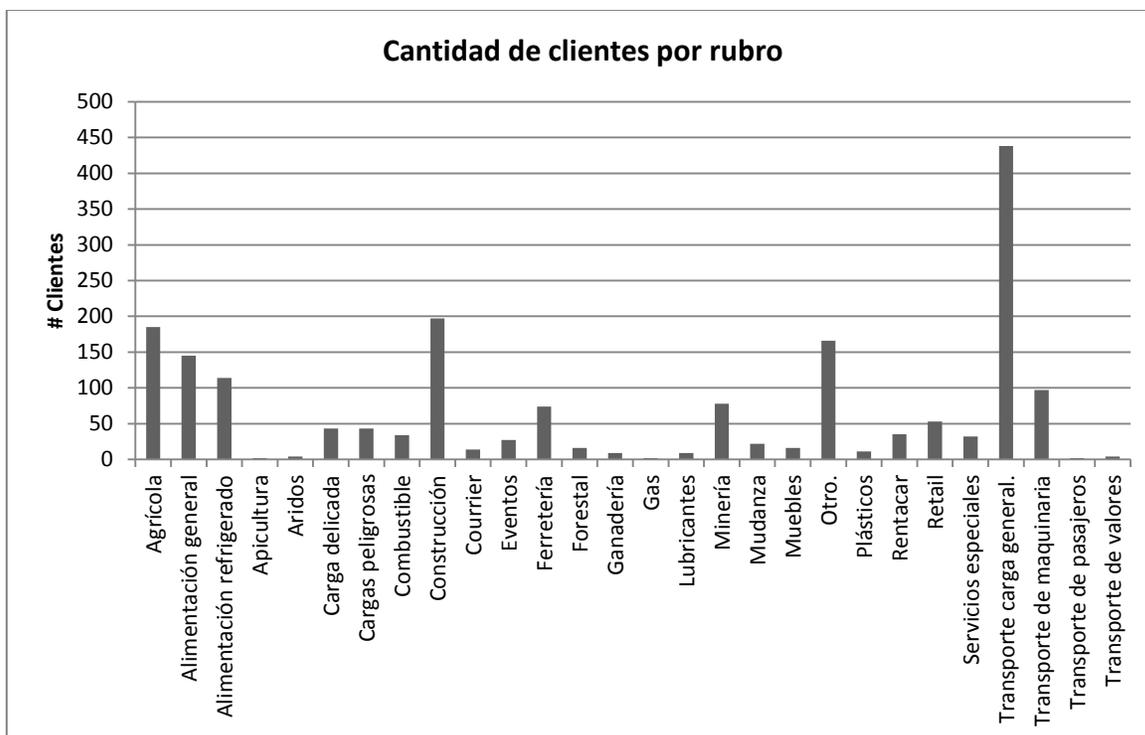


Tabla 51: Frecuencia de clientes según su rubro

**Anexo D: Frecuencia de observaciones del número de transacciones realizadas por los clientes en los años que deciden comprar.**

Número de transacciones	Frecuencia	%
1	345	75,66%
2	75	16,45%
3	21	4,61%
4	5	1,10%
5	4	0,88%
6	5	1,10%
7	1	0,22%
<b>Total general</b>	<b>456</b>	<b>100,00%</b>

Tabla 52: Frecuencia de observaciones del número de transacciones realizadas por los clientes que deciden comprar

**Anexo E: Frecuencia de observaciones de acuerdo a decisión de compra y recencia**

Recencia	Decisión			
	Compra		No compra	
	# Observaciones	%	# Observaciones	%
1	232	12,64%	1604	87,36%
2	90	7,38%	1130	92,62%
3	55	5,88%	881	94,12%
4	21	3,20%	636	96,80%
5	18	3,79%	457	96,21%
6	26	7,37%	327	92,63%
7	3	1,56%	189	98,44%
8	4	3,13%	124	96,88%
9	2	2,20%	89	97,80%
10	4	6,35%	59	93,65%
11		0,00%	51	100,00%
12	1	2,50%	39	97,50%
13		0,00%	32	100,00%
14		0,00%	18	100,00%

Tabla 53 Frecuencia de observaciones según decisión de compra y recencia real

**Anexo F: Resultados Anova para la variable Monto por Transacción**

Variable		DF	Sum sq	Mean sq	F value	pr(>F)
Tipo de cliente	Inter-grupos	1	1,091e+16	1,091e+16	27,1	2,14e-07
	Intra-grupos	1870	7,528e+17	4,026e+14		
	Total	1871				
Rubro	Inter-grupos	7	9,414e+15	1,345e+15	3,324	0,0016
	Intra-grupos	1864	167,53	4,047e+14		
	Total	1871				
Zona	Inter-grupos	2	2,436e+15	1,218e+15	2,991	0,0505
	Intra-grupos	1869	7,612e+17	4,023e+14		
	Total	1871				
Tamaño de Flota	Inter-grupos	3	8,065e+16	2,688e+16	73,52	<2e-16
	Intra-grupos	1868	6,830e+17	3,657e+14		
	Total	1871				
Tiene vehículo de otro tipo	Inter-grupos	1	8,666e+15	8,666e+15	21,46	3,86e-06
	Intra-grupos	1870	7,550e+17	4,038e+14		
	Total	1871				

**Anexo G: Cambios en las matrices de transición entre un año y otro**

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	(1;1)	(1;2)	(1;3)	(2;1)	(2;2)	(2;3)	(3;1)	(3;2)	(3;3)	(4;1)	(4;2)	(4;3)	(5;1)	(5;2)	(5;3)	(6;1)	(6;2)	(6;3)
1 (1;1)	0	0,275	0	0,725	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2 (1;2)	0	0	0,7949	0	0,205	0	0	0	0	0	0	0	0	0	0	0	0	0
3 (1;3)	0	0	0,8228	0	0	0,177	0	0	0	0	0	0	0	0	0	0	0	0
4 (2;1)	0	0,2005	0	0	0	0	0,8	0	0	0	0	0	0	0	0	0	0	0
5 (2;2)	0	0	0,7192	0	0	0	0	0,281	0	0	0	0	0	0	0	0	0	0
6 (2;3)	0	0	0,7543	0	0	0	0	0	0,246	0	0	0	0	0	0	0	0	0
7 (3;1)	0	0,1284	0	0	0	0	0	0	0	0,8716	0	0	0	0	0	0	0	0
8 (3;2)	0	0	0,6007	0	0	0	0	0	0	0	0,3993	0	0	0	0	0	0	0
9 (3;3)	0	0	0,6432	0	0	0	0	0	0	0	0	0,3568	0	0	0	0	0	0
10 (4;1)	0	0,1156	0	0	0	0	0	0	0	0	0	0	0,8844	0	0	0	0	0
11 (4;2)	0	0	0,5718	0	0	0	0	0	0	0	0	0	0	0,42821	0	0	0	0
12 (4;3)	0	0	0,6154	0	0	0	0	0	0	0	0	0	0	0	0,3846	0	0	0
13 (5;1)	0	0,0501	0	0	0	0	0	0	0	0	0	0	0	0	0	0,94991	0	0
14 (5;2)	0	0	0,3501	0	0	0	0	0	0	0	0	0	0	0	0	0	0,64989	0
15 (5;3)	0	0	0,3923	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,6077
16 (6;1)	0	0,1801	0	0	0	0	0	0	0	0	0	0	0	0	0	0,81994	0	0
17 (6;2)	0	0	0,6917	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,30831
18 (6;3)	0	0	0,7289	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,27113

Ilustración 34: Matriz de Transición para el año 2012 de clientes Tipo Empresa, Rubro R7 y que poseen vehículos de otros segmentos.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	(1;1)	(1;2)	(1;3)	(2;1)	(2;2)	(2;3)	(3;1)	(3;2)	(3;3)	(4;1)	(4;2)	(4;3)	(5;1)	(5;2)	(5;3)	(6;1)	(6;2)	(6;3)
1 (1;1)	0	0,2511	0	0,7489	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2 (1;2)	0	0	0,774	0	0,226	0	0	0	0	0	0	0	0	0	0	0	0	0
3 (1;3)	0	0	0,8041	0	0	0,196	0	0	0	0	0	0	0	0	0	0	0	0
4 (2;1)	0	0,1814	0	0	0	0	0,819	0	0	0	0	0	0	0	0	0	0	0
5 (2;2)	0	0	0,6936	0	0	0	0	0,306	0	0	0	0	0	0	0	0	0	0
6 (2;3)	0	0	0,7307	0	0	0	0	0,269	0	0	0	0	0	0	0	0	0	0
7 (3;1)	0	0,1152	0	0	0	0	0	0	0,8848	0	0	0	0	0	0	0	0	0
8 (3;2)	0	0	0,5708	0	0	0	0	0	0	0,4292	0	0	0	0	0	0	0	0
9 (3;3)	0	0	0,6144	0	0	0	0	0	0	0	0,3856	0	0	0	0	0	0	0
10 (4;1)	0	0,1036	0	0	0	0	0	0	0	0	0	0,89644	0	0	0	0	0	0
11 (4;2)	0	0	0,5413	0	0	0	0	0	0	0	0	0	0,45868	0	0	0	0	0
12 (4;3)	0	0	0,5858	0	0	0	0	0	0	0	0	0	0	0,41422	0	0	0	0
13 (5;1)	0	0,0445	0	0	0	0	0	0	0	0	0	0	0	0	0,95547	0	0	0
14 (5;2)	0	0	0,3226	0	0	0	0	0	0	0	0	0	0	0	0	0,67745	0	0
15 (5;3)	0	0	0,3633	0	0	0	0	0	0	0	0	0	0	0	0	0	0,63672	0
16 (6;1)	0	0,1625	0	0	0	0	0	0	0	0	0	0	0	0	0,83746	0	0	0
17 (6;2)	0	0	0,6647	0	0	0	0	0	0	0	0	0	0	0	0	0,33526	0	0
18 (6;3)	0	0	0,7038	0	0	0	0	0	0	0	0	0	0	0	0	0	0,29621	0

Ilustración 35: Matriz de Transición para el año 2013 de clientes Tipo Empresa, Rubro R7 y que poseen vehículos de otros segmentos.

**Anexo H: Predicción Acumulada a 2018 considerando los perfiles formados por las variables “Naturaleza”, “Rubro” y “Tiene Vehículo Otro Segmento”.**

Naturaleza	Rubro F	Otro T Vehículo	# Clientes	Predicción a 2018	Promedio	Desviación Estándar	Valor Máximo	Valor Mínimo
Empresa	R1	NO	49	\$ 86.661.485	\$ 1.768.602	\$ 664.509	\$ 4.720.906	\$ 383.011
		SI	7	\$ 45.291.732	\$ 6.470.247	\$ 1.493.142	\$ 8.214.253	\$ 3.965.559
	R2	NO	105	\$ 7.996.722.273	\$ 76.159.260	\$ 221.751.853	\$ 2.281.823.273	\$ 7.142.588
		SI	32	\$ 5.527.546.771	\$ 172.735.837	\$ 176.341.274	\$ 930.447.471	\$ 27.489.885
	R3	NO	34	\$ 665.328.313	\$ 19.568.480	\$ 50.168.197	\$ 287.754.259	\$ 4.217.799
		SI	19	\$ 1.190.420.309	\$ 62.653.700	\$ 74.110.600	\$ 281.695.166	\$ 9.743.868
	R4	NO	165	\$ 1.882.115.636	\$ 11.406.761	\$ 8.639.002	\$ 71.067.421	\$ 2.016.988
		SI	37	\$ 935.492.526	\$ 25.283.582	\$ 10.581.028	\$ 72.075.967	\$ 16.871.442
	R5	NO	45	\$ 830.806.664	\$ 18.462.370	\$ 16.388.507	\$ 94.432.887	\$ 5.052.974
		SI	16	\$ 551.636.486	\$ 34.477.280	\$ 11.784.245	\$ 58.348.506	\$ 24.649.252
	R6	NO	167	\$ 1.759.178.377	\$ 10.534.002	\$ 6.791.061	\$ 68.454.921	\$ 2.013.686
		SI	20	\$ 610.791.170	\$ 30.539.559	\$ 24.027.050	\$ 123.773.775	\$ 4.106.996
	R7	NO	65	\$ 669.077.517	\$ 10.293.500	\$ 11.153.856	\$ 70.768.435	\$ 2.942.657
		SI	9	\$ 230.907.041	\$ 25.656.338	\$ 26.762.949	\$ 75.021.139	\$ 10.032.345
	R8	NO	264	\$ 4.748.310.049	\$ 17.986.023	\$ 12.446.686	\$ 140.201.396	\$ 2.743.087
		SI	44	\$ 2.124.685.168	\$ 48.288.299	\$ 26.985.454	\$ 148.346.665	\$ 26.602.611
Persona	R1	NO	122	\$ 185.012.868	\$ 1.516.499	\$ 339.769	\$ 2.108.917	\$ 90.767
		SI	7	\$ 23.894.597	\$ 3.413.514	\$ 730.692	\$ 4.544.441	\$ 2.512.704
	R2	NO	103	\$ 378.609.755	\$ 3.675.823	\$ 1.377.155	\$ 9.185.169	\$ 1.364.045
		SI	19	\$ 163.887.279	\$ 8.625.646	\$ 2.122.970	\$ 14.410.554	\$ 5.820.181
	R3	NO	20	\$ 233.751.800	\$ 11.687.590	\$ 2.915.342	\$ 17.222.218	\$ 7.794.717
		SI	4	\$ 20.300.809	\$ 5.075.202	\$ 1.705.092	\$ 7.438.827	\$ 3.564.286

	R4	NO	62	\$ 151.505.716	\$ 2.443.641	\$ 637.712	\$ 3.476.701	\$ 122.393
		SI	7	\$ 35.978.166	\$ 5.139.738	\$ 2.775.915	\$ 8.879.409	\$ 335.654
	R5	NO	17	\$ 477.202.226	\$ 28.070.719	\$ 12.558.885	\$ 41.214.426	\$ 1.600.391
	R6	NO	224	\$ 588.149.001	\$ 2.625.665	\$ 1.215.331	\$ 8.696.480	\$ 516.876
		SI	27	\$ 196.340.836	\$ 7.271.883	\$ 4.332.086	\$ 18.925.926	\$ 2.038.689
	R7	NO	21	\$ 59.632.090	\$ 2.839.623	\$ 604.443	\$ 4.037.343	\$ 1.725.694
		SI	2	\$ 11.678.219	\$ 5.839.109	\$ 1.068.398	\$ 6.594.581	\$ 5.083.638
	R8	NO	148	\$ 243.408.574	\$ 1.644.653	\$ 420.286	\$ 3.150.398	\$ 275.497
		SI	11	\$ 37.987.848	\$ 3.453.441	\$ 1.068.446	\$ 5.548.079	\$ 1.035.330
	<b>Total general</b>			<b>1872</b>	<b>\$ 32.662.311.302</b>	<b>\$ 17.447.816</b>	<b>\$ 64.656.511</b>	<b>\$ 2.281.823.273</b>

Tabla 54: Resumen Predicción CLV acumulado a 2018 a nivel de Perfil "Naturaleza"-"Rubro"-"Vehículo de otro segmento"