



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

**EVALUACIÓN DE PROMOCIONES EN LA RETENCIÓN DE UN SEGMENTO DE
CLIENTES DE UNA TIENDA POR DEPARTAMENTO**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

LUIS ALBERTO CARDOZA LANDABUR

PROFESOR GUÍA:

ALEJANDRA PUENTE CHANDÍA

MIEMBROS DE LA COMISIÓN:

RICARDO MONTOYA MOREIRA

CAROLINA NAVARRETE CAEROLS

SANTIAGO DE CHILE

2018

EVALUACIÓN DE PROMOCIONES EN LA RETENCIÓN DE UN SEGMENTO DE CLIENTES DE UNA TIENDA POR DEPARTAMENTO

Una organización busca generar una relación de largo plazo con sus clientes, pues el éxito de ellas no sólo se centra en captar nuevos consumidores, sino que también, en retener y satisfacer a los actuales. La industria del retail ha mostrado un gran interés por la retención de los clientes, debido a los altos costos asociados a la adquisición de clientes.

Este trabajo de título busca proponer campañas de email marketing para gestionar a los clientes de una tienda por departamento, con el fin de aumentar la retención. Para lograr lo señalado, se comienza con un modelo de propensión para dos segmentos de clientes (antiguos y nuevos), que permitirá seleccionar a aquellos clientes con mayor riesgo de fuga. Luego, se realiza un diseño experimental, que tiene por objetivo evaluar el efecto de un conjunto de promociones, buscando maximizar la contribución de cada campaña, ya sea mediante el aumento de visitas o de gasto de los clientes. De esta manera, se plantearon diversas formas de abordar a los clientes, de acuerdo a tres segmentos relevantes: antigüedad del cliente, tenencia de la tarjeta de crédito del holding y recency.

Con respecto a los modelos de propensión para clientes antiguos y nuevos, se obtiene un *accuracy* de 76% y 68% y, un *lift* de 2.6 y 1.6, respectivamente. Las variables más importantes que explican el fenómeno estudiado son: cantidad de compras realizadas en el último año, antigüedad del cliente, estado de la tarjeta de crédito, recency, entre otras.

Se evidencia que las promociones asociadas a montos de descuentos logran una mayor activación que otro tipo de incentivos (entre 2% y 5%). Asimismo, destaca que aquellas que proporcionan un descuento sujeto a un ticket mínimo, son consideradas como una barrera para los clientes que no han comprado hace más de siete meses en la tienda.

Con respecto a las propuestas de trabajos futuros, se evidencia la necesidad de incorporar información de encuestas de lealtad y comentarios que realizan los clientes en redes sociales para encontrar nuevas características que permitan mejorar el modelo realizado, junto con proporcionar *insights* importantes asociados a la calidad de servicio y motivos por los cuales un cliente deja de comprar en la tienda por departamento.

Finalmente, se sugiere probar otros canales de contacto como mensajes de texto (SMS) e, incluso, redes sociales, lo que aumentaría considerablemente el alcance de la gestión realizada, en especial para aquellos clientes donde no se posee información de contactabilidad.

AGRADECIMIENTOS

Quiero agradecerle a mis padres, Angélica y Luis, por todo el apoyo incondicional que me han brindado en las etapas de mi vida, tanto personal como profesional, entregándome confianza y valores para superarme cada día y ser una mejor persona. Siempre recordaré todos esos momentos de alegrías que vivimos juntos y aquellos donde me animaban y me daban fuerza para seguir adelante día a día. Son y siempre serán un pilar fundamental para mí.

También quiero agradecer a mi familia, en especial a mi abuelita y tía, por creer en mí, por el ánimo que siempre me daban y apoyarme, aunque sea a la distancia, por cada desafío que tuve que enfrentar.

A mis amigos de la Universidad, Álvaro, Cristian, Iván P., Iván Q., Mario y Osvaldo, espero que estos cuatro años que estuvimos juntos, sean el comienzo de una gran amistad y de muchas historias venideras.

A mis profesores, Alejandra y Ricardo, quiénes me prestaron toda la ayuda posible para llevar a cabo mi trabajo de título. Gracias por la paciencia y por todos los consejos que me entregaron sobre distintas formas de abordar el trabajo. En especial a mi profesora Alejandra, quién me brindó apoyo dentro y fuera de la sala de clases.

Al equipo de Inteligencia de Clientes, Carolina, Camila, Ignacio, Jorge, Pablo y Valentina, quiénes hicieron que esta experiencia estuviese llena de aprendizajes y buenos recuerdos. Agradezco que me hayan incorporado al área y brindado la mejor ayuda posible cada vez que lo necesitaba.

Gracias a todos ustedes con quiénes compartí esta etapa universitaria.

TABLA DE CONTENIDO

1. Introducción.....	1
1.1. Antecedentes generales de la industria	1
1.2. Antecedentes generales de la empresa	2
2. Descripción del proyecto.....	3
2.1. Justificación del estudio.....	3
2.2. Objetivos.....	5
2.2.1. Objetivo general.....	5
2.2.2. Objetivos específicos	5
2.3. Resultados esperados	5
2.4. Alcances	5
3. Marco conceptual.....	6
3.1. Criterio de fuga	6
3.1.1. Criterio comercial	6
3.1.2. Criterio RFM.....	6
3.2. Modelo de fuga.....	7
3.2.1. Árboles de decisión.....	7
3.3. Distancia de Mahalanobis	9
3.4. Diseño de experimentos	10
3.4.1. Promociones	10
3.4.2. Validez de experimentos	10
3.4.3. Muestreo.....	11
4. Metodología	11
4.1. Selección de datos	12
4.2. Pre-procesamiento	12
4.3. Transformación.....	12
4.4. Análisis exploratorio	13
4.5. Árboles de decisión.....	13
4.6. Diseño experimental	13
4.6.1. Objetivo	13
4.6.2. Hipótesis.....	13
4.6.3. Variable experimental	14
4.6.4. Condiciones experimentales.....	14
4.7. Propuesta de gestión	14
5. Desarrollo metodológico.....	14
5.1. Selección de datos	15
5.2. Pre-procesamiento de datos	16
5.3. Transformación.....	18

5.4. Análisis exploratorio	18
5.5. Modelo de fuga.....	28
5.5.1. Árboles de decisión.....	28
5.6. Diseño experimental	35
5.6.1. Muestra de clientes	35
5.6.2. Tamaño muestral	36
5.6.3. Promociones realizadas.....	36
5.6.4. Grupo control y Tratamiento	37
5.6.5. Tasa de activación de clientes	38
5.6.6. Venta incremental	39
5.6.7. Caracterización de clientes.....	40
5.6.1. Test de Hipótesis	45
5.6.2. Análisis de Asistencia Incremental	50
5.6.3. Análisis de Regresión Lineal	57
5.7. Propuesta de gestión	64
5.7.1. Según tipo de cliente	64
5.7.2. Según tenencia de la tarjeta de crédito	65
5.7.3. Según recency	66
6. Impacto económico	68
7. Conclusiones.....	69
8. Trabajos futuros.....	70
9. Bibliografía.....	71
10. Anexos.....	74
10.1. Anexo 1. Árbol de decisión para clientes antiguos	74
10.1. Anexo 2. Árbol de decisión para clientes nuevos.....	75
10.2. Anexo 3. Análisis de sensibilidad de la probabilidad de corte.	76
10.3. Anexo 4. Sistema de promociones	77
10.4. Anexo 5. Gráfica	80
10.1. Anexo 6. Impacto económico.....	81

ÍNDICE DE TABLAS

Tabla 1: Características del método de crecimiento	9
Tabla 2: Esquema de condiciones experimentales.....	14
Tabla 3: Variables disponibles.	15
Tabla 4: Variables calculadas	16
Tabla 5: Variables creadas para evitar missing values.....	17
Tabla 6: Casos evaluados para la determinación de valores extremos para clientes antiguos y nuevos.....	17
Tabla 7: Casos evaluados para la definición de valores extremos de clientes nuevos. .	17
Tabla 8: Distribución de clientes antes y después de la eliminación de valores extremos.	18
Tabla 9: Distribución de fuga para los clientes del año móvil 2016.....	19
Tabla 10: Distribución de clientes fugados para el año móvil 2016.	19
Tabla 11: Distribución de variables RFM para clientes antiguos.....	24
Tabla 12: Distribución de variables RFM para clientes nuevos.....	25
Tabla 13: Análisis de RFM para clientes antiguos.	26
Tabla 14: Análisis de variables RFM para clientes nuevos.	27
Tabla 15: Métricas de desempeño para los modelos realizados.	34
Tabla 16: Subject de promociones realizadas	36
Tabla 17: Test de medias y varianzas para las promociones de monto de descuento. .	37
Tabla 18: Test de medias y varianzas para las promociones de montos con restricción de ticket.	38
Tabla 19: Test de medias y varianzas para las promociones de porcentajes de descuentos.....	38
Tabla 20: Resultados de las promociones enviadas.....	39
Tabla 21: Venta incremental por promoción enviada.....	40
Tabla 22: Costos y utilidad generada por cada una de las promociones.....	40
Tabla 23: Test de medias para clientes nuevos.....	45

Tabla 24: Test de medias para clientes antiguos.....	45
Tabla 25: Test de hipótesis para clientes nuevos que recibieron una promoción conservadora.....	46
Tabla 26: Test de hipótesis para clientes antiguos que recibieron una promoción conservadora.....	46
Tabla 27: Test de hipótesis para clientes nuevos que recibieron una promoción agresiva.....	47
Tabla 28: Test de hipótesis para clientes antiguos que recibieron una promoción agresiva.....	47
Tabla 29: Test de hipótesis para clientes nuevos que compraron hace menos de 6 meses.....	47
Tabla 30: Test de hipótesis para clientes nuevos que compraron hace 6 o más meses.....	47
Tabla 31: Test de hipótesis para clientes antiguos que compraron hace menos de 7 meses.....	48
Tabla 32: Test de hipótesis para clientes antiguos que compraron hace más de 7 meses.....	48
Tabla 33: Test de hipótesis para clientes nuevos retornados con propensión alta	49
Tabla 34: Test de hipótesis para clientes antiguos retornados con propensión alta.....	49
Tabla 35: Resumen de hipótesis y su resultado.....	50
Tabla 36: Asistencia incremental para clientes con tarjeta de crédito abierta	51
Tabla 37: Asistencia incremental para clientes con tarjeta de crédito cerrada.....	51
Tabla 38: Asistencia incremental para clientes sin tarjeta de crédito.....	52
Tabla 39: Asistencia incremental para clientes nuevos.....	53
Tabla 40: Asistencia incremental para clientes antiguos.....	53
Tabla 41: Asistencia incremental para clientes que no han comprado entre 1 a 3 meses.....	54
Tabla 42: Asistencia incremental para clientes que no han comprado entre 4 a 6 meses.....	55
Tabla 43: Asistencia incremental para clientes que no han comprado entre 7 a 9 meses.....	56

Tabla 44: Asistencia incremental para clientes que no han comprado entre 10 a 12 meses.	57
Tabla 45: Coeficientes para la regresión de clientes con tarjeta abierta.....	58
Tabla 46: Coeficientes para la regresión de clientes con tarjeta cerrada.	59
Tabla 47: Coeficientes para la regresión de clientes sin tarjeta de crédito del holding. .	59
Tabla 48: Coeficientes para la regresión de clientes nuevos.	60
Tabla 49: Coeficientes para la regresión de clientes antiguos.	61
Tabla 50: Coeficientes para la regresión de clientes con compras hace menos de tres meses.....	62
Tabla 51: Coeficientes para la regresión de clientes con compras entre cuatro a seis meses.....	62
Tabla 52: Coeficientes para la regresión de clientes con compras entre siete y nueve meses.....	63
Tabla 53: Coeficientes para la regresión de clientes con compras hace más de 10 meses.	63
Tabla 54: Propuesta de gestión para clientes nuevos.	64
Tabla 55: Propuesta de gestión para clientes antiguos.	65
Tabla 56: Propuesta de gestión para clientes con tarjeta de crédito abierta o activa.....	65
Tabla 57: Propuesta de gestión para clientes con tarjeta de crédito cerrada.	66
Tabla 58: Propuesta de gestión para clientes sin tarjeta de crédito del holding.	66
Tabla 59: Coeficientes para la regresión de clientes con compras hace menos de tres meses.....	66
Tabla 60: Coeficientes para la regresión de clientes con compras entre cuatro a seis meses.....	67
Tabla 61: Coeficientes para la regresión de clientes con compras entre siete y nueve meses.....	67
Tabla 62: Coeficientes para la regresión de clientes con compras hace más de 10 meses.	67
Tabla 63: Impacto económico (en unidades monetarias) en la activación de clientes antiguos con mayor gasto, según la propensión.	68

Tabla 64: Impacto económico (en unidades monetarias) en la activación de clientes nuevos con mayor gasto, según la propensión. 68

Tabla 65: Métricas de desempeño con distintos puntos de corte para el modelo de clientes antiguos..... 76

Tabla 66: Impacto económico (en unidades monetarias) en la activación de clientes antiguos, según la propensión. 81

Tabla 67: Impacto económico (en unidades monetarias) en la activación de clientes nuevos, según la propensión. 81

Tabla 68: Impacto económico (en unidades monetarias) en la activación de clientes antiguos con menor gasto, según la propensión..... 82

Tabla 69: Impacto económico (en unidades monetarias) en la activación de clientes nuevos con menor gasto, según la propensión..... 82

ÍNDICE DE GRÁFICOS

Gráfico 1: Participación de mercado en Retail en 2015, por tipo de empresa.....	1
Gráfico 2: Outliers para visitas y cantidad de productos.....	18
Gráfico 3: Distribución de clientes antiguos por género.....	19
Gráfico 4: Distribución de clientes nuevos por género.....	19
Gráfico 5: Distribución de clientes antiguos por zona de residencia y fuga.....	20
Gráfico 6: Distribución de clientes nuevos por zona de residencia y fuga.....	20
Gráfico 7: Distribución de clientes antiguos por rango etario y fuga.....	21
Gráfico 8: Distribución de clientes nuevos por rango etario y fuga.....	21
Gráfico 9: Distribución de clientes antiguos por visitas.....	21
Gráfico 10: Distribución de clientes nuevos por visitas.....	21
Gráfico 11: Visitas promedio por tipo de cliente y fuga.....	22
Gráfico 12: Distribución de clientes antiguos por tenencia de tarjeta y fuga.....	22
Gráfico 13: Distribución de clientes nuevos por tenencia de tarjeta y fuga.....	22
Gráfico 14: Distribución de clientes antiguos por compras anteriores y fuga.....	23
Gráfico 15: Distribución de clientes nuevos por compras anteriores y fuga.....	23
Gráfico 16: Distribución de clientes antiguos por tipo de tarjeta y fuga.....	24
Gráfico 17: Distribución de clientes nuevos por tipo de tarjeta y fuga.....	24
Gráfico 18: Evaluación de distintas métricas, en función de la probabilidad de corte para clientes antiguos.....	30
Gráfico 19: Distribución del tipo de tarjeta CMR con respecto a la probabilidad de fuga.	32
Gráfico 20: Distribución de la antigüedad del cliente con respecto a la probabilidad de fuga.....	32
Gráfico 21: Distribución del rango etario con respecto a la probabilidad de fuga.....	33
Gráfico 22: Distribución de género con respecto a la probabilidad de fuga.....	33
Gráfico 23: Distribución del tipo de cuenta con respecto a la probabilidad de fuga.....	34

Gráfico 24: Comparación de Lift tanto para el modelo de clientes antiguos como para el de nuevos.....	35
Gráfico 25: Distribución de clientes que canjearon, por rango etario y tipo de promoción.	41
Gráfico 26: Distribución de clientes que canjearon, por género y tipo de promoción.....	41
Gráfico 27: Distribución de clientes que canjearon, por zona de residencia y tipo de promoción.	42
Gráfico 28: Distribución de clientes que canjearon, por tipo de cliente y tipo de promoción.	42
Gráfico 29: Distribución de clientes que canjearon, por estado de tarjeta de crédito y tipo de promoción.	43
Gráfico 30: Distribución de clientes que canjearon, por nivel de visitas y tipo de promoción.	44
Gráfico 31: Distribución de canjes por nivel de aperturas.	44
Gráfico 32: Rama del árbol que presenta la mayor cantidad de clientes antiguos fugados.	74
Gráfico 33: Rama del árbol de decisión que presenta mayor cantidad de clientes nuevos fugados.	75

ÍNDICE DE ILUSTRACIONES

Ilustración 1: Funnel con apertura de clientes potenciales.	3
Ilustración 2: Partición realizada para la generación del modelo.	28
Ilustración 3: Variables más importantes para el modelo de clientes antiguos.....	29
Ilustración 4: Variables más importantes para el modelo de clientes nuevos.....	31
Ilustración 5: Configuración de una promoción.	77
Ilustración 6: Creación de cupones seriados.....	77
Ilustración 7: Resumen de promociones enviadas a los clientes.	78
Ilustración 8: Ejemplo de la gráfica enviada	80

1. INTRODUCCIÓN

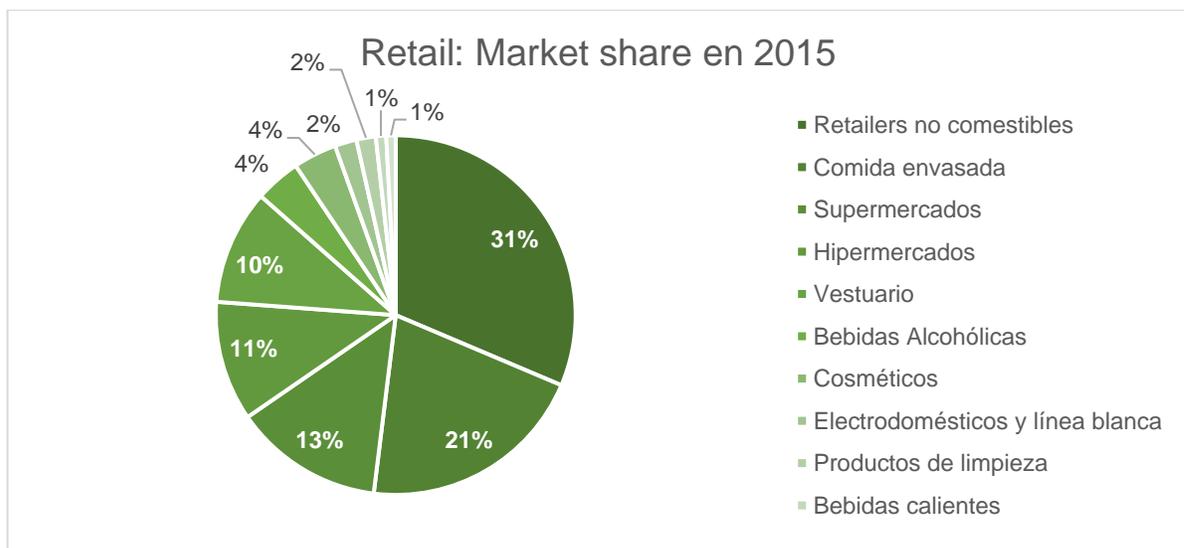
1.1. ANTECEDENTES GENERALES DE LA INDUSTRIA

La industria del retail ha experimentado grandes cambios a lo largo del tiempo, lo que se refleja en las fluctuaciones que ha presentado su crecimiento. En esta línea, se evidencia que, entre 2010 y 2013, las ventas minoristas crecieron, en promedio, a razón de dos dígitos. No obstante, a partir de 2014, la industria ingresó a una etapa de desaceleración, conocida como el súper ciclo del consumo, que disminuyó el crecimiento de las ventas a menos del 3%. (Cámara de Comercio de Santiago, 2016).

Adicionalmente, se observa que el crecimiento de la industria en el año 2016 fue de 2.5%. Sin embargo, la Cámara de Comercio de Santiago (CCS) señala que el comercio minorista crecerá a una tasa de 3% durante el 2017, debido al “pronunciado crecimiento de las ventas de bienes durables tanto de automóviles como de productos electrónicos y tecnológicos”. (America Retail, 2017).

Por otro lado, el PIB de Chile en 2016 fue de US\$247 mil millones, donde el sector del comercio minorista alcanzó un 20.4% de dicho indicador, equivalente a US\$50 mil millones. (Economía y Negocios, 2017).

Gráfico 1: Participación de mercado en Retail en 2015, por tipo de empresa.



Fuente: Estimaciones CCS sobre antecedentes de Euromonitor Internacional.

Del Gráfico 1, se evidencia que existen distintos y diversos actores presentes en la industria, lo que genera una alta competitividad y dinamismo en el sector económico. Del mismo modo, se observa que *retailers* no comestibles, comida envasada y supermercados, son quienes lideran la industria con el 65% de participación de mercado.

De esta manera, se hace más evidente que las empresas puedan tomar acciones proactivas y no reactivas que permitan adelantarse a las necesidades y preferencias de los consumidores, mediante la utilización de diversos modelos que entreguen *insights* relevantes para enfocar las estrategias de marketing que una compañía realiza. Un claro ejemplo de lo anterior, responde a que una organización busca generar una relación de

largo plazo con sus clientes, pues el éxito de ellas no sólo se centra en captar nuevos consumidores, sino que también, en retener y satisfacer a los actuales. Por esta razón, el diseño de una estrategia de retención ad-hoc al público objetivo de la empresa, mediante el uso de técnicas de *data mining*, se torna cada vez más relevante para mantener una posición líder en la industria.

El estudio se realizará en la industria del retail, enfocado en el sector que abarca a las tiendas por departamento, donde los clientes poseen una relación no contractual con la compañía, es decir, no se puede saber, con certeza, cuando un consumidor dejará de relacionarse con la empresa.

1.2. ANTECEDENTES GENERALES DE LA EMPRESA

La organización cuenta con operaciones en distintas líneas de negocio: tiendas por departamento; mejoramiento del hogar; supermercados; inmobiliario y, servicios financieros distribuidos en Chile, Argentina, Perú, Colombia, Uruguay, Brasil y próximamente México con mejoramiento del hogar y servicios financieros.

A diciembre de 2016¹, la compañía contaba con 107.361 colaboradores, MM\$ 609.024 de utilidad neta, 476 tiendas, equivalentes a una superficie de 2.838.000 m² y 39 centros comerciales, equivalentes a una superficie arrendable de 1.943.000 m².

El trabajo de título se realizará en la línea de negocio correspondiente a tiendas por departamento que tiene presencia en Chile, Argentina, Perú y Colombia.

Cabe destacar que, la línea de negocios más importante está constituida por las tiendas por departamento y la creciente importancia que ha alcanzado la omnicanalidad, creando una sinergia entre los distintos canales donde comercializan sus productos. Dicho esto, la compañía ha enfocado gran parte de sus inversiones en potenciar el canal online y los servicios asociados que permiten conectar la tienda virtual con la física. De lo anterior, han surgido propuestas como quioscos en las tiendas físicas que permiten al vendedor asegurar la venta, en caso que no exista stock disponible en la tienda; un servicio que permite al cliente comprar en la tienda virtual y retirar en la tienda más cercana, sin costo alguno, entre otros.

La organización clasifica a sus clientes mediante la cantidad de visitas que realizan a la tienda y el monto de las compras efectuadas en ella en dos segmentos: clientes de alto valor y clientes normales. De esta manera, se evidencia que los clientes más valiosos para la organización son aquellos que tienen transacciones sobre el tercil de compras, acumulando operaciones que superen el quintil de venta durante el año calendario, mientras que los clientes normales son aquellos que no cumplen esta categoría. En cuanto a la distribución de clientes, se sabe que, por cada cliente frecuente, existen nueve clientes normales.

Adicionalmente, se evidencia que los clientes normales se dividen en dos tipos de clientes: antiguos, quiénes han comprado en dos años consecutivos y, nuevos, quiénes

¹ Fuente: Memoria Anual 2016 de la empresa.

no compraron en el año anterior. Se evidencia que la mayoría de sus clientes son antiguos.

2. DESCRIPCIÓN DEL PROYECTO

2.1. JUSTIFICACIÓN DEL ESTUDIO

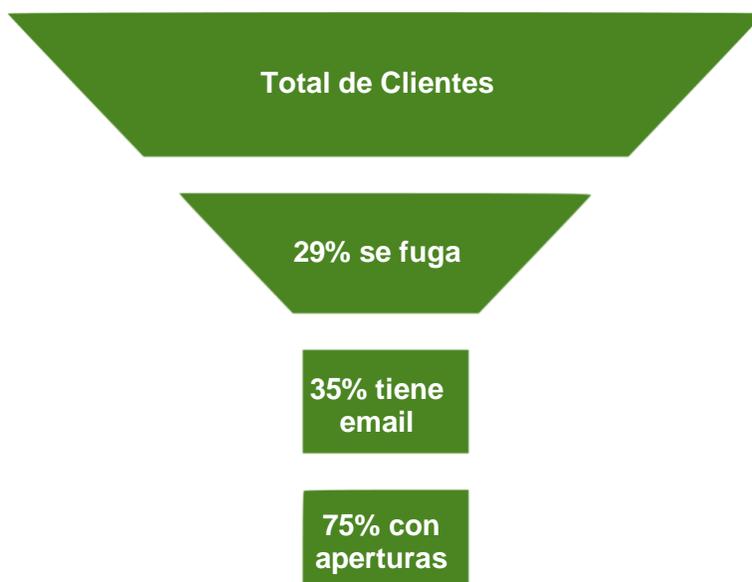
A lo largo de los años, la adquisición de nuevos clientes se ha encarecido debido a los altos costos que están asociados a los programas de captación y fidelización. Por esta razón, mantener a un cliente que presenta una relación tanto contractual como no contractual con una compañía, se vuelve cada vez más importante.

De acuerdo a lo anterior, si un cliente se encuentra satisfecho con el servicio que una organización brinda, es probable que continúe relacionándose con ella, lo que conlleva a un aumento considerable de su valor para la empresa, independiente del tipo de relación que mantengan (contractual y no contractual). Más aún, la retención es uno de los pilares fundamentales para aumentar el valor de la cartera de clientes que una empresa posee.

Por lo pronto, la organización cuenta con una definición comercial, la cual señala que si un cliente deja de comprar después de 12 meses consecutivos, se considerará como fugado. Bajo este punto de vista, se evidencia que el porcentaje de clientes fugados anualmente asciende, aproximadamente, al 29% del total de clientes de la tienda².

Si se estudian aquellos clientes que se fugan, en términos de contactabilidad, se evidencia que sólo se puede enviar una campaña por correo electrónico al 35% de ellos, donde el 75% tiene, al menos, una apertura. Cabe mencionar que, de los clientes fugados, el 26% no presenta datos demográficos (género, edad ni lugar de residencia).

Ilustración 1: Funnel con apertura de clientes potenciales.



Fuente: Elaboración propia.

² La proporción de clientes fugados se calculó con los datos transaccionales del año móvil 2015 y 2016.

Lo anterior, se traduce en que la fuga de consumidores está compuesta por dos componentes; una fija, correspondiente al 18% que son los clientes sin correo electrónico registrado (65% de los clientes fugados) y, una variable, que equivale al 9% del total de clientes (35% de los clientes fugados) con los que se pueden generar acciones de email marketing enfocadas en la retención.

De esta manera y, bajo el supuesto de que todos los clientes con correo electrónico volverán a comprar en el año siguiente, manteniendo un comportamiento similar, se evidencia que las ventas de la organización crecerían, en promedio, en un 4%. Sin embargo, para concretizar el impacto que se puede lograr al estudiar en esta área, se realiza un análisis de sensibilidad, de acuerdo al porcentaje de clientes que se lograría retener. En esta línea, se obtiene que, si se retiene al 1% de los clientes se obtendría una utilidad que bordea los 12.64 [u.m]³, mientras que, si se retiene al 10% de ellos, se obtendría una utilidad que asciende a los 126 [u.m.].

Adicionalmente, destaca que, si bien son clientes de bajo valor o de bajo relacionamiento con la tienda por departamento, la cantidad de consumidores incluidos en esta categoría, lo hace ser un segmento relevante de gestionar y retener.

Con el objetivo de evitar que los clientes dejen de comprar en un año, la tienda por departamento cuenta con una política de retención que se activa para clientes que han dejado de comprar en los últimos ocho meses, en categorías blandas. En particular, dicha estrategia consiste en tres campañas sucesivas que se envían durante tres meses consecutivos, donde cada una de ellas posee una duración promedio de 14 días desde que se emite la promoción. Cabe destacar que, el plan se interrumpe, una vez que el cliente realiza una transacción en la tienda.

Con lo anterior, se evidencia que no existe distinción entre los clientes, es decir, todos reciben las mismas acciones promocionales, independiente si son clientes antiguos que han comprado continuamente o si son clientes nuevos que sólo poseen una compra en la tienda.

Es importante señalar que el estudio se concentra sólo en los clientes normales y no en los de alto valor de la compañía, porque estos últimos poseen una tasa de fuga que no supera el 1% al año, lo que se debe, principalmente, al cierre de la tarjeta de crédito del holding⁴.

De este modo, se busca realizar un modelo de fuga tanto para clientes antiguos y nuevos, en conjunto con una propuesta de gestión de clientes que permita responder a ¿quiénes son los consumidores que se fugan?, ¿cuáles son las principales variables que explican cuando se fuga? y, ¿cuáles son las acciones promocionales más apropiadas de realizar para cada uno de ellos?

³ Se define u.m. como unidades monetarias.

⁴ Fuente: Información proporcionada por el área Inteligencia de Clientes de la empresa.

2.2. OBJETIVOS

2.2.1. OBJETIVO GENERAL

Desarrollar un modelo de propensión de fuga que, mediante el uso eficiente de acciones promocionales, permita disminuir el abandono de un segmento de clientes.

2.2.2. OBJETIVOS ESPECÍFICOS

1. Definir las variables relevantes que explican la fuga de los clientes normales.
2. Desarrollar un modelo que determine la probabilidad de fuga.
3. Definir y evaluar el efecto de promociones en la probabilidad de fuga de clientes.
4. Sugerir una política de retención para los clientes, de acuerdo a los aspectos relevantes para el negocio (tipo de cliente, intensidad de visitas, entre otros).

2.3. RESULTADOS ESPERADOS

Generación de un modelo que permita predecir qué tan propenso es un cliente a la fuga, dado ciertas variables independientes. Lo anterior, se traduce en obtener un modelo de propensión que permita predecir cuándo un cliente dejará de comprar, permitiendo focalizar la comunicación en tratar de que dichos clientes concreten una compra en la tienda.

Una vez realizado el modelo, se realizarán experimentos que permitan evaluar qué clientes responden positivamente a ciertos estímulos realizados por la organización. En particular, se espera probar distintas promociones y, caracterizar a los clientes que utilizaron el incentivo brindado. De esta manera, se sugerirá una estrategia de marketing directo asociado a las acciones promocionales que pueden ser utilizadas por cada tipo de cliente, basado en rangos de probabilidad de fuga.

2.4. ALCANCES

Dado que el trabajo de título considera la utilización de una gran cantidad de registros transaccionales, se trabajará sólo con los clientes identificables, es decir, aquellos que cuentan con, al menos, una transacción en la organización durante el período de estudio.

Adicionalmente, como se menciona anteriormente, se limitará el estudio a aquellos clientes normales de la tienda por departamento porque son los que concentran la fuga de la compañía.

El canal a utilizar para el envío de las promociones será el correo electrónico. Dada la naturaleza del problema, no se incorpora la utilización de cupones POS, ni de SMS debido a su alto costo.

Con respecto a las campañas, se utilizarán variaciones tanto en montos de regalos como en porcentajes de descuento en las categorías de vestuario y calzado hombre, mujer y niño, las que serán válidas para compras en la tienda física y tienda online.

Finalmente, debido a que se trabajará con clientes de bajo relacionamiento con la tienda por departamento (baja frecuencia de compras), no se podrá evaluar si los efectos de la gestión de clientes basada en email marketing son consecuentes a largo plazo, es decir,

si un cliente que se activó por la campaña, continúa comprando y, por consiguiente, su probabilidad de fuga disminuye en períodos siguientes.

3. MARCO CONCEPTUAL

3.1. CRITERIO DE FUGA

3.1.1. CRITERIO COMERCIAL

La organización reconoce la importancia de retener y fidelizar a los clientes. De este modo, han definido un criterio comercial para determinar cuándo un cliente está propenso a fugarse y cuándo se considera como fugado. Dichas reglas son válidas para los clientes que realizan compras en categorías blandas de la tienda por departamento. Dichas reglas son válidas para los clientes que realizan compras en categorías blandas de la tienda por departamento.

En particular, la caracterización es idéntica para todo tipo de cliente y se basa, únicamente, en el tiempo que transcurre desde la última compra, es decir, de acuerdo al *recency* de cada cliente. Dicho esto, se observa que, si un cliente no ha realizado una transacción en los últimos ocho meses, se activa una campaña de retención con el objetivo de incentivar una compra en la tienda y, de esta manera, se busca retener al consumidor. Más aún, si el comportamiento persiste y no concreta una interacción con la compañía en un año, se considerará como un cliente fugado.

En el siguiente punto, se muestra una forma alternativa de caracterizar el fenómeno en estudio, lo que permite definir reglas que puedan comprender cuándo un cliente se puede clasificar como un fugado.

3.1.2. CRITERIO RFM

El criterio RFM se suele utilizar como una métrica relevante para determinar cuánto tiempo debe transcurrir para considerar a un cliente como fugado. A continuación, se describen las variables asociadas al criterio.

1. Recency (R): corresponde al tiempo transcurrido desde la última transacción realizada en la tienda.
2. Frequency (F): corresponde al tiempo promedio que transcurre entre transacciones.
3. Monetary value (M): es el ticket promedio de las transacciones realizadas en un período de tiempo.

Si bien puede ser una forma trivial de caracterizar a los consumidores, este criterio proporciona información relevante para el negocio. Mediante la interacción entre las tres variables permite generar puntuaciones o score para identificar cuáles son los segmentos de clientes de alto valor para la organización, donde aquellos con mayor score (cercaos a 555) corresponden a los clientes leales, mientras que quienes poseen un menor score (cercaos a 111), corresponden a clientes esporádicos que no generan grandes utilidades para la compañía.

3.2. MODELO DE FUGA

Las compañías que presentan relaciones no contractuales con sus clientes, a diferencia de aquellas contractuales, no pueden determinar irrefutablemente cuando el consumidor deja de estar activo o se fuga. Por esta razón, se han desarrollado distintos modelos que permiten pronosticar o predecir el comportamiento futuro de ellos, los cuales se basan, principalmente, en el historial de compra de cada cliente. A continuación, se describirán el modelo que se utilizará para el desarrollo del trabajo de título.

3.2.1. ÁRBOLES DE DECISIÓN

Los árboles de decisión buscan explicar una variable dependiente que puede ser de dos tipos: categórica (árboles de clasificación) o continua (árboles de regresión). Se componen por un nodo inicial, nodos no terminales y terminales que dan cuenta el comienzo del árbol, la variable con la que se realiza la partición de la base y la clase en la que se clasifica, respectivamente y ramas que son los posibles valores que puede tomar una variable.

Por lo pronto, el desarrollo de un árbol corresponde a una partición consecutiva de los datos, mediante la ramificación de cada nodo utilizando la variable independiente con mayor relación a la dependiente y un punto de corte óptimo. El proceso se realiza de forma sucesiva hasta cumplir con el criterio de detención definido. Cabe mencionar que, en cada nodo no terminal, se efectúa un test que permite entregar una decisión, representada por el nodo terminal.

3.2.1.1. CRITERIOS DE DIVISIÓN

En la literatura, existen distintos criterios que permiten crear particiones de los datos. A continuación, se presentarán aquellos más importantes, identificados en (Rokach & Maimon, 2005).

3.2.1.1.1. GANANCIA DE INFORMACIÓN

La ganancia de información es un criterio que se basa en la medida de impureza y utiliza la medida de entropía para su representación. A continuación, se evidencia la fórmula correspondiente.

$$\text{Information Gain}(a_i, S) = \text{Entropy}(y, S) - \sum_{v_{i,j} \in \text{dom}(a_i)} \frac{|\sigma_{a_i=v_{i,j}} S|}{|S|} \cdot \text{Entropy}(y, \sigma_{a_i=v_{i,j}} S)$$

Donde la entropía se define como una medida de variabilidad o desorden en los datos:

$$\text{Entropy}(y, S) = \sum_{c_j \in \text{dom}(y)} \frac{|\sigma_{y=c_j} S|}{|S|} \cdot \log_2 \frac{|\sigma_{y=c_j} S|}{|S|}$$

3.2.1.1.2. ÍNDICE DE GINI

Es un criterio basado en la impureza que mide la divergencia entre la distribución de probabilidad de los valores de los atributos objetivos. A continuación, se expresa la fórmula correspondiente.

$$Gini(y, S) = 1 - \sum_{c_j \in dom(y)} \left(\frac{|\sigma_{y=c_j} S|}{|S|} \right)^2$$

De lo anterior, se desprende que, si se selecciona el atributo a_i , la fórmula de la ganancia del índice de Gini, se expresa a continuación.

$$Gini\ Gain(a_i, S) = Gini(y, S) - \sum_{v_{i,j} \in dom(a_i)} \frac{|\sigma_{a_i=v_{i,j}} S|}{|S|} \cdot Gini(y, \sigma_{a_i=v_{i,j}} S)$$

3.2.1.1.3. RATIO DE MÁXIMA VEROSIMILITUD

Se utiliza para medir la significancia estadística del criterio de ganancia de información. Se constituye por un test de hipótesis, donde la nula (H_0) corresponde a que el atributo ingresado y el objetivo son condicionalmente independientes. Si se acepta H_0 , el test se distribuye como χ^2 con grados de libertad equivalentes a $(dom(a_i) - 1) \cdot (dom(y) - 1)$.

$$G^2(a_i, S) = 2 \cdot \ln(2) \cdot |S| \cdot Information\ Gain(a_i, S)$$

3.2.1.1.4. RATIO DE GANANCIA

Es una medida que normaliza la ganancia de información. Su determinación se desarrolla en dos etapas: se calcula la ganancia de información para todos los atributos y, de aquellos en que lo calculado es mayor que el promedio de la ganancia de información, se calcula el ratio de ganancia. De esta manera, destaca que aquel atributo con mejor ratio de ganancia, es el que se selecciona.

$$Gain\ Ratio(a_i, S) = \frac{Information\ Gain(a_i, S)}{Entropy(a_i, S)}$$

3.2.1.2. CRITERIOS DE PARADA

Como se ha mencionado, el árbol continúa su crecimiento hasta que se cumplen los criterios de detención definidos. En este sentido, para que un árbol complete la ramificación, obteniendo un nodo terminal, se identifican ciertas condiciones que permiten su interrupción.

- La profundidad máxima del árbol se ha alcanzado.
- El número de casos en el nodo terminal es menor que el número de casos mínimo para los nodos padres.
- Si un nodo se divide, el número de casos en uno o más nodos hijos será menor que el número de casos mínimo para los nodos hijos.

3.2.1.3. ALGORITMOS DE ÁRBOLES DE DECISIÓN

Para que el árbol pueda agregar nodos y ramificaciones, se requieren de técnicas enfocadas en el crecimiento del árbol. En particular, se evidencian cuatro algoritmos que permiten realizar la creación del árbol con sus nodos y ramificaciones respectivas. A continuación, basado en (IBM, 2017) se explicita cada una de ellas.

- CHAID (*CHI-square Automatic Interaction Detection*): un algoritmo de árbol estadístico que explora los datos de forma rápida y eficiente, generando segmentos y perfiles en relación con el resultado deseado. Para generar los nodos y ramificaciones, CHAID escoge la variable independiente que tiene una mayor relación con la dependiente.
- CHAID exhaustivo: una modificación de CHAID, que examina todas las posibles particiones para cada predictor.
- CRT (*Classification and Regression Trees*): un completo algoritmo de árbol binario, que particiona datos y genera subconjuntos homogéneos. CRT proporciona segmentos homogéneos con respecto a la variable dependiente. Un nodo terminal, en el que todos los casos toman el mismo valor en la variable dependiente, es un nodo homogéneo.
- QUEST (*Quick, Unbiased, Efficient Statistical Tree*): un algoritmo estadístico que selecciona variables sin parcialidades y genera árboles binarios precisos de forma rápida y eficaz. El método se puede utilizar si la variable dependiente es nominal.

Tabla 1: Características del método de crecimiento

Feature	CHAID*	CRT	QUEST
Basado en chi-cuadrado**	X		
Variables (predictoras) independientes sustitutas		X	X
Poda de árboles		X	X
División de nodos multinivel	X		
División de nodos binarios		X	X
Variables de influencia	X	X	
Probabilidades previas		X	X
Costes de clasificación errónea	X	X	X
Cálculo rápido	X		X

* Incluye CHAID exhaustivo.

** QUEST también utiliza una medida de chi-cuadrado para variables independientes nominales.

Fuente: Creación de árboles de decisión, IBM Knowledge Center (IBM, 2017).

3.3. DISTANCIA DE MAHALANOBIS

La distancia de Mahalanobis, a diferencia de la euclidiana, toma en consideración la covarianza existente entre variables. De esta forma, es posible conseguir un resultado más completo que con un análisis univariado, pues determina la distancia de cada punto, respecto a la media de los datos y al mismo tiempo toma en cuenta la correlación entre variables (Rosenmai , 2013) (Steffen, 2016).

En términos matemáticos, la distancia de Mahalanobis queda formulada por la siguiente expresión: $D(\vec{x}, \vec{y}) = \sqrt{(\vec{x} - \vec{y})^T S^{-1} (\vec{x} - \vec{y})}$. Donde \vec{x} e \vec{y} son vectores aleatorios de una misma distribución y S la matriz de covarianza. No obstante, para su uso como método de detección de *outliers* se asume la normalidad de las características o atributos en

estudio. Así, el cálculo queda definido como la distancia de los datos (multivariada) respecto a su media (Orlov, s.f.), tal que:

$$D = \sqrt{(R - \mu_1)S_1^{-1}(R - \mu_1)}.$$

Con R la realización de una de las variables aleatorias y μ_1 y S_1 la media y matriz de covarianzas del otro atributo en comparación. Finalmente, una vez determinadas las distancias respecto al centro, se establece un corte con el cual se terminan por identificar los registros que serán considerados *outliers*. (Eidgenössische Technische Hochschule Zürich, 2012).

3.4. DISEÑO DE EXPERIMENTOS

3.4.1. PROMOCIONES

Una promoción es un estímulo que se le entrega a los clientes de una compañía para incentivar la adquisición de ciertos productos, definidos por la promoción enviada.

3.4.1.1. GRUPO CONTROL

Corresponde al grupo de personas que tiene las mismas características que el grupo tratamiento, pero no están expuestas a la acción promocional que se desea evaluar. El objetivo de crear este grupo es utilizarlo como una comparación con el grupo de tratamiento para medir el impacto del efecto estudiado.

3.4.1.2. GRUPO TRATAMIENTO

Corresponde al grupo de personas que recibe la acción promocional que se desea evaluar. Se utiliza para estudiar el cambio en la respuesta con respecto al grupo control.

3.4.2. VALIDEZ DE EXPERIMENTOS

Para que un experimento sea válido y pueda explicar el efecto observado en la población, debe tener validez tanto interna como externa.

3.4.2.1. VALIDEZ INTERNA

La validez interna consiste en que el experimento sea capaz de mostrar relaciones causales representativas de la interacción del factor que se está estudiando y que no estén condicionadas a otros externos al experimento.

3.4.2.2. VALIDEZ EXTERNA

La validez externa consiste en si el experimento realizado y los resultados obtenidos de este, pueden ser aplicados a situaciones reales con un contexto distinto al realizado. Un ejemplo de este tipo de validez corresponde a los experimentos de campo que tienen una alta validez externa, pero no tienen validez interna.

Para efectos del trabajo de título, las promociones que se entregarán variarán en montos de descuento y porcentaje de descuento en las líneas que generan mayor volumen del retail, vestuario y calzado hombre, mujer y niño.

3.4.3. MUESTREO

3.4.3.1. PROPORCIÓN

El tamaño de la muestra se puede determinar construyendo un intervalo de confianza alrededor de la media. Sabemos que:

$$error = K \cdot \hat{S}_{\bar{Y}} = K \cdot \sqrt{\frac{N-n}{N-1} \cdot \frac{p \cdot q}{n}} \quad \text{donde } q = 1 - p$$

Despejando, se obtiene que, el tamaño de la muestra es:

$$n = \frac{K^2 pqN}{e^2(N-1) + K^2 pq}$$

donde: N: tamaño de la población.

K: nivel de confianza.

s: variabilidad de los datos.

De esta manera, se evidencia que el tamaño de la muestra depende de tres factores:

1. Nivel de confianza (K).
2. El error estadístico que se está dispuesto a tolerar.
3. Variabilidad de los datos.

Dado que la variabilidad de los datos (s^2) suele ser desconocida, se puede asumir que los casos tienen variabilidad máxima, lo que ocurre cuando $p = q = 0.5$.

No obstante, se reconoce que para tamaños de población muy grandes (N), se puede asumir población finita, con lo que el tamaño de la muestra, se calcula como sigue:

$$n_{\infty} = \frac{K^2 pq}{e^2}$$

4. METODOLOGÍA

Para el desarrollo del trabajo de título, se utilizará la metodología KDD (*Knowledge Discovery in Databases*) que consiste en selección de datos, pre-procesamiento, transformación, *data mining* e interpretación y evaluación de resultados. A continuación, se presenta una descripción de cada uno de los aspectos que abarca.

4.1. SELECCIÓN DE DATOS

La organización cuenta con una base de datos que consolida todas las transacciones realizadas en su negocio, incluyendo tanto tiendas físicas como la tienda online. En particular, se trabajará con dicha base de datos, utilizando la información registrada para el período que comprende entre los años móviles 2015 y 2017.

Además, la empresa registra, para cada uno de los clientes, los datos socio-demográficos y de navegación que han realizado en la tienda online. Dicho esto, se incorporará la información en la base a utilizar para el modelo.

Finalmente, se incorporará la información que la compañía consolida respecto al estado actual de la tarjeta de crédito asociada al negocio y los reclamos que se han realizado tanto en la tienda como en la división financiera.

4.2. PRE-PROCESAMIENTO

Con las bases de datos a utilizar seleccionadas y consolidadas, se procederá a limpiar y realizar un procesamiento de la información con el objetivo de reducir posibles errores que podrían generarse al no realizar un correcto tratamiento de los datos.

En la misma línea, se reconoce que la base de datos puede contar con registros que se alejan de la distribución de los datos para ciertas variables. Por lo anterior, se decide excluir a aquellos clientes que presenten valores atípicos o fuera de lo común. Se estudiarán las características más importantes, entre las cuales se encuentran: venta, visitas y cantidad de productos comprados.

Por la definición del estudio, se evidencia la existencia de un problema de *missing values* para las variables sociodemográficas de los clientes que recién comienzan una relación no contractual con la compañía. En particular, para reducir el error que podría ocasionar no tratar dichas variables, se propone definir variables categóricas que incluyan un nivel asociado a la falta de información permitiendo caracterizar al cliente, basado en dichas variables.

4.3. TRANSFORMACIÓN

El proceso de transformación de los datos, permite construir variables a partir de otras presentes en la base de datos. Lo anterior, se torna importante, pues permite incluir variables como *input* del modelo de fuga, que no son registradas como tal, en la base de datos transaccional. A continuación, se describen algunas de las variables que fueron generadas.

Se construirá una variable denominada visitas que refleja la cantidad de veces que un cliente asistió y compró en la tienda por departamento. En particular, la variable se define al agrupar el identificador del cliente, la tienda en la que realizó la compra y el día cuando lo realizó. Adicionalmente, se calcula la cantidad de visitas a vestuario, deco y electro.

Además, se construyen variables categóricas que representan rangos de edad, género, zona y qué tipo de datos sociodemográficos son faltantes para cada cliente.

Por otro lado, se calcularán variables binarias que indican en qué meses realiza la compra junto con una variable numérica que indica la cantidad de meses que compra cada cliente.

Finalmente, se construyen variables asociadas al tipo de tarjeta de crédito que posee el cliente, el estado de la cuenta (abierta, cerrada, otras) y la cantidad de años que ha comprado anteriormente el cliente, es decir, si ha efectuado compras en años anteriores, distintos de los datos utilizados para la evaluación.

4.4. ANÁLISIS EXPLORATORIO

Se realizará un análisis exploratorio que permita caracterizar el fenómeno en estudio para el caso particular de la compañía. Así como también, comprender qué tipo de variables son más relevantes de incluir en el modelo para explicar la fuga de clientes.

4.5. ÁRBOLES DE DECISIÓN

Con la base de datos construida y las variables independientes definidas, se procede a realizar un árbol de decisión para predecir la probabilidad de fuga de los clientes normales, tanto para antiguos como nuevos. Se analizará la capacidad predictiva de cada modelo generado, a través del cálculo de *accuracy*, *precision*, *recall* y *f-score*.

Adicionalmente, se comparará el desempeño del modelo realizado con un aleatorio basado en la probabilidad de fuga actual de la tienda por departamentos.

4.6. DISEÑO EXPERIMENTAL

4.6.1. OBJETIVO

Se busca evaluar la efectividad de distintas promociones realizadas a los clientes normales que se encuentran dentro del conjunto con mayor probabilidad de fuga y, caracterizar quiénes son los que responden ante la exposición a los incentivos, con el fin de determinar distintas formas de gestionar a los clientes.

4.6.2. HIPÓTESIS

Para el desarrollo del experimento, se reconocen ciertas hipótesis que surgen tanto del análisis descriptivo y de los principales resultados de los modelos generados.

Sin embargo, es necesario mencionar que las promociones conservadoras son aquellas que otorgan bajos descuentos (\$5.000, 20%, \$5.000 de dcto. por compras sobre \$15.000), mientras que las promociones no conservadoras son las que otorgan un alto descuento (\$10.000, 30%, \$10.000 de dcto. por compras sobre \$30.000).

Por medio del desarrollo de la investigación aplicada, se busca probar las hipótesis que se detallan a continuación.

- H1. Promociones tienen un impacto positivo en el nivel de visitas de clientes antiguos y nuevos, independiente de la pensión.
- H2. Promociones conservadoras tienen mejor respuesta en clientes antiguos que en clientes nuevos.

- H3. Promociones agresivas tienen mayor impacto que una conservadora tanto en clientes antiguos como nuevos.
- H4. Promociones tienen mejor respuesta en clientes nuevos con propensión alta que no han comprado hace 6 o más meses.
- H5. Promociones tienen mejor respuesta en clientes antiguos con una propensión alta que no han comprado hace 7 o más meses.
- H6. Promociones tienen un impacto significativo en el nivel de visitas en clientes nuevos retornados con propensión alta.
- H7. Promociones tienen un impacto significativo en el nivel de visitas en clientes antiguos retornados con propensión alta.

4.6.3. VARIABLE EXPERIMENTAL

Se evaluará sólo una variable experimental que corresponde al incentivo que recibe el cliente. Se evidencia que dicha variable se desglosa en seis niveles o condiciones experimentales distintas.

4.6.4. CONDICIONES EXPERIMENTALES

A continuación, se dan a conocer las condiciones experimentales que se buscan evaluar, a través del correo electrónico como canal de envío.

Cabe destacar que, en secciones posteriores, se detalla cómo se determina la cantidad de clientes que recibirán cada promoción y quiénes conformarán el grupo control.

Tabla 2: Esquema de condiciones experimentales

1. \$5.000 de descuento por compras sobre \$15.000.	Grupo Control
2. \$10.000 de descuento por compras sobre \$30.000.	
3. 20% de descuento.	
4. 30% de descuento.	
5. \$5.000 de descuento.	
6. \$10.000 de descuento.	

Fuente: Elaboración propia.

4.7. PROPUESTA DE GESTIÓN

Una vez que el tiempo de las campañas realizadas haya expirado, se procederá a evaluar los resultados obtenidos. En particular, se caracterizarán a los clientes que respondieron positivamente a los estímulos generados, según la propensión de fuga que indicó el modelo, perfil demográfico y de compra, entre otros factores.

Finalmente, se propondrá una estrategia de marketing directo enfocada en la retención de los clientes normales.

5. DESARROLLO METODOLÓGICO

A continuación, se explicita la metodología *KDD* llevada a cabo para la resolución del problema planteado precedentemente.

5.1. SELECCIÓN DE DATOS

Se consolida la información presente en el negocio correspondiente al año móvil de 2015, 2016 y 2017. En particular, se construyen dos bases de datos: una para el periodo (1) 2015 – 2016 y, otra para (2) 2016 – 2017. Con la primera, se busca generar el modelo de propensión y, con la segunda, se utiliza el modelo creado anteriormente con el objetivo de determinar la probabilidad de fuga de los clientes en el próximo año.

Por otro lado, se cuenta con variables preliminares o básicas recopiladas en cada transacción realizada por los clientes en las distintas tiendas de la empresa. En la Tabla 3 se detallan las que fueron utilizadas.

Tabla 3: Variables disponibles.

Variable	Tipo	Descripción
ID_CLTE	Numérica	Identificador del cliente.
SEXO_DEF	Nominal	Indica el sexo del cliente (Femenino, Masculino).
EDAD_DEF	Numérica	Indica la edad, en años, del cliente.
COMUNA_DEF	Nominal	Indica la comuna donde reside el cliente.
REGION_DEF	Nominal	Indica la región donde reside el cliente.
ZONA_RESIDENCIA	Nominal	Zona donde vive (Norte, Sur, Oriente, Poniente).
VISITAS_AGNO	Numérica	Visitas acumuladas en el año móvil.
VENTA_TOTAL_AGNO	Numérica	Venta acumulada en el año móvil.
IND_CLTE_WEB_AGNO	Binaria	1 si el cliente compró en la tienda online en el año móvil.
IND_CLTE_TIENDA_AGNO	Binaria	1 si el cliente compró en la tienda física en el año móvil.
IND_CLTE_OMNISCANAL_AGNO	Binaria	1 si el cliente compra en tienda online y física en el año móvil
CTA_HOLDING	Nominal	Indica el tipo de relacionamiento (Abierta, Cerrada, Sin tarjeta)..

Fuente: Elaboración propia.

Cabe mencionar que, si no existe información de alguna de las variables mencionadas previamente, se obtendrá un registro 0, para el caso de las variables numéricas, mientras que, para las variables nominales, se registra como *Sin información*.

Por lo pronto, un cliente nuevo se define como aquel que realiza una compra en el mes actual, pero no registra compras en los últimos doce meses móviles. Mientras que, el cliente antiguo es quién posee, al menos, una compra durante el período actual y, además, posee una o más compras en los últimos doce meses móviles.

Como se menciona en la justificación del problema, se excluyeron a los clientes frecuentes o de alto valor de la tienda porque su fuga no alcanza el 1% de los clientes. Más aún, su inclusión distorsionaría los resultados a obtener en etapas posteriores de la metodología, porque presentan características diametralmente opuestas con respecto a los clientes normales de la organización.

Tabla 4: Variables calculadas

Variables	Tipo	Descripción
IND_CLTE_SOLO_WEB_AGNO	Binaria	1 si el cliente sólo compra en la tienda online.
IND_CLTE_SOLO_TIENDA_AGNO	Binaria	1 si el cliente sólo compra en la tienda física.
IND_VESTUARIO_AGNO	Binaria	1 si el cliente compra en vestuario.
IND_ELECTRO_AGNO	Binaria	1 si el cliente compra en electro.
IND_DECO_AGNO	Binaria	1 si el cliente compra en deco.
IND_NAVEG_AGNO	Binaria	1 si el cliente navega en la tienda online.
IND_USA_CMR_AGNO	Binaria	1 si el cliente usa la tarjeta de crédito.
IND_RECLAMOS_AGNO	Binaria	1 si el cliente realizó un reclamo.
IND_BAJA_FREQ_AGNO	Binaria	1 si el cliente compra productos de baja frecuencia.
IND_MEDIA_FREQ_AGNO	Binaria	1 si el cliente compra productos de media frecuencia.
IND_ALTA_FREQ_AGNO	Binaria	1 si el cliente compra productos de alta frecuencia.
IND_LEALTAD_AGNO	Binaria	1 si el cliente canjea puntos de lealtad.
IND_OTROS_AGNO	Binaria	1 si el cliente compra en regalos y servicios.
IND_BOLETAS_AGNO	Numérica	Indica cuántas boletas fueron emitidas en el año móvil.
IND_CANT_VENTA_AGNO	Numérica	Indica la cantidad de productos comprados.
IND_CANT_DEVOL_AGNO	Numérica	Indica la cantidad de productos devueltos.
VENTA_WEB_AGNO	Numérica	Venta anual del año móvil en la tienda online.
VENTA_TIENDA_AGNO	Numérica	Venta anual del año móvil en la tienda física.
VENTA_VEST_AGNO	Numérica	Venta anual del año móvil en vestuario.
VENTA_ELECT_AGNO	Numérica	Venta anual del año móvil en electro.
VENTA_DECO_AGNO	Numérica	Venta anual del año móvil en deco.
VISITA_VEST_AGNO	Numérica	Cantidad de visitas con compras a vestuario.
VISITA_ELECT_AGNO	Numérica	Cantidad de visitas con compras a electro.
VISITA_DECO_AGNO	Numérica	Cantidad de visitas con compras a deco.
LINEAS_COMPRADAS_AGNO	Numérica	Cantidad de líneas distintas compradas.
CANT_MESES_AGNO	Numérica	Cantidad de meses que realiza una compra.
REGENCY_DIAS_AGNO	Numérica	Recency en días al último día del período.
FRECUENCIA_DIAS_AGNO	Numérica	Frecuencia en días.

Fuente: Elaboración propia.

5.2. PRE-PROCESAMIENTO DE DATOS

Se consolida la información de los clientes, como fue indicado precedentemente, en dos bases de datos que incluyen data transaccional, demográfica, de navegabilidad, entre otras, correspondientes al año móvil de septiembre de 2015, 2016 y 2017.

Se observa que existen *missing values* asociados a la información sociodemográfica de los clientes, quienes, en su mayoría, corresponden a nuevos ingresos para la tienda. Como se observa en la Tabla 5, se define una variable categórica que busca capturar la falta de información presente en dichos campos.

Tabla 5: Variables creadas para evitar missing values.

Variable	Tipo	Descripción
RANGO_EDAD	Ordinal	Indica el rango de edad del cliente (0 a 25, 26 a 45, 46 años en adelante, sin información)
RANGO_GENERO	Ordinal	Indica el sexo del cliente (Femenino, Masculino, Sin información)
RANGO_ZONA	Ordinal	Indica la zona de residencia del cliente (Oriente, Poniente, Norte, Sur, Sin información)
CLTE_SIN_INFO	Ordinal	Indica qué información no posee el cliente (edad, género, zona y sus combinaciones)

Fuente: Elaboración propia.

Adicionalmente, en los datos transaccionales, se observa la presencia de clientes empresas que aumentan considerablemente la varianza de variables importantes para el negocio tales como venta, visitas y cantidad de productos comprados. De esta manera, se decide eliminar los valores extremos basado en el cálculo de la distancia de Mahalanobis que permite evaluar si un registro correspondería a un *outlier* de forma multivariada y no univariada al utilizar la herramienta box-plot.

Se evalúan los distintos casos asociados a las variables más relevantes para el negocio; venta, visitas y cantidad de productos. Dicho esto, en las Tabla 6 y Tabla 7 se detallan las diferentes combinaciones de cálculo para cada tipo de cliente.

Tabla 6: Casos evaluados para la determinación de valores extremos para clientes antiguos y nuevos.

Clientes Antiguos	
Variable 1	Variable 2
N_VENTA_ACT	VENTA_ACT
N_VENTA_ANT	VENTA_ANT
VENTA_ACT	VENTA_ANT
VISITAS_ACT	VISITAS_ANT
VENTA_ACT	VISITAS_ACT
VENTA_ANT	VISITAS_ANT
N_VENTA_ANT	VENTA_ANT
N_VENTA_ACT	VISITAS_ACT
N_VENTA_ANT	N_VENTA_ACT

Fuente: Elaboración propia.

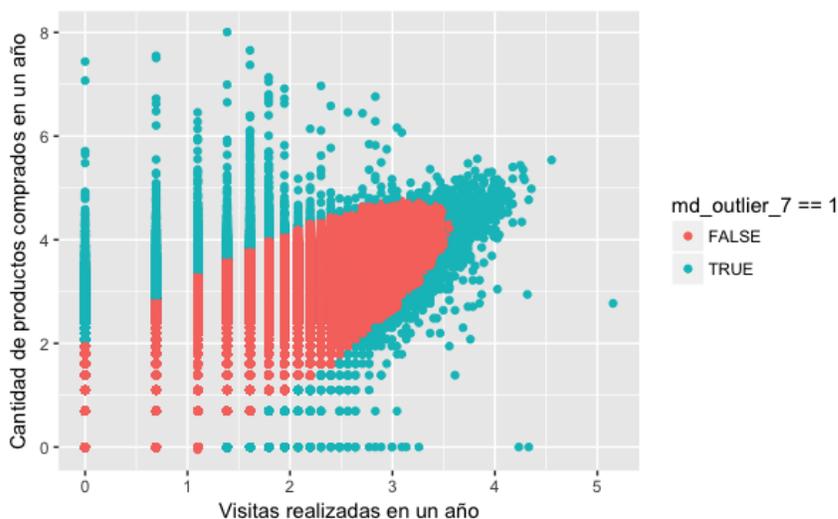
Tabla 7: Casos evaluados para la definición de valores extremos de clientes nuevos.

Clientes Nuevos	
Variable 1	Variable 2
N_VENTA_ACT	VENTA_ACT
VENTA_ACT	VISITAS_ACT
N_VENTA_ACT	VISITAS_ACT

Fuente: Elaboración propia.

A partir del Gráfico 2, se puede observar que la distancia de *Mahalanobis* elimina aquellos registros que más se alejan de los datos. De esta manera, se eliminan un total de 160.132 y 155.167 clientes, respectivamente.

Gráfico 2: Outliers para visitas y cantidad de productos.



Fuente: Elaboración propia.

Tabla 8: Distribución de clientes antes y después de la eliminación de valores extremos.

Distribución de clientes	Período móvil 2015		Período móvil 2016	
	Antiguos	Nuevos	Antiguos	Nuevos
Inicial	2.922.782	1.961.118	3.206.110	1.778.747
Final	2.815.130	1.908.638	3.091.829	1.737.861
Diferencia	107.652 (3.7%)	52.480 (2.7%)	114.281 (3.6%)	40.886 (2.3%)

Fuente: Elaboración propia.

Cabe destacar que, se conservan aquellos registros con venta total acumulada negativa, pues reflejan un fenómeno asociado a la devolución de productos por parte de los clientes que podría ser relevante de estudiar al generar el modelo de fuga.

5.3. TRANSFORMACIÓN

Se calculan variables asociadas a la frecuencia de compra de cada cliente, denominadas visitas, las que dependen de las compras que realiza un cliente en un día y tienda determinada.

Por otro lado, se generan variables que dependen de la línea o categoría que el cliente realiza compras, es decir, se crean variables binarias, de si compró en electro, deco, vestuario, y del monto que invierte o consume en cada una de esas líneas.

Adicionalmente, se incluyen variables asociadas a las compras que realiza en eventos particulares del retail: navidad, día del niño, día de la madre, liquidación, entre otros.

5.4. ANÁLISIS EXPLORATORIO

Para concretar lo mencionado en secciones anteriores, en la Tabla 9, se calcula el porcentaje de clientes fugados, definidos como aquellos que compraron en el año móvil 2016, pero no en el año móvil 2017 y categorizados por clientes antiguos y nuevos. A

partir de ello, se observa que el 44% de los clientes nuevos y el 19% de los antiguos se fuga de la tienda por departamento.

Por otro lado, en la Tabla 10, se puede observar que, del total de clientes fugados para el año móvil 2016, el 61% de ellos son clientes nuevos, mientras que el 39% corresponde a clientes antiguos.

Tabla 9: Distribución de fuga para los clientes del año móvil 2016.

Tipo de clientes	Cantidad de Clientes ⁵	Porcentaje de clientes fugados
Nuevos	71.574	44%
Antiguos	105.567	19%

Fuente: Elaboración propia.

Tabla 10: Distribución de clientes fugados para el año móvil 2016.

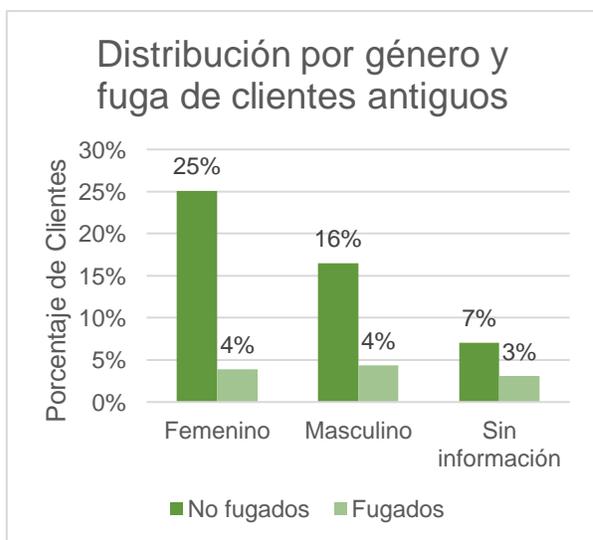
Tipo de clientes	Cantidad de Clientes ⁶	Porcentaje de clientes fugados
Nuevos	31.493	61%
Antiguos	20.058	39%

Fuente: Elaboración propia.

A continuación, se evidenciarán distintas distribuciones para diversas variables relevantes para el período móvil de 2016 (considera 12 meses, distribuidos entre 2015 y 2016) y se utiliza la información del año móvil de 2017 para determinar qué cliente se fuga y poder encontrar diferencias antes de la fuga.

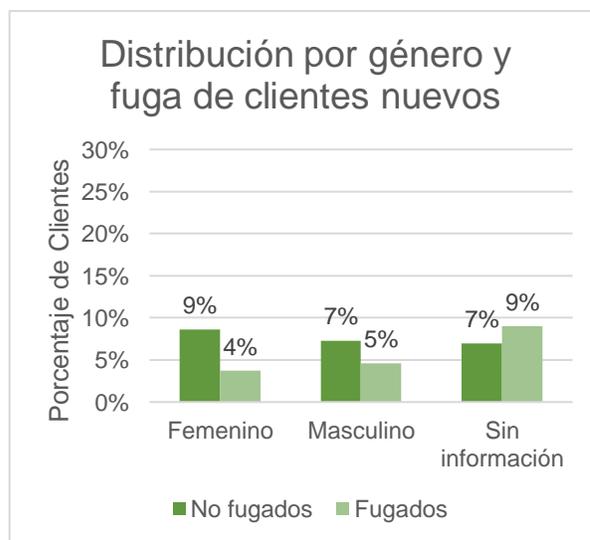
A partir de los Gráfico 3 y Gráfico 4, se observa que los consumidores de la tienda por departamento son mayoritariamente mujeres (42%), tanto para clientes nuevos (9%) como para antiguos (25%) que no se fugan, mientras que de los clientes que se fugan, la proporción es similar tanto entre mujeres (8%), hombres (9%) y aquellos en que la organización desconoce el género de la persona (12%).

Gráfico 3: Distribución de clientes antiguos por género



Fuente: Elaboración propia.

Gráfico 4: Distribución de clientes nuevos por género

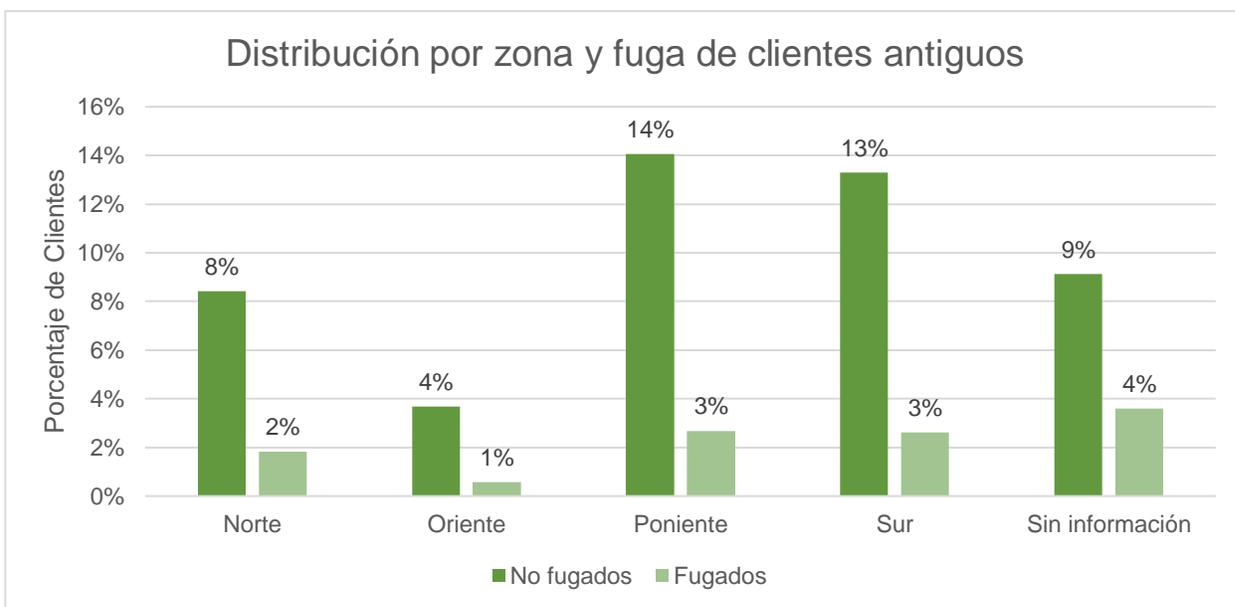


Fuente: Elaboración propia.

^{5 5} El número representa una proporción del total de clientes.

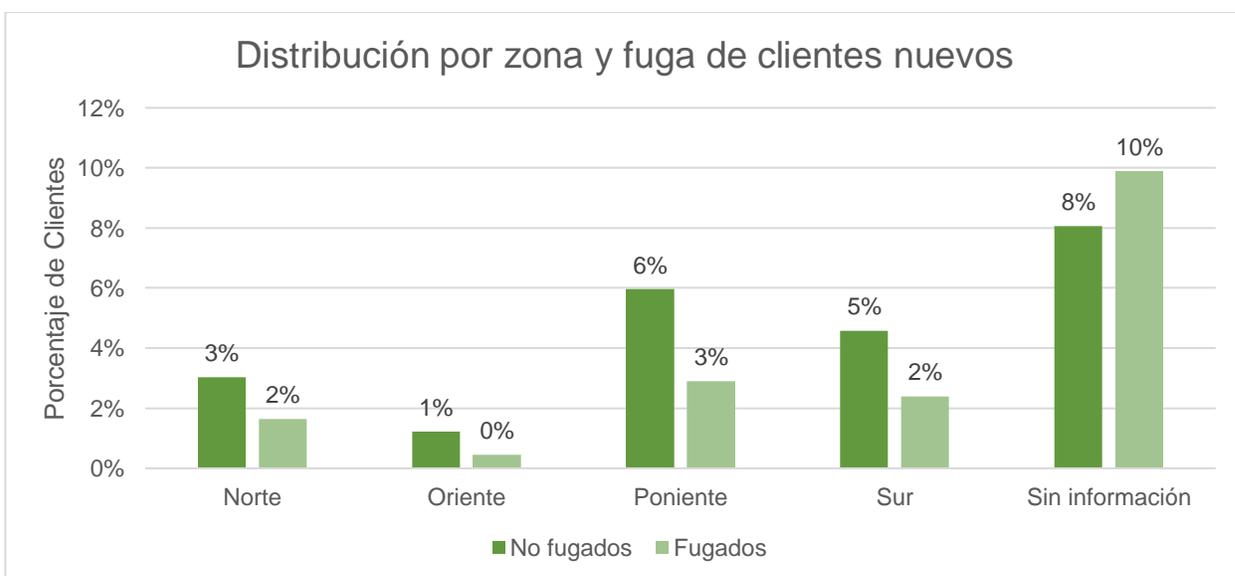
En cuanto a la procedencia de los clientes, se observa, a partir del Gráfico 5 y Gráfico 6, que los clientes de la tienda provienen de la zona poniente de la Región Metropolitana (26%), así como también de la zona sur del país (23%). No obstante, la procedencia de los clientes nuevos es mayoritariamente desconocida para la tienda (18%). Con respecto a los clientes que se fugan, se observa que viven en la zona poniente y sur del país, mientras que la componente asociada a la falta de información alcanza una gran relevancia, en especial para los clientes nuevos (10%), respecto del total de clientes.

Gráfico 5: Distribución de clientes antiguos por zona de residencia y fuga.



Fuente: Elaboración propia.

Gráfico 6: Distribución de clientes nuevos por zona de residencia y fuga.

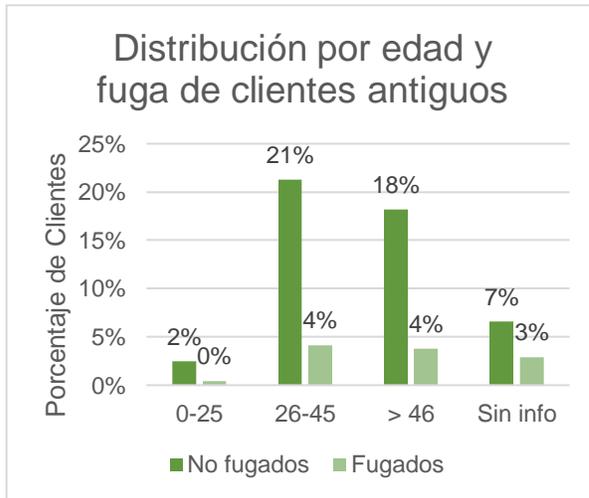


Fuente: Elaboración propia.

De los Gráfico 7 y Gráfico 8, se observa que los clientes se encuentran, mayoritariamente, entre 26 a 45 años (37%) seguido por quienes superan los 46 años (32%). Por otro lado,

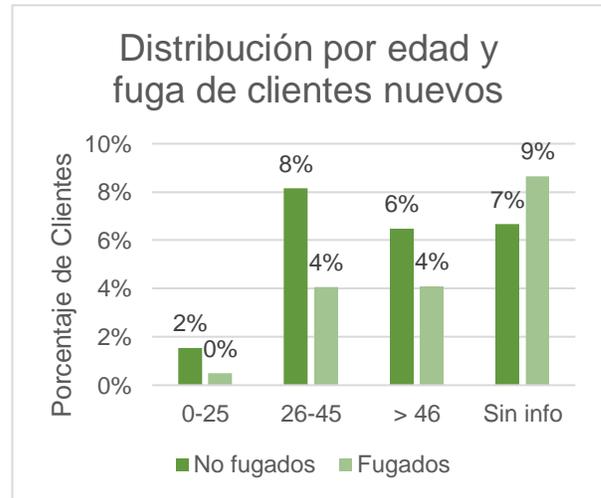
se reconoce que los fugados tienen la misma proporción tanto para clientes antiguos como para nuevos (entre 0% y 4%). No obstante, la categoría asociada a la falta de información tiene mayor concentración para clientes nuevos (9%) que para antiguos (3%).

Gráfico 7: Distribución de clientes antiguos por rango etario y fuga.



Fuente: Elaboración propia.

Gráfico 8: Distribución de clientes nuevos por rango etario y fuga.

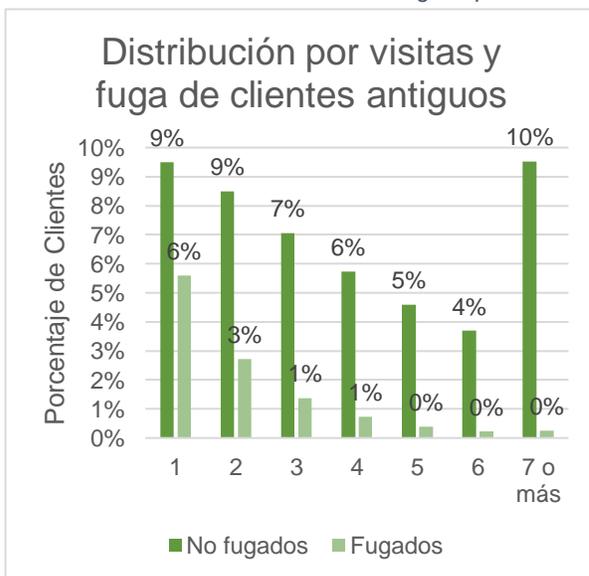


Fuente: Elaboración propia.

Como ya se conoce el perfil de los clientes de la tienda por departamento, resulta interesante comprender cómo se distribuyen las visitas con respecto a los clientes fugados definido como aquellos consumidores que dejaron de comprar en un año.

En los Gráfico 9 y Gráfico 10, se observa que los clientes nuevos suelen realizar una visita en los últimos doce meses (22%), mientras que los antiguos presentan una concentración en el intervalo de 1 (9%), 2 (9%) y 7 o más visitas (10%). Mientras que, los clientes fugados suelen acumular una visita en los últimos doce meses (6%).

Gráfico 9: Distribución de clientes antiguos por visitas.



Fuente: Elaboración propia

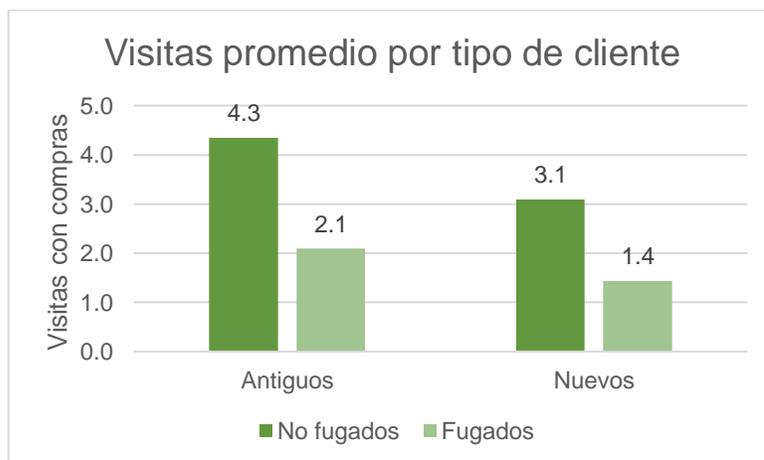
Gráfico 10: Distribución de clientes nuevos por visitas.



Fuente: Elaboración propia

Del Gráfico 11, se observa que los clientes antiguos y nuevos que no se fugan tienen, en promedio, 4.3 y 3.1 visitas, respectivamente, mientras que los clientes antiguos y nuevos que se fugan tienen, en promedio, 2.1 y 1.4 visitas, respectivamente. De esta manera, se evidencia que los clientes fugados tienen menor actividad que aquellos que continúan comprando en la tienda por departamento.

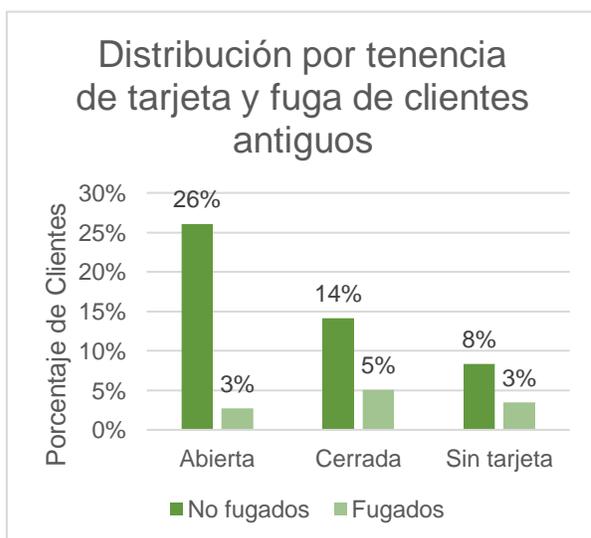
Gráfico 11: Visitas promedio por tipo de cliente y fuga.



Fuente: Elaboración propia

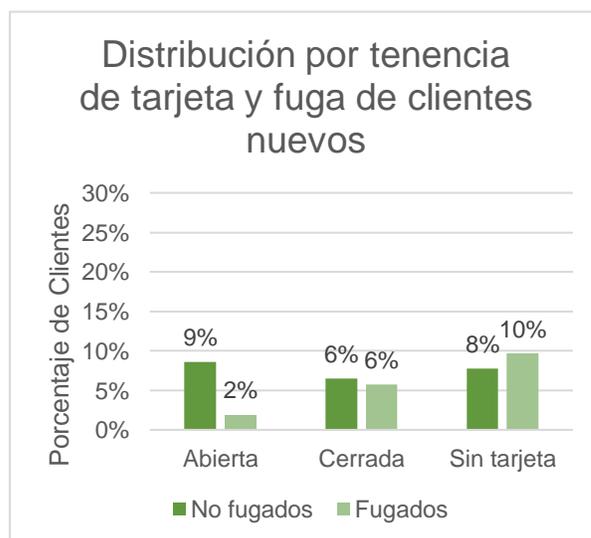
Se busca comprender si los clientes fugados presentan diferencias, al ser distribuidos por la tenencia de la tarjeta de crédito. De esta manera, en los Gráfico 12 y Gráfico 13, se observa que un 29% de los clientes antiguos están asociados con el producto, seguido por un 19% de clientes con cuenta cerrada debido a morosidad en el pago de sus obligaciones financieras. En cuanto a los clientes nuevos, la distribución es similar, donde predomina la cuenta cerrada (12%) y aquellos que no han contratado el producto (18%). Por otro lado, los clientes fugados son, en su mayoría, clientes que tienen la cuenta cerrada (11%) o bien, a la fecha de evaluación no habían contratado el producto (13%).

Gráfico 12: Distribución de clientes antiguos por tenencia de tarjeta y fuga.



Fuente: Elaboración propia.

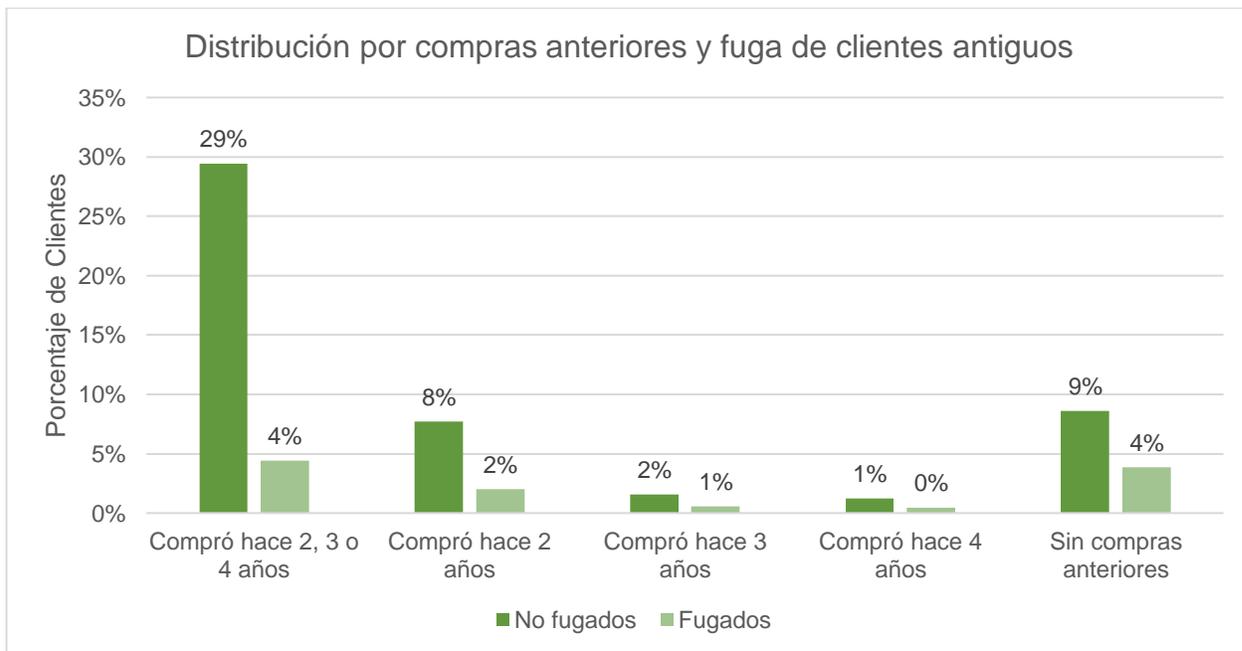
Gráfico 13: Distribución de clientes nuevos por tenencia de tarjeta y fuga.



Fuente: Elaboración propia.

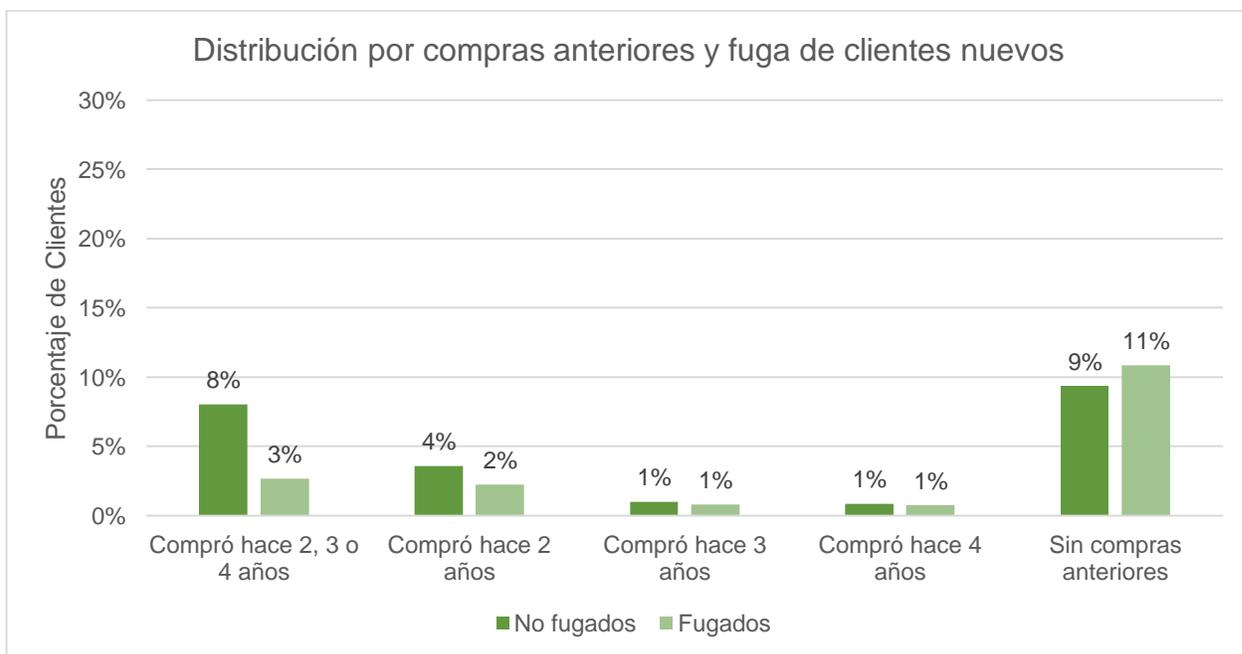
De los Gráfico 14 y Gráfico 15, se observa que la fuga de clientes (del año móvil 2017) se concentra en aquellos clientes que no tienen compras fuera del período de evaluación (entre 1 y 3 años antes del año móvil 2016), cifra que alcanza el 15% del total de clientes, mientras que el 14% restante se encuentra distribuido similarmente entre las categorías y por tipo de cliente.

Gráfico 14: Distribución de clientes antiguos por compras anteriores y fuga.



Fuente: Elaboración propia.

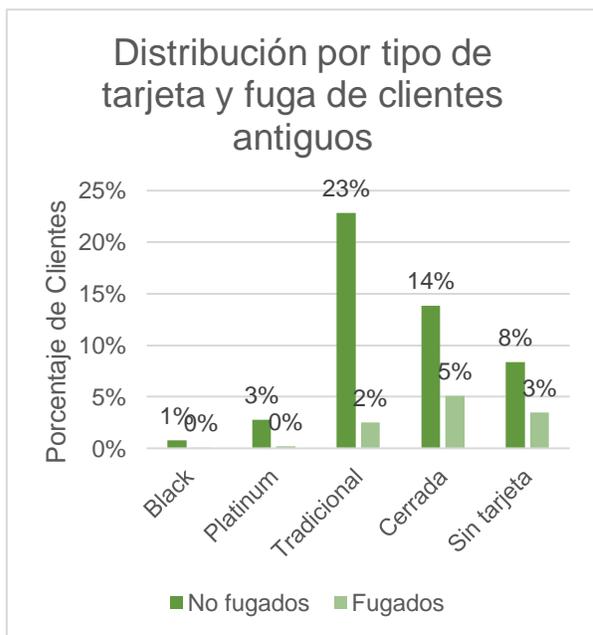
Gráfico 15: Distribución de clientes nuevos por compras anteriores y fuga.



Fuente: Elaboración propia.

Adicionalmente, se observa que los clientes fugados suelen no tener la tarjeta de crédito (13%) o bien la tuvieron en algún momento (11%), pero también existe una composición pequeña de clientes de la categoría Tradicional (4%), siendo este segmento, el más masivo de la tienda por departamento.

Gráfico 16: Distribución de clientes antiguos por tipo de tarjeta y fuga.



Fuente: Elaboración propia.

Gráfico 17: Distribución de clientes nuevos por tipo de tarjeta y fuga.



Fuente: Elaboración propia.

Por otro lado, se busca comprender cómo se relacionan las variables de recency, frecuencia y valor monetario. En este sentido, se realiza un análisis de RFM para comprender si alguna o un conjunto de dichas variables tendrá mayor relevancia al momento de realizar el modelo.

En la Tabla 11 y Tabla 12, se presentan los rangos asociados para construir la puntuación de RFM. De ellas, se observa que a menor score de Recency, más tiempo ha transcurrido desde la última compra del cliente; a menor score de Frecuencia, se observa que es un cliente que ha visitado la tienda en sólo una oportunidad y, a menor score de Valor Monetario, menos ha gastado el cliente.

Tabla 11: Distribución de variables RFM para clientes antiguos.

Recency [días]			Frecuencia [cantidad de compras]			Monetary value [\$]		
1	>= 221	<= 365	1	>= 1	<= 1	1	>= 1.080	<= 29.890
2	>= 122	< 221	2	> 1	<= 2	2	> 29.890	<= 66.960
3	>= 65	< 122	3	> 2	<= 4	3	> 66.960	<= 134.340
4	>= 27	< 65	4	> 4	<= 6	4	> 134.340	<= 264.790
5	>= 0	< 27	5	> 6	<= 32	5	> 264.790	<= 4.970.425

Fuente: Elaboración propia.

Tabla 12: Distribución de variables RFM para clientes nuevos.

Recency [días]			Frecuencia [cantidad de compras]			Monetary value [\$]		
1	>= 269	<= 365	1	>= 1	<= 1	1	>= 239	<= 13.980
2	>= 181	< 269	2	> 1	<= 2	2	> 13.980	<= 27.960
3	>= 100	< 181	3	> 2	<= 3	3	> 27.960	<= 52.990
4	>= 45	< 100	4	> 3	<= 5	4	> 52.990	<= 139.960
5	>= 0	< 45	5	> 5	<= 16	5	> 139.960	<= 7.069.755

Fuente: Elaboración propia.

De lo anterior, se desprende que aquellos clientes que obtienen la puntuación máxima (555) o cercana a ella corresponden a clientes de alto valor para la categoría, mientras que los clientes con la puntuación más baja (111) son aquellos que no proporcionan un valor a la organización.

A continuación, en la Tabla 13 y Tabla 14, se presenta el análisis de RFM tanto para clientes antiguos como nuevos, los cuales indican cómo distribuyen los clientes fugados, de acuerdo a las variables en estudio.

De la Tabla 13, se desprende que los clientes fugados antiguos se concentran entre el score 1 y 2 de Frecuencia, así como también en el score 1 de Recency, independiente del score del Valor Monetario. De esta manera, se puede concluir que los clientes antiguos que realizaron sólo una visita podrían ser más propensos a la fuga que aquellos con mayor frecuencia. Por otro lado, se observa que los clientes que no han realizado una compra en un período extenso de tiempo (sobre 7 meses), podría ser un indicio de que el cliente se va a fugar.

Cabe mencionar que los segmentos con mayor cantidad de clientes, son aquellos que se ubican cercanos al cuadrante (1,1,1) y al (5,5,5), mientras que los cercanos al (5,1,5) y 1,5,1) son los que tienen una menor cantidad de clientes.

Tabla 13: Análisis de RFM para clientes antiguos.

Porcentaje de fuga								
		Frecuencia						
		1	2	3	4	5		
Recency	5	27%	17%	9%	4%	1%	5	Monetary Value
		29%	18%	9%	4%	1%	4	
		29%	18%	10%	5%	2%	3	
		31%	19%	11%	5%	4%	2	
		32%	21%	14%	6%	4%	1	
Recency	4	30%	18%	10%	5%	2%	5	Monetary Value
		30%	19%	10%	5%	2%	4	
		31%	20%	11%	6%	3%	3	
		32%	21%	13%	7%	5%	2	
		33%	22%	15%	9%	3%	1	
Recency	3	29%	21%	12%	7%	4%	5	Monetary Value
		33%	21%	12%	7%	4%	4	
		32%	21%	13%	7%	4%	3	
		33%	22%	15%	8%	5%	2	
		34%	23%	16%	10%	10%	1	
Recency	2	33%	23%	16%	11%	9%	5	Monetary Value
		35%	24%	17%	11%	9%	4	
		35%	25%	17%	11%	8%	3	
		37%	25%	17%	11%	11%	2	
		37%	26%	19%	14%	16%	1	
Recency	1	37%	29%	21%	19%	20%	5	Monetary Value
		39%	30%	24%	19%	18%	4	
		40%	30%	24%	18%	18%	3	
		40%	29%	22%	19%	20%	2	
		40%	29%	23%	21%	10%	1	

Fuente: Elaboración propia.

De la Tabla 14, se observa que los clientes nuevos presentan comportamientos diferentes, en términos de las variables en estudio, donde la concentración de clientes fugados se genera en el score 1 o 2 de Frecuencia y score 1 de Recency, independiente del score asignado al Valor Monetario. Adicionalmente, destaca que los segmentos con mayor cantidad de clientes, se encuentran concentrados entre los score 1 y 2 de Frecuencia, independiente del Valor Monetario o del Recency.

Tabla 14: Análisis de variables RFM para clientes nuevos.

Porcentaje de fuga								
		Frecuencia						
		1	2	3	4	5		
Recency	5	48%	28%	18%	9%	2%	5	Monetary Value
		49%	30%	20%	12%	5%	4	
		51%	32%	21%	15%	11%	3	
		51%	34%	25%	18%	9%	2	
		53%	37%	28%	19%	5%	1	
Recency	4	52%	33%	21%	12%	4%	5	Monetary Value
		52%	34%	24%	15%	8%	4	
		54%	36%	26%	18%	10%	3	
		54%	38%	28%	21%	12%	2	
		56%	39%	32%	21%	0%	1	
Recency	3	56%	36%	26%	17%	10%	5	Monetary Value
		56%	39%	27%	20%	12%	4	
		57%	39%	28%	21%	16%	3	
		56%	38%	28%	26%	17%	2	
		57%	41%	31%	19%	20%	1	
Recency	2	63%	43%	32%	24%	19%	5	Monetary Value
		60%	43%	33%	25%	23%	4	
		60%	44%	32%	27%	20%	3	
		60%	43%	32%	28%	20%	2	
		60%	46%	38%	40%	0%	1	
Recency	1	62%	46%	38%	34%	38%	5	Monetary Value
		62%	48%	38%	35%	30%	4	
		61%	46%	38%	33%	50%	3	
		60%	47%	36%	32%	67%	2	
		60%	47%	39%	32%	100% ⁷	1	

Fuente: Elaboración propia.

De lo anterior, se desprende que la variable correspondiente a la frecuencia de compra, entendida como visitas con transacciones realizadas a la tienda, es capaz de discriminar la fuga tanto para clientes antiguos como nuevos.

⁷ En la casilla sólo se ubica un cliente y que, además, se fuga.

5.5. MODELO DE FUGA

Para la generación del modelo, es importante realizar una partición de la base de datos con la que se está trabajando, pues permite evaluar la capacidad predictiva del modelo que se está construyendo. Usualmente se realiza la división en dos bases, una llamada entrenamiento con el 80% de los datos y otra denominada testeo con el 20% de los registros restantes. Lo anterior, se realiza con el objetivo de generar un modelo con la partición de entrenamiento y evaluar la capacidad predictiva del mismo con la partición de testeo. No obstante, existen distintas variaciones, en términos de porcentajes de datos, así como también la inclusión de una tercera base llamada validación que posibilita la evaluación real de la predicción del modelo generado, donde la partición de entrenamiento mantiene su función, pero la partición de testeo se utiliza para evitar el sobreajuste de los parámetros del modelo utilizado.

En particular, se opta por crear tres particiones de la base de datos para entrenar el modelo, puesto que se puede obtener una evaluación más certera de cómo se está prediciendo el fenómeno en estudio, en las bases de testeo y validación. En la Ilustración 2, se detalla que la separación consiste en 50% para datos de entrenamiento, 30% para testeo y el 20% restante para validación.

Ilustración 2: Partición realizada para la generación del modelo.

Partición	Tamaño (%)	Etiqueta	Valor
Entrenamiento	50	Entrenamiento	"1_Entrenamiento"
Comprobación	30	Comprobación	"2_Comprobación"
Validación	20	Validación	"3_Validación"

Fuente: Elaboración propia, a partir de nodo Partición de SPSS Modeler.

Cabe destacar que, para construir un modelo predictivo, se requiere de una variable dependiente que sea de interés para estudiar y predecir, con el objetivo de focalizar estrategias comunicacionales del negocio. En particular, la generación de la variable dependiente que señala cuándo un cliente se considera fugado fue determinada de acuerdo al negocio, es decir, si un cliente no ha comprado en doce meses, está fugado.

5.5.1. ÁRBOLES DE DECISIÓN

Para estimar una probabilidad de fuga, se utiliza un árbol de decisión, pues es un modelo de *white-box* que permite entender las reglas condicionales que genera y comprender el fenómeno particular que se está estudiando que, en este caso, corresponde a la fuga de clientes y, de esta manera, verificar cuáles son las principales variables que lo explican.

En particular, se utiliza el algoritmo CHAID por sobre los demás tipos de árboles porque permite tener una ramificación no binaria, según las variables que el modelo considera que son relevantes para explicar el fenómeno.

Por otro lado, se definieron reglas de parada del árbol de decisión, de tal manera que no se genere un árbol tan extenso, lo que podría aumentar considerablemente el sobreajuste a los datos, disminuyendo la capacidad predictiva y, por consiguiente, reduciendo la posible gestión a realizar con reglas condicionales extensas.

Con lo anterior, las condiciones para generar el árbol de decisión fueron las siguientes:

- Algoritmo utilizado: CHAID.
- Ramas de profundidad máximas: 10.
- Mínimo de registros rama padre: 2%.
- Mínimo de registros rama hijo: 2%.

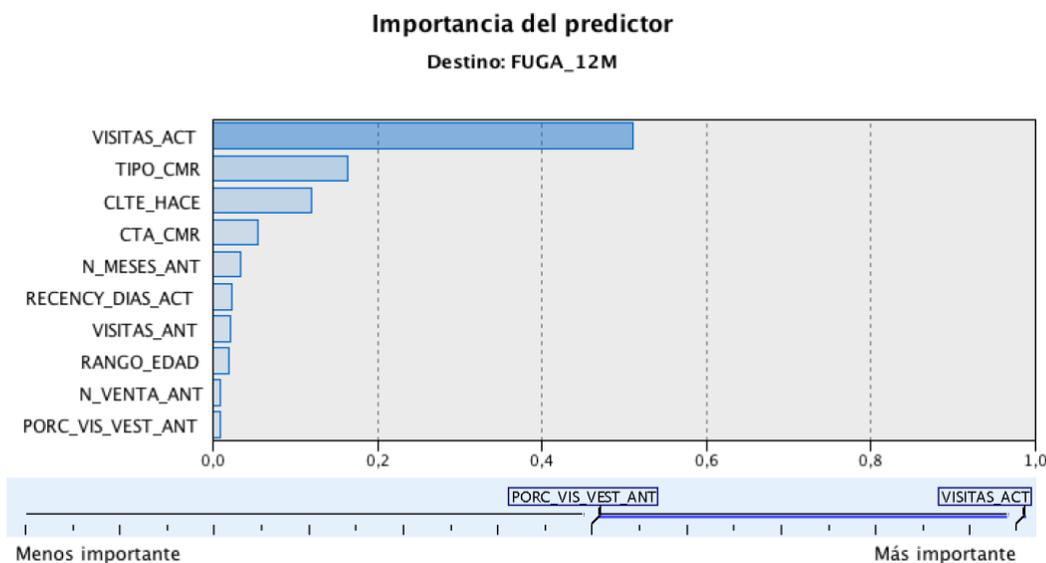
5.5.1.1. CLIENTES ANTIGUOS

Se realizó un modelo de propensión para predecir cuándo un cliente antiguo, es decir, una persona que ha comprado, al menos, una vez en cada uno de los períodos de evaluación, se fuga. Como se menciona precedentemente, el porcentaje de clientes antiguos fugados no supera el 19%, lo que da cuenta de que es una base no balanceada.

En la Ilustración 3, se pueden observar las variables más relevantes para determinar la probabilidad de fuga de cada cliente, las cuales se encuentran ordenadas por importancia de cada predictor, la que se evalúa, a partir de la partición de comprobación e indica la importancia relativa de cada variable para el modelo realizado.

Cabe destacar que la variable asociada a la frecuencia de compra (visitas), tiene una importancia relativa del 51% respecto al fenómeno estudiado. Sin embargo, si se consideran las variables del tipo de tarjeta de crédito y la antigüedad del cliente, la importancia aumenta al 79%.

Ilustración 3: Variables más importantes para el modelo de clientes antiguos.



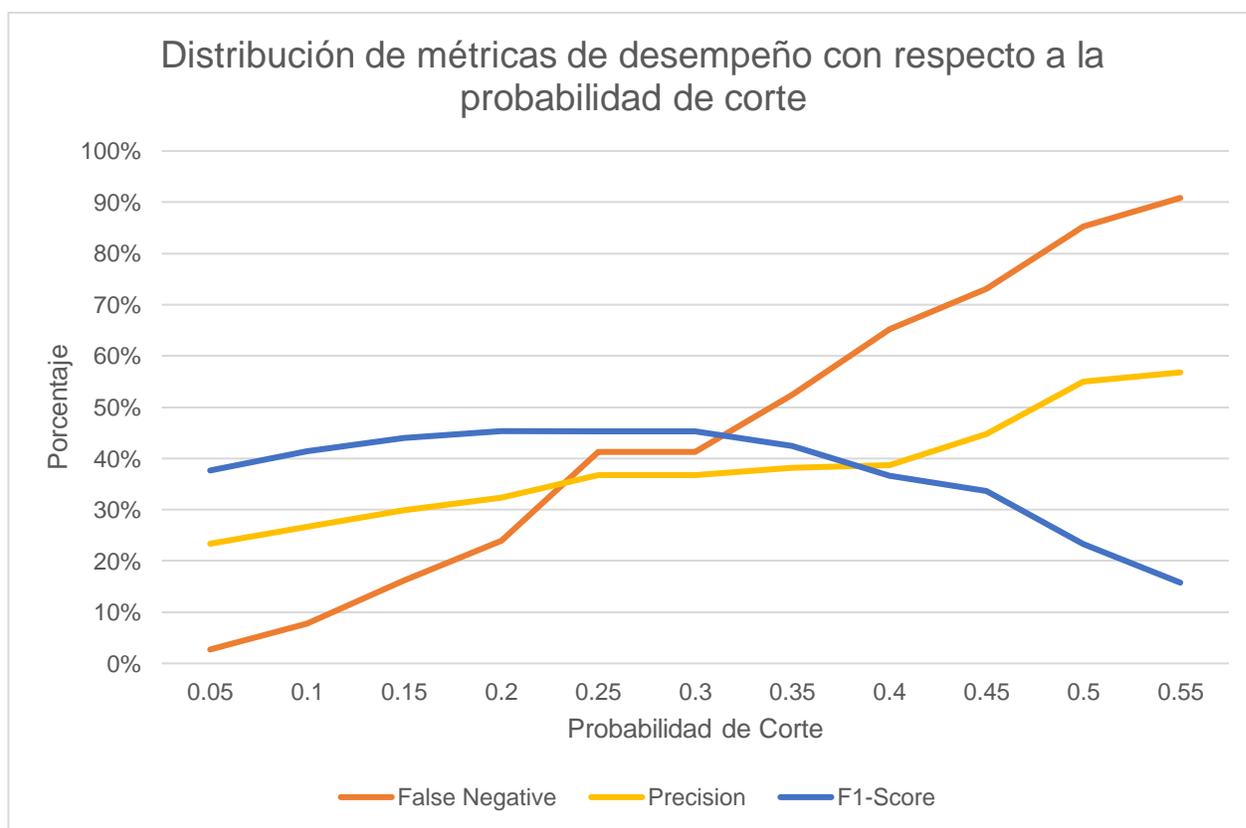
Fuente: Elaboración propia.

Por otro lado, cabe mencionar que, comúnmente los modelos están configurados para definir un caso de éxito cuando la probabilidad calculada supera el umbral del 50%, lo que proporciona un modelo predictivo de baja calidad cuando la proporción de casos de éxitos y no éxitos no está balanceada.

Por lo anterior, se evalúan tres métricas de desempeño relevantes (falsos negativos, precisión y F1-score) en intervalos de 5%, con el objetivo de determinar cuál es el punto de corte más apropiado.

De esta manera, se escoge el punto que maximice la medida de armonía entre la precisión y *recall*, representado por F1-score. A partir del Gráfico 18 y de la Tabla 65, se observa que el máximo valor ocurre cuando el corte equivale a 0.25. De esta manera, se define un caso de éxito (un cliente fugado), cuando la probabilidad es igual o superior al 25%.

Gráfico 18: Evaluación de distintas métricas, en función de la probabilidad de corte para clientes antiguos.



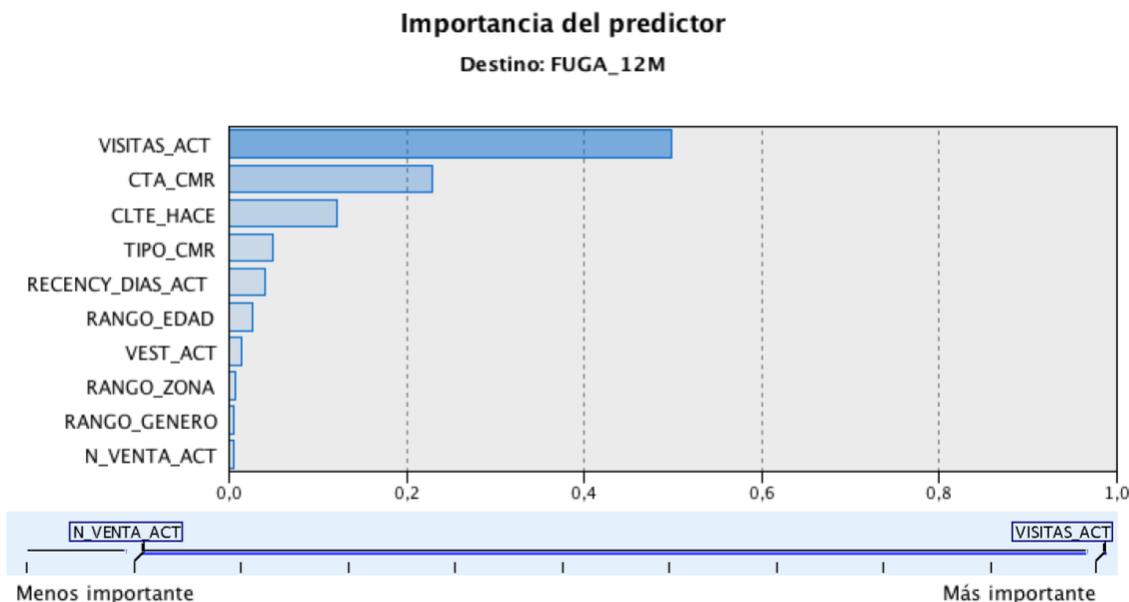
Fuente: Elaboración propia.

5.5.1.1. CLIENTES NUEVOS

Adicionalmente, se realizó un modelo de propensión para predecir cuándo un cliente nuevo, es decir, aquel individuo que ha comprado en los últimos doce meses, pero no en los doce meses anteriores a ese período, se fuga. En secciones anteriores, se menciona que el porcentaje de clientes fugados asciende a 44%, cifra muy superior a la cantidad de clientes antiguos fugados.

En la Ilustración 3, se pueden observar las variables más importantes para determinar la probabilidad de fuga de cada cliente, ordenadas por importancia. Cabe destacar que la variable asociada a la frecuencia de compra, es capaz de explicar el 50% del fenómeno por sí sola. Sin embargo, si se consideran las variables del estado de la cuenta de la tarjeta de crédito y la antigüedad del cliente, se logra explicar el 85%.

Ilustración 4: Variables más importantes para el modelo de clientes nuevos.



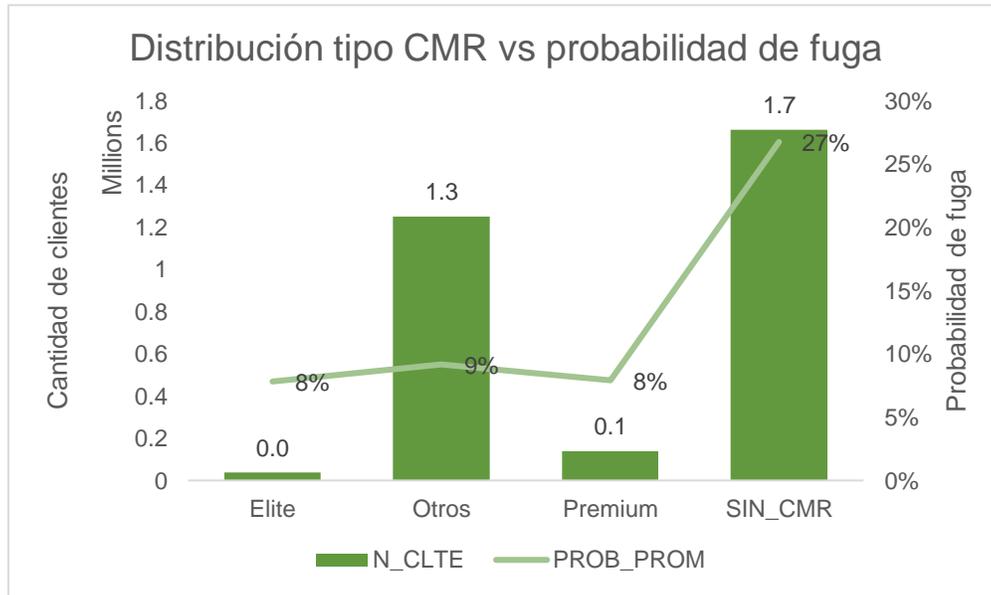
5.5.1.2. INTERPRETACIÓN DEL MODELO

Con el modelo generado y la probabilidad de corte modificada, se procede a comprender cómo afectan cada una de las variables tales como: tenencia de la tarjeta de crédito, rangos etarios, sociodemográficos, género, entre otros, en la propensión de fuga de un cliente de la tienda por departamento.

De esta forma, en el Gráfico 19, se observa que no tener la tarjeta de crédito asociada al negocio aumenta a 27% la probabilidad de fuga, mientras que clientes con alguna de las diferentes categorías de la tarjeta presentan una probabilidad promedio que fluctúa entre el 8 y 9%.

De lo anterior, se desprende que los clientes que no tienen una relación contractual con la organización, mediante la contratación de la tarjeta de crédito del holding, son clientes con mayor probabilidad de fuga. Esto puede dar evidencia de que son clientes esporádicos que persiguen comprar, mayoritariamente, productos con descuentos, más que vestirse en la tienda por departamento.

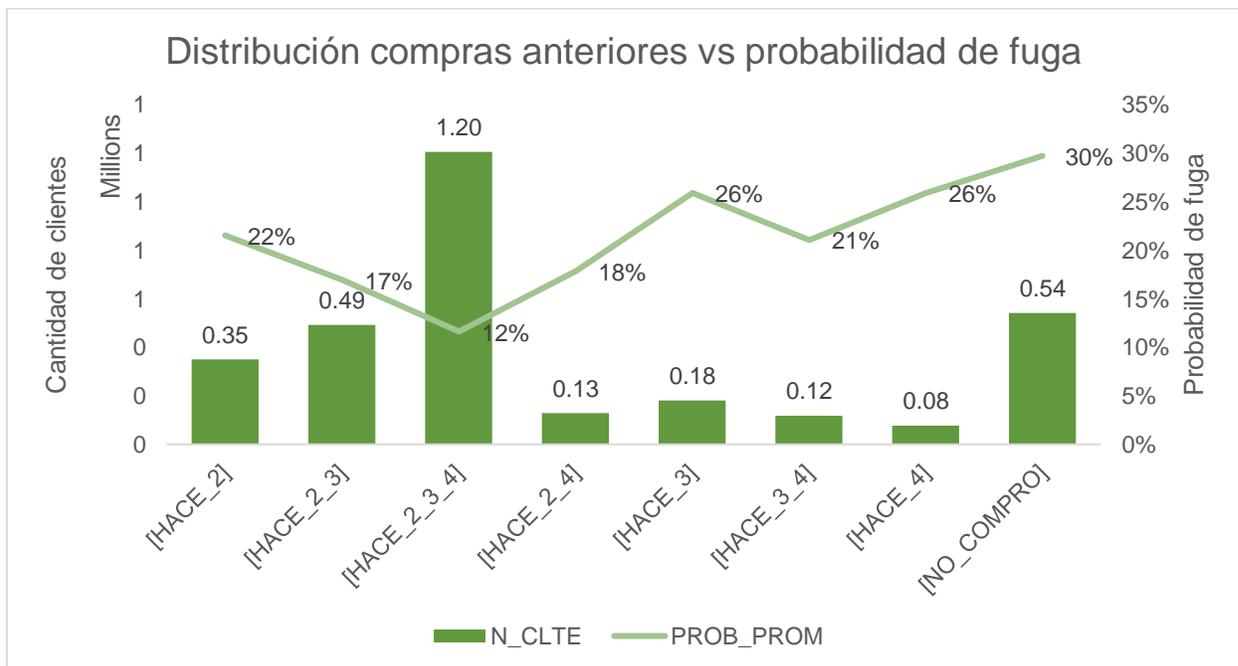
Gráfico 19: Distribución del tipo de tarjeta CMR con respecto a la probabilidad de fuga.



Fuente: Elaboración propia.

Por otro lado, si analizamos la variable asociada a la antigüedad que posee el cliente, se observa del Gráfico 20 que los clientes que no han comprado en períodos anteriores a la evaluación (dos, tres o cuatro años), tienen un 30% de probabilidad de fuga, a diferencia de los clientes que han comprado constantemente, donde su probabilidad de fuga disminuye a un 12%. Del mismo modo, se reconoce que existen clientes que han presentado compras en sólo un período o que se han fugado en el pasado, tienen una probabilidad superior que fluctúa entre 22 y 26%.

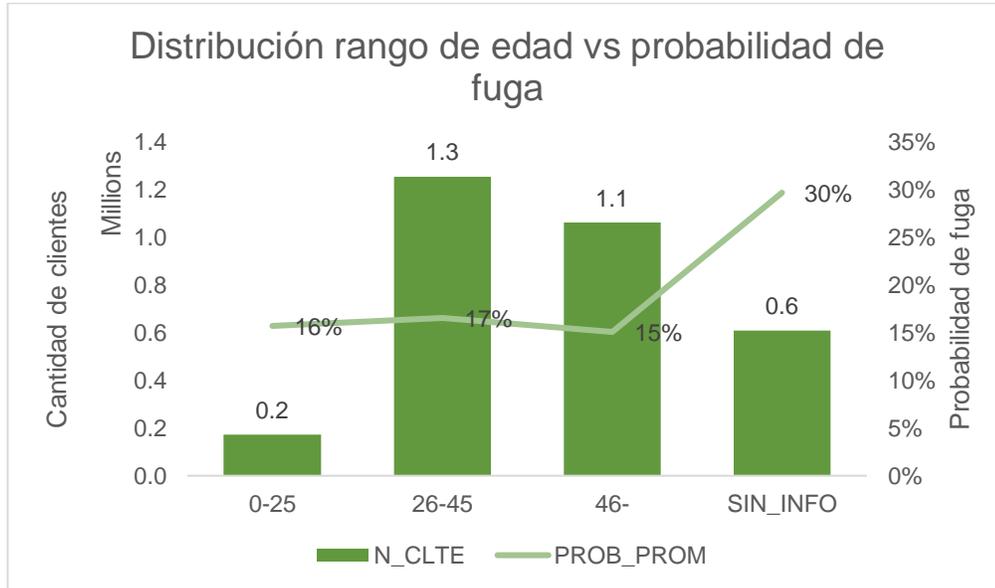
Gráfico 20: Distribución de la antigüedad del cliente con respecto a la probabilidad de fuga.



Fuente: Elaboración propia.

Por otro lado, del Gráfico 21, se observa que los clientes de los cuales la tienda no posee información del rango etario, tienen un 30% de probabilidad de fuga, a diferencia de los clientes con los que cuentan con la información, la probabilidad disminuye a 15 y 17%, siendo los clientes pertenecientes al rango [26-45] los que presentan un 17% de probabilidad de fuga.

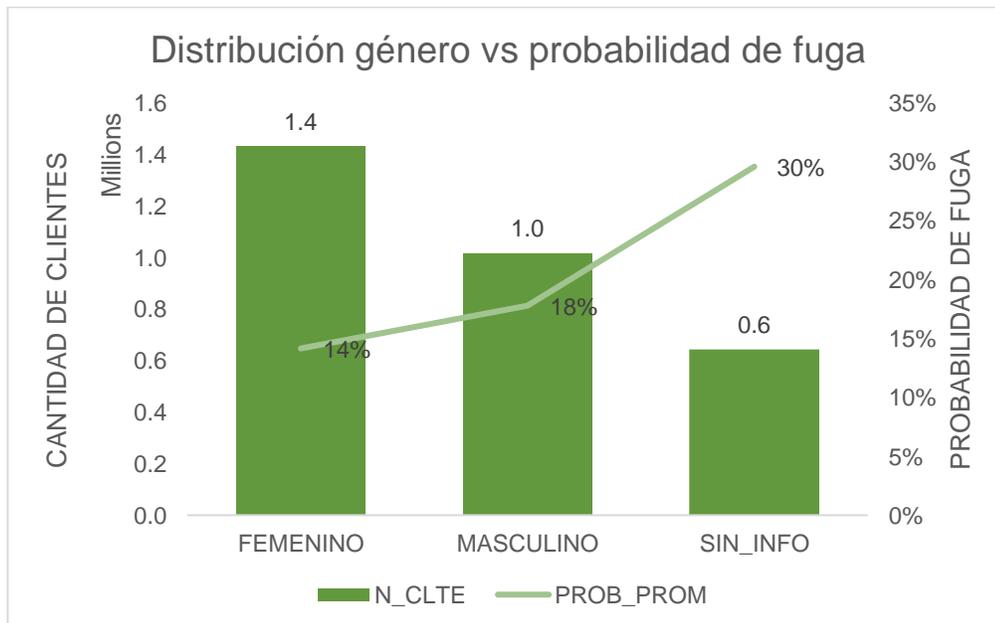
Gráfico 21: Distribución del rango etario con respecto a la probabilidad de fuga.



Fuente: Elaboración propia.

En cuanto al género, se observa, del Gráfico 22 que los clientes del sexo masculino, tienen un 18% de probabilidad, mientras que las mujeres alcanzan un 14%. Además, se evidencia que aquellos consumidores que no presentan información de género tienen un 30% de probabilidad de fuga.

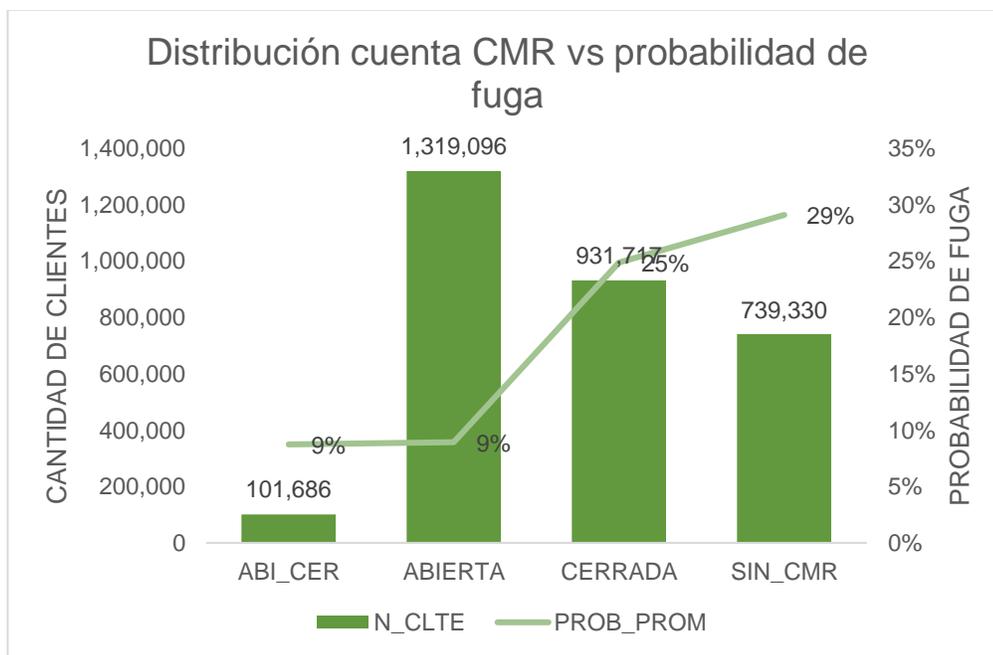
Gráfico 22: Distribución de género con respecto a la probabilidad de fuga.



Fuente: Elaboración propia.

Finalmente, del Gráfico 23, se observa que los clientes con cuenta cerrada tienen un 25% de probabilidad de fuga, mientras que los clientes que no poseen la tarjeta de crédito tienen un 29% de probabilidad de fuga.

Gráfico 23: Distribución del tipo de cuenta con respecto a la probabilidad de fuga.



Fuente: Elaboración propia.

5.5.1.3. MÉTRICAS DE DESEMPEÑO

Para evaluar la capacidad predictiva de un modelo, se calculan distintas métricas de desempeño, entre las cuales se encuentran *accuracy*, *precision*, *recall* y F1-Score.

En la Tabla 15, se pueden observar dichos resultados tanto para el modelo de clientes antiguos como el de nuevos. De ella, se evidencia el efecto que tienen las bases desbalanceadas, entre casos de éxito y no éxito (fuga y no fuga), donde la precisión, es decir, cuánto logra predecir el modelo de los casos de éxito, es inferior para el caso del modelo de clientes antiguos (40%), no así para los nuevos, donde la precisión asciende a 63%.

Por otro lado, el *accuracy*, es decir, cuánto logra predecir de los casos idénticos (éxito, éxito; no éxito, no éxito), es superior para el modelo de clientes antiguos, donde alcanza un 76%, mientras que para el de nuevos, es de 68%.

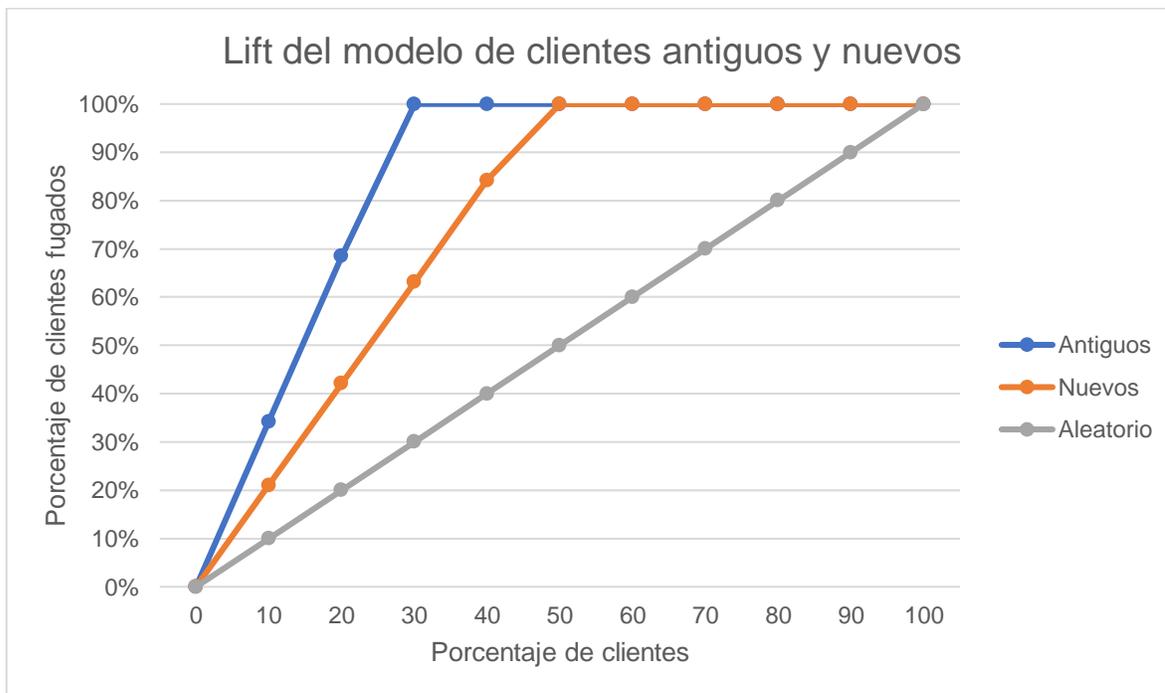
Tabla 15: Métricas de desempeño para los modelos realizados.

Métrica	Antiguos	Nuevos
Accuracy	76%	68%
Precision	40%	63%
Recall	58%	64%
F1-Score	48%	64%

Fuente: Elaboración propia.

Por otro lado, en el Gráfico 24, se evalúa el *lift* para ambos modelos, donde se puede observar que con el 10% de los clientes antiguos, se alcanza un ~34% de clientes fugados, lo que equivale a un *lift* de 2.6, mientras que con el 10% de los clientes nuevos, se obtiene un ~21% de clientes fugados, lo que equivale a un *lift* de 1.6.

Gráfico 24: Comparación de Lift tanto para el modelo de clientes antiguos como para el de nuevos.



Fuente: Elaboración propia.

5.6. DISEÑO EXPERIMENTAL

En esta sección se incluyen los experimentos realizados y se detallan los resultados obtenidos de cada uno de ellos. Se pone énfasis en el cálculo del tamaño muestral, de tal manera, de quedar con una base suficientemente grande de clientes para evitar que existan promociones sin respuesta, pues, dado el problema, se trabajará con una cartera deteriorada de clientes (bajas compras y tasas de respuesta).

5.6.1. MUESTRA DE CLIENTES

Para determinar quiénes son los clientes de la tienda por departamento que recibirán el incentivo, se excluyen ciertos segmentos de consumidores que, por políticas de la compañía, se filtran para cada envío de promociones, independiente del tipo o categoría de campaña.

1. Grupo control anual.
2. Lista negra SERNAC.

Adicionalmente, se excluyeron clientes que pertenecían a los siguientes segmentos de clientes.

1. Funcionarios Falabella.

2. Clientes que habían recibido el primer incentivo de la campaña de retención del mes de octubre.
3. Clientes con baja propensión de fuga, es decir, aquellos que se encontraban dentro de los cinco deciles de probabilidad más baja (menos probable que se fuguen).

Posterior a la exclusión de dichos segmentos, se disponía de una base potencial aproximada de 353.000 clientes.

5.6.2. TAMAÑO MUESTRAL

Para calcular el tamaño muestral, se utiliza la fórmula enunciada en 3.4.3 Muestreo, donde se definen los valores de cada una de las variables utilizadas, según lo indicado a continuación.

- Nivel de confianza (K): 95%, equivalente a 1.96
- Se asume un tamaño de población muy grande (N = 353.000)
- Se asume variabilidad máxima, es decir, $p = q = 0.5$.

De esta manera, se obtiene un tamaño muestral, en primera instancia, de 384 clientes. Sin embargo, debido a que los clientes que serán contactados no tienen un alto relacionamiento con la organización, se estima, con la contraparte, la respuesta esperada de los clientes, la que fluctúa entre 1.0-1.2%. Para el cálculo posterior, se considerará el promedio del rango mencionado, es decir, 1.1%.

Basado en (Bartlett, Kotrlík, & Higgins, 2001), se puede ajustar el tamaño muestral obtenido anteriormente, de acuerdo a la respuesta esperada, según el cociente entre ambos números, obteniendo, por tanto, un tamaño muestral de ~35.000 clientes.

5.6.3. PROMOCIONES REALIZADAS

Como se menciona en 4.6.4 Condiciones experimentales, se generaron seis campañas que variaban en el beneficio entregado. En particular, se proporcionaron los descuentos que se enuncian a continuación:

Tabla 16: Subject de promociones realizadas

Promoción	Subject
\$5.000 de dcto. sobre \$15.000.	Te regalamos \$5.000 en tu próxima compra.
\$10.000 de dcto. sobre \$30.000.	Te regalamos \$10.000 en tu próxima compra.
20% de descuento.	Te regalamos un 20% dcto. en tu próxima compra.
30% de descuento.	Te regalamos un 30% dcto. en tu próxima compra.
\$5.000 de descuento.	Te regalamos \$5.000 en tu próxima compra.
\$10.000 de descuento.	Te regalamos \$5.000 en tu próxima compra.

Fuente: Elaboración propia.

Por otro lado, las promociones estuvieron vigentes desde el 10 de noviembre hasta el 19 de noviembre de 2018.

5.6.4. GRUPO CONTROL Y TRATAMIENTO

Para llevar a cabo el experimento y evaluar el efecto correspondiente, se construye, mediante una muestra aleatoria, un grupo control y un grupo tratamiento por cada promoción a realizar, poniendo énfasis en que sean grupos similares estadísticamente, a través del test de Levene (varianzas) y test t (medias).

De lo anterior, se obtiene un grupo control de 35.000 clientes y cada grupo tratamiento de 53.000 clientes.

5.6.4.1. TEST DE LEVENE Y TEST T

Con el objetivo de validar que las muestras son comparables en medias y varianzas, se realizan dos test de hipótesis: test de Levene y test t.

El test de Levene plantea, en su hipótesis nula, que si el estadístico o *p-valor* supera la barrera del 5%, se puede suponer varianzas iguales para la variable en estudio.

Por otro lado, el test t busca validar la existencia de medias iguales para los grupos en comparación, es decir, si el estadístico supera el 0,05 significa que es posible suponer igualdad de medias para la variable y los grupos en estudio.

En las Tabla 17, Tabla 18 y Tabla 19, se presentan los resultados de los estadísticos y las medias tanto para el grupo control como los grupos tratamientos generados. De ellas, se puede concluir que todos los grupos promocionales son comparables, con excepción del incentivo de \$5.000 de descuento que, para la variable *recency*, no es posible afirmar igualdad de medias. Sin embargo, la diferencia entre el grupo tratamiento y control es de 0.94 días, más aún, no se evidencian grandes diferencias entre las tres variables restantes, por lo que es posible determinar que todos son grupos comparables .

Tabla 17: Test de medias y varianzas para las promociones de monto de descuento.

Promoción	Media GC	\$5.000 de descuento			\$10.000 de descuento		
		Media GM	Sig. Test Levene	Sig. Test T	Media GM	Sig. Test Levene	Sig. Test T
Probabilidad	0,33	0,33	0,35	0,59	0,33	0,44	0,41
Venta	90.564,35	91.120,60	0,73	0,43	90.517,25	0,75	0,95
Visitas	1,96	1,97	0,87	0,20	1,97	0,18	0,36
Recency	155,83	154,89	0,08	0,01	155,08	0,51	0,29

Fuente: Elaboración propia.

Tabla 18: Test de medias y varianzas para las promociones de montos con restricción de ticket.

Promoción	Media GC	\$5.000 sobre \$15.000			\$10.000 sobre \$30.000		
		Media GM	Sig. Test Levene	Sig. Test T	Media GM	Sig. Test Levene	Sig. Test T
Probabilidad	0,33	0,33	0,85	0,84	0,33	0,60	0,79
Venta	90.564,35	90.686,66	0,47	0,86	90.841,49	0,28	0,70
Visitas	1,96	1,96	0,56	0,84	1,97	0,11	0,19
Recency	155,83	155,82	0,93	0,98	155,30	0,91	0,45

Fuente: Elaboración propia.

Tabla 19: Test de medias y varianzas para las promociones de porcentajes de descuentos.

Promoción	Media GC	20% de descuento			30% de descuento		
		Media GM	Sig. Test Levene	Sig. Test T	Media GM	Sig. Test Levene	Sig. Test T
Probabilidad	0,33	0,33	0,71	0,52	0,34	0,24	0,49
Venta	90.564,35	90.301,04	0,73	0,71	91.357,01	0,05	0,26
Visitas	1,96	1,97	0,32	0,55	1,97	0,17	0,35
Recency	155,83	155,74	0,18	0,90	155,66	0,53	0,81

Fuente: Elaboración propia.

5.6.5. TASA DE ACTIVACIÓN DE CLIENTES

Para contactar a los clientes mediante un email, se utiliza una plataforma de envíos de correos electrónicos masiva. De ella, es posible realizar un seguimiento de aquellos que recibieron el correo, lo abrieron e incluso, si hicieron clic en él.

En la Tabla 20, se evidencia que alrededor del 94% de la base de clientes recibió el correo promocional en su bandeja de entrada. Lo anterior se explica debido a exclusiones adicionales que se generan en la plataforma de envíos masivos, donde se excluyen a clientes que se desuscribieron; correos que rebotaron en campañas previas; y, como doble verificación, se vuelve a eliminar a clientes que se encuentren en el grupo control anual y que hayan reclamado ante Sernac por el envío desproporcionado de publicidad⁸.

Por otro lado, se observa que la tasa de apertura de cada una de las campañas bordea el 20%. Destaca que, dicho porcentaje se encuentra en el margen de apertura de las promociones de gestión de clientes de bajo relacionamiento que realiza la organización⁹. Cabe mencionar que, los incentivos focalizados en clientes frecuentes o de alto valor, alcanzan tasas de apertura que, en ocasiones, superan el 45%, lo que se explica debido a que son consumidores de alto relacionamiento con la tienda por departamento.

De esta manera, se observa que las promociones de montos de descuento sin limitación (\$5.000 y \$10.000) son capaces de activar a más clientes que las de porcentajes de

⁸ El último filtro se considera por políticas de la organización. No obstante, se validó que dicha exclusión no eliminó más clientes, puesto que ya habían sido extraídos de la base que recibirían la promoción.

⁹ En junio, el porcentaje de apertura fue de 18.9%. Durante agosto, la campaña se filtró por porcentaje de apertura y se alcanzó un 24.2%.

descuento o montos sujeto a cierta compra. En particular, se evidencia que las con restricción de monto son las que tuvieron la menor activación, lo que puede deberse a que los clientes de bajo relacionamiento consideran que dicha condición es una barrera muy alta para optar por el descuento, no así con la de montos de regalo que, independiente de lo que compre, accederá al descuento.

Tabla 20: Resultados de las promociones enviadas.

Tipo promoción	Total	Enviados	Vistos	Aperturas	Canjes	Tasa de activación (sobre vistos)
20% dcto.	53.000	49.951	10.035	20.09%	111	1.11%
30% dcto.	53.000	49.889	10.199	20.44%	211	2.07%
\$5.000	53.000	49.968	10.320	20.65%	217	2.10%
\$10.000	53.000	49.969	10.397	20.81%	572	5.50%
\$5.000 sobre \$15.000	53.000	50.006	10.355	20.71%	105	1.01%
\$10.000 sobre \$30.000	53.000	49.899	10.459	20.96%	106	1.01%
Total	318.000	299.682	61.765	20.61%	1.322	2.14%

Fuente: Elaboración propia.

5.6.6. VENTA INCREMENTAL

Además de lo enunciado precedentemente, existe otra forma de evaluar el desempeño de un incentivo, que consiste en estimar la venta incremental que se genera por cliente contactado para cada una de las promociones realizadas.

De esta manera, se evalúa la activación y la venta de cada uno de los grupos tratamientos y se compara con la obtenida por el grupo control definido. Así es como, en la Tabla 21, se puede observar que se genera una venta incremental por cliente contactado que varía entre \$55 y \$178.

Por otro lado, si se ordenan las promociones, según la venta incremental que generaron, se obtiene:

1. \$10.000 dcto.
2. 30% dcto.
3. \$10.000 sobre \$30.000.
4. \$5.000 sobre \$15.000.
5. \$5.000 dcto.
6. 20% dcto.

De lo anterior, se concluye que los incentivos más agresivos o que más beneficio proporcionan, son aquellos que más venta incremental generan. Lo que se explica, pues si un cliente con bajo relacionamiento recibe un incentivo, será más propenso a utilizarlo, a medida que mayor sea el descuento proporcionado.

En la misma línea, se observa que la promoción de \$10.000 de descuento fue la que más venta incremental genera, puesto que, corresponde a un alto beneficio monetario, sin restricciones como lo son las promociones que requieren cumplir cierto monto para utilizarlas.

Tabla 21: Venta incremental por promoción enviada.

Promoción	Grupo Mail	Grupo Control	Tasa rpta. GM	Rpta. Incr.	Venta Incr.	Venta Incr. por Clte
20% dcto.	53.000	35.000	3,2%	0,2%	\$ 2.877.818	\$ 55
30% dcto.	53.000	35.000	3,4%	0,4%	\$ 7.961.294	\$ 153
\$5.000 dcto.	53.000	35.000	3,4%	0,3%	\$ 3.315.291	\$ 64
\$10.000 dcto.	53.000	35.000	3,9%	0,9%	\$ 9.278.813	\$ 178
\$5.000 sobre \$15.000	53.000	35.000	3,2%	0,2%	\$ 3.636.360	\$ 70
\$10.000 sobre \$30.000	53.000	35.000	3,1%	0,1%	\$ 4.317.649	\$ 83

Fuente: Elaboración propia.

De esta manera, se obtiene una venta incremental por cliente contactado promedio, entre todas las promociones, de \$101.

Con el objetivo de comparar los resultados mencionados anteriormente, se considerará la venta incremental generada por el primer toque de la campaña de retención que realizó la organización desde febrero hasta septiembre. De esto, se obtiene un promedio de \$57 de venta incremental por cliente contactado, el cual corresponde a un 56% de la venta incremental obtenida por las promociones realizadas.

Por otro lado, si incorporamos los costos de cada una de las promociones, como se observa en la Tabla 22, se obtiene que el incentivo de \$10.000 deja de ser el más rentable, superándolo por el de 30% de descuento.

En este sentido, se evidencia que, si bien la promoción de \$10.000 de descuento fue la que generaba mayor activación entre los clientes que fueron contactados, no es la más rentable, debido a sus altos costos asociados (\$ 93).

Tabla 22: Costos y utilidad generada por cada una de las promociones.

Promoción	Venta Incremental por cliente contactado	Costos por cliente contactado	Utilidad
20% dcto.	\$ 55	\$ 15	\$ 40
30% dcto.	\$ 153	\$ 52	\$ 101
\$5.000 dcto.	\$ 64	\$ 18	\$ 46
\$10.000 dcto.	\$ 178	\$ 93	\$ 85
\$5.000 sobre \$15.000	\$ 70	\$ 9	\$ 61
\$10.000 sobre \$30.000	\$ 83	\$ 18	\$ 65

Fuente: Elaboración propia.

5.6.7. CARACTERIZACIÓN DE CLIENTES

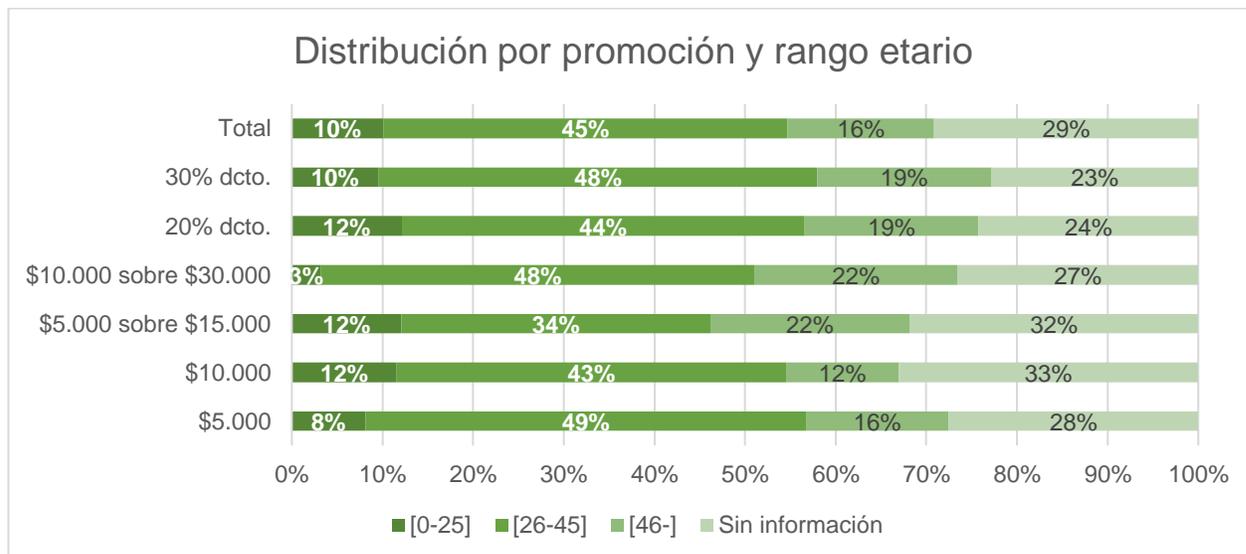
Se realiza un análisis descriptivo respecto a los clientes que se activaron, en términos de variables sociodemográficas y otras relevantes para el negocio.

5.6.7.1. POR RANGO ETARIO

Del Gráfico 25, se puede observar que los clientes que más se activaron se encontraban en el rango etario [26-45], independiente del tipo de promoción que recibieron. Lo anterior, representa, en cierta medida, lo que se observa en la tienda por departamentos, pues, durante 2017, el 43% de los clientes que compraban pertenecía a dicho rango etario.

Se evidencia que los clientes mayores a 46 años fueron activados, en mayor medida, por las promociones de montos de descuentos sujetos a ticket mínimo, mientras que aquellos entre 26 y 45 años se activaron por las promociones agresivas (\$10.000, \$10.000 sobre \$30.000 y 30% de descuento).

Gráfico 25: Distribución de clientes que canjearon, por rango etario y tipo de promoción.

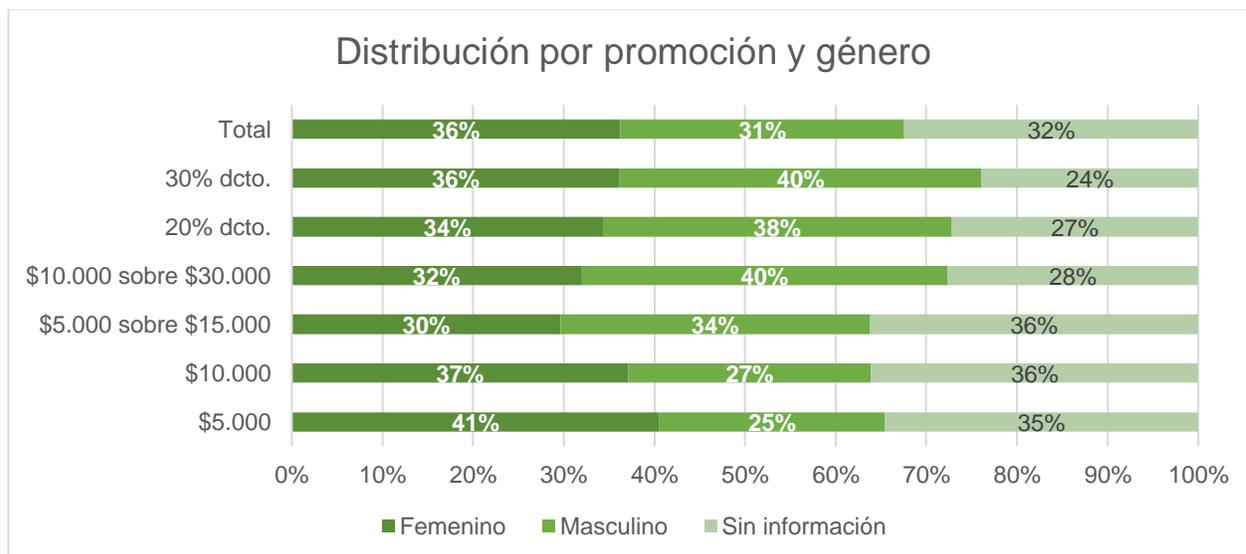


Fuente: Elaboración propia.

5.6.7.2. POR GÉNERO

Del Gráfico 26, se evidencia que las promociones que otorgan un monto de regalo fueron canjeadas por más mujeres, mientras que las que hablaban de porcentaje de descuento y de monto de regalo sujeto a una compra mínima fueron canjeados, en su mayoría, por hombres.

Gráfico 26: Distribución de clientes que canjearon, por género y tipo de promoción.



Fuente: Elaboración propia.

5.6.7.3. POR ZONA DE RESIDENCIA

Del Gráfico 27, se observa que los clientes que se activaron por las promociones provienen, en su mayoría, de la zona poniente de la Región Metropolitana. Adicionalmente, se evidencia que los clientes que viven en la zona oriente de la Región Metropolitana se activan, en mayor medida, por las promociones que proporcionan porcentaje de descuento que las de montos de descuento, lo que está relacionado con la posibilidad de comprar más y tener un ticket promedio más alto.

Gráfico 27: Distribución de clientes que canjearon, por zona de residencia y tipo de promoción.

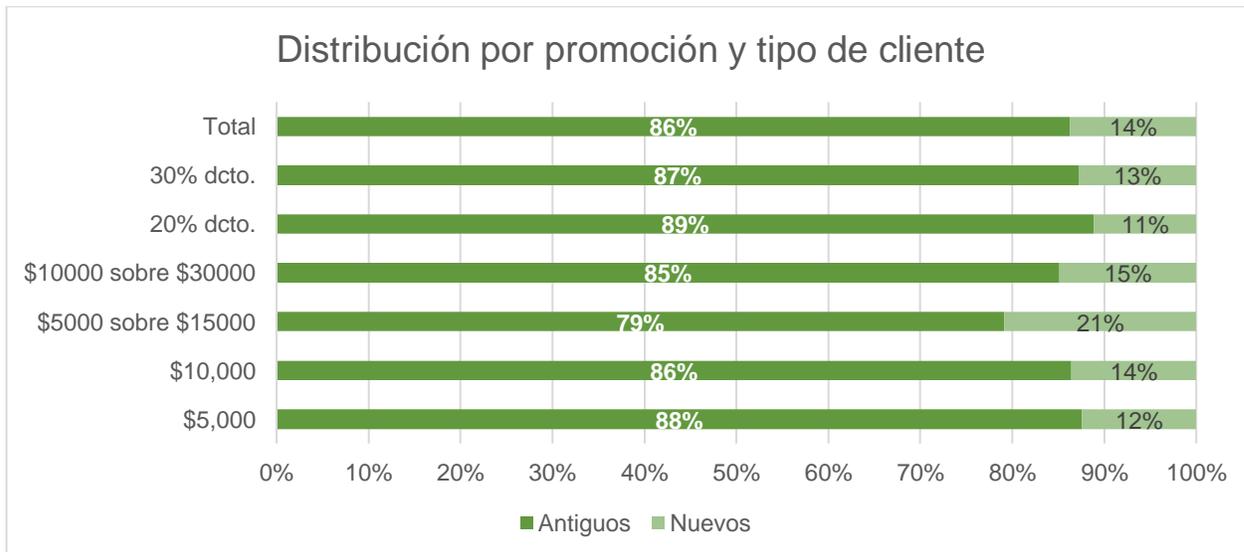


Fuente: Elaboración propia.

5.6.7.4. POR TIPO DE CLIENTE

Del Gráfico 28, se evidencia que el canje de las promociones fue realizado por clientes antiguos, más que los clientes nuevos. En particular, estos últimos, se activaron, mediante las promociones que entregan montos de descuento sujeto a la compra de un ticket mínimo.

Gráfico 28: Distribución de clientes que canjearon, por tipo de cliente y tipo de promoción.



Fuente: Elaboración propia.

5.6.7.5. POR TENENCIA DE LA TARJETA DE CRÉDITO

Del Gráfico 29, se observa que, en general, las promociones permitieron activaron clientes que tuvieron la tarjeta de crédito del holding, pero que actualmente se encuentra cerrada por morosidad.

Adicionalmente, se percibe que los clientes que tienen la cuenta abierta de la tarjeta de crédito fueron activados, en su mayoría, por las promociones asociadas a porcentajes de descuento, junto con la promoción de \$10.000 sobre \$30.000.

Por otro lado, aquellos clientes que no se encuentran relacionados con la tarjeta de crédito, fueron activados por promociones que hablan de montos de regalo y descuento (\$5.000 sobre \$15.000).

Gráfico 29: Distribución de clientes que canjearon, por estado de tarjeta de crédito y tipo de promoción.



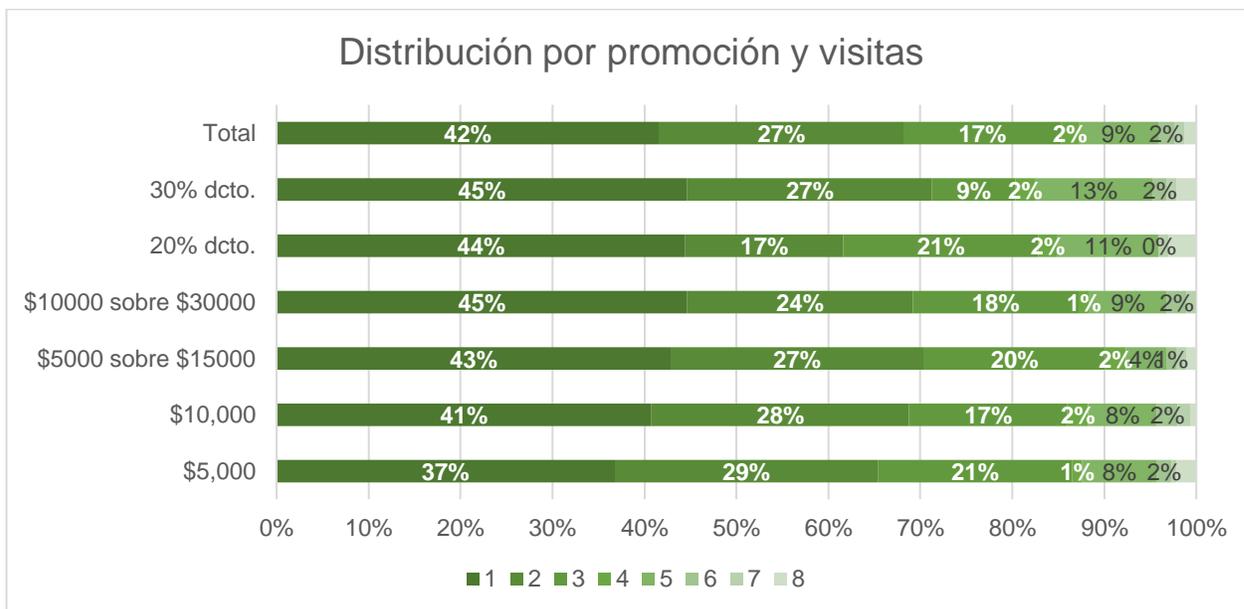
Fuente: Elaboración propia.

5.6.7.6. POR NIVEL DE VISITAS

Del Gráfico 30, se observa que, para todas las promociones enviadas, entre el 60 y 70% de los clientes que se activaron, tenían entre 1 a 2 visitas antes del envío de la campaña.

Por otro lado, se evidencia que los clientes con una visita se activaron, en menor medida, producto de las promociones con montos de descuento, donde los clientes con dos visitas tuvieron una mayor activación respecto a la población total.

Gráfico 30: Distribución de clientes que canjearon, por nivel de visitas y tipo de promoción.



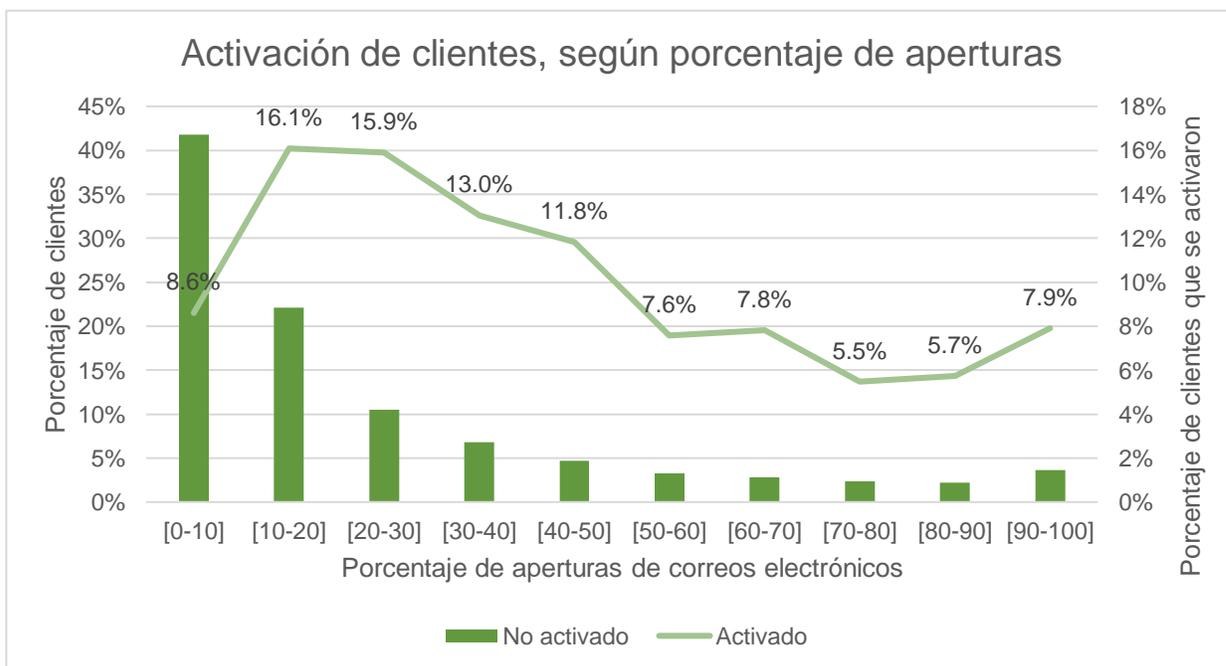
Fuente: Elaboración propia.

5.6.7.7. POR NIVEL DE APERTURAS

Del Gráfico 31, se puede evidenciar que el 50% de los clientes se activaron, tenía una apertura promedio de correos electrónicos que fluctúa entre 0 y 40%.

Lo anterior nos muestra que, a pesar de que se contactaron clientes que tienen distintos porcentajes de aperturas históricos, fueron aquellos que tenían una apertura promedio entre 10 y 30%, los que más se activaron cuando recibieron la promoción entregada.

Gráfico 31: Distribución de canjes por nivel de aperturas.



Fuente: Elaboración propia.

5.6.1. TEST DE HIPÓTESIS

En la sección 4.6.2 Hipótesis, se mencionaron diversas hipótesis relevantes a testear, enfocadas principalmente en el nivel de visitas o respuesta de distintos tipos de clientes, diferenciados por la propensión de fuga asociada a cada uno de ellos. A continuación, se prueban dichas hipótesis, mediante un test t para igualdad de medias.

5.6.1.1. HIPÓTESIS 1

Se busca probar que las *“Promociones tienen un impacto positivo en el nivel de visitas de clientes antiguos y nuevos, independiente de la propensión”*.

Para testear la hipótesis se segmentaron a los clientes por su antigüedad (nuevos y antiguos), comparándolos a nivel de visitas obtenidas en el período de vigencia de la promoción.

De la Tabla 23, se observa que los clientes nuevos presentaron un 3.7% (0.026 vs 0.025) más de visitas que el grupo control, pero no es un resultado estadísticamente significativo.

Por otro lado, de la Tabla 24, se evidencia que los clientes antiguos en el grupo tratamiento alcanzaron una media de visitas superior al grupo control (0.058 vs 0.054, respectivamente), es decir, un 7.2% más de visitas que el grupo control, siendo significativo con un 98% de confianza.

Tabla 23: Test de medias para clientes nuevos.

Variable	Grupo	Cantidad	Media	Significancia
Visitas	GM	50.747	0.026	0.699
	GC	5.570	0.025	

Fuente: Elaboración propia.

Tabla 24: Test de medias para clientes antiguos.

Variable	Grupo	Cantidad	Media	Significancia
Visitas	GM	261.375	0.058	0.012
	GC	29.147	0.054	

Fuente: Elaboración propia.

De esta manera, se rechaza la hipótesis planteada, puesto que no se cumple lo propuesto cabalmente, sino que sólo es posible afirmar que, para los clientes antiguos, el nivel de visitas aumenta al recibir una promoción o incentivo.

5.6.1.2. HIPÓTESIS 2

En esta sección, se busca probar que las *“Promociones conservadoras tienen mejor respuesta en clientes antiguos que en clientes nuevos”*.

Para realizar el test apropiadamente, se debe mencionar que una promoción conservadora se considera como tal si no proporciona un incentivo monetario elevado para el cliente. De esta manera, las promociones conservadoras son: \$5.000 de dcto. por compras sobre \$15.000, 20% de dcto. y \$5.000 de dcto.

De lo anterior, se segmentaron a los clientes por su antigüedad y se seleccionaron aquellos que recibieron una promoción conservadora y se realiza el test de hipótesis comparándolo con el grupo control.

De la Tabla 25, se afirma que, para los clientes nuevos, tanto el grupo tratamiento como el grupo control presentaron la misma media, pero no es un resultado estadísticamente significativo.

En cuanto a los clientes antiguos, de la Tabla 26, se observa que el grupo tratamiento presenta una mayor media de visitas (0.057) que el grupo control (0.054), siendo esta diferencia de un 4%, no estadísticamente significativa con una confianza de 95%.

Tabla 25: Test de hipótesis para clientes nuevos que recibieron una promoción conservadora.

Variable	Grupo	Cantidad	Media	Significancia
Visitas	GM	25.432	0.025	0.828
	GC	5.570	0.025	

Fuente: Elaboración propia.

Tabla 26: Test de hipótesis para clientes antiguos que recibieron una promoción conservadora.

Variable	Grupo	Cantidad	Media	Significancia
Visitas	GM	130.656	0.057	0.148
	GC	29.147	0.054	

Fuente: Elaboración propia.

De lo anterior, se desprende que la hipótesis de que las promociones conservadoras tienen mejor respuesta en los clientes antiguos que en los nuevos, se rechaza, pues no se registró un efecto significativo al 95% de confianza.

5.6.1.3. HIPÓTESIS 3

La hipótesis que se busca probar corresponde a que las *“Promociones agresivas tienen mayor impacto que una conservadora tanto en clientes antiguos como nuevos”*.

A diferencia de lo anterior, se definen promociones agresivas como aquellas que proporcionan un incentivo monetario superior por una compra realizada. De esta manera, son catalogadas como tal, las promociones de \$10.000 de descuento por compras sobre \$30.000, 30% de descuento y \$10.000 de descuento.

En la misma línea, se segmentaron a los clientes por su antigüedad y se seleccionaron aquellos que recibieron una promoción agresiva.

De la Tabla 27, se percibe que los clientes nuevos del grupo tratamiento tuvieron un 9% más de visitas que el grupo control, no siendo este resultado significativo al 95%, sino que al 67%.

Respecto a los clientes antiguos, se observa, en la Tabla 28, que el grupo tratamiento logró un 10% más de visitas que su respectivo grupo control, significativo con un 99% de confianza.

Tabla 27: Test de hipótesis para clientes nuevos que recibieron una promoción agresiva.

Variable	Grupo	Cantidad	Media	Significancia
Visitas	GM	25.315	0.028	0.326
	GC	5.570	0.025	

Fuente: Elaboración propia

Tabla 28: Test de hipótesis para clientes antiguos que recibieron una promoción agresiva.

Variable	Grupo	Cantidad	Media	Significancia
Visitas	GM	130.719	0.060	0.001
	GC	29.147	0.054	

Fuente: Elaboración propia

Por otro lado, al comparar cada uno de los segmentos y considerando el resultado obtenido en 5.6.1.2 Hipótesis 2, se observa que, en términos de media, los clientes que recibieron promociones agresivas alcanzaron un mayor nivel de visitas. No obstante, se rechaza la hipótesis de que las promociones agresivas tienen mayor impacto tanto en clientes nuevos como antiguos, puesto que no se encontraron efectos significativos para ambos segmentos planteados, sino que sólo se obtuvo un resultado estadísticamente significativo para los clientes antiguos que recibieron promociones agresivas.

5.6.1.4. HIPÓTESIS 4

Se busca probar que las *“Promociones tienen mejor respuesta en clientes nuevos con propensión alta que no han comprado hace 6 o más meses”*.

Para realizar el test, se seleccionaron aquellos clientes nuevos con propensión alta (dentro de los dos deciles de mayor probabilidad) que, por un lado, habían comprado hace menos de 6 meses y por otro, que habían comprado hace más de 6 meses.

De la Tabla 29, se evidencia que los clientes nuevos que compraron hace menos de 6 meses y recibieron un incentivo, lograron un 16% más de visitas, pero no obteniendo un resultado estadísticamente significativo.

No obstante, de la Tabla 30, se observa que los clientes nuevos que compraron hace más de 6 meses, incrementaron su nivel de visitas en un 52%, siendo este resultado estadísticamente significativo con ~92% de confianza.

Tabla 29: Test de hipótesis para clientes nuevos que compraron hace menos de 6 meses.

Variable	Grupo	Cantidad	Media	Significancia
Visitas	GM	2.921	0.023	0.714
	GC	303	0.020	

Fuente: Elaboración propia.

Tabla 30: Test de hipótesis para clientes nuevos que compraron hace 6 o más meses.

Variable	Grupo	Cantidad	Media	Significancia
Visitas	GM	8.329	0.022	0.084
	GC	901	0.014	

Fuente: Elaboración propia.

Por lo anterior, se acepta la hipótesis de que las promociones o incentivos tienen una mejor respuesta en clientes nuevos con propensión alta con un ~92% de confianza.

5.6.1.5. HIPÓTESIS 5

Se desea evaluar la hipótesis de que las *“Promociones tienen mejor respuesta en clientes antiguos con una propensión alta que no han comprado hace 7 o más meses”*.

Para probar la hipótesis planteada, se seleccionaron aquellos clientes antiguos que tenían una propensión alta (dentro de los dos deciles de mayor probabilidad de fuga) y que no habían comprado hace 7 o más meses.

De la Tabla 31, se evidencia que los clientes antiguos que compraron hace menos de 7 meses y recibieron una promoción, aumentaron su nivel de visitas en un 6% con respecto al grupo control, pero no de manera significativa.

Por otro lado, de la Tabla 32, se observa que los clientes antiguos que compraron hace más de 7 meses, alcanzaron un efecto menor que el del grupo control, pero al igual que el caso anterior, no es estadísticamente significativo.

Tabla 31: Test de hipótesis para clientes antiguos que compraron hace menos de 7 meses.

Variable	Grupo	Cantidad	Media	Significancia
Visitas	GM	41.357	0.045	0.483
	GC	4.675	0.043	

Fuente: Elaboración propia.

Tabla 32: Test de hipótesis para clientes antiguos que compraron hace más de 7 meses.

Variable	Grupo	Cantidad	Media	Significancia
Visitas	GM	40.036	0.031	0.663
	GC	4.616	0.032	

Fuente: Elaboración propia.

Por lo tanto, no se puede aceptar la hipótesis de que las promociones tienen un mejor efecto en los clientes antiguos con propensión alta que no han comprado hace 7 o más meses, puesto que los resultados no alcanzaron un nivel de significancia aceptable, más aún, el grupo tratamiento logró un medias más bajas que el grupo control.

5.6.1.6. HIPÓTESIS 6

Se busca probar que las *“Promociones tienen un impacto significativo en el nivel de visitas en clientes nuevos retornados con propensión alta”*.

Se define un cliente nuevo retornado como aquel consumidor clasificado como nuevo, pero que posee, al menos, una compra histórica, fuera del horizonte de evaluación que considera la definición por tipo de cliente. En particular, se define retornado a un cliente, si tiene una o más compras dentro de un período de cuatro años, anteriores a la definición por tipo de cliente.

De esta manera, en la Tabla 33, se observa que los clientes nuevos retornados con propensión alta, que recibieron una promoción aumentaron su nivel de visitas en un 110% más que el grupo control, siendo el resultado significativo con una confianza del ~84%.

Tabla 33: Test de hipótesis para clientes nuevos retornados con propensión alta

Variable	Grupo	Cantidad	Media	Significancia
Visitas	GM	2.429	0.017	0.162
	GC	243	0.008	

Fuente: Elaboración propia.

De lo anterior, se desprende que no es posible aceptar la hipótesis de que las promociones tienen un mejor efecto en los clientes nuevos retornados con propensión alta, pues, a pesar de que se evidencia una media más alta, no es estadísticamente significativo.

5.6.1.7. HIPÓTESIS 7

Se desea probar la hipótesis de que las *“Promociones tienen un impacto significativo en el nivel de visitas en clientes antiguos retornados con propensión alta”*.

Para probar lo señalado, se tiene un caso similar al anterior, donde la diferencia recae en que en esta hipótesis no se consideran los clientes nuevos, sino que los antiguos, manteniendo la misma definición de retornado y de propensión alta.

Así es como, en la Tabla 34, se evidencia que los clientes antiguos retornados tienen un 6% más de visitas que el grupo control, pero no es un resultado significativo.

Tabla 34: Test de hipótesis para clientes antiguos retornados con propensión alta.

Variable	Grupo	Cantidad	Media	Significancia
Visitas	GM	51.842	0.041	0.391
	GC	5.898	0.038	

Fuente: Elaboración propia.

Se concluye que la hipótesis de que las promociones tienen mejor impacto en el nivel de visitas de clientes antiguos retornados con propensión alta generan un aumento en el promedio de visitas, pero no se logra un resultado estadísticamente significativo, por lo que la hipótesis debe ser rechazada.

Finalmente, en la Tabla 35, se muestra un resumen de las hipótesis y el resultado de los test de hipótesis realizados para cada una de ellas, con el objetivo de evaluar si la hipótesis nula se puede o no rechazar.

Tabla 35: Resumen de hipótesis y su resultado.

Hipótesis	Hipótesis nula
H1. <i>Promociones tienen un impacto positivo en el nivel de visitas de clientes antiguos y nuevos, independiente de la propensión</i>	No se puede rechazar
H2. <i>Promociones conservadoras tienen mejor respuesta en clientes antiguos que en clientes nuevos</i>	No se puede rechazar
H3. <i>Promociones agresivas tienen mayor impacto que una conservadora tanto en clientes antiguos como nuevos</i>	No se puede rechazar
H4. <i>Promociones tienen mejor respuesta en clientes nuevos con propensión alta que no han comprado hace 6 o más meses</i>	No se puede rechazar
H5. <i>Promociones tienen mejor respuesta en clientes antiguos con una propensión alta que no han comprado hace 7 o más meses</i>	No se puede rechazar
H6. <i>Promociones tienen un impacto significativo en el nivel de visitas en clientes nuevos retornados con propensión alta</i>	No se puede rechazar
H7. <i>Promociones tienen un impacto significativo en el nivel de visitas en clientes antiguos retornados con propensión alta</i>	No se puede rechazar

Fuente: Elaboración propia.

5.6.2. ANÁLISIS DE ASISTENCIA INCREMENTAL

Uno de los principales objetivos del negocio responde a que sus clientes mantengan o aumenten su nivel de visitas durante un año. De esta manera, se busca evaluar el impacto que tuvo la promoción realizada, focalizándose en medir la asistencia incremental que presentó el grupo tratamiento respecto al grupo control respectivo.

Por consiguiente, se evalúa el impacto de cada una de las promociones, de acuerdo a distintas variables, entre ellas, la tenencia de la tarjeta de crédito, la antigüedad del cliente y el período desde su última compra o *recency*.

5.6.2.1. TENENCIA DE LA TARJETA DE CRÉDITO

Para este caso, se calcula la asistencia incremental que tuvieron los clientes en el período en promoción, según la propensión (alta, media y baja) y la clasificación que tienen en la

tenencia de la tarjeta de crédito, es decir, si tienen la cuenta abierta, cerrada o no tienen una relación contractual con la emisora de tarjetas de crédito del holding.

De esta manera, en la Tabla 36, se puede evidenciar que no fueron contactados clientes con tarjeta de crédito abierta y con propensión alta. De ello y como se menciona anteriormente, un cliente con la tarjeta de crédito del holding, tiene menos probabilidad de fugarse de la tienda por departamento.

Tabla 36: Asistencia incremental para clientes con tarjeta de crédito abierta

Abierta	Alta	Media	Baja
\$5.000 de dcto.	-	1,4%**	0,4%
\$10.000 de dcto.	-	0,8%	1,2%
20% de dcto.	-	0,4%	-0,7%
30% de dcto.	-	-0,7%	1,1%
\$5.000 de dcto. sobre \$15.000	-	0,3%	0,6%
\$10.000 de dcto. sobre \$30.000	-	-0,6%	1,8%

* p < 0.1, ** p < 0.05, *** p < 0.01

Fuente: Elaboración propia.

De la Tabla 36, se evidencia que:

1. Dado que no existían clientes con propensión alta y con tarjeta de crédito abierta o activa, no se generan recomendaciones para este segmento.
2. Para los clientes con propensión media, se obtiene que la promoción de \$5.000 de descuento genera una asistencia incremental de 1.4%, significativa al 95%.
3. Para el segmento de clientes con tarjeta activa y propensión baja, no se obtuvieron promociones significativas.

De lo anterior, se concluye que si se desea aumentar el nivel de visitas para los clientes con tarjeta abierta, se recomienda realizarlo sólo en consumidores con propensión media, utilizando la promoción de \$5.000 de descuento.

Con respecto a clientes que no tienen la tarjeta de crédito activa, se observa, de la Tabla 37, que existen promociones las cuales no son apropiadas de enviar para algunos segmentos de clientes, los que se detallan a continuación.

Tabla 37: Asistencia incremental para clientes con tarjeta de crédito cerrada.

Cerrada	Alta	Media	Baja
\$5.000 de dcto.	-0,5%	0,5%	0,2%
\$10.000 de dcto.	-0,3%	0,9%*	1,1%**
20% de dcto.	-1,3%***	0,4%	0,3%
30% de dcto.	-0,6%*	1,2%**	0,2%
\$5.000 de dcto. sobre \$15.000	-0,8%**	1,9%***	0,7%
\$10.000 de dcto. sobre \$30.000	-1,2%*	0,2%	0,1%

* p < 0.1, ** p < 0.05, *** p < 0.01

Fuente: Elaboración propia.

De la Tabla 37, se muestra que:

1. Para los clientes con propensión alta, tanto las promociones de porcentaje de descuento como las de montos de descuentos con restricción, son significativas y generan asistencia incremental inferior que el grupo control. De esta manera, no es recomendable enviar dichos tipos de promociones a este segmento de clientes con tarjeta cerrada.
2. Para los clientes con propensión media, se reconoce que las promociones que generan asistencia incremental, en orden ascendente de confianza son:
 - a. \$10.000 de descuento.
 - b. 30% de descuento.
 - c. \$5.000 de descuento sobre \$15.000.
3. Para los clientes con propensión baja, se recomienda enviar la promoción de \$10.000, siendo estadísticamente significativa al 95%.

De lo anterior, se desprende que para los clientes con tarjeta cerrada, no es recomendable contactar a aquellos con propensión alta, sino que es mejor realizarlo cuando se encuentran en una etapa menos riesgosa de fuga, la que se alcanza cuando tienen propensión media, donde se podría incentivar una visita a la tienda con una promoción agresiva (\$10.000 y 30% de descuento) y conservadora (\$5.000 de descuento sobre \$15.000).

Finalmente, los clientes sin relacionamiento con el holding, son aquellos que no poseen la tarjeta de crédito, pero que aún así realizan compras en la tienda por departamento con otros medios de pago.

Tabla 38: Asistencia incremental para clientes sin tarjeta de crédito.

Sin tarjeta	Alta	Media	Baja
\$5.000 de dcto.	0,3%	0,7%	-2,2%**
\$10.000 de dcto.	0,8%	1,7%***	1,0%
20% de dcto.	0,3%	1,1%*	-1,1%
30% de dcto.	-0,8%	0,4%	-1,1%
\$5.000 de dcto. sobre \$15.000	-0,6%	1,0%*	-0,7%
\$10.000 de dcto. sobre \$30.000	-0,7%	0,9%*	0,7%

* p < 0.1, ** p < 0.05, *** p < 0.01

Fuente: Elaboración propia.

De la Tabla 38, se concluye que:

1. Los clientes con propensión alta, no debiesen recibir promociones, debido a que los efectos que generan, no son significativos estadísticamente.
2. Para los clientes con propensión media, se evidencia que las promociones que generan asistencia incremental, en orden de confianza son:
 - a. \$10.000 de descuento sobre \$30.000.
 - b. \$5.000 de descuento sobre \$15.000.
 - c. 20% de descuento.
 - d. \$10.000 de descuento.
3. Para los clientes con propensión baja, sólo una promoción es significativa y corresponde a la que proporciona \$5.000 de descuento. No obstante, el efecto que genera es negativo, respecto al grupo control.

Como se menciona anteriormente, para este caso, se evidencia que los clientes, no deberían ser contactados cuando alcanzan un nivel de propensión alto (dos deciles con mayor probabilidad de fuga), sino que la acción de marketing directo debería ocurrir en etapas previas, es decir, cuando el cliente tiene una propensión media, cuyo caso se obtienen cuatro de seis promociones estadísticamente significativas que generan efectos positivos.

5.6.2.2. ANTIGÜEDAD DEL CLIENTE

En esta sección, se evalúa la asistencia incremental obtenida, según la antigüedad del cliente; antiguos y nuevos, diferenciados por la propensión que tenían al momento previo del envío de la promoción.

Tabla 39: Asistencia incremental para clientes nuevos.

Nuevos	Alta	Media	Baja
\$5.000 de dcto.	-1,4%*	-0,5%	-0,2%
\$10.000 de dcto.	-1,5%**	-0,1%	0,6%
20% de dcto.	-1,6%**	-0,2%	-0,3%
30% de dcto.	-1,4%*	-0,2%	-0,3%
\$5.000 de dcto. sobre \$15.000	-0,8%	0,1%	-0,5%
\$10.000 de dcto. sobre \$30.000	-1,6%**	-0,4%	0,0%

* p < 0.1, ** p < 0.05, *** p < 0.01

Fuente: Elaboración propia.

De la Tabla 39, se observa que:

1. Los clientes con propensión alta obtuvieron promociones significativas, pero en todas ellas tienen efectos negativos. Adicionalmente, la promoción que no fue significativa, también posee un efecto negativo. Por lo que no es recomendable enviar promociones a este segmento de clientes.
2. Para los clientes con propensión media, no se obtuvieron promociones significativas.
3. Para los clientes con propensión baja, no se obtuvieron promociones significativas.

De lo anterior, se concluye que los clientes nuevos no debiesen recibir promociones, pues, en el nivel de propensión alto, los incentivos generan menos asistencia incremental que el grupo control. El caso es distinto para los clientes con propensión media y baja, donde ninguna promoción genera efectos significativos.

Tabla 40: Asistencia incremental para clientes antiguos.

Antiguos	Alta	Media	Baja
\$5.000 de dcto.	0,0%	1,0%***	0,1%
\$10.000 de dcto.	0,3%	1,2%***	1,3%**
20% de dcto.	-0,7%*	0,6%**	-0,1%
30% de dcto.	-0,5%	0,5%*	0,4%
\$5.000 de dcto. sobre \$15.000	-0,7%**	1,3%***	0,9%*
\$10.000 de dcto. sobre \$30.000	-0,9%**	0,2%	0,7%

* p < 0.1, ** p < 0.05, *** p < 0.01

Fuente: Elaboración propia.

De la Tabla 40, se muestra que:

1. Los clientes con propensión alta tuvieron promociones significativas (20% de descuento, \$5.000 de descuento sobre \$15.000 y \$10.000 de descuento sobre \$30.000), pero ninguna obtuvo efectos positivos. Por otro lado, las promociones de montos de descuentos, si bien no son significativas, lograron efectos positivos.
2. Los clientes con propensión media, obtuvieron cinco de seis promociones significativas, las cuales se ordenan, según el nivel de confianza obtenido.
 - a. 30% de descuento.
 - b. 20% de descuento.
 - c. \$5.000 de descuento.
 - d. \$10.000 de descuento.
 - e. \$5.000 de descuento sobre \$15.000.
3. Los clientes con propensión baja, tuvieron asistencia incremental significativa en dos promociones, las cuales se ordenan, según el nivel de confianza obtenido.
 - a. \$5.000 de descuento sobre \$15.000.
 - b. \$10.000 de descuento.

De lo anterior, se concluye que para los clientes antiguos existen promociones que contribuyen a la asistencia incremental para consumidores con propensión media y baja, dentro de las cuales se encuentran las que hablan de porcentajes de descuentos (20% y 30% de descuento), montos de descuento (\$5.000 y \$10.000 de descuento) y montos de descuento sujeto a restricciones (\$5.000 de descuento sobre \$15.000).

Cabe mencionar que, tres de las seis promociones enviadas fueron significativas para los clientes antiguos con propensión alta, pero lograron efectos negativos, respecto al grupo control. Por lo que, dichas promociones no son recomendables de enviar a dicho segmento de consumidores.

5.6.2.3. REGENCY

Para esta sección, se calcula la asistencia incremental que tuvieron los clientes, según el período desde su última compra o *recency*, clasificado por 1 a 3 meses, 4 a 6 meses, 7 a 9 meses y 10 a 12 meses.¹⁰

Tabla 41: Asistencia incremental para clientes que no han comprado entre 1 a 3 meses.

[1-3]	Alta	Media	Baja
\$5.000 de dcto.	-1,8%**	1,8%***	0,7%
\$10.000 de dcto.	-1,2%*	2,9%***	0,6%
20% de dcto.	-2,6%***	1,9%***	0,5%
30% de dcto.	-2,0%**	1,5%***	1,0%*
\$5.000 de dcto. sobre \$15.000	-2,6%***	2,5%***	1,3%*
\$10.000 de dcto. sobre \$30.000	-2,7%***	1,4%**	0,5%

* p < 0.1, ** p < 0.05, *** p < 0.01

Fuente: Elaboración propia.

De la Tabla 41, se puede observar que:

¹⁰ El peso de los segmentos sobre el total de clientes corresponde a 34%, 28%, 20% y 18%, respectivamente.

1. Para los clientes con propensión alta, todas las promociones fueron estadísticamente significativas, pero con efectos negativos en todas sus variaciones. Por lo que para el segmento en estudio que tiene compras hace menos de 3 meses, no se recomienda enviar promociones.
2. Para los clientes con propensión media, todas las promociones fueron significativas y además, tuvieron efectos positivos. De ello, al ordenarlas de menor a mayor, según la asistencia incremental, se obtiene:
 - a. \$10.000 de descuento sobre \$30.000.
 - b. 30% de descuento.
 - c. \$5.000 de descuento.
 - d. 20% de descuento.
 - e. \$5.000 de descuento sobre \$15.000.
 - f. \$10.000 de descuento.
3. Para los clientes con propensión baja, sólo dos promociones fueron significativas, las cuales se ordenan, según la asistencia incremental que generan.
 - a. 30% de descuento.
 - b. \$5.000 de descuento sobre \$15.000.

De lo anterior, se concluye que no es recomendable contactar a clientes con compras hace menos de tres meses y con propensión alta, pues las promociones generan efectos inferiores que el grupo control. Asimismo, los clientes con propensión media son clientes que obtuvieron asistencia incremental positiva significativa, en todas las promociones que se enviaron.

Para este caso, se evidencia, nuevamente, un mejor desempeño de las promociones en clientes con propensión media y no alta. Esto, se puede explicar debido a que cuando un consumidor alcanza un nivel de propensión alto, es más costoso para la organización incentivar una compra porque es un cliente que ya tiene nulo relacionamiento con la tienda por departamento.

Tabla 42: Asistencia incremental para clientes que no han comprado entre 4 a 6 meses.

[4-6]	Alta	Media	Baja
\$5.000 de dcto.	0,0%	0,0%	-1,1%*
\$10.000 de dcto.	0,6%	0,3%	1,8%**
20% de dcto.	-0,1%	-0,1%	0,1%
30% de dcto.	-1,7%**	0,2%	-0,2%
\$5.000 de dcto. sobre \$15.000	-0,2%	0,1%	-0,5%
\$10.000 de dcto. sobre \$30.000	-1,1%*	-1,0%*	0,9%

* p < 0.1, ** p < 0.05, *** p < 0.01

Fuente: Elaboración propia.

De la Tabla 42, se evidencia que:

1. Para los clientes con propensión alta, sólo dos promociones son significativas, pero generan una menor asistencia incremental. Por lo que, no es recomendable enviar las promociones de 30% de descuento y \$10.000 de descuento sobre \$30.000 para este segmento de clientes.
2. Los clientes con propensión media, sólo obtuvieron una promoción significativa que, además, no genera asistencia incremental positiva. De esta forma, no se

recomienda enviar la promoción de \$10.000 de descuento sobre \$30.000 para este segmento de clientes.

3. Para los clientes con propensión baja, se obtuvieron una promoción significativa positiva y otra negativa. Por consiguiente, no se recomienda enviar la promoción de \$5.000 de descuento, pero sí el incentivo de \$10.000 de descuento.

De lo anterior, se concluye que el segmento de clientes que no han comprado hace cuatro y seis meses, no es relevante de estudiar, debido a que las promociones no generan un efecto significativo mayor, en asistencia incremental, para los consumidores con propensión alta y media, no así para baja, donde el ticket de \$10.000 de descuento es capaz de generar asistencia incremental positiva y significativa.

Tabla 43: Asistencia incremental para clientes que no han comprado entre 7 a 9 meses.

[7-9]	Alta	Media	Baja
\$5.000 de dcto.	0,2%	1,2%*	0,3%
\$10.000 de dcto.	0,0%	1,1%*	1,8%*
20% de dcto.	-0,7%	0,5%	-1,0%
30% de dcto.	0,9%*	0,2%	-1,2%
\$5.000 de dcto. sobre \$15.000	-1,7%***	1,0%*	1,7%*
\$10.000 de dcto. sobre \$30.000	-0,1%	0,3%	0,3%

* p < 0.1, ** p < 0.05, *** p < 0.01

Fuente: Elaboración propia.

De la Tabla 43, se muestra que:

1. Los clientes con propensión alta, alcanzaron dos promociones significativas, una de ellas genera un efecto positivo, mientras que la otra es negativo. De esta manera, es recomendable enviar la promoción de 30% de descuento, pero no se debería ofrecer la promoción de \$5.000 de descuento sobre \$15.000 a este segmento de clientes.
2. Los clientes con propensión media, obtuvieron tres promociones significativas, las cuales se ordenan, según su asistencia incremental, obteniendo:
 - a. \$5.000 de descuento sobre \$15.000.
 - b. \$10.000 de descuento.
 - c. \$5.000 de descuento.
3. Los clientes con propensión baja, obtuvieron dos promociones significativas, las cuales se ordenan, según su asistencia incremental, obteniendo:
 - a. \$5.000 de descuento sobre \$15.000.
 - b. \$10.000 de descuento.

De lo anterior, se concluye que los clientes que no han comprado hace siete o nueve meses, deberían recibir promociones que hablan de montos de descuentos libres o sujetos a una restricción para aquellos con propensión media y baja, mientras que para los clientes con propensión alta, se recomienda enviar un incentivo asociado a un porcentaje de descuento y no de un monto con restricción porque, al ser clientes mayor propensión de fuga y con una compra registrada hace más de siete meses, es posible que perciban la restricción de los \$15.000 como una barrera muy alta para utilizar el descuento, lo que los incentiva a no usarlo.

Tabla 44: Asistencia incremental para clientes que no han comprado entre 10 a 12 meses.

[10-12]	Alta	Media	Baja
\$5.000 de dcto.	0,0%	-0,7%	-1,6%*
\$10.000 de dcto.	0,4%	-1,5%**	0,1%
20% de dcto.	-0,1%	-1,0%	-2,2%**
30% de dcto.	-0,1%	-1,2%*	-0,7%
\$5.000 de dcto. sobre \$15.000	1,0%*	0,1%	-1,3%
\$10.000 de dcto. sobre \$30.000	-0,7%	-1,0%	0,8%

* p < 0.1, ** p < 0.05, *** p < 0.01

Fuente: Elaboración propia.

De la Tabla 44, se manifiesta que:

1. Los clientes con propensión alta obtuvieron sólo una promoción significativa y con efecto positivo que corresponde a la de \$5.000 de descuento sobre \$15.000.
2. Para los clientes con propensión media, se obtuvieron promociones significativas, pero cuyo efecto no genera mayor asistencia incremental. En este sentido, no se recomienda enviar las promociones de \$10.000 de descuento ni 30% de descuento.
3. Al igual que en el caso anterior, para los clientes con propensión baja, se obtuvieron promociones significativas, pero que tienen un efecto menor, en términos de asistencia incremental. De este modo, no se recomienda enviar promociones de \$5.000 de descuento ni de 20% de descuento para el segmento de clientes.

De esto, se concluye que para los clientes que no han comprado hace más de 10 meses, sólo la promoción de \$5.000 de descuento sobre \$15.000 es significativa y además, genera asistencia incremental positiva. Por otro lado, para los clientes con propensión media o baja, las promociones que son significativas, no generan asistencia incremental positiva, por lo que no se recomendaría enviar este tipo de promociones al segmento de clientes.

5.6.3. ANÁLISIS DE REGRESIÓN LINEAL

Con el objetivo de complementar el análisis de asistencia incremental, se realiza un modelo de regresión lineal para determinar qué promociones, para cada uno de los segmentos, son más apropiadas de enviar, de acuerdo al efecto que cada una de ellas genera en la visita o compra de un cliente en la tienda por departamento.

En este sentido, se realiza una regresión lineal que estima el efecto en las visitas de quienes efectivamente estuvieron expuestos al estímulo generado, permitiendo evaluar la decisión que toma un cliente de usar o no el cupón en las visitas o compras que realiza en la tienda por departamento.

Además, se incorporaron variables sociodemográficas (género, edad y zona de residencia), que definen el tipo de propensión que tenía cada cliente al momento del envío (alta, media o baja) y la interacción entre si recibió o vio la promoción, teniendo una determinada propensión.

Como se menciona anteriormente, una alternativa relevante de estudiar, corresponde al funcionamiento de las promociones para los clientes que efectivamente estuvieron expuestos a cada incentivo, es decir, determinar el efecto estadístico que tuvo cada promoción en los clientes que vieron el correo electrónico y pudieron tomar la decisión de usar el cupón o no, en una compra durante el período de vigencia de la campaña.

De esta forma, y al igual que la información proporcionada en 5.6.2 Análisis de Asistencia Incremental, se realiza una regresión para cada segmento de clientes, separado por tenencia de la tarjeta de crédito, antigüedad del cliente y *recency*.

5.6.3.1. TENENCIA DE LA TARJETA DE CRÉDITO

Para esta sección, se realiza una regresión para cada categoría de tenencia de tarjeta de crédito, es decir, se tiene un modelo de regresión lineal para clientes con tarjeta abierta o activa, otra para consumidores con tarjeta cerrada y, por último, una para clientes sin contrato.

De esta forma, se procede a indicar los resultados obtenidos para la regresión de clientes con tarjeta abierta, cerrada y sin cuenta, respectivamente.

Tabla 45: Coeficientes para la regresión de clientes con tarjeta abierta.

Promoción	Alta	Media	Baja
\$5.000 de dcto.	-	0%	6%
\$10.000 de dcto.	-	1%	0%
20% de dcto.	-	-1%***	4%
30% de dcto.	-	-2%	-2%
\$5.000 de dcto. sobre \$15.000	-	0%***	-3%
\$10.000 de dcto. sobre \$30.000	-	-1%	8%***

* p < 0.1, ** p < 0.05, *** p < 0.01

Fuente: Elaboración propia

De la Tabla 45, se puede apreciar que:

1. No hay clientes con tarjeta de crédito abierta o activa que, además, posean propensión alta. Por lo anterior, no se realiza una recomendación para este segmento.
2. Para los clientes con propensión media, existen dos promociones significativas, de las cuales, generan efectos nulos (0%) o negativos (-1%), por esta razón, no se recomienda entregar incentivos, en forma de una promoción de 20% de descuento o un ticket de \$5.000 de descuento por compras sobre \$15.000.
3. Para los clientes con propensión baja, sólo la promoción de \$10.000 de descuento sobre \$30.000 fue capaz de generar un efecto positivo (8%) y significativo para el segmento de clientes.

Tabla 46: Coeficientes para la regresión de clientes con tarjeta cerrada.

Promoción	Alta	Media	Baja
\$5.000 de dcto.	10%***	14%***	31%***
\$10.000 de dcto.	5%	14%***	38%***
20% de dcto.	2%	10%***	21%***
30% de dcto.	2%	14%***	21%***
\$5.000 de dcto. sobre \$15.000	7%	13%***	35%***
\$10.000 de dcto. sobre \$30.000	4%	13%***	27%***

* p < 0.1, ** p < 0.05, *** p < 0.01

Fuente: Elaboración propia

De la Tabla 46, se evidencia que:

1. Sólo la promoción de \$5.000 de descuento es capaz de generar un efecto positivo (10%) y significativo para los clientes con propensión alta.
2. Para los clientes con propensión media, todas las promociones fueron significativas y generan un efecto positivo, que fluctúa entre 10 y 14%. De esta manera, es posible ordenar, según el beta obtenido.
 - a. 20% de descuento.
 - b. \$5.000 de descuento por sobre \$15.000.
 - c. \$10.000 de descuento por sobre \$30.000.
 - d. 30% de descuento.
 - e. \$5.000 de descuento.
 - f. \$10.000 de descuento.
3. El caso es similar para los clientes con propensión baja. No obstante, el efecto es mucho mayor, variando desde 21 a 38%. Asimismo, al ordenar las promociones, según su efecto, se tiene:
 - a. 20% de descuento.
 - b. 30% de descuento.
 - c. \$10.000 de descuento sobre \$30.000.
 - d. \$5.000 de descuento.
 - e. \$5.000 de descuento sobre \$15.000.
 - f. \$10.000 de descuento.

Tabla 47: Coeficientes para la regresión de clientes sin tarjeta de crédito del holding.

Promoción	Alta	Media	Baja
\$5.000 de dcto.	11%***	20%***	19%***
\$10.000 de dcto.	14%***	16%***	28%***
20% de dcto.	9%***	8%***	9%
30% de dcto.	9%***	17%***	26%***
\$5.000 de dcto. sobre \$15.000	10%***	19%***	12%***
\$10.000 de dcto. sobre \$30.000	16%***	20%***	21%***

* p < 0.1, ** p < 0.05, *** p < 0.01

Fuente: Elaboración propia

De la Tabla 47, se muestra que:

1. Las promociones enviadas a los clientes con propensión alta, tuvieron efectos positivos y significativos, en cada una de sus variaciones. De este modo, se ordenan, de acuerdo a su efecto (beta) alcanzado.

- a. 20% de descuento.
 - b. 30% de descuento.
 - c. \$5.000 de descuento sobre \$15.000.
 - d. \$5.000 de descuento.
 - e. \$10.000 de descuento.
 - f. \$10.000 de descuento sobre \$30.000.
2. Al igual que el caso anterior, todas las promociones tuvieron efectos positivos y significativos en el nivel de visitas de los clientes en el período promocional. De esta forma, se evidencia un orden que permite ir aumentando progresivamente el efecto en las visitas de los clientes.
 - a. 20% de descuento.
 - b. \$10.000 de descuento.
 - c. 30% de descuento.
 - d. \$5.000 de descuento sobre \$15.000.
 - e. \$5.000 de descuento.
 - f. \$10.000 de descuento sobre \$30.000.
 3. Para los clientes con propensión baja, sólo una promoción de las seis enviadas, no generó significancia estadística. No obstante, si se ordenan las promociones, según el efecto que genera en el nivel de visitas, se tiene:
 - a. \$5.000 de descuento sobre \$15.000.
 - b. \$5.000 de descuento.
 - c. \$10.000 de descuento sobre \$30.000.
 - d. 30% de descuento.
 - e. \$10.000 de descuento.

5.6.3.2. ANTIGÜEDAD DEL CLIENTE

En esta sección, se realiza una regresión por tipo de cliente; antiguos y nuevos. En la misma línea, se enunciarán los resultados obtenidos de los coeficientes de la regresión que representan el efecto que tienen las promociones para cada segmento estudiado.

Tabla 48: Coeficientes para la regresión de clientes nuevos.

Promoción	Alta	Media	Baja
\$5.000 de dcto.	8%***	5%***	6%***
\$10.000 de dcto.	8%***	8%***	4%***
20% de dcto.	1%	5%***	7%***
30% de dcto.	4%	11%***	4%***
\$5.000 de dcto. sobre \$15.000	7%***	10%***	2%
\$10.000 de dcto. sobre \$30.000	5%***	4%***	12%***

* p < 0.1, ** p < 0.05, *** p < 0.01

Fuente: Elaboración propia

De la Tabla 48, se evidencia que:

1. Para los clientes nuevos con propensión alta, las promociones de porcentajes de descuentos no generan efectos significativos. Por esto, el orden, de menor a mayor contribución al nivel de visitas, se ajusta como sigue:
 - a. \$10.000 de descuento sobre \$30.000.
 - b. \$5.000 de descuento sobre \$15.000.

- c. \$5.000 de descuento.
 - d. \$10.000 de descuento.
2. El orden de promociones, según el efecto que generan al nivel de visitas, se manifiesta de la siguiente manera:
 - a. \$10.000 de descuento sobre \$30.000.
 - b. 20% de descuento.
 - c. \$5.000 de descuento.
 - d. \$10.000 de descuento.
 - e. \$5.000 de descuento sobre \$15.000.
 - f. 30% de descuento.
 3. En cuanto a los clientes con propensión baja, el orden de promociones, quedaría como sigue:
 - a. 30% de descuento.
 - b. \$10.000 de descuento.
 - c. \$5.000 de descuento.
 - d. 20% de descuento.
 - e. \$10.000 de descuento sobre \$30.000.

Tabla 49: Coeficientes para la regresión de clientes antiguos.

Promoción	Alta	Media	Baja
\$5.000 de dcto.	9%***	10%***	-4%
\$10.000 de dcto.	10%***	10%***	6%**
20% de dcto.	9%***	6%***	-9%***
30% de dcto.	6%	9%***	0%
\$5.000 de dcto. sobre \$15.000	7%***	9%***	-6%**
\$10.000 de dcto. sobre \$30.000	9%***	6%***	-4%

* p < 0.1, ** p < 0.05, *** p < 0.01

Fuente: Elaboración propia

De la Tabla 49, se muestra que:

1. Sólo una promoción no genera efectos significativos en los clientes con propensión alta. De las restantes, se ordenan, según el efecto generado, obteniendo:
 - a. \$5.000 de descuento sobre \$15.000.
 - b. \$10.000 de descuento sobre \$30.000.
 - c. 20% de descuento.
 - d. \$5.000 de descuento.
 - e. \$10.000 de descuento.
2. Todas las promociones enviadas a los clientes con propensión media alcanzaron significancia estadística. De este modo, se ordenan las promociones, según el efecto o beta generado en el nivel de visitas.
 - a. \$10.000 de descuento sobre \$30.000.
 - b. 20% de descuento.
 - c. \$5.000 de descuento sobre \$15.000.
 - d. \$5.000 de descuento.
 - e. 30% de descuento.
 - f. \$10.000 de descuento.
3. Para los clientes con propensión baja, se observan tres de seis promociones significativas. No obstante, los incentivos de 20% de descuento y \$5.000 de descuento sobre \$15.000, no son recomendados de usar, pues generan efectos

negativos (entre -6 y -9%) en el nivel de visitas de los clientes del segmento. No así, la promoción de \$10.000 que genera un aumento de 6% en la variable estudiada.

5.6.3.3. RECENCY

En esta sección, se realiza una regresión por agrupación de *recency*, según lo indicado anteriormente. De esta manera, se estudian los efectos que tienen las promociones en el nivel de visitas para los clientes con compras entre cero y tres meses; cuatro y seis meses; siete y nueve meses; y, más de 10 meses.

Tabla 50: Coeficientes para la regresión de clientes con compras hace menos de tres meses.

Promoción	Alta	Media	Baja
\$5.000 de dcto.	9%	17%***	-1%
\$10.000 de dcto.	15%***	18%***	12%***
20% de dcto.	8%	6%	-13%***
30% de dcto.	3%	7%	10%
\$5.000 de dcto. sobre \$15.000	4%	3%	-5%
\$10.000 de dcto. sobre \$30.000	17%***	11%***	-1%

* p < 0.1, ** p < 0.05, *** p < 0.01

Fuente: Elaboración propia

De la Tabla 50, se observa que:

1. Para los clientes con propensión alta, las promociones de \$10.000 de descuento y \$10.000 de descuento sobre \$30.000, alcanzaron resultados estadísticamente significativos y con efectos positivos en el nivel de visitas de los clientes en el segmento.
2. Para los clientes con propensión media, se obtuvieron tres promociones significativas y con efectos positivos, las cuales, ordenadas, según su efecto.
 - a. \$10.000 de descuento sobre \$30.000.
 - b. \$5.000 de descuento.
 - c. \$10.000 de descuento.
3. Las promociones significativas enviadas a los clientes con propensión baja fueron 20% de descuento y \$10.000 de descuento. No obstante, la promoción que habla de porcentajes de descuentos, si bien es significativa, no es recomendable de utilizar debido a que genera efectos negativos (-13%) en el nivel de visitas del segmento en estudio.

Tabla 51: Coeficientes para la regresión de clientes con compras entre cuatro a seis meses.

Promoción	Alta	Media	Baja
\$5.000 de dcto.	6%	6%	-11%
\$10.000 de dcto.	14%***	9%	1%
20% de dcto.	6%	0%	-3%
30% de dcto.	-6%	11%***	-15%***
\$5.000 de dcto. sobre \$15.000	6%	8%	6%
\$10.000 de dcto. sobre \$30.000	12%	7%	-3%

* p < 0.1, ** p < 0.05, *** p < 0.01

Fuente: Elaboración propia

De la Tabla 51, se percibe que:

1. Es recomendable enviar sólo una promoción (\$10.000 de descuento) para los clientes con propensión alta, pues su efecto asciende a 14%, siendo el resultado estadísticamente significativo.
2. Similar es el caso para los clientes con propensión media, donde la promoción de 30% de descuento fue capaz de generar un efecto del 11%, siendo significativo al 95%.
3. Para los clientes con propensión baja, sólo una promoción es significativa (30% de descuento), pero no es recomendable enviar este incentivo al segmento de clientes, pues genera un efecto negativo (-15%) en las visitas.

Tabla 52: Coeficientes para la regresión de clientes con compras entre siete y nueve meses.

Promoción	Alta	Media	Baja
\$5.000 de dcto.	2%	11%***	7%
\$10.000 de dcto.	-2%	8%	3%
20% de dcto.	1%	0%	-4%
30% de dcto.	6%	-6%	2%
\$5.000 de dcto. sobre \$15.000	2%	8%	12%
\$10.000 de dcto. sobre \$30.000	-5%	0%	-2%

* p < 0.1, ** p < 0.05, *** p < 0.01

Fuente: Elaboración propia

De la Tabla 52, se evidencia que:

1. Si bien se evidencian efectos positivos y negativos para las distintas promociones, ninguna de ellas alcanza la significancia estadística. Por lo que no es recomendable abordar este segmento de clientes, mediante incentivos por email marketing.
2. Para los clientes con propensión media, sólo una de las seis promociones alcanza la significancia estadística y además, genera un efecto positivo (11%) en el nivel de visitas del segmento de clientes.
3. Las promociones enviadas a los clientes con propensión baja no alcanzan la significancia estadística para recomendar su envío, a pesar de que algunas de ellas tienen efectos positivos en las visitas.

Tabla 53: Coeficientes para la regresión de clientes con compras hace más de 10 meses.

Promoción	Alta	Media	Baja
\$5.000 de dcto.	-6%	-6%	0%
\$10.000 de dcto.	12%***	4%	5%
20% de dcto.	1%	-9%***	-2%
30% de dcto.	1%	3%	33%***
\$5.000 de dcto. sobre \$15.000	12%***	10%***	-7%
\$10.000 de dcto. sobre \$30.000	2%	-2%	7%

* p < 0.1, ** p < 0.05, *** p < 0.01

Fuente: Elaboración propia

De la Tabla 53, se muestra que:

1. Para los clientes con propensión alta, sólo dos promociones alcanzaron efectos positivos (~12%) estadísticamente significativas, las cuales se ordenaron, según el efecto que producían en las visitas.

- a. \$10.000 de descuento.
 - b. \$5.000 de descuento sobre \$15.000.
2. Para los clientes con propensión media, una de las promociones significativas, corresponde a la de 20% de descuento, sin embargo, proporciona un efecto negativo, es decir, reduce en un 9% el nivel de visitas. No siendo recomendable utilizar esta promoción para el segmento de clientes en estudio. Por otro lado, la promoción de \$5.000 de descuento sobre \$15.000, también es estadísticamente significativa y genera un efecto positivo, del orden del 10% en las visitas.
 3. Sólo una promoción alcanza efectos positivos y significativos en los clientes con propensión baja; 30% de descuento, aumentando un 33% el nivel de visitas del segmento de consumidores en estudio.

5.7. PROPUESTA DE GESTIÓN

Para determinar una propuesta de gestión de clientes normales con el objetivo de aumentar la retención anual, se determina, junto a la organización, un aspecto general que debería cumplir cada campaña, según las distintas aperturas mencionadas precedentemente, el que consiste en que el incentivo o beneficio entregado debe ser progresivo, es decir, no se debería otorgar un descuento muy alto, como primer toque, para luego, entregar rebajas mejores. Además, la campaña de retención consistirá en una serie de tres toques progresivos y con un claro incremento en el beneficio entregado.

De esta manera, la sugerencia de toques se realizará según tipo de cliente (antiguos y nuevos), tenencia de la tarjeta de crédito (abierta, cerrada y sin tarjeta) y *recency* (1 a 3, 4 a 6, 7 a 9 y 10 a 12 meses), propuesta que podrá variar, según lo requerido por el negocio y la situación de dinamismo que enfrente.

5.7.1. SEGÚN TIPO DE CLIENTE

Para determinar el orden de las promociones, según el tipo de cliente, se tomaron en consideración los resultados obtenidos del análisis de asistencia incremental y de la regresión lineal.

En este sentido, en la Tabla 54, se evidencia la distribución de toques para los clientes nuevos, donde se pone de manifiesto un orden lógico que permite tanto ir en crecimiento del efecto en el nivel de visitas, como mantener un aumento en el beneficio proporcionado.

Tabla 54: Propuesta de gestión para clientes nuevos.

Promoción	Nuevos		
	Alta	Media	Baja
\$5.000 de dcto.	TOQUE 3	TOQUE 2	TOQUE 2
\$10.000 de dcto.	-	-	-
20% de dcto.	-	-	-
30% de dcto.	-	TOQUE 3	TOQUE 1
\$5.000 de dcto. sobre \$15.000	TOQUE 1	-	-
\$10.000 de dcto. sobre \$30.000	TOQUE 2	TOQUE 1	TOQUE 3

Fuente: Elaboración propia

Por otro lado, en la Tabla 55, se muestra la distribución de toques propuesta para los clientes antiguos. Cabe mencionar que para clientes con propensión baja, se recomienda enviar sólo una promoción (\$10.000 de descuento) y no enviar dos tipos de incentivos (20% de descuento y \$5.000 de descuento sobre \$15.000), debido a la obtención de efectos negativos y significativos en el modelo de regresión lineal.

Tabla 55: Propuesta de gestión para clientes antiguos.

Promoción	Antiguos		
	Alta	Media	Baja
\$5.000 de dcto.	TOQUE 3	-	-
\$10.000 de dcto.	-	TOQUE 3	TOQUE 1
20% de dcto.	TOQUE 2	-	NO ENVIAR
30% de dcto.	-	TOQUE 2	-
\$5.000 de dcto. sobre \$15.000	TOQUE 1	TOQUE 1	NO ENVIAR
\$10.000 de dcto. sobre \$30.000	-	-	-

Fuente: Elaboración propia

5.7.2. SEGÚN TENENCIA DE LA TARJETA DE CRÉDITO

Del mismo modo, se realiza una propuesta de toques para los clientes, según el estado de la tenencia de la tarjeta de crédito del holding; abierta, cerrada o sin contrato, basada en la información proporcionada del análisis de asistencia incremental y de regresión lineal.

De la Tabla 56, se evidencia que y como se menciona anteriormente, no existían clientes con tarjeta abierta y propensión alta, lo que no permite recomendar toques a realizar para ellos. No obstante, tanto para los clientes con propensión media y baja, no se recomiendan tres toques, debido a que no todas las promociones alcanzaron la significancia estadística.

Además, se observa que existen tanto recomendaciones de promociones como sugerencias de promociones a no enviar, puesto que reducen el nivel de visitas para dicho segmento de clientes.

Tabla 56: Propuesta de gestión para clientes con tarjeta de crédito abierta o activa

Promoción	Alta	Media	Baja
\$5.000 de dcto.	No existen casos	TOQUE 2	-
\$10.000 de dcto.		-	-
20% de dcto.		NO ENVIAR	-
30% de dcto.		-	-
\$5.000 de dcto. sobre \$15.000		TOQUE 1	-
\$10.000 de dcto. sobre \$30.000		-	TOQUE 1

Fuente: Elaboración propia

De la Tabla 56, se observa la distribución de toques para los clientes con tarjeta de crédito cerrada. De ella, destaca que los clientes con propensión alta, sólo una promoción es efectiva, generando un efecto positivo significativo en el nivel de visitas de dichos clientes.

Tabla 57: Propuesta de gestión para clientes con tarjeta de crédito cerrada.

Promoción	Alta	Media	Baja
\$5.000 de dcto.	TOQUE 1	-	-
\$10.000 de dcto.	-	TOQUE 3	TOQUE 3
20% de dcto.	-	-	TOQUE 1
30% de dcto.	-	TOQUE 2	-
\$5.000 de dcto. sobre \$15.000	-	TOQUE 1	TOQUE 2
\$10.000 de dcto. sobre \$30.000	-	-	-

Fuente: Elaboración propia

Finalmente, en la Tabla 58, se evidencia la distribución de toques para los clientes que no poseen la tarjeta de crédito del holding. Destaca la variación entre los distintos niveles de propensión, dando cuenta que es más apropiado adoptar una estrategia un poco más agresiva para aquellos que tienen propensión media que con los que tienen propensión alta. Lo anterior, tiene sentido al suponer que cuando los clientes alcanzan el nivel de propensión alta, es mucho más costoso retenerlos con los incentivos monetarios utilizados.

Tabla 58: Propuesta de gestión para clientes sin tarjeta de crédito del holding.

Promoción	Alta	Media	Baja
\$5.000 de dcto.	TOQUE 3	-	TOQUE 2
\$10.000 de dcto.	-	TOQUE 2	-
20% de dcto.	TOQUE 1	TOQUE 1	-
30% de dcto.	-	TOQUE 3	TOQUE 3
\$5.000 de dcto. sobre \$15.000	TOQUE 2	-	TOQUE 1
\$10.000 de dcto. sobre \$30.000	-	-	-

Fuente: Elaboración propia

5.7.3. SEGÚN RECENCY

Por último, se complementa la propuesta de gestión de clientes, basada en el período de última compra (*recency*) de los clientes, de acuerdo a cuatro categorías: 1 a 3, 4 a 6, 7 a 9 y 10 a 12 meses.

De la Tabla 59, se evidencia la distribución de toques para los clientes que realizaron una compra hace menos de tres meses. De ella, se reconoce que para los clientes con propensión alta, predomina el beneficio de un ticket de \$10.000 de descuento libre o con restricción, mientras que para los clientes con propensión media, la sugerencia sigue siendo basada en un monto de descuento.

Tabla 59: Coeficientes para la regresión de clientes con compras hace menos de tres meses.

Promoción	Alta	Media	Baja
\$5.000 de dcto.	-	TOQUE 2	-
\$10.000 de dcto.	TOQUE 2	TOQUE 3	TOQUE 1
20% de dcto.	-	-	NO ENVIAR
30% de dcto.	-	-	-
\$5.000 de dcto. sobre \$15.000	-	-	-
\$10.000 de dcto. sobre \$30.000	TOQUE 1	TOQUE 1	-

Fuente: Elaboración propia

En la Tabla 60, se muestra la distribución de toques para los clientes con compras hace cuatro o seis meses. De ella, se plantea que sólo se encontró una promoción significativa para cada nivel de propensión, destacando el segmento de clientes con propensión baja, donde se recomienda no enviar la promoción de 30% de descuento, pues reduce las visitas que obtienen los clientes.

Tabla 60: Coeficientes para la regresión de clientes con compras entre cuatro a seis meses.

Promoción	Alta	Media	Baja
\$5.000 de dcto.	-	-	-
\$10.000 de dcto.	TOQUE 1	-	-
20% de dcto.		-	-
30% de dcto.	-	TOQUE 1	NO ENVIAR
\$5.000 de dcto. sobre \$15.000	-	-	-
\$10.000 de dcto. sobre \$30.000	-	-	-

Fuente: Elaboración propia

Del mismo modo, en la Tabla 61, se detalla la distribución de toques para los clientes con compras hace siete o nueve meses. Se observa que, si bien, en la mayoría de los casos no se manifiestan sugerencias de gestión, sólo se recomienda contactar a los clientes con propensión media utilizando la promoción de \$5.000 de descuento.

Tabla 61: Coeficientes para la regresión de clientes con compras entre siete y nueve meses.

Promoción	Alta	Media	Baja
\$5.000 de dcto.	-	TOQUE 1	-
\$10.000 de dcto.	-		-
20% de dcto.	-	-	-
30% de dcto.	-	-	-
\$5.000 de dcto. sobre \$15.000	-	-	-
\$10.000 de dcto. sobre \$30.000	-	-	-

Fuente: Elaboración propia

Por otro lado, en la Tabla 62, se presenta la distribución de toques para los clientes que no han comprado hace más de diez meses.

Tabla 62: Coeficientes para la regresión de clientes con compras hace más de 10 meses.

Promoción	Alta	Media	Baja
\$5.000 de dcto.	-	-	-
\$10.000 de dcto.	TOQUE 2	-	-
20% de dcto.	-	NO ENVIAR	-
30% de dcto.	-	-	TOQUE 1
\$5.000 de dcto. sobre \$15.000	TOQUE 1	TOQUE 1	-
\$10.000 de dcto. sobre \$30.000	-	-	-

Fuente: Elaboración propia

6. IMPACTO ECONÓMICO

Dado que se propone realizar distintas acciones de marketing directo para diferentes tipos de clientes, es importante cuantificar lo que se puede lograr producto de dicha gestión. En esta línea, se realiza un análisis de sensibilidad basado en la cantidad de clientes activados y en la propensión que tiene cada uno, considerando aquellos clientes que tienen mayor y menor gasto en su clasificación.

En la Tabla 66 y Tabla 67, se evidencia que al retener al 1% de los clientes antiguos, independiente del nivel de propensión (alta, media o baja), se obtienen 10.005 [u.m.], mientras que al retener al 1% de los clientes nuevos, en el mismo caso, se obtendrían 4.471 [u.m.].

Por otro lado, en la Tabla 63, se evidencia que al activar un 1% de los clientes antiguos con propensión alta de fuga en un año, la organización podría alcanzar 2.398 [u.m.], mientras que si ocurre el mismo escenario para los clientes nuevos, según lo señalado en la Tabla 64, se sumarían 2.094 [u.m.], generando 4.492 [u.m.] producto de la gestión realizada.

Tabla 63: Impacto económico (en unidades monetarias) en la activación de clientes antiguos con mayor gasto, según la propensión.

Activación de Clientes	Alta [35-55%]	Media [20-35%]	Baja [16-20%]
1%	2.398	5.242	1.619
2%	3.816	8.435	2.611
3%	4.919	10.991	3.415
4%	5.831	13.180	4.135
5%	6.611	15.125	4.799
6%	7.293	16.870	5.414
7%	7.902	18.452	5.990
8%	8.452	19.905	6.532
9%	8.965	21.249	7.043
10%	9.441	22.501	7.530

Fuente: Elaboración propia.

Tabla 64: Impacto económico (en unidades monetarias) en la activación de clientes nuevos con mayor gasto, según la propensión.

Activación de Clientes	Alta [35-55%]	Media [20-35%]	Baja [16-20%]
1%	2.094	1.144	1.159
2%	3.294	1.726	1.784
3%	4.198	2.154	2.264
4%	4.918	2.499	2.655
5%	5.513	2.790	2.988
6%	6.013	3.050	3.277
7%	6.437	3.286	3.531
8%	6.810	3.504	3.759
9%	7.146	3.708	3.967
10%	7.452	3.898	4.159

Fuente: Elaboración propia.

Ahora bien, al estudiar a los clientes con menor gasto, se evidencia, de las Tabla 68 y Tabla 69 que, el impacto disminuye, a un total de 13 [u.m.] para el mismo caso señalado precedentemente. Lo anterior, muestra la existencia de un *trade-off* entre activar aquellos clientes con alto riesgo de fuga que tienen un ticket promedio elevado versus clientes con un ticket promedio bajo, donde los costos de la gestión se vuelven importantes para definir si contactar a todos los clientes que pueden fugarse o sólo a un segmento de ellos.

7. CONCLUSIONES

El trabajo de título se realiza en una empresa del retail, enfocada en el negocio de las tiendas por departamento. De ello, destaca la creciente necesidad de mantener a los clientes, puesto que la adquisición de nuevos es altamente costosa. De modo que, aspectos como la calidad de servicio y experiencia de clientes, toman relevancia para que dichos clientes continúen comprando y no dejen de consumir los productos que se ofrecen.

De esta manera, se realiza un estudio con el objetivo de generar una campaña de retención de clientes que permita disminuir el nivel de fuga de los consumidores de la tienda por departamento, mediante la generación de un modelo de fuga que entregue un *target* potencial de clientes, basado en la probabilidad que, por medio de acciones de email marketing, serán contactados, puesto que son clientes con un alto riesgo de fuga.

Por medio de la campaña realizada, se pudo evaluar distintos segmentos de clientes (antigüedad del cliente, tenencia de tarjeta de crédito y *recency*), permitiendo, mostrar diferencias entre ellos, donde ciertas promociones pueden generar efectos positivos e incluso mayores en ciertas categorías de clientes, así como también, se evidenciaron incentivos que disminuyen el nivel de visitas de un consumidor. En esta línea, se evidencia que las que otorgan un descuento sujeto a montos mínimos de compra suelen ser consideradas como una barrera para aquellos clientes que no han realizado compras hace más de siete meses, lo que, en vez de re-activarlos, los incentiva a no utilizar el beneficio proporcionado.

Un aspecto relevante evidenciado tanto en el análisis de asistencia incremental como de regresión lineal, consiste en que es posible que los clientes que tienen una propensión alta, sea mucho más difícil de retener, pues alcanzaron un estado terminal considerado como fugado. De este modo, se concluye que abordar a los clientes en un estado previo, el que se puede denominar como propensos a la fuga, obtendría resultados mucho más efectivos, lo que ocurriría con los clientes que tienen propensión media, quienes, en la gran mayoría de los casos, obtuvieron más resultados significativos que aquellos con propensión alta.

Finalmente, se proporciona una forma de gestionar a los clientes, basado en los resultados obtenidos de tanto el análisis de asistencia incremental como de regresión lineal, sujetos a que el beneficio proporcionado vaya creciendo conforme aumenta el número de toques realizado y que la promoción haya generado un efecto positivo y significativo. Adicionalmente, se realiza una apertura por tres variables relevantes para el negocio: antigüedad del cliente, tenencia de la tarjeta de crédito y *recency*, sugerencias que podrá adoptar el negocio, dependiendo de las necesidades que posea en cada período.

8. TRABAJOS FUTUROS

Como se define precedentemente, el objetivo principal del trabajo de título consistía en generar una propuesta de gestión de clientes, basada en los resultados obtenidos del diseño experimental, los cuales, a su vez, dependían del modelo de fuga construido.

Un aspecto que podría ser relevante de incorporar, consiste en comprender por qué los clientes dejan de comprar en la tienda. De esta manera, sería interesante realizar una encuesta de satisfacción a aquellos clientes con propensión alta de fuga, con el objetivo de encontrar nuevas características o variables que puedan ser incorporadas al modelo y mejorar, de cierta forma, la capacidad predictiva del modelo.

Debido al creciente aumento en el uso de redes sociales, sería interesante utilizar la información que los clientes generan en la *Fanpage* de la tienda por departamento. Lo anterior, con el objetivo de identificar quiénes son los clientes que manifiestan públicamente que tienen una mala imagen de marca y así, poder incorporar dicha indagación al modelo de propensión, puesto que, dada las condiciones del negocio, un consumidor puede dejar de comprar, sin previo aviso ni tampoco generar un reclamo formal que permita a la organización conocer las razones de su decisión o el momento exacto en que deja de consumir en la organización.

Para evaluar la efectividad del modelo, asociada al target de clientes que señala con mayor riesgo de fuga y de las promociones proporcionadas para cada uno de los segmentos, se propone realizar un evolutivo histórico de la probabilidad de fuga de los clientes que fueron seleccionados para cada campaña, comparando la probabilidad al inicio y la probabilidad que posee después de tres, seis, nueve y doce meses. De esta forma, se podría validar si los efectos de la gestión de clientes generan cambios significativos en la propensión de fuga de los clientes de la tienda por departamento.

Adicionalmente, la utilización de otros canales de contacto, tales como mensajes de textos (SMS), redes sociales o estados de cuenta (si aplica), permitiría aumentar considerablemente el alcance de la gestión realizada, focalizada en aquellos clientes que poseen un riesgo mayor de fuga o bien que poseen un porcentaje de apertura muy bajo o nulo o bien que no tenga asociado un correo electrónico.

Finalmente, debido a la alta utilización de campañas de email marketing de distinta índole (gestión de clientes, acumula compra, entre otras), se vuelve interesante poder predecir qué segmento de clientes canjearán con mayor probabilidad, un cierto tipo de promoción, con el objetivo de determinar el incentivo más apropiado para cada cliente. En este sentido, se podría diseñar una campaña de retención particular para cada uno de ellos, de manera tal, de maximizar la probabilidad de canje o activación.

Finalmente, producto de que los clientes en riesgo de fuga, suelen tener una o dos visitas en un período de 12 meses, se vuelve relevante el hecho de conocer si existen clientes que comenzaron con un comportamiento similar en la tienda por departamento (clones), con el objetivo de identificar cuál o cuáles son los eventos que incentivaron que dichos clientes siguieran comprando y enviarle comunicación o promociones similares a los clones en riesgo de fuga.

9. BIBLIOGRAFÍA

- America Retail. (julio de 2017). Recuperado en septiembre de 2017, de Chile: Industria del retail prevé crecimiento de 3,3% real para 2018: <http://www.america-retail.com/chile/chile-industria-del-retail-preve-crecimiento-de-33-real-para-2018/>
- Ascarza, E. (2016). *Retention futility: targeting high risk customers might be ineffective*. Working paper.
- Bartlett, J. E., Kotrlik, J. W., & Higgins, C. C. (2001). Organizational research: Determining appropriate sample size in survey research. *Information Technology, Learning and Performance Journal*, 43-50.
- Bastías, P. (2009). *Metodología para realizar predicción de fuga de clientes en una empresa de retail*. Memoria para optar al título de Ingeniera Civil Industrial, Universidad de Chile, Departamento de Ingeniería Industrial.
- Cámara de Comercio de Santiago. (enero de 2016). (Departamento de Estudios) Recuperado en junio de 2017, de Tendencias del Retail en Chile: https://www.ccs.cl/prensa/2016/01/tendencias_retail_2016_WEB.PDF
- CERET. (s.f.). Recuperado el 06 de abril de 2017, de Industria del Retail: <http://www.ceret.cl/industria-del-retail/>
- Economía y Negocios. (marzo de 2017). Recuperado en octubre de 2017, de Retail chileno crecería 2,4% anual a 2021, uno de los menores avances en la región: <http://www.economiaynegocios.cl/noticias/noticias.asp?id=348293>
- Eidgenössische Technische Hochschule Zürich. (2012). *Finding Multivariate Outlier*. Recuperado el 15 de septiembre de 2017, de <https://stat.ethz.ch/education/semesters/ss2012/ams/slides/v2.2.pdf><https://stat.ethz.ch/education/semesters/ss2012/ams/slides/v2.2.pdf>
- Fader, P. S., Hardie, B. G., & Lee, K. L. (2005). "Counting Your Customers" the Easy Way: An Alternative to the Pareto/NBD Model (Vol. 24). Marketing Science.
- Fader, P. S., Hardie, B. G., & Lee, K. L. (2005). "Counting Your Customers" the Easy Way: An Alternative to the Pareto/NBD Model. *Marketing Science*, 24, 275-284.
- Fritis, N. (s.f.). *Estadística para la Investigación de Mercados - Cap. 7. Apunte de curso IN5625*. Universidad de Chile, Departamento de Ingeniería Industrial.
- Hadden, J., Tiwari, A., Roy, R., & Ruta, D. (2005). Computer assisted customer churn management: State-of-the-art and future trends. *Computers & Operations Research*, 2902-2917.

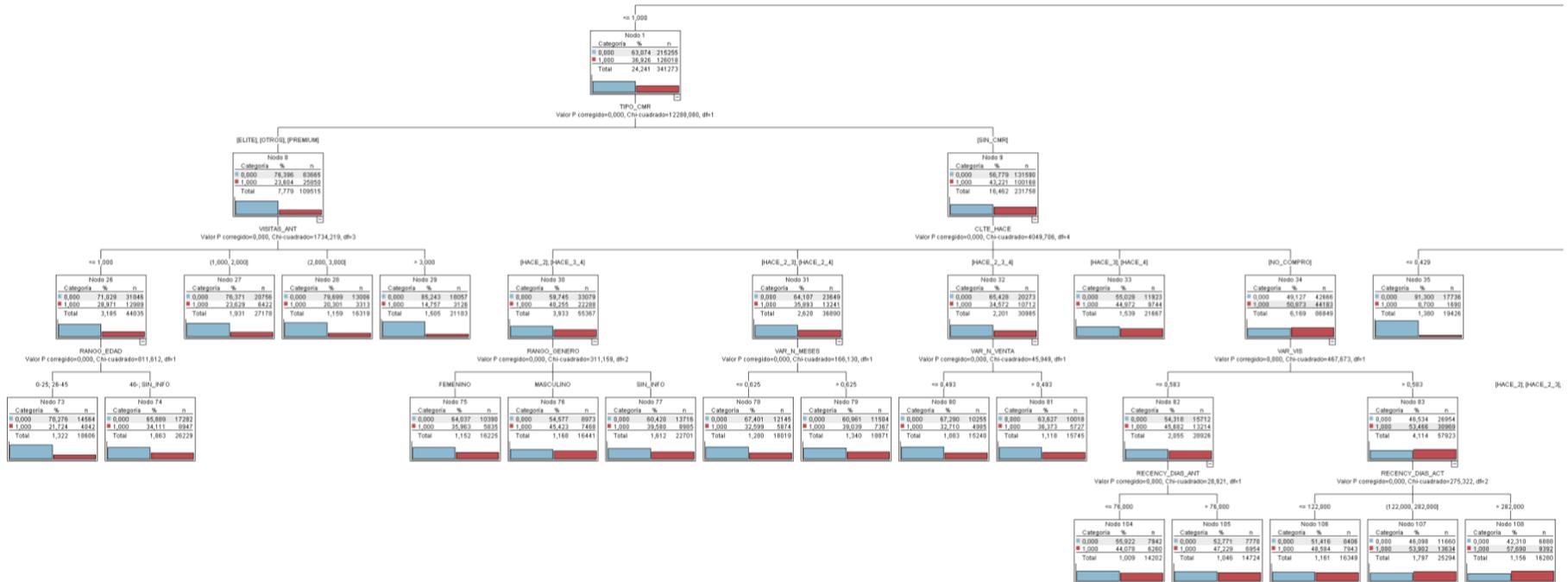
- HG DATA. (s.f.). Recuperado el 09 de abril de 2017, de IBM SPSS Modeler Market Share, Customers and Competitors: <https://discovery.hgdata.com/product/ibm-spss-modeler>
- HG DATA. (s.f.). Recuperado el 09 de abril de 2017, de Microsoft SQL Server Market Share, Customers and Competitors: <https://discovery.hgdata.com/product/microsoft-sql-server>
- IBM. (mayo de 2017). *Creación de árboles de decisión*. Recuperado el 02 de julio de 2017, de https://www.ibm.com/support/knowledgecenter/es/SSLVMB_22.0.0/com.ibm.spss.statistics.help/spss/tree/idh_idd_treegui_main.htm
- IBM. (mayo de 2017). *SPSS Decision Trees - IBM SPSS Decision Trees*. Recuperado el 02 de julio de 2017, de <http://www-03.ibm.com/software/products/es/spss-decision-trees>
- Jerath, K., Fader, P. S., & Hardie, B. G. (2011). New Perspectives on Customer “Death” Using a Generalization of the Pareto/NBD Model. *Marketing Science*, 30, 866-880.
- Lemmens, A., & Gupta, S. (2013). *Managing churn to maximize profits*. Harvard Business School. Working paper.
- Orlov, A. (s.f.). *Mahalanobis distance*. Recuperado en septiembre de 2017, de http://www.encyclopediaofmath.org/index.php?title=Mahalanobis_distance&oldid=17720
- Riquelme, F. (2017). *Evaluación de la efectividad de promociones personalizadas en la retención y aumento de los clientes de alto valor de una tienda por departamento*. Memoria para optar al título de Ingeniera Civil Industrial, Universidad de Chile, Departamento de Ingeniería Industrial.
- Rokach, L., & Maimon, O. (2005). Top-down induction of decision trees classifiers - A survey. *IEEE Transactions on systems, man, and cybernetics*, 35(4), 476-487.
- Rosenmai, P. (25 de noviembre de 2013). *Using Mahalanobis Distance to Find Outliers*. Recuperado el 15 de septiembre de 2017, de <http://eurekastatistics.com/using-mahalanobis-distance-to-find-outliers/>
- Segovia, C. (2005). *Caracterización del proceso de fuga de clientes de un retail banking utilizando información transaccional*. Memoria para optar al título de Ingeniera Civil Industrial, Universidad de Chile, Departamento de Ingeniería Industrial.

Steffen. (8 de diciembre de 2016). *Outlier Detection with Mahalanobis Distance*. Recuperado el 10 de septiembre de 2017, de <https://www.r-bloggers.com/outlier-detection-with-mahalanobis-distance/>

10. ANEXOS

10.1. ANEXO 1. ÁRBOL DE DECISIÓN PARA CLIENTES ANTIGUOS

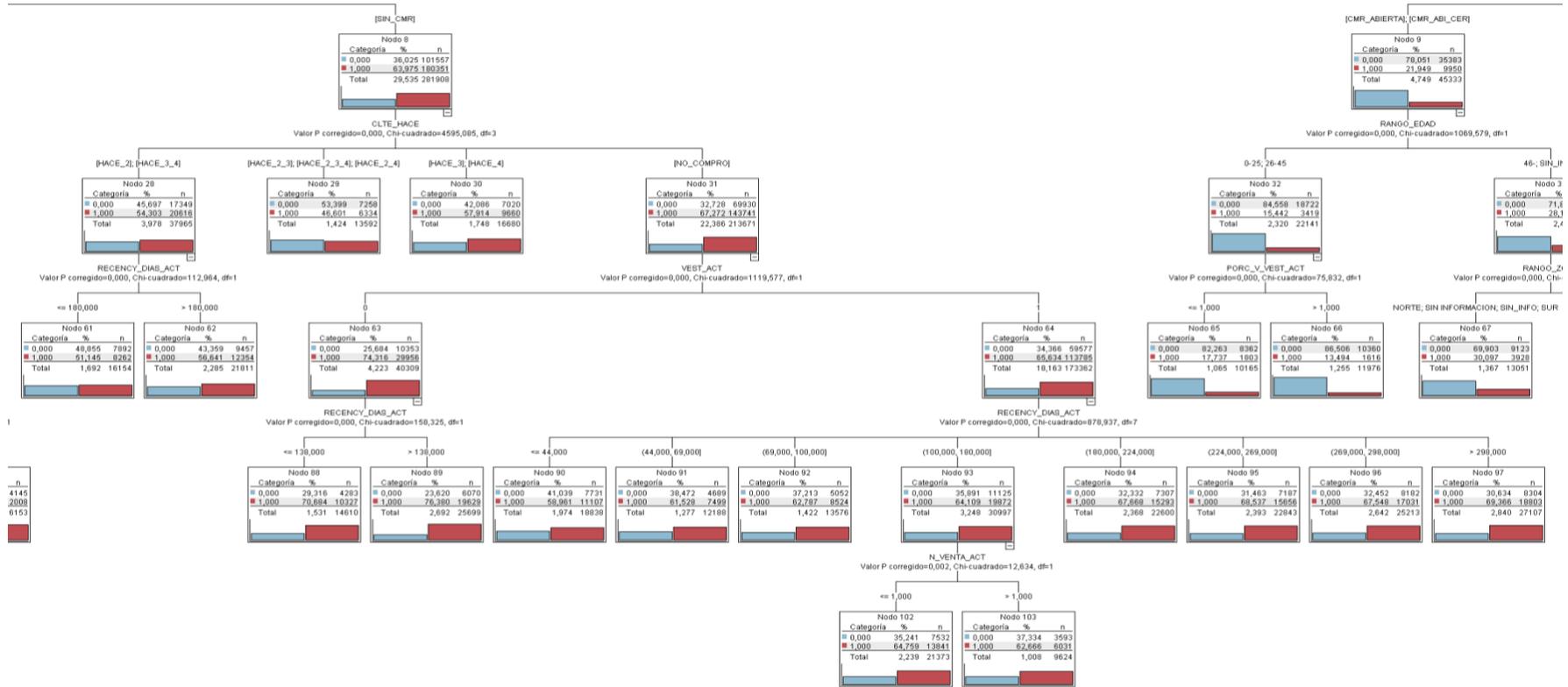
Gráfico 32: Rama del árbol que presenta la mayor cantidad de clientes antiguos fugados.



Fuente: Elaboración propia.

10.1. ANEXO 2. ÁRBOL DE DECISIÓN PARA CLIENTES NUEVOS

Gráfico 33: Rama del árbol de decisión que presenta mayor cantidad de clientes nuevos fugados.



Fuente: Elaboración propia.

10.2. ANEXO 3. ANÁLISIS DE SENSIBILIDAD DE LA PROBABILIDAD DE CORTE.

Para evaluar cuál es el punto de corte más apropiado para los clientes antiguos, se opta por buscar aquel que maximice la medida de armonía entre la precisión y el *recall*, evidenciada por el F1-Score.

Como se menciona anteriormente, se hace necesario realizar una modificación como la planteada, debido a que, para este tipo de cliente, los casos favorables no superan un 19%, a diferencia del modelo de clientes nuevos, donde los casos de éxito bordean un 49%.

De este modo, en la Tabla 65, se observa que se obtiene el máximo F1-Score para el corte de 25%, es decir, se asumirá, como caso de éxito (fuga), cuando un cliente antiguo supere una probabilidad de 25%.

Tabla 65: Métricas de desempeño con distintos puntos de corte para el modelo de clientes antiguos.

Punto de corte	Falsos Negativos	Accuracy	Precision	Recall	F1-Score
0.05	3%	42%	24%	97%	38%
0.1	8%	54%	28%	92%	42%
0.15	17%	63%	31%	83%	45%
0.2	26%	69%	35%	74%	47%
0.25	42%	76%	40%	58%	48%
0.3	46%	77%	42%	54%	47%
0.35	58%	79%	44%	42%	43%
0.4	71%	81%	48%	29%	36%
0.45	80%	81%	50%	20%	29%
0.5	91%	82%	55%	9%	15%
0.55	96%	82%	57%	4%	7%

Fuente: Elaboración propia.

10.3. ANEXO 4. SISTEMA DE PROMOCIONES

Para enviar promociones, se deben configurar todas y cada una de ellas en el sistema de promociones de la tienda por departamento.

En la Ilustración 5: Configuración de una promoción., se observa cómo se seleccionan las características que debe cumplir una boleta para aplicar el descuento en las categorías deseadas, en este caso, se muestra la configuración para la promoción de \$5.000 de descuento sobre \$15.000. En ella, se observa que el monto debe ser mínimo de \$15.000 y comprar dentro de las líneas de vestuario y calzado seleccionadas para que el descuento se haga efectivo.

Ilustración 5: Configuración de una promoción.

Editor de Reglas

Código	: 1041223	Estado	: Programacion	Tiene Horario Definido	: N
Nombre	: Retencion 5/15	Prioridad	: 3	Descuento a Aplicar	: Prorrateo Automatico
Campaña	: P. Consumer Insight Falabella	Acumulable	: Acumulable (5)	Usuario	: cdelaprida
Repetible	: 1	Fecha Desde	: 2017-11-10	Fecha Hasta	: 2017-11-26
Tipo de Ejecución	: Futuro	Tipo de Promoción	: Precio		
Descripción	: Retencion 5/15 VC Mujer, Hombre y Niños				

Definición :

Agregar Condición	Monto minimo	Monto = 15000 [15,000]	
		(Línea en lista 1,129,388,486,1265 [301-HOMBRES,302-RINCON JUVENIL HOMBRES,304-DAMAS,305-RINCON JUVENIL DAMAS,310-CALZADO]	
		O Sub Línea en lista 1065,1105,2266724,234,811,862,868777,910,976,2266732 [30906-RECIENTE NACIDOS,30905-VESTUARIO BEBE,30302-VESTUARIO DEPORTIVO MUJER,30301-VESTUARIO DEPORTIVO]	
	Si Cumple que	Local no en lista 3524,3546,3440,3537,3120,8529,3547,3570,3670,3688,3125,8502,3821,2544,8525,3529,8510,3831,3118,4600. [Aldo Estado] FALABELLA RETAIL S.A.,(Mossimo Plaza Vespucio) Falabella Retail S.A.,(Americano Alto)	
Agregar Beneficio	ENTONCES		
	Aplicar Desc a Items	desc al grupo = 5000 [5,000]	
		(Línea en lista 1,129,388,486,1265 [301-HOMBRES,302-RINCON JUVENIL HOMBRES,304-DAMAS,305-RINCON JUVENIL DAMAS,310-CALZADO]	
		O Sub Línea en lista 1065,1105,2266724,234,811,862,868777,910,976,2266732 [30906-RECIENTE NACIDOS,30905-VESTUARIO BEBE,30302-VESTUARIO DEPORTIVO MUJER,30301-VESTUARIO DEPORTIVO]	

Pasar a Vigente

Fuente: Elaboración propia.

Para que cada cliente reciba un código único válido en tienda, se generan cupones seriadados, los cuales se asocian a la promoción recién configurada. Una vez creados, se pueden subir a la plataforma de ExactTarget, para completar el campo dinámico asociado al cupón de tienda.

Ilustración 6: Creación de cupones seriadados.

Editor Tipo Cupón Seriado

Ingrese datos

Tipo Cupón :

Tipo Función :

Nombre :

tipo identificación :

Tipo : Nominativo Portador

Origen :

Promoción :

Fecha Desde :

Fecha Hasta :

Texto :

Retención 5/15 Vestuario Calzado con TMP

Máximo Transacciones Aplicar :

Días de vigencia cupón :

Veces cliente :

Fuente: Elaboración propia.

En la Ilustración 7: Resumen de promociones enviadas a los clientes., se observan todas las promociones creadas para el diseño experimental. En lo mismo, se evidencian distintos códigos, los que indican que el código que utilizan los clientes será válido en una boleta y se descontará del total de la misma, no siendo acumulable por producto, para el caso de promociones de montos de descuentos y siendo acumulable para promociones de porcentaje de descuentos.

Ilustración 7: Resumen de promociones enviadas a los clientes.

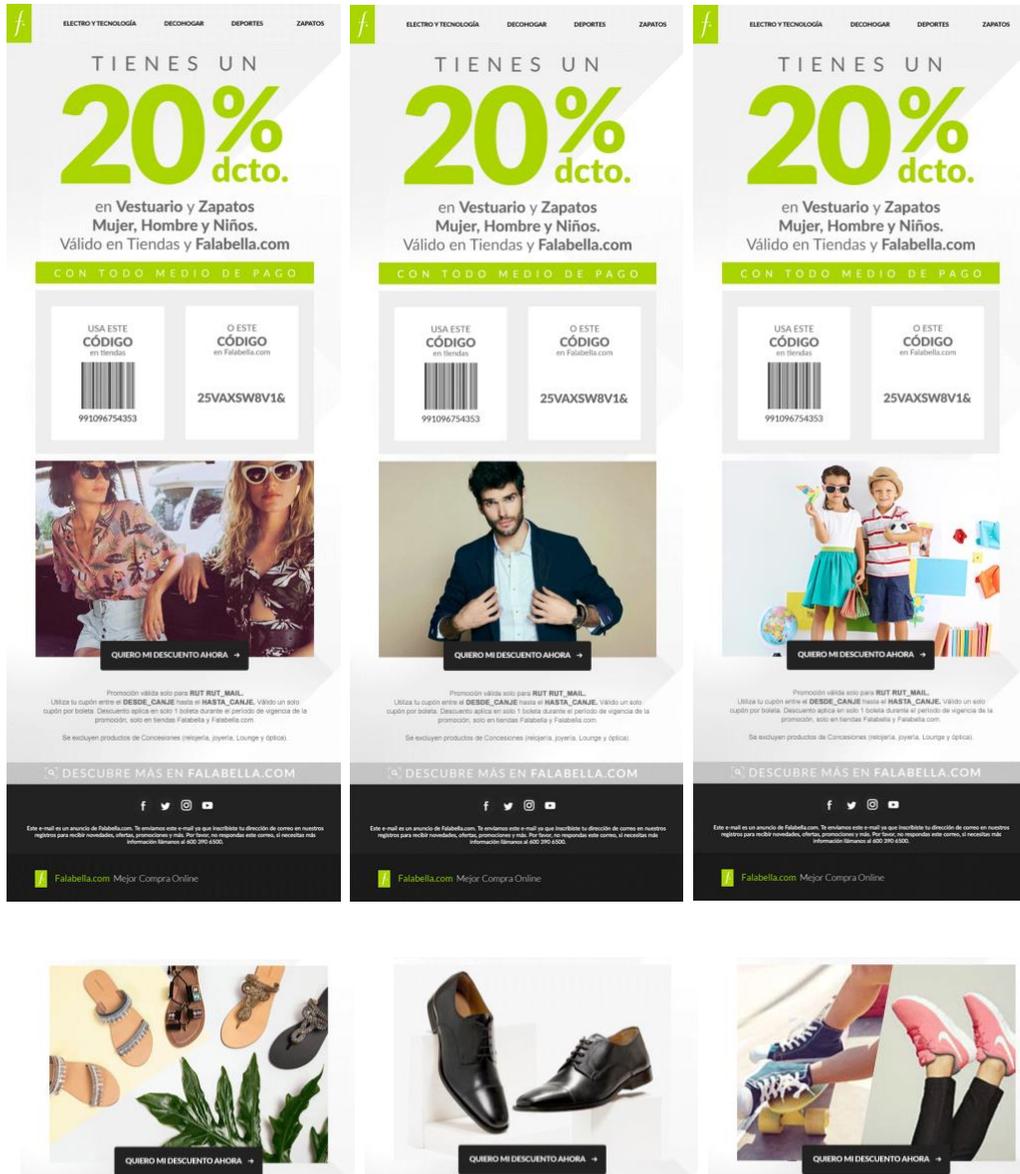
Ver	Código	Nombre Promoción	Nombre Campaña	Estado	Desde	Hasta	Tipo Desc	Acum	Prioridad	Repetible	Tiene Horario	Usuario	Fecha Mod.	Usuario Mod.	Fecha Crea.
Q	1041223	Retencion 5/15	P. Consumer Insight	Vigente	10/11/2017	26/11/2017	A	S	3	1	N	cdelaprida	06/11/2017 02:21 PM	cdelaprida	06/11/2017
Q	1041222	Retencion 10/30	P. Consumer Insight	Vigente	10/11/2017	26/11/2017	A	S	3	1	N	cdelaprida	06/11/2017 02:07 PM	cdelaprida	06/11/2017
Q	1041220	Retencion Clientes 5mil	P. Consumer Insight	Vigente	10/11/2017	26/11/2017	A	A	3	1	N	cdelaprida	06/11/2017 01:47 PM	cdelaprida	06/11/2017
Q	1041219	Retencion Clientes 10mil	P. Consumer Insight	Vigente	10/11/2017	26/11/2017	A	A	3	1	N	cdelaprida	06/11/2017 01:40 PM	cdelaprida	06/11/2017
Q	1041179	Retencion Clientes 20%	P. Consumer Insight	Vigente	10/11/2017	26/11/2017	A	S	3	0	N	cdelaprida	06/11/2017 12:37 PM	cdelaprida	06/11/2017
Q	1041194	Retencion Clientes 30%	P. Consumer Insight	Vigente	10/11/2017	26/11/2017	A	S	3	0	N	cdelaprida	06/11/2017 12:31 PM	cdelaprida	06/11/2017

Fuente: Elaboración propia.

10.4. ANEXO 5. GRÁFICA

Los clientes recibieron un correo electrónico con un *template* que proporciona información sobre el incentivo, las categorías donde es válido, un cupón en tienda y otro en la tienda online (sólo se puede utilizar uno de ellos) y un GIF alusivo a vestuario y zapatos de mujer, hombre y niños. En la Ilustración 8, se evidencian las seis imágenes que contiene el GIF.

Ilustración 8: Ejemplo de la gráfica enviada



10.1. ANEXO 6. IMPACTO ECONÓMICO

A continuación, se detallan dos tablas que muestran el alcance, en unidades monetarias, de la activación entre un 1% y un 10% tanto de clientes antiguos como de clientes nuevos.

Tabla 66: Impacto económico (en unidades monetarias) en la activación de clientes antiguos, según la propensión.

Activación de Clientes	Alta [35-55%]	Media [20-35%]	Baja [16-20%]	Total
1%	935	5.693	3.377	10.005
2%	1.639	8.836	5.565	16.040
3%	2.229	11.397	7.272	20.898
4%	2.920	13.675	8.476	25.071
5%	3.505	15.658	9.601	28.764
6%	3.975	17.441	10.650	32.066
7%	4.415	19.054	11.583	35.051
8%	4.930	20.590	12.274	37.795
9%	5.314	21.990	13.018	40.321
10%	5.746	23.297	13.616	42.658

Fuente: Elaboración propia.

Tabla 67: Impacto económico (en unidades monetarias) en la activación de clientes nuevos, según la propensión.

Activación de Clientes	Alta [59-69%]	Media [53-59%]	Baja [48-53%]	Total
1%	2.373	780	1.319	4.471
2%	3.664	1.251	2.005	6.920
3%	4.524	1.632	2.583	8.740
4%	5.209	1.967	3.026	10.202
5%	5.747	2.269	3.397	11.413
6%	6.175	2.553	3.721	12.449
7%	6.520	2.802	4.022	13.344
8%	6.833	3.042	4.283	14.158
9%	7.131	3.273	4.498	14.902
10%	7.400	3.493	4.694	15.587

Fuente: Elaboración propia.

Tabla 68: Impacto económico (en unidades monetarias) en la activación de clientes antiguos con menor gasto, según la propensión.

Activación de Clientes	Alta [35-55%]	Media [20-35%]	Baja [16-20%]
1%	10	17	12
2%	24	44	34
3%	41	79	61
4%	63	122	92
5%	87	172	127
6%	114	228	166
7%	143	289	207
8%	175	359	252
9%	210	431	298
10%	247	504	348

Fuente: Elaboración propia.

Tabla 69: Impacto económico (en unidades monetarias) en la activación de clientes nuevos con menor gasto, según la propensión.

Activación de Clientes	Alta [35-55%]	Media [20-35%]	Baja [16-20%]
1%	3	3	2
2%	9	8	6
3%	17	15	11
4%	28	23	18
5%	39	34	26
6%	54	46	35
7%	69	59	45
8%	87	73	56
9%	106	89	68
10%	126	106	81

Fuente: Elaboración propia.

