

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

**VALOR ESPERADO DEL CLIENTE
EN EL MERCADO DE SEGUROS PATRIMONIALES**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

IGNACIO DOMINGO TRUJILLO SILVA

PROFESOR GUÍA:

ALEJANDRA PUENTE CHANDÍA

MIEMBROS DE COMISIÓN:

**PABLO MARIN VICUÑA
MARCEL GOIC FIGUEROA**

**SANTIAGO DE CHILE
MARZO 2013**

VALOR ESPERADO DEL CLIENTE EN EL MERCADO DE SEGUROS PATRIMONIALES

En la actualidad, las compañías aseguradoras operaran en un mercado cada vez más competitivo. Por ejemplo, el uso de Internet ha permitido disminuir los costos de información en los que incurría un asegurado que desea cambiar de compañía. Por otra parte, como la gestión de los productos se ha realizado tradicionalmente de forma separada por áreas de negocio y para cada contrato suscrito individualmente, se han exhibido niveles de cross selling muy bajos. Como consecuencia de esto, ha surgido una creciente preocupación por distinguir, entender y fidelizar a los clientes.

El primer paso es conocer su valor, de modo de saber tanto a quienes debo retener como suscribir. En esta línea, este trabajo busca determinar el valor del cliente asociado a una compañía de Seguros Generales. La memoria considera el cálculo individual y multidimensional de la relación entre el asegurador y sus clientes. En el sentido individual, pues las tendencias modernas han permitido hacer una evaluación particular y no agregada del cliente. Multidimensional, ya que hoy se busca evaluar al cliente por su portafolio (múltiples productos) actual/potencial y no por la rentabilidad de productos particulares.

La metodología a utilizar se basa en el concepto de Customer Lifetime Value (CLV), utilizado frecuentemente en el marketing relacional. La memoria parte por analizar los datos históricos, de modo de generar perfiles explicativos de los modelos de renovación y suscripción de nuevos productos. Luego, calcula el CLV histórico, esperado y potencial, a partir de modelos Logit de Renovación y Suscripción.

Este trabajo concluye que la variable que más condiciona la no renovación es *Tener un siniestro* (experiencia traumática). Por otro lado, las variables que más lo incentivan son la *Antigüedad del Cliente* y *Tener una póliza del ramo Incendios*. La variable *Prima Directa* va en esta misma línea, la evidencia asegura que los clientes que cuentan con productos personalizados (*Primas Directas* diferenciadas) tienden a permanecer mayor tiempo en la compañía.

En términos de valor del cliente se observa que para grupos de *Edad* mayores el valor del cliente aumenta, este aumento en el valor está explicado por un aumento en la permanencia de dichos clientes. El *Género* femenino en prácticamente todos los modelos aumenta el valor esperado, sólo el modelo asociado al ramo Incendio deja a esta variable fuera de la ecuación.

Finalmente, se recomienda mejorar este trabajo con un cambio en la unidad de análisis. Hoy el objeto de estudio debe ser el hogar y no el contratante, los últimos trabajos en la materia se han centrado en la familia como tomador de productos de seguro, pues todos los miembros adultos participan en la decisión de la cobertura de riesgos comunes como son el hogar, el automóvil u objetos valiosos que deseen proteger. Sin embargo, la existencia de múltiples contratos asociados a un mismo hogar complejiza drásticamente el análisis

Tabla de contenido

1.	INTRODUCCIÓN.....	7
2.	DESCRIPCIÓN DEL PROYECTO Y JUSTIFICACIÓN	10
2.1.	Rentabilidad a nivel de póliza y producto.....	11
2.2.	Horizontes de evaluación muy cortos	12
2.3.	Venta a través de intermediarios	12
3.	OBJETIVOS	14
3.1.	Objetivo General	14
3.2.	Objetivos Específicos.....	14
4.	METODOLOGÍA.....	15
4.1.	Metodología para estimar Renovación y Suscripción de Productos.....	15
4.2.	Metodología para estimar Siniestralidad (Riesgo)	15
4.3.1.	Selección de Muestra y creación de variables iniciales:.....	16
4.3.2.	Variables y conclusiones de la muestra	16
4.3.3.	Herramientas.....	16
4.4.	Variables Explicativas	17
4.4.1.	Determinación de variables socio-demográficas y transaccionales significativas para los modelos	17
4.4.2.	Herramientas.....	17
4.5.	Modelo	17
	Modelos Logísticos:	17
	El alcance de esta sección, está acotada a correr los modelos logísticos asociados a la probabilidad de renovación y suscripción. Lo central de este capítulo es el ajuste empírico realizado a los modelos.....	17
	Finalmente, se realiza un análisis e interpretación del vector de coeficientes y de las variables significativas y no significativas.....	18
4.6.	CLV.....	18
4.6.1.	Resultados Agregados.....	18
4.6.2.	Herramientas.....	18
4.7.	Perfiles y Recomendaciones de Marketing	18
5.	MARCO CONCEPTUAL.....	19
6.	MODELO Y NOTACION.....	20

6.1.	Valor Histórico del Cliente:	20
6.2.	Valor Esperado del Cliente:	21
6.3.	Valor Potencial del Cliente:.....	22
6.4.	Probabilidades de Modelos:	23
6.4.1.	Probabilidad de Renovación:.....	23
6.4.2.	Probabilidad de Suscripción	24
6.5.	Modelo Logit:	24
7.	ALCANCES	26
8.	RESULTADOS ESPERADOS	27
9.	ANÁLISIS DE DATOS Y ESTADÍSTICA DESCRIPTIVA.....	28
9.1.	Estadística Descriptiva.....	32
9.1.1.	Número de Pólizas Vigentes:	32
9.1.3.	Monto de primas pagadas.....	37
9.1.4.	Cantidad de Siniestros	38
9.1.5.	Monto de Siniestros Pagados	39
9.2.	VARIABLES A EXPLICAR	40
9.2.1.	Tasa de Renovación (Modelo Renovación):	40
9.2.2.	Cross Selling o tenencia de productos (Modelo Suscripción):	41
9.2.3.	Permanencias (Duración).....	42
10.	TRANSFORMACIÓN, AGRUPACIÓN Y SELECCIÓN DE VARIABLES	47
10.1.	Cluster y Categorización de Variables:.....	47
10.1.1.	Edad:	47
10.1.2.	Género	48

10.1.3. Antigüedad:	48
10.1.4. Número de Productos Contratados:	49
10.1.5. Prima Directa:.....	50
10.1.6. Monto de Siniestro:.....	52
10.1.7. Cantidad de Siniestros:	54
10.2. Test ANOVA	56
11. RESULTADOS Y VALIDACIÓN DE MODELOS	58
11.1. Modelo de Renovación	58
Renovación Vehículos	59
Renovación Incendio	64
Renovación Varios.....	68
11.2. Modelo de Suscripción.....	71
Suscripción Vehículos.....	72
Suscripción Incendio.....	74
Suscripción Varios	75
12. VALOR ESPERADO DEL CLIENTE PARA DIFERENTES PERFILES:	78
13. CONCLUSIONES GENERALES	85
14. CITAS BIBLIOGRÁFICAS	88
ANEXOS	89
ANEXO A: FRECUENCIAS DE MUESTRA	89

Índice de Tablas

Tabla 1: Estadísticas Socio demográficas – Corte 2009	29
Tabla 2: Estadísticas Permanencia	30
Tabla 3: Variables de interés base de datos	31
Tabla 4: Variables de interés base de datos	32
Tabla 5: Variables de interés base de datos	33
Tabla 6: Número de Pólizas en vigor de Varios	33
Tabla 7: Promedio Prima Directa por Ramo (Cifras en UF)	38
Tabla 8: Promedio Prima Directa por Ramo (Cifras en UF)	39
Tabla 9: Promedio Costo de Siniestros por Ramo (Cifras en UF)	40
Tabla 10: Permanencias por Ramo	43
Tabla 11: Conglomerados Edad para Modelo “Renovación de una póliza de Vehículos”	47
Tabla 12: Frecuencia de Permanencia por Cliente	48
Tabla 13: Frecuencia de Número de Pólizas por Cliente	49
Tabla 14: Conglomerados Prima para Modelo “Renovación de una póliza de Vehículos”	52
Tabla 15: Test Anova para la variable independiente Edad	57
Tabla 16: Modelo logístico para el ajuste de la probabilidad de renovar una póliza de Vehículos si estaba contratada el año anterior	59
Tabla 17: Modelo logístico para el ajuste de la probabilidad de renovar más de una póliza de Vehículos si estaban contratadas el año anterior	62
Tabla 18: Modelo logístico para el ajuste de la probabilidad de renovar	64
Tabla 19: Modelo logístico para el ajuste de la probabilidad de renovar más de una póliza de Incendio si estaban contratadas el año anterior	66
Tabla 20: Modelo logístico para el ajuste de la probabilidad de renovar una póliza de Varios si estaba contratada el año anterior	68
Tabla 21: Modelo logístico para el ajuste de la probabilidad de renovar más de una póliza de Varios si estaban contratadas el año anterior	70
Tabla 22: Modelo logístico para el ajuste de la probabilidad de Suscribir una póliza de Vehículos si tenía al menos una póliza de Incendio o Varios	72
Tabla 23: Modelo logístico para el ajuste de la probabilidad de Suscribir una póliza de Incendio si tenía al menos una póliza de Vehículos o Varios	74
Tabla 24: Modelo logístico para el ajuste de la probabilidad de Suscribir una póliza de Varios si tenía al menos una póliza de Vehículos o Incendio	76
Tabla 25: Valor Histórico Promedio al año 2009	78
Tabla 26: Valor Esperado para Renovación de una o más póliza de Vehículos	79
Tabla 27: Valor Esperado para Renovación de una o más pólizas de Incendio	80
Tabla 28: Valor Esperado para Renovación de una o más pólizas de Varios	81
Tabla 29: Valor potencial de Suscripción de una póliza de Vehículos	82
Tabla 30: Valor potencial de Suscripción de una póliza de Incendio	83
Tabla 31: Valor potencial de Suscripción de una póliza de Varios	83
Tabla 32: Valor del Cliente Esperado más Potencial	84

Índice de Gráficos e Ilustraciones

Gráfico 1: Ciclo de Vida del Cliente	8
Gráfico 3: Distribución de Estado Civil para cortes 31/12/2009, 31/12/2010 y 31/12/2011	36
Gráfico 4: Distribución de Ubicación y Ramo para cortes 31/12/2009, 31/12/2010 y 31/12/2011	37
Gráfico 5: Tendencia de la Renovación de Prima por Ramo	41
Gráfico 6: Tendencia de Cross Selling por Ramo	42
Gráfico 7: Siniestralidad por permanencia.....	44
Gráfico 8: Participación por cluster de siniestralidad (clientes que dejan la Cía)	45
Gráfico 9: Clientes que dejan la compañía por Siniestralidad (LR).....	46
Gráfico 10: Histograma de Número de Pólizas por Cliente y Ramo	49
Gráfico 11: Histograma de Frecuencia de Prima Directa (UF)- Ramo Vehículos.....	50
Gráfico 12: Distribución de Frecuencia de PD (UF) - Incendio	51
Gráfico 13: Histograma de Frecuencia de Prima Directa (UF)- Ramo Varios	51
Gráfico 14: Histograma de Frecuencia de Costo de Siniestro- Ramo Vehículos.....	53
Gráfico 15: Histograma de Frecuencia de Costo de Siniestro- Ramo Incendio	53
Gráfico 16: Histograma de Frecuencia de Costo de Siniestro- Ramo Varios.....	54
Gráfico 17: Frecuencia de Cantidad de Siniestros - Ramo Vehículos	54
Gráfico 18: Frecuencia de Cantidad de Siniestros - Ramo Incendio	55
Gráfico 19: Frecuencia de Cantidad de Siniestros - Ramo Varios	55
Gráfico 20: Curva COR para modelo logístico para el ajuste de la probabilidad de renovar una póliza de Vehículos si estaba contratada el año anterior	61
Gráfico 21: COR – Vehículos más de una Póliza	63
Gráfico 22: COR – Incendio de una Poliza	66
Gráfico 23: Gráfico COR – Incendio más de una Póliza	67
Gráfico 24: COR – Varios una Póliza	69
Gráfico 25: Gráfico COR – Varios una Póliza	71
Gráfico 26: COR – Suscripción Vehículo	73
Gráfico 27: COR – Suscripción Incendio	75
Gráfico 28: COR – Suscripción Varios.....	77

1. INTRODUCCIÓN

Las compañías de seguros generales en Chile actúan en un mercado con baja concentración¹ y alto grado de competencia. El paradigma clásico de las compañías de seguros aseguraba que “un cliente que conocían era un cliente que no querían retener”, pues normalmente el cliente que conocían necesitaba reportar un siniestro. Clásicamente ese era el único punto de contacto² entre el asegurado y la compañía, contacto que no era del todo grato para ninguno de los implicados. Es decir, existía un incentivo a que existiera un escasísimo estudio del cliente (Guillén, 2009), la interacción era mínima, considerando que la mayor parte de la venta se realiza través de corredores (intermediarios).

El uso de internet y la disminución de las barreras de información han estimulado el aumento de la venta directa (sin intermediarios). Esto ha creado la necesidad de conocer al asegurado, orientado la gestión aseguradora hacia el cliente-consumidor dado que las principales dificultades, no sólo se dan en la creación y lanzamiento al mercado de productos equilibrados técnicamente, sino en conseguir su éxito comercial (Guillén, 2009).

Los principales efectos esperados de un mayor conocimiento del cliente, son por ejemplo: la *disminución de los costos de adquisición* dado que el pago de comisiones a corredores es el mismo en la venta nueva como en la renovación, siendo que en la renovación el cliente puede ser capturado por venta directa; un *aumento de los ingresos* como consecuencia de la extensión del tiempo de relación con el cliente y una *disminución de la siniestralidad* gracias a la segmentación y tarificación de la cartera ajustada por riesgo asociadas a las características del asegurado.

Las compañías de seguros se ven obligadas a invertir importantes sumas para mantener el flujo de adquisición de clientes, siendo que estos clientes alcanzan un mayor potencial de

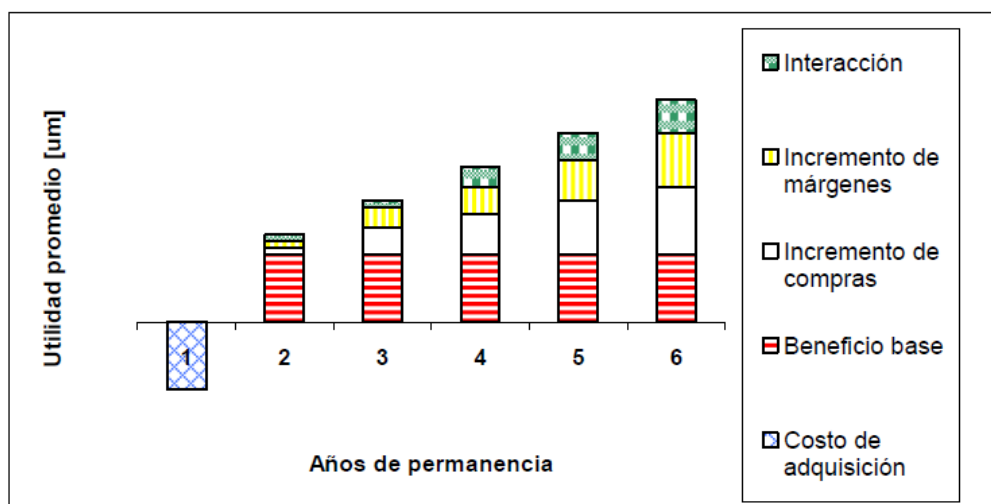
¹ Concentración de mercado medida según el Índice de Herfindahl alcanzó 0,085 puntos, que según la normativa Chilena se considera baja concentración, pues es menor a 0,1 (Fuente: Generación Propia. Véase gráfico en anexo III).

² Clásicamente las compañías de seguros generales han realizado su venta a través de corredores de seguros, lo que genera una mayor distancia entre las compañías y los clientes.

rentabilidad a mayores años de permanencia. Este hecho es especialmente importante en la industria de los seguros generales, pues la gama de productos ofertada al cliente no es muy amplia, de modo que el driver de rentabilidad se basa en la duración de la relación con el cliente y su siniestralidad³. Es decir, existe poco espacio para el incremento de compras, prácticamente con 4 o 5 productos un cliente alcanza su potencial máximo.

A medida que transcurre el tiempo y la compañía gana experiencia con sus clientes, se hace más eficiente el servicio que la compañía provee a ellos. Estos clientes al sentirse satisfechos, darán buenas referencias de la compañía a otros potenciales clientes. Estas referencias son centrales en el mercado asegurador, pues las compañías aseguradoras basan su negocio en la confianza.

Gráfico 1: Ciclo de Vida del Cliente



Fuente: Zero defections: Quality comes to service. Harvard Business Review, Sept - Oct 1990

El Gráfico 1 detalla las ganancias percibidas a medida que aumenta la permanencia. En esta línea Rayals y Know (2005) concluyen que un aumento en la retención de los clientes de un 85% al 90% incrementaba la utilidad como mínimo en un 35% y podía llegar a ser del 95% para las compañías estudiadas.

³ Siniestralidad: Proporción de la prima que se destina al pago de siniestros

En dicha grafica se observa la utilidad genérica para cada año de permanencia de un cliente, en él se distingue:

- Costos de intermediación: Son las comisiones pagadas a la fuerza de venta propia o externa, gastos de campañas, etc.
- Beneficio base: corresponde a la utilidad reportada por los primeros productos adquiridos por el cliente.
- Incremento de ventas: como dice el concepto corresponde al cross o up-selling de un mismo cliente.
- Incremento de márgenes: la experiencia adquirida con el cliente ayuda a mejorar la interacción con él, siendo más eficiente en los *servicios prestados* como en el *proceso de renovación*, donde puede retener al cliente sin necesidad de pagar la intermediación al corredor.
- Interacción: Los clientes que han creado confianzas recomiendan a la compañía, desarrollando el mejor marketing posible para una aseguradora, la recomendación directa. La recomendación es especialmente importante en este mercado, dado que la diferenciación está dada por el servicio prestado en caso de siniestro

En este mismo sentido, cuando un cliente queda insatisfecho éste realiza en promedio 11 opiniones negativas respecto de la compañía a potenciales clientes (Kotler 1990). Debido a esto, la importancia de conocer sus necesidades tanto para controlar al cliente insatisfecho como para ofrecer los bienes correctos a cada cliente.

En este entorno es fundamental revisar el grado de vinculación del cliente con la empresa y ser capaces de mantenerlo fiel. El objetivo es transformar al cliente de valor real o potencial, en un cliente rentable y fiel. Para conseguir lo anterior se debe gestionar toda la información que se tiene del cliente, detectar el grado y las causas de su vinculación y conocer el valor que aporta para la compañía para, finalmente, tomar la decisión de fidelizar o prescindir de él.

2. DESCRIPCIÓN DEL PROYECTO Y JUSTIFICACIÓN

El enfoque tradicional de suscripción basa su evaluación en la rentabilidad por póliza, valoriza el contrato individual que existe entre el contratante y la compañía. Con la misma perspectiva, el análisis agregado se hace por líneas de productos (Ramos: Vehículos, Incendio y Terremoto, SOAP y Varios), considerando en ningún momento el valor del cliente con la diversidad de productos que podría tener.

El modo habitual de venta practicado por las compañías de seguros en Chile se ha basado en la intermediación a través de un corredor, quien hace el papel de fuerza de venta y contraparte permanente para el cliente. Este modelo de venta contempla muy pocos puntos de contacto entre el cliente y la aseguradora.

Finalmente, dado que la duración de los contratos es normalmente igual a un año, el horizonte temporal de análisis para la renovación de una póliza se hace para el mismo periodo. La decisión de renovación, cancelación, aumento o disminución de prima, se estructura en base al estudio de la última póliza contratada, sesgo que muchas veces tiene un importante costo en términos de la fidelidad de los clientes⁴.

En resumen, el enfoque tradicional tranza y rentabiliza contratos anuales (pólizas) y no a sus clientes. En este sentido, existen 3 elementos críticos para el análisis moderno: *rentabilidad a nivel de póliza o producto, venta a través de intermediarios y horizontes de evaluación muy cortos.*

Este trabajo considera entre sus objetivos cambiar dicho foco de análisis y centrarlo en el estudio del cliente. Busca rentabilizar la cartera del cliente, crear herramientas para vender directamente y considerar las permanencias históricas del cliente como horizontes de evaluación. La justificación asociada al cambio de enfoque se detalla a continuación:

⁴ Dependiendo del horizonte de análisis los indicadores de siniestralidad y cobranzas cambian. Por ejemplo, un cliente durante 5 años no tuvo siniestros ni morosidad, pero éste último año choco tres veces. Este cliente recibirá una evaluación negativa para su renovación, siendo que su historia completa es positiva.

2.1. Rentabilidad a nivel de póliza y producto

Normalmente las compañías seguros suben la prima de renovación o derechamente deciden no renovar un contrato, que cierra la anualidad con alta siniestralidad o con un número de siniestros alto. En ningún momento se considera la cartera de productos que tiene ese cliente y las consecuencias que podría tener el no renovar esa única póliza.

Caso concreto:

Francisco Javier Molinet entro a la compañía el año 2007 cuando compro por medio del corredor Falabella un seguro de auto, el año 2009 adquirió un auto para su esposa y otro para su hijo, asociado a dichos autos agregó dos seguros a su cartera. Ese mismo año compro su casa con un crédito hipotecario, por lo tanto estuvo obligado a tomar un seguro de Incendio y por supuesto lo tomo con la misma compañía que aseguraba sus autos.

El año 2010 Francisco decide renovar el seguro de su esposa tomado el 2009, pero le informan que no renovarían su contrato, pues registro una siniestralidad del 113%. Francisco muestra su descontento pero le informan que es una decisión del departamento de riesgo.

Lo que no esperaba la compañía es que Francisco cancelara sus otros dos seguros de autos y el seguro de Incendio. El departamento de riesgo se defendió argumentando que era un mal cliente, lo que era correcto en el enfoque tradicional. Pero si Francisco habría sido evaluado por sus 3 contratos de autos agrupadamente o más aun, por sus contratos de autos e incendio, la siniestralidad habría sido de un 35% y 28% respectivamente⁵. Es decir, las rentabilidades de la cartera de Francisco eran excelentes pero se perdieron por análisis centrados en el contrato y no en el cliente.

Por lo tanto, este trabajo estudia el comportamiento de los asegurados con uno o múltiples contratos en lo que respecta a las decisiones de renovación y suscripción de nuevos productos, mejorando así la cuantificación de su valor histórico, esperado y potencial.

Finalmente, como un mismo cliente puede tener pólizas de diferentes ramos (productos en seguros), entonces también estamos rompiendo con el paradigma que analiza la rentabilidad por producto.

⁵ Siniestralidades sobre un 70% se consideran deficientes.

2.2. Horizontes de evaluación muy cortos

Usualmente en Chile los Departamentos de Riesgo ocupan horizontes de un año para evaluar los ajustes de prima en la renovación de las pólizas. Esta práctica se ha justificado por la dificultad que genera a las compañías de seguros construir las cadenas de contratos (renovaciones) asociadas a un cliente, dificultad asociada a que los corredores muchas veces no distinguen si el nuevo contrato es una renovación o venta nueva. Los avances en programación y construcción de bases de datos han permitido construir cadenas de renovaciones y aumentar los horizontes de análisis de riesgos.

Un ejemplo práctico que grafica la potenciabilidad del nuevo enfoque, es reflejada en el caso de Francisco que se expuso en el apartado anterior. Las siniestralidades de 35% y 28% que calculamos para las carteras agregadas, bajarían a 23% y 26% respectivamente si se usara un horizonte de evaluación de tres años. Por lo tanto, se considera central para este trabajo usar horizontes más largos.

En términos prácticos, este trabajo considerará una extensión de 5 años de vida asociada a los clientes de la compañía de seguros en cuestión.

2.3. Venta a través de intermediarios

La vieja guardia del mercado de los seguros generales en Chile, aseguraría que todos los argumentos expuestos no tienen mayor importancia, pues el corredor es el encargado de fidelizar y contactar al cliente. Por lo tanto, los deseos de conocer y fidelizar no tienen sentido si es una actividad externalizada.

La posición contraria, plantea que los aumentos en la venta a través de sucursales e internet, han acrecentado en los últimos años la venta directa (sin intermediarios). Las proyecciones de la industria aseguran que debería seguir aumentando (Hurtado, 2010), debido a la alta penetración en el uso de internet de la población más joven. Eso sí, dicha demanda no se traspasará al mercado asegurador en el corto plazo, pues los clientes más jóvenes todavía no tienen la edad en que el seguro se convierte en una necesidad (Ej: Hipotecarios, Resp. Civil, Vehículos, etc.).

Por otra parte, las compañías de seguro pagan comisiones a los corredores sin distinguir entre renovaciones o venta nueva. Por tanto, existe la oportunidad para que el cliente capte al asegurado en el momento de la renovación o pague comisiones diferenciadas por perfiles más convenientes para la compañía.

Todo esto sólo es posible si la compañía de seguros conoce las características y comportamiento de sus clientes, no sólo fidelizar por fidelizar, sino más bien distinguir y fidelizar.

Finalmente, se debe destacar que fidelizar y distinguir a los clientes es hoy uno de los desafíos más importantes a los que se enfrentan las compañías de seguros, considerando el nivel de competitividad anteriormente descrito y el sesgo que han producido históricamente los intermediarios (corredores). El análisis del valor del cliente es el punto de partida en una carrera por conocer a los asegurados, la identificación de los más rentables permitirá que la compañías estrechen sus lazos, de modo de aumentar la probabilidad de sobrevivencia como de contactarlos para acceder a ellos de forma directa.

3. OBJETIVOS

3.1. Objetivo General

Determinar el valor de los cliente asociados a una Cía. de Seguros Generales

3.2. Objetivos Específicos

- Dar una visión global del mercado de los seguros generales a través de la estadística descriptiva.
- Interpretar la sensibilidad e incentivos que generan las variables explicativas.
- Segmentar de acuerdo a las variables relevantes.
- Determinar el valor esperado y potencial de los clientes
- Brindar recomendaciones de mejoras para apoyar la forma de toma de decisiones en retención y suscripción
- Recomendar mejoras y posibles estudios futuros.

4. METODOLOGÍA

Para cumplir con los objetivos propuestos dividiremos el análisis en dos líneas de investigación:

4.1. Metodología para estimar Renovación y Suscripción de Productos

Se basa en el Customer Lifetime Value (CLV), métrica utilizada frecuentemente en el marketing relacional. La memoria parte por analizar los datos históricos, de modo de generar perfiles explicativos de los modelos de renovación y suscripción de nuevos productos. Luego, se calcula CLV histórico, esperado y potencial, a partir de los modelos de probabilidad de renovación y suscripción.

4.2. Metodología para estimar Siniestralidad (Riesgo)

Este aparatado se abordará desde un punto de vista puramente bibliográfico, ya que no es el foco de la presente memoria estudiar los modelos de siniestralidad y de gestión del riesgo. Es decir, se ocuparan estándares aportados por el departamento de riesgo de la compañía estudiada.

Más aun, históricamente la industria aseguradora ha estudiado considerablemente la optimización del riesgo. En toda compañía de seguros existe un departamento de riesgo y actuariado especializado en la tarificación y optimización de la siniestralidad. Eso sí, cada vez que se haga uso de los modelos o estadística ad hoc se explicitará la metodología o la referencia a ella.

A continuación se detallan los pasos metodológicos para el desarrollo de la presente memoria:

4.3. Datos

4.3.1. Selección de Muestra y creación de variables iniciales:

Se toman cinco cortes, cada uno al 31 de diciembre de cada año desde el 2007 al 2011. Sólo se consideran los vigentes en el corte del 2009 y se estudia su comportamiento en los demás.

La actividad más dura está asociada al procesamiento, limpieza y transformación de datos, de modo que sea aplicable el modelo a considerar. En la compañía estudiada no existe un enfoque en el cliente sino en el contrato, por lo tanto la data fue modificada y reestructura a nivel de cliente.

A partir de las reestructuraciones de datos propuesta, se crearan variables determinantes para el modelo, como: antigüedad, aumentos de prima, binarias de siniestros, etc.).

4.3.2. Variables y conclusiones de la muestra

Este apartado tiene por objetivo afinar las variables creadas en el punto anterior e investigar la pertinencia de agregar otras a la muestra. Además, de realiza un análisis descriptivo de dichas variables

Finalmente, se explican alcances y supuestos generales de la base de datos.

4.3.3. Herramientas

- Data Warehouse (Business Objects – SAP)
- Microsoft Excel 2007
- Microsoft Access
- MATrix LABoratory (MATLAB R2010a)

4.4. Variables Explicativas

4.4.1. Determinación de variables socio-demográficas y transaccionales significativas para los modelos

El objeto de esta sección es la determinación de las variables socio-demográficas a incorporar y descartar en la metodología del cálculo de CLV y duración a nivel de clientes. Se inicia con la aplicación del test Anova sobre las posibles variables para determinar si éstas son significativas estadísticamente. Luego, las variables categorizables se agrupan bajo el método de K-means.

Se realiza un completo análisis de las variables explicativas detallando su estadística descriptiva relacionada a las variables explicativas. Finalmente, se exponen las conclusiones, alcances y supuestos generales asociados a las variables explicativas.

4.4.2. Herramientas

- Microsoft Excel 2007
- Microsoft Access 2007
- SPSS for Windows 15 (2006)
- MATrix LABORatory (MATLAB R2010a)

4.5. Modelo

Modelos Logísticos:

El alcance de esta sección, está acotada a correr los modelos logísticos asociados a la probabilidad de renovación y suscripción. Lo central de este capítulo es el ajuste empírico realizado a los modelos.

Finalmente, se realiza un análisis e interpretación del vector de coeficientes y de las variables significativas y no significativas.

4.5.1. Herramientas

- SPSS for Windows 15 (2006)
- MATrix LABoratory (MATLAB R2010a)
- Microsoft Excel 2007

4.6. CLV

4.6.1. Resultados Agregados

- Agregar resultados a partir de resultados bibliográficos.
- Análisis e interpretación de la comparación de modelos.
- Conclusiones y discusión de resultados.

4.6.2. Herramientas

- Microsoft Word y Excel 2007

4.7. Perfiles y Recomendaciones de Marketing

- Análisis e interpretación de perfiles de clientes.
- Definir reglas de negocio asociadas a la fidelización y suscripción.
- Recomendaciones y desarrollos futuros.

5. MARCO CONCEPTUAL

Para cuantificar el valor del cliente asociado a una compañía de seguros generales este trabajo usa el CLV, el cual fue definido hace aproximadamente 30 años por Kotler (1974), como “el valor actual de los beneficios que se espera obtener a lo largo del tiempo que vaya a durar la relación con el cliente”.

En términos más modernos se define como la suma de los flujos descontados que un consumidor, o un segmento de consumidores, genera durante su relación con la empresa (Berger y Nasr, 1998), o dicho de otro modo, “el valor presente neto de los beneficios ligados a cada consumidor, una vez que este ha sido adquirido, menos los costes asociados a cada consumidor a lo largo de la vida de dicho consumidor” (Dywer, 1997; Blattberg, 2008). La necesidad asociada a negocios específicos, generó que se fuera complejizando, por ejemplo Haenlein (2007), utiliza un modelo basado en una combinación de una cadena de markov de primer orden y un árbol de clasificación y regresión (CART) para definir los posibles estados probabilísticos en que se puede encontrar un cliente.

Donkers (2007) realiza una comparación de los modelos formulados para el cálculo del CLV. En términos simples compara la capacidad predictiva de los modelos cuando son aplicados al mercado asegurador. El trabajo de Donkers es uno de los más complejos en el cálculo del CLV en este mercado. Donkers concluye que los modelos sencillos proporcionan buenas predicciones del CLV y que los más complejos consiguen mejoras marginales. Se usa como guía los papers de Donkers (2007) en el sentido de centrar el análisis en dos modelos logísticos (Renovación y Suscripción de clientes) y de Guillén (2009) que usa la perspectiva de Donkers en un caso aplicado a España y que agrega modelos de sobrevivencia de Cox como complemento.

6. MODELO Y NOTACION

La perspectiva del CLV que usa Guillén y que se desarrolla, distingue tres tipos análisis: el valor histórico, el esperado y el potencial:

6.1. Valor Histórico del Cliente:

Representa la suma de los márgenes históricos aportados por el cliente en todos los productos que ha tenido suscritos con la compañía. No importa aquí qué ramos tenga suscritos ni las características de las carteras que forme parte. Este enfoque es puramente descriptivo, pues no aporta información de forma directa al CLV, centra el análisis en los resultados históricos y no en los esperados o potenciales.

$$CLV_{i,t} = \sum_{\tau=0}^T Profit_{i,(t-\tau)}$$

$$Profit_{i,t} = \sum_{j=1}^J Ramo_{i,j,t} Uso_{i,j,t} Margen_{i,j,t}$$

$Ramo_{i,j,(t-\tau)}$ = *Dummy* indicando 1 si el cliente i tiene el ramo j en el tiempo $(t - \tau)$

$Uso_{i,j,t}$ = *Cantidad de productos* del ramo j que tiene el cliente i en el tiempo $(t - \tau)$

$Margen_{i,j,t}$ = *Margen del ramo j* en el tiempo $(t - \tau)$ para el cliente i

Este análisis recoge una visión agregada para todos los ramos y anualidades suscritas por el cliente.

6.2. Valor Esperado del Cliente:

Aporta el beneficio esperado que tiene el cliente asociado a los productos que tiene suscritos hoy. Para ello, el valor esperado debe tener en cuenta dos factores: la probabilidad que renueve su póliza y el beneficio esperado en la cartera para el tipo de póliza que tiene suscrita. El comportamiento histórico es relevante en cuanto ayuda a elaborar un pronóstico. Pero no es relevante como valor individual, lo determinante es el comportamiento en la siguiente anualidad.

Se realiza un análisis relacionado a la siguiente anualidad, es decir que esperamos que ocurra en el siguiente periodo. En este caso no existe un aumento de pólizas, pues ese caso corresponde al Valor Potencial. En consecuencia, se analizará sólo que el cliente renueve o cancele.

Interesa pronosticar si la póliza será renovada en el año siguiente, condicionado a que está contratada el año anterior.

$$\Pr (Re_{i(T+1),j,l} = 1 | Re_{i,T,j,l} = 1)$$

Donde:

$$Re_{i,T,j,l} = \begin{cases} 1 & \text{si en } T, \text{ el individuo "i", posee la póliza "l" para el ramo "j"} \\ 0 & \text{si no} \end{cases}$$

$$E(CLV_{i,j,T+1}) = E(Profit_{i,j,T+1}) \cdot \Pr (Re_{i(T+1),j,l} = 1 | Re_{i,T,j,l} = 1)$$

Si el cliente renueva la póliza, entonces se espera obtener un beneficio que corresponde al promedio de los beneficios obtenidos por los clientes con el ramo “j” en el año anterior

$$E(Profit_{i,j,T+1}) = \frac{1}{M_{T,k}} \sum_{i=1}^{M_{T,j}} Ramo_{i,j,T} Uso_{i,j,T} Margen_{i,j,T}$$

Donde:

$$M_{T,j} = \text{Numero de clientes con el ramo "j" en el tiempo "T"}$$

Finalmente, el valor esperado asociado al cliente "i", es la suma del valor esperado de sus ramos multiplicado por la probabilidad que mantenga cada una de sus pólizas.

$$E(CLV_{i,T+1}) = \sum_{j=1}^J E(CLV_{i,j,T+1}) = \sum_{j=1}^J E(Profit_{i,j,T+1}) \cdot \Pr(Re_{i(T+1),j,l} = 1 | Re_{i,T,j,l} = 1)$$

6.3. Valor Potencial del Cliente:

Aquí se calcula el valor potencial asociado a que el cliente adquiera más pólizas. Es decir, es necesario contar con la probabilidad que el cliente adquiera una nueva póliza como también con el beneficio que aportan dichos tipos de seguros a la compañía.

Se tendrá en cuenta sólo la posibilidad que un cliente suscriba una póliza sólo en un ramo distinto al que ya tiene alguna póliza.

$$\Pr(Us_{o_{i,j,(t+1)}} > 0 | Us_{o_{i,j,t}} = 0)$$

Dónde:

$Us_{o_{i,j,t}}$ = cantidad de productos del ramo j que tiene el cliente i en el tiempo t

$$E(CLV_{i,j,T+1}) = E(Profit_{i,j,T+1}) \cdot \Pr(Us_{o_{i,j,(t+1)}} > 0 | Us_{o_{i,j,t}} = 0)$$

Si el cliente adquiere una nueva póliza, entonces se espera obtener un beneficio que corresponde al promedio de los beneficios obtenidos por los clientes con el ramo "j" en el anterior.

$$E(Profit_{i,j,T+1}) = \frac{1}{M_{T,k}} \sum_{i=1}^{M_{T,j}} Ramo_{i,j,T} Us_{o_{i,j,T}} Margen_{i,j,T}$$

Donde:

$M_{T,k}$ = Numero de clientes con el ramo "k" en el tiempo "T"

Finalmente, el valor potencial $P(CLV_{i,T+1})$ asociado al cliente “i”, es la suma de los valores potenciales asociado a todas sus pólizas nuevas

$$\begin{aligned}
 P(CLV_{i,T+1}) &= \sum_{j=1}^J E(CLV_{i,j,T+1}) \\
 &= \sum_{j=1}^J E(Profit_{i,j,T+1}) \cdot \Pr(Us_{i,j,(t+1)} > 0 | Us_{i,j,t} = 0)
 \end{aligned}$$

6.4. Probabilidades de Modelos:

Para la aplicación de la metodología anterior es central estimar las siguientes probabilidades:

6.4.1. Probabilidad de Renovación:

Se utiliza un modelo de regresión logística para ajustar la probabilidad de renovación dentro de cada ramo.

$$\Pr(Re_{i(T+1),j,l} = 1 | Re_{i,T,j,l} = 1)$$

Para ello, en primer lugar se analiza por separado los clientes que tienen sólo una póliza del ramo planteado y las que tienen más. Es decir, se ajustan dos modelos logísticos, uno para los clientes que tienen una póliza del ramo “j” y otro para los que tienen más de una. El fundamento de esta distinción, se basa en que los clientes que tienen un producto tienen un comportamiento muy diferente a los que tienen más de uno. La intuición planteada fue evidenciada en los continuos ajustes que sufrió el modelo, es decir, el modelo logro un mejor ajuste al separar las muestras.

6.4.2. Probabilidad de Suscripción

Se utiliza un modelo de regresión logística para ajustar la probabilidad de suscribir una o más pólizas en un ramo en que el cliente no contaba con pólizas.

$$\Pr (Uso_{i,j,(t+1)} > 0 | Uso_{i,j,t} = 0)$$

Las variables explicativas a utilizar se obtendrán de los perfiles de análisis de datos. Por supuesto, analizando su significancia.

6.5. Modelo Logit:

Como se explicaba en el punto anterior, se utilizará un modelo logit para determinar la probabilidad de renovación y suscripción. Este deriva de la formación de un modelo de variable latente:

$$Y^* = \beta_0 + \beta'X_i + e_i$$

Donde X_i es un vector columna con K dimensiones de variables explicativas, β_0 es la ordenada en el origen, β' es el vector de parámetros desconocidos y e_i es el término de perturbación aleatoria del modelo. La variable Y^* se asocia normalmente con la utilidad que proporciona el individuo i al decidirse por una de las dos posibilidades contempladas en el modelo (Renovar o no renovar, suscribir o no suscribir).

En la realidad Y^* no puede ser observada. No obstante permite formular un modelo de elección discreta donde sólo el resultado final es observado, de modo que si la variable Y^* supera un determinado valor observaremos que se produce el suceso, mientras que de no superarlo, éste no sucederá.

$$Y_i = 1, \quad \text{si } Y_{i1}^* > Y_{i2}^*$$

$$Y_i = 0, \quad \text{en otros casos}$$

Donde Y_i es una variable binaria observable. De este modo, podemos calcular las siguientes probabilidades:

$$P(Y_i = 1|X_i) = P(Y_i^* > 0|X_i) = P(e_i > -\beta_0 - \beta'X_i|X_i) = F(\beta_0 + \beta'X_i)$$

$$P(Y_i = 0|X_i) = P(Y_i^* \leq 0|X_i) = P(e_i \leq -\beta_0 - \beta'X_i|X_i) = 1 - F(\beta_0 + \beta'X_i)$$

Donde $F(\cdot)$ representa la función de distribución (simétrica) del término aleatorio e_i . Cuando se asume la función logística para este término, el modelo resultante se denomina modelo Logit.

$$P(Y_i = 1|X_i) = \frac{e^{\beta_0 + \beta'X_i}}{1 + e^{\beta_0 + \beta'X_i}}$$

$$P(Y_i = 0|X_i) = \frac{1}{1 + e^{\beta_0 + \beta'X_i}}$$

Una vez especificado, el modelo se estima usando el método de máxima verosimilitud.

7. ALCANCES

Los alcances de este informe se circunscriben a la estimación del valor del cliente perteneciente a una compañía de seguros generales, bajo las siguientes simplificaciones:

Probabilidad de Suscripción: Se considera sólo en los casos que el cliente no tiene una póliza de ese ramo y tiene al menos una de otro

.

Individuo de análisis: Las compañías de seguro tienen dos tipos de clientes, el contratante de la póliza y el asegurado, bajo este escenario se considera al “contratante”. De todas maneras, las coincidencias entre contratante y asegurado están en torno al 87,5%.

Cortes anuales: Se consideran cortes anuales para evaluar la renovación o no renovación de un producto. Es decir, se evalúa que un cliente que tiene una póliza en un corte determinado la tenga en doce meses más. Esta simplificación se produce porque la compañía estudiada no tiene afinada la relación de renovación póliza a póliza.

Tipos de Clientes: Sólo se considerarán las líneas personales, pues el mercado de las empresas es muy distinto. Un estudio paralelo aumentaría la dificultad al doble, pues se necesitarían análisis paralelos.

Tipos de Ramos: Para simplificar el análisis se consideran sólo los principales ramos o productos que vende la compañía: Vehículos, Incendio y Terremoto, y Varios.

8. RESULTADOS ESPERADOS

Se espera definir por lo menos 5 características del cliente que permitan explicar los modelos de probabilidad de retención y suscripción. Entre los cuales deberían estar: la antigüedad de conducción, estado civil, lugar de residencia (urbana, rural), género, nivel de prima, edad, permanencia o antigüedad, etc.

Se espera obtener modelos globalmente significativos para las probabilidades de retención y suscripción.

Una vez que se cuente con el valor individual de cada cliente de la compañía, se espera tener valores agregados por tipo de perfil. La expectativa es que los clientes hombres, más antiguos, con nivel de prima más alta y más viejos tengan mayor valor esperado como clientes. Estos resultados se justifican en la literatura existente.

9. ANÁLISIS DE DATOS Y ESTADÍSTICA DESCRIPTIVA

Este capítulo tiene por objetivo realizar una breve descripción de la muestra de asegurados escogida, detallar sus características y dar los primeros fundamentos de inclusión de estas; socio-demográficas, productos contratados (Vehículos, Incendio y Varios), permanencia, Cross-Selling, etc.

En la misma línea, se aborda la estadística descriptiva asociada a la variable Siniestralidad. En este caso la descripción es más simple y agregada, pues se ocuparan modelos existentes y aportados por el departamento de riesgo de la compañía estudiada.

Se comienza por definir y describir los elementos comunes para el análisis: el *Cliente*, la *Muestra* y las *Variables*.

Cliente:

El primer paso es definir al individuo al que se le calculará su valor (cliente). La compañía aseguradora que provee los datos define al cliente como el *Contratante*, es decir el que adquiere la relación de pago con la compañía, que no necesariamente es el *Asegurado*.

Se ha escogido una muestra de *Líneas Personales* en el que 87,5% de los casos el *Contratante* es el *Asegurado*, así que no existe mayor diferencia entre el individuo que contrata la póliza que es el que justifica la Renovación/Suscripción y el que explica la Siniestralidad (*Asegurado*).

Muestra:

Se obtuvo información mensual a nivel individual para un horizonte de tiempo que va desde el 31 de diciembre del 2006 al 31 de diciembre del 2011. A continuación, se realizaron 6 cortes anuales al 31 de diciembre del 2006, 2007, 2008, 2009, 2010 y 2011.

La muestra considera a los 135.603 clientes vigentes al 31 de diciembre del 2009 y observa su comportamiento de renovación, suscripción, antigüedad, prima y siniestros en los 5 cortes restantes.

Tabla 1: Estadísticas Socio demográficas – Corte 2009

Datos Descriptivos		Participación
Edad	Menor de 18	0.1%
	18 - 24	0.3%
	25 - 29	5.9%
	30 - 39	26.2%
	40 - 49	27.4%
	50 - 59	21.2%
	Más de 60	18.8%
	Total	100.0%
Genero	Hombre	62.00%
	Mujer	37.06%
	Sin información	1%
	Total	100.00%
Estado Civil*	Casado	38.1%
	Solteros	13.4%
	Divorciados	0.2%
	Sin información	48.3%
	Total	100.00%
Residencia	Urbana	91.50%
	Rural	8.50%
	Total	100.00%
Nacionalidad	Chilena	51.3%
	Extranjeros	0.7%
	Sin información	48.0%
	Total	100.00%

Fuente: Elaboración Propia. A partir de datos aportados por la compañía estudiada.

Con respecto a las características demográficas descritas en la Tabla 1, tenemos que el 62% son hombre y un 37.06% son mujeres, el resto no son clasificables. La edad de los asegurados se concentra en el grupo de 30 a 50 años, quienes concentran el 53.6% de la muestra, la jóvenes van desde 18 a 30 años y corresponden a un 6.4% de la muestra. El 91.5% de los asegurados viven en zona Urbana y un 8.5% en zona rural.

El estado civil de los asegurados al 31/12/2007 se distribuye de la siguiente forma, un 38.1% son casados, un 13,4% solteros y un 48.1% no es clasificables.

La Tabla 2 confirma lo expresado en la justificación de este trabajo, con respecto a la baja tasa de permanencia que tienen los productos de seguros generales. Prácticamente el 50% de los clientes

Tabla 2: Estadísticas Permanencia

Permanencias	Frecuencia	Participación
1	75.998	56.20%
2	32.760	24.20%
3	15.838	11.70%
4	7.391	5.50%
>4	3.168	2%
Total	135.155	100.00%

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

La permanencia de los asegurados en la compañía se resume en lo siguiente: un 56.2% tiene una sola permanencia (12 meses), un 24,2% dos permanencias y aproximadamente un 20% tiene más de dos permanencias. La permanencia promedio para el ramo Vehículos, Incendio y Varios es de 18.3, 63.2 y 20.7 meses respectivamente.

Variables:

La metodología que se usa para la elección de las variables explicativas se funda en cuatro tipos de argumentos: bibliográfico, voz experta, disponibilidad de datos e intuición/prueba error propia.

Bibliográfico: Se usa como guía los papers de Donkers (2007) y de Guillén (2009). Con estos trabajos se obtuvo la primera aproximación a las variables consideradas determinantes tanto en la *Renovación y Suscripción* como para la *Siniestralidad*.

Voz Experta: El Departamento de riesgo de la compañía analizada contribuyó con el conocimiento de las variables críticas que justifican la Siniestralidad.

Disponibilidad de variables: La información a nivel de clientes recogida por las compañías de seguro es bastante pobre, pues han mantenido históricamente un bajo contacto. Desde este punto de vista, necesariamente existe una restricción de acceso a la información. Por ejemplo, el estado civil que bibliográficamente es definida como una variable relevante

Empíricamente en base a ajuste: Por supuesto, el argumento final para la elección de las variables está justificado en este trabajo. Es parte central, reconocer y justificar las variables que son los driver de la siniestralidad, renovación, suscripción y duración.

Finalmente, considerando los cortes especificados en el apartado anterior y las variables definidas según la metodología expuesta. La muestra de 135,155 clientes en vigor al 31 de diciembre del año 2009 fue estructura de la siguiente forma:

Tabla 3: Variables de interés base de datos

Variables de Interés
Número de Pólizas vigentes al corte del 31/12/2009, 31/12/2010 y 31/12/2011 separado por tipo de ramo: Vehículos, Incendio, Varios.
Variables Demográficas como la edad, sexo, estado civil, lugar de residencia y antigüedad como cliente de la compañía.
Monto de primas pagadas durante 2010 y 2011 para cada producto del ramo Vehículos, Incendio y Varios
Cantidad de siniestros ocurridos durante 2006 y 2007 para cada producto de autos o diversos que tenga el asegurado.
Monto de los siniestros durante 2010 y 2011 para cada producto del ramo Vehículos, Incendio y Varios

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

Las variables o información disponible para el modelo contemplado contienen algunas limitaciones, que se centran principalmente en los siguientes supuestos:

- No conocemos el costo particular de los siniestros, sino el agregado de la vida de la póliza.

- La siniestralidad de Vehículos, Incendio y Varios es muy diferente, pues existen diferencias implícitas en el modelo de negocios y en los costos de distribución de los ramos. Por ejemplo en el caso de los productos Incendio y Varios la varianza de la siniestralidad es enorme, pues no es un negocio de grande números sino de administración del riesgo. Por otro lado, el producto Vehículo evidencia tasas de siniestralidad muy estables dado el volumen en cartera.

De modo de contextualizar y proporcionar mayor grado de información sobre el tipo de asegurados y particularidades de los productos comercializados, se detallaran las principales características asociados a dichos productos.

9.1. Estadística Descriptiva

La presente sección proporciona una guía y contextualización de las Variables de Interés expuestas anteriormente:

9.1.1. Número de Pólizas Vigentes:

En la tabla 4 se observa que aproximadamente el 83% de los asegurados al 31/12/2011 tiene sólo una póliza de Vehículos, un 10% dos y cerca de 7% más de dos. Los porcentajes son estables en los tres cortes realizados.

Tabla 4: Variables de interés base de datos

VEHICULOS	Al 31/12/2009		Al 31/12/2010		Al 31/12/2011	
N° de Pólizas	Frecuencia	Peso	Frecuencia	Peso	Frecuencia	Peso
1	28,620	77%	48,214	80%	74,413	83%
2	4,424	12%	6,570	11%	8,596	10%
3	1,359	4%	1,779	3%	2,073	2%
4	884	2%	1,312	2%	1,616	2%
5	1,845	5%	2,570	4%	2,910	3%
	37,132	100%	60,445	100%	89,608	100%

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

Al igual que las estadísticas del ramo Vehículos, las pólizas de Incendio presentan un comportamiento similar en la proporción de número de pólizas por clientes. El ramo

Incendio presenta una mayor proporción en relación a Vehículos en la tenencia de dos pólizas, de la tabla 5 se desprende que un 16% de los clientes vigentes a finales del año 2009 tenían dos pólizas de Incendio y un 65% una (ver tabla 5).

Tabla 5: Variables de interés base de datos

INCENDIO	Al 31/12/2009		Al 31/12/2010		Al 31/12/2011	
	Frecuencia	Peso	Frecuencia	Peso	Frecuencia	Peso
1	13,459	65%	20,133	70%	29,454	75%
2	3,350	16%	4,088	14%	4,752	12%
3	1,536	7%	1,701	6%	1,977	5%
4	848	4%	944	3%	1,004	3%
5	1,665	8%	1,880	7%	1,990	5%
	20,858	100%	28,746	100%	39,177	100%

Fuente: Elaboración Propia. A partir de datos aportados por la compañía estudiada.

La Tabla 6 detalla la frecuencia de clientes por número de pólizas. Al igual que en Vehículos e Incendio los altos porcentajes de clientes que tienen sólo una póliza. Estadísticamente el 68% de los clientes al 31/12/2012 tienen 1 póliza vigente, un 16% dos y alrededor de un 16% más de dos.

Tabla 6: Número de Pólizas en vigor de Varios

VARIOS	Al 31/12/2009		Al 31/12/2010		Al 31/12/2011	
	Frecuencia	Peso	Frecuencia	Peso	Frecuencia	Peso
1	6,010	68%	9,547	72%	20,095	78%
2	1,380	16%	1,828	14%	3,160	12%
3	324	4%	438	3%	522	2%
4	292	3%	424	3%	628	2%
5	775	9%	980	7%	1,280	5%
	8,781	100%	13,217	100%	25,685	100%

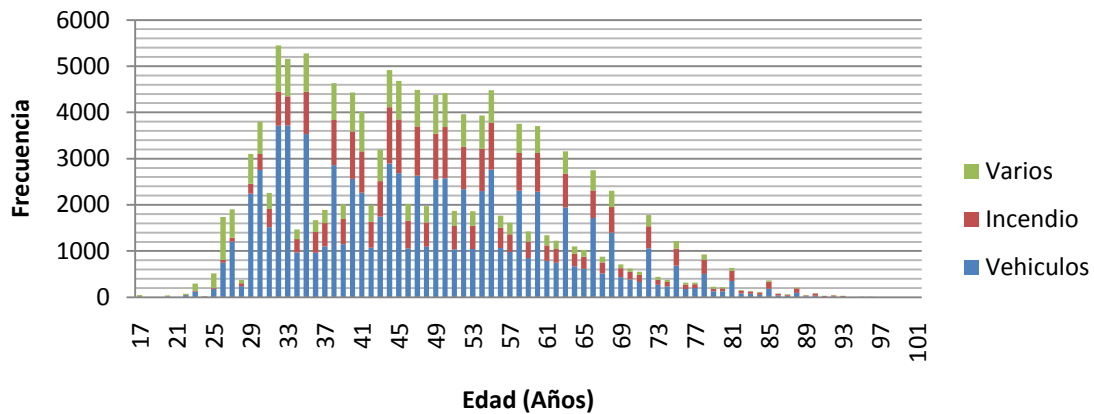
Fuente: Elaboración Propia. A partir de datos aportados por la compañía estudiada.

9.1.2. Variables Demográficas

Edad:

Estadísticamente la mayor proporción de los seguros generales son adquiridos por personas mayores a 30 años que han adquirido un grado de responsabilidad mayor, tanto sentimental (familia) como patrimonial (auto y casa propia). El Grafico 2 evidencia dicho supuesto, pues la mayor parte de las pólizas vigentes al 31/12/2009.

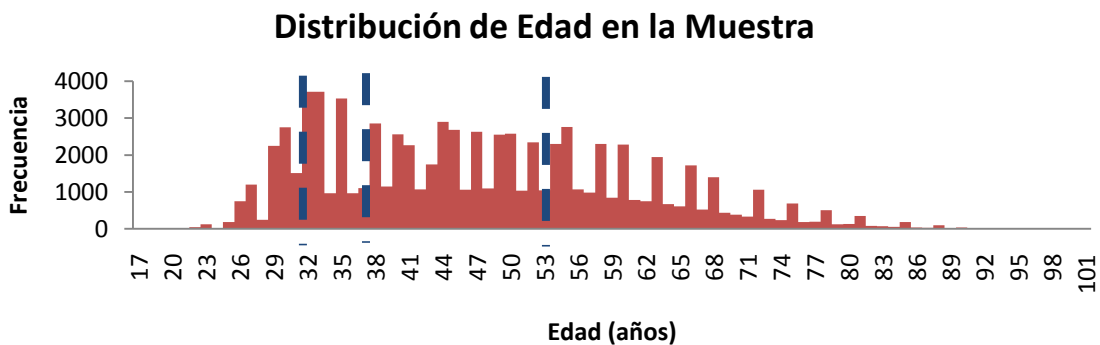
Gráfico 2: Distribución de Edad por Ramo al corte del 31/12/2009



Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

A partir del gráfico 3 se agrupa la edad en cuatro clúster. Esta agrupación tiene como fin la visualización de los modelos agregados, por lo tanto no constituye la agrupación que se usa en los modelos específicos de renovación y suscripción.

Gráfico 3: Distribución de Edad



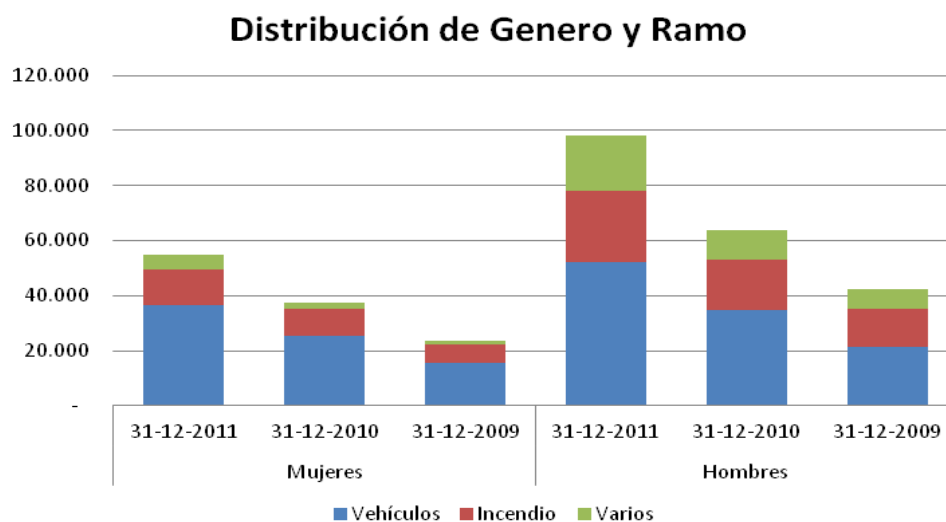
Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

Género:

El gráfico 4 evidencia que los hombres adquieren una mayor proporción de pólizas. Esto se debe a que una mayor proporción de ellos son el jefe de hogar, por lo tanto es el hombre, el que contrata el seguro asociado a un activo en particular, sea Vehículo, Casa, etc.

En el último tiempo se ha evidenciado un aumento en la proporción de mujeres que toman seguros. Esto está fundamentado en la participación laboral de la mujer y en el aumento de mujeres jefes de hogar.

Gráfico 4: Distribución de Genero para cortes 31/12/2009, 31/12/2010 y 31/12/2011

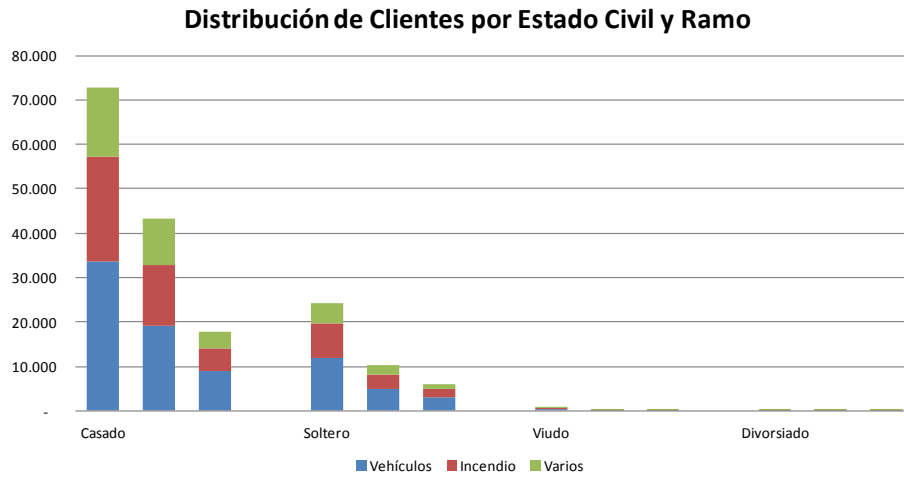


Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

Estado Civil:

En la misma línea que el Género y la Edad la aversión al riesgo es determinante a la hora de adquirir uno de estos productos. Debido a ello, la distribución de Cliente por estado civil está claramente marcada por el hecho que los casados toman una proporción mayor de seguros.

Gráfico 2: Distribución de Estado Civil para cortes 31/12/2009, 31/12/2010 y 31/12/2011



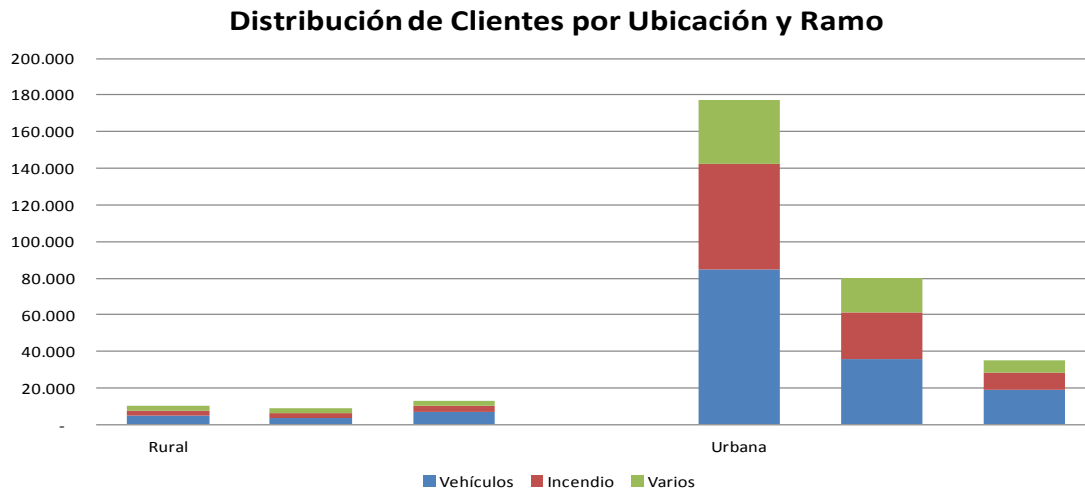
Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

Lugar de Residencia:

El 91% de los clientes vigentes en los tres cortes realizados tienen domicilio en zona Urbana. Al ser el seguro una necesidad de sociedades más complejas, el mayor grado de penetración de dichos productos se da en urbes y ciudades donde las personas se ven muy expuestas ante la ocurrencia de un siniestro. Por otro lado, el marketing y los puntos de venta de dichos productos se concentran enteramente en las zonas urbanas, lo que genera un mayor conocimiento de estos productos.

Se incluyó como variable explicativa, pues se contaba con dicha información. Luego, se evalúa la pertinencia y significancia de dicha variable para los modelos.

Gráfico 3: Distribución de Ubicación y Ramo para cortes 31/12/2009, 31/12/2010 y 31/12/2011



Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

9.1.3. Monto de primas pagadas

El monto en primas que el cliente tiene contratada con la compañía fue incluido como variable explicativa, fundamentado en que a mayor prima debería generarse un mayor compromiso tanto en la renovación como en la suscripción de nuevos productos.

El nivel de primas de los productos estudiados es muy diferente, esto lo refleja la tabla 7 donde las primas promedio de vehículos, Incendio y Varios están en torno a las 12.5, 6 y 3 UF respectivamente.

Dicha diferencia es una de las razones metodológicas para dividir el enfoque en el modelo de CLV que veremos más adelante.

Tabla 7: Promedio Prima Directa por Ramo (Cifras en UF)

Prima Directa por Ramo y Corte (UF)						
Ramos	Estadístico	31/12/2011	31/12/2010	31/12/2009	31/12/2008	31/12/2007
Vehículos	Promedio	12.13	12.80	12.65	12.92	13.09
	Desv. Estándar	18.45	18.70	6.59	12.42	7.59
Incendio	Promedio	6.82	7.04	5.64	5.65	5.23
	Desv. Estándar	9.55	8.75	7.26	9.90	9.43
Varios	Promedio	2.26	2.75	2.84	2.82	3.13
	Desv. Estándar	6.76	8.04	9.70	8.40	8.43

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

Las alzas de primas del ramo Incendio para el corte del 2011 y 2010 se producen, porque post terremoto 27/F los costos de reaseguro⁶ aumentan en todo el mercado. Por el contrario, para el ramo Vehículos sistemáticamente han ido disminuyendo las primas promedio.

9.1.4. Cantidad de Siniestros

La *cantidad de siniestros* al igual que el *monto en primas* fue incluida como variable explicativa. Intuitivamente debería ser una variable crítica, ya que normalmente el siniestro es un evento traumático donde es muy difícil que el cliente quede conforme con la liquidación. Por tanto, debería ser determinante a la hora de tomar la decisión de renovar o suscribir una nueva póliza

Al igual que en los niveles de prima descritos en el apartado anterior, la cantidad de siniestro difiere por ramo. En la tabla 8 se observa que en el ramo Vehículos 3 de cada 10 clientes presenten siniestros, por otro lado en el ramo Incendio 3 de cada 100 y en el

⁶ El costo de reaseguro se refiere a el seguro catastrófico que toman todas las compañías de seguros en Chile.

producto varios 1 de cada 100 aproximadamente. Es decir, el ramo Vehículos es el que presenta las más altas frecuencias entre los tres ramos⁷ descritos.

Tabla 8: Promedio Prima Directa por Ramo (Cifras en UF)

Cantidad de Siniestros por Ramo y Corte						
Ramos	Estadístico	31/12/2011	31/12/2010	31/12/2009	31/12/2008	31/12/2007
Vehículos	Promedio	29.2%	31.5%	33.3%	33.3%	36.7%
	Desv. Estándar	60.0%	58.8%	61.0%	61.0%	66.1%
Incendio	Promedio	3.2%	3.4%	12.3%	3.3%	3.8%
	Desv. Estándar	23.6%	23.7%	37.3%	24.8%	27.2%
Varios	Promedio	0.6%	1.2%	1.1%	1.3%	1.6%
	Desv. Estándar	8.0%	11.1%	9.7%	11.4%	13.0%

Fuente: Elaboración Propia. A partir de datos aportados por la compañía estudiada.

9.1.5. Monto de Siniestros Pagados

Es claro que la variable *Monto de Siniestros* está correlacionada con la variable *cantidad de Siniestros*, pero se han agregado ambas como variables explicativas para luego evaluar cual aporta más al modelo.

El principal driver de rentabilidad en seguros generales es monto de siniestros pagados sobre la venta (Prima Directa). Si observamos la tabla 8 (Primas) y la tabla 9 (Monto Siniestros), es claro que el monto de siniestro representa aproximadamente entre un 65% y 70% de la prima directa.

⁷ La Frecuencia de los siniestros de Vehículos es mucho mayor a la asociada a Incendio, etc.

Tabla 9: Promedio Costo de Siniestros por Ramo (Cifras en UF)

Pago de Siniestros por Ramo y Corte (UF)						
Ramos	Estadístico	31/12/2011	31/12/2010	31/12/2009	31/12/2008	31/12/2007
Vehículos	<i>Promedio</i>	7.38	6.15	6.18	5.53	6.56
	<i>Desv. Estándar</i>	28.68	23.44	24.01	147.66	26.17
Incendio	<i>Promedio</i>	2.62	3.41	16.06	0.96	0.87
	<i>Desv. Estándar</i>	102.20	79.74	119.35	28.97	17.72
Varios	<i>Promedio</i>	0.72	0.77	2.43	2.27	0.62
	<i>Desv. Estándar</i>	22.67	13.72	165.49	106.15	12.77

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

El monto promedio de pago de siniestros asociado al ramo Incendio para el corte 31 de diciembre del 2009, está fuera de los rangos históricos y por sobre la prima directa. La cifra se justifica porque la mayoría de las pólizas vigentes en dicho corte estarían luego vigentes el 27 de febrero del año 2010, fecha en que se produjo un importante terremoto en Chile.

9.2. VARIABLES A EXPLICAR

En el apartado anterior se ha abordado la estadística descriptiva para las variables socio-demográficas y transaccionales. En esta sección se expondrán tasas que son un muy buen aproximador de las variables dependientes que utilizaremos en el modelo de valor esperado: Renovación, Suscripción y Permanencia⁸.

9.2.1. Tasa de Renovación (Modelo Renovación):

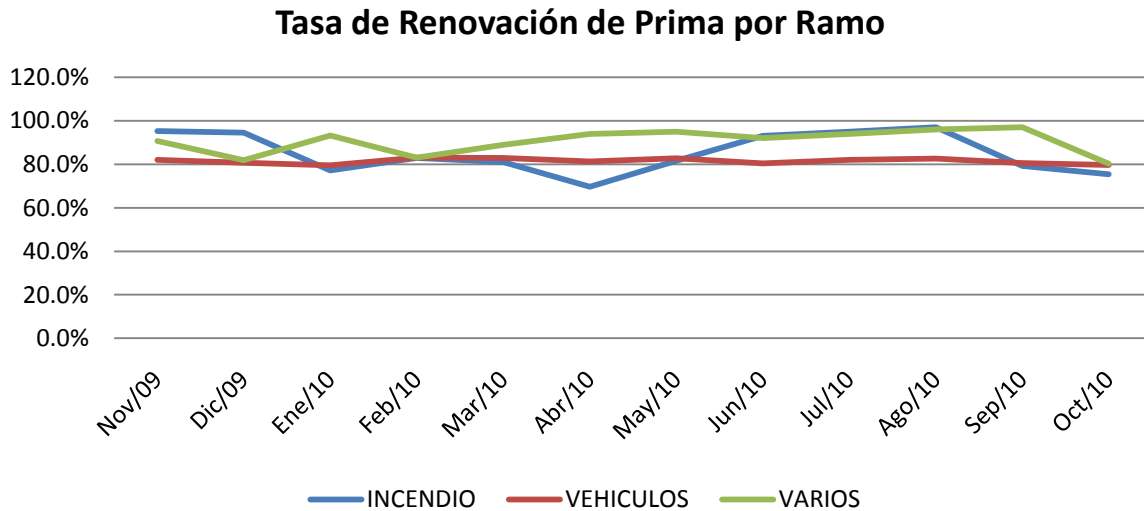
Corresponde a la proporción de clientes ponderados por su prima que siguen vigentes en una ventana de tiempo de 12 meses⁹. El Grafico 5 muestra la tendencia para los Ramos

⁸ Las variables estudiadas no necesariamente son las variables dependientes en el modelo, pues su objetivo es descriptivo.

⁹ Se usa una ventana de tiempo de 12 meses, debido a que el 98.2% de los seguros de auto que vende la cía. de seguro proveedora de los datos tienen una vigencia de 12 meses.

Vehículos, Incendio Y Varios para fines del 2009 y principios del 2010. Los países desarrollados¹⁰ presentan tasas muy superiores que en el caso del ramo vehículos ronda el 95%.

Gráfico 4: Tendencia de la Renovación de Prima por Ramo



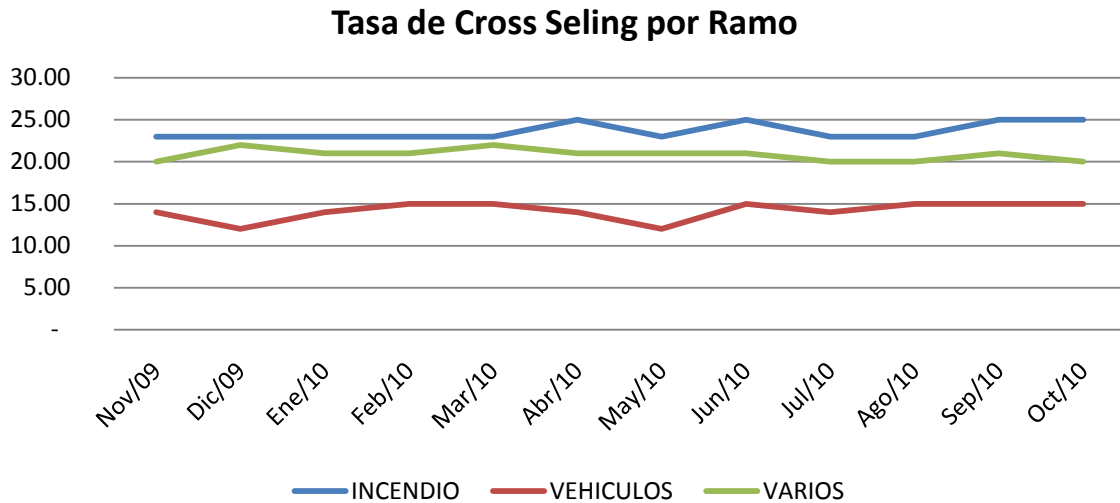
Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

9.2.2. Cross Selling o tenencia de productos (Modelo Suscripción):

La suscripción de un nuevo producto es el segundo elemento necesario para el cálculo del CLV, pues es la manera de medir la potencialidad de un determinado cliente. Con el fin de describir el nivel de cross selling existente, se graficará la tendencia asociada al porcentaje de clientes que tienen más de un producto.

¹⁰ MAPFRE renueva cerca del 95% de sus pólizas en España.

Gráfico 5: Tendencia de Cross Selling por Ramo



Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

El grafico 6 se interpreta por Ramo de la siguiente manera; de todos los clientes que tienen el Ramo Vehículos es el porcentaje que tiene al menos otro producto.

En términos generales al 31 de diciembre del 2009 un 86.2% tiene pólizas de un solo ramo, un 10.7% tiene de dos ramos y un 3.1% de los tres ramos estudiados.

Los índices de cross selling en el mercado de Seguros Generales son bastante pobres, esto se ha debido históricamente a que los esfuerzos comerciales no han tenido este foco.

9.2.3. Permanencias (Duración)

La variable dependiente en el modelo de Cox es la duración o sobrevivencia del cliente en la Compañía. Por lo tanto, describiremos estadísticamente la permanencia para tener una aproximación en el promedio a dicha variable.

Tabla 10: Permanencias por Ramo

Permanencia Años	Vehículos		Incendio		Varios	
	Frecuencia	%	Frecuencia	%	Frecuencia	%
1	26,270	32%	5,106	18%	11,223	47%
2	19,215	24%	6,323	22%	3,990	17%
3	10,689	13%	4,252	15%	2,146	9%
4	12,875	16%	2,735	9%	2,434	10%
5 o más	12,113	15%	10,738	37%	4,102	17%
TOTAL	81,162	100%	29,154	100%	23,895	100%

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

Intuitivamente el ramo Incendio debería evidenciar permanencias mayores. Al ser menor la frecuencia de siniestros, cambio de materia asegurada y prima, el cliente debería tener menos incentivos a cambiaren en el corto plazo. En la vereda contraria el producto Varios, genera menor fidelidad, pues la materia asegurada no necesariamente es un producto que conservaré a largo plazo, ni reemplazaré en el caso de siniestro.

Prácticamente la mitad de los clientes tienen una o dos permanencias en el ramo vehículos, mayor a eso la distribución es relativamente pareja en este ramo. Por otro lado, en el caso del producto Incendio prácticamente un 40% de la muestra tiene más de 5 permanencias, menor a eso, la distribución es pareja. Finalmente en el ramo Varios las permanencias promedio son más bajas, pues cerca del 50% tiene una permanencia, para permanencias mayores la distribución de frecuencias es pareja.

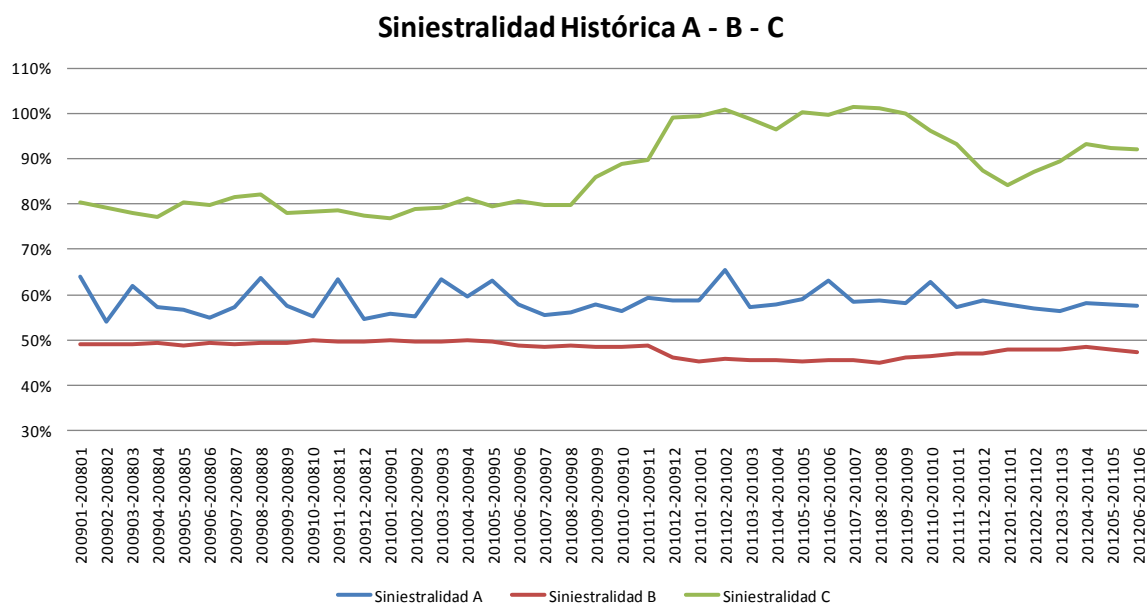
VALOR O RENTABILIDAD

Son dos los elementos necesarios para calcular la esperanza de un suceso; la probabilidad de dicho suceso y el valor asociado. Se cuenta con la metodología para estimar la probabilidad, por tanto este capítulo se centra en identificar el *valor* asociado a la Renovación y Suscripción de pólizas.

La variable clave para estimar el margen o la utilidad de un producto en seguros generales es la *Siniestralidad*. En términos simples corresponde al Monto o pago de siniestros

dividido por la prima ganada¹¹, es una muy buena aproximación del porcentaje de la venta que se lleva el pago de siniestros. El gráfico 7 muestra la siniestralidad por permanencia, donde el Grupo A son los asegurados que están en su primer año de permanencia y en promedio tienen una siniestralidad del 59.3% con una desviación estándar del 2,3%. El grupo B está al menos en segunda permanencia, este grupo tiene una siniestralidad promedio del 49.1% y una desviación estándar del 1.8%, en términos simples son los mejores clientes entre los tres grupos. Finalmente, el grupo C corresponde a los clientes que dejan la compañía, este grupo tiene una siniestralidad promedio del 86% y una desviación estándar del 7.3%, la alta siniestralidad con respecto a los demás grupos se justifica, ya que muchas veces la cía. Cancela pólizas a clientes que han tenido una siniestralidad muy alta en su última permanencia, además que en la última permanencia se incluyen las pérdidas totales donde desaparece la materia asegurada.

Gráfico 6: Siniestralidad por permanencia



Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

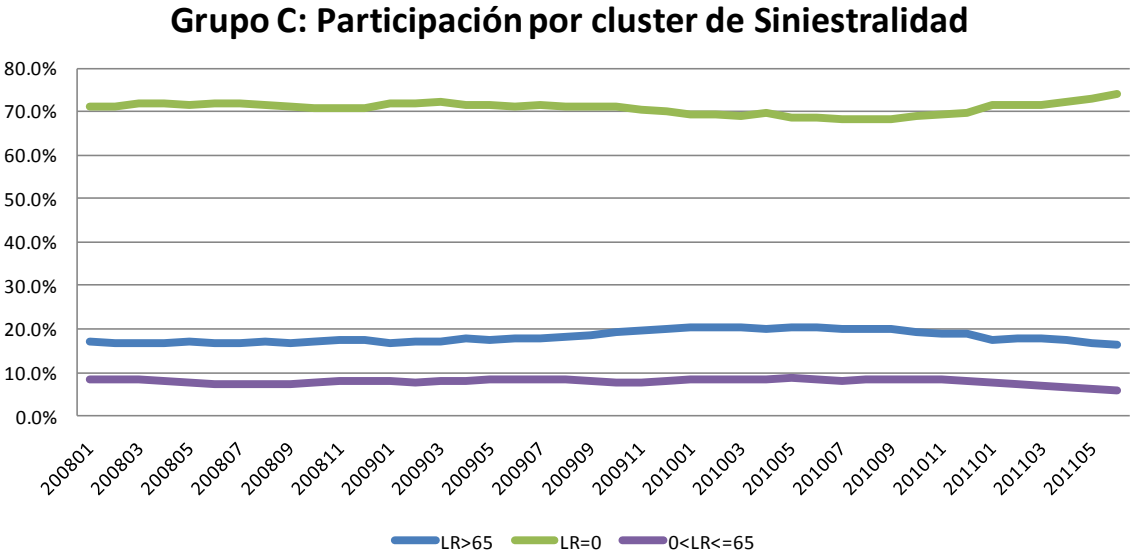
Renovación y Siniestralidad

¹¹ Prima devengada linealmente en periodo de vigencia

Con el objeto de presentar los resultados menos agregados, se cruzará el número clientes que dejan la compañía con su siniestralidad. El Grafico 7 divide a los clientes que salen de la compañía en tres grupos de siniestralidad (LR): siniestralidad (LR) igual a 0%, de 0 a 65% y mayor a 65%. En el Grafico 7 se puede observar que en promedio un 70% de los clientes que dejan la compañía no tenían siniestros (LR=0), un 10% tenía una siniestralidad aceptable entre 0 y 65% y un 20% una siniestralidad mayor al 65%.

Al aplicar un primer cruce obtenemos que prácticamente el 70% de los clientes que dejan la cía. lo hacen sin haber tenido siniestros (Grafico 7). Es decir, el cliente que se ha fugado constituyó una rentabilidad para la cía. de entorno al 65%¹².

Gráfico 7: Participación por cluster de siniestralidad (clientes que dejan la Cía)

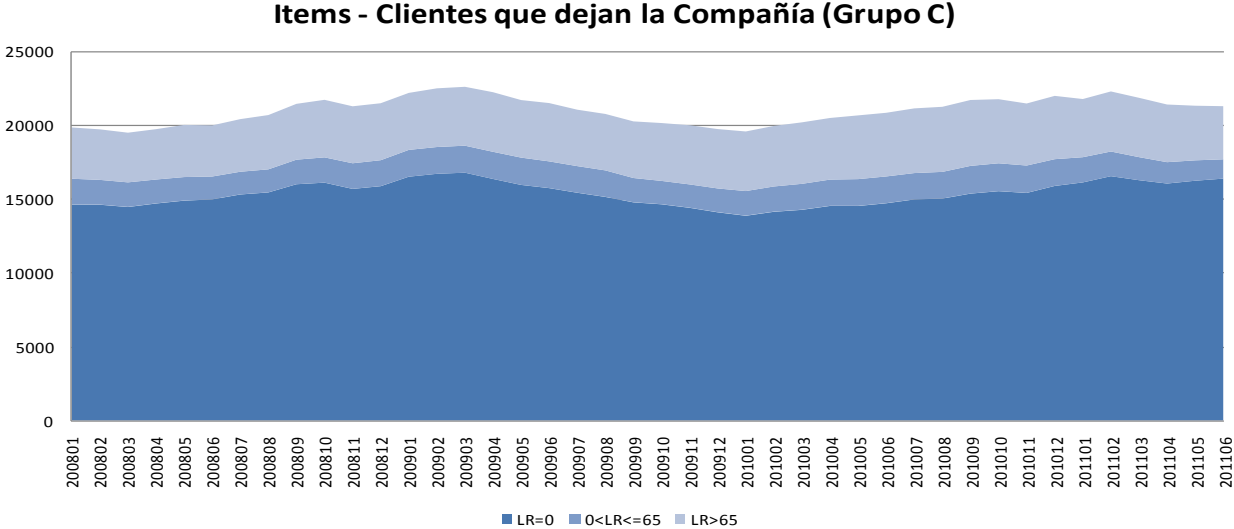


Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

¹² Los costos de distribución, gastos fijos y ROE se estiman alrededor del 35% de las ventas.

El gráfico 9 da cuenta de la masa que deja la compañía, que en promedio está en torno a los 16.000 clientes mensuales. Este gráfico detalla el potencial de fidelización que tiene la compañía estudiada.

Gráfico 8: Clientes que dejan la compañía por Siniestralidad (LR)



Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

10. TRANSFORMACIÓN, AGRUPACIÓN Y SELECCIÓN DE VARIABLES

La elección de las variables a incorporar se realiza en 3 etapas. En un inicio se hacen las transformaciones necesarias, se categorizan variables continuas mediante el método de k-means o Grafico y finalmente se normalizan las variables no categóricas. Una vez logrado lo anterior, se hace una selección evaluando que las variables sean estadísticamente significativas sobre la variable explicada.

10.1. Cluster y Categorización de Variables:

Algunas de las variables que se consideran en el modelo necesitan ser caracterizadas, pues de esa forma mejora tanto la significancia de la variable como la predicción del modelo. Las variables que fue necesario segmentar son:

10.1.1. Edad:

La variable **edad** fue segmentada en cuatro conglomerados según los métodos de k-means y grafico¹³. La tabla 11 muestra los rangos de edad que fueron calculados para el modelo *Renovación de una póliza de Vehículos*. En el Anexo se detallan la estadística el número de clientes por rango de edad para los demás modelos desarrollados.

Tabla 11: Conglomerados Edad para Modelo “Renovación de una póliza de Vehículos”

Conglomerado	Rango de Edad	N° Clientes	Peso
1	Menor a 38 años	16,457	30%
2	De 38 a 52 años	17,851	32%
3	De 52 a 66 años	14,763	27%
4	Más de 66 años	5,997	11%
Total		55,068	100%

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

¹³ Ver Estadística Descriptiva de la variable Edad

Esta segmentación fue desarrollada para el modelo completo. En términos específicos cada modelo cuenta con una segmentación de la variable Edad propia que se ajusta a la realidad del Modelo, Ramo y de la cantidad de pólizas renovadas.

10.1.2. Género

Esta variable posee sólo dos valores, hombre o mujer. Por lo tanto, se considerará como variable categórica.

10.1.3. Antigüedad:

Dado que estamos midiendo la antigüedad en años, con un rango de 1 a 5, no se agrupará más la variable. Los grupos son bien diferentes entre sí, Eso sí, se evaluará el ajuste como variable categórica o continúa en el modelo.

Tabla 12: Frecuencia de Permanencia por Cliente

Permanencia Años	Vehículos		Incendio		Varios	
	Frecuencia	%	Frecuencia	%	Frecuencia	%
1	26,270	32%	5,106	18%	11,223	47%
2	19,215	24%	6,323	22%	3,990	17%
3	10,689	13%	4,252	15%	2,146	9%
4	12,875	16%	2,735	9%	2,434	10%
5 o más	12,113	15%	10,738	37%	4,102	17%
TOTAL	81,162	100%	29,154	100%	23,895	100%

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

10.1.4. Número de Productos Contratados:

Se observa el grafico 10, la conclusión directa es que debemos agrupar el número de productos contratados en 1 o más pólizas, debido a que la frecuencia es demasiado baja cuando consideramos más de Dos.

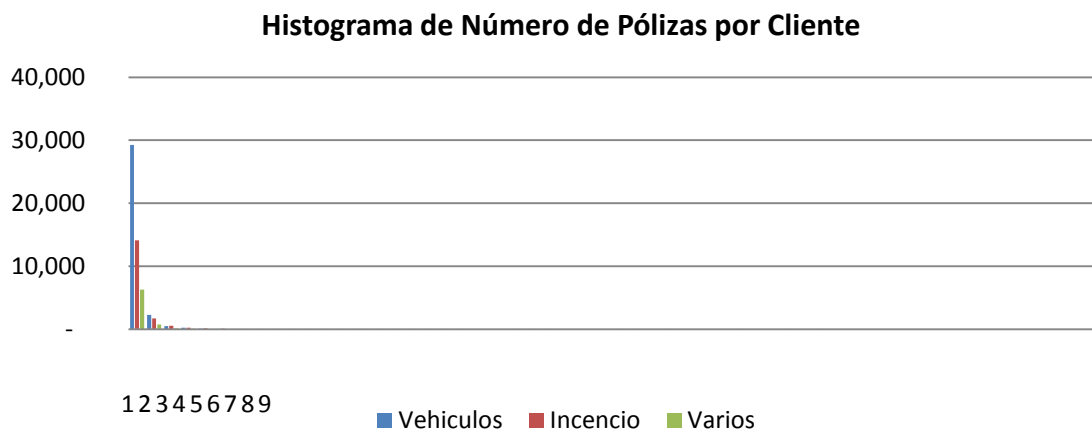
Tabla 13: Frecuencia de Número de Pólizas por Cliente

Frecuencia Número de Pólizas por Cliente			
	Vehículos	Incendio	Varios
1 Póliza	72,120	18,863	20,046
> 1 Póliza	5,958	3,037	2,257

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

La Variable número de productos quedaría agrupada en dos grupos, tener una o tener más de una.

Gráfico 9: Histograma de Número de Pólizas por Cliente y Ramo

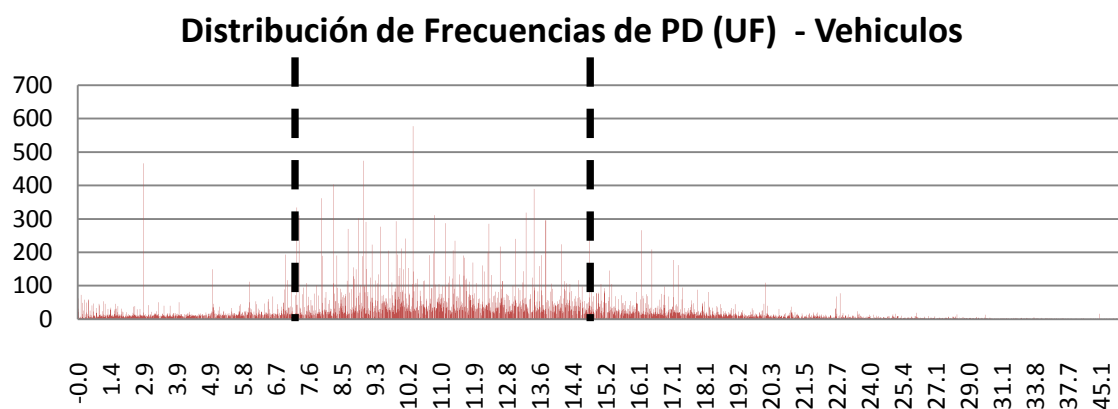


Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

10.1.5. Prima Directa:

La prima directa asociada a un cliente es el valor monetario de los productos adquiridos por él. La Segmentación de dicha variable se evalúa gráficamente, los siguientes tres gráficos corresponden a la frecuencia de de las primas, cada uno para cada ramo. Los cortes transversales dan una idea visual de los clúster escogidos.

Gráfico 10: Histograma de Frecuencia de Prima Directa (UF)- Ramo Vehículos

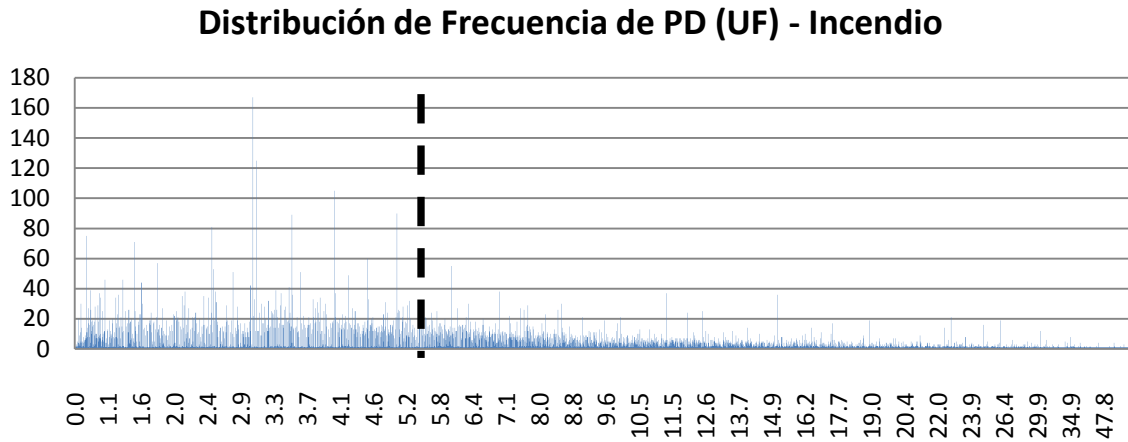


Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

Para el Ramo vehículos se divide la muestra de Prima Directa en tres grupos: Primas bajas, medias y Altas. El Grafico 11 da una visión grafica de la segmentación.

Por otro lado, la prima asociada al ramo Incendio se divide en dos clúster. El primero representa a los clientes con baja prima expuesta (clientes masivos) y un segundo menos masivo representa a los clientes con una prima expuesta mayor.

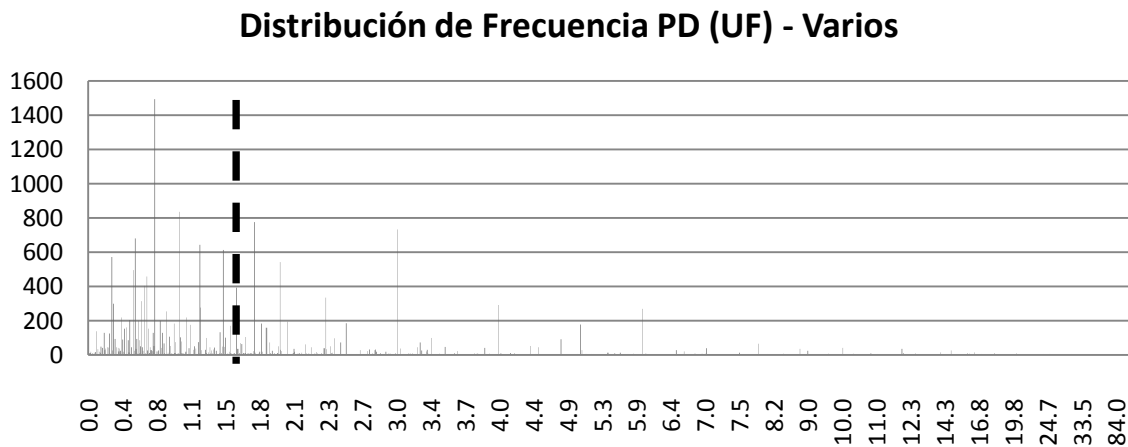
Gráfico 11: Distribución de Frecuencia de PD (UF) - Incendio



Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

La Prima Directa del ramo Varios recibe un trato similar al de la póliza de Incendio. Es decir, ha sido agrupada en dos grupos uno que recibe a la masividad de pólizas (commodities) y las primas más específicas que son las asociadas al clúster de mayor valor monetario.

Gráfico 12: Histograma de Frecuencia de Prima Directa (UF)- Ramo Varios



Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

La siguiente Tabla resume el número de conglomerados, los rangos de primas y números de clientes asociados al modelo *Renovación de una póliza de Vehículos*.

Tabla 14: Conglomerados Prima para Modelo “Renovación de una póliza de Vehículos”

Conglomerado	Rangos de Prima	N° Clientes	Peso
1	Menor a 7 UF	1,208	2%
2	De 7 a 15 UF	40,337	75%
3	Más de 15 UF	13,522	23%
Todos		55,067	100%

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

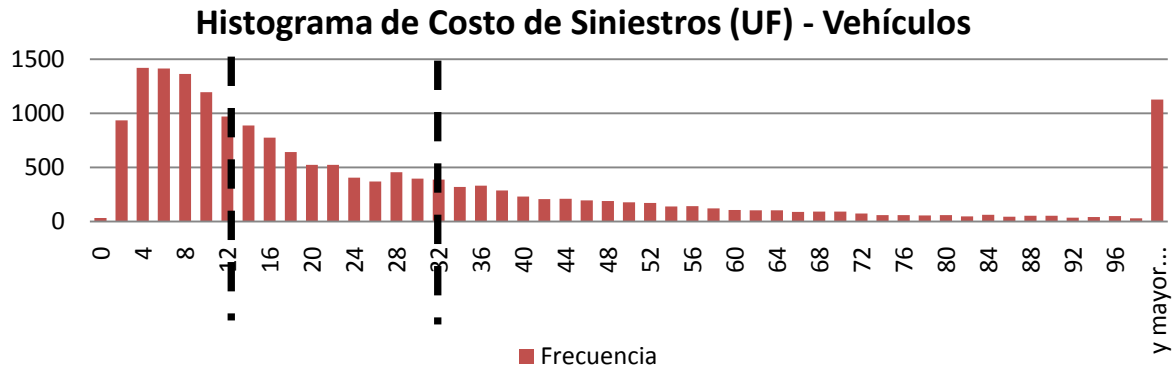
Las segmentaciones asociadas a los demás ramos y modelos, se explicitan en los anexos.

10.1.6. Monto de Siniestro:

El carácter del siniestro es muy diferente según el monto de liquidación. Los siniestros de baja cuantía están asociados a pequeños servicios que el cliente pide a la compañía, por ejemplo topones en el automóvil, amagos de Incendio, etc. Por otro lado, los siniestros de mayor cuantía están asociados a la pérdida de valor del bien asegurado, por lo tanto la actitud que toma el cliente frente a la liquidación del siniestro es muy distinta.

La agrupación del ramo vehículos está dividida en tres grupos: los siniestros bajo, medios y altos. Los bajos son los menos a 12 UF, los medios entre 12 y 32 UF y los altos sobre las 31 UF.

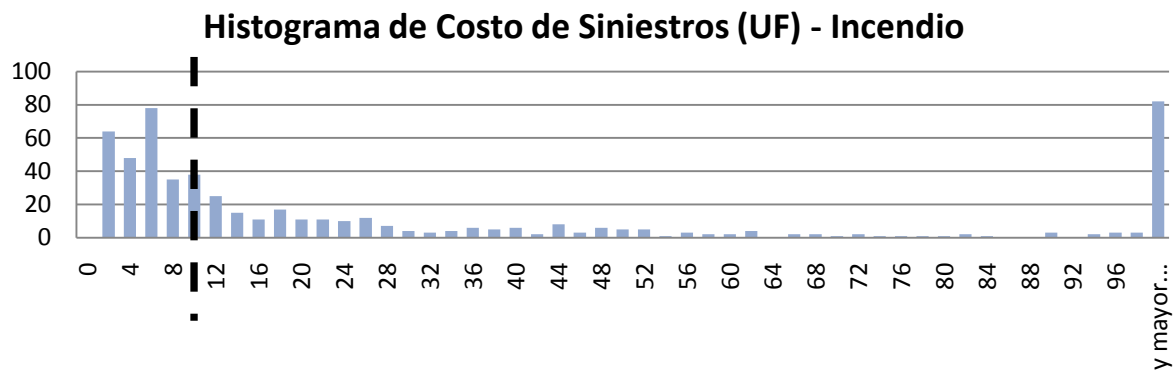
Gráfico 13: Histograma de Frecuencia de Costo de Siniestro- Ramo Vehículos



Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

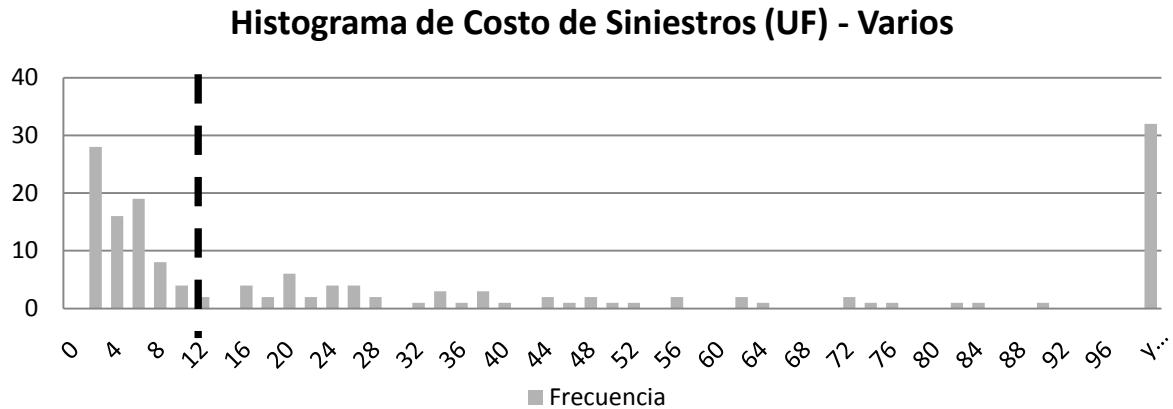
Basado en argumento planteado a comienzos de este apartado, el monto de siniestros para el ramo Incendio y Varios fue agrupado en dos clúster. El primero que agrupa a los siniestros de baja cuantía (mayor frecuencia) y los siniestros de cuantía mayor que son menos frecuentes. Los gráficos 18 y 19 evidencian de forma grafica los fundamentos de dicha agrupación.

Gráfico 14: Histograma de Frecuencia de Costo de Siniestro- Ramo Incendio



Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

Gráfico 15: Histograma de Frecuencia de Costo de Siniestro- Ramo Varios



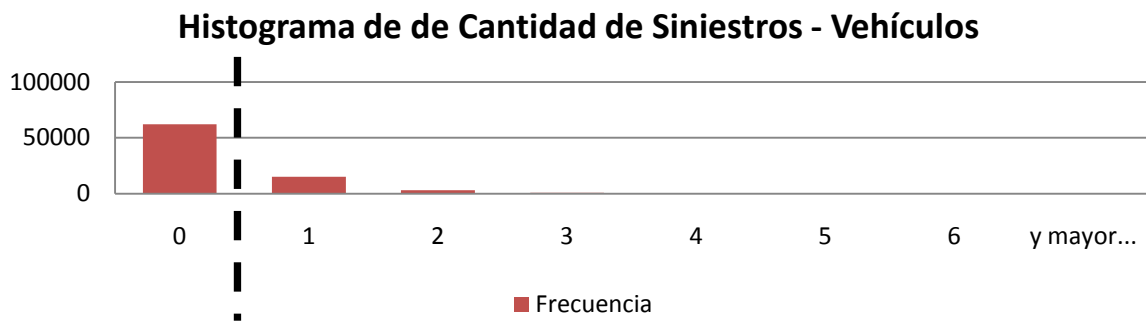
Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

10.1.7. Cantidad de Siniestros:

Los gráficos 20, 21 y 22 grafican la frecuencia de número de siniestros para los ramos Vehículos, Incendio y Varios. A simple vista, se observa que existe una clara distinción entre tener o no tener siniestros.

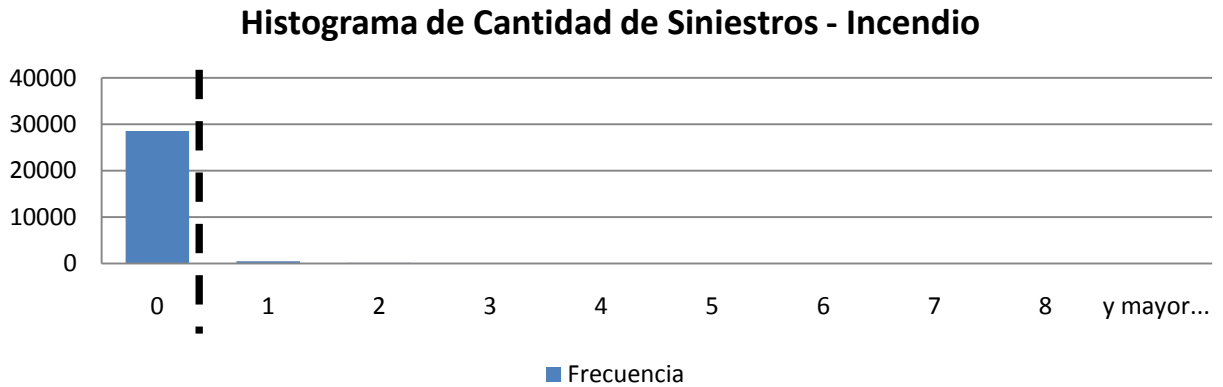
Por tanto, es directo concluir que la variable cantidad de siniestros se debe transformar en una variable binaria: *Tiene Siniestro* o *No tiene Siniestro*. De todas maneras, esta definición mejora la significancia a la hora de ser incluido en los modelos.

Gráfico 16: Frecuencia de Cantidad de Siniestros - Ramo Vehículos



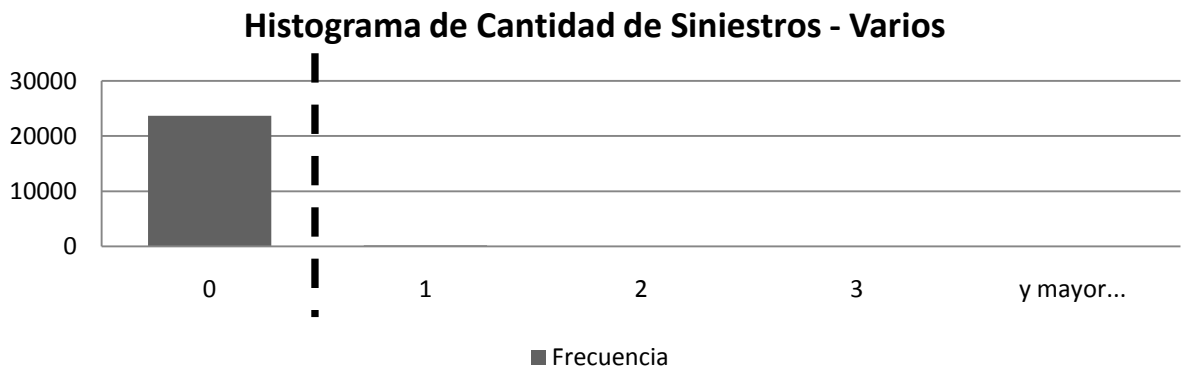
Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

Gráfico 17: Frecuencia de Cantidad de Siniestros - Ramo Incendio



Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

Gráfico 18: Frecuencia de Cantidad de Siniestros - Ramo Varios



Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

10.2. Test ANOVA

Para el test ANOVA definimos en principio a las variables dependientes e independientes:

Variables Dependientes:

- Renovación una póliza Vehículos, Incendio y Varios
- Renovación más de una póliza Vehículos, Incendio y Varios
- Suscripción Vehículos, Incendio y Varios

Variables Independientes: Género, Estado Civil, Edad, Antigüedad por Ramo, Número de Siniestros, etc.

Al aplicar el test ANOVA a la variable **Edad**, se observa que es estadísticamente significativa sobre las variables **Renovación de una póliza** y **Suscripción**, para los productos Vehículos, Incendio y Varios. Esto implica que la variable Edad es una variable que aporta información para mejorar la estimación de la variable renovación.

Por otro lado, la variable **edad** no es estadísticamente significativa sobre la variable Renovación más de una póliza en los ramos Vehículos e Incendio.

La tabla 15 resume el test ANOVA aplicado a la variable Edad y confirma los resultados expuestos.

Tabla 15: Test Anova para la variable independiente Edad

TEST ANOVA: VARIABLE EDAD

Variable Dependiente		Suma de cuadrados gl		Media cuadrática F		Sig.
Renovación una póliza Vehículos	Inter-grupos	59.4	81	0.734	3.301042266	0.00
	Intra-grupos	12,220.0	54,986	0.222		
	Total	12,279.4	55,067			
Renovación más de una póliza Vehículos	Inter-grupos	8.3	70	0.119	0.763433642	0.93
	Intra-grupos	607.9	3,895	0.156		
	Total	616.2	3,965			
Renovación una póliza Incendio	Inter-grupos	18.7	80	0.234	1.219557711	0.05
	Intra-grupos	2,850.3	14,867	0.192		
	Total	2,869.0	14,947			
Renovación más de una póliza Incendio	Inter-grupos	17.5	74	0.237	1.823372158	0.000
	Intra-grupos	303.2	2,332	0.130		
	Total	320.7	2,406			
Renovación una póliza Varios	Inter-grupos	87.6	80	1.095	4.506280352	0.000
	Intra-grupos	3,224.2	13,268	0.243		
	Total	3,311.8	13,348			
Renovación más de una póliza Varios	Inter-grupos	11.3	62	0.182	0.867620737	0.757
	Intra-grupos	242.0	1,152	0.210		
	Total	253.3	1,214			
Suscripción una póliza Vehículos	Inter-grupos	4.0	89	0.045	1.397328353	0.008
	Intra-grupos	917.5	28,745	0.032		
	Total	921.5	28,834			
Suscripción una póliza Incendio	Inter-grupos	0.9	82	0.011	1.292932301	0.039
	Intra-grupos	605.8	70,405	0.009		
	Total	606.7	70,487			
Suscripción una póliza Varios	Inter-grupos	0.6	87	0.007	1.265803677	0.048
	Intra-grupos	414.0	73,191	0.006		
	Total	414.6	73,278			

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

11. RESULTADOS Y VALIDACIÓN DE MODELOS

El capítulo 11 de este trabajo tiene por objetivo ajustar dos modelos logísticos Probabilidad de Renovación y de Suscripción

En los tres modelos se trabajó con el 80% de la muestra para la calibración y el 20% restante para la validación. En general, se usó la matriz de *observación/pronóstico* para medir la calibración y validación del modelo. Además, se agrega como validación la curva COR para estimar la ganancia del modelo independiente del punto de corte seleccionado.

Finalmente, se evaluó la inclusión o descarte de variables según significancia y ajuste al modelo total (método de prueba error).

11.1. Modelo de Renovación

Con el modelo de Renovación se desea determinar la probabilidad que el cliente mantenga una anualidad las pólizas que tiene hasta el momento. La potencialidad de adquirir nuevas pólizas se tratará en el modelo de suscripción.

Se corrió un modelo logístico para cada ramo. Considerando separadamente los asegurados que sólo tienen una póliza de ese ramo y los que tienen más de una. La separación por Ramo y número de pólizas, se fundamenta en que el corte para el modelo logístico (Renueva o no renueva) se definió en el promedio, y dicho promedio es muy distinto para los clientes por ramo y por clúster de número de pólizas.

En resumen, para todos los clientes que tienen una póliza de Vehículos a fines del 2009, se pronosticó la probabilidad de que renueven esa póliza durante el año 2010. Paralelamente se corrió un segundo modelo logístico para los clientes que tengan dos o más pólizas de Vehículos a fines del 2009, de manera que se tenga la probabilidad de renovación de todas esas pólizas en el siguiente año.

De la misma forma, se realizó el proceso análogo para el ramo Incendio y Varios.

Renovación Vehículos

La tabla 16 muestra los resultados para el ramo Vehículos para el caso en que el cliente tiene contratada sólo una póliza. La muestra tiene un número total de observaciones de 55,068 casos, correspondiente a los clientes que tenían una póliza del ramo Vehículos a finales del 2009. Las renovaciones a finales del 2010 alcanzan 36,585 casos, lo que representa un 66.4% de la muestra.

Tabla 16: Modelo logístico para el ajuste de la probabilidad de renovar una póliza de Vehículos si estaba contratada el año anterior

Variables en la ecuación	Variables Categorizadas	Beta	gl	Sig.	Exp(B)
Genero	Femenino	0,089	1	0,0000	1,0927
Edad	Mayor a 51 años		3	0,0000	
	Hasta 33 años	-0,136	1	0,0000	0,8726
	Desde 33 hasta 40	-0,192	1	0,0000	0,8257
	Desde 40 hasta 51	-0,123	1	0,0000	0,8841
Prima Directa	Hasta 6.5 UF		2	-	
	Desde 6.5 a 15 UF	-1,385	1	-	0,2503
	Desde 15 UF	-0,058	1	0,0161	0,9439
Antigüedad en vehículos	Mayor a 2 años		2	0,0000	
	Un año	-0,485	1	0,0000	0,6156
	Dos años	-0,260	1	0,0000	0,7710
Tiene Póliza de Incendio	No tiene	-0,262	1	0,0001	0,7692
Tiene Siniestro	No tiene	0,290	1	0,0000	1,3365
Constante	Constante	1,313	1	0,0000	3,7159

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

Una vez eliminadas las variables no significativas para el modelo. Han quedado como variables significativas las expuesta en la tabla 16.

De la tabla 16 se desprende que el género masculino disminuye la probabilidad de renovar. De la categorización de prima directa, se concluye que el grupo de prima que van de 6.5 a 15 UF (primas estándar o commodities) disminuyen de forma importante la probabilidad de renovar en contraposición a las primas de los extremos (productos más específicos).

El tener una póliza de Incendio aumenta la probabilidad de renovar en este modelo, por el contrario el tener una póliza de Varios no fue significativo (variable quedó fuera de ecuación).

Intuitivamente se ha planteado que el siniestro genera rechazo a la hora de la renovación, esto es confirmado a partir modelo propuesto, donde la variable *Tiene Siniestro* apunta en la dirección planteada.

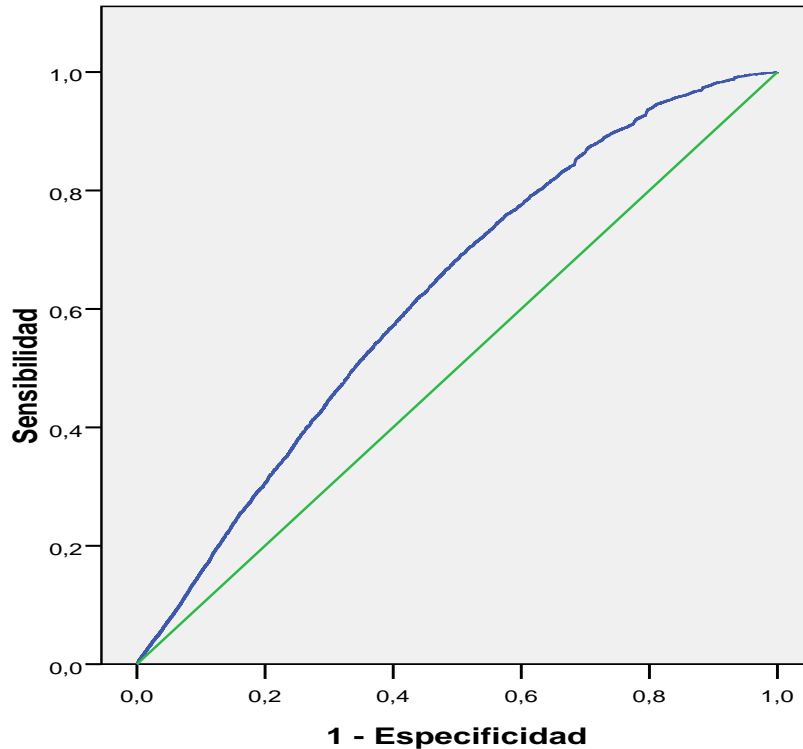
Finalmente, la variable *Antigüedad* aumenta la fidelidad, pues a antigüedades menores disminuye la probabilidad de renovación como se puede apreciar en la tabla 16.

Para medir la sensibilidad y pronóstico de un modelo logístico, se ha recurrido a la curva de ROC, ya que es una medida global e independiente del punto de corte.

Para cada uno de los modelos se graficará e interpretará la curva, de modo de tener una idea del aporte que representa el modelo.

El área bajo la curva azul que detalla el gráfico 19 es igual a 0.66. Es decir, el modelo gana 0.16 sobre la probabilidad de lanzar una modelo aleatoriamente

Gráfico 19: Curva COR para modelo logístico para el ajuste de la probabilidad de renovar una póliza de Vehículos si estaba contratada el año anterior



Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

La tabla 17 muestra los resultados para el ramo Vehículos en el caso en que el cliente tiene contratada más de una póliza. La muestra tiene un número total de observaciones de 3,966 casos, correspondiente a los clientes que tenían más de una póliza del ramo Vehículos a finales del 2009. Las renovaciones a finales del 2010 alcanzan 3,203 casos, lo que representa un 80% de la muestra.

Las variables *Género* y *Tiene pólizas de Varios* no fueron significativas para el modelo, por lo tanto no fueron incluidas en la ecuación. Intuitivamente la variable *Género* no es significativa, pues la decisión de renovar un mayor número de pólizas es una medida más racional y no dependiente necesariamente del género.

Tabla 17: Modelo logístico para el ajuste de la probabilidad de renovar más de una póliza de Vehículos si estaban contratadas el año anterior

Variables en la ecuación	Variables Categorizadas	Beta	gl	Sig.	Exp(B)
Edad	Mayor a 60 años		3	0,0073	0,835
	hasta 37 años	-0,1799	1	0,2266	0,653
	desde 37 a 46 años	-0,4262	1	0,0005	0,840
	desde 46 a 60 años	-0,1746	1	0,0905	0,637
Prima Directa	Menor a 35 UF	-0,4505	1	0,0000	0,648
Tiene Siniestro	No tiene	-0,4341	1	0,0000	0,712
Tiene Póliza de Incendio	No tiene	-0,3395	1	0,0491	
Antigüedad en Vehículos	Mayor a 2 años		2	0,0000	0,499
	Un año	-0,6960	1	0,0000	0,810
	Dos años	-0,2109	1	0,0373	16,205
Constante	Constante	2,7853	1	0,0000	

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

Las variables *Antigüedad* y *Edad* apuntan en la misma dirección que el modelo descrito para Vehículos una póliza. Es decir, a mayor antigüedad mayor es el compromiso del cliente con la compañía, del mismo modo para grupos de edades mayores la probabilidad de renovación aumenta.

La variable *Prima Directa* está segmentada en dos grupos, uno congrega a los clientes que tienen primas expuestas menores a 35 UF y otro a los mayores. El modelo predice que los clientes con una cartera mayor a 35 UF tienen mayor probabilidad de renovar. La intuición asociada, es que el cliente con mayores primas tiene más que perder al cambiarse, por lo tanto evalúa con mayor detalle dicho cambio.

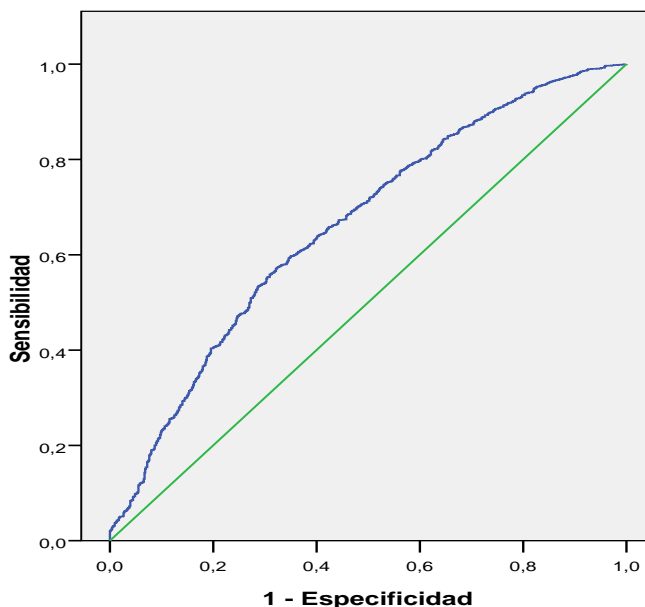
La variable *Tiene Siniestro* apunta en la dirección contraria que el modelo anterior. Esto puede deberse, a que el cliente que tiene un mayor número de pólizas, no tiene incentivo al

cambio de su cartera, pues es muy probable que reciba un aumento en su prima si su historial de siniestro es negativo¹⁴.

La curva de COR para el modelo de renovación de más de una póliza de Vehículos se detalla a continuación.

El área bajo la curva azul que detalla el grafico 20 es igual a 0.61. Es decir, el modelo gana 0.11 sobre la probabilidad de lanzar una modela aleatoriamente. El presente modelo es menos preciso que el anterior pero gana en sensibilidad.

Gráfico 20: COR – Vehículos más de una Póliza



Nota: Modelo logístico para el ajuste de la probabilidad de renovar más de una póliza de Vehículos si estaban contratadas el año anterior

¹⁴ Las compañías de Seguros Generales comparten el historial de siniestros. Este historial es evaluado especialmente a la hora de suscribir clientes con varias pólizas.

Renovación Incendio

La tabla 18 muestra los resultados para el ramo Incendio para el caso en que el cliente tiene contratada sólo una póliza. La muestra tiene un número total de observaciones de 14,948 casos, correspondiente a los clientes que tenían una póliza del ramo Incendio a finales del 2009. Las renovaciones a finales del 2010 alcanzan 11,076 casos, lo que representa un 74% de la muestra.

La variable *Tiene Póliza de Varios* y el *Género* no entran en la ecuación. Es muy interesante que la variable *Género* no haya sido significativa, esto puede justificarse porque la decisión de renovar o no renovar es compartida por la pareja dentro de un hogar, no es una decisión sobre la propiedad personal como es el auto.

Tabla 18: Modelo logístico para el ajuste de la probabilidad de renovar una póliza de Incendio si estaba contratada el año anterior

Variables en la ecuación	Variables Categorizadas	Beta	gl	Sig.	Exp(B)
Antigüedad en Incendio	Mayor a 2 años	0,5600	2	0,0000	
	Un año	-0,0679	1	0,0000	1,7506
	Dos años	0,2916	1	0,1879	0,9343
Tiene Siniestro	No tiene		1	0,0000	1,3385
Edad	Mayor a 54 años	-0,2914	3	0,0000	
	Hasta 40 años	-0,1294	1	0,0000	0,7472
	Desde 40 hasta 47	-0,0864	1	0,0145	0,8786
	Desde 47 hasta 54	1,0268	1	0,1094	0,9172
Prima Directa	Hasta 5,57 UF	-0,2716	1	0,0000	2,7922
Tiene Póliza de Vehículo	No tiene	0,3401	1	0,0002	0,7622
Constante	Constante	0,5600	1	0,0002	1,4051

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

En este modelo el comportamiento de la variable *Antigüedad* es algo más complejo. En este caso si un cliente tiene una *Antigüedad* de un año aumenta su probabilidad de renovar relativo al segundo año, esto puede deberse a que la segunda anualidad constituye el punto de inflexión en la decisión de renovar. La póliza de Incendio es un producto más barato y con un horizonte de permanencia mayor al de Vehículos, por lo tanto la decisión de

mantener o no la póliza no se toma en la primera anualidad, sino con más calma en la segunda¹⁵.

Las variables *Tiene Siniestros* y *Tiene póliza de Vehículos* apuntan en igual dirección que en el modelo de Vehículos (1 Póliza). Es decir, el tener siniestro disminuye la probabilidad de renovar y el tener una póliza de Vehículos aumenta la probabilidad de Renovar.

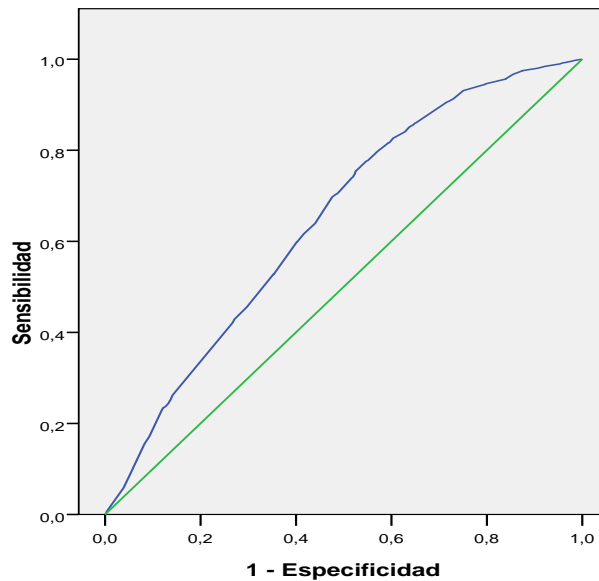
Al ir avanzado en los grupos de *Edad* va aumentando la probabilidad de renovar. La intuición que está detrás de esto, dice relación con que los grupos de mayor Edad tienen definido su Hogar y es poco probable que lo cambien, en la vereda contraria están los grupos más jóvenes que tienen mayor probabilidad a cambiar de hogar, por crecimiento de la familia, mayores ingresos, etc.

La curva de COR para el modelo de renovación de una póliza de Incendio se detalla a continuación.

El área bajo la curva azul que detalla el Grafico 21 es igual a 0.68. Es decir, el modelo gana 0.18 sobre la probabilidad de lanzar una modelo aleatoriamente.

¹⁵ La intuición planteada la confirma el departamento de Suscripción de la compañía analizada.

Gráfico 21: COR – Incendio de una Poliza



Nota: Modelo logístico para el ajuste de la probabilidad de renovar una póliza de Incendio si estaba contratada el año anterior

La tabla 19 muestra los resultados del ramo Incendio para el caso en que el cliente tiene contratada más de una póliza. La muestra tiene un número total de observaciones de 2,407 casos, correspondiente a los clientes que tenían más de una póliza del ramo Incendio a finales del 2009. Las renovaciones a finales del 2010 alcanzan 2,026 casos, lo que representa un 84% de la muestra.

Las variables *Tiene Póliza de Varios*, *Género*, *Antigüedad* y *Prima* no entran en la ecuación.

Tabla 19: Modelo logístico para el ajuste de la probabilidad de renovar más de una póliza de Incendio si estaban contratadas el año anterior

Variables en la ecuación	Variables Categorizadas	Beta	gl	Sig.	Exp(B)
Edad	Mayor a 60 años		2	0,000	
	Hasta 50 años	-0,8268	1	0,000	0,4374
	Desde 50 hasta 60	-0,2621	1	0,073	0,7695
Tiene Siniestro	No tiene	0,4856	1	0,003	1,6251
Tiene póliza de Vehículos	No tiene	-0,9201	1	0,000	0,3985
Constante	Constante	2,4016	1	0,000	11,0406

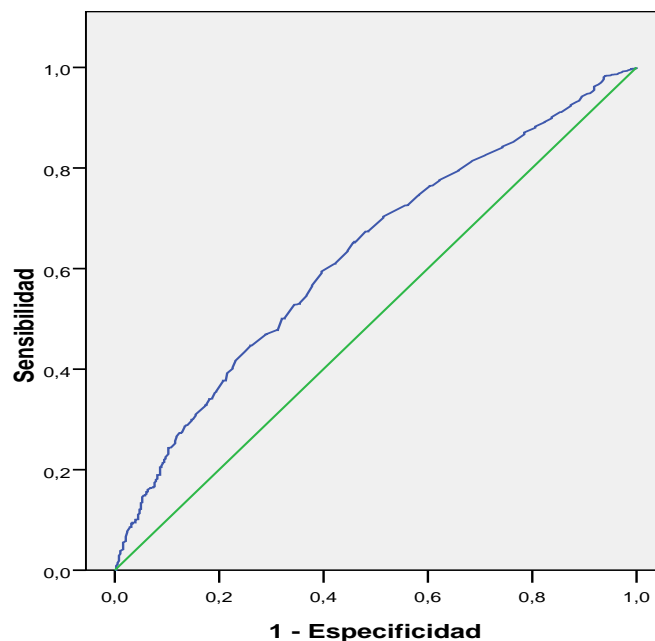
Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

Las tres variables en la ecuación *Edad*, *Tiene Siniestro* y *Tiene Póliza de Vehículos* asientan el modelo de Incendio una póliza, pues describen los mismos incentivos. Es decir, pasar a grupos de edad mayores, Tener póliza extra de Vehículos y no tener siniestros durante el periodo aumenta la probabilidad de renovar.

La curva COR que detalla el ajuste del modelo de renovación de más de una póliza en el ramo Incendio se grafica a continuación.

El área bajo la curva azul que detalla el grafico 22 es igual a 0.63. Es decir, el modelo gana 0.13 sobre la probabilidad de lanzar una modela aleatoriamente. En general los modelos que pronostican la renovación de más de una póliza se ajustan peor que los que pronostican la renovación de una póliza. Las razones pasan por el volumen de datos en cada modelo y por la ocurrencia del suceso.

Gráfico 22: Gráfico COR – Incendio más de una Póliza



Nota: Modelo logístico para el ajuste de la probabilidad de renovar más de una póliza de Incendio si estaban contratadas el año anterior

Renovación Varios

La tabla 20 muestra los resultados para el ramo Varios para el caso en que el cliente tiene contratada sólo una póliza. La muestra tiene un número total de observaciones de 13,349 casos, correspondiente a los clientes que tenían una póliza del ramo Varios a finales del 2009. Las renovaciones a finales del 2010 alcanzan 7,257 casos, lo que representa un 54.4% de la muestra.

La variable *Tiene Póliza de Vehículos* no entra en la ecuación. Se debe notar que en ningunos de los modelos precedentes la variable *Tiene pólizas de Varios* fue significativa, ésta póliza es la que genera menos fidelidad entre los tres productos analizados. Por el contrario, la variable *Tiene Póliza Incendio* aumenta considerablemente la probabilidad de renovar a las demás pólizas en cartera.

Tabla 20: Modelo logístico para el ajuste de la probabilidad de renovar una póliza de Varios si estaba contratada el año anterior

Variables en la ecuación	Variables Categorizadas	Beta	gl	Sig.	Exp(B)
Género	Femenino	-0,1609	1	0,0011	0,8514
Edad	Mayor a 49 años		3	0,0000	
	hasta 31 años	-0,4639	1	0,0000	0,6288
	desde 31 a 40 años	-0,1202	1	0,0132	0,8867
	desde 40 a 49 años	-0,1403	1	0,0028	0,8691
Prima Directa	Menor a 1 UF	-0,4146	1	0,0000	0,6606
Tiene Siniestro	No tiene	0,6910	1	0,0007	1,9957
Tiene Póliza de Incendio	No tiene	-0,4178	1	0,0002	0,6585
Antigüedad en Varios	Mayor a 2 años		2	0,0000	
	Un año	-1,1835	1	0,0000	0,3062
	Dos años	-1,0990	1	0,0000	0,3332
Constante	Constante	1,0882	1	0,0000	2,9688

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

La variable *Género* en este modelo tiene el comportamiento contrario al que tiene en el ramo Vehículos. En este caso el Género Femenino disminuye la probabilidad de renovar la póliza.

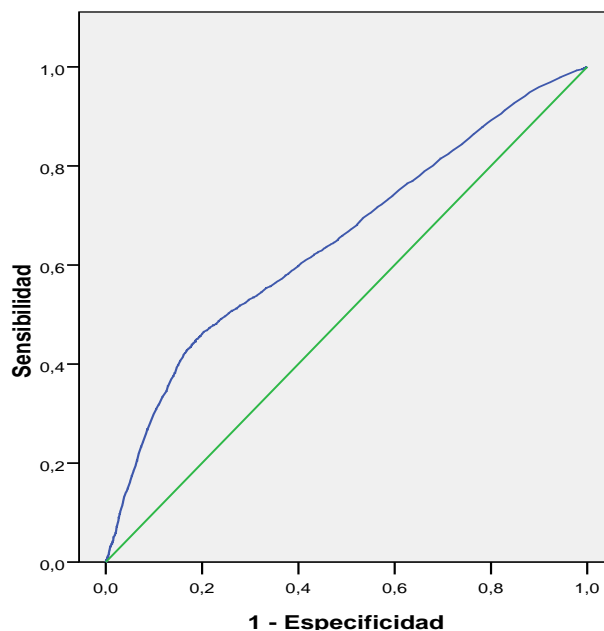
Por otro lado, las variables *Edad*, *Tiene póliza de Incendio* y *Tiene Siniestro*, tienen una interpretación muy similar a los modelos precedentes. Es decir, al aumentar en los grupos de Edad, tener una póliza de Incendio y no tener siniestro durante la anualidad aumenta la probabilidad de renovar la póliza.

En algún grado la póliza de Varios también es una póliza fidelizable, pues la variable *Antigüedad en Varios* aumenta la probabilidad de renovar.

La curva de COR para el modelo de renovación de una póliza de Varios se detalla a continuación.

El área bajo la curva azul que detalla el gráfico 23 es igual a 0.649. Es decir, el modelo gana 0.149 sobre la probabilidad de lanzar una modelo aleatoriamente.

Gráfico 23: COR – Varios una Póliza



Nota: Modelo logístico para el ajuste de la probabilidad de renovar una póliza de Varios si estaba contratada el año anterior

La tabla 21 muestra los resultados del ramo Varios para el caso en que el cliente tiene contratada más de una póliza. La muestra tiene un número total de observaciones de 1,215 casos, correspondiente a los clientes que tenían más de una póliza del ramo Varios a finales del 2009. Las renovaciones a finales del 2010 alcanzan 855 casos, lo que representa un 70% de la muestra.

Las variables *Tiene Siniestro*, *Género* y *Prima Directa* no entra en la ecuación. Que las variables *Género* y *Prima Directa* salgan del modelo de más de una póliza está en la línea de los resultados anteriores. Cuando la decisión renovación implica una cartera o un nivel de prima mayor parece ser independiente del Género, pues es una decisión evaluada económicamente.

Tabla 21: Modelo logístico para el ajuste de la probabilidad de renovar más de una póliza de Varios si estaban contratadas el año anterior

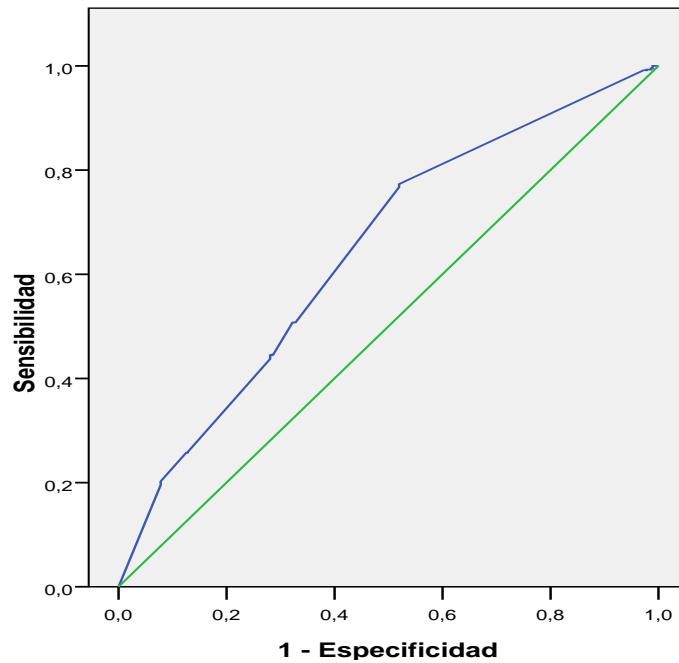
Variables en la ecuación	Variables Categorizadas	Beta	gl	Sig.	Exp(B)
Edad	Mayor a 58 años		2	0,088	
	Hasta 44	-0,1246	1	0,485	0,8829
	Desde 44 a 58 años	-0,3498	1	0,039	0,7048
Antigüedad en Varios	Mayor a 2 años		2	0,000	
	Un año	-0,9136	1	0,000	0,4011
	Dos años	-0,0679	1	0,714	0,9343
Tiene Póliza de Vehículos	No tiene	-0,7095	1	0,000	0,4919
Constante	Constante	2,0062	1	0,000	7,4354

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

De las Variables *Antigüedad en Varios* y *Tiene Póliza de Vehículos* se concluye

La curva de COR para el modelo de renovación de más de una póliza de Varios se detalla a continuación. El área bajo la curva azul que detalla el gráfico 24 es igual a 0.644. Es decir, el modelo gana 0.144 en sensibilidad y pronóstico sobre la probabilidad de lanzar una modelo aleatoriamente.

Gráfico 24: Gráfico COR – Varios una Póliza



Nota: Modelo logístico para el ajuste de la probabilidad de renovar más de una póliza de Varios si estaban contratadas el año anterior

11.2. Modelo de Suscripción

El modelo de Suscripción considera la posibilidad que un cliente adquiriera una o más pólizas de un ramo distinto a aquellos en los que tiene contratada una póliza. Por ejemplo, este modelo ajusta la probabilidad que adquiriera una o más pólizas de vehículos durante el 2010 dado que tenía contratada al menos una póliza de Incendio o Varios a fines del 2009.

Análogamente, se corrió un segundo y tercer modelo para estimar la probabilidad de contratación de una o más póliza de Incendio y Varios respectivamente. Por supuesto, sujeto a que el cliente tenga contratada al menos una póliza en cualquiera de los demás ramos.

Suscripción Vehículos

La tabla 22 muestra los resultados del ajuste de la probabilidad de contratación de una póliza de vehículos condicionada al hecho de no poseerla en el año anterior. La muestra tiene un número total de observaciones de 28,835 casos, correspondiente a los clientes que tenían una póliza del ramo Incendio o Varios a finales del 2009. De éstos, 953 contratan alguna póliza de Vehículos durante el año 2009, lo cual representa un 3,3% del total. Las variables incluidas en el modelo son *Edad*, *Tiene póliza de Incendio*, *Tiene Póliza de Varios* y *Prima Directa en Cartera*.

Tabla 22: Modelo logístico para el ajuste de la probabilidad de Suscribir una póliza de Vehículos si tenía al menos una póliza de Incendio o Varios

Variables en la ecuación	Variables Categorizadas	Beta	gl	Sig.	Exp(B)
Edad	Mayor a 54 años		2	0.021	
	Hasta 39 años	-0,2704	1	0.005	0.7630
	Desde 39 a 54 años	-0,0666	1	0.367	0.935
Tiene Siniestro en Incendio	No tiene	-0,3894	1	0.000	0.677
Tiene una póliza de Varios	No tiene	-0,5885	1	0.001	0.555
Prima Directa	Menor a 16 UF	-0,6758	1	0.000	0.508
Tiene Póliza de Incendio	No tiene	-1,1537	1	0.000	0.315
Constante	Constante	-1,5196	1	0.000	0.218

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

En este modelo la variable *Edad* esta agrupada para edades mayores a los demás modelos. El grupo de edades que toma una póliza de Vehículos siendo que ya tenía pólizas de Incendio o Varios aumenta, pues el promedio de edad que toma las pólizas de dichos ramos es mucho mayor. De todas maneras, la variable *Edad* apunta en la dirección clásica, es decir para aumentos en el clúster de edad aumenta la probabilidad de suscribir una póliza.

Tener pólizas de Incendio o Varios obviamente aumenta la probabilidad de suscribir una póliza de Vehículos. El peso relativo es mayor en el caso del ramo Incendio, es decir tener una póliza de Incendio genera mayor probabilidad de suscribir que tener una de Varios.

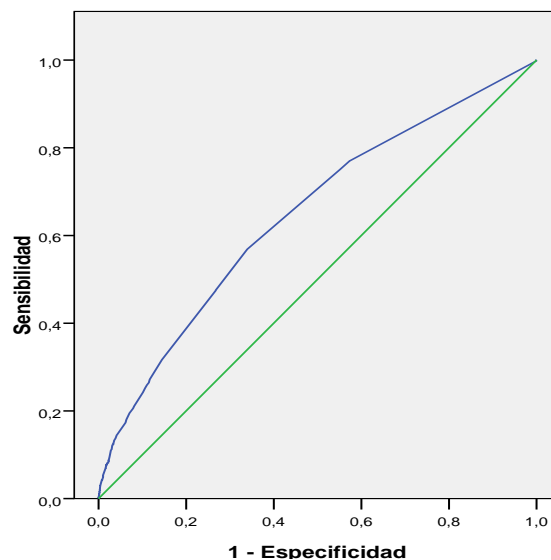
En el caso que la variable *Tener Siniestro* es negativa (No tener siniestro), la probabilidad de suscribir aumenta. Es decir, se confirma en la intuición asociada a que el siniestro necesariamente genera una distancia a la hora de renovar con la compañía.

Finalmente, la variable *Prima Directa* aumenta la probabilidad de suscribir cuando la prima en cartera del cliente es menor, cuando el cliente tiene poca exposición y por lo tanto no toma la decisión de forma estrictamente económica, sino asociada a la compañía que conoce. Por otro lado, los clientes con exposición de primas mayores tienden a realizar una evaluación puramente económica al momento de suscribir un nuevo producto.

La evaluación de los beneficios del modelo tanto desde el punto de vista de la sensibilidad y el pronóstico se detalla en la curva COR presentada a continuación.

La curva de COR para el modelo de suscripción de una póliza de Vehículos se detalla a continuación. El área bajo la curva azul que detalla el gráfico 25 es igual a 0.66. Es decir, el modelo gana 0.16 en sensibilidad y pronóstico sobre la probabilidad de lanzar una modelo aleatoriamente.

Gráfico 25: COR – Suscripción Vehículo



Nota: Modelo logístico para el ajuste de la probabilidad de Suscribir una póliza de Vehículos si tenía al menos una póliza de Incendio o Varios

Suscripción Incendio

La tabla 23 muestra los resultados del ajuste de la probabilidad de contratación de una póliza de Incendio condicionada al hecho de no poseerla en el año anterior. La muestra tiene un número total de observaciones de 70,488 casos, correspondiente a los clientes que tenían una póliza del ramo Vehículos o Varios a finales del 2009. De éstos, 612 contratan alguna póliza de Incendio durante el año 2009, lo cual representa un 1% del total. Las variables incluidas en el modelo son Género, Edad, Tiene póliza de Vehículos, Antigüedad y Prima Directa en Cartera.

Tabla 23: Modelo logístico para el ajuste de la probabilidad de Suscribir una póliza de Incendio si tenía al menos una póliza de Vehículos o Varios

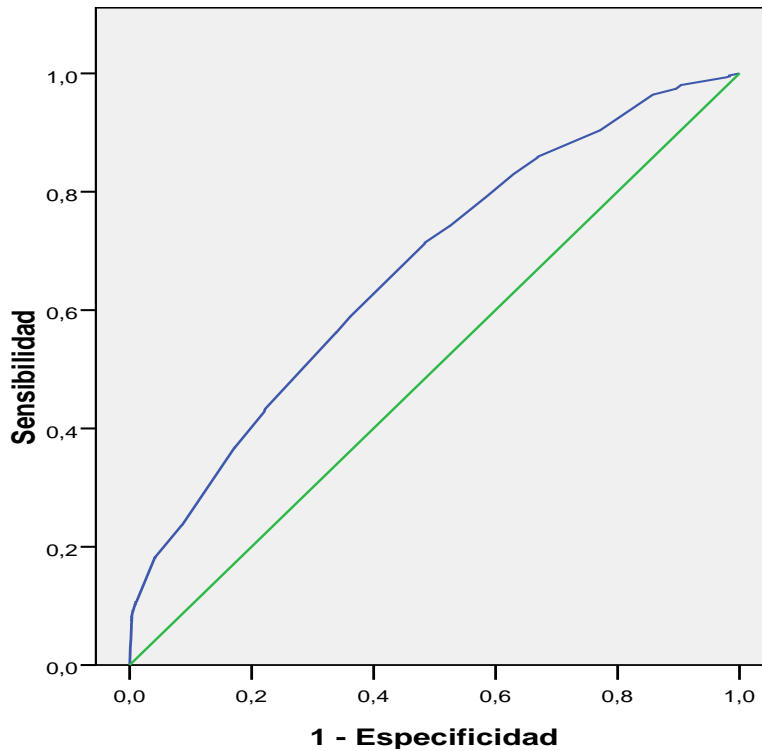
Variables en la ecuación	Variables Categorizadas	Beta	gl	Sig.	Exp(B)
Genero	Femenino	-0,5553	1	0.000	0.573
	Mayor a 55 años		2	0.000	
Edad	Hasta 37	-0,2262	1	0.072	0.797
	Desde 37 hasta 55 años	0,1340	1	0.147	1.144
Tiene pólizas de Vehículos	No tiene	-1,2242	1	0.000	0.294
Antigüedad Vehículos	Continua	0,1744	1	0.001	1.190
Antigüedad Varios	Continua	0,3558	1	0.000	1.427
Prima Directa	Menor a 16 UF	-0,7251	1	0.000	0.482
Constante	Constante	-4,2790	1	0.000	0.012

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

De la tabla 23 se concluye que los hombres tienen mayor probabilidad de suscribir una nueva póliza de Incendio. Las variables *Edad* y *Antigüedad* indican que para clúster de edad y antigüedad mayor aumenta la probabilidad de suscribir.

La curva de COR para el modelo de suscripción de una póliza de Incendio se detalla a continuación. El área bajo la curva azul que detalla el gráfico 26 es igual a 0.668. Es decir, el modelo gana 0.168 en sensibilidad y pronóstico sobre la probabilidad de lanzar una modelo aleatoriamente.

Gráfico 26: COR – Suscripción Incendio



Nota: Modelo logístico para el ajuste de la probabilidad de Suscribir una póliza de Incendio si tenía al menos una póliza de Vehículos o Varios

Suscripción Varios

La tabla 24 muestra los resultados del ajuste de la probabilidad de contratación de una póliza de Varios condicionada al hecho de no poseerla en el año anterior. La muestra tiene un número total de observaciones de 73,279 casos, correspondiente a los clientes que tenían una póliza del ramo Vehículos o Incendio a finales del 2009. De éstos, 612 contratan alguna póliza de Varios durante el año 2009, lo cual representa un 1% del total. Las variables incluidas en el modelo Género, *Edad*, *Tiene póliza de Vehículos*, *Antigüedad* y *Prima Directa en Cartera*.

Tabla 24: Modelo logístico para el ajuste de la probabilidad de Suscribir una póliza de Varios si tenía al menos una póliza de Vehículos o Incendio

Variables en la ecuación	Variables Categorizadas	Beta	gl	Sig.	Exp(B)
Genero	Femenino	-0,511	1	0,000	0,600
	Mayor a 58		2	0,000	
Edad	Hasta 38 años	-0,724	1	0,000	0,485
	Desde 38 a 58 años	-0,250	1	0,020	0,779
Prime Directa	Menor a 16 UF	-0,302	1	0,006	0,739
Antigüedad Total	Mayor a 2 años		2	0,005	
	Un año	-0,107	1	0,387	0,898
	Dos años	-0,419	1	0,001	0,658
¿Tiene una póliza de Vehículos?	No tiene	-1,856	1	0,000	0,156
¿Tiene una póliza de Incendio?	No tiene	-1,564	1	0,000	0,209
Constante	Constante	-2,916	1	0,000	0,054

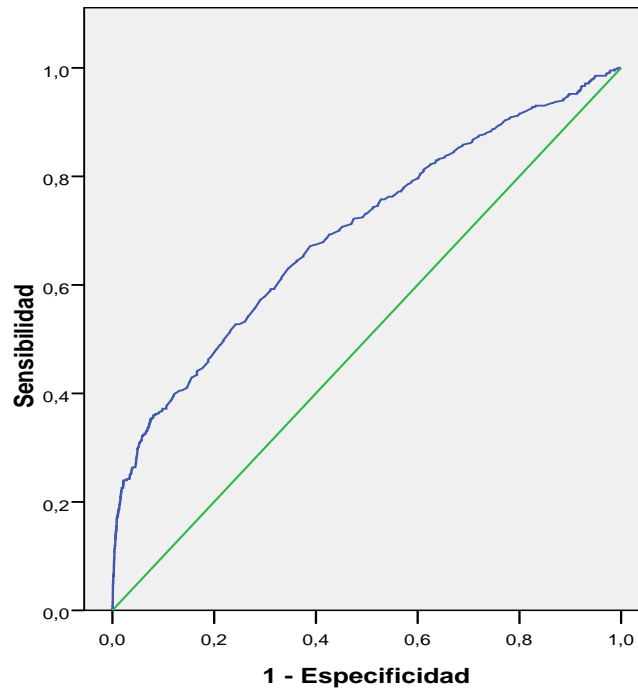
Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

De la tabla 24 se concluye que los hombres tienen mayor probabilidad de suscribir una nueva póliza de Varios. Las variables *Edad* y *Antigüedad* indican que para clúster de edad y antigüedad mayor aumenta la probabilidad de suscribir, es decir se premia la historia con la compañía en este caso.

Finalmente, el tener por lo menos una póliza de vehículos o Incendio aumenta la probabilidad de suscribir. Eso sí, relativamente la póliza de Incendio genera mayor probabilidad que la póliza de Vehículos.

La curva de COR para el modelo de suscripción de una póliza de Varios se detalla a continuación. El área bajo la curva azul que detalla el grafico 27 es igual a 0.696. Es decir, el modelo gana 0.196 en sensibilidad y pronóstico sobre la probabilidad de lanzar una modela aleatoriamente.

Gráfico 27: COR – Suscripción Varios



Nota: Modelo logístico para el ajuste de la probabilidad de renovar más de una póliza de Varios si estaban contratadas el año anterior

12. VALOR ESPERADO DEL CLIENTE PARA DIFERENTES PERFILES:

El presente capítulo detalla el valor histórico, esperado y potencial para diferentes perfiles de clientes. Estos perfiles se definen en función del Género, Antigüedad, Prima, etc.

Una vez ajustada la probabilidad de renovación y suscripción para cada ramo, se necesita conocer el valor histórico de estos clientes en la anualidad anterior, es decir, para el corte del año 2009. A partir de estos dos elementos se calcula el valor esperado de cada asegurado en cada ramo, multiplicando la probabilidad condicionada de renovación de las pólizas contratadas en ese ramo por el valor histórico del cliente en ese ramo durante el año anterior.

La Tabla 25 detalla el valor esperado por ramo y por modelo (1 o más pólizas). Los valores esperados para los ramos Vehículos e Incendio en el modelo una póliza son muy parecidos, por una parte las primas del ramo vehículo son más caras pero esto se compensa con la frecuencia siniestral que en el caso de Vehículos es mucho mayor. Por otro lado, el valor esperado asociado a las pólizas de Varios es prácticamente un tercio de los demás ramos, esto se debe a que las primas de Varios son generalmente más bajas y su siniestralidad en los últimos años no ha sido la mejor.

Tabla 25: Valor Histórico Promedio al año 2009

CLV (\$) Esperado	Ramo	Número de Pólizas	Media	Desviación Estándar	Mínimo	Máximo	N° de Clientes
General	Vehículos	Una póliza	67,838	10,834	12	102,111	55,068
		Dos o más pólizas	214,320	22,533	101,707	264,891	3,966
	Incendio	Una póliza	65,280	7,843	-	80,787	14,948.0
		Dos o más pólizas	90,954	6,350	67,424	104,442	2,407.0
	Varios	Una póliza	20,477	5,004	575	33,505	13,342.0
		Dos o más pólizas	111,765	17,115	16,411	137,449	1,215.0

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

Otra conclusión destacable de la tabla 25 es la diferencia de valor que tienen los clientes que tienen una póliza en contraposición a los que tienen más de una. Al ser los seguros patrimoniales un negocio probabilístico, las carteras con mayor número de productos aumenta el valor histórico de dicho cliente, pues la siniestralidad disminuye (grandes números).

La tabla 26 resume el valor esperado para la renovación de una póliza en Vehículo. En términos de Género la mujer es más rentable que el hombre, esta conclusión está en la línea de lo planteado en el capítulo anterior donde se observa que el género femenino aumenta la probabilidad de renovar la póliza de vehículos. Por otro lado, para grupos de Edad mayores aumenta el valor esperado.

Tener una póliza de Incendio o Varios y la Antigüedad en el ramo aporta valor esperado al cliente, pues crea incentivos sobre la renovación al tener una mayor cartera expuesta

Tabla 26: Valor Esperado para Renovación de una o más póliza de Vehículos

CLV Esperado (\$)	Media	Desviación Estándar	Mínimo	Máximo	N° de Clientes
Hombres	66.573	11.156	12	102.111	30.624
Menores de 35	63.424	10.832	1.916	101.770	8.568
De 35 a 45	65.698	10.830	12	102.065	9.923
De 45 a 55	68.691	11.025	7.601	101.466	8.407
Más de 55	71.370	10.417	347	102.111	3.726
Mujeres	69.423	10.199	3.727	102.077	24.444
Menores de 35	67.231	9.835	3.727	101.307	7.889
De 35 a 45	68.670	10.324	5.472	102.077	7.928
De 45 a 55	71.466	9.797	9.524	101.887	6.356
Más de 55	73.952	9.601	15.525	101.061	2.271
1 año antigüedad en VM	63.490	10.973	347	102.065	20.651
2 año antigüedad en VM	69.050	9.232	12	101.743	20.729
3 año antigüedad en VM	72.564	10.413	1.119	102.111	13.688
No tiene póliza de Incendio	67.715	10.788	12	102.111	53.777
Tiene póliza de Incendio	72.976	11.507	25.412	100.855	1.291
No tiene póliza de Varios	67.816	10.781	12	102.111	54.470
Tiene póliza de Varios	69.854	14.803	13.929	100.191	598

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

La tabla 27 resume el valor esperado para la renovación de una póliza de Incendio. Se han considerado cuatro variables para el análisis *Género, Edad, Antigüedad y Tener otros productos*. Al igual que en el punto anterior el género femenino aumenta el valor del cliente del mismo modo para aumentos en los grupos edad aumenta también el valor.

Para la póliza Incendio Tener una póliza extra de Vehículos o Varios no es determinante. En general la póliza de Incendio crea incentivos a la renovación de las demás pero no necesariamente las demás aumentan la probabilidad de renovación de ella.

Por el contrario, a lo evidenciado de la tabla 26 la antigüedad no genera necesariamente aumentos de valor sino que a mayor antigüedad disminuye.

Tabla 27: Valor Esperado para Renovación de una o más pólizas de Incendio

CLV Esperado (\$)	Media	Desviación Estándar	Mínimo	Máximo	N° de Clientes
Hombres	64.959	8.246	-	80.787	10.012
Menores de 35	65.809	6.510	39.344	78.950	395
De 35 a 45	65.149	8.586	-	80.787	1.974
De 45 a 55	64.748	7.894	78	80.000	3.475
Más de 55	64.965	8.508	3	80.496	4.168
Mujeres	65.930	6.909	289	80.395	4.936
Menores de 35	66.575	6.856	36.775	76.584	198
De 35 a 45	66.456	6.661	12.099	80.395	1.010
De 45 a 55	65.439	6.922	289	78.924	1.822
Más de 55	66.054	7.005	4.564	78.909	1.906
1 año antigüedad en IN	69.231	7.106	78	80.787	4.533
2 año antigüedad en IN	64.305	9.095	51	78.272	2.576
3 año antigüedad en IN	63.315	6.909	-	77.575	7.839
No tiene póliza de Vehículos	65.158	7.802	-	80.496	13.763
Tiene póliza de Vehículos	66.689	8.175	2.546	80.787	1.185
No tiene póliza de Varios	65.254	7.749	-	80.000	14.538
Tiene póliza de Varios	66.183	10.640	15.827	80.787	410

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

La tabla 28 resume el valor esperado para la renovación de una póliza de Varios. Se consideran cuatro variables para el análisis *Género, Edad, Antigüedad y Tener otros productos*.

Para este modelo no es claro que el género femenino tenga mayor valor, la diferencia no es concluyente, incluso el valor asociado al género masculino es levemente mayor. En la misma línea, la variable *Edad* tampoco es concluyente en el análisis, pues los valores esperados.

Finalmente, *Tener una póliza de Incendio* genera mayor valor que el *Tener una póliza de Vehículos* en este modelo, esto confirma las intuiciones planteadas en el modelo inmediatamente anterior.

Tabla 28: Valor Esperado para Renovación de una o más pólizas de Varios

CLV Esperado (\$)	Media	Desviación Estándar	Mínimo	Máximo	N° de Clientes
Hombres	20.512	4.954	575	33.505	11.083
Menores de 35	21.450	4.646	13.674	30.135	474
De 35 a 45	21.972	5.155	11.257	33.505	1.342
De 45 a 55	19.988	5.209	9.474	33.458	4.128
Más de 55	20.466	4.622	575	30.240	5.139
Mujeres	20.300	5.238	5.161	33.098	2.259
Menores de 35	21.326	4.790	11.153	28.631	89
De 35 a 45	21.734	5.399	12.436	33.098	287
De 45 a 55	19.665	5.504	6.101	32.601	949
Más de 55	20.407	4.833	5.161	30.161	934
1 año antigüedad en Varios	16.093	1.956	4.391	28.512	6.444
2 año antigüedad en Varios	21.394	1.766	11.853	31.833	2.843
3 año antigüedad en Varios	26.799	1.735	575	33.505	4.055
No tiene póliza de Vehículos	20.260	4.915	575	33.458	12.619
Tiene póliza de Vehículos	24.251	5.033	11.573	33.505	723
No tiene póliza de Incendio	20.168	4.798	575	31.068	12.846
Tiene póliza de Incendio	28.471	3.289	14.707	33.505	496

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

Valor Potencial:

En este caso, se considera la posibilidad de crecimiento de un cliente en un ramo distinto a aquéllos en los que tiene contratada alguna póliza. Se comenzó ajustando a través de un modelo logístico la probabilidad de contratación de una póliza de Vehículos condicionada que el cliente no tuviera ninguna suscrita en ese ramo el año anterior. Se utiliza como variables explicativas, la *Antigüedad*, *Género*, *Tenencia de Siniestros*, *Prima en cartera*, etc.

Se debe notar que el volumen de las observaciones para los modelos de suscripción está muy lejos de lo evidenciado en los modelos precedentes. Por lo tanto, el análisis apunta a la sensibilidad y la interpretación de la dirección de las variables más que al pronóstico exacto.

En general de la tabla 29 se concluye que el *Tener una póliza de Incendio* prácticamente dobla al valor de *Tener una póliza de Varios* a la hora de cuantificar el valor de la suscripción de una póliza de Vehículos. La variable *Edad* genera un aumento en el valor potencial en el segundo grupo (35 a 45 años), esto puede deberse a que dicho grupo de edad tiene mayor probabilidad de comprar sus primeros activos como son una casa, un auto, etc.

Tabla 29: Valor potencial de Suscripción de una póliza de Vehículos

CLV (\$) Potencial	Genero	Edad	Media	Desviación Estándar	Mínimo	Máximo	N° de Clientes
Tiene póliza de Incendio	Mujer	Menores de 35	3.898	1.739	3.285	15.592	216
		De 35 a 45	5.052	3.226	3.285	31.790	1.137
		De 45 a 55	4.320	2.454	3.285	46.093	1.896
		Más de 55	4.039	1.755	3.285	22.056	1.916
	Hombre	Menores de 35	4.047	2.114	3.285	30.302	389
		De 35 a 45	5.017	2.997	3.285	42.276	2.323
		De 45 a 55	4.421	2.622	3.285	35.419	3.692
		Más de 55	4.099	1.968	3.285	22.056	4.162
Tiene póliza de Varios	Mujer	Menores de 35	2.131	899	1.450	7.976	482
		De 35 a 45	2.273	1.153	1.450	11.719	380
		De 45 a 55	2.391	1.479	1.606	16.904	453
		Más de 55	2.043	615	1.970	7.976	767
	Hombre	Menores de 35	2.050	644	1.450	7.976	2.653
		De 35 a 45	2.197	963	1.450	14.465	1.662
		De 45 a 55	2.438	1.556	1.606	32.391	2.442
		Más de 55	2.016	509	1.450	11.719	4.265

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

Tabla 30: Valor potencial de Suscripción de una póliza de Incendio

CLV (\$) Potencial	Genero	Edad	Media	Desviación Estándar	Mínimo	Máximo	N° de Clientes
Tiene póliza de Varios	Mujer	Menores de 35	206	291	135	7.520	1.928
		De 35 a 45	206	208	135	3.988	4.802
		De 45 a 55	187	168	135	7.520	8.462
		Más de 55	190	145	135	3.988	9.879
	Hombre	Menores de 35	381	548	238	11.095	2.705
		De 35 a 45	411	628	238	11.095	7.827
		De 45 a 55	340	394	238	11.095	9.562
		Más de 55	340	362	238	11.095	12.219
Tiene póliza de Vehículos	Mujer	Menores de 35	124	63	71	902	463
		De 35 a 45	128	168	71	2.773	1.266
		De 45 a 55	116	125	71	2.773	4.187
		Más de 55	122	126	71	2.773	5.106
	Hombre	Menores de 35	67	30	40	106	100
		De 35 a 45	69	68	40	994	225
		De 45 a 55	62	64	40	1.639	790
		Más de 55	71	91	40	1.639	967

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

Tabla 31 Valor potencial de Suscripción de una póliza de Varios

CLV (\$) Potencial	Genero	Edad	Media	Desviación Estándar	Mínimo	Máximo	N° de Clientes
Tiene póliza de Varios	Mujer	Menores de 35	104	89	40	578	216
		De 35 a 45	137	166	38	3.228	1.113
		De 45 a 55	114	166	36	3.999	1.829
		Más de 55	120	158	40	2.504	1.913
	Hombre	Menores de 35	199	291	65	3.533	386
		De 35 a 45	232	376	51	11.520	2.249
		De 45 a 55	162	184	58	3.073	3.470
		Más de 55	172	189	61	4.144	4.161
Tiene póliza de Vehículos	Mujer	Menores de 35	160	55	100	1.135	1.828
		De 35 a 45	244	658	87	22.859	5.420
		De 45 a 55	121	54	88	2.076	8.206
		Más de 55	132	35	103	1.555	9.851
	Hombre	Menores de 35	259	143	153	3.384	2.788
		De 35 a 45	518	1.123	136	27.177	8.084
		De 45 a 55	194	106	136	5.172	9.518
		Más de 55	209	95	162	3.905	12.247

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

La tabla 32 que se detalla a continuación contiene los resultados agregados del modelo Esperado y Potencial

Se observa que para grupos de *Edad* mayores el valor del cliente aumenta, este aumento en el valor está explicado por un aumento en la permanencia de dichos clientes. El *Género* femenino tiene mayor valor que el masculino.

Finalmente, el valor monetario de un cliente sin diferenciar está en torno a los 89,614 pesos, esta cifra es una muy buena herramienta para cuantificar lo que es rentable de invertir en posibles campañas de marketing. De todas maneras, la clasificación y perfiles planteados ayudaran a optimizar lo invertido.

Tabla 32: Valor del Cliente Esperado más Potencial

CLV (\$) Esperado	Media	Desviación Estándar	Mínimo	Máximo
General	89.614	42.488	-	264.891
Mujeres	97.556	35.267	289	264.811
Hombres	84.945	45.568	-	264.891
Menores de 35	99.492	16.070	2.921	155.180
De 35 a 45	102.189	16.332	19	155.649
De 45 a 55	106.563	16.167	11.590	155.359
Más de 55	110.317	15.542	529	155.700

Nota: Todos los productos tanto los que tienen una póliza como los que tienen más de una

Fuente: Elaboración Propia a partir de datos aportados por la compañía estudiada.

13. CONCLUSIONES GENERALES

El análisis de la fidelidad y el valor del cliente es el punto de partida para la gestión de las carteras en las compañías de seguros. En este trabajo, se ha propuesto una primera aproximación a esta cuantificación.

Los modelos desarrollados en este trabajo apuntan al análisis de las sensibilidades y dirección de las variables explicativas. Lo ganado en términos de pronóstico no es lo esperado pero constituye el primer paso en la tarea de construir mejores bases centradas en el cliente por la compañía de seguro estudiada.

A partir de los modelos propuestos se concluye que la variable que condiciona más la no renovación es *Tener un siniestro*. Por lo tanto, este contacto entre la compañía y el cliente es central para su fidelidad, debería ser tarea primordial de la compañía mejorar la experiencia frente a este hecho. A modo de ejemplo, en los años 90 la compañía Estadounidense Progressive identificó este punto, y creó toda una división centrada en la mejora de la experiencia y procesos asociados al siniestro. Una década después la compañía multiplicó por tres su participación de mercado en EEUU y su valorización se multiplicó por 10 [9].

La póliza que tiene la mayor capacidad de generar Cross selling es la asociada al producto Incendio, los clientes que tienen esta póliza tienen mayores probabilidades de adquirir otras. Por lo tanto, se recomienda que la compañía estudiada ofrezca nuevos productos a estos clientes y observe las tasas de interés y compra de dichas campañas.

La diferenciación de Prima es un elemento determinante en la carrera por fidelizar al cliente. Los modelos desarrollados han arrojado que los clientes que tienen primas más específicas tienen a renovar con mayor probabilidad. Es decir, cuando el producto recibe diferenciación entonces el cliente le cuesta más dejar la compañía. Desde este punto de vista, la compañía estudiada debería estar muy abierta e incluso incentivar las cotizaciones de bienes específicos que piden los clientes.

La variable *Antigüedad* es considerada determinante, ya que es significativa en prácticamente todos los modelos. De todas maneras, el *Tener Siniestro* distorsiona en alguna medida los grados de fidelidad alcanzados por medio de la Antigüedad del cliente.

En términos de valor del cliente se observa que para grupos de *Edad* mayores el valor del cliente aumenta, este aumento en el valor está explicado por un aumento en la permanencia de dichos clientes. El *Género* femenino en prácticamente todos los modelos aumenta el valor esperado, sólo el modelo asociado al ramo Incendio deja a esta variable fuera de la ecuación.

Una práctica muy interesante a desarrollar una vez que hayan mejorado los modelos de proyección, es que la compañía pague comisiones diferenciadas por perfil al corredor de seguros. Esto permite evadir el sesgo que produce hoy el intermediador y dota a la compañía de perfiles que le son más gratos en sobrevivencia como siniestralidad.

Finalmente, es importante resaltar que la rentabilidad del cliente tráficamente en seguros patrimoniales a estado asociada a la Siniestralidad. Las compañías de seguros en Chile han avanzado muchísimo en el estudio de las variables que justifican la siniestralidad pero se han quedado cortos en el estudio de las variables que justifican la retención, suscripción y sobrevivencia. Por tanto, existe un gran espacio para la optimización de los modelos de riesgo al incluir dichas variables.

Trabajos Futuros

Finalmente, se recomienda mejorar este trabajo con un cambio en la unidad de análisis. Hoy el objeto de estudio debe ser el hogar y no el contratante, los últimos trabajos en la materia se han centrado en la familia como tomador de productos de seguro, pues todos los miembros adultos participan en la decisión de la cobertura de riesgos comunes como son el hogar, el automóvil u objetos valiosos que deseen proteger. Sin embargo, la existencia de múltiples contratos asociados a un mismo hogar complejiza drásticamente el análisis, por la

numero combinaciones de productos asegurados y por la variedad de fechas de vigencia que podrían tener una cartera de pólizas.

Otra línea de investigación interesante seria estudiar el comportamiento del asegurado en las distintas pólizas contratadas. Definir probabilidades de transición entre distintas composiciones de pólizas, quizás con procesos de markov.

14. CITAS BIBLIOGRÁFICAS

[1] Berger, P.D. Y Nasr, N.L. (1998). “Customer Lifetime Value: Marketing Models and Applications”, *Journal of Interactive Marketing*, Vol. 12, n° 1, pgs. 18–30.

[2] Dwyer, F.R. (1997). “Customer lifetime valuation to support marketing decision making”, *Journal of Direct Marketing*, Vol. 11, n° 4, pgs. 8-15.

[3] Donkers, B.; Verhoef, P. Y de Jong, M. (2007). “Modeling CLV: A test of competing models in the insurance industry”, *Quantitative Marketing and Economics*, Vol. 5, n° 2, pgs. 163–190.

[4] Kotler, Philip; Armstrong, Gary (1990). “Principles of Marketing”. Prentice Hall (4ª edición) . Parte II, pg 158

[5] Guillén, M., Nielsen, J.P. and Pérez-Marín, A.M. (2006) “La duración de distintos contratos de seguros en los hogares. Un enfoque integrado” *Gerencia de Riesgos y Seguros*, 96, 4, 23-32.

[6] Guillén, M., Nielsen, J.P. and Pérez-Marín, A.M. (2009) “Cross-buying behaviour and customer loyalty in the insurance sector” *ESIC Market*, 132, 77-136.

[7] Hurtado, C. Abril 2005. Aprendizaje en motores de búsqueda en la web. Congreso de Ciencia y Tecnología Computacional. Tulúa.

[8] Lewis, M. 2005. Incorporating Strategic Consumer Behavior into Customer Valuation. *Journal of Marketing*, Vol. 69 Issue 4. p 230-138.

[9] Lutova, Marina; Sender, Ilan; Wells J. R. The Progressive Corp. HARVARD BUSINESS PUBLISHING

[10] Schargrotsky, Ernesto 2008. “Innovación y Segmentación de Mercado: El caso Progressive Corporation”. Universidad Torcuato Di Tella

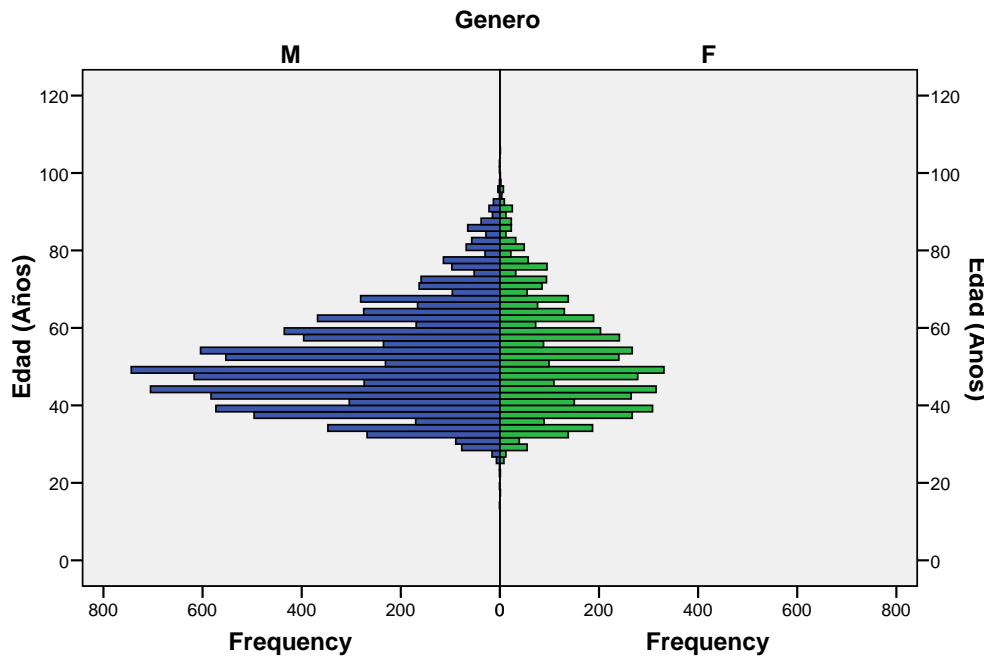
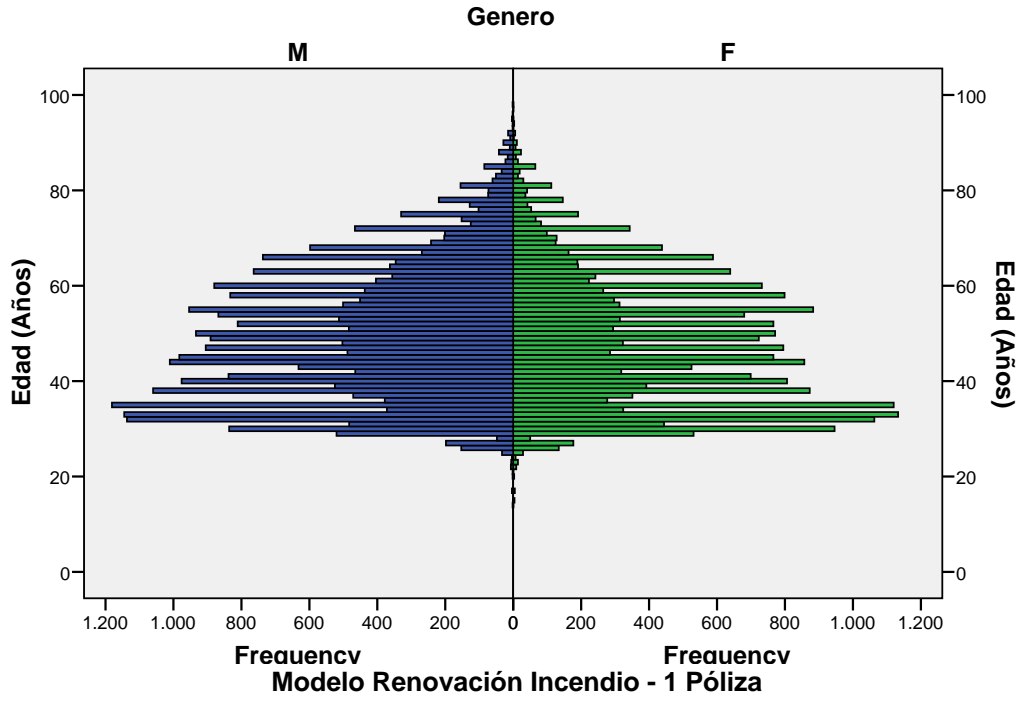
[11] Marín, P. 2005. Estimación de lifetime value basado en comportamiento transaccional. Memoria Ingeniería Civil Industrial. Universidad de Chile.

[12] Ryals, L. J. & Knox, S. (2005). Measuring risk- adjusted customer lifetime value and its impact on relationship marketing strategies and shareholder value. *European Journal of Marketing*, 39, 5/6, 456-472

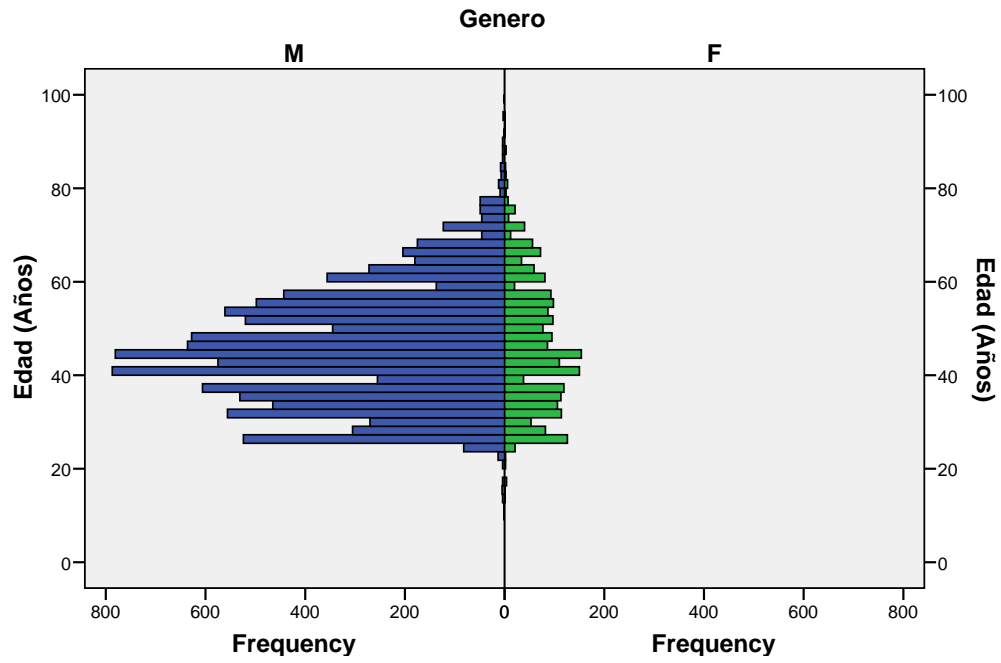
ANEXOS

ANEXO A: FRECUENCIAS DE MUESTRA

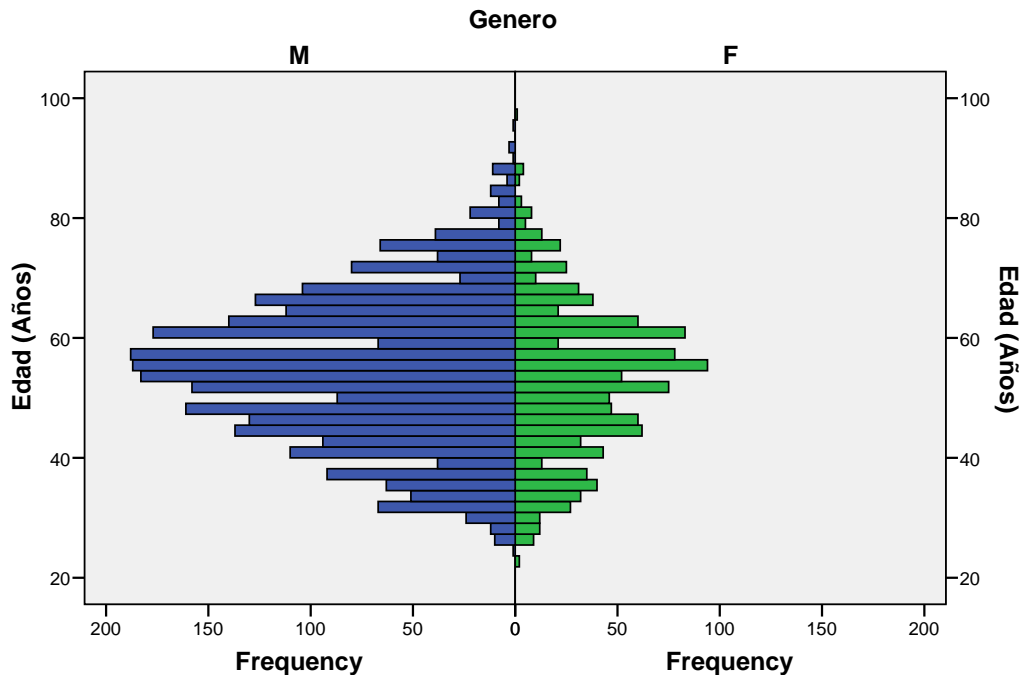
Modelo Renovación Vehículo - 1 Póliza



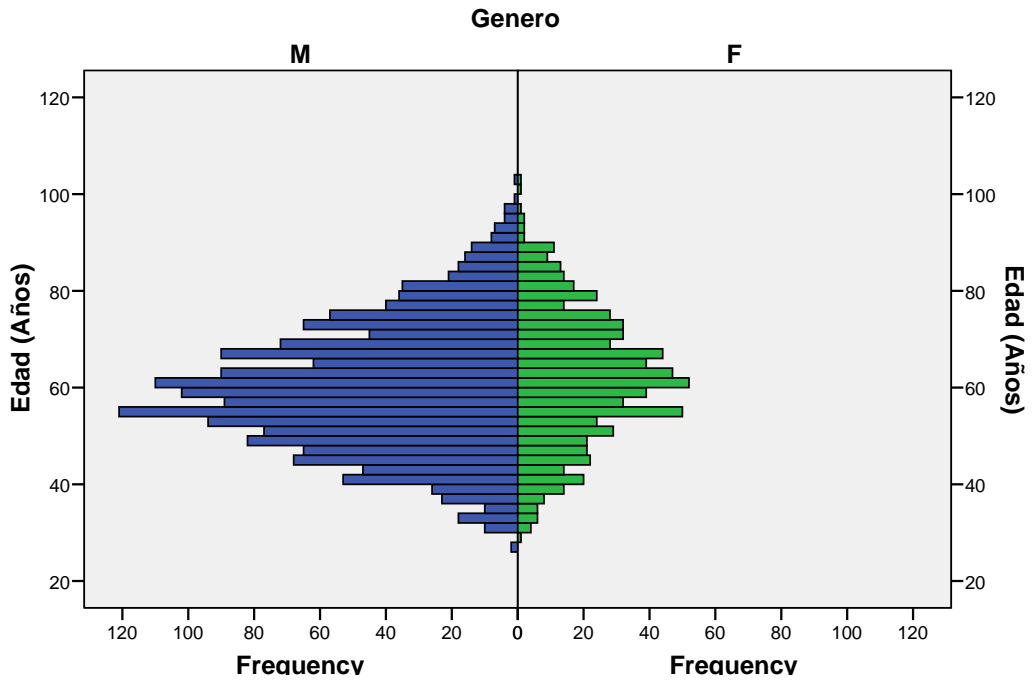
Modelo Renovación Varios - 1 Póliza



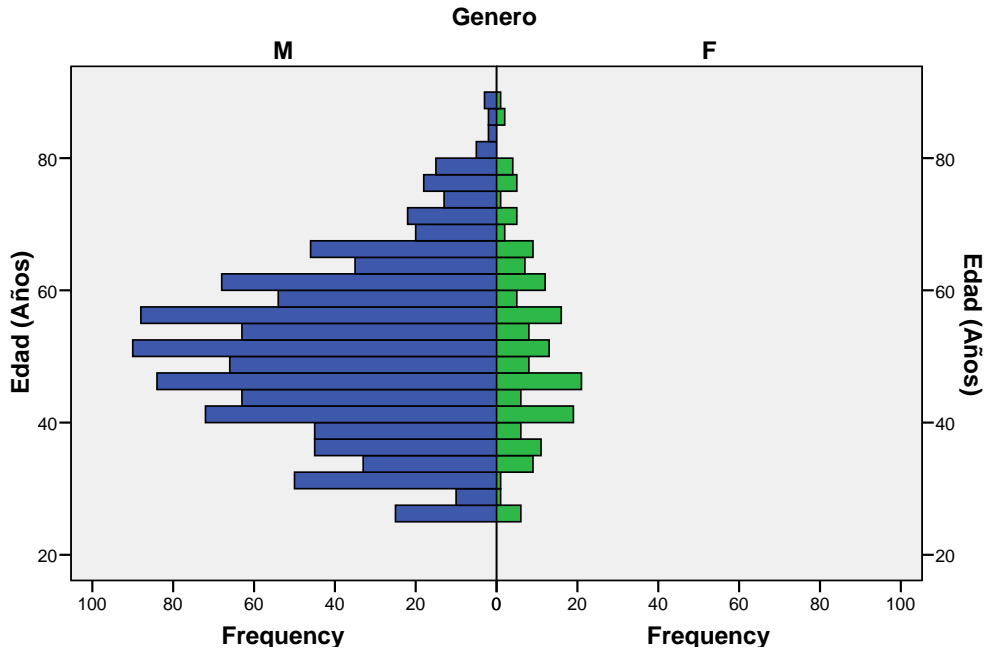
Modelo Renovación Vehículos - Más de 1 Póliza



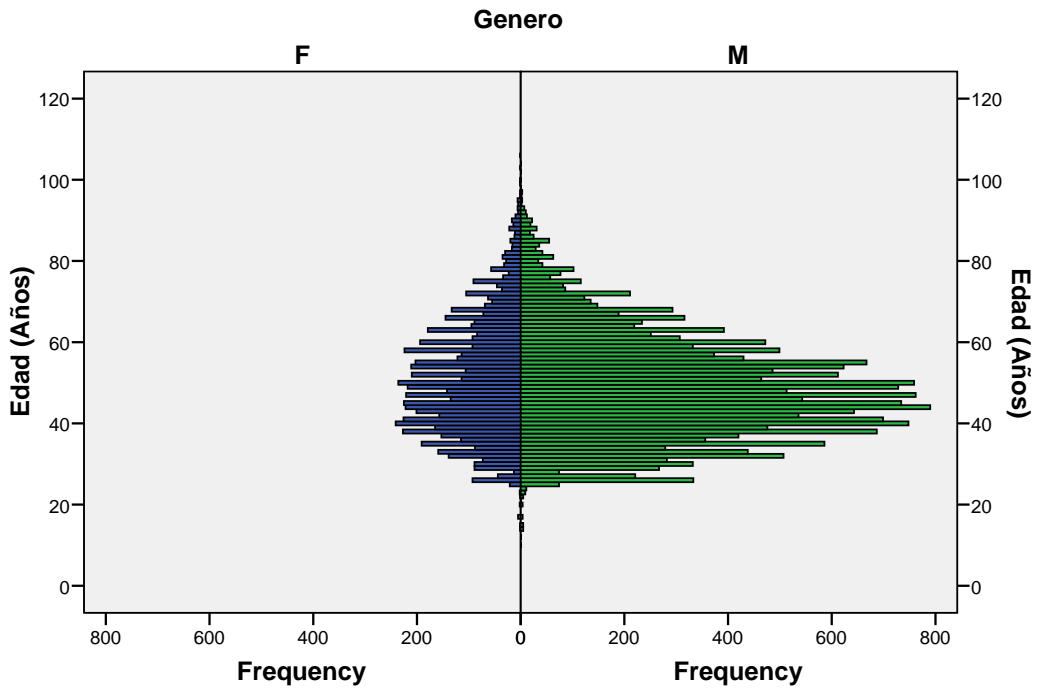
Modelo Renovación Incendio - Más de 1 Póliza



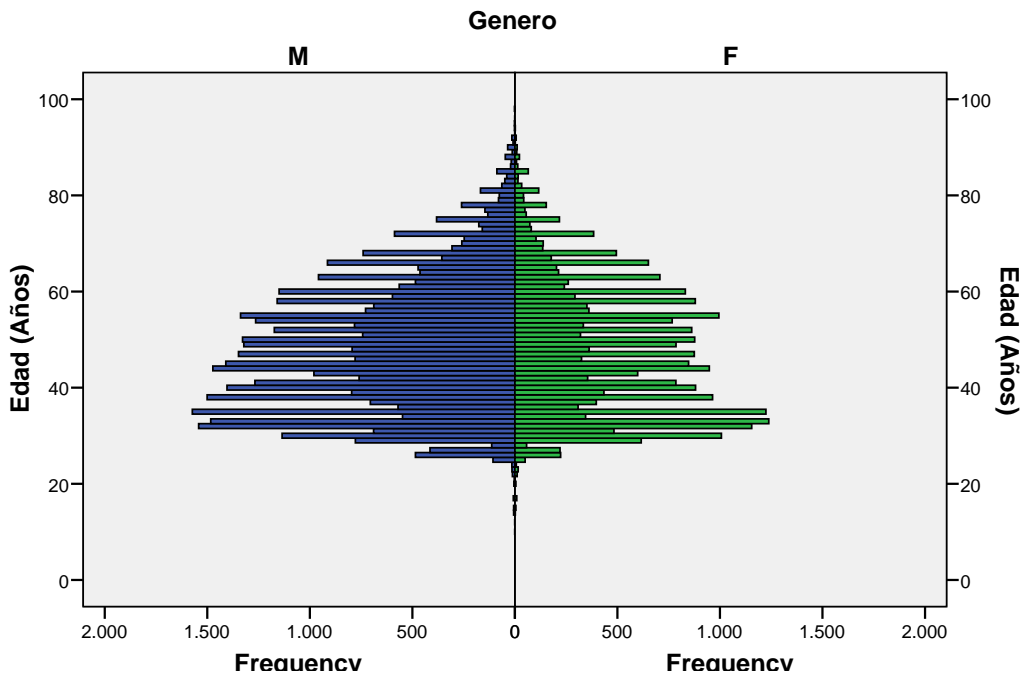
Modelo Renovación Varios - Más de 1 Póliza



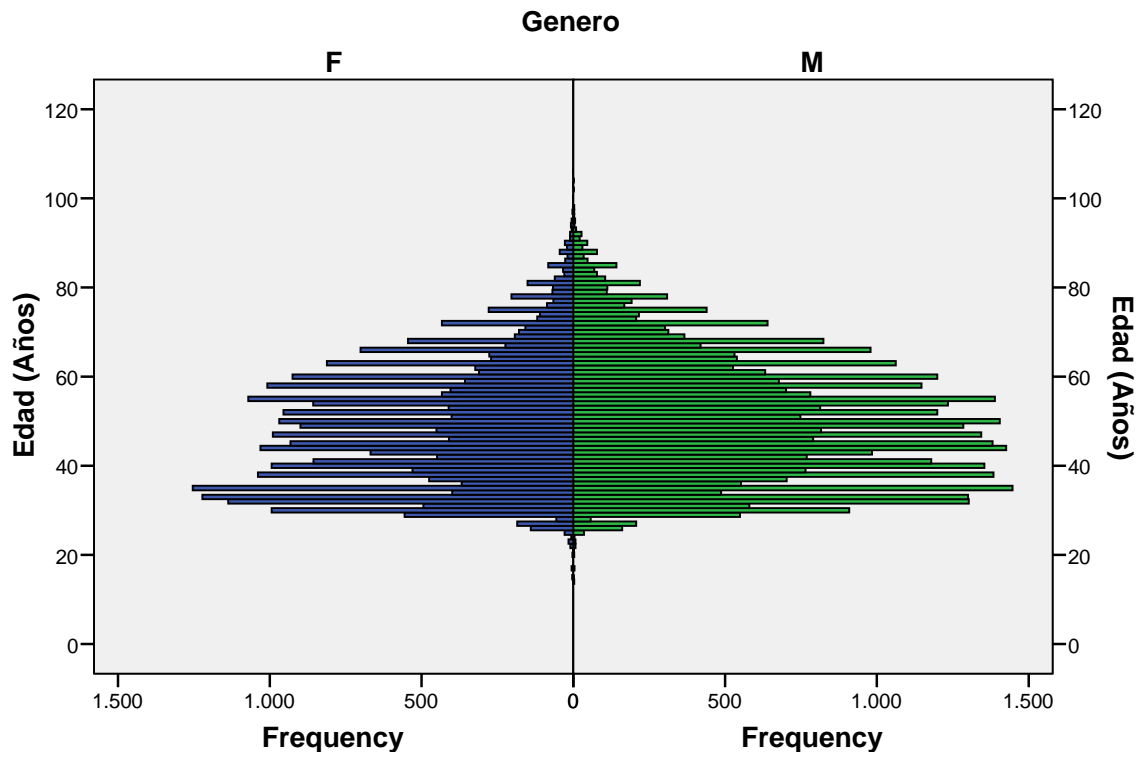
Modelo Suscripción Vehículos



Modelo Suscripción Incendio

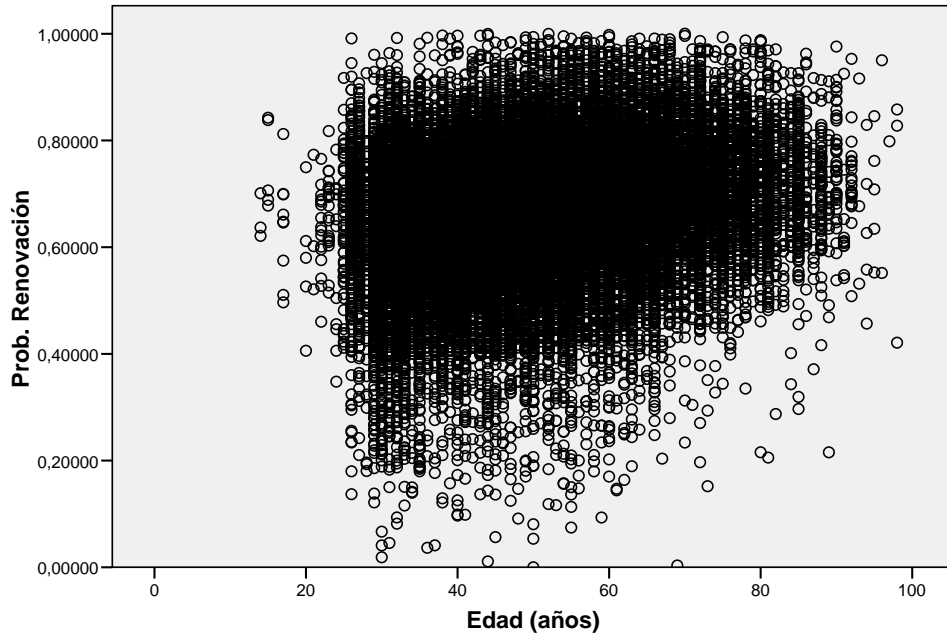


Modelo Suscripción Varios

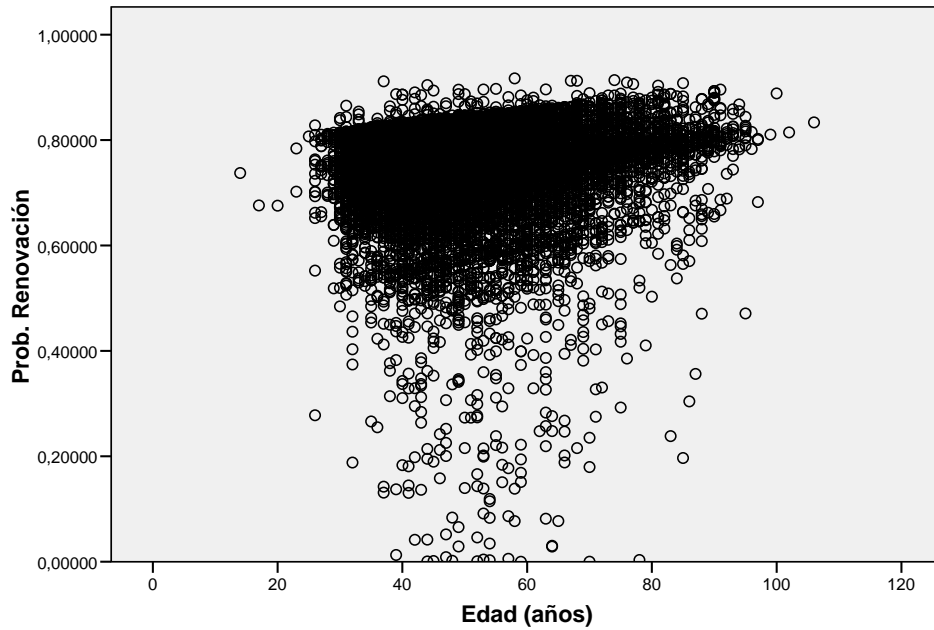


ANEXO B: GRAFICOS DE DISPERSIÓN – VARIABLE PROBABILIDAD

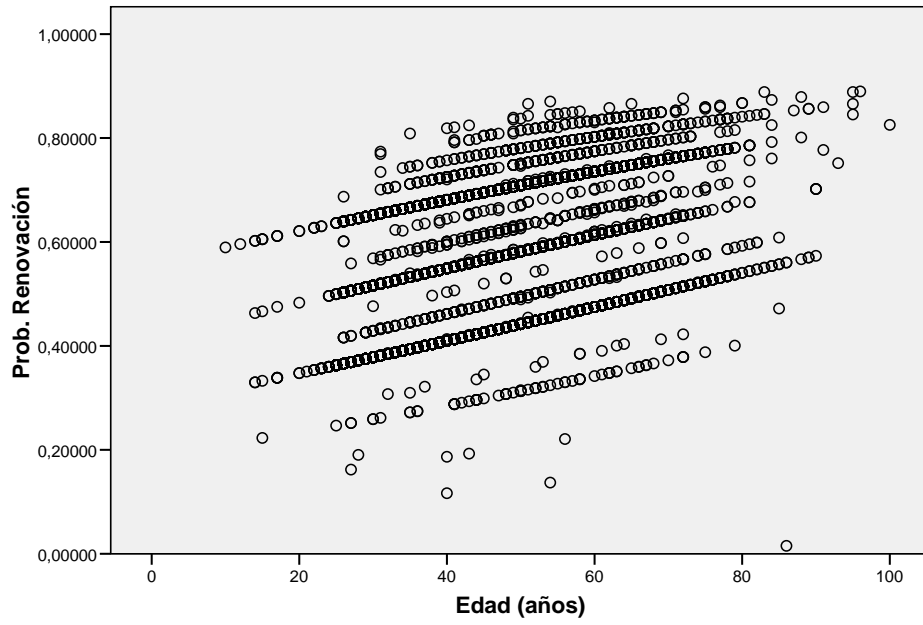
Renovación Vehículos - 1 Póliza



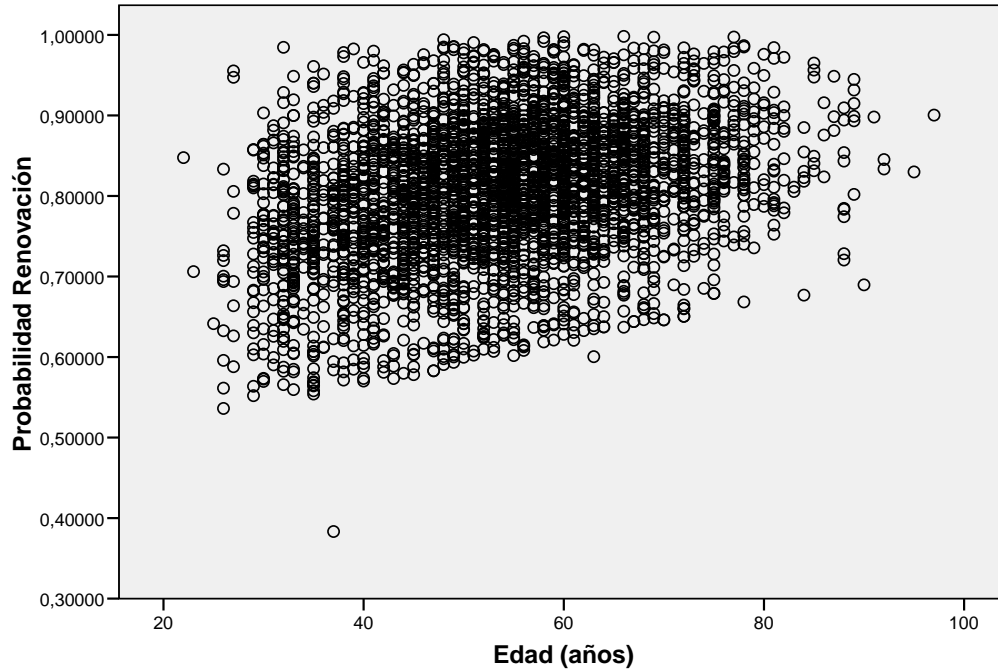
Renovación Incendio - 1 Póliza



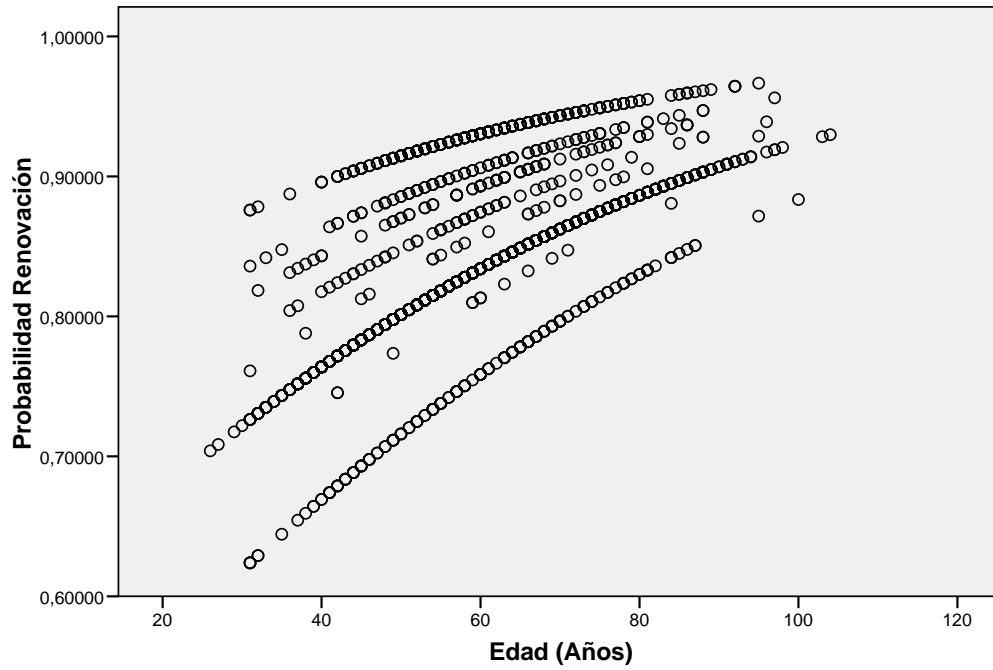
Renovación Varios - 1 Póliza



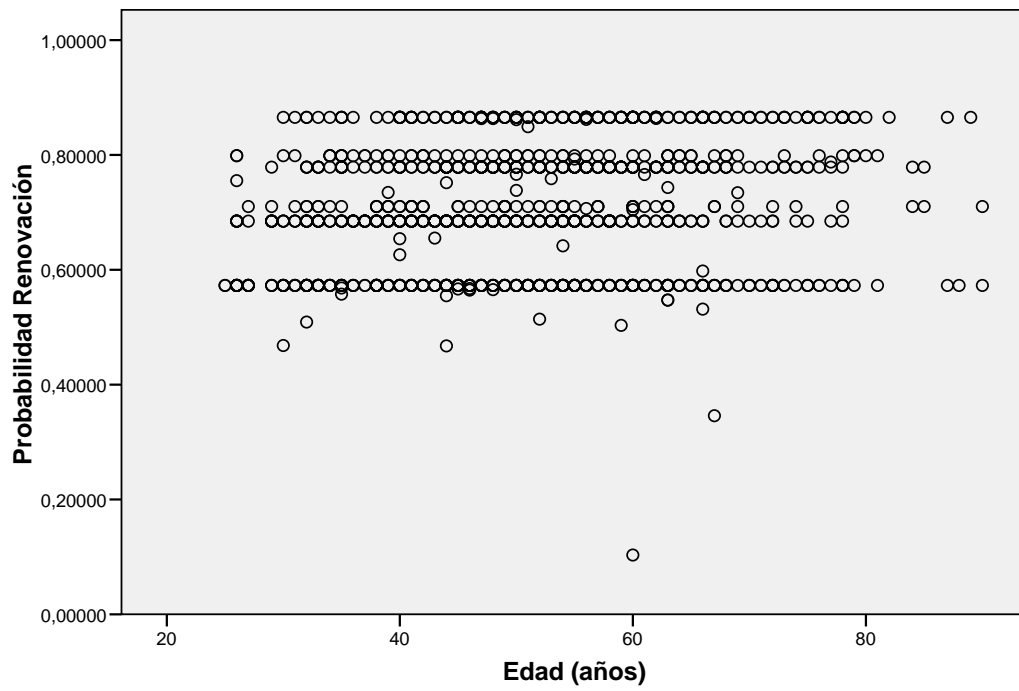
Renovación Vehículos - Más de una Póliza



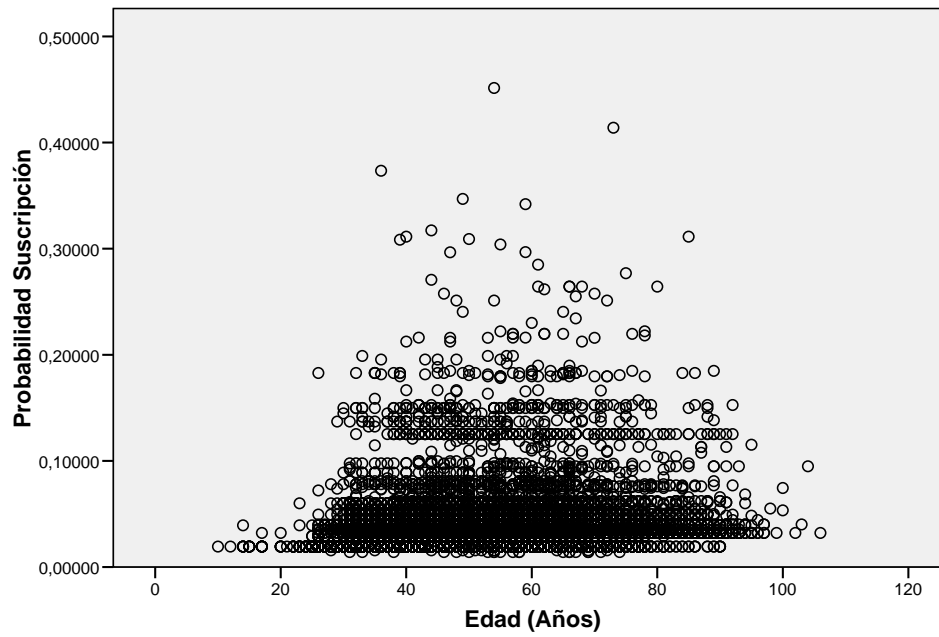
Renovación Incendio - Más de una Póliza



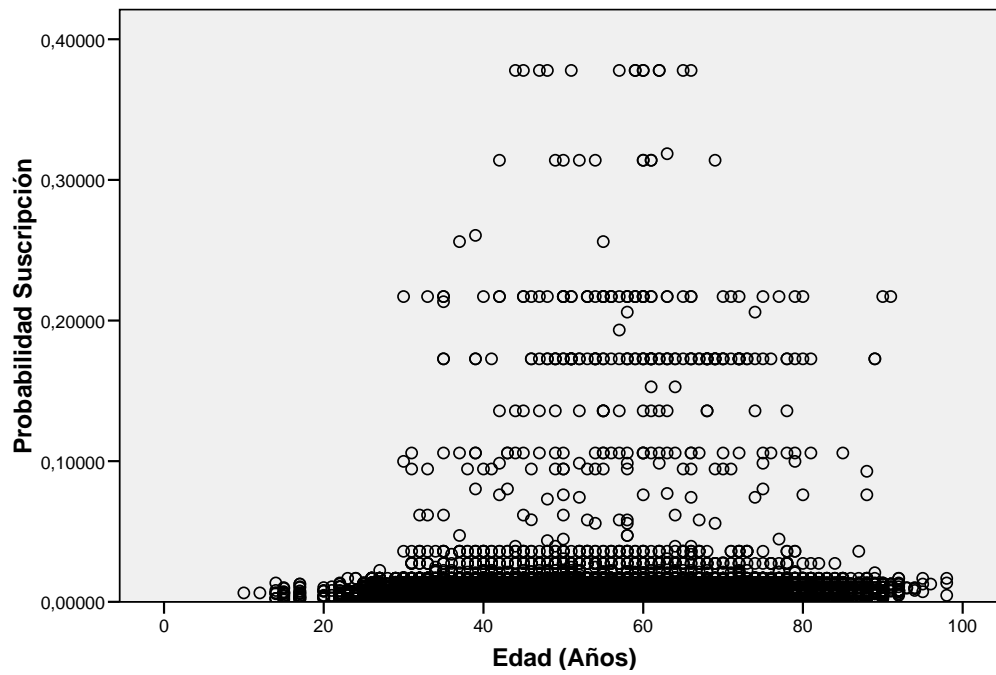
Renovación Varios - Más de una Póliza



Suscripción Vehículos - Al menos 1 Póliza



Suscripción Incendio - Al menos 1 Póliza



Suscripción Varios - Al menos 1 Póliza

