



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

**ESTIMACIÓN DEL CUSTOMER LIFETIME VALUE A NIVEL DE  
CLIENTES, DE UN BANCO USANDO VARIABLES TRANSACCIONALES  
Y SOCIODEMOGRÁFICAS**

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

**VALENTINA ANAÍS RODRÍGUEZ HERRERA**

**PROFESOR GUÍA:**

**LUIS ABURTO LAFOURCADE**

**MIEMBRO DE LA COMISIÓN**

**ALEJANDRA PUENTE CHANDÍA**

**PABLO MARÍN VICUÑA**

**SANTIAGO DE CHILE**

**ENERO 2013**



## RESUMEN EJECUTIVO

RESUMEN DE LA MEMORIA PARA OPTAR AL  
TÍTULO DE INGENIERO CIVIL INDUSTRIAL  
POR: VALENTINA RODRÍGUEZ HERRERA  
FECHA: 21/01/2013  
PROF. GUÍA: SR. LUIS ABURTO LAFOURCADE

Los consumidores están siendo cada día más exigentes, presionando un cambio del enfoque de negocios desde los productos hacia los clientes. Además, cada individuo es diferente y tiene una evolución distinta de su consumo en una empresa, por lo que tomar decisiones basados en su conducta actual y pasada no es suficiente.

El objetivo de este proyecto es estimar el CLV (customer lifetime value) de los clientes de un banco, usando variables transaccionales y socio-demográficas, y así focalizar las acciones de promoción para retención. Los individuos de estudio son poco más de 17 mil cuenta-correntistas nuevos. El cálculo del CLV se realiza en un horizonte de 4 años y es aplicable a un 41% del total de los clientes. Los objetivos específicos son: evaluar las variables transaccionales y sociodemográficas que se incluirán en los modelos, estimar la renta, estimar tres modelos para predecir de dos formas distintas el CLV y validarlos. Por último, identificar al segmento más valioso y proponer acciones de promoción para fidelizar a este grupo, por medio de una asignación más eficiente del presupuesto de promociones que supera los \$1.000 MM de pesos.

La metodología utilizada considera las etapas de pre-procesamiento y limpieza de los datos, un análisis exploratorio a fin de observar relaciones significativas entre las variables transaccionales (deuda, saldo, tenencia, etc.) y sociodemográficas (género, edad, e. civil etc.). Luego, se estima la renta de los clientes con una regresión lineal para utilizarla en la estimación del CLV de los clientes. Se abarcan dos enfoques, el primero consiste en realizar una regresión lineal sobre la variable CLV calculada a partir de los datos, y el segundo sobre las variables CLV promedio (CLV/duración) y duración por separado, las cuales se multiplican para obtener un CLV final. Los modelos son validados con una muestra de datos nueva. El horizonte de tiempo utilizado es de 4 años y la tasa de descuento de 12%. Finalmente, se realiza una caracterización de los clientes más valiosos y se entregan propuestas de promoción para su fidelización.

En la predicción de la renta se obtiene un MAPE (Mean Absolute Percentage Error) de 46%. Se observan zonas donde los clientes rentan un 21% más que un cliente promedio. Lo mismo ocurre con el género masculino. Para ambos modelos de valoración se obtiene un WAPE (Weighted Absolute Percentage Error) superior al 80%. El mejor desempeño lo tiene el enfoque con VP promedio y duración, que acierta en un 85% de los clientes que realmente pertenecen al mejor quintil de valor. Los clientes más valiosos en el plazo de 48 meses son los altos ejecutivos de edades mayores a 50 años.

Se propone promocionar la adquisición de créditos de consumo y aumento del cupo en la tarjeta de crédito. Para trabajos futuros se recomienda incorporar otras variables transaccionales como cantidad de transacciones y montos con la tarjeta de débito y crédito. Además se propone evaluar los drivers de lealtad de los clientes de un banco.



*Dedicado a mi FAMILIA.*

## AGRADECIMIENTOS

Agradezco enormemente a mi familia que siempre me ha apoyado en todas mis actividades.

A mis papás Jaime y Marcia que me ensaaron lo que es el bien y siempre me han empujado a lograr mis objetivos.

A mis hermanas Carolina, Alejandra y Mabel que han sido un gran ejemplo para mí. A mis cuñados Eduardo, Cristian y Diego que me regalonean como a una hermana menor.

A mis sobrinos Damián, Laura, Isidora y Alfonso que llegaron a chasconear mis estudios con sus gritos, risas y llantos. Apenas podía escuchar mis pensamientos mientras estudiaba, pero me recordaban cuáles son las cosas importantes de la vida. A mi abuelita Alicia, tías, primas y suegros que sé que me desean lo mejor.

A Patrick, gracias por siempre regalarme tu amor y apoyo.

A mis amigas Aintzane y Mariel, que en estos seis años siempre siguieron siendo mis grandes compañeras.

A mis amigos y amigas de la Escuela, a Giorgio, Claudia, Javier, Rocío, Daniel y Hernán, sin su apoyo me hubiera costado, más de lo que fue, llegar hasta aquí.

A los profesores de la Universidad Mattia Makovec, Claudio Pizarro y Marcel Goic que, además de entregar su conocimiento, se notó su preocupación por el aprendizaje de los alumnos.

A todos los funcionarios de la Facultad, bibliotecarias de la Escuela e Industrias, encargados del CEC y de equipamiento, guardias del 1er piso de Blanco y señoras de la limpieza, que se encargaron de cada día entregarme su cariño. Gracias Karen, Gaby y Yunny por su preocupación.

Agradezco a Leopoldo, Rodrigo, Daniela y Mauricio por abrirme las puertas y confiar en mí aún sin conocerme.

A los profesores Guía, Co-guía e Integrante, que me ayudaron a hacer lo mejor de la memoria.

Finalmente, le agradezco a Dios. Que el llegar a ser ingeniera, sea una herramienta más que pueda usar para construirme cada día como una mejor persona.

## I. Tabla de contenido

I.	ANTECEDENTES GENERALES DE LA INDUSTRIA.....	1
1.	CUSTOMER RELATIONSHIP MANAGEMENT .....	1
A.	Ciclo de Vida del Cliente .....	2
2.	SISTEMA BANCARIO CHILENO .....	3
A.	Monto colocaciones.....	3
B.	Mercado Personas .....	4
C.	Cuentas Corrientes.....	5
D.	Productos financieros .....	7
II.	DESCRIPCIÓN DEL PROYECTO Y JUSTIFICACIÓN .....	7
1.	Acciones de Marketing .....	7
A.	Tarjetas de Crédito .....	7
B.	Campañas Generales.....	9
2.	Justificación .....	10
	Esto se convierte en una motivación para buscar otro medio de segmentar a los clientes por valor. ....	11
A.	Quintiles.....	11
B.	Plan de Toques .....	13
III.	OBJETIVOS .....	15
1.	OBJETIVO GENERAL.....	15
2.	OBJETIVOS ESPECÍFICOS.....	16
IV.	ALCANCES.....	16
V.	RESULTADOS ESPERADOS .....	16
VI.	MARCO TEÓRICO.....	17
1.	CUSTOMER LIFETIME VALUE .....	17
A.	Horizonte de Vida.....	18
B.	Cross-selling .....	18
C.	Tasa de Retención .....	18
2.	Modelo de regresión lineal .....	19
3.	Calibración y validación .....	19
4.	Testeo.....	19
A.	Tests de análisis exploratorio.....	19
B.	Tests de predicción .....	20
5.	Modelo Gamma-Weibull.....	21
VII.	DESARROLLO METODOLÓGICO.....	22

1.	Explicación de estimación de <i>CLV</i> (Valor Presente VP)	22
2.	Selección de Datos	23
3.	Pre-procesamiento y transformación	24
4.	Análisis Exploratorio	25
5.	Modelamiento	26
6.	Evaluación	26
VIII.	APLICACIÓN DE LA METODOLOGÍA	26
1.	Selección de Datos	27
2.	Pre-procesamiento y Transformación	28
3.	Análisis descriptivo	30
A.	Valor Presente	32
B.	Quintil Final y variables	37
C.	Duración	38
D.	Valor Presente Promedio	42
4.	Resultados Modelamiento	42
A.	Renta	42
B.	Valor Presente	45
C.	Duración	46
D.	Valor Presente Promedio	48
E.	Resultados de Ajuste modelos de predicción de Valor Presente	51
F.	Evaluación de la Regresión según tipo de datos	51
G.	Conclusiones	53
IX.	RANKING Y PROPUESTAS DE FIDELIZACIÓN	54
1.	Ranking	54
2.	Propuestas de promoción para Fidelización	55
X.	TRABAJOS FUTUROS	57
XI.	BIBLIOGRAFÍA	59
XII.	ANEXOS	62
	Anexo A. Participación en Monto de colocaciones de los bancos “Mediados” del Sistema bancario chileno.	62
	Anexo B. Distribución cartera principales bancos	62
	Anexo C. Explicación de Quintiles de valor.	63
	Anexo D. Porcentaje de toques realizados para cada grupo según Quintil Inicial.	64
	Anexo E. Cálculo del tamaño muestral	64
	Anexo F. Asignación Zona-Comuna.	64

<b>Anexo G. Test Anova edad vs. género.....</b>	<b>67</b>
<b>Anexo H. Frecuencia Estado Civil.....</b>	<b>67</b>
<b>Anexo I. Gráfico Box-Plot estado civil vs. Edad. ....</b>	<b>67</b>
<b>Anexo J. Test Anova duración vs. Quintil Final.....</b>	<b>68</b>
<b>Anexo K. Tests Anova's realizados sobre Productos de Colocación.....</b>	<b>68</b>
<b>Anexo L. Quintil de mayor valor del modelo de predicción con duración. ....</b>	<b>69</b>

## Índice de Tablas

Tabla 1. Acciones de marketing unidad encargada de todos los segmentos. ....	9
Tabla 2. Segmentación de clientes según Renta. Fuente: Banco.....	10
Tabla 3. Contribución Acumulada (suma) de los primeros 5 meses del año 2012 de los cuenta-correntistas del Banco (714.000), según Segmento y Estado Civil. Fuente: Banco. ....	11
Tabla 4. Fórmulas de cálculo de las denominaciones de Valor Inicial (Contribución de los primeros tres meses en el banco) y Valor Final (horizonte máximo de 48 meses), de los clientes de estudio. ....	11
Tabla 5. Matriz de Quintiles de Valor Inicial y Final de los que presenta cambio de valor de cada Quintil Inicial al finalizar su relación con el banco.....	12
Tabla 6. Matriz de Quintiles que representa el Valor Total de cada Grupo Inicial-Grupo Final. ....	12
Tabla 7. Cantidad de clientes por grupo. ....	13
Tabla 8. Toques realizados para cada grupo. ....	14
Tabla 9. Probabilidad que individuo dado que está en un Quintil Inicial, se mueva a un Quintil Final dado que no se realizan toques. ....	15
Tabla 10. Probabilidad que individuo dado que está en un Quintil Inicial, se mueva a un Quintil Final dado que se realizan toques. ....	15
Tabla 11. Tabla de Frecuencia de Zonas por Comunas.....	27
Tabla 12. Tabla de Frecuencia del género.....	28
Tabla 13. Tabla de Frecuencia de tramos de edad. ....	28
Tabla 14. Tabla de Frecuencia de ocupación y asignación de etiquetas. ....	29
Tabla 15. Porcentaje de participación en estado civil por nivel de tramo etario. ...	30
Tabla 16. Porcentaje de participación en género por nivel de tramo etario.....	30
Tabla 17. Participación en tramo etario por ocupación. ....	31
Tabla 18. Correlaciones de variable CLV vs. Renta. ....	31
Tabla 19. Correlación de VP con variables de Contribución del Mes 3.....	31
Tabla 20. Cuartiles de los clientes por variable duración.....	38
Tabla 21. Resultados del modelamiento de duración de los clientes utilizando una función Gamma-Weibull. ....	39
Tabla 22. Test Anova para testear diferencia de medias de VP entre cohortes. ....	41
Tabla 23. Resultados globales de ajuste del modelo de predicción de la renta.....	43

<b>Tabla 24. Correlación VP con Renta estimada.....</b>	<b>44</b>
<b>Tabla 25. Resultado estimación con regresión lineal de la variable duración.....</b>	<b>47</b>
<b>Tabla 26. a. Resultado estimación no significativa con regresión lineal de la variable duración con regresor Contribución Gasto Mes 3. b. Resultado estimación significativa con regresión lineal de la variable duración con los regresores Contribución Gasto Mes 3 y Tenencia de LDC en Mes 3. ....</b>	<b>48</b>
<b>Tabla 27. Estimaciones de regresión lineal con VP-promedio como variable dependiente. ....</b>	<b>49</b>
<b>Tabla 28. Estadísticos de ajuste de modelos Con y Sin duración. ....</b>	<b>51</b>
<b>Tabla 29. Descriptivo de VP para los Quintiles 3, 4 y 5.....</b>	<b>51</b>
<b>Tabla 30. Resultados WAPE de estimación del modelo de regresión lineal sobre la variable VP con distintos tipo de datos.....</b>	<b>52</b>
<b>Tabla 31. Gráfico Box-Plot de cada Quintil vs. Desviación porcentual absoluta individual.....</b>	<b>53</b>
<b>Tabla 32. Gráfico Box-Plot de cada Tramo Etario vs. Desviación porcentual absoluta individual.....</b>	<b>53</b>
<b>Tabla 33. Matriz de individuos de participación de Quintiles reales en la agrupación de quintiles obtenida a partir de los valores estimados con los modelos de regresión lineal. ....</b>	<b>55</b>
<b>Tabla 34. Explicación breve de Quintiles de valor inicial y final. ....</b>	<b>63</b>
<b>Tabla 35. Porcentaje de toques realizados para cada grupo según Quintil Inicial.</b>	<b>64</b>
<b>Tabla 36. Gráfico Box-Plot estado civil vs. Edad.....</b>	<b>67</b>
<b>Tabla 37. Test ANOVA para evaluar significatividad de tenencia de productos de consumo sobre el VP.....</b>	<b>68</b>
<b>Tabla 38. Tabla de Frecuencia de Tenencia de Consumo Mes 3 para Quintil de mayor valor del modelo de predicción con duración. ....</b>	<b>69</b>

## Índice de Ilustraciones

Ilustración 1. Crecimiento de la rentabilidad del cliente a lo largo de su ciclo de vida. Fuente: Trabajo de Tesis [37].....	2
Ilustración 2. Participación en Sistema Financiero. Julio 2012. Fuente: SBIF, Elaboración propia.....	3
Ilustración 3. Participación en Monto de colocaciones de los primeros cuatros bancos del Sistema bancario chileno. Fuente: Banco de Chile.....	4
Ilustración 4. Participación en Monto de colocaciones del mercado Personas de los bancos “Mediados” del Sistema bancario chileno. Fuente: Banco de Chile.....	5
Ilustración 5. Participación en Cuenta Corriente. Finales del 2011. ....	6
Ilustración 6. Montos de Cuentas corrientes de personales naturales por Entidad. Elaboración propia. Fuente: SBIF.....	6
Ilustración 7. Customer Lifetime Value. ....	17
Ilustración 8. Gráficas que describen la agrupación de cada grupo de entreda en uno grande. No considera como información el mes de entrada los clientes.....	23
Ilustración 9. Variables a utilizar en el modelamiento. ....	24
Ilustración 10. Definición de contribuciones. ....	24
Ilustración 11. Tabla de Frecuencia de estado civil versión 2.....	29
Ilustración 12. Gráfico de Dispersión de la Contribución Neta de los primero 3 meses vs. VP. ....	33
Ilustración 13. Gráfico Box-Plot Zona por Comunas vs. Valor Presente.....	33
Ilustración 14. Gráfico Box-Plot Tramo Etario vs. Valor Presente.....	34
Ilustración 15. Tabla que presenta los Coeficientes de Variación del VP por tramo Etario.....	35
Ilustración 16. Gráfico Box-Plot ocupación vs. Valor Presente. ....	35
Ilustración 17.a. Porcentaje hombre y mujer en los Quintiles de valor. b. Gráfico Box-Plot del género según su renta. ....	36
Ilustración 18. Gráfico Box-Plot de VP respecto a Tenencia de productos Línea de Crédito y Crédito de Consumo en Mes 3.....	36
Ilustración 19. Gráfico Box-Plot de Quintiles por Deuda en TDC del Mes 2. ....	37

<b>Ilustración 20. a. Tablas de porcentaje Quintil vs. Tenencia de Crédito Hipotecario Mes 3. b. Tablas de porcentaje Quintil vs. Tenencia de Crédito de Consumo Mes 3.</b>	<b>38</b>
.....	
<b>Ilustración 21. Gráfico de frecuencia de duración.</b>	<b>38</b>
<b>Ilustración 22. Gráfico de frecuencia de clientes que permanecen en el banco en cada mes.</b>	<b>39</b>
<b>Ilustración 23. Gráfico que compara la trayectoria de la frecuencia de abandono de los clientes con respecto a su estimación.</b>	<b>40</b>
<b>Ilustración 24. Gráfico Box-Plot de Quintiles vs. Duración en meses.</b>	<b>40</b>
<b>Ilustración 25. Tenencia de Línea de Crédito en Mes 3 vs. Duración del Cliente.</b>	<b>41</b>
<b>Ilustración 26. Gráfico Box-Plot del VP promedio por Zonas.</b>	<b>42</b>
<b>Ilustración 27. Resultados estimación de la variable renta con regresión lineal.</b>	<b>43</b>
<b>Ilustración 28. Grafico Box-Plot de la renta y renta estimada según la ocupación.</b>	<b>44</b>
.....	
<b>Ilustración 29. Resultados de Ajuste global del modelo con variable dependiente VP.</b>	<b>46</b>
<b>Ilustración 30. Gráfico Box-Plot de Quintiles Finales vs. Número de Cheques girados en el tercer mes de relación con el banco.</b>	<b>50</b>
<b>Ilustración 31. Gráfico Box-Plot tenencia de LDC en Mes 3 y VP promedio</b>	<b>50</b>
<b>Ilustración 32. Contribución basilea Mes 3 entregada por cada Quintil de valor final.</b>	<b>52</b>
<b>Ilustración 33. Participación por ocupación en Quintil de mayor valor predicho a partir del Modelo Con duración.</b>	<b>56</b>
<b>Ilustración 34. Participación en Monto de colocaciones de los bancos “Mediados” del Sistema bancario chileno. Fuente: Banco de Chile.</b>	<b>62</b>
<b>Ilustración 35. Distribución sectorial de la cartera de colocaciones (%). Diciembre 2011. Fuente: Elaboración propia, SBIF.</b>	<b>63</b>

# I. ANTECEDENTES GENERALES DE LA INDUSTRIA

## 1. CUSTOMER RELATIONSHIP MANAGEMENT

La alta competencia entre las empresas y el gran desarrollo tecnológico han hecho que la captación de clientes y el mantenerlos fidelizados sea cada vez más complejo. Por esto, entender el comportamiento del cliente es fundamental. Se ha reconocido que el verdadero valor no está en los productos, sino que en el valor que se entrega a los clientes [40], de que es necesario entenderlos y así hacer una asignación efectiva y eficiente de los recursos.

El costo de adquisición puede ser al menos cinco veces mayor que el de captación [1],[2]. Por lo que resulta más conveniente, en términos económicos, priorizar la retención. Esto exige un cambio en la gestión de la relación con el cliente, pasando de un marketing transaccional (lanzamiento de productos al mercado esperando que éste sea el que capte a los clientes), a un marketing relacional [3].

Este último tiene como fin, generar relaciones estables y de continuidad con los clientes más rentables para la compañía, fortalecer la relación y acrecentar el nivel de satisfacción y lealtad. Con esto, se espera maximizar las oportunidades de ventas cruzadas (*cross-selling*) y de ventas de productos de mayor valor agregado (*up-selling*).

Este tipo de marketing se basa en el hecho de que los costos operativos asociados al servicio del cliente se reducen a lo largo del tiempo. Esto sucede gracias a que este último va adquiriendo mayor conocimiento de los procesos de la empresa [4].

El marketing relacional se desarrolla a través del concepto *Customer Relationship Management (CRM)*. Éste, es un modelo de gestión orientado a seleccionar y gestionar los clientes, con el objeto de optimizar su valor a largo plazo. No está restringida solamente a las relaciones de compra y venta, también se ocupa de la gestión y registro de la información obtenida del cliente (información transaccional, reclamos, sugerencias, etc.). Existen cuatro tipos de *CRM* según sus funciones [5]:

- *CRM* Estratégico: Se enfoca en la estrategia central del negocio, en el cliente. Tiene como objetivo capturar y mantener clientes rentables.
- *CRM* Operacional: Se centra en la automatización de los procesos de cara al cliente, tales como ventas, marketing y servicio al cliente.
- *CRM* Analítico: Se dedica a la explotación inteligente de los datos relacionados con los clientes. Se puede desarrollar con objetivos estratégicos o tácticos.
- *CRM* Colaborativo: Consiste en la aplicación de la tecnología, a través de las barreras organizativas, a fin de optimizar el valor de la empresa, de los socios y los clientes.

Este trabajo trata un subtema del *CRM Analítico*, el cual, consiste en el desempeño de las actividades de captura, almacenamiento, extracción, integración,

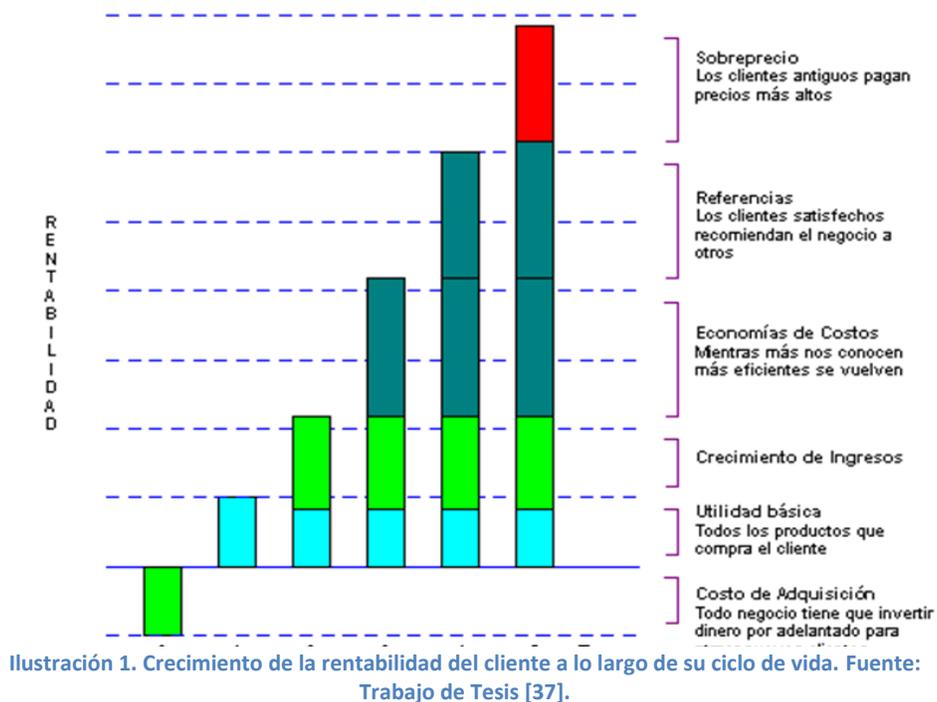
procesamiento, interpretación, distribución, uso y reporte de los datos relacionados con los clientes. Su propósito es mejorar tanto el valor del cliente, como el de la empresa [5].

A partir del análisis de la información se puede obtener conocimiento nuevo como: cuáles son los clientes más valiosos, cuáles son los más propensos a abandonar la empresa y cuáles son más probables de responder a cierta oferta. Precisamente la primera pregunta es la que se desea responder, en el caso específico de un banco.

### A. Ciclo de Vida del Cliente

A través del ciclo de vida de un cliente, se puede observar que éste inicialmente genera solo pérdidas (llamado costo de captación). Al capturarlo, inherentemente entrega un valor base debido a la entrega misma del servicio. Luego, incrementa su valor con las compras que realiza y el mayor conocimiento que alcanza el cliente sobre los productos y servicios ofrecidos.

A medida que el cliente va conociendo a la empresa, reduce los costos operativos, pues sabe cómo interactuar y evita operaciones innecesarias. Finalmente, la fidelización genera la promoción del negocio por parte del propio cliente con referencias positivas [4].



## 2. SISTEMA BANCARIO CHILENO

El sistema bancario de Chile está concentrado. Seis bancos conforman el 80% de las colocaciones. Banco Santander se encuentra en primer lugar con una participación de un 20%, seguido de cerca por Banco de Chile con 19%. Luego se encuentran Banco Estado, BCI, Corpbanca y BBVA, con 14%, 12% y 7% respectivamente. Otros bancos representan el 18% y están representados por Itaú, Security, Bice, Falabella, entre otros.

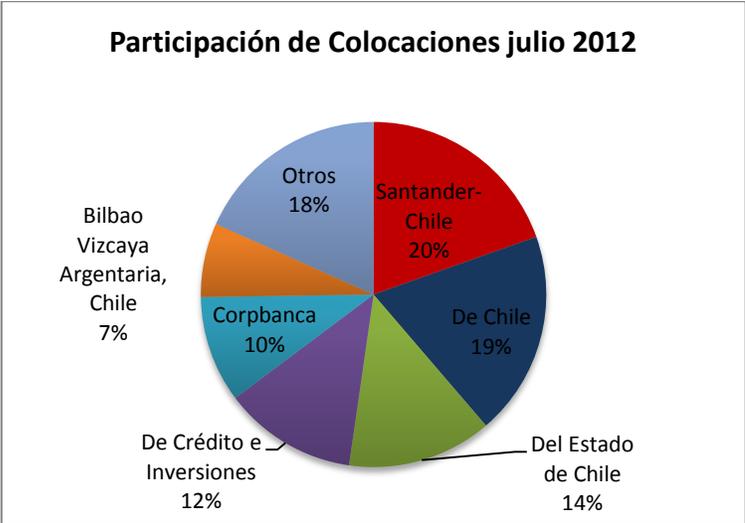
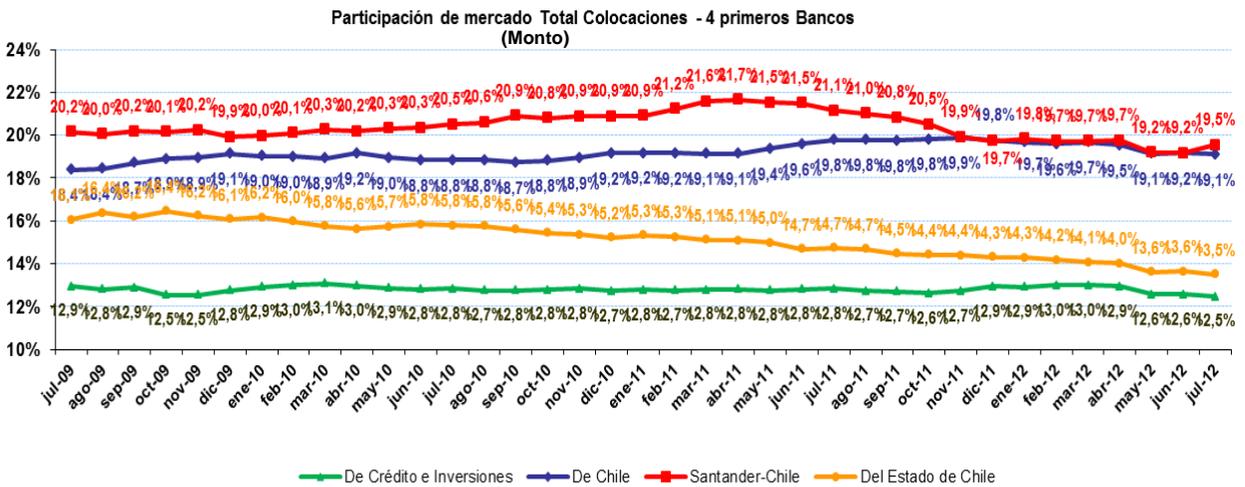


Ilustración 2. Participación en Sistema Financiero. Julio 2012. Fuente: SBIF, Elaboración propia.

### A. Monto colocaciones

En cuanto a la evolución del monto en colocaciones, en diciembre de 2011, Banco de Chile supera por primera vez desde el inicio del análisis a Santander, quien vuelve a retomar su liderazgo al siguiente mes, pero con una diferencia estrecha. Banco Estado pierde de manera continua cuota de mercado. Se puede observar además la estabilidad de BCI, cuya participación se mantiene relativamente inmutable en el tiempo.

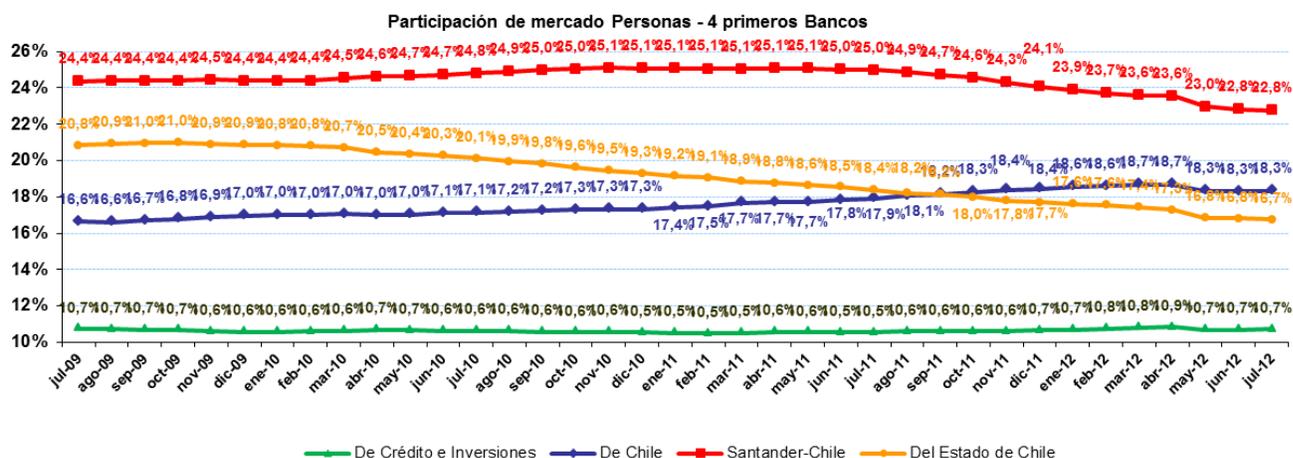


**Ilustración 3. Participación en Monto de colocaciones de los primeros cuatro bancos del Sistema bancario chileno. Fuente: Banco de Chile.**

Respecto a la participación de los bancos medianos (Ver Anexo A), a partir de agosto de 2011, CorpBanca se aleja sustantivamente de BBVA (6,9%), alcanzando un 10% de participación. Itaú (4%) presenta un aumento sostenido de su participación acercándose a Scotiabank (5%), quien, al contrario, ha disminuido su porcentaje de mercado desde julio de 2009.

## B. Mercado Personas

En el mercado de Personas (Consumo y Vivienda), Banco de Chile ha tenido un crecimiento continuo, subiendo un 0,43% de su participación en un año (julio 2011 a 2012) y 1,64% en dos. La institución pasa a ser el segundo competidor en Personas en Octubre del año pasado. Su alto desempeño se debe al aumento de participación del mercado Hipotecario, con altos montos por persona. Al contrario, Banco Santander ha tenido la mayor variación negativa con una caída en su participación de -2,22% en un año, seguido de Banco Estado con un -1,66% (-3,28% en dos años). Las tendencias de estos últimos siguen a la baja.



**Ilustración 4. Participación en Monto de colocaciones del mercado Personas de los bancos “Mediados” del Sistema bancario chileno.**  
Fuente: Banco de Chile.

El mercado de Consumo se encuentra relativamente estable. No obstante, Banco Estado tuvo una alta caída de -1,97% desde marzo de 2011 a marzo de 2012, permaneciendo invariable desde entonces. Junto a esto, Banco Santander, a pesar de tener un aumento de 1,94% en un año (2010 al 2011), pasó a disminuir su cuota en un -2,11% al año siguiente.

Distinto es el escenario del mercado de la Vivienda, donde Banco Estado lideró hasta diciembre de 2010, al ser superado por Banco Santander, en mayor medida debido a una baja del primero (caída de -4,49% desde el 2010 hasta julio 2012). Santander también ha visto reducida su participación desde octubre de 2011. Al mismo tiempo, Banco de Chile está acercándose a los bancos recién mencionados, al aumentar en 2,84% su cuota en tres años.

Para mayor profundización, en el Anexo B. se describe la distribución de la cartera de los bancos más importantes.

### C. Cuentas Corrientes

Se observa que Santander y Chile se disputan el primer lugar en el número de cuentas corrientes que tienen a disposición. Los siguen BCI (16%) y Banco Estado (11%). Al interior de Otros se encuentran Banco Falabella (6%) y BBVA (5%).

### Participación Número en Cuentas Corrientes Personas Naturales

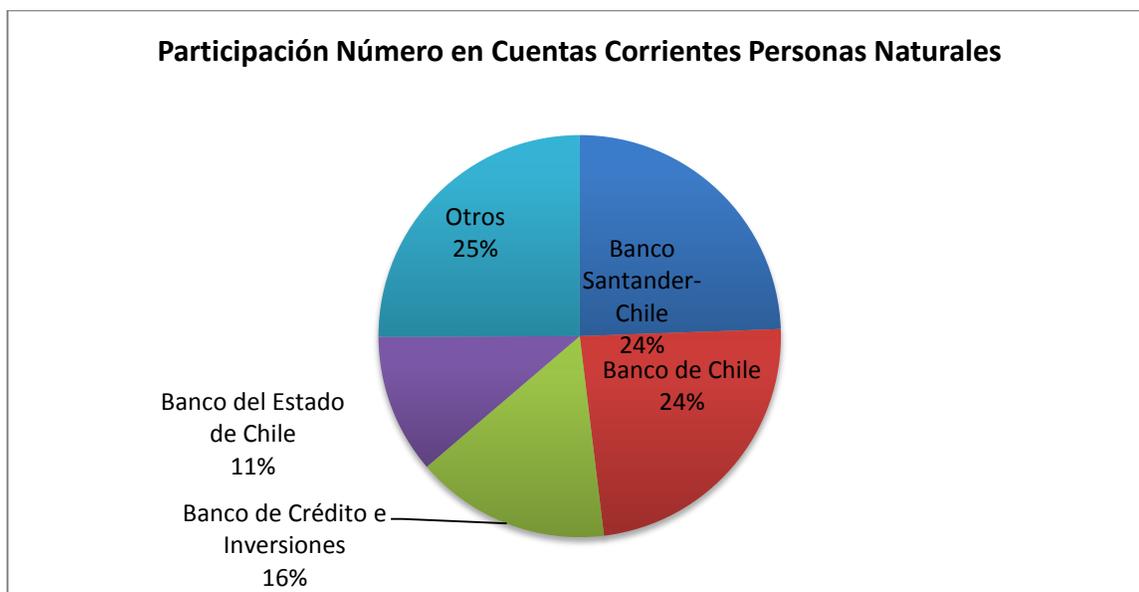


Ilustración 5. Participación en Cuenta Corriente. Finales del 2011.

A pesar de la estrecha competencia en número de cuentas, Banco de Chile posee a disposición el mayor monto total del sistema, superando a Santander, BCI y Estado. Se observa una evolución similar entre los tres primeros.

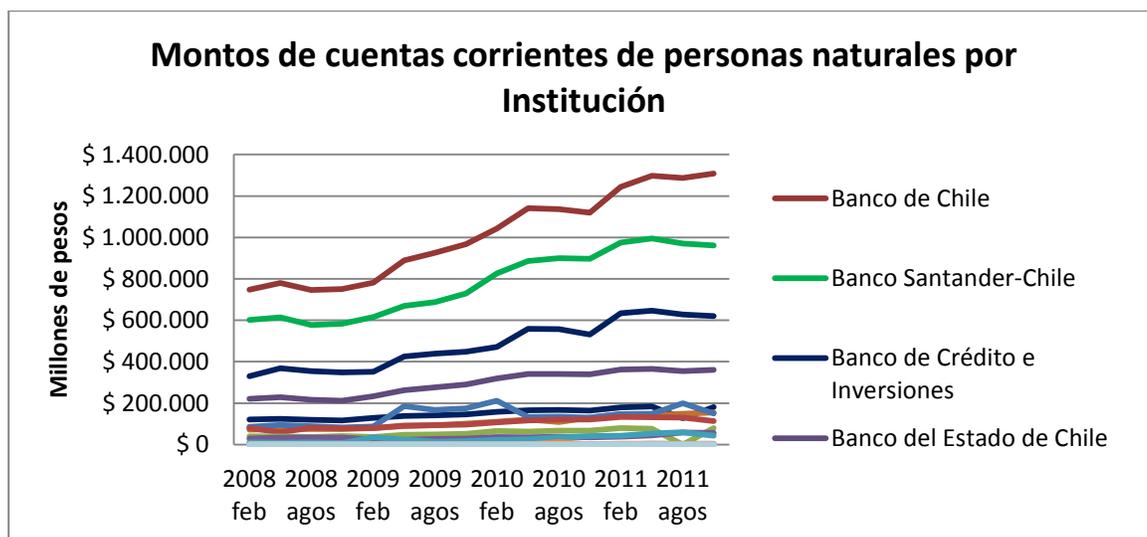


Ilustración 6. Montos de Cuentas corrientes de personales naturales por Entidad. Elaboración propia. Fuente: SBIF.

## D. Productos financieros

### o *Tarjeta de Crédito*

Se manifiesta una alta competencia entre Santander y Banco de Chile en cuanto a montos de Tarjetas de Crédito, en donde evolucionan de manera similar alcanzando alrededor de los \$MM 550 mil en diciembre de 2011.

### o *Tarjeta de Débito y Ahorro*

El banco Estado destaca de sobremanera en estos productos. En marzo del 2008, el ente mantenía montos de operaciones totales de tarjeta de débito cercanos a los valores de los demás bancos. Sin embargo, en septiembre de 2011 se escapa en poco más de un billón de pesos de su competidor más cercano. En Ahorro, el banco estatal mantiene más del 80% de participación de los montos desde al menos noviembre de 2010.

### o *Depósitos a Plazo*

Nuevamente se observa una alta disputa por el liderazgo entre Santander y Banco de Chile con montos cercanos a los \$MM 8 millones en agosto de 2011.

## II. DESCRIPCIÓN DEL PROYECTO Y JUSTIFICACIÓN

A continuación se describen las distintas áreas y alguna acciones de marketing que realizados por la empresa.

### 1. Acciones de Marketing

Existen dos áreas principales destinadas a diseñar las acciones de adquisición, fidelización y retención de los clientes. Una unidad observa a los clientes en general y la otra división trabaja en la acciones de fidelización a través de las tarjetas de crédito (involucra tanto a los clientes cuenta-correntistas, como mono-productos que solamente tienen tarjeta). Estas dos sub-áreas trabajan en conjunto con el área de Publicidad.

La unidad enfocada al desarrollo de tarjetas de crédito ha dispuesto de variadas tarjetas con beneficios comunes y distintos. Abajo se describe brevemente las cualidades y beneficios de promoción de este producto:

#### A. Tarjetas de Crédito

Las tarjetas de crédito, además de su servicio financiero, entregan beneficios en Vestuario y Calzado, Restaurantes y Entretenimiento, Tecnología y Hogar, Viajes, etc. Existen distintos tipos para distintas ocasiones e intereses del usuario. No es necesario tener una Cuenta Corriente para disponer de ellas. Se listan a continuación [20]:

- Banco de Chile / AAdvantage

Cada compra que se efectúe con Tarjeta de Crédito Banco de Chile / AAdvantage, ya sea en dólares o pesos, tanto en Chile como en el extranjero, acumula millas.

- Tarjeta Crédito Universal

Se puede comprar en más de 70 mil comercios en el país y realizar avances en todos los cajeros automáticos. Se accede a todos los descuentos y beneficios que tienen las Tarjetas del Banco de Chile en restaurantes, entretenimiento, tecnología, vestuario y calzado.

- Tarjeta Entel Visa

Al igual que la anterior, se puede comprar en más de 70.000 comercios. Por cada \$100 de compras con la Tarjeta Entel Visa se acumula 1 punto Zona Entel (equivale a \$1). Se pueden acumular y canjear por equipos, bolsas de minutos, recargas, Tarjetas SIM, reparaciones en Servicio Técnico, entre otros.

- Tarjeta de Crédito *Travel Club*

Programa de beneficios que permite a los socios del programa Travel Club acumular Dólares-Premio para canjearlos por viajes y productos.

- Tarjeta Club La Tercera Visa

Se puede utilizar como tarjeta de crédito y acceder a los beneficios que entrega el diario La Tercera.

- MasterCard *Black*

Posee beneficios en Chile y el extranjero. No cobra comisión en compras en el extranjero y posee variados seguros de cobertura en el exterior. Además, esta tarjeta pertenece al programa de beneficios Travel Club mencionado arriba.

- Visa *Signature*

Dista de MasterCard Black en los tipos de seguros.

- Pago Automático de Cuentas (PAT)

Con este sistema se pueden pagar las cuentas de manera automática utilizando el saldo de la cuenta corriente o la tarjeta de crédito. El beneficio para el banco es el poder aumentar el grado de relación del cliente con la empresa, elevando la barrera de salida.

## B. Campañas Generales

En la unidad encargada de todos los clientes se realizan distintas acciones que se mencionan a continuación:

Acción	Tipo
Campaña de Bienvenida	Fidelización
Campaña de Inactivos (nueva)	Retención Semi-reactiva
Campaña de Retención	Retención Reactiva
Campaña Invita y Gana	Captación
Plan de Toques	Fidelización
Campaña de Adquisición	Captación
Campaña de Vinculación	Cross-selling / Up-selling

Tabla 1. Acciones de marketing unidad encargada de todos los segmentos.

De modo general, radican en:

- **“Campaña” de Bienvenida:** Call Center/Ejecutivo realiza bienvenida vía teléfono al cliente, confirmando productos y asegurándose de que esté todo en orden.
- **Campaña de Inactivos (nueva):** Como filtro principal considera a los clientes que no han tenido movimiento voluntario (actividades que no sean efectos del banco, por ejemplo, cobro de comisiones o movimiento de la deuda debido a los intereses), en al menos 3 meses a 6 meses o en 7 meses en adelante. No existe ningún criterio dependiente del valor actual ni potencial del cliente.
- **Campaña de Retención:** Esta es una medida reactiva que se efectúa cuando el cliente demuestra interés en cerrar su cuenta corriente. Tiene dos alternativas de procedimiento dependiendo de la acción inicial del cliente. Cuando el cliente entrega una carta física, se tienen dos días para cerrar la cuenta, por lo que no hay interacción con el cliente. Si el cliente manifiesta de forma oral su deseo de abandono, con el ejecutivo o mesón de ayuda al cliente, se le solicita que llame a un Call Center, el cual, tiene cargado un Pull de ofertas dependiente del estado del cliente.
- **Campaña Invita y Gana:** Un cliente obtiene un regalo en efectivo por recomendaciones exitosas y participa en un sorteo de regalos.
- **Plan de Toques:** Este es un plan anual, que considera invitaciones a distintos eventos, unos son reservados y otros más masivos. Existe un presupuesto definido y según eso se asigna la cantidad de clientes que serán invitados a cada evento. Los clientes se pueden repetir entre eventos y las condiciones de invitación consideran: Segmento (mencionado en tabla 2.), Tipo de Banca (es el modelo de atención, existen la Banca Privada, Preferencial, Selecta, Tradicional, Joven, entre otros), Comportamiento Verde (el cliente paga al día), Rentabilidad (Contribución acumulada anual) y Marca (Banco al que pertenece). Además, existen otros filtros a los que se recurre según el evento:

género, edad, intereses, región, cliente invitado en años anteriores, validación de Red (ejecutivos). Los eventos incluyen: invitación a ópera, Torneo de Golf, regalo de cumpleaños y Navidad, etc.

- **Campaña de Adquisición:** Actualmente se enfoca únicamente en la captación de los universitarios con carrera de rentabilidad futura alta, con visitas y creación de cuentas corrientes en sus facultades. Esta rentabilidad se identifica por la Carreras que en información pública se sabe que tendrán los más altos salarios promedios durante su periodo laboral activo.
- **Campaña de Vinculación:** Ofrece productos que generan una mayor vinculación del cliente con la empresa, pero no necesariamente fidelización. Entre los productos están PAT (Pago automático con cargo a la tarjeta de crédito), PAC (Pago automático con cargo a la cuenta corriente), seguros, créditos de consumo, etc.

## 2. Justificación

El planteamiento del problema nace del reconocimiento actual de mejorar la toma de decisiones de fidelización, debido a que se desean realizar medidas más eficientes de tiempo, recursos y horas-hombre dedicadas. El presupuesto disponible son poco más de mil millones de pesos por lo que es importante utilizarlo lo mejor posible.

El banco actualmente segmenta en siete grupos a los clientes, basándose en la renta “actual” declarada. En la tabla 2. se puede observar la etiqueta de cada segmento.

Segmentación Banco	Tramo Renta (en miles)
Joven	<400
Clásico-Activo	400-500
Clásico	500-650
Personal	650-1300
Selecto	1300-2200
Preferencial	2200-4000
Privado	4000>

Tabla 2. Segmentación de clientes según Renta. Fuente: Banco

Junto con esto, las acciones de marketing se realizan en base a la contribución neta acumulada anual de los clientes. La Contribución Neta se calcula mensualmente y es la resta entre los ingresos y los gastos totales de cada individuo. Este valor se obtiene directamente de la base de datos de la compañía.

Se considera que las acciones basadas únicamente en la contribución neta acumulada son “miopes”, es decir, dado que se basa en información pasada o actual,

no observa las posibles contribuciones futuras (tanto positivas como negativas), que pueda tener un persona.

La tabla 3. se presenta para mostrar cómo la variable renta no es un indicador suficiente del valor acumulado de un cliente (datos de 700 mil registros). La renta no es un predictor único del valor del usuario. Se puede observar cómo varía la contribución según el Estado Civil del individuo.

Segmentación Banco	Soltero	Casado	Separado	Separado y Casado	Divorciado	Viudo
Joven	1.916	37.329	141.172		21.009	
Clásico-Activo	82.801	95.247	21.402	-9.254	43.442	107.351
Clásico	106.571	99.405	93.402	169.201	78.516	116.892
Personal	126.438	180.923	119.392	118.115	136.853	129.459
Selecto	167.143	192.422	228.613	149.170	182.731	254.969
Preferencial	238.023	246.704	278.462	560.807	272.311	340.525
Privado	848.992	914.685	697.940	836.575	765.708	1.050.325

Tabla 3. Contribución Acumulada (suma) de los primeros 5 meses del año 2012 de los cuenta-correntistas del Banco (714.000), según Segmento y Estado Civil. Fuente: Banco.

Esto se convierte en una motivación para buscar otro medio de segmentar a los clientes por valor.

#### A. Quintiles

Se crearon dos grupos de quintiles de valor para comprender el movimiento del valor de los clientes del banco en el plazo de estudio. Un quintil corresponde a un quinto de una población ordenada de menor a mayor valor. Sin embargo, en este caso el ordenamiento es de mayor a menor y se realiza sobre el valor (contribuciones) del cliente. Así, dado el tamaño de la población (17.425 registros), cada quintil está conformado 3.485 individuos.

Uno de los grupos de quintiles creados se llama “Quintil Inicial” y se basa en el valor que entrega el cliente en sus primeros tres meses de relación con el banco (“Valor inicial”). La otra agrupación de quintiles se llama “Quintil Final” y corresponde a quintiles calculados según el valor de vida en un plazo máximo de 48 meses (“Valor Final”). En la tabla 4 se muestran la forma de computar el valor para cada caso.

<b>Valor Inicial</b>	$Contribución_1 + \frac{Contribución_2}{(1 + \delta)^1} + \frac{Contribución_3}{(1 + \delta)^2}$
<b>Valor Final</b>	$\sum_t^T \frac{Contribución Neta_t}{(1 + \delta)^t}$

Tabla 4. Fórmulas de cálculo de las denominaciones de Valor Inicial (Contribución de los primeros tres meses en el banco) y Valor Final (horizonte máximo de 48 meses), de los clientes de estudio.

Para ambas separaciones (Inicial y Final), el Quintil 1 corresponde a los clientes que se encuentran dentro del 20% de los individuos de mayor valor. En contraposición, el Quintil 5 indica a las personas que se encuentran dentro del 20% de peor valor (Ver Anexo C).

La tabla 5. presenta el movimiento de los clientes desde un Quintil Inicial a uno Final. Se observa que el 32% de los individuos que en sus primeros tres meses en la empresa se encuentran dentro del grupo de mayor valor, a los 4 años continúan siendo los clientes más valiosos.

A la vez, se puede observar cómo individuos que inicialmente no tienen tan alto valor (Quintil Inicial 5, 4 y 3), terminan integrando el grupo de los clientes de mayor valor (Quintil Final 1 y 2). Esto constituye al menos el 30% de las personas que componen cada uno de esos quintiles iniciales.

A raíz de esto se observa, que si las decisiones se tomaran únicamente con la información del valor acumulado actual de los clientes, se omitiría el hecho de que por ejemplo, el 23% de los clientes que inicialmente pertenecen al segundo quintil de mayor valor (tabla 5), disminuyen su valor final pasando al segundo peor quintil de valor, perdiendo esfuerzos en ellos.

Participación en Quintil Final dado el Quintil Inicial		Quintil Final					Total
		1	2	3	4	5	
Quintil Inicial	1	32%	24%	19%	12%	14%	100%
	2	19%	20%	25%	28%	7%	100%
	3	20%	20%	21%	23%	16%	100%
	4	13%	20%	20%	20%	27%	100%
	5	15%	16%	16%	16%	37%	100%

Tabla 5. Matriz de Quintiles de Valor Inicial y Final de los que presenta cambio de valor de cada Quintil Inicial al finalizar su relación con el banco.

Al analizar el valor total que genera cada grupo en la tabla 6, se descubre que existe un grupo de clientes que comienza perteneciendo al mejor Quintil de valor y a su salida genera las mayores pérdidas para la empresa (grupo inicial 1 al grupo final 5).

Valor Grupo		Al Término de Relación con el Banco				
		1	2	3	4	5
Primeros 3 meses	1	3.101.245.190	707.061.085	220.021.428	16.344.405	-1.195.624.709
	2	1.535.272.990	571.410.862	288.785.316	19.273.096	-339.435.928
	3	1.592.545.430	599.732.997	239.224.404	-1.638.124	-831.320.586
	4	919.382.287	602.031.252	223.572.248	-3.620.491	-597.446.349
	5	1.193.619.760	465.462.936	185.186.825	3.554.743	-1.045.803.229

Tabla 6. Matriz de Quintiles que representa el Valor Total de cada Grupo Inicial-Grupo Final.

Si las decisión de retención se basara únicamente en la contribución que generan las personas en sus primeros tres meses en el banco, entonces, se permitiría que los individuos de los grupos de montos iniciales bajos y finales altos (bloques celestes) se fugaran. Esto provocaría que dejara de percibirse un 34% de las ganancias (no se consideraron las pérdidas) que se tendrían en un periodo de 48 meses.

Cuando se actúa en base a esto, pueden estar ocurriendo dos casos de interés que se están perdiendo de vista: clientes con características homogéneas que están entregando distintas rentabilidades por diferentes motivos o clientes con similar contribución acumulada actual, pero que tienen distinto valor futuro y están siendo manejados de la misma forma pese a que el valor de uno es más atractivo que el del otro.

Estos clientes se podrían manejar de manera diferente y los esfuerzos de marketing podrían ser mejor focalizados. Por ejemplo, para el primer caso, los clientes podrían estar teniendo a disposición distintos productos, y se le podría proponer un up-selling a uno de los clientes. En el segundo caso, se debieran hacer mayores esfuerzos de fidelización al de mayor valor futuro.

## B. Plan de Toques

Como se menciona en el punto II.1.B., el plan de toques consiste en realizar promociones para la fidelización del cliente. En la tabla 7. se observa la frecuencia de clientes que se mueven de un Quintil Inicial a un Quintil Final dados.

Se ve que caen menos individuos del Quintil Inicial 2 al Quintil Final 5 con respecto al Quintil Inicial 1. Los del Quintil I. 2 caen en mayor proporción al Quintil F. 4. Un cliente que parte muy bien en cuanto a su valor, puede llegar a caer más fuertemente que uno comienza un valor un poco más bajo.

Quintil Inicial\Quintil Final	1	2	3	4	5
1	1.129	823	649	413	471
2	679	688	873	989	256
3	692	708	721	817	546
4	457	710	685	705	928
5	528	556	556	561	1.284

Tabla 7. Cantidad de clientes por grupo.

En la tabla 8 se tienen la cantidad de toques que se realizaron por cada Quintil Inicial/Final (tabla en porcentaje en Anexo D) el año siguiente a su ingreso. Se realizaron más toques sobre los clientes de peor valor que sobre los de valor regular. Solamente el 34% de los clientes del Quintil I. 4 “tocados” pasaron al Quintil F. 1. Un 18% de los clientes del segundo quintil inicial “tocados” dejaron de entregar valor (quintil 4 tiene Valores finales cercanos a cero). Tal vez para estos grupos se debe cambiar la forma de hacerles promoción.

Quintil Inicial\Quintil Final	1	2	3	4	5	Total
1	56	31	17	5	1	110
2	30	17	19	15	1	82
3	21	8	6	3	2	40
4	11	6	6	1	4	28
5	32	7	6	3	1	49

Tabla 8. Toques realizados para cada grupo.

El banco acierta en un alto porcentaje a los clientes valiosos al realizar las campañas de fidelización. Sin embargo, existen un porcentaje de personas que generan grandes pérdidas y que es preferible no retenerlos. Estos constituyen un 15% de los clientes “tocados” (no se contabiliza la intensidad del toque, solamente si el cliente *fue tocado*). Este porcentaje de pérdida corresponde a poco más de \$170 millones de pesos.

Los invitados a los eventos del Plan de Toques, no necesariamente son los clientes más valiosos a futuro. La misma idea ocurre con la captación de los jóvenes universitarios, que, a pesar de que pueden poseer un alto valor futuro, poco se hace en la fidelización de estos individuos. Por ejemplo, en el Call Center de Servicio al Cliente, la Banca Joven se encuentra como última prioridad de atención, siendo que pueden existir jóvenes que a futuro tendrán una alta rentabilidad.

#### o Efectividad del plan de toques

Con respecto al funcionamiento de las actividades de toque, se tiene que, gracias a esto, al menos el 37% de clientes de cada Quintil Inicial pasan al Quintil Final 1. Para evaluarlo, calcularon las probabilidades condicionales con respecto a utilizar y no utilizar “toques”:

$$\mathbb{P}[\text{Estado}_{\text{inicial}} \text{ a } \text{Estado}_{\text{final}} | \text{Sin toques}] = \frac{\mathbb{P}[\text{Sin toques} | \text{Edo}_{\text{inicial}} \text{ a } \text{Edo}_{\text{final}}] \cdot \mathbb{P}[\text{Edo}_{\text{inicial}} \text{ a } \text{Edo}_{\text{final}}]}{\mathbb{P}[\text{Sin toques}]}$$

(dado que se está en un Quintil Inicial específico)

$$\mathbb{P}[\text{Estado}_{\text{inicial}} \text{ a } \text{Estado}_{\text{final}} | \text{Con toques}] = \frac{\mathbb{P}[\text{Con toques} | \text{Edo}_{\text{inicial}} \text{ a } \text{Edo}_{\text{final}}] \cdot \mathbb{P}[\text{Edo}_{\text{inicial}} \text{ a } \text{Edo}_{\text{final}}]}{\mathbb{P}[\text{Con toques}]}$$

(dado que se está en un Quintil Inicial específico)

Los siguientes números se obtuvieron calculando la probabilidad de pasar de un Quintil Inicial (dado que se está en ese quintil), a un Quintil Final dado que se realizan/no se realizan toques. Los resultados son:

Sin Toques	1	2	3	4	5
1	32%	23%	19%	12%	14%
2	19%	20%	25%	29%	7%
3	19%	20%	21%	24%	16%
4	13%	20%	20%	20%	27%
5	14%	16%	16%	16%	37%

Tabla 9. Probabilidad que individuo dado que está en un Quintil Inicial, se mueva a un Quintil Final dado que no se realizan toques.

Con Toques	1	2	3	4	5
1	51%	1%	0%	0%	0%
2	37%	0%	1%	0%	0%
3	53%	0%	0%	0%	0%
4	39%	0%	0%	0%	0%
5	65%	0%	0%	0%	0%

Tabla 10. Probabilidad que individuo dado que está en un Quintil Inicial, se mueva a un Quintil Final dado que se realizan toques.

A partir de esto, se comprueba la importancia de las acciones de fidelización, pero a los clientes correctos. Es necesario identificar el valor que puede entregar la persona a la compañía.

La intención de la memoria es apoyar medidas de fidelización que en primera instancia son más rentables, más que de retención del cliente cuando desea abandonar, pues en la primera se tiene más tiempo para reaccionar de mejor manera tanto para el cliente, como para la compañía.

Para lograr esto, se busca calcular un indicador como el *Customer Lifetime Value* (CLV), en español, Valor del Tiempo de Vida del cliente, y convertir esa información en acciones.

Específicamente, en esta investigación se desean realizar propuestas de fidelización (promoción) para los clientes o segmento de clientes nuevos más valiosos.

### III. OBJETIVOS

#### 1. OBJETIVO GENERAL

- Estimar el *Customer Lifetime Value* de clientes de un banco usando variables transaccionales y socio-demográficas para focalizar las acciones de fidelización, específicamente de promoción.

## 2. OBJETIVOS ESPECÍFICOS

- Identificar las variables transaccionales y demográficas que se incluirán en la estimación del valor del cliente.
- Estimar la Renta de los clientes.
- Desarrollar un modelo de estimación del *CLV* de cada individuo.
- Calcular el *CLV* a nivel individual y validarlo.
- Identificar las características de los clientes del segmento más valioso.
- Proponer acciones de promoción para la retención del segmento más rentable.

## IV. ALCANCES

Este trabajo considera las siguientes limitaciones:

- La estimación utilizada para el cálculo del *CLV* será aplicada a los clientes cuenta-correntistas nuevos (de enero a junio del 2008), de un banco particular, y se realizará a nivel individual. Por lo que los resultados serán válidos solo para este caso de estudio.
- El horizonte de estudio será de no más de 48 meses, que es la data disponible.
- La información utilizada será transaccional y socio-demográfica y los registros serán de clientes a nivel nacional.
- No se levantarán nuevos datos no transaccionales.
- Clientes con datos faltantes serán imputados o eliminados.
- Se realizarán propuestas de fidelización únicamente sobre el mix de marketing de promoción y sobre los clientes más valiosos.

## V. RESULTADOS ESPERADOS

Los resultados esperados asociados a los objetivos específicos son:

- Evaluación de variables transaccionales y socio-demográficas en la estimación del *CLV*.
- Desarrollo de Modelos para predecir la Renta y el Valor Presente de los clientes con cuenta corriente del banco.
- Obtención del *CLV* de la base de clientes de estudio.

- Acciones de promoción de fidelización para el segmento más valioso de clientes.

## VI. MARCO TEÓRICO

En el Marco Teórico se explican detalladamente en qué consiste el *Customer Lifetime Value*, el Modelo a utilizar y los instrumentos de análisis de los datos.

### 1. CUSTOMER LIFETIME VALUE

El concepto de *CLV* fue introducido por primera vez hace más de treinta años por Philip Kotler. Es una fórmula que mide el valor de un cliente de una empresa. La fórmula es la siguiente:

$$CLV = \sum_{t \in T} \frac{(p_t - c_t) \cdot r(t)}{(1 + \delta)^t}$$

Ilustración 7. *Customer Lifetime Value*.  
Valor presente de los flujos futuros que puede entregar un cliente.

Donde  $(p_t - c_t)$  consiste en el margen que entrega el cliente en el tiempo  $t$ ,  $r(t)$  a la tasa de retención del cliente en  $t$ , y delta  $\delta$  a la tasa de descuento. Si a esta ecuación se le incluye el costo de adquisición del cliente, se estará hablando del *Equity* del cliente [18].

La realización de este cálculo para toda la cartera de clientes de una empresa, sirve para clasificarlos sobre su aporte a la utilidad de la compañía. Esta identificación sirve para entregar el servicio en función de la contribución de cada cliente, diferenciando. Con esto, se evita incurrir en ineficiencias de uso de los recursos.

La valoración de los clientes puede servir para [19]:

- Formular e implementar estrategias específicas de los clientes. Esta acción se realiza con el fin de maximizar las ganancias que se obtienen a lo largo de la vida del individuo, y aumentar su tiempo de permanencia  $T$ .
- Saber cuánto se puede invertir en conservar al cliente con el fin de obtener retorno sobre la inversión. Los recursos son limitados, por lo que se prefiere invertir en aquellos clientes que traerán la máxima rentabilidad a la empresa. Esto sólo es posible calculando el *CLV*.

En la literatura existen variadas metodologías para el cálculo del *Customer Lifetime Value*, gran porcentaje están aplicados a datos de retail. Específicamente en los Trabajos de Título, se tiene que algunos son enfocados al modelo RFM [15] y modificaciones de este adicionando el uso de variables demográficas [11]. Otros emplean metodologías que consideran la estimación del monto promedio a gastar por transacción con Modelo Jerárquico bayesiano [12] o Modelos Estructurales con Clase Latente [25]. También, existen modelos de migración con Cadenas de Markov de primer orden [18].

En [25] se estima el valor de vida de los clientes, pero se enfoca en la obtención del nivel de inversión óptimo que se destina a la retención. El proyecto utiliza programación dinámica a partir de la definición de estados del cliente en su relación con la empresa (Cadenas de Markov), y considera como inputs variables de satisfacción.

#### A. Horizonte de Vida

Existen diferentes variaciones del *CLV* según el horizonte de tiempo. Unos sugieren considerar un tiempo arbitrario, o tiempo esperado del cliente [35]. Mientras que en [36] proponen que el tiempo sea considerado infinito pues argumentan que el horizonte de tiempo limitado sobre-estima el *CLV*.

Para el caso del modelamiento se supondrá el tiempo de vida del cliente igual al tiempo promedio, cuyo valor es de 15 años aproximadamente. El análisis considerará al cliente como lost-for-good (perdido-para-bien), en donde se supone que el cliente abandona la empresa de manera definitiva. En caso de que vuelva a contratar servicios, se considerará como un cliente nuevo [21]. No ocurre lo mismo en el caso de considerar al cliente como always-a-share, utilizado comúnmente Cadenas de Markov, donde el cliente solamente va cambiando de estado [29].

#### B. Cross-selling

La posibilidad de elección del cliente entre distintos productos es un tema no menor dependiendo de la profundidad que se desee entregar al modelo. Se puede modelar el servicio como uno solo, agregado, hasta representar las inter-dependencias entre los productos. Para lo último, se podría estimar la probabilidad de elección de cada producto o definir la secuencia en la que consumirán los productos (que dependerán de la evolución de las necesidades del individuo durante su relación con la compañía).

Estos temas se tratan en [30], en donde se asigna una decisión de compra binaria de cada uno de los productos y se considera la evolución de la demanda del cliente en el tiempo, utilizando Probit Multivariado. Anteriormente a esto, se ha planteado un modelo secuencial de tiempo de adopción de un nuevo producto [31], y un modelo en el que se determina qué producto es más probable de comprar y cuándo [32].

#### C. Tasa de Retención

La tasa de retención es otro factor que afecta el *CLV* del cliente, se debe tener cuidado al utilizar, ya que [22] muestran que cambios en las tasas de retención pueden dar lugar a grandes cambios en el valor del cliente.

Además que esta tasa tiene variados elementos que pueden hacer que se dude en aplicar un mismo modelo de predicción para cohorte distinto, ya que puede haber

diferencias sistemáticas en su comportamiento en comparación con los grupos de otras características sociodemográficas (es decir, efectos de cohorte transversal).

A continuación se enumeran los factores que influyen en la gran disparidad de la tasa de retención entre cohortes: (1) dependencia de la duración (la tasa de abandono decrece a medida que la tenencia del suscriptor aumenta, es decir, se tiene una dependencia en la duración negativa (Reichheld 1996; Hughes 2006), (2) la actividad de marketing es variable en el tiempo (por ejemplo, efectos de promociones), (3) la heterogeneidad del contratante, (4) los efectos de cohorte transversal (características sociodemográficas), y (5) los efectos tiempo calendario.

## 2. Modelo de regresión lineal

Regresión lineal permite modelar las interacciones entre variables. Actúa minimizando los errores cuadrados que se calculan al restar el valor estimado con el valor real.

$$y = \alpha + \beta x + \varepsilon$$

En la ecuación anterior,  $y$  representa la variable a explicar,  $x$  la variable que tiene una relación con  $y$ ,  $\beta$  el estimador que expresa la relación directa entre  $x$  e  $y$ ,  $\alpha$  el intercepto y  $\varepsilon$  el error que se comete al definir  $y$  como una combinación de  $x$ .

## 3. Calibración y validación

La calibración y validación del modelo se hará con tres muestras de datos distintas. La primera para estimar los parámetros, la segunda para testarlos y evitar el ajuste a la primera, y la última, debido a que se vuelve a minimizar el error con respecto a la segunda muestra, para probar si los datos se ajustaron a la segunda muestra, se vuelve a validar con una muestra nueva.

## 4. Testeo

### A. Tests de análisis exploratorio

Los tests de análisis que serán usados son el de Anova y regresión univariada, para identificar si las variables de estudio explican varianza (entregan información).

#### ○ Tests Anova

El Test Anova (Análisis de Varianza), se utiliza para conocer si dos variables son independientes de forma estadísticamente significativa. El método que utiliza es el contraste entre medias de dos o más poblaciones. Se presenta como:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_n$$
$$H_1: \text{al menos una es distinta}$$

Donde  $\mu_i$  es la media de cada población.

Para rechazar la hipótesis nula  $H_0$ , tiene que ocurrir que la variabilidad entre grupos debe ser significativamente grande respecto a la intra-grupos. La regla de decisión consiste en rechazar  $H_0$  si el nivel de significancia es menor a cierto número. En este caso se considerará 0,05 (Nivel de confianza del 95%).

## B. Tests de predicción

Se utilizan para revisar el desempeño de los modelos, compararlos entre ellos y decidir cuál explica mejor los datos [26].

### ○ $R^2$ Ajustado

Es una medida de la calidad global del conjunto de variables explicativas utilizadas. Se llama ajustado porque castiga por la cantidad de parámetros (o grados de libertad), tratando de favorecer la parsimonia.

$$R^2 = 1 - \frac{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - k - 1}}{\frac{\sum_{i=1}^n (y_i - \bar{Y})^2}{n - 1}}$$

Donde  $y_i$  es la variable dependiente del registro  $i$ ,  $\hat{y}_i$  es la estimación de la variable endógena,  $\bar{Y}$  es el promedio de los  $y_i$ ,  $n$  el total de registros y  $k$  el número de parámetros estimados.

### ○ Media Cuadrática (RMS)

Es una medida que permite evaluar las diferencias existentes entre el valor real y el estimado. Éste consiste en el cálculo de la raíz cuadrada del promedio de los errores de estimación al cuadrado.

$$RMS = \frac{100\%}{N} \sum_t (Valor Real_t - Valor Estimado_t)^2$$

### ○ Error porcentual absoluto medio (MAPE)

Es una medida de la precisión de un método. Se usa para la evaluación individual de cada estimación.

$$MAPE = \frac{100\%}{N} \sum_t \left| \frac{Valor Real_t - Valor Estimado_t}{Valor Real_t} \right|$$

- *Error porcentual absoluto ponderado (WAPE)*

Es una modificación del MAPE, trata de salvar los altos valores de desviación causados por valores de predicción pequeños. La idea es entrega la importancia adecuada a cada valor a estimar.

$$WAPE = 100\% \frac{\sum_i^n \left| \frac{Valor Real_i - Valor Estimado_i}{Valor Real_i} \right| \cdot Valor Real_i}{\sum_i^n Valor Real_i}$$

## 5. Modelo Gamma-Weibull

La función Gamma-Weibull se utiliza en el modelamiento de la duración de un evento en tiempo continuo, con dependencia en la duración. Esto último quiere decir que la tasa de ocurrencia del evento no es uniforme en el tiempo.

Esta tasa de ocurrencia se llama *Hazard Rate* (Tasa de Riesgo) y es la probabilidad de que un evento de estudio ocurra ahora, dado que aún no ha ocurrido. Su fórmula es:

$$h(t) = \frac{f(t)}{1 - F(t)}$$

Donde  $f(t)$  es la función de densidad de probabilidad y  $F(t)$  la función de distribución acumulada de una variable aleatoria no negativa, que en este caso es el tiempo de ocurrencia.

La función de riesgo más utilizada es la Weibull, descrita como:

$$h(t) = c\lambda t^{c-1} \quad F(t) = 1 - e^{-\lambda t^c}$$

Donde  $\lambda$  representa la escala de la función y  $c$  la forma. Cuando  $c$  es igual a 1, la distribución Weibull se convierte en una Exponencial.

Una forma de introducir heterogeneidad no observable al modelo de duración es definiendo una distribución para  $\lambda$ . Se puede asumir que esta tasa se distribuye Gamma en la población.

$$g(\lambda) = \frac{\alpha^r \lambda^{r-1} e^{-\alpha\lambda}}{\Gamma(r)}$$

Con  $\alpha$  parámetro de escala y  $r$  de forma.

Así, la distribución acumulada del tiempo de ocurrencia de un evento es [39]:

$$\mathbb{P}(T \leq t) = \int_0^{\infty} (1 - e^{-\lambda t^c}) \frac{\alpha^r \lambda^{r-1} e^{-\alpha \lambda}}{\Gamma(r)} d\lambda$$

$$\mathbb{P}(T \leq t) = 1 - \left( \frac{\alpha}{\alpha + t^c} \right)^r$$

## VII. DESARROLLO METODOLÓGICO

### 1. Explicación de estimación de CLV (Valor Presente VP)

Se desea estimar el CLV de un periodo 4 años de los clientes que ingresaron entre enero y junio del 2008. Para esto, primero se calcula el CLV real o VP real (como se llamará en adelante).

Para esto, se desarrollan tres modelos con los que evalúan dos formas de predicción del VP. Una de ellas (modelo sin duración), consiste en desarrollar una regresión con el VP como variable endógena. La otra (modelo con duración), consiste en estimar el VP promedio (VP dividido por la duración) y la duración de los clientes de manera separada, para luego multiplicarlas y obtener un valor predicho del VP.

El segundo método nace con el objeto de levantar el supuesto de tasa de retención constante e igual a uno del primero. Como se verá a continuación, el cálculo del VP real no incluye la tasa de retención.

Una vez calibrado el modelo, la validación de la regresión se realiza con una muestra de datos nueva, esto es, con un grupo de registros distintos y seleccionados aleatoriamente.

#### ○ Cálculo del VP real

El cálculo del VP real consiste en traer a valor presente la contribución neta que genera el cliente mensualmente (Ver ecuación de VP real), descontado a una tasa anual del 12%. Como se menciona anteriormente, la contribución neta es un dato ya calculado por el banco, disponible en la base de datos, por lo que no se presenta ninguna fórmula de su cálculo.

$$VP \text{ real} = \sum_{t \in 48 \text{ meses}} \frac{\text{ContribuciónNeta}_t}{(1 + 0,00948)^t}$$

El periodo de estudio es de 48 meses, por lo que según esta fórmula, si se tiene un cliente que cierra su cuenta corriente en el mes 24, sus contribuciones a partir del mes 25 hasta el 48 serán cero y no sumaran ningún valor a la fórmula.

- *Individuos de Estudio*

Para el modelamiento, se unen en mismo grupo de individuos los clientes que ingresaron en los meses de estudio. Se consideran como si hubieran ingresado en el mismo mes (Ver Ilustración 8). La duración de corto plazo no es diferente entre grupos de entrada, por lo que finalmente no se considera en el modelamiento la variable que los diferencia por su mes de ingreso (Ver VIII.3.C.).

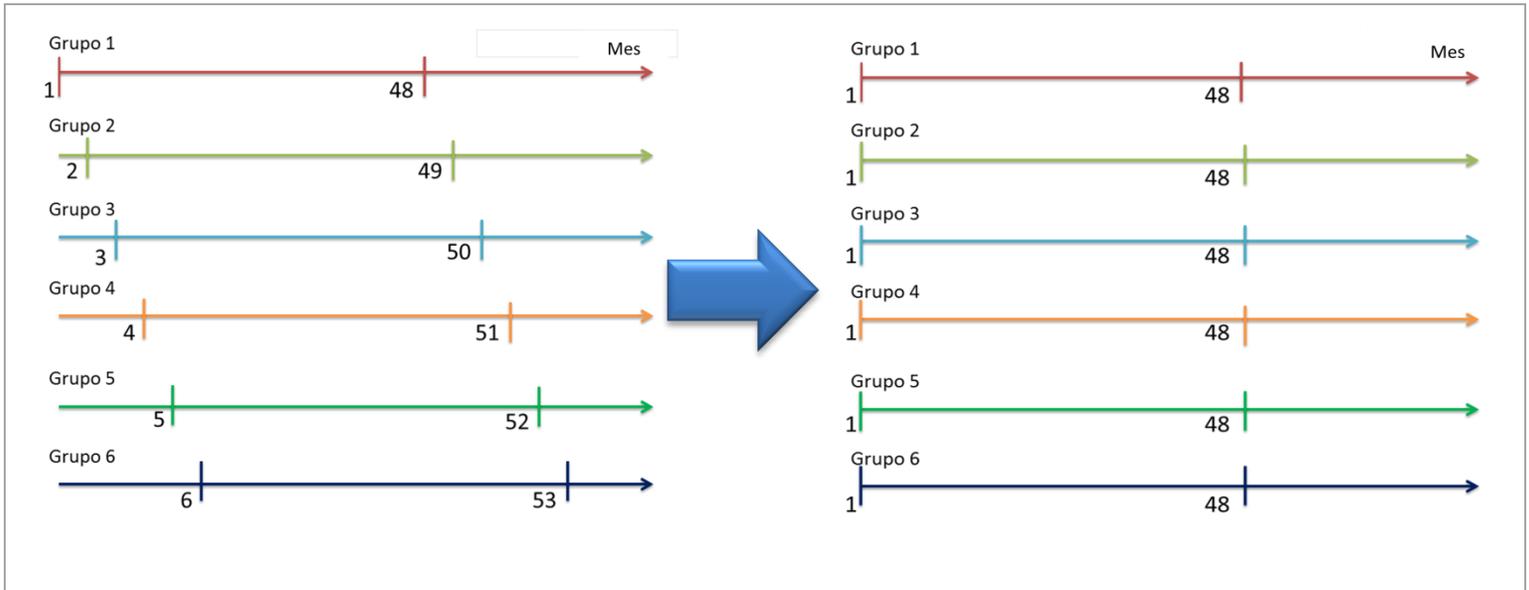


Ilustración 8. Gráficas que describen la agrupación de cada grupo de entrada en uno grande. No considera como información el mes de entrada los clientes.

Los tres modelos estiman la variable dependiente con variables transaccionales: saldos, deudas, contribuciones, y variables socio-demográficas: género, edad, estado civil, ocupación, etc. Este paso se realiza sobre una muestra de datos seleccionados al azar.

- *Validación de los datos*

Para la validación de los datos, se actúa de manera distinta a otras memorias donde la validación se desempeña sobre los mismos registros, en un periodo de tiempo posterior. En este caso, la data se particiona en porcentajes, en este caso de 80%-20%, de tal modo de evitar sobre-ajuste de los datos.

## 2. Selección de Datos

Se seleccionan los datos transaccionales y socio-demográficos de un grupo de clientes de un banco: individuos que hayan iniciado su relación con el banco con la apertura de una cuenta corriente.

Las variables disponibles son:

Socio-demográfica	Transaccional
Edad (años)	Tenencia de Productos <sup>1</sup>
Renta (pesos)	Deudas <sup>1</sup>
Estado Civil (dummy)	Saldos <sup>1</sup>
Ocupación (dummy)	Contribuciones: captación, colocación, gastos, filiales, provisiones.
Género (dummy)	Duración
Comuna (dummy)	<sup>1</sup> Crédito de Consumo (pesos), Hipotecario (UF), Línea de Crédito (pesos), Tarjeta de crédito (pesos).
Ciudad (dummy)	
Región (dummy)	

Ilustración 9. Variables a utilizar en el modelamiento.

	valor	Definición
Contribución Captación	+	Contribución obtenida de los productos de colocación: tarjeta de crédito, crédito de consumo, hipotecario, etc.
Contribución Colocación	+	Contribución obtenida de los productos de captación: depósito a plazo, cuenta de ahorro, fondos mutuos, etc.
Contribución Comisiones	+	Ganancias obtenidas por comisiones.
Contribución Filiales	+	Contribución proveniente de otros bancos de la empresa.
Contribución Gastos	-	Gastos en que se incurren por mantener al cliente.
Contribución Provisiones	-	Provisión (dinero almacenado) por posible no pago del cliente.

Ilustración 10. Definición de contribuciones.

### 3. Pre-procesamiento y transformación

Esta etapa consiste en pre-procesar los datos, eliminando registros y variables que no entreguen información. Además, eliminar outliers y realizar imputación de datos. Con esto se busca dejar los datos con el menor ruido posible, a fin de evitar las distorsiones en los resultados de los modelos.

Las tareas involucran:

- Eliminar registros cuyo valor de la variable antigüedad sea mayor a la edad.
- Seleccionar únicamente registros de clientes cuyo primer servicio con el banco haya sido la apertura de una cuenta corriente.
- Seleccionar individuos que entran y salen una vez de la empresa (*lost-for-good*).
- Seleccionar los clientes que iniciaron su cuenta corriente entre los meses de enero y junio del 2008.

También, se ejecutan las transformaciones en los datos para poder trabajar con ellos, generar nueva información y reducir dimensionalidad. Esto requiere:

- Calcular la edad a partir de la fecha de nacimiento.
- Obtener el tiempo de pertenencia del cliente en la empresa a partir de la tenencia de los productos.
- Convertir variables categóricas a dummies.
- Calcular ratios para evaluar la aversión al riesgo del individuo, o características intrínsecas a su conducta financiera. Se calcularon los siguientes ratios:
  - Promedio de las contribuciones de los primeros tres meses.
  - Ratios entre Saldos de productos de colocación: Tarjeta de crédito (TDC) y Línea de Crédito (LDC).

$$\text{Ratio}_1 = \frac{\text{Monto producto captación}_1}{\text{Monto producto captación}_2}$$

- Variaciones entre Contribuciones de los primeros tres meses: Mes 1-Mes 2, Mes 1-Mes 3, Mes 2-Mes 3.

$$\text{Ratio}_2 = \frac{\text{Contribución}_{\text{MES 3}} - \text{Contribución}_{\text{MES 1}}}{\text{Contribución}_{\text{MES 3}}}$$

- Ratios entre variables transaccionales (saldos, deudas, montos en cuenta) vs. Renta (la idea de este último es (como lo sugiere el conocimiento de expertos), evaluar el rasgo: aversión al riesgo).

$$\text{Ratio}_3 = \frac{\text{Deuda}}{\text{Renta}} \quad \text{Ratio}_4 = \frac{\text{Deuda}}{\text{Saldo}}$$

#### 4. Análisis Exploratorio

En este punto se realiza un análisis exploratorio de los datos para identificar relaciones entre variables y seleccionar las que sean significativas para los modelos a emplear.

El análisis contempla los siguientes pasos:

- Analizar correlaciones entre variables de estudio.
- Histograma de clientes según variable antigüedad.
- Histograma de clientes según variable edad.

- Analizar relación entre Margen colocación o captación con Ratios de saldos, montos y deudas.

## 5. Modelamiento

Se desarrollan los modelos propuestos y se validan. Los dos formas de modelamiento son:

1. *Modelo de Sin duración*: Regresión Lineal sobre la variable dependiente, Valor Presente (CLV), calculada de antemano con una tasa de descuento del 12% anual, según indicación del banco.

$$VP = \beta_0 + \beta_1 \cdot \text{género} + \beta_2 \cdot \text{contribución}_{\text{colocaciónM3}} + \beta_3 \cdot \text{tenencia}_{\text{consumoM3}} + \dots$$

2. *Modelo de Con duración*: Estimación de dos modelos con Regresión Lineal: uno de duración y otro de predicción del Valor Promedio del individuo. Finalmente se multiplican las dos predicciones para obtener el Valor Presente total.

$$VP_{\text{prom}} = \beta_0 + \beta_1 \cdot \text{género} + \beta_2 \cdot \text{contribución}_{\text{colocaciónM3}} + \beta_3 \cdot \text{tenencia}_{\text{consumoM3}} + \dots$$

$$\text{duración} = \beta_0 + \beta_1 \cdot \text{género} + \beta_2 \cdot \text{comuna} + \beta_3 \cdot \text{tenencia}_{\text{consumoM3}} + \dots$$

$$VP = VP_{\text{prom}} \cdot \text{duración}$$

## 6. Evaluación

El método de evaluación será distinto al utilizado en memorias como [15], en que se particiona el tiempo: un periodo se utiliza para calibrar el modelo y un periodo de tiempo posterior para validarlo.

En este caso, se seleccionan aleatoriamente tres muestras de datos y se corre el modelo sobre dos de ellos, con el fin de que no se sobre-ajuste el modelo a esos datos.

## VIII. APLICACIÓN DE LA METODOLOGÍA

Los datos utilizados corresponden al de clientes iniciaron su vida en la empresa con un contrato de cuenta corriente. Se tiene información de los saldos, deudas y montos de los distintos productos para un periodo de 53 meses: desde enero del 2008 a mayo del 2012.

## 1. Selección de Datos

Se seleccionan los datos transaccionales y socio-demográficos de un grupo de clientes de un banco: individuos que hayan iniciado su relación con el banco con la apertura de una cuenta corriente, entran y salen una sola vez (*lost-for-good*) e ingresaron el año 2008 entre los meses de enero y junio.

Inicialmente se tenían 714.500 registros de cuenta-correntistas potenciales para trabajar (clientes con número de rut mayores a tres millones y menores a 50 millones). Se realizaron los filtros mencionados en el párrafo anterior, quedando únicamente 18.240 clientes. Finalmente, al eliminar los outliers, se concluyó con una muestra de 17.425 registros.

Los clientes de estudio son individuos de todo el país. Se dispone de las variables comuna, ciudad y región. Los municipios fueron agrupados por Zonas. Esta agrupación se basó en un estudio de Nivel Socio-Económico de Adimark y cuya asignación fue entregada por el Banco.

A continuación se presentan las Zonas y su porcentaje sobre el total de los registros. Las comunas que integran cada Zona se encuentran en Anexo F. La etiqueta Z representa “Sin Información”. No se realizó una imputación de estos datos.

Etiqueta Zona	Comunas representativas	Frecuencia	Porcentaje
A	Antofagasta, Calama, Chuquicamata.	467	2,7%
B	Copiapó, La Serena, Coquimbo.	318	1,8%
C	Algarrobo, Santo Domingo, Con-Cón, Olmué.	1.185	6,8%
D	Angol, Chaitén, Osorno.	1.601	9,2%
E	Aisén, Punta Arenas, Coyhaique.	175	1,0%
F	Curacaví, Talagante, Colina.	409	2,3%
G	La Reina, Las Condes, Lo Barnechea, Vitacura.	899	5,2%
H	Ñuñoa, Pedro de Valdivia, Providencia.	715	4,1%
I	Peñalolén, Independencia, Huechuraba.	2.027	11,6%
J	Independencia, San Bernardo, Recoleta.	1.155	6,6%
K	Cerro Navia, El Bosque, Pedro Aguirre Cerda.	648	3,7%
Z	Sin Información.	7.825	44,9%
Total		17.424	100%

Tabla 11. Tabla de Frecuencia de Zonas por Comunas.

El género está repartido prácticamente entre 60%-40% de hombres y mujeres respectivamente.

Etiqueta	Género	Frecuencia	Porcentaje
1	Hombre	10.258	59,1%
2	Mujer	7.090	40,9%
	<b>Total</b>	<b>17.424</b>	<b>100%</b>

Tabla 12. Tabla de Frecuencia del género.

## 2. Pre-procesamiento y Transformación

Para trabajar con la edad, se hicieron tramos de ésta, agrupándola de a cinco años para los tramos intermedios y para el último tramo se escogió la edad en la que se hacía un cambio drástico en la dispersión del VP.

Etiqueta	Tramo Etario	Frecuencia	Porcentaje
1	- 30	3.114	17,9%
2	31 - 35	4.150	23,8%
3	36 - 40	3.288	18,9%
4	41 - 45	2.507	14,4%
5	46 - 50	1.775	10,2%
6	51 - 55	1.212	7,0%
7	56 - 63	934	5,4%
8	64 -	444	2,5%
	<b>Total</b>	<b>17.424</b>	<b>100%</b>

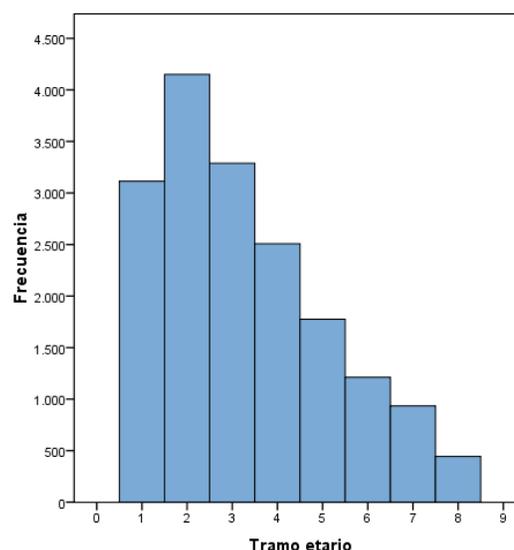


Tabla 13. Tabla de Frecuencia de tramos de edad.

También se procedió a imputar la variable estado civil (frecuencia en Anexo H). En los datos de estudio, existe un 43% de registros sin información de su estado civil, por lo que se procedió a estimarlo, manteniendo la proporción con el género y la lógica con la ocupación y edad. Para esto, se observó un gráfico box-plot y tabla de descriptiva de edad vs. estado civil (Anexo I), de donde se obtuvo la proporción del estado civil para distintos tramos de edades:

- Entre 23 y 27 años: No existían más que solteros
- Entre 28 y 36 años: El 33% se repartía entre soltero, casados y separados.
- Entre 37 y 52 años: Se repartían la mitad para casados y separados.
- 53 años en adelante: La división consistió en casados, separados y viudos.

A partir de esta separación, se reemplazó en cada nivel de estado civil de manera proporcional al género. Con esta separación se acrecentó la correlación del estado civil con la edad.

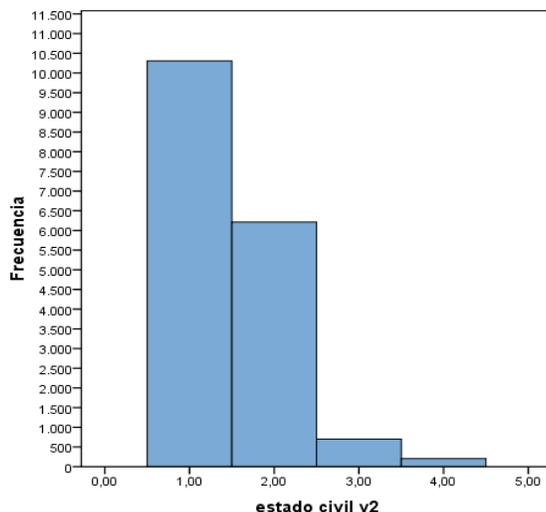


Ilustración 11. Tabla de Frecuencia de estado civil versión 2.

Para la ocupación se tienen un 43,6% de datos sin información pero, como se verá más adelante, no implica que esto no entregue información.

Etiqueta	Ocupación	Frecuencia	Porcentaje
0	Sin Información	7.590	43,6%
1	Alto ejecutivo - Propietario GE o ME	201	1,2%
2	Ejecutivo medio - Profesional Independiente carrera Tradicional	1.368	7,9%
3	Empleado administrativo medio y bajo - Prof. Indep. Técnico	6.470	37,1%
4	Obrero calificado	487	2,8%
5	Oficio menor-obrero no calificado	221	1,3%
6	Jubilado	172	1,0%
7	Sin Ocupación	915	5,3%
	<b>Total</b>	<b>17.424</b>	<b>100%</b>

Tabla 14. Tabla de Frecuencia de ocupación y asignación de etiquetas.

### 3. Análisis descriptivo

Como primera medida, se calcularon tablas de contingencia y correlaciones para identificar variables que puedan estar explicando la misma información y se pueda reducir la cantidad de variables a utilizar.

Se realizó un test de significancia y existe diferencia entre la edad y el género y la edad y el estado civil. Por lo que ambas se incluyen en el modelo (Ver Anexo G).

Tramo Etario	Estado civil					
	Sin inf.	Soltero	Casado	Separado	Viudo	
- 30	32%	64%	3%	0%	0%	100%
31 - 35	42%	46%	11%	1%	0%	100%
36 - 40	48%	29%	21%	2%	0%	100%
41 - 45	47%	21%	29%	4%	0%	100%
46 - 50	47%	15%	33%	5%	0%	100%
51 - 55	47%	14%	35%	4%	1%	100%
56 - 63	47%	12%	34%	6%	1%	100%
64 -	43%	10%	36%	2%	9%	100%

Tabla 15. Porcentaje de participación en estado civil por nivel de tramo etario.

Tramo Etario	Género			
	Sin inf.	Masculino	Femenino	
- 30	0%	55%	45%	100%
31 - 35	0%	57%	43%	100%
36 - 40	0%	61%	39%	100%
41 - 45	0%	60%	40%	100%
46 - 50	1%	63%	36%	100%
51 - 55	1%	61%	38%	100%
56 - 63	2%	57%	41%	100%
64 -	0%	62%	38%	100%

Tabla 16. Porcentaje de participación en género por nivel de tramo etario.

Tramo Etario	ocupación							
	Sin Inf.	Alto ejecutivo	Ejecutivo medio - Profesional Independiente carrera Tradicional	Empleado administrativo o medio y bajo	Obrero calificado	Oficio menor-obrero no calificado	Jubilado	Sin Ocupación
- 30	13%	4%	34%	18%	11%	4%	0%	44%
31 - 35	23%	8%	30%	26%	20%	10%	1%	14%
36 - 40	21%	15%	13%	20%	19%	18%	1%	8%
41 - 45	16%	16%	10%	15%	18%	17%	1%	8%
46 - 50	11%	18%	6%	10%	14%	19%	2%	9%
51 - 55	7%	13%	4%	6%	9%	13%	9%	7%
56 - 63	6%	15%	2%	4%	6%	11%	17%	6%
64 -	2%	9%	1%	1%	2%	9%	68%	3%
	100%	100%	100%	100%	100%	100%	100%	100%

Tabla 17. Participación en tramo etario por ocupación.

Se calculó además la correlación entre VP y renta, resultando significativa.

Coeficiente de Correlación de Pearson Sig. (2 colas)		
<b>vp vs. Renta</b>	<b>0,239</b>	<b>0,000</b>

Tabla 18. Correlaciones de variable CLV vs. Renta.

Las variables Región y Comuna entregan la misma información, por lo que se utilizará solamente una de ellas: comuna (agrupada por Zonas).

Tomando como referencia el Mes 3, se calculó la correlación del VP con las diferentes Contribuciones. Todos los estadísticos resultaron significativos.

VP vs.:	Coeficiente de Correlación de Pearson	Sig. (2 colas)
Contribución Neta Mes 3	<b>0,108</b>	<b>0,000</b>
Contribución Captación Mes 3	<b>0,070</b>	<b>0,000</b>
Contribución Colocación Mes 3	<b>0,041</b>	<b>0,000</b>
Contribución Comisiones Mes 3	<b>0,063</b>	<b>0,000</b>
Contribución Filiales 3	<b>0,148</b>	<b>0,000</b>
Contribución Gastos Mes 3	<b>0,037</b>	<b>0,000</b>
Contribución Provisiones Mes 3	<b>0,034</b>	<b>0,000</b>

Tabla 19. Correlación de VP con variables de Contribución del Mes 3.

Se realizó un Test Anova sobre el VP utilizando como factor las variables de tenencia de productos: Línea de Crédito, Tarjeta de Crédito, Consumo e Hipotecario. Estas generan de forma individual un efecto significativo en el Valor Presente. Los resultados de los test se encuentran en Anexo J.

## A. Valor Presente

- *Distribución del VP en la población*

Se tiene que el 30% de los individuos posee un valor presente negativo. Se piensa que las razones son: horizonte de tiempo de estudio corto o tasa de descuento muy alta.

- *Análisis Tasa de Descuento*

Con respecto a la tasa de descuento, se realizó una comparación utilizando tasas de descuento de 10% y 8% anual, convertidas a mensual (recordar que la tasa que se utiliza en la compañía es de un 12% anual). Los registros con VP menor a cero continúan rodeando el 30% y la variación del VP respecto de la tasa inicial, es de 5% para la tasa de 8% y de 3% para la de 10%. Se descarta que el bajo valor sea debido a este factor.

- *Duración*

Al considerar a los individuos que se mantienen por menos de 47 meses, el promedio de duración es de 21,9 meses, es decir, 1,8 años. Por lo que se piensa que esta es la principal razón del valor de vida negativos para algunos individuos, no se alcanza a recuperar la inversión inicial.

Para efectos de la memoria, se indicará si el análisis incluye a los individuos de valor negativo o no. No se perderá generalidad, ya que de todas maneras se estará estudiando a los clientes de todos los tramos etarios y a los de mayor valor.

- *VP vs. Contribución acumulada*

En el siguiente gráfico se puede observar como una primera impresión, que el utilizar la Contribución Neta Acumulada de un individuo no es un referente suficiente para predecir el Valor del Cliente a lo largo de su vida (periodo de estudio de 4 años). Existe un alto porcentaje de individuos de una misma contribución con distintos valores finales.

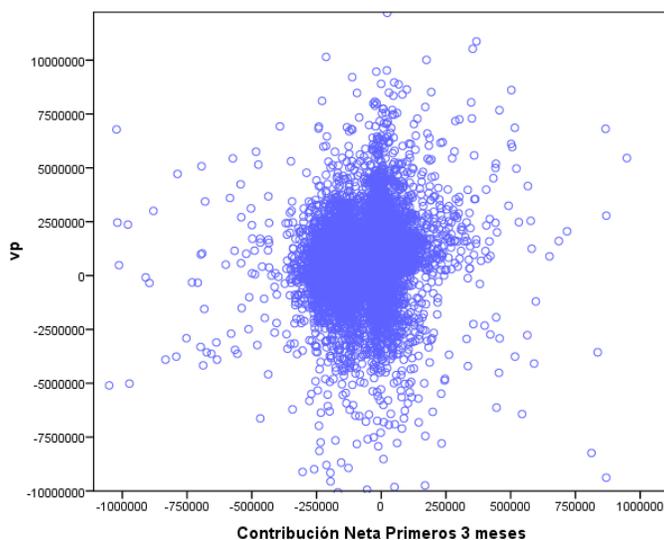
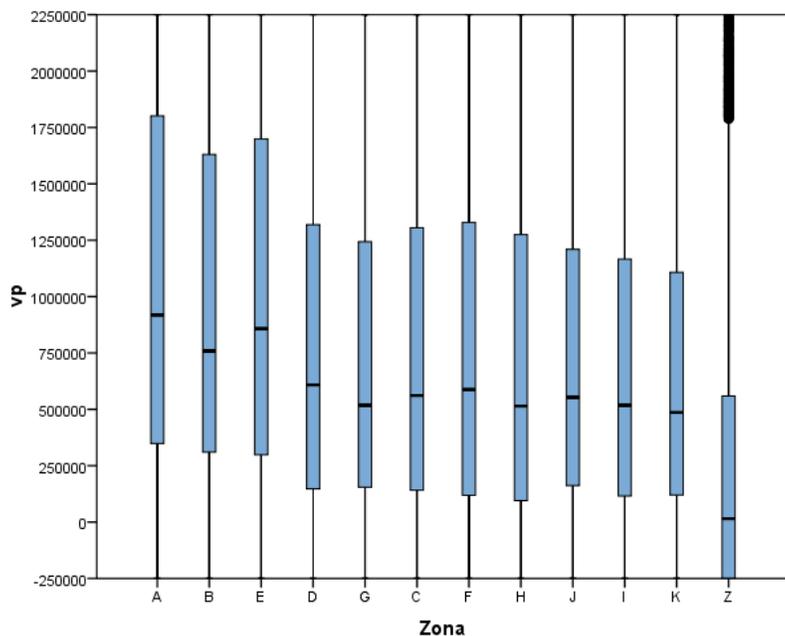


Ilustración 12. Gráfico de Dispersión de la Contribución Neta de los primero 3 meses vs. VP.

- *VP vs. Zonas*

En el sector A se encuentran comunas que destacan por su riqueza como Antofagasta, Calama, Chuquicamata, entre otros. Le sigue la Zona B con las comunas de Chañaral, La Serena y Coquimbo. Se observa que los sectores A, B y E (Ver Anexo D), poseen un 50% de individuos con VP superior a los \$750 mil pesos.



Zona	Comunas representativas
A	Antofagasta, Calama, Chuquicamata.
B	Copiapó, La Serena, Coquimbo.
C	Algarrobo, Santo Domingo, Con-Cón, Olmué.
D	Angol, Chaitén, Osorno.
E	Aisén, Punta Arenas, Coyhaique.
F	Curacaví, Talagante, Colina.
G	La Reina, Las Condes, Lo Barnechea, Vitacura.
H	Nuñoa, Pedro de Valdivia, Providencia.
I	Peñalolén, Independencia, Huechuraba.
J	Independencia, San Bernardo, Recoleta.
K	Cerro Navia, El Bosque, Pedro Aguirre Cerda.
Z	Sin Información.

Ilustración 13. Gráfico Box-Plot Zona por Comunas vs. Valor Presente.

Se puede observar además, que los individuos que no entregan información de su localidad, generan un valor más bajo relativo a los clientes de los que sí se tiene

información. A partir de esto se presume que Zona es una variable *missing at random* o *missing not at random*.

- *VP vs. Tramo Etario*

El valor de los clientes aumenta con su edad y cae en el último tramo en donde la mayor parte está jubilada. A la vez, en este último tramo se tienen todos valores positivos, es decir, no genera pérdidas.

Sin embargo, no hay que perder de vista que se tiene la limitación de los 4 años de datos de los clientes nuevos. Los clientes de menor edad no han alcanzado su madurez económica, por lo que es de esperar que su valor sea menor. Sin embargo, al ingresar más tempranamente, si son leales con la empresa, debieran entregar mayor valor que las personas de mayor edad.

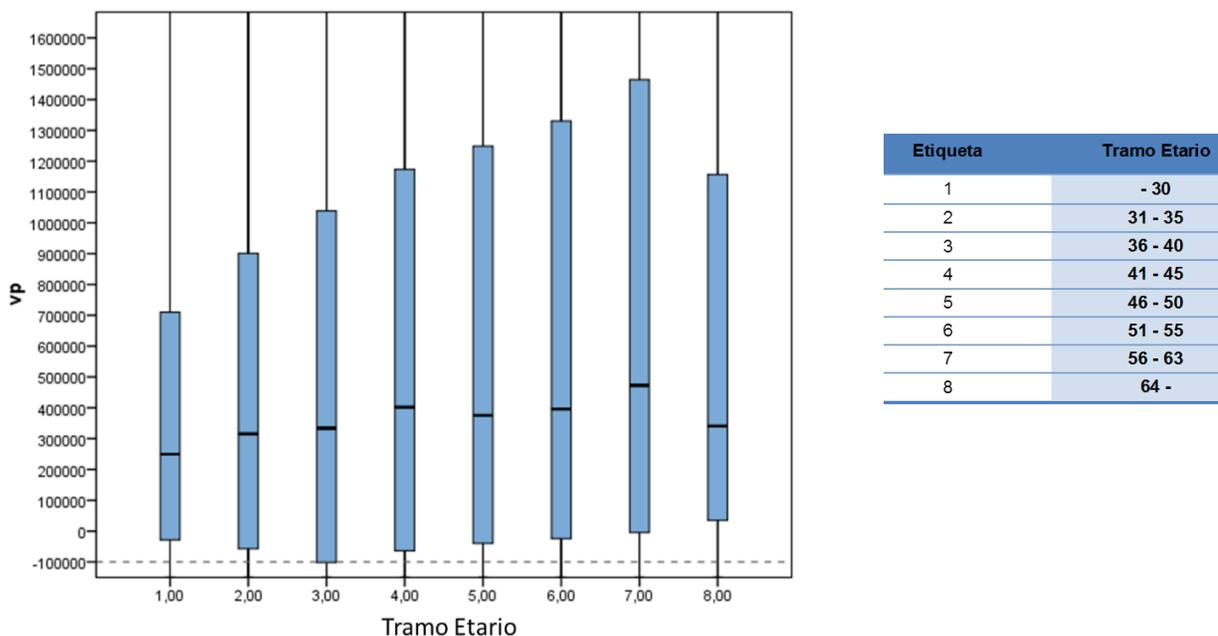


Ilustración 14. Gráfico Box-Plot Tramo Etario vs. Valor Presente.

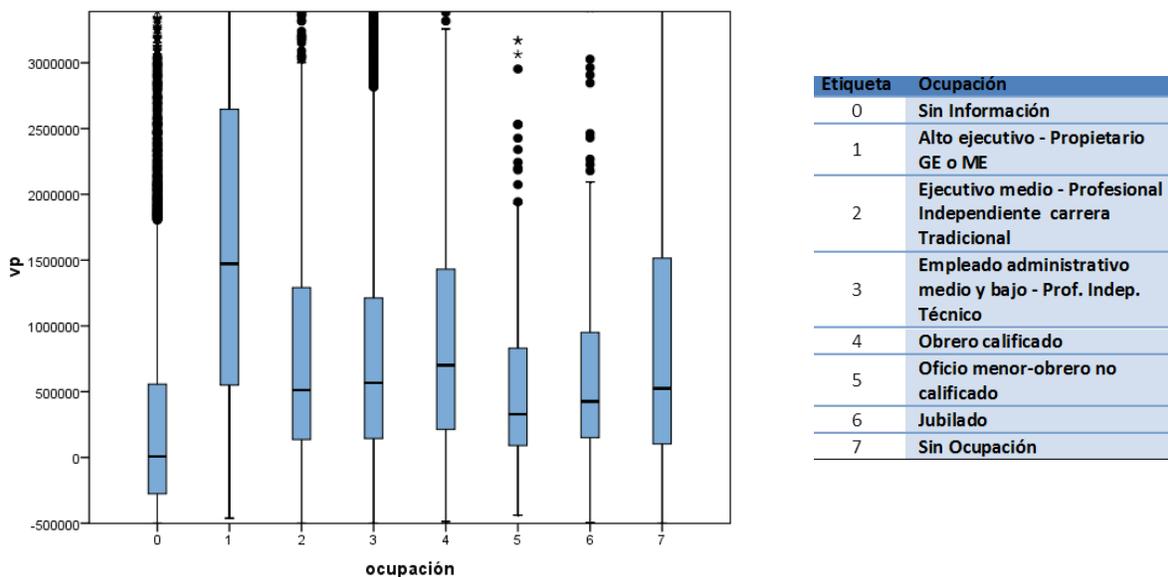
El percentil 75 de los clientes con edad entre los 56 y 63 años duplica en valor a los adultos menores de 30 años y la dispersión aumenta con la edad.

Tramo Etario	Coefficiente de Variación
- 30	0,99
31 - 35	1,03
36 - 40	1,05
41 - 45	1,12
46 - 50	1,27
51 - 55	1,30
56 - 63	1,21
64 -	1,90

Ilustración 15. Tabla que presenta los Coeficientes de Variación del VP por tramo Etario.

- *VP vs. Ocupación*

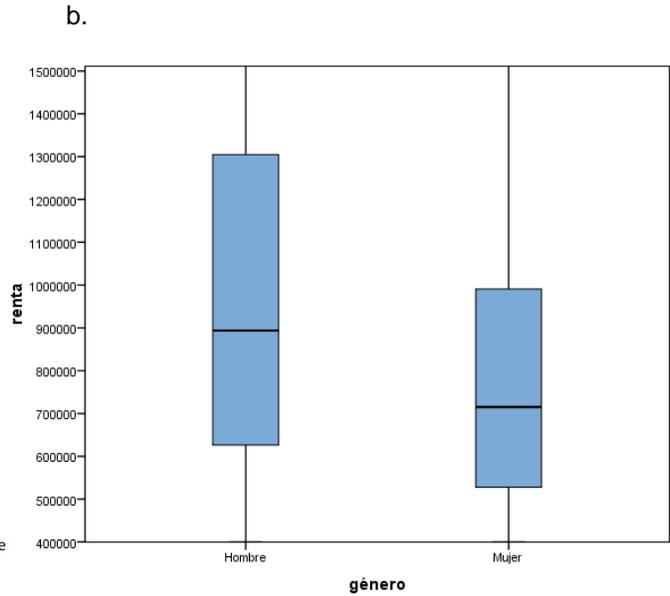
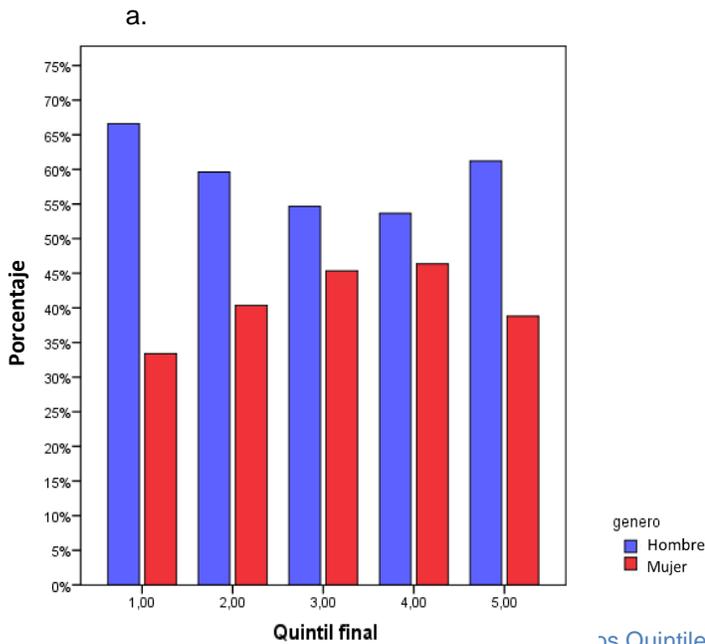
Tal como ocurre con la variable Zona, el 75% de los individuos que no entregan información de su ocupación, poseen un valor inferior al percentil 50 de la mayoría de los otros niveles.



Como es de esperar, los clientes de altos cargos entregan un mayor valor. El 50% de los Ejecutivos y dueños de empresas superan en valor al 75% de cada una de las otras ocupaciones.

- *VP vs. % Composición género.*

A partir del gráfico 17. se infiere que el género masculino es el que contribuye un mayor valor respecto a la mujer tanto para el Quintil Final (quintiles calculados según el valor presente de los clientes con un horizonte de tiempo de 48 meses) de más alto valor como para el de menor. Estos datos se podrían deber a la menor aversión al riesgo y a la mayor Renta promedio que poseen los hombres.



a. Quintiles de valor. b. Gráfico Box-Plot del género según su renta.

A partir de los siguientes gráficos Box-Plot (Ilust.18), la interpretación que se logra es que la tenencia de Línea de Crédito y de Crédito de Consumo en el Mes 3 de vida del cliente, aporta de forma positiva al valor presente de más del 75% de los individuos que tienen estos productos.

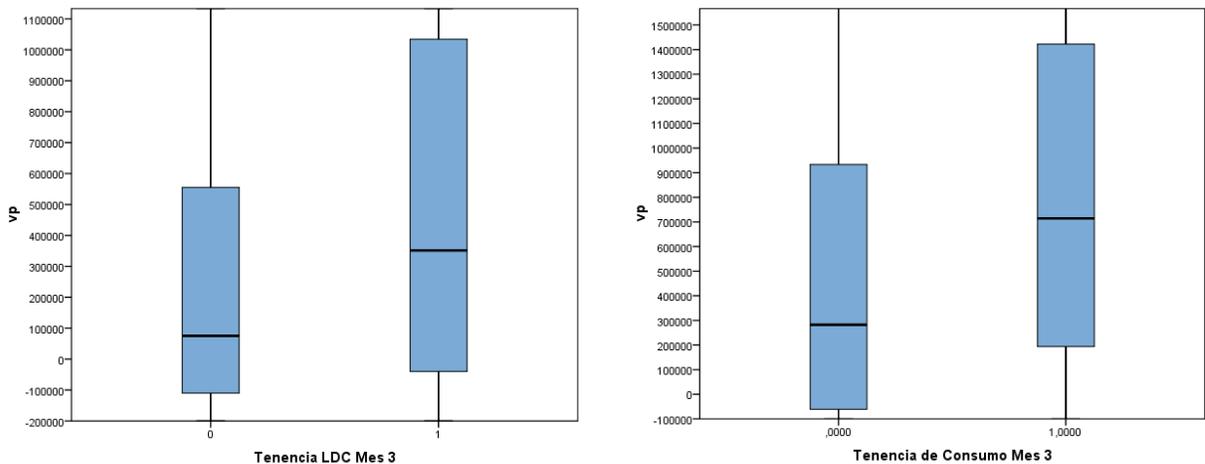


Ilustración 18. Gráfico Box-Plot de VP respecto a Tenencia de productos Línea de Crédito y Crédito de Consumo en Mes 3.

## B. Quintil Final y variables

Primero se recuerda que Quintil Final consiste en la partición en quintiles de los clientes, según su valor de vida (4 años). El Quintil 1 representa a los individuos de mayor valor.

En la Ilustración 19 se observa que la Deuda en la tarjeta de crédito (TDC) posee efectos positivos en el Valor Final de los individuos. Los individuos que se encuentran en el Quintil Final de mayor valor (número 1), tienen una deuda promedio más alta. Sin embargo, también se observa los clientes de valor negativo que también posee una deuda alta. Por esto, se puede concluir que la deuda en pesos con la TDC es un ponderador importante del valor de los clientes, pero no se puede concluir su signo.

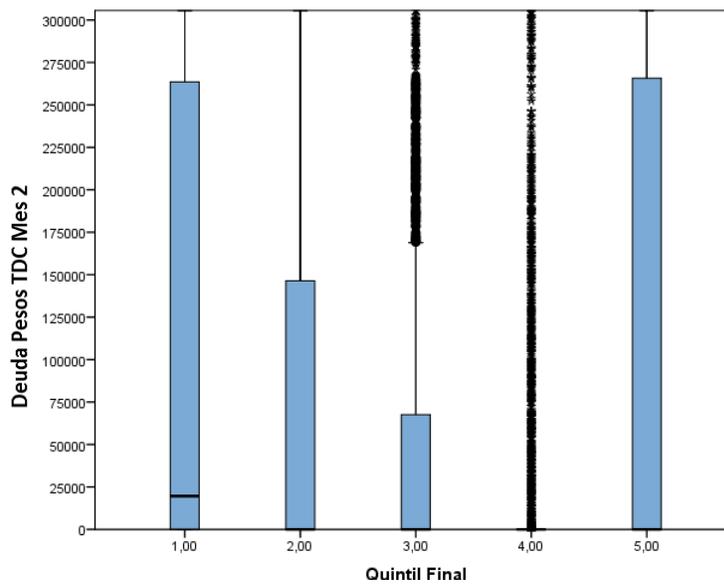


Ilustración 19. Gráfico Box-Plot de Quintiles por Deuda en TDC del Mes 2.

El gráfico 23 muestra que el porcentaje de hipotecarios adquiridos por Quintil Final de valor en el Mes 3, tiene una estrecha relación con el valor de los clientes. El test ANOVA que demuestra su significatividad se encuentra en Anexo G. Lo mismo ocurre con la tenencia de Créditos de Consumo.

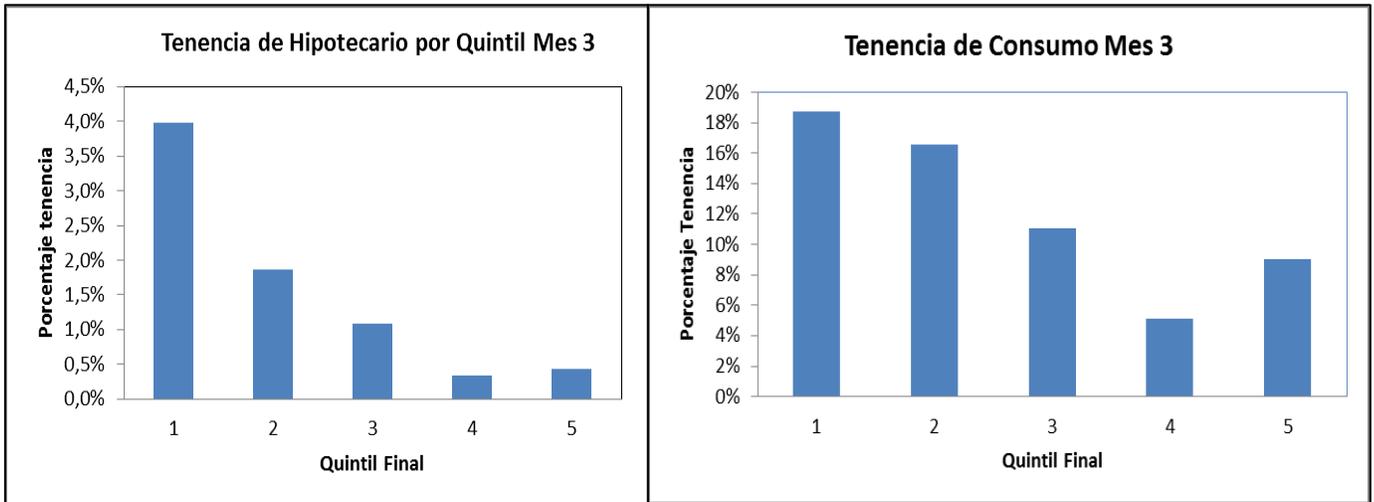
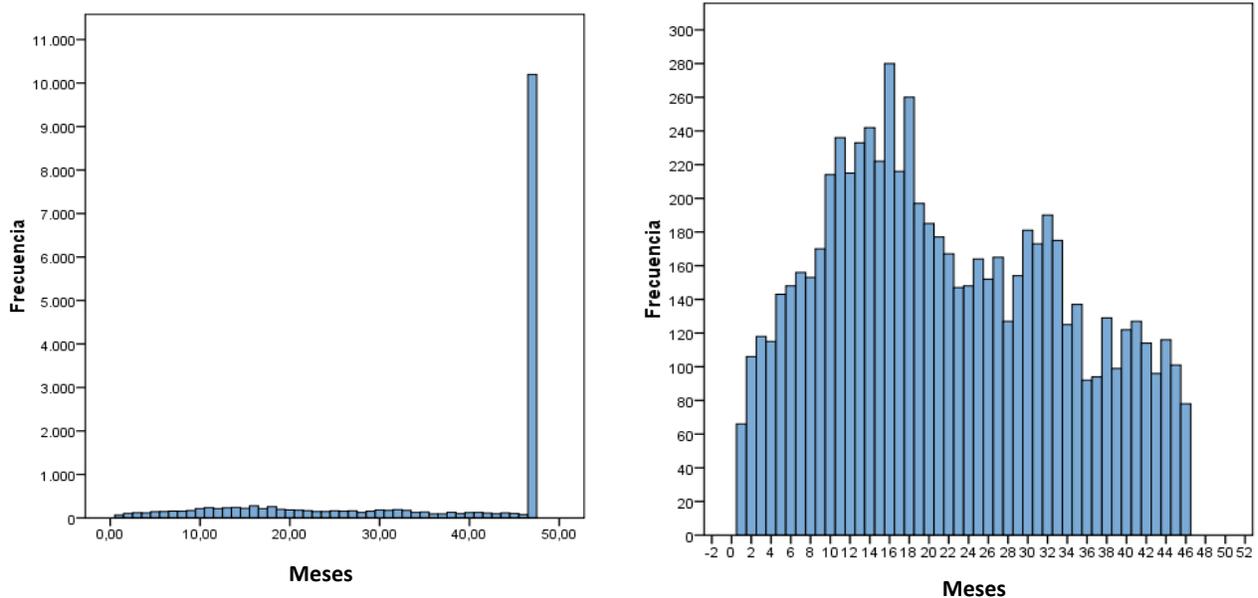


Ilustración 20. a. Tablas de porcentaje Quintil vs. Tenencia de Crédito Hipotecario Mes 3. b. Tablas de porcentaje Quintil vs. Tenencia de Crédito de Consumo Mes 3.

### C. Duración

El gráfico 21.1 representa la frecuencia de duración (fuga) de todos los clientes, la cola final consiste en los individuos que se mantienen en la empresa. En la figura 22.2. se realiza un zoom a los clientes fugados. Se observa un peak entre los 16 y 18 meses. Además, el 50% de los clientes dura más de 48 meses en el banco.



Debido a que no se tiene información de la duración de los clientes para más allá de 48 meses. Se procede a estimar su evolución a través del tiempo con método de estimación paramétrico.

Para lo anterior, se utiliza una distribución Gamma-Weibull. En el siguiente gráfico se muestra la curva de permanencia de los clientes durante los 48 meses.

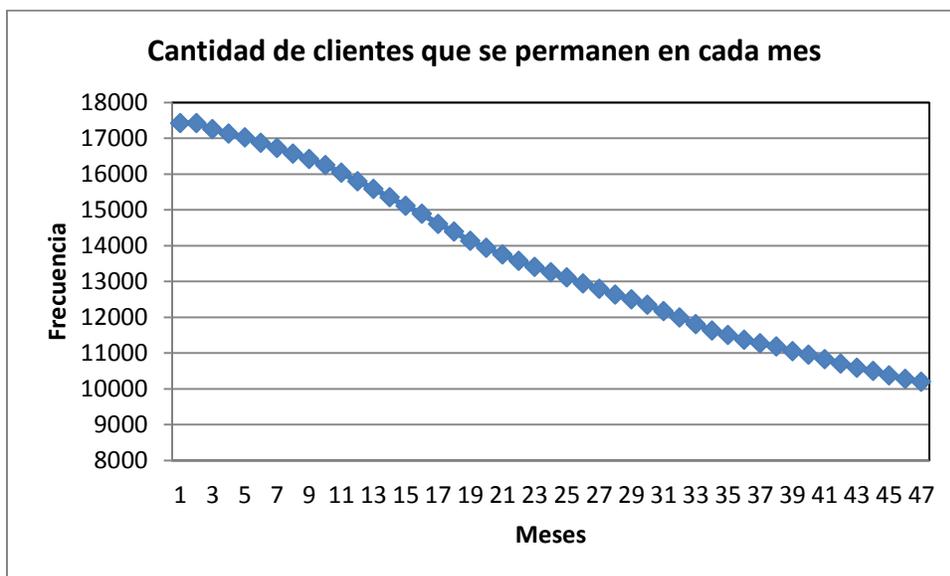


Ilustración 22. Gráfico de frecuencia de clientes que permanecen en el banco en cada mes.

Como resultado de la estimación se obtuvo:

Parámetro	Valor
<b>r</b>	<b>92,46827801</b>
<b>Alpha</b>	<b>20.152,41552</b>
<b>c</b>	<b>1,24251382</b>
<b>LL</b>	<b>-39.456,754</b>

Tabla 21. Resultados del modelamiento de duración de los clientes utilizando una función Gamma-Weibull.

De la figura 23 que se muestra a continuación, se observa un mediano ajuste de los datos y sobre-estimación en el pronóstico de la cantidad de abandono de por día, ya que tendería a disminuir la tasa de riesgo, pues va disminuyendo la heterogeneidad (fieles-no fieles) y quedan los clientes más fieles [38].

Sin embargo, se puede obtener un estimado de la duración promedio de los cuenta-correntistas del banco: el promedio permanece alrededor de 8,3 años. Esta cifra no es muy alejada de la información real que se tiene para todos los cuenta-correntistas de 7,5 años.

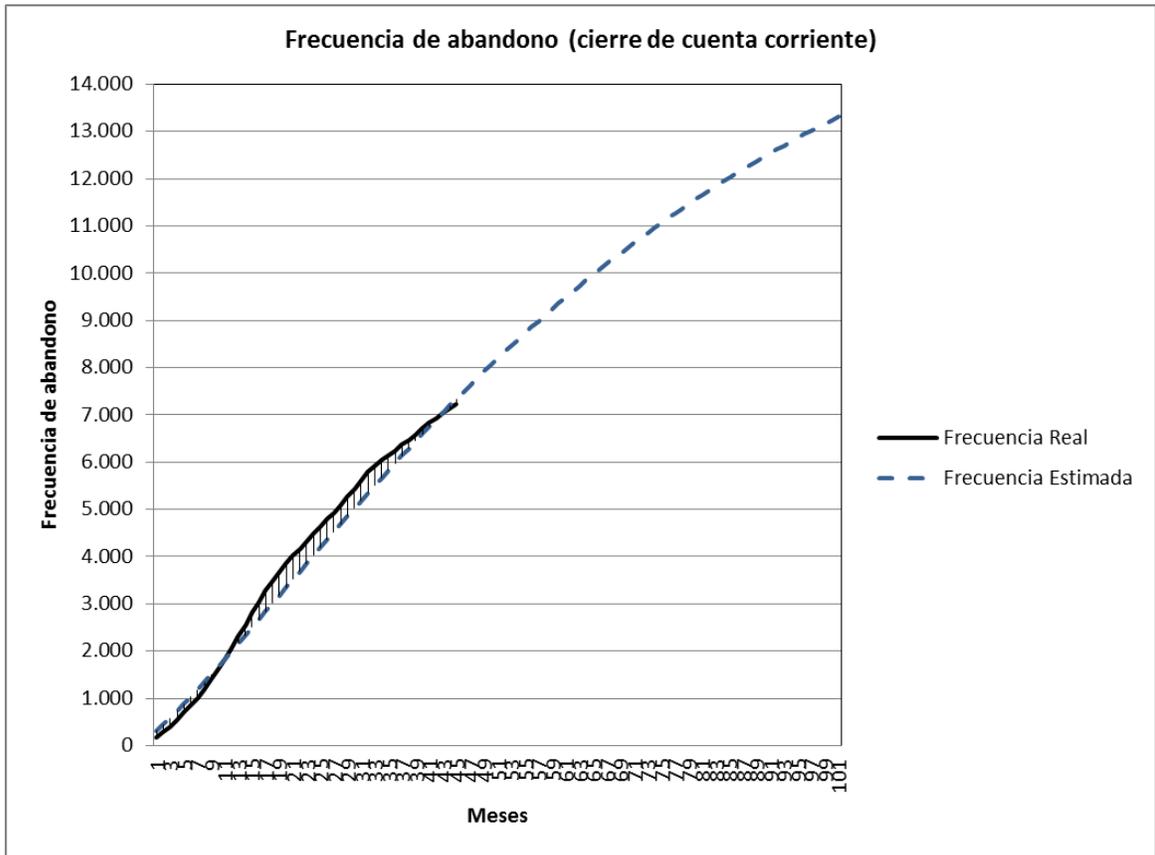


Ilustración 23. Gráfico que compara la trayectoria de la frecuencia de abandono de los clientes con respecto a su estimación.

La ilustración 24. muestra la relación existente entre Quintil al que pertenece el cliente y su duración. Las personas que se encuentran el quintil 1, de mayor valor, son quienes tienen una mayor duración promedio (Ver Anexo J Test Anova).

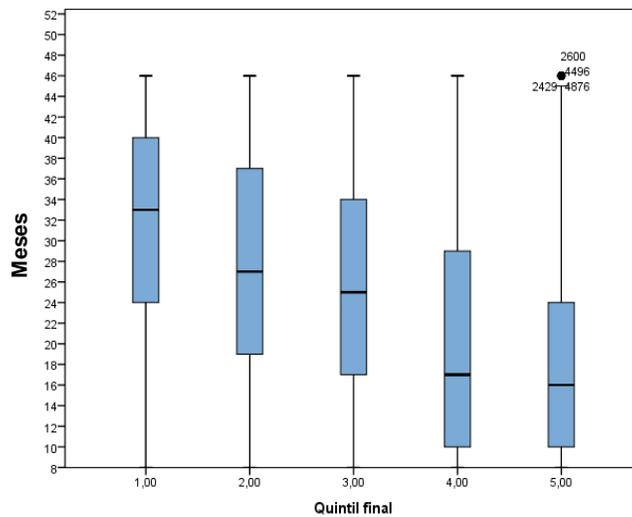
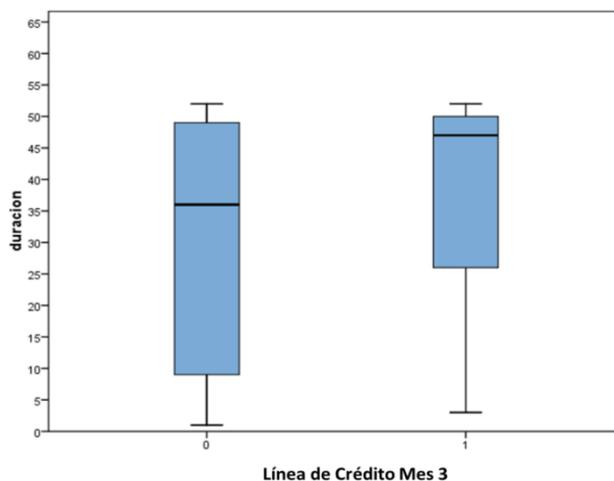


Ilustración 24. Gráfico Box-Plot de Quintiles vs. Duración en meses.

Al realizar un análisis por producto, se puede ver que la tenencia de la Línea de Crédito aumenta en promedio 10 meses la duración del cliente.



**Ilustración 25. Tenencia de Línea de Crédito en Mes 3 vs. Duración del Cliente.**

Se realizó un Test Anova para evaluar si el mes de entrada del cliente influía en su duración. El resultado fue no significativo, es decir, la duración promedio no varía según el tiempo de entrada.

Cohorte – Duración					
	Suma de Cuadrados	df	Media Cuadrática	F	Sig.
<b>Entre grupos</b>	<b>259,950</b>	<b>5</b>	<b>51,990</b>	<b>,365</b>	<b>,873</b>
<b>Dentro del Grupo</b>	<b>1028947,888</b>	<b>7219</b>	<b>142,533</b>		
<b>Total</b>	<b>1029207,838</b>	<b>7224</b>			

Tabla 22. Test Anova para testear diferencia de medias de VP entre cohortes.

## D. Valor Presente Promedio

Las Zonas mantienen casi el mismo orden que el del VP sin duración. La Zona Z (Sin Información), el único grupo con un valor presente promedio negativo pronunciado de -\$25.114 pesos (la Zona B tiene un promedio de \$24.806 pesos) y un coeficiente de variación de -6,8.

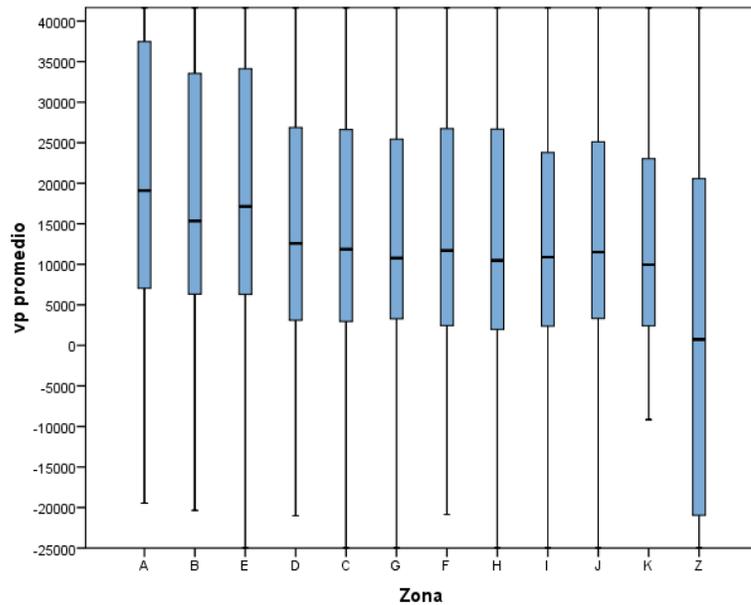


Ilustración 26. Gráfico Box-Plot del VP promedio por Zonas.

## 4. Resultados Modelamiento

### A. Renta

Se estimó la renta con una regresión lineal, utilizando el método *Backward* en el Software SPSS V.17, sobre una población de 4.090 registros (individuos de los que se tenía información de la renta). Se obtuvieron los siguientes resultados:

Variables Explicativas	Betas Estimados	Desv. Estándar	t	Significancia
<b>Constante</b>	745.025,094	41.878,154	17,790	,000
<b>Tenencia TDC Mes 2</b>	-117.989,155	19.730,896	-5,980	,000
<b>Cupo Pesos TDC Mes 2</b>	,359	,019	19,056	,000
<b>Contribución Captación Mes 2</b>	4,420	1,343	3,291	,001
<b>Deuda Pesos TDC Mes 3</b>	0,118	,044	2,686	,007
<b>Consumo Monto Mes 3</b>	0,037	,013	2,805	,005
<b>Ln Saldo Promedio en Cta.Cte. Mes 3</b>	6.182,823	1.977,363	3,127	,002
<b>Pertenece Zona A</b>	192.210,888	34.919,701	5,504	,000
<b>Pertenece Zona B</b>	295.691,112	45.836,918	6,451	,000
<b>Pertenece Zona C</b>	-59.202,090	24.634,888	-2,403	,016
<b>Pertenece Zona G</b>	212.308,077	30.443,202	6,974	,000
<b>Pertenece Zona H</b>	221.266,783	30.558,467	7,241	,000
<b>Pertenece Zona K</b>	-135.539,708	33.220,739	-4,080	,000
<b>Ocupación 1</b>	966.472,170	100.875,525	9,581	,000
<b>Ln Cupo Dólar TDC Mes 2</b>	-15.527,149	7.521,307	-2,064	,039
<b>Ln Consumo Saldo Mes 3</b>	-5.983,721	3.026,543	-1,977	,048
<b>Hombre</b>	200.264,542	16.067,700	12,464	,000
<b>Ocupación 2</b>	387.167,570	22.268,370	17,386	,000

Ilustración 27. Resultados estimación de la variable renta con regresión lineal.

Los estadísticos de ajuste conseguidos son:

Análisis Global	Valor
R2 ajustado Calibración	23,1%
MAPE	42%
WAPE	34%
R2 ajustado validación	18,5%
MAPE	46%
WAPE	36%

Tabla 23. Resultados globales de ajuste del modelo de predicción de la renta.

La renta estimada resulta tener una correlación similar a la de la renta real.

VP vs. Renta Estimada	Coefficiente de Correlación de Pearson	Sig. (2 colas)
	0,218	0,000

Tabla 24. Correlación VP con Renta estimada

- *Análisis de resultado Estimación*

Respecto a la variable ocupación, la renta sub-estima entre 200 mil y 20 mil pesos a la renta real. Sin embargo, a excepción del tramo 5 (entre 50 y 55 años), se observa una tendencia similar entre la renta real y la estimada

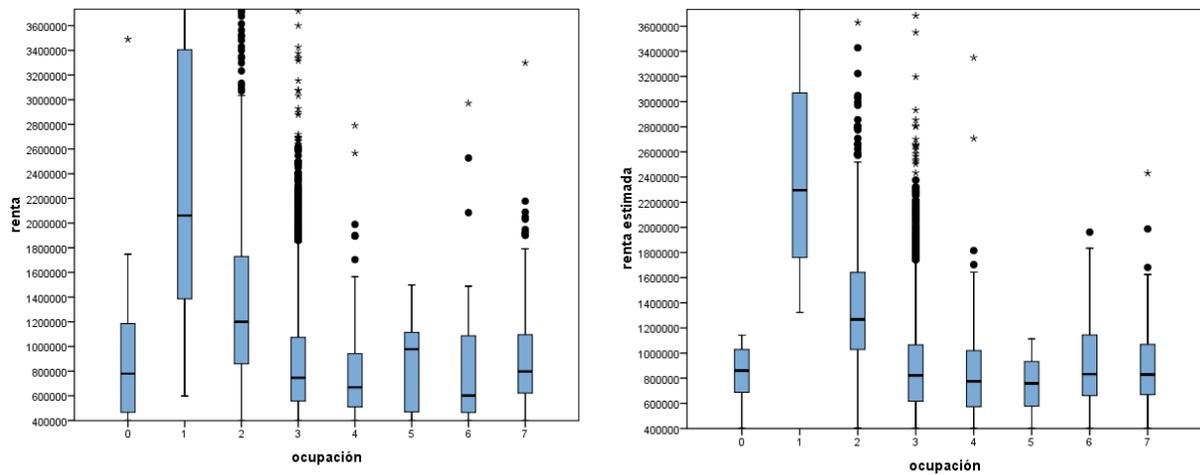


Ilustración 28. Grafico Box-Plot de la renta y renta estimada según la ocupación.

En cuanto a la edad, donde más se equivoca en predecir el modelo, es en el tramo sobre los 64 años, ya que solamente se tiene la renta para 41 individuos.

A partir del modelo se puede concluir que los clientes con la ocupación de ejecutivos de altos rangos ganan alrededor 966 mil pesos más que el resto de los clientes, seguidos de los profesionales tradicionales con 387 mil sobre el promedio. Lo mismo ocurre con el género, el hombre tiene en promedio ingresos superiores a la mujer en alrededor de 200 mil pesos.

La zona B (La Serena, Coquimbo, Vallenar, etc.), dispone de un mayor salario promedio que el resto, por casi 300 mil pesos. Las comunas de La Reina, Las Condes, Lo Barnechea y Vitacura (Zona G) y Ñuñoa, Pedro Valdivia Norte, Providencia (Zona H) también contribuyen con un alto valor a la renta por sobre los demás municipios, con poco más de 200 mil pesos.

Es esperable que la Contribución en Captación explique la Renta, ya que tiene que ver con los Fondos Mutuos y Depósitos a Plazo que explican de forma indirecta el patrimonio disponible del cliente.

Además, el saldo promedio que tiene el cliente en su cuenta corriente evidencia una renta promedio positiva. Al contrario, el disponer un alto saldo (por pagar) en un crédito de consumo al mes 3, implica una renta promedio menor.

La Tenencia de Tarjeta de Crédito en el mes 2, evidencia una baja renta presumiéndose una necesidad por obtener crédito inmediato.

## B. Valor Presente

Los resultados de este modelo son:

<b>Variables Explicativas</b>	<b>Betas Estimados</b>	<b>Desv. Estándar</b>	<b>t</b>	<b>Significancia</b>
<b>Constante</b>	-156.189,701	47.144,453	-3,313	,001
<b>Pertenece Zona A</b>	1.194.825,749	78.563,411	15,208	,000
<b>Pertenece Zona B</b>	1.095.349,398	93.764,117	11,682	,000
<b>Pertenece Zona C</b>	865.549,364	51.400,658	16,839	,000
<b>Pertenece Zona D</b>	969.448,576	45.730,628	21,199	,000
<b>Pertenece Zona E</b>	1.055.259,516	124.789,821	8,456	,000
<b>Pertenece Zona F</b>	892.916,425	82.713,657	10,795	,000
<b>Pertenece Zona G</b>	818.814,772	59.771,824	13,699	,000
<b>Pertenece Zona H</b>	810.693,509	64.940,611	12,484	,000
<b>Pertenece Zona I</b>	791.253,752	41.486,580	19,073	,000
<b>Pertenece Zona J</b>	810.996,093	51.748,588	15,672	,000
<b>Pertenece Zona K</b>	799.460,306	66.683,407	11,989	,000
<b>Ln Cupo Dólar TDC Mes 2</b>	-27.790,471	8.261,346	-3,364	,001
<b>Ocupación 7</b>	463.635,454	58.007,921	7,993	,000
<b>Renta Estimada</b>	,320	,036	8,809	,000
<b>Razón Renta Est. Vs. Numero de Cheque Mes 1</b>	,305	,076	4,016	,000
<b>Contribución Filiales Mes 3</b>	26,134	1,561	16,745	,000
<b>Tenencia de TDC Mes 2</b>	118.200,319	29.949,241	3,947	,000
<b>Contribución Neta Mes 2</b>	1,727	,184	9,406	,000
<b>Contribución Colocación Mes 2</b>	5,455	,337	- 16,204	,000
<b>Contribución Comisiones Mes 2</b>	-2,646	1,260	-2,100	,036
<b>Contribución Provisiones Mes 2</b>	-3,105	,350	-8,872	,000

Variables Explicativas	Betas Estimados	Desv. Estándar	t	Significancia
<b>Tramo Edad 1</b>	-239.810,052	34.504,724	-6,950	,000
<b>Tramo Edad 5</b>	46.056,963	8.383,581	5,494	,000
<b>Tramo Edad 6</b>	40.561,458	8.246,458	4,919	,000
<b>Tramo Edad 7</b>	44.301,277	7.964,653	5,562	,000
<b>Tramo Edad 8</b>	53.803,185	9.919,223	5,424	,000

Los ajuste globales del modelo de VP:

Análisis Global	Valor
R2 ajustado Calibración	26%
WAPE	85%
R2 ajustado validación	22,2%
WAPE	87%

Ilustración 29. Resultados de Ajuste global del modelo con variable dependiente VP.

- *Análisis de resultado Estimación*

Los adultos jóvenes hasta los 30 años, entregan entre \$239.810 pesos menos en promedio de VP que los demás clientes. Los clientes de menor edad no han alcanzado su madurez económica, por lo que es de esperar que su valor sea menor.

Como era de esperar, la Contribución en colocaciones aumenta del valor del cliente. Lamentablemente la variable ocupación 7 corresponde a Sin Ocupación, por lo que es difícil de comprender el que estos individuos aporten de forma positiva al valor presente.

Mientras más se provisione, mayor valor entrega el cliente. Esto se puede suponer debido a que mientras más endeudado el cliente, mayor posibilidad de generar ganancias sobre eso. No obstante, aporta muy poco al VP del cliente.

Las Zonas definen mayormente el VP del cliente con casi 800 mil pesos de base. La tenencia de Tarjeta de crédito también es un signo positivo para el VP.

### C. Duración

Variables Explicativas	Betas Estimados	Desv. Estándar	t	Significancia
<b>Constante</b>	16,0718808	0,5453353	-3,313	0,0000000
<b>Pertenece Zona A</b>	19,9351439	1,3219479	15,208	0,0000000
<b>Pertenece Zona B</b>	19,2426158	1,6358174	11,682	0,0000000
<b>Pertenece Zona C</b>	19,1574331	0,7666040	16,839	0,0000000
<b>Pertenece Zona D</b>	19,3332826	0,6600187	21,199	0,0000000

<b>Variables Explicativas</b>	<b>Betas Estimados</b>	<b>Desv. Estándar</b>	<b>t</b>	<b>Significancia</b>
<b>Pertenece Zona E</b>	22,5160002	2,1391655	8,456	0,0000000
<b>Pertenece Zona F</b>	18,1894467	1,3025273	10,795	0,0000000
<b>Pertenece Zona G</b>	18,8785826	0,9158408	13,699	0,0000000
<b>Pertenece Zona H</b>	19,4857024	0,9958240	12,484	0,0000000
<b>Pertenece Zona I</b>	19,3584081	0,6257825	19,073	0,0000000
<b>Pertenece Zona J</b>	18,6241072	0,7765141	15,672	0,0000000
<b>Pertenece Zona K</b>	19,6838997	1,0859298	11,989	0,0000000
<b>Renta Estimada</b>	-0,0000029	0,0000006	-3,364	0,0000079
<b>Razón Núm. Cheques / Renta Estimada Mes 2</b>	0,0000023	0,0000007	7,993	0,0012234
<b>Razón Núm. Cheques / Renta Estimada Mes 3</b>	0,0000042	0,0000007	8,809	0,0000000
<b>Razón Cupo Pesos TDC / Renta Estimada Mes 3</b>	1,0988697	0,1478411	4,016	0,0000000
<b>Contribución Provisiones Netas Mes 3</b>	0,0000089	0,0000016	16,745	0,0000000
<b>Tenencia de Seguro Mes 1</b>	0,3421038	0,1651543	3,947	0,0383478
<b>Tenencia de Seguro Mes 2</b>	-0,8961318	0,2072262	9,406	0,0000155
<b>Contribución Gastos Mes 3</b>	0,0000242	0,0000055	- 16,204	0,0000124
<b>Tenencia de LDC Mes 3</b>	6,7619725	0,5128171	-2,100	0,0000000
<b>Tenencia de TDC Mes 3</b>	1,2002160	0,2476955	-8,872	0,0000013
<b>Cupo en Pesos TDC Mes 3</b>	0,0000036	0,0000004	-6,950	0,0000000
<b>Cupo en Dólar TDC Mes 3</b>	-0,0009191	0,0001836	5,494	0,0000006
<b>Deuda en Pesos TDC Mes 2</b>	-0,0000063	0,0000006	4,919	0,0000000

Tabla 25. Resultado estimación con regresión lineal de la variable duración.

Sus estimadores globales:

<b>Análisis Global</b>	<b>Valor</b>
R2 ajustado Calibración	36,1%
WAPE	28%
R2 ajustado validación	36,7%
WAPE	33%

Tabla. Resultados globales de ajuste del modelo de predicción de la duración.

- *Análisis de resultado Estimación*

Mientras más se tenga que provisionar en los primeros tres meses, la estadía del cliente será menor. Esto es algo que se desconocía ya que el provisionar puede estar relacionado con un crédito hipotecario o de consumo que incentive a una permanencia más prolongada del cliente.

Al mismo tiempo, un mayor número de cheques girados respecto a la renta, en el Mes 2 y 3 y un Cupo más grande en la cuenta corriente, aumentan el tiempo de vida del individuo. Se piensa que el cliente está más involucrado con la empresa.

Con respecto a la Contribución Gasto en el Mes 3, no se puede concluir de forma directa ya que esta variable no tiene una correlación significativa con la duración, se presume que entonces que esta variable en conjunto con otra u otras genera una correlación significativa con la duración.

Si se realiza una regresión lineal univariada de esta variable y una bivariada con la variable ejemplo Tenencia de LDC Mes 3, también incorporada por el modelo, se puede observar cómo cambia la significancia de la variable dentro de la regresión.

a.

	Duración	Beta	Error Estándar	Signif.
Modelo Univariado	Constante1	22,036	,170	,000
	Contribución Gasto Mes 3	0,000008	,000	,152

b.

	Duración	Beta	Error Estándar	Signif.
Modelo Bivariado	Constante2	15,095	,448	,000
	Contribución Gasto Mes 3	0,000012	,000	,026
	LDC_N_3	7,768	,465	,000

Tabla 26. a. Resultado estimación no significativa con regresión lineal de la variable duración con regresor Contribución Gasto Mes 3. b. Resultado estimación significativa con regresión lineal de la variable duración con los regresores Contribución Gasto Mes 3 y Tenencia de LDC en Mes 3.

Se puede interpretar que dada la diversidad de los clientes, el que dos individuos requieran un alto gasto para la empresa no logra explicar de forma individual su duración, pero sí el que posea línea de crédito y al mismo tiempo esté generando un alto gasto afecta de manera positiva a la duración.

#### D. Valor Presente Promedio

Variables Explicativas	Betas Estimados	Desv. Estándar	t	Significancia
<b>Constante</b>	-2082,111	814,578	-2,556	,011
<b>Cupo LDC Mes 1</b>	-,004	,001	-4,412	,000
<b>Cupo TDC Dólar Mes 1</b>	-1,672	,241	-6,932	,000
<b>Tenencia CAP Mes 1</b>	4818,824	1612,746	2,988	,003
<b>Tenencia LDC Mes 2</b>	-6121,115	1336,519	-4,580	,000

Variables Explicativas	Betas Estimados	Desv. Estándar	t	Significancia
<b>Cupo LDC Mes 2</b>	,006	,001	7,242	,000
<b>Tenencia Consumo Mes 2</b>	-11274,349	1674,615	-6,733	,000
<b>Contribución Neta Mes 2</b>	,770	,018	41,859	,000
<b>Contribución Captación Mes 2</b>	-,678	,049	-13,807	,000
<b>Contribución Comisiones Mes 2</b>	-,705	,045	-15,643	,000
<b>Contribución Filiales Mes 2</b>	-,770	,070	-11,083	,000
<b>Contribución Gastos Mes 2</b>	-,718	,019	-38,576	,000
<b>Contribución Provisiones 2</b>	-,751	,020	-37,966	,000
<b>Contribución Basilea Mes 2</b>	5,195	,895	5,807	,000
<b>Núm. de Cheques Mes 3</b>	346,179	43,167	8,020	,000
<b>Tenencia LDC Mes 3</b>	11691,691	1446,606	8,082	,000
<b>Tenencia TDC Mes 3</b>	3455,860	400,392	8,631	,000
<b>Deuda TDC Pesos Mes 3</b>	,017	,001	15,297	,000
<b>Tenencia Consumo Mes 3</b>	10101,343	1510,418	6,688	,000
<b>Contribución Neta Mes 3</b>	,038	,005	7,401	,000
<b>Contribución Basilea Mes 3</b>	-2,698	,493	-5,470	,000
<b>Tenencia SEG 1</b>	908,553	289,415	3,139	,002

Tabla 27. Estimaciones de regresión lineal con VP-promedio como variable dependiente.

- *Análisis de resultado Estimación*

Los resultados obtenidos esperados son que la Contribución en Gastos del Mes 2 (número negativo), implique un aumento en el Valor presente, el cliente genera más movimiento en la empresa. También, mientras mayor sea la cantidad de cheques que gire el cliente en el Mes 3, mayor será su valor. Esta interpretación recae sobre los valores positivos, como se e observa en el siguiente gráfico:

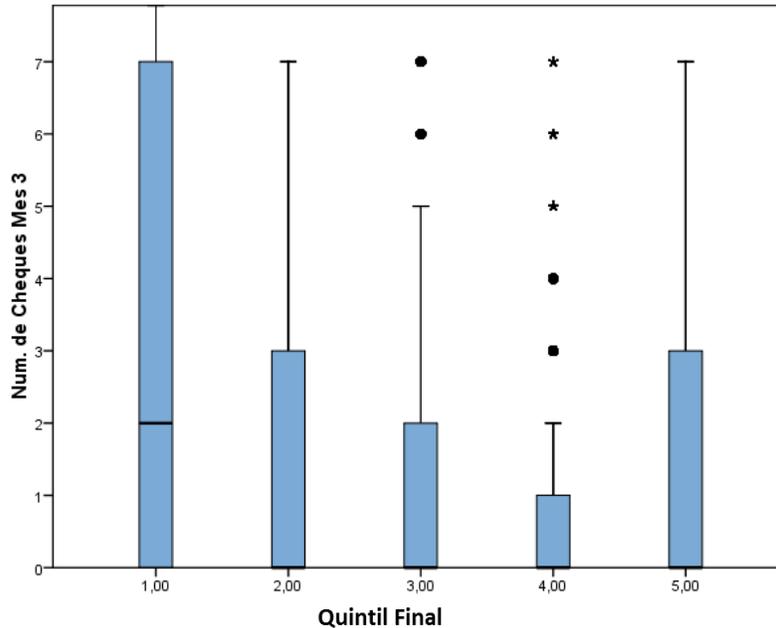


Ilustración 30. Gráfico Box-Plot de Quintiles Finales vs. Número de Cheques girados en el tercer mes de relación con el banco.

En el modelo se obtiene que una persona que gira 5 cheques más sobre otra, en promedio agrega \$1.730 pesos, esto supera al 34% del VP-promedio de todos los clientes.

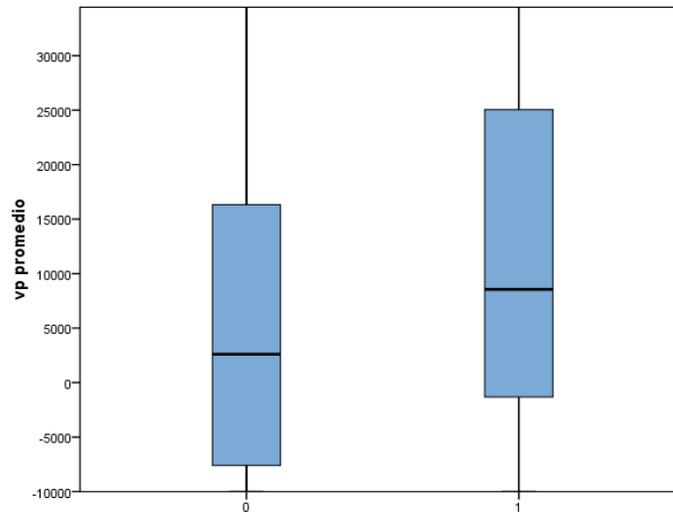


Ilustración 31. Gráfico Box-Plot tenencia de LDC en Mes 3 y VP promedio

La tenencia de la Línea de Crédito en el Mes 3 genera una contribución media al VP promedio de \$11.691 pesos. En contraposición a la tenencia de Crédito de Consumo en el Mes 2 que reduce el valor casi en la misma cantidad.

Un variable importante es que ya en el primer Mes, mientras el cliente tenga más seguros, mayor será su valor (en 908,55 veces su cantidad).

## E. Resultados de Ajuste modelos de predicción de Valor Presente

Se esperaba que la información total de las contribuciones de los primeros tres meses del cliente entregara alguna información, pero el modelo no las incluyó. Lo mismo ocurrió con las variaciones porcentuales de las contribuciones de un mes para otro calculadas. También se descubrió que en gran parte contribuye más a generar información el Mes 3.

Los resultados de ajuste de los modelos VP y VP promedio multiplicado por las duración son:

	Modelo VP	Modelo VP promedio y Duración
<b>Media Cuadrática</b>	<b>1.735.280</b>	<b>1.809.896</b>
<b>WAPE</b>	<b>85%</b>	<b>84%</b>

Tabla 28. Estadísticos de ajuste de modelos Con y Sin duración.

Se puede decir según estos indicadores que los modelos se comportan de manera similar.

## F. Evaluación de la Regresión según tipo de datos

Se decidió evaluar el modelo con los estadísticos WAPE y Error cuadrático medio. El WAPE tiene el inconveniente de entregar valores muy altos para casos en que la predicción es de un valor muy pequeño (un valor partido por un número chico genera un número muy grande).

Se realizó un intervalo de confianza, al 95%, del Valor Presente para los tres quintiles de menor valor. Se puede observar que la media del quintil 4 es poco menos de 10 mil pesos y a pesar de tener un intervalo de confianza distinto de cero, su coeficiente de variación es de 7,5, es decir, la media no es un buen indicador del VP.

		Final		
		Quintil 3	Quintil 4	Quintil 5
	Media	\$332.029	\$9.731	-\$1.150.540
<b>Intervalo de Confianza al 95%</b>	Límite inferior	\$328.310	\$7.298	-\$1.229.337
	Límite superior	\$335.749	\$12.165	-\$1.071.743
	Desv. Estándar	\$111.980	\$73.266	\$2.372.537
	Coef. De Variación	0,3	7,5	2,1

Tabla 29. Descriptivo de VP para los Quintiles 3, 4 y 5.

El modelo ejecutado sobre todos los clientes (valor negativo y positivo) no resulta adecuado para un cierto tipo de registros: su VP se encuentra en torno al cero.

Nuevamente las altas desviaciones porcentuales absolutas se deben a valores muy cercanos a cero.

Dado el alto Valor del WAPE, se probó además el modelo con distintos tipos de registros, obteniendo los WAPEs:

Datos	WAPE VP
Valor Presente	85%
Valor Presente Absoluto mayor a 100 mil	83%
Valor Presente Mayor a Cero	64%
Valor Presente sobre 100 mil	58%

Tabla 30. Resultados WAPE de estimación del modelo de regresión lineal sobre la variable VP con distintos tipo de datos.

Se encuentra que el quintil 4 posee los valores presentes más cercanos a cero y a la vez sus variables explicativas también rodean el cero. Esto se puede ver en los gráficos Quintil final vs. Alguna variable transaccional. Por ejemplo, en la siguiente figura (Ilustración 3), se observa cómo la Contribución Basilea Mes 3 se encuentra en torno al cero para el grupo 4. De modo resumido muestra lo poco que debe provisionarse (según las normativas SBIF), para cubrirse de alguna pérdida. Esto se puede explicar por el poco movimiento al tercer mes de los individuos de ese grupo.

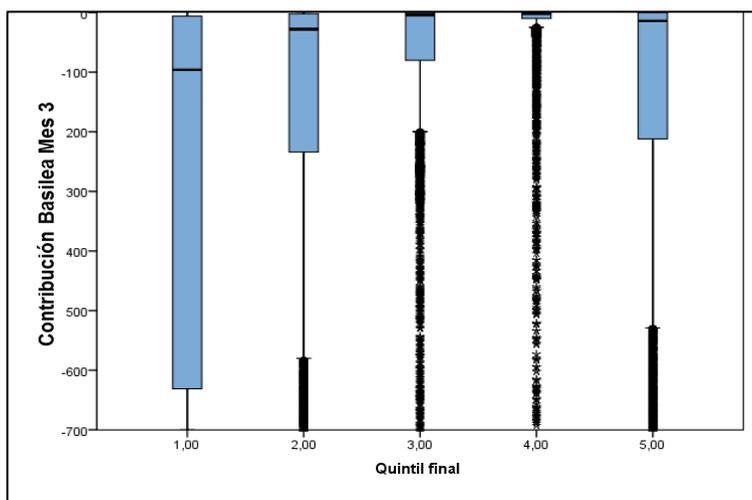


Ilustración 32. Contribución basilea Mes 3 entregada por cada Quintil de valor final.

Se puede concluir que la regresión se ajusta mejor para valores grandes de las variables explicativas que se expresan por medio de valores grandes de la variable dependiente.

Como se puede observar en la figura el modelo Sin duración se comporta mejor para los primeros tres quintiles y se dispara donde los VP son cercanos o muy inferiores a cero. Al contrario, el modelo Con duración tiene un peor desempeño relativo en quintiles de mayor valor, la desviación porcentual absoluta para el Quintil 5 es por lo

menos 35% inferior respecto al modelo Sin duración y el 25% de su cuarto Quintil posee una desviación menor al 150% (que aún sigue siendo alto).

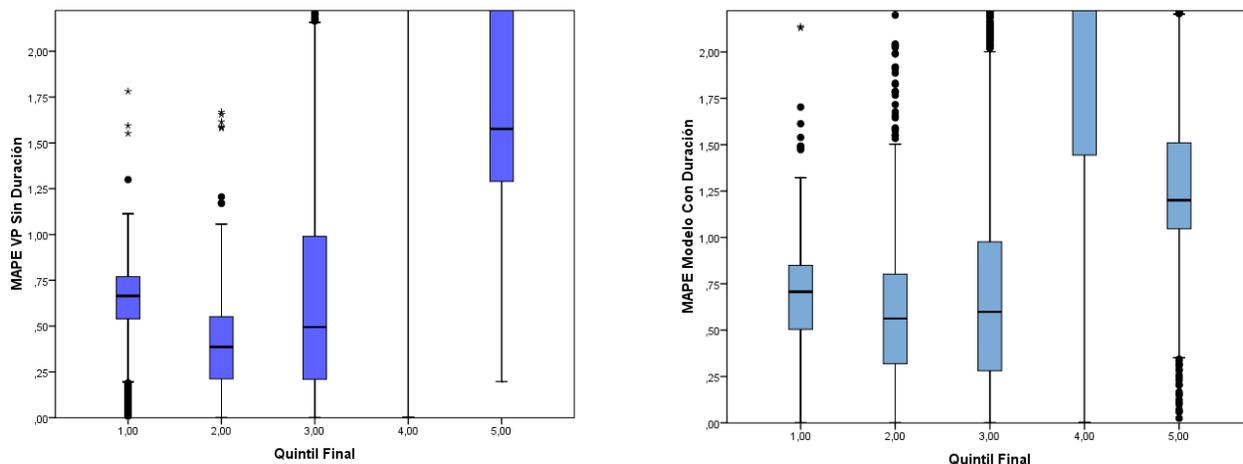


Tabla 31. Gráfico Box-Plot de cada Quintil vs. Desviación porcentual absoluta individual.

Con respecto a la Edad se observa también que el modelo Con duración es menos disperso y las medias de sus MAPEs individuales son similares a la del modelo Con duración.

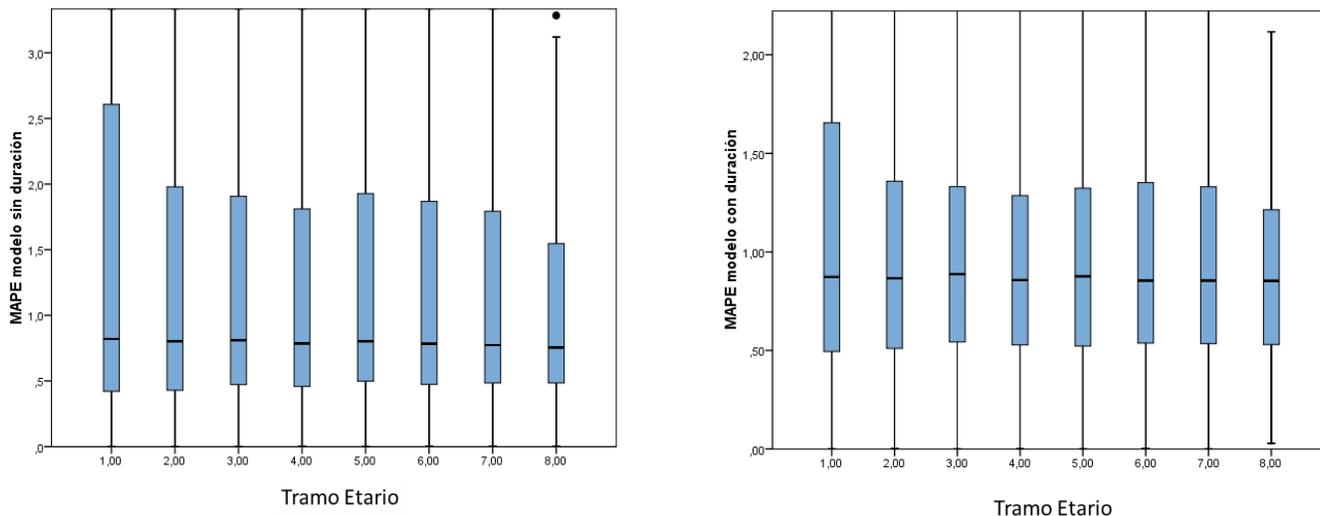


Tabla 32. Gráfico Box-Plot de cada Tramo Etario vs. Desviación porcentual absoluta individual.

## G. Conclusiones

De los resultados se debe rescatar:

- Destinar recursos de fidelización a clientes de las regiones II, III, IV y Metropolitana, pues la zona representa el 21% de la renta promedio y la renta aumenta en 0,32 veces en promedio el VP.
- Los Ejecutivos de grandes empresas representan un 99% de la renta promedio, debe ser el enfoque principal.
- Los clientes aumentan su valor a medida que aumenta su edad, y los jóvenes en promedio no generan ganancias en los primeros cuatro años, pero es conveniente retener a quienes tengan una contribución neta positiva al mes 2, ya que esta variable afecta positivamente el VP.
- Toda variable relacionada con el Cupo TDC en dólares afecta de manera negativa al VP, la renta y la duración.

## IX. RANKING Y PROPUESTAS DE FIDELIZACIÓN

### 1. Ranking

Se ordenaron los clientes en base a los resultados de las predicciones de cada modelo de mayor a menor valor final. Luego, se obtuvieron los Quintiles Finales 1 y 2, sobre los valores estimados de ambos enfoques. Por último, para cada Quintil Final estimado, se obtuvo el porcentaje de clientes de cada Quintil Final al que pertenecen realmente.

Para el modelo con duración se tiene que, de los primeros 310 clientes que se estiman que son de mayor valor (del Quintil 1), el 85% realmente es del Quintil 1 y el 93% de los clientes pertenece a los quintiles más valiosos (1 y 2). Se comete un error de Tipo II para el Quintil 1 (casos que no son de valor y son etiquetados como tal), de un 25% de los casos para el modelo con Duración y del 36 % para el modelo sin duración.

Lo mismo ocurre con un tamaño de valores ordenados de mayor a menor de 1.500 individuos. El Quintil Final 1 predicho acierta en que el 86% de esos clientes son del Quintil Final real 1 y 2.

Tamaño	Modelo	Quintil Final Predicho	Quintil Final Real					Total
			1	2	3	4	5	
310	Con duración	1	85%	8%	5%	1%	2%	100%
	Sin duración	2	64%	6%	3%	1%	26%	100%
1.500	Con duración	1	63%	23%	9%	2%	3%	100%
	Sin duración	2	58%	16%	6%	1%	19%	100%
3.485	Con duración	1	48%	23%	10%	3%	16%	100%
	Sin duración	2	35%	34%	18%	7%	7%	100%

Tabla 33. Matriz de individuos de participación de Quintiles reales en la agrupación de quintiles obtenida a partir de los valores estimados con los modelos de regresión lineal.

\*Recordatorio: Cada Quintil está compuesto por 3.485 individuos.

A partir de estos datos y los resultados de desempeño de los modelos mencionados más atrás, se opta por realizar un Ranking utilizando los resultados del modelo con duración, aunque la otra estimación es más interpretable debido a sus estimadores.

## 2. Propuestas de promoción para Fidelización

A partir de la información obtenida se proponen acciones de fidelización orientadas a los clientes principalmente Ejecutivos de altos cargos que son los que tienen mayor participación en el Quintil 1 (Ver gráfico 27), cuya edad se encuentre entre los 45 y 60 años.

Estos segmentos son padres, por lo que la idea es incluir a toda la familia como usuarios del Banco.

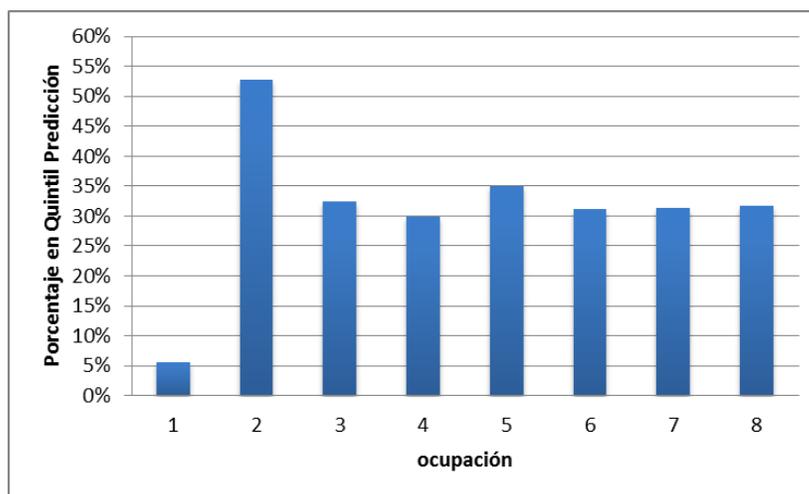


Ilustración 33. Participación por ocupación en Quintil de mayor valor predicho a partir del Modelo Con duración.

Dentro de las acciones de promoción se proponen:

- Como se mencionó al inicio de este informe, el Banco tiene distintos beneficios con diversas tarjetas de crédito. Una idea es incorporar tarjetas adicionales de débito o giftcard que pueda usar toda la familia para canjear beneficios, que los hijos porten una tarjeta y también vivan los servicios que entrega el banco en tiendas juegos, bowling, juguetes, etc.
- Un 24,2% de los clientes de mayor valor al tercer mes tiene un Crédito de Consumo (Ver Anexo F), en general del total de los clientes poseen solo un 12%. Se puede promocionar periodo de tiempo un crédito de consumo de un tipo determinado (ej. compra de un auto) y sortear la devolución de un porcentaje del crédito entre los que lo hayan solicitado. Con esto además se obtendría información adicional del objeto de los créditos. El ya posee este tipo de promoción, al suscribirse al pago automático de la colegiatura se participa en el sorteo de una anualidad gratis.
- Los regalos de Navidad se utilizan como acciones de fidelización y “captura” (no utilizan la cuenta-corriente). Actualmente se entregan regalos al 72% de los clientes del mejor quintil, el 28% restante significa el gasto de un 12% del presupuesto del Plan de Toques.

Se propone reducir la cantidad de regalos de Navidad y que sea una acción de fidelización que esté dirigida a toda la familia.

- En la sección de modelamiento se puede observar que la tenencia de LDC y su Deuda contribuyen de manera positiva al Valor del Cliente.

Se propone realizar una promoción comunitaria, inscribirse con un grupo de amigos clientes del Banco y si sus transacciones con tarjeta superan cierto número y monto, forman parte del sorteo de jugar un partido de futbol, con un jugador de la selección definido de antemano por la empresa. Esta promoción incluiría a todos los tramos etarios del grupo de alto valor.

- La duración promedio aumenta con el cupo en la tarjeta de crédito. Además, la duración está relacionada positivamente con el valor final (Quintil Final) del cliente. El Quintil 1 tiene una mayor duración promedio.

Por lo que se propone incentivar el tamaño del Cupo con invitaciones o descuentos a eventos de interés de los clientes.

- Los eventos de Golf son uno de los preferidos para los clientes de alta contribución neta. Se puede ofrecer crédito con la tarjeta para la compra exclusiva a de artículos de golf o arriendo de canchas.

Todas estas son medidas de corto plazo y consideran el alcance de 4 años del valor del cliente, se debe tener presente que lo que distinguirá a Largo Plazo a un banco de otro, será la comunicación que se tenga con los clientes, con diferentes niveles de comunicación para cada segmento de clientes. Para los clientes de mayor valor se debe tener un equipo ejecutivo que se adelante a las necesidades y requerimientos del cliente, que está acostumbrado a una pronta respuesta y le gusta saber que su dinero y gestión en buenas manos.

## X. TRABAJOS FUTUROS

- Los clientes trabajados en este proyecto fueron personas que iniciaron su relación con el banco con Cuenta corriente. No se incluyó información de la tenencia de cuentas en otros bancos. Esto podría cambiar la conducta del uso de productos de los individuos, ya que puede tener una cuenta primaria con la cual realice sus transacciones más importantes y tener una adicional en el banco. Se propone incorporar esta información en trabajos futuros.
- Se podría realizar un modelo que compita con la regresión, basado en las transacciones diarias con cheques, tarjetas de crédito y transferencia que se base en el RFM con variables sociodemográficas con el de la memoria [11].
- Se reconoce que los datos de rentabilidad, cuando se basan sólo en los productos que tiene el cliente, no proporcionan una visión completa sobre el valor del cliente a un banco. Se debe considerar la existencia de inter-relaciones entre los clientes, estos es, relación entre cónyuges, entre cliente-dueño de empresa, padre-hijo, etc. Se desea incluir solamente la primera asociación (se definirá si la segunda también) en la valoración del cliente, puesto que es la más importante al considerar las acciones de gestión comercial que se empleen para algunos clientes [9].
- Para mejor entendimiento del comportamiento del cliente es de gran utilidad la información declarada por ellos. Se sugiere incorporar variables de lealtad y satisfacción de clientes [24],[28].

- Se deberían incluir variables como el costo de cambio de banco, considerando las reglas para cerrar un contrato, las cuales afectan la conducta del cliente y alargar el horizonte de vida estudiado.
- Evaluar el impacto de las acciones de fidelización, tanto en la idoneidad de la actividad así como en la intensidad más efectiva y eficiente (cantidad de toques que se realiza al cliente). Evaluar además si la transparencia institucional y las ventajas de inversión, transparencia en cuanto a cobros y los costos de las tasas de interés son drivers de la fidelidad del cliente.

## XI. BIBLIOGRAFÍA

- [1] C.B. Bhattacharya. When customers are members: Customer retention in paid membership context. *Journal of the Academy of Marketing Science*, 26(1):31–44. (1998).
- [2] E. Ramusson. Complaints can build relationships. *Sales and Marketing Management*, 151(9):89–90. (1999).
- [3] L. G. R. Cava. IESE Business School University of Navarra. 2004. [En línea]. Disponible: [http://www.iese.edu/en/files/6\\_13439.pdf](http://www.iese.edu/en/files/6_13439.pdf). [Último acceso: 02 Julio 2012].
- [4] F. Reichheld, E. Sasser. Zero defections. Quality comes to services. *Harvard Business Review*, 1990:105–111, September–October, 2000.
- [5] Introduction to customer relationship management. <http://buttleassociates.com/doc/Chapter1CRMbook2e.pdf>
- [6] Michael Haenlein. Andreas M. Kaplan. Anemone J. Beeser. A Model to Determine Customer Lifetime Value in a Retail Banking Context. *European Management Journal*. Vol. 25, No. 3, pp. 221–234. (2007). doi:10.1016/j.emj.2007.01.004
- [7] Cristhian Herrera. CRM: gestión de las relaciones con clientes. (2006). <http://www.adictosaltrabajo.com/tutoriales/tutoriales.php?pagina=crm1>
- [8] Lista de Bancos, América del Sur, Chile. [En línea]. Disponible: <http://www.listadebancos.com/bancos-de-chile.htm> [Último acceso: 02 Julio 2012].
- [9] Haitham Salawdeh. CRM and Relationship Profitability in Banking. Corporate Technology Solutions.
- [10] Nan-Chen Hsieh. An integrated data mining and behavioral scoring model for analyzing bank customers. Department of Information Management, National Taipei College of Nursing, No. 365, Min Te Road 11257, Taipei, Taiwan, ROC
- [11] 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 (Octubre 2006). Código: 2006 B144.
- [12] Felipe Reuse. Estimación del valor de clientes fidelizados en una cadena de supermercados utilizando modelos jerárquicos bayesianos. Memoria para optar al título de ingeniero civil industrial. Santiago, Chile (Octubre 2011). Código: 2011 R319.
- [13] Hao Su-li. The customer marketing strategy of commercial banks based on customer lifetime value. International Conference on Industrial Engineering and Engineering Management. 2009. doi = 10.1109/ICIEEM.2009.5344412.
- [14] Hao Su-li. Appraisal of the Customer Lifetime Value of Commercial Banks Based on Unascertained Measurement. International Conference on Information Management, Innovation Management and Industrial Engineering. 2009. doi = 10.1109/ICIII.2009.253.

[15] Pablo Marín. Estimación de lifetime value basada en comportamiento transaccional. Memoria para optar al título de ingeniero civil industrial. Santiago, Chile (Agosto 2005). Código: 2005 M338e.

[16] Pedro Urzúa. Estimación de customer lifetime value mediante técnicas supervisadas de data mining en una empresa de retail. Memoria para optar al título de ingeniero civil industrial. Santiago, Chile (Octubre 2007). Código: 2007 Ur9.

[17] Bas. de Jong Donkers, Verhoef Martijn. Predicting Customer Lifetime Value in Multi-Service Industries. ERIM Report Series Reference No. ERS-2003-038-MKT. (Abril 2003). <http://repub.eur.nl/res/pub/325/ERS-2003-038-MKT.pdf>

[18] Robert C. Blattberg, Byung-Do Kim, Scott A. Neslin. Database Marketing: Analyzing and Managing Customers (International Series in Quantitative Marketing). (Enero 2008)

[19] V. Kumar. Chapter 29. Customer Lifetime Value. University of Connecticut . [http://www.terry.uga.edu/~rgrover/chapter\\_29.pdf](http://www.terry.uga.edu/~rgrover/chapter_29.pdf)

[20] Banco de Chile. Tarjetas.

<http://ww3.bancochile.cl/wps/wcm/connect/BancoDeChile/Personas/productos-y-servicios/tarjetas-de-credito/>

[21] Xavier Dréze y André Bonfrer. Moving from Customer Lifetime Value to Customer Equity. Working paper. University of Pennsylvania. University Park. 2005. <http://jsr.sagepub.com/content/9/2/139.refs>

[22] Sunil Gupta, Donald Lehmann y, Jennifer Stuart. Valuing Customers. Journal of Marketing Research, pp. 7-18. 2004. [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=459595](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=459595)

[23] C. Fuentelsaz Gallego. Cálculo del tamaño de la muestra. [http://www.metodologiasytecnicas.ecaths.com/archivos/metodologiasytecnicas/calculo\\_muestra.pdf](http://www.metodologiasytecnicas.ecaths.com/archivos/metodologiasytecnicas/calculo_muestra.pdf)

[24] Juan Ignacio Vélez. Antecedentes de la satisfacción de los clientes del retail. Memoria para optar al título de ingeniero civil industrial. 2012. Santiago, Chile.

[25] Diana Gropellier. Modelo para estimar el valor de vida de los clientes de un banco. Memoria para optar al título de ingeniero civil industrial. 2002. IBN: 2002 G913. Santiago, Chile.

[26] Knott, Hayes y Neslin. Next product to buy Models for cross-selling applications. JIM. 16 (3), 59-75. 2002.

[27] Kamakura, Kossar y Wedel (2004). Identifying innovators for the cross-selling of new products. Management Science, 50(8) , 1120-33.

[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, Volumen 9, N. 2. Páginas 139-155. 2006. DOI:

10.1177/1094670506293810

<http://www.anderson.ucla.edu/faculty/dominique.hanssens/content/JSR2006.pdf>

[30] Shibo Li, Baohong Sun, Ronald Wilcox. Cross-selling sequentially ordered products: an application to Consumer Banking. *Journal of Marketing Research*. 2005.

[31] Kamakura, Kossar y Wedel. "Identifying innovators for the cross-selling of new products," *Management Science* 50. 2004.

[32] Aaron Knott, Andrew Hayes, and Scott Neslin (2002), "Next-Product-to-Buy Models for Cross-Selling Applications," *Journal of Interactive Marketing*, 16 (3), 59-75.

[33] Wagner Kamakura y Michel Wedel. Exploratory Tobit Factor Analysis for Multivariate Censored Data. 2001. *Multivariate Behavioral Research*. 36 (1), 53-82.

[34] In Jae Myung. Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology* 47 (2003) 90-100. 2001.

[35] Reinartz, Werner J., and V. Kumar. The impact of customer relationship characteristics on profitable lifetime duration. 2003. *Journal of Marketing*, 67 (January), 77-99.

[36] Gupta y Lehmann. *Managing Customers as Investments*.. 2005. ISBN: 013148950.

[38] Fader, Peter y Hardie. Customer-Base valuation in a contractual setting: The perils of ignoring heterogeneity. *Marketing Science*, Vol 29, No 1, pp 85-93. 2010.

[39] Dipak, Jain y Naufel. Investigating household purchase timing decisions: A conditional Hazard Function Approach," *Marketing Science*, Vol. 10, No. 1, pp. 1-23. 1991.

[40] Earl Naumann. *Creating Customer Value*. Thomson E. P. ISBN: 0538838477. Vol. XV. pp 279. 1995.

## XII. ANEXOS

### Anexo A. Participación en Monto de colocaciones de los bancos “Mediados” del Sistema bancario chileno.

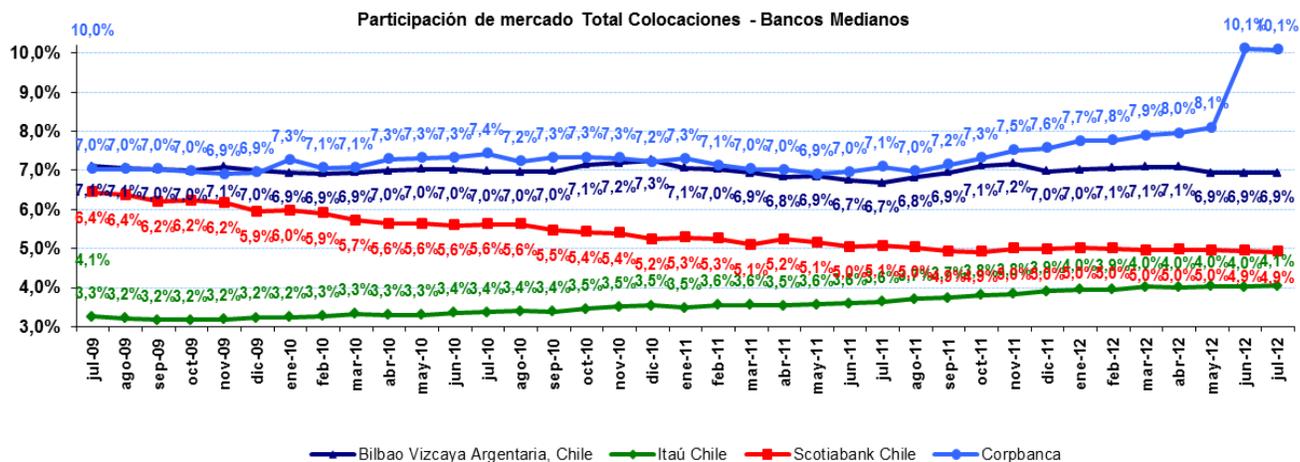


Ilustración 34. Participación en Monto de colocaciones de los bancos “Mediados” del Sistema bancario chileno. Fuente: Banco de Chile.

### Anexo B. Distribución cartera principales bancos

Las entidades Bice, Security y Corpbanca presentan de manera evidente un enfoque Empresarial, mientras que Falabella uno de Consumo. Banco Estado y Scotiabank tienen una alta proporción de su cartera orientada a Empresas y a Vivienda. Lo mismo ocurre con BBVA y Santander, pero estos últimos tienen un mayor porcentaje de Consumo. Banco de Chile e Itaú tienen una composición similar. Su diferencia radica en que Banco Chile posee menor porcentaje en Empresas y mayor en Adeudado por Bancos. Finalmente, BCI posee menor proporción en Personas, que los dos bancos anteriores y una mayor orientación a las Empresas.

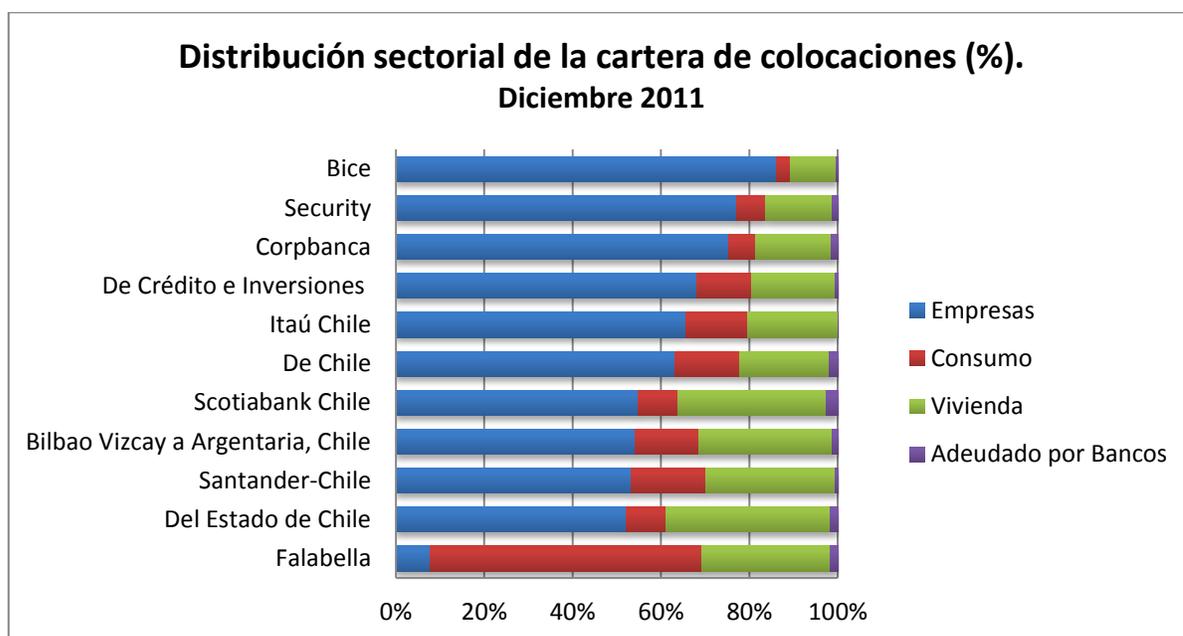


Ilustración 35. Distribución sectorial de la cartera de colocaciones (%). Diciembre 2011. Fuente: Elaboración propia, SBIF.

### Anexo C. Explicación de Quintiles de valor.

	Valor Inicial	Valor Final
<b>Quintil 1</b>	Grupo que posee el 20% de los clientes de mayor valor en los primeros tres meses	Grupo que posee el 20% de los clientes de mayor valor a lo largo de su vida máxima de 48 meses.
<b>Quintil 2</b>		
<b>Quintil 3</b>	⋮	⋮
<b>Quintil 4</b>		
<b>Quintil 5</b>	Grupo que posee el 20% de los clientes de menor valor en los primeros tres meses	Grupo que posee el 20% de los clientes de menor valor a lo largo de su vida máxima de 48 meses.

Tabla 34. Explicación breve de Quintiles de valor inicial y final.

## Anexo D. Porcentaje de toques realizados para cada grupo según Quintil Inicial.

Quintil Inicial\Quintil Final	1	2	3	4	5
1	51%	28%	15%	5%	1%
2	37%	21%	23%	18%	1%
3	53%	20%	15%	8%	5%
4	39%	21%	21%	4%	14%
5	65%	14%	12%	6%	2%

Tabla 35. Porcentaje de toques realizados para cada grupo según Quintil Inicial.

## Anexo E. Cálculo del tamaño muestral

Muestreo aleatorio. No se considera el muestreo aleatorio estratificado debido a que lo sugiere la experiencia de los expertos del banco y se plantea que se cumple la ley de los Grande Números.

$$n = \frac{N \cdot z^2_{\alpha/2} p \cdot (1 - p)}{(N - 1) \cdot e^2 + z^2_{\alpha/2} p \cdot (1 - p)}$$

Donde:

$N = 714.000$ , tamaño de la población. Asumiendo que se distribuye normalmente.

$z_{\alpha/2} = 1,64$  nivel de confianza elegido.

$p = 0,5$  proporción de una categoría

$e = 0,6$  error máximo.

## Anexo F. Asignación Zona-Comuna.

A	B	C	C	D	D	D
ALTO HOSPICIO	ALTO DEL CARMEN	ALGARROBO	SAN FERNANDO	ACHAO	LOTA	TOLTEN
ANTOFAGASTA	ANDACOLLO	CABILDO	SAN FRANCISCO DE MOSTAZAL	ALTO BIOBIO	LUMACO	TOME
ARICA	CALDERA	CALERA	SAN VICENTE	ANCUD	MAFIL	TRAIQUEN
CALAMA	CANELA	CALLE LARGA	SANTA CRUZ	ANGOL	MARIQUINA	TREGUACO
CAMARONES	CHANARAL	CARTAGENA	SANTA MARIA	ANTUCO	MAULE	TREHUACO
CAMINA	COMBARBALA	CASABLANCA	SANTO DOMINGO	ARAUCO	MAULLIN	TUCAPEL
CHUQUICAMATA	COPIAPO	CATEMU	SN VICENTE DE TAGUA TAGUA	BAQUEDANO	MELIPEUCO	VALDIVIA
COLCHANE	COQUIMBO	CHEPICA	VALPARAISO	BULNES	MOLINA	VICHUQUEN
GENERAL LAGOS	DIEGO DE ALMAGRO	CHIMBARONGO	VENTANAS	CABRERO	MULCHEN	VICTORIA
HUARA	EL SALVADOR	CHORRILLOS	VILLA ALEMANA	CALBUCO	NACIMIENTO	VILCUN
IQUIQUE	FREIRINA	CODEGUA	VINA DEL MAR	CANETE	NEGRETE	VILLA ALEGRE

MARIA ELENA	HUASCO	COINCO	ZAPALLAR	CAPITAN PASTENE	NINHUE	VILLARRICA
MEJILLONES	ILLAPEL	COLTAUCO		CARAHUE	NIQUEN	YERBAS BUENAS
OLLAGUE	LA HIGUERA	CON CON		CASTRO	NUEVA IMPERIAL	YUMBEL
PICA	LA SERENA	CURAUMA		CAUQUENES	OSORNO	YUNGAY
PLAYA BRAVA	LOS VILOS	CURIMON		CHACAO	PADRE LAS CASAS	
POTRERILLOS	MINCHA	DONIHUE		CHAITEN	PAILAHUEQUE	
POZO AL MONTE	MONTE PATRIA	EL MELON		CHANCO	PAILLACO	
POZO ALMONTE	OVALLE	EL QUISCO		CHIGUAYANTE	PALENA	
PUTRE	PAIGUANO	EL TABO		CHILLAN	PANGUIPULLI	
SAN PEDRO DE ATACAMA	PAIHUANO	GRANEROS		CHILLAN VIEJO	PARRAL	
SIERRA GORDA	PUNITAQUI	HIJUELAS		CHOLCHOL	PELARCO	
TALTAL	RIO HURTADO	ISLA DE PASCUA		CHONCHI	PELLUHUE	
TOCOPILLA	SALAMANCA	JUAN FERNANDEZ		COBQUECURA	PEMUCO	
<b>E</b>	TIERRA AMARILLA	LA CALERA		COCHAMO	PENCAHUE	
AISEN	TONGOY	LA CRUZ		COELEMU	PENCO	
ANTARTICA	VALLENAR	LA ESTRELLA		COIHUECO	PERQUENCO	
AYSÉN	VICUNA	LA LIGUA		COLBUN	PINTO	
CABO DE HORNOS	<b>G</b>	LAS CABRAS		COLLIPULLI	PITRUFQUEN	
CHILE CHICO	LA REINA	LIMACHE		CONCEPCION	PORTEZUELO	
CISNES	LAS CONDES	LITUECHE		CONSTITUCION	PUCON	
COCHRANE	LO BARNECHEA	LLAILLAY		CONTULMO	PUERTO MONTT	
	VITACURA	<b>LLO LLEO</b>		CORONEL	PUERTO OCTAY	
COYHAIQUE	<b>H</b>	<b>LOLOL</b>		CORRAL	PUERTO VARAS	
GUAITECAS	NUNOA	<b>LOS ANDES</b>		CUNCO	PUQUELDON	
LAGO VERDE	PEDRO VALDIVIA NORTE	MACHALI		CURACAUTIN	PUREN	
LAGUNA BLANCA	PROVIDENCIA	MALLOA		CURACO DE VELEZ	PURRANQUE	
NATALES	<b>I</b>	MARCHIGUE		CURANILAHUE	PUYEHUE	
NAVARINO	HUECHURABA	MARCHIHUE		CURARREHUE	QUEILEN	
O HIGGINS	LA CISTERNA	MOSTAZAL		CUREPTO	QUELLON	
O'HIGGINS	LA FLORIDA	NANCAGUA		CURICO	QUEMCHI	
O'HIGGINS	MACUL	NAVIDAD		DALCAHUE	QUILACO	
PORVENIR	MAIPU	NOGALES		EL CARMEN	QUILLECO	
PRIMAVERA	PENALOLEN	OLIVAR		EMPEDRADO	QUILLON	
PUERTO AISEN	SAN MIGUEL	OLMUE		ENTRE LAGOS	QUINCHAO	
PUNTA ARENAS	SANTIAGO	PALMILLA		ERCILLA	QUIRIHUE	
RIO IBANEZ	<b>J</b>	PANQUEHUE		FLORIDA	RANQUIL	
RIO VERDE	CERRILLOS	PAPUDO		FREIRE	RAUCO	
SAN GREGORIO	ESTACION CENTRAL	PAREDONES		FRESIA	RENAICO	
TIMAUKEL	INDEPENDENCIA	PERALILLO		FRUTILLAR	RETIRO	
TORRES DEL PAINE	PUENTE ALTO	PETORCA		FUTALEUFU	RIO BUENO	
TORTEL	QUILICURA	PEUMO		FUTRONO	RIO CLARO	
<b>F</b>	QUINTA NORMAL	PICHIDEGUA		GABRIELA	RIO NEGRO	
ALHUE	RECOLETA	PICHILEMU		GALVARINO	ROMERAL	
BUIN	SAN BERNARDO	PLACILLA		GORBEA	SAAVEDRA	

CALERA DE TANGO	SAN JOAQUIN	PUCHUNCAVI		HUALAIHUE	SAGRADA FAMILIA	
CHOCALAN	K	PUMANQUE		HUALANE	SAN CARLOS	
COLINA	CERRO NAVIA	PUTAENDO		HUALPEN	SAN CLEMENTE	
CURACAVI	CONCHALI	QUILLOTA		HUALQUI	SAN FABIAN	
EL MONTE	EL BOSQUE	QUILPUE		HUEPIL	SAN IGNACIO	
ISLA DE MAIPO	LA GRANJA	QUINTA DE TILCOCO		LA UNION	SAN JAVIER	
LAMPA	LA PINTANA	QUINTERO		LAGO RANCO	SAN JOSE DE LA MARIQUINA	
MARIA PINTO	LO ESPEJO	RANCAGUA		LAJA	SAN JOSE DE MARIQUINA	
MELIPILLA	LO PRADO	RECREO		LANCO	SAN JUAN DE LA COSTA	
PADRE HURTADO	PEDRO AGUIRRE CERDA	RENACA		LAUTARO	SAN NICOLAS	
PAINE	PUDAHUEL	RENGO		LEBU	SAN PABLO	
PENAFLORE	RENCA	REQUINOA		LICANTEN	SAN PEDRO DE LA PAZ	
PIRQUE	SAN RAMON	RINCONADA		LINARES	SAN RAFAEL	
SAN JOSE DE MAIPO		ROSARIO		LLANQUIHUE	SAN ROSENDO	
SAN PEDRO		SAN ANTONIO		LONCOCHE	SANTA BARBARA	
TALAGANTE		SAN ESTEBAN		LONGAVI	SANTA JUANA	
TIL TIL		SAN FELIPE		LONQUIMAY	TALCA	
				LOS ALAMOS	TALCAHUANO	
				LOS ANGELES	TEMUCO	
				LOS LAGOS	TENO	
				LOS MUERMOS	TEODORO SCHMIDT	
				LOS SAUCES	TIRUA	

### Anexo G. Test Anova edad vs. género.

EDAD – GÉNERO					
	Suma de Cuadrados	df	Media Cuadrática	F	Sig.
Entre grupos	2620,983	1	2620,983	24,999	,000
Dentro del Grupo	1818635,710	17346	104,845		
Total	1821256,693	17347			

### Anexo H. Frecuencia Estado Civil.

Etiqueta	Estado Civil	Frecuencia	Porcentaje
0	Sin información	7.530	43,2
1	Soltero	5.971	34,3
2	Casado	3.454	19,8
3	Separado	399	2,3
4	Viudo	70	,4
Total		17424	100,0

### Anexo I. Gráfico Box-Plot estado civil vs. Edad.

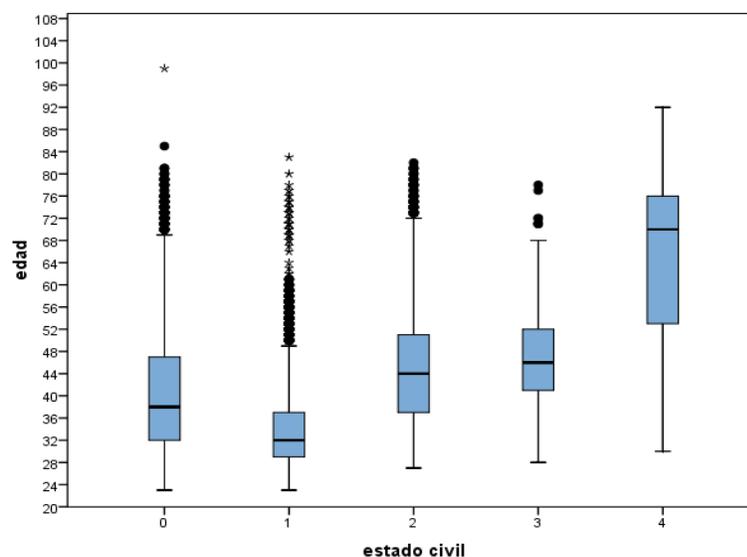


Tabla 36. Gráfico Box-Plot estado civil vs. Edad.

Anexo J. Test Anova duración vs. Quintil Final.

DURACIÓN – QUINTIL FINAL					
	Suma de Cuadrados	df	Media Cuadrática	F	Sig.
Entre grupos	1078838,447	4	269709,612	1495,470	,000
Dentro del Grupo	3141535,726	17419	180,351		
Total	4220374,174	17423			

Anexo K. Tests Anova's realizados sobre Productos de Colocación.

LÍNEA DE CRÉDITO - VP					
	Suma de Cuadrados	df	Media Cuadrática	F	Sig.
Entre grupos	54.764.906.302.310	1	54.764.906.302.310	18	0,000
Dentro del Grupo	52.900.990.056.758.900	17.422	3.036.447.598.253		
Total	52.955.754.963.061.200	17.423			
TARJETA DE CRÉDITO - VP					
	Suma de Cuadrados	df	Media Cuadrática	F	Sig.
Entre grupos	43.300.910.222.530	2	21.650.455.111.265	7	0,013
Dentro del Grupo	52.912.454.052.838.600	17.421	3.037.279.952.519		
Total	52.955.754.963.061.100	17.423			
HIPOTECARIO - VP					
	Suma de Cuadrados	df	Media Cuadrática	F	Sig.
Entre grupos	277.574.446.792.188	1	277.574.446.792.188	92	0,000
Dentro del Grupo	52.678.180.516.268.500	17.422	3.023.658.622.217		
Total	52.955.754.963.060.700	17.423			
CONSUMO - VP					
	Suma de Cuadrados	df	Media Cuadrática	F	Sig.
Entre grupos	74.131.013.400.607	1	74.131.013.400.607	24	0,000
Dentro del Grupo	52.881.623.949.660.600	17.422	3.035.336.009.050		
Total	52.955.754.963.061.200	17.423			

Tabla 37. Test ANOVA para evaluar significatividad de tenencia de productos de consumo sobre el VP.

Anexo L. Quintil de mayor valor del modelo de predicción con duración.

Tenencia de Crédito de Consumo		
	Frecuencia	Porcentaje
0	68	75,6
1	22	24,4
<b>Total</b>	<b>90</b>	<b>100,0</b>

Tabla 38. Tabla de Frecuencia de Tenencia de Consumo Mes 3 para Quintil de mayor valor del modelo de predicción con duración.