



UNIVERSIDAD DE CHILE

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

DESARROLLO Y EVALUACIÓN DE PROGRAMA DE INCENTIVOS PARA AUMENTAR FIDELIZACIÓN EN CLIENTES EN CUPONERA ONLINE

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

FELIPE ALEJANDRO PERALTA ALVAREZ

PROFESOR GUÍA:

TODD PEZZUTI LLOYD

MIEMBROS DE LA COMISIÓN:

DANIEL SCHWARTZ PELROTH

ALEJANDRA PUENTE CHANDIA

SANTIAGO DE CHILE

2016

Desarrollo y evaluación de programa de incentivos para aumentar fidelización en clientes en cuponera online

El presente proyecto se desarrolla dentro del ambiente del comercio *online* (*e-commerce*), que se caracteriza por representar todas las transacciones mediante plataformas *online* (internet y *apps*). Dentro de este ámbito se han desarrollado los sitios *online* de ventas de cupones, que ofrecen productos y servicios a un menor costo para los usuarios. A pesar de la creciente expansión de estos sitios, en general estos perciben un problema de fidelización pues la cartera de clientes se caracteriza por estar compuesta de clientes que solo generan una o dos compras en toda su historia de vida en el sitio.

Considerando lo anterior, el objetivo principal del presente proyecto consiste en desarrollar intervenciones experimentales y determinar cuál de todas posee un mayor impacto en el aumento del número de compras de los clientes y cuál genera una mayor activación de los clientes dentro del sitio.

Para poder concretar el objetivo mencionado, se utilizará una metodología que consiste en el cálculo inicial de la probabilidad de que un cliente llegue a tener más de tres compras en su historia, mediante un modelo *logit*, y con esta probabilidad se generarán segmentos de clientes: baja, media y alta probabilidad. A cada segmento de clientes se le aplicarán dos experimentos que podrían estimular su actividad y, luego de medir estadísticamente el cambio de comportamiento, se desarrollará un modelo de venta incremental para determinar los sub-segmentos de clientes que hayan obtenido una mejor respuesta a las intervenciones experimentales.

En el primero de los dos experimentos que se realizan, se enviará un *mail con oferta y asuntos personalizados*. Por otra parte, el segundo experimento consiste en un sistema de acumulación de compras con recompensa por cumplimiento de la meta estipulada.

Se concluye que las intervenciones experimentales no generan un cambio de comportamiento significativo en los clientes para ninguna de las variables de interés (número de compras y activación de clientes). A pesar de lo anterior, sí se puede establecer que la segmentación realizada es pertinente, puesto que sí existen diferencias de comportamiento para los distintos segmentos de clientes determinados. Finalmente, se encuentran dos deciles de clientes que sí responden positivamente a las intervenciones y que sí son influenciados por estas, de los cuales se obtiene un perfil susceptible de intervención en futuras campañas de *marketing*.

Se recomienda a la empresa segmentar a sus clientes para realizar un buen enfoque en sus procesos de captación y retención y se sugiere además intervenir a los clientes que cumplan un perfil de “clientes influenciados”. Finalmente se recomienda la realización e implementación del segundo experimento, pues, a pesar de que no produce un cambio significativo en el comportamiento, sí puede generar beneficios intangibles que van alineados con la misión y visión de la empresa.

Agradecimientos

Al finalizar esta etapa de mi vida, es necesario poder agradecer a todos los actores que han hecho esto posible.

En primer lugar, quisiera agradecer a mi padre Alex y a mis abuelitos Mariela y Julio, pues han sido desde siempre mis guías y mentores, mis pilares y consejeros; han sido quienes no me han dejado caer y a quienes les debo la vida. Muchas gracias por sus esfuerzos y apoyo incondicional en cada etapa de mi vida.

Agradezco también a mis hermanos, Sandra, Matías y Marcelo, quienes también han sido parte de mi vida y entregan su amor y apoyo incondicional, en especial a Matías, mi compañero de vida.

Agradezco también a mis equipos de vóleybol de Ingeniería y de la Universidad de Chile, pues siempre han sido ese grupo de amigos que juegan vóley y disfrutan del deporte. Han sido mis amigos y, sin saberlo, han ayudado a mejorar mi vida en los peores momentos de estrés, gracias a todos.

Agradezco a mis amigos de la Universidad, Pablo, Carlos, Ary, Felipe Vergara, Felipe Ñancupil, Camilo, Pedro y Caleb. Jamás olvidaré esas tardes de estudio cuando éramos mechones, ni los almuerzos o carretes que pasamos, sin duda son grandes recuerdos que llevaré por siempre.

Finalmente, no puedo dejar de agradecer a Dios.

Tabla de contenido

Índice de contenidos	iv
Índice de ilustraciones.....	ix
Índice de tablas	vii
1 Introducción	1
2 Justificación	2
3 Objetivos	3
3.1 Objetivo general.....	3
3.2 Objetivos específicos.....	3
4 Alcances.....	3
5 Resultados esperados	4
6 Marco conceptual	5
6.1 Modelos de respuesta:.....	5
6.1.1 Regresión logística:.....	5
6.2 Segmentación:.....	6
6.3 Diseño experimental	6
6.4 Grupos de control	7
6.5 Evaluación estadística	8
6.6 Modelos de venta incremental (<i>uplift</i>)	9
7 Marco metodológico	9
7.1 Revisión bibliográfica	9
7.2 Estructura y pre-procesamiento de base de datos	10
7.3 Cálculo de probabilidad de compra (<i>Propensity Score</i>)	10
7.4 Segmentación.....	10
7.5 Diseño de experimentos	10
7.6 Experimentación	10
7.7 Evaluación y análisis de resultados de intervenciones experimentales	11
7.8 Confeción de modelo <i>uplift</i>	11
7.9 Conclusiones y recomendaciones	11
8 Descripción de la situación actual.....	12
8.1 Antecedentes generales.....	12
8.2 Análisis descriptivo de clientes	13
8.2.1 Análisis de clientes por número de compras	13
8.2.2 Análisis de clientes por género	14
8.2.3 Análisis de clientes por edad.....	15
8.2.4 Análisis de clientes por tiempo de activación	16

8.2.5	Análisis de clientes por tiempo entre compras.....	17
8.2.6	Análisis de clientes por categoría de primera compra.....	18
8.2.7	Análisis de clientes por subcategoría de primera compra.....	19
9	Desarrollo de modelos.....	20
9.1	Estructura y pre-procesamiento de base de datos.....	20
9.2	Cálculo de <i>propensity score</i>	21
9.2.1	Nomenclatura de variables en modelo.....	22
9.2.2	Interpretación de coeficientes.....	25
9.2.3	Evaluación del modelo.....	25
9.3	Segmentación.....	28
9.4	Diseño de experimentos.....	29
9.4.1	Grupos de experimentación.....	29
9.4.2	Experimentos que se realizarán.....	30
9.5	Desarrollo de experimentos.....	31
9.5.1	Nivel 1: Personalización y recomendación simple.....	31
9.5.2	Nivel 2: Carrera por premio (Club de fidelización).....	32
9.6	Evaluación y análisis de experimentos.....	34
9.7	Indicadores de correo (Open rate y Click through rate).....	34
9.7.1	Activación de clientes.....	35
9.7.2	Impacto en número de compras.....	41
9.8	Determinación de clientes con mayor cambio de comportamiento debido a intervenciones.....	50
9.8.1	Modelo <i>Uplift</i>	50
10	Conclusiones.....	62
11	Recomendaciones y trabajo futuro.....	60
11.1	Recomendaciones.....	60
11.2	Trabajo futuro.....	60
12	Bibliografía.....	63
13	Anexos.....	64
13.1	Anexo 1: Histograma de compras.....	64
13.2	Anexos 2: Resultados regresión logística 1.....	65
13.3	Anexos 3: Resultados regresión logística 2.....	66
13.4	Anexos 4: Resultados regresión logística 3.....	67
13.5	Anexos 5: Resultados regresión logística 4.....	68
13.6	Anexos 6: Resultados regresión logística 5.....	69
13.7	Anexos 7: Resultados regresión logística 6.....	70
13.8	Anexos 8: Resultados regresión logística 7.....	71

13.9	Anexos 9: Resultados test de medias. Caracterización clientes deciles 1, 2, 9 y 10. Variable Edad	72
13.10	Anexos 10: Resultados test de medias. Caracterización clientes deciles 1, 2, 9 y 10. Variable Frecuencia de compras	72
13.11	Anexos 11: Resultados test de medias. Caracterización clientes deciles 1, 2, 9 y 10. Variable Tiempo de activación	72
13.12	Anexos 12: Tabla de contingencia Categoría productos.	72
13.13	Anexos 13: Test de proporciones categoría Productos.....	73
13.14	Anexos 14: Tabla de contingencia, Subcategoría Productos de gran consumo.....	73
13.15	Anexos 15: Test de proporciones, Subcategoría Productos de gran consumo.....	74

Índice de tablas

Tabla 1: Resumen coeficientes regresiones.....	23
Tabla 2: Resumen coeficientes regresiones.....	24
Tabla 3: Matriz de confusión, umbral 30%.	26
Tabla 4: Área bajo la curva ROC.	28
Tabla 5: Distribución de propensity score para segmentos de clientes.	29
Tabla 6: Distribución de número de clientes para los grupos de experimentación según nivel de tratamiento y segmento de propensity score.....	30
Tabla 7: Distribución de proporción de clientes para los grupos de experimentación según nivel de tratamiento y segmento de propensity score.....	30
Tabla 8: Caracterización de niveles experimentales.	33
Tabla 9: Indicadores de rendimiento para correo electrónico, Open rate y Click through rate (CTR).	35
Tabla 10: Estimaciones de medias para niveles de tratamiento. Variable dependiente: Activación. Efectos fijos: Nivel de tratamiento.	35
Tabla 11: Prueba de efectos inter sujetos. Variable dependiente: Activación. Efectos fijos: Nivel de tratamiento.	36
Tabla 12: Comparación de pares entre niveles de tratamiento. Variable dependiente: Activación. Efectos fijos: Nivel de tratamiento.	36
Tabla 13: Tabla de contingencia para variable activación en los niveles de tratamiento.	37
Tabla 14: Prueba Chi-cuadrado para variable activación en niveles de tratamiento.	37
Tabla 15: Estimaciones de medias para segmento de clientes. Variable dependiente: Activación. Efectos fijos: Segmento de clientes	37
Tabla 16: Prueba de efectos inter sujetos. Variable dependiente: Activación. Efectos fijos: Segmento de clientes.	38
Tabla 17: Comparación de pares entre segmento de clientes. Variable dependiente: Activación. Efectos fijos: Segmento de clientes.	38
Tabla 18: Estimaciones de medias para segmento de clientes dentro de cada nivel de tratamiento. Variable dependiente: Activación. Efectos fijos: Nivel de tratamiento y Segmento de clientes.....	39
Tabla 19: Prueba de efectos inter sujetos. Variable dependiente: Activación. Efectos fijos: Nivel de tratamiento y Segmento de clientes.....	39
Tabla 20: Comparación de pares entre segmento de clientes para los distintos segmentos de clientes. Variable dependiente: Activación. Efectos fijos: Nivel de tratamiento y Segmento de clientes.....	40
Tabla 21: Estimaciones de medias para nivel de tratamiento dentro de segmento de clientes. Variable dependiente: Activación. Efectos fijos: Segmento de clientes y Nivel de tratamiento.	41
Tabla 22: Comparación de pares entre nivel de tratamiento dentro de cada segmento de clientes. Variable dependiente: Activación. Efectos fijos:Segmento de clientes y Nivel de tratamiento.....	41
Tabla 23: Estimaciones de medias para niveles de tratamiento. Variable dependiente: Número de compras. Efectos fijos: Nivel de tratamiento.....	42
Tabla 24: Prueba de efectos inter sujetos. Variable dependiente: Número de compras. Efectos fijos: Nivel de tratamiento.	42
Tabla 25: Comparación de pares para nivel de tratamiento. Variable dependiente: Número de compras. Efectos fijos: Nivel de tratamiento.....	43
Tabla 26: Estimaciones de medias para Segmento de clientes. Variable dependiente: Número de compras. Efectos fijos: Segmento de clientes.....	43
Tabla 27: Prueba de efectos inter sujetos. Variable dependiente: Número de compras. Efectos fijos: Segmento de clientes.....	43
Tabla 28: Comparación de pares para segmentos de cliente. Variable dependiente: Número de compras. Efectos fijos: Segmento de clientes.....	44

Tabla 29: Estimaciones de medias para segmento de clientes dentro de cada nivel de tratamiento. Variable dependiente: Número de compras. Efectos fijos: Nivel de tratamiento y Segmento de clientes.	45
Tabla 30: Prueba de efectos inter sujetos. Variable dependiente: Número de compras. Efectos fijos: Nivel de tratamiento y Segmento de clientes.	45
Tabla 31: Comparación de pares entre segmento de clientes para los distintos segmentos de clientes. Variable dependiente: Número de compras. Efectos fijos: Nivel de tratamiento y Segmento de clientes.	46
Tabla 32: Estimaciones de medias para nivel de tratamiento dentro de segmento de clientes. Variable dependiente: Número de compras. Efectos fijos: Segmento de clientes y Nivel de tratamiento.	47
Tabla 33: Comparación de pares entre nivel de tratamiento dentro de cada segmento de clientes. Variable dependiente: Número de compras. Efectos fijos: Segmento de clientes y Nivel de tratamiento.	48
Tabla 34: Q de Qini para distintos método de separación para Random Forest.	51
Tabla 35: Tabla resumen de resultado modelo Uplift basado en Random Forest ordenado por deciles. ..	52
Tabla 36: Resumen de estadísticos descriptivos para edad.	54
Tabla 37: Estadísticos descriptivos para variable Frecuencia de compra.	55
Tabla 38: Estadísticos descriptivos para variable Tiempo de activación.	56
Tabla 39: Distribución de órdenes de compra según categoría para deciles 1, 2, 9 y 10.	58
Tabla 40: Distribución de órdenes de compra según subcategoría para deciles 1, 2, 9 y 10.	60

Índice de ilustraciones

Ilustración 1: Esquema de desarrollo metodológico.....	12
Ilustración 2: Histograma de compra clientes activos.	13
Ilustración 3: Distribución de género en clientes activos.	14
Ilustración 4: Distribución de clientes según tramo de edad.....	15
Ilustración 5: Distribución de clientes según tiempo de activación.....	16
Ilustración 6: Distribución de clientes según tiempo entre compras.	17
Ilustración 7: Benchmark de categorías y subcategorías del sitio.	18
Ilustración 8: Distribución de clientes según categoría de primera compra.	19
Ilustración 9: Distribución de clientes según categoría de primera compra.	20
Ilustración 10: Curva ROC. Fuente: Elaboración propia en software.....	27
Ilustración 11: Histograma de distribución propensity score.	28
Ilustración 12: Correo enviado a clientes nivel 1 de tratamiento: Personalización y recomendación simple.	32
Ilustración 13: Mail invitación experimento nivel 2: Carrera por compras (club de fidelización).	33
Ilustración 14: Correos nivel 2: Carrera por compras (club de fidelización). Izquierda: Correo semanal de estado de avance. Derecha: Correo de ganadores que completaron la meta estipulada.....	34
Ilustración 15: Gráfico de media de número de compras realizada en el periodo de experimentación segmento por segmento de clientes. APS: Alto propensity score, MPS: Medio propensity score, BPS: Bajo propensity score.....	48
Ilustración 16: Curva de ganancia de información para modelos Uplift. Q de Qini para modelo basado en Random Forest: 2,76%. Q de Qini para modelo basado en Regresión: 0,07%.	51
Ilustración 18: Curva uplift, modelo basado en Random Forest ordenado por deciles.	52
Ilustración 18: Histograma de edades para deciles de interés. Deciles 1, 2, 9 y 10.....	54
Ilustración 19: Histograma de frecuencia de compras para deciles 1, 2, 9 y 10.....	55
Ilustración 20: Histograma de frecuencia de compras para deciles 1, 2, 9 y 10.....	56
Ilustración 21: Distribución de categorías de órdenes de compras generadas por los deciles 1 y 2 durante proceso de experimentación.....	57
Ilustración 22: Distribución de categorías de órdenes de compras generadas por los deciles 9 y 10 durante proceso de experimentación.....	57
Ilustración 23: Distribución de subcategorías de órdenes de compras generadas por los deciles 1 y 2 durante proceso de experimentación.....	59
Ilustración 24: Distribución de subcategorías de órdenes de compras generadas por los deciles 9 y 10 durante proceso de experimentación.....	59
Ilustración 25: Distribución de número de compras históricas en la empresa.....	64
Ilustración 26: Resultado modelo 1 – Categorías: producto, health & beauty y gastronomía.....	65
Ilustración 28: Resultado modelo 2 – Categorías: productos, health & beauty, gastronomía y subcategorías: hogar, otros restaurantes, productos de gran consumo y tratamientos reductivos	66
Ilustración 28: Resultado modelo 3: Género, edad, categorías: productos, health & beauty, gastronomía y subcategorías: hogar, otros restaurantes, productos de gran consumo y tratamientos reductivos	66
Ilustración 29: Resultado modelo 4 – Monto de primera compra, género, edad, categorías: productos, health & beauty, gastronomía y subcategorías: hogar, otros restaurantes, productos de gran consumo y tratamientos reductivos.....	67
Ilustración 30: Resultado modelo 5 - Frecuencia, edad, categorías: productos, health & beauty, gastronomía y subcategorías: hogar, otros restaurantes, productos de gran consumo y tratamientos reductivos sólo si género es femenino.....	68
Ilustración 31: Resultado modelo 6 – Activación, frecuencia, edad, categorías: productos, health & beauty sólo si género es femenino, gastronomía y subcategorías: hogar, otros restaurantes, productos de gran consumo y tratamientos reductivos	69

Ilustración 32: Resultado modelo 7 – Recency, activación, frecuencia, edad, categorías: productos, health & beauty sólo si género es femenino, gastronomía y subcategorías: hogar, otros restaurantes, productos de gran consumo y tratamientos reductivos	70
Ilustración 33: Resultado modelo 8 – Recency, activación, frecuencia, logaritmo de edad, categorías: productos, health & beauty sólo cuando género es femenino, gastronomía y subcategorías: hogar, otros restaurantes, productos de gran consumo y tratamientos reductivos	71
Ilustración 34: Test de medias para variable: Edad para deciles 1, 2, 9 y 10.....	72
Ilustración 35: Test de medias para variable: Frecuencia de compras para deciles 1, 2, 9 y 10.	72
Ilustración 36: Test de medias para variable: Tiempo de activación para deciles 1, 2, 9 y 10.	72
Ilustración 37: Tabla de contingencia: Categoría "Productos" para deciles 1, 2, 9 y 10.....	72
Ilustración 38: Prueba Chi-Cuadrado de proporciones para dos muestras independientes. Categoría "Productos" para deciles 1, 2, 9 y 10.	73
Ilustración 39: Tabla de contingencia: Subcategoría "Productos gran consumo" para deciles 1, 2, 9 y 10.	73
Ilustración 40: Prueba Chi-Cuadrado de proporciones para dos muestras independientes. Subcategoría "Productos gran consumo" para deciles 1, 2, 9 y 10.	74

1 Introducción

El presente trabajo se enmarca dentro de una empresa de venta de cupones de descuento *online*, que se circunscribe en el rubro del *e-commerce*.

En los últimos años, el comercio electrónico (*e-commerce*) ha aumentado considerablemente su participación de mercado, cuyas ventas el año 2013 ascendieron a US\$ 1.600 millones y se estimó que estas superarían los US\$ 2.000 millones el 2014, con un crecimiento anual de entre un 20 y 30%, según el presidente de la cámara de comercio de Santiago (CSS), Peter Hill en mayo del 2014 (Hill, 2014).

Lo anterior da muestras de que el rubro dentro del cual se encuentra inserto el desarrollo del proyecto es de gran importancia por su creciente aumento en el nivel de ventas dentro del país.

A pesar de esto, la empresa donde se desarrolla el proyecto enfrenta un gran problema, pues se aprecia que la cartera de clientes que poseen tiene un comportamiento particular e indeseable para ellos, puesto que el 72% de los clientes tienen una o dos compras dentro de su historia¹. Esto representa una gran dificultad, ya que adquirir a un nuevo cliente constituye un gasto mayor que retener a un cliente ya existente (Wansink, 2003), por lo que se tiene un especial interés por parte de la empresa en determinar cuáles serían las mejores acciones posibles de realizar para poder prolongar el ciclo de vida de un cliente en la empresa y aumentar el número de compras y rentabilidad de ellos.

Anteriormente se ha analizado el proceso de recompra y la lealtad de un cliente en comercio electrónico, pero este ha sido de tipo cualitativo y con un enfoque más orientado a analizar a las empresas oferentes. No se ha realizado, hasta el momento, un estudio acerca de las acciones realizadas a clientes y el efecto que tiene en el número de compras, y con menor énfasis al segmento de clientes poco frecuentes en los sitios.

Debido a lo anterior es surge la idea de realizar dos experimentos para poder así evaluar cuál de ellos posee un mayor impacto en el aumento del número y activación de los clientes, con el fin de mejorar la situación actual del número de compras de los clientes y rentabilizar mejor las inversiones.

Los experimentos que se realizan consisten en el envío de un mail con ofertas personalizadas (experimento 1) y elaboración de un sistema de acumulación de compras con recompensas (experimento 2).

¹ Fuente: Elaboración propia

2 Justificación

El tema de esta memoria se desarrolla dentro del sitio *web* de comercio *online* *Cuponatic*, empresa dedicada a la venta de cupones *online*, cuyo modelo de negocio se basa en ofrecer productos y servicios a sus clientes con un descuento llamativo para ellos.

Dado su carácter de comercio *online*, su principal medio de *marketing* opera a través de plataformas como las redes sociales y, particularmente, el correo electrónico, pues es su principal canal de comunicación con sus clientes. El *mail* enviado actualmente ofrece una serie de descuentos que podrían ser llamativos para los suscritos.

A pesar del alto flujo de personas que visitan el sitio y el número de compras realizadas, la empresa enfrenta una problemática: aproximadamente el 72% de los clientes se ubican en la categoría de “poco frecuentes”, dada la baja recurrencia de compra en el sitio² (53% tiene una compra y 19% tienen dos compras), lo que demuestra un patrón de comportamiento claro y poco rentable para la empresa debido a que se percibe un bajo nivel de apego y fidelidad de los clientes con el sitio.

Lo anterior cobra vital relevancia, pues si se considera que la inversión por adquisición de clientes puede llegar hasta los US\$ 10 por persona y en promedio se observa que un cliente entrega US\$ 6 de comisión por compra realizada, en muchos casos se pierde dinero, puesto que la inversión es mayor al ingreso percibido por clientes, más aún si se considera que la mayoría de estos solo registra una compra en su historial.

Por otro lado, los clientes que tienen más de dos compras, generan una comisión promedio de US\$ 32, lo que incentiva a modificar el comportamiento de compra de los estos, dado que en este caso sí se aprecia un claro beneficio para la empresa.

Adicionalmente, es posible ver que si se modifica el comportamiento de compra de un 5% de los clientes objetivos (aumentando en una compra su historial), se pueden generar beneficios correspondientes al 11,37%³ del ingreso mensual promedio de la empresa, es decir, existe un claro beneficio económico para la compañía al modificar el comportamiento de compra de sus clientes y aumentar su relación con la empresa en términos del número de compras realizadas.

Finalmente, se destaca que hoy en día la empresa no posee una clara política de retención de clientes, a diferencia de lo que ocurre con su política de adquisición, donde sí cuentan con estrategias y campañas de activación y captura de nuevos usuarios.

De acuerdo con lo anterior, el presente estudio pretende evaluar el impacto de intervenciones experimentales a clientes del sitio para poder determinar si estas intervenciones modifican el comportamiento de los consumidores y aumentan su número de compras realizadas, mejorando así su frecuencia de compra, es decir, que exista un menor intervalo de tiempo entre compras.

Considerando el impacto de las intervenciones, se propondrá a la empresa algunas iniciativas para implementar mejoras en la fidelidad de clientes y potenciar su política de retención de clientes, que hoy en día es prácticamente inexistente.

² Véase Anexo 1: Histograma de compras

³ Fuente: Elaboración propia

3 Objetivos

El proyecto tiene por finalidad evaluar experimentalmente la respuesta de un grupo de clientes frente a dos niveles de experimentación distintos (experimento 1 y experimento 2) en dos variables relevantes: número de compras y activación de clientes. La primera variable hace referencia a la cantidad de compras realizadas por los clientes, gracias a la experimentación, y la segunda busca medir la cantidad de clientes que fueron incentivados a comprar, independiente de la cantidad de compras realizadas, gracias al proceso experimental.

Además de medir experimentalmente dichos comportamientos, se buscará entender cuáles son las variables relevantes que influyen en la decisión de compra de los clientes y determinar qué clientes responden mejor a la experimentación, con el fin de seleccionar la mejor cartera de intervenciones y el perfil de clientes ideales para poder focalizar los recursos de *marketing* y capturar así a los consumidores que potencialmente presenten un comportamiento de compra positivo, es decir, usuarios que puedan tener más de tres compras en su historia como clientes.

Sobre la base de lo anterior, el proyecto posee los siguientes objetivos generales y específicos:

3.1 Objetivo general

- Determinar las intervenciones que tengan un mayor impacto en el aumento de número de compras y activación de clientes dentro de un sitio *web* de venta de cupones.

3.2 Objetivos específicos

1. Identificar variables relevantes que inciden en la decisión de compra de los clientes.
2. Medir y evaluar el efecto de intervenciones experimentales en todos los segmentos de clientes.
3. Determinar los sub-segmentos de clientes con mayor propensión de ser incentivados por medio de intervenciones.
4. Definir el conjunto de acciones que tengan el mayor impacto en el número de compras y activación de clientes en todos los segmentos.

4 Alcances

La empresa de cupones online donde se desarrollará el proyecto, Cuponatic, posee actividades comerciales en cuatro países de América latina, pero el presente trabajo se limitará a trabajar con los clientes de las actividades en Chile.

Es necesario señalar que en este estudio se considera un grupo de clientes activos, entendidos como aquellos que han realizado al menos una compra entre el 01 de noviembre de 2014 y el 30 de junio de 2015.

Actualmente en Cuponatic, se cuenta con estrategias de activación que consisten en realizar concursos donde se sorteá algún premio a aquellos clientes que “compran” un cupón de adherencia al sorteo. Dicho cupón posee costo cero, debido a que es un sorteo y la idea subyacente del concurso es que el cliente se registre y conozca el sistema de compra del sitio y no que ellos desembolsen dinero por su participación. Considerando esto, es que para el presente desarrollo no se considerarán este tipo de transacciones, pues no representan el comportamiento que se desea evaluar, al no representar a un cliente en sí.

El proyecto contempla la realización de dos experimentos: uno de ellos (experimento número 1) consiste en el envío de un correo con mayor personalización a los clientes y el otro (experimento número 2), en un sistema de acumulación de compras con recompensa si se logra la meta.

Para el experimento número uno la personalización se presenta de dos formas: asunto del correo personal, incluyendo en nombre del cliente, y una recomendación de oferta simple basada en su historial de compra. Para este punto se destaca que la recomendación no se basa en un modelo estructurado, debido a que no es atingente al objetivo del presente proyecto.

El proceso de experimentación tendrá un tiempo limitado de desarrollo de dos meses, entre el 26 de julio de 2015 y el 26 de septiembre del mismo año, enviando recordatorios semanales a los clientes. Para los usuarios del experimento 1, se enviará un correo con una nueva recomendación que variará dependiendo de las ofertas del sitio, mientras que para el experimento 2, se enviará su estado de cuenta hasta la fecha de envío, donde se incluirán sus nuevas compras (actualizadas). El grupo de control no será contactado por ningún medio adicional.

La empresa cuenta con seis principales categorías de ofertas: Productos, *health & beauty*, gastronomía, entretenimiento, viajes y otros. Dentro de estas categorías, se excluirán del análisis las correspondientes a viajes y oferta de *deals*, para el experimento uno y dos, debido a políticas de la empresa.

5 Resultados esperados

Se pretende tener una medición de las variables relevantes a la hora de tomar la decisión de compra de los clientes de la empresa.

Además, se pretende tener un resultado claro del proceso de experimentación realizado dentro de la empresa, identificando claramente qué experimento tuvo un mayor impacto para las variables: número de compras generadas y número de clientes activados, en cada segmento de clientes y si estos fueron realmente significativos para modificar el comportamiento de compra del grupo de usuarios intervenidos.

Junto con la determinación de los experimentos con mayor impacto en las variables declaradas, se pretende contar con la identificación del segmento de clientes que tengan un mayor cambio en su probabilidad de compra, dependiendo de si son intervenidos o no para poder informar a la empresa el perfil de consumidores que posee el mejor comportamiento, con el fin de enfocar sus campañas de *marketing* para aumentar la retención de clientes.

Finalmente se pretende realizar un resumen informativo para la empresa con los resultados de la experimentación, entregando información clara de los resultados del proceso, con las intervenciones de mayor impacto y los segmentos de clientes que deberían ser objeto de intervención en el futuro, para, de este modo, maximizar su rentabilidad de inversión.

6 Marco conceptual

6.1 Modelos de respuesta:

Corresponden a modelos que buscan explicar y entender el comportamiento de los clientes, con el fin de poder dar información respecto de las variables relevantes que modelan la respuesta de ellos. Son altamente relevantes en el diseño de promociones, pues buscan poder pronosticar el comportamiento de los clientes frente a la promoción en cuestión. Dentro de los *modelos de respuesta*, se encuentran los de *propensión* que explican la probabilidad de respuesta de un cliente frente a un fenómeno en particular, en este caso se modela la respuesta del cliente frente a la intervención, no el cambio en esta.

6.1.1 Regresión logística:

La *regresión logística* es utilizada como método de *modelo de propensión* y consiste en el cálculo de la *probabilidad de ocurrencia* de un evento en función de variables relevantes. Esta probabilidad proviene de la predicción de decisión en función de regresores asociados a dichas variables y ponderados por la importancia relativa de estos sobre la base de la información de datos observacionales. La fórmula para el cálculo de la probabilidad se presenta a continuación:

$$P(Y_i = 1|X) = \frac{e^{(\beta_0 + \sum_k(\beta_{ik} * X_{ik}))}}{1 + e^{(\beta_0 + \sum_k(\beta_{ik} * X_{ik}))}}$$

La fórmula anterior calcula la probabilidad de que Y tome el valor 1 dado el vector X de variables. En este caso, corresponde a que el cliente compra (Y corresponde a decisión de comprar), dado el vector de variables X determinado anteriormente (edad, género, monto de compra, etc.). Los parámetros β del lado derecho de la ecuación corresponden a la valoración de cada variable ingresada en el cálculo de la probabilidad requerida, con excepción de β_0 que corresponde a la valoración y probabilidad intrínseca de un cliente a comprar.

Se destaca que este método es altamente utilizado, pues es un modelo de fácil interpretación y rápida modelación, lo que facilita el cálculo de propensión de compra y, además, el modelo tiene un buen ajuste a los datos, por lo que es adecuado al problema que se intenta resolver, pues el objetivo consiste principalmente en medir el efecto de las intervenciones y no el desarrollo de un modelo que explique el comportamiento de compra de clientes. Esto solo sirve para poder clasificar y segmentar a los clientes.

Finalmente, esta probabilidad obtenida es la que se denomina *Propensity score* y corresponde al valor entregado por el modelo de predicción, *regresión logística*.

6.2 Segmentación:

Para efectos del proyecto, se realizará una segmentación simple de los clientes sobre la base de su *Propensity Score*, calculado anteriormente. Este proceso de segmentación busca separar o dividir un grupo de clientes en grupos más pequeños donde los clientes dentro del grupo son homogéneos entre sí, y donde los grupos sean heterogéneos entre sí con el fin de poder clasificar a los clientes y medir el comportamiento de cada grupo y en particular, ver si algún segmento posee una mejor respuesta frente a los distintos niveles de experimentación.

El proceso de segmentación puede ser realizado mediante técnicas estructuradas, con el fin de lograr que los clientes dentro de los segmentos sean homogéneos y que los segmentos sean heterogéneos entre sí, buscando distintas medidas de similitud y criterios de separación.

Dentro de las técnicas más comunes de segmentación no jerárquicas, se encuentran las siguientes:

- *K-Means*
 - Método iterativo que busca reducir la desviación estándar, agrupando n observaciones dentro de k grupos. El proceso busca particionar el conjunto en el cual cada observación se acerque más a la media del grupo.
- *Expectation maximization (EM)*
 - Este proceso de categorización se basa en la alternación de pasos de cálculo de esperanza, donde se computa la esperanza de la verosimilitud en función de las variables latentes incorporadas, y un paso de maximización que busca maximizar dicha esperanza calculada.

Si bien, este tipo métodos de segmentación determinan de mejor manera a los clientes que se parecen más entre sí, mediante procesos estructurados, la segmentación que se realizará en este proyecto no se basa en un modelo estructurado; sino mediante un indicador secundario, ya que no es el objetivo del proyecto catalogar a los clientes; sino que poder buscar espacios de mejora para las intervenciones y medir si existen predisposiciones distintas a ser buenos clientes y si esta predisposición cambia a partir del proceso experimental.

6.3 Diseño experimental

El diseño experimental busca estructurar el proceso de experimentación para obtener una investigación bien realizada. Dicho proceso se basa en la manipulación de alguna variable de interés, con el fin de analizar el cambio de comportamiento producido por la modificación de esa variable. El diseño experimental opera sobre la base de la asignación aleatoria de las unidades experimentales en diversos niveles o categorías de las variables de interés que se manipulan.

Dentro de las variables que interactúan en el diseño experimental, aparecen tres tipos principales:

- *Variable independiente*, la cual se pretende manipular según sea el interés del investigador.

- *Variable dependiente*, la cual se medirá posteriormente para analizar las variaciones según las modificaciones en las *variables independientes*.
- *Variables exógenas* que servirán como control y escapan del control del experimentador.

Lo anterior puede traducirse en que la *variable independiente* es la causa de la *variable dependiente* y la *variable exógena* representa confusión en la medición de la *variable dependiente*.

La *variable dependiente* o de respuesta es aquel aspecto que se desea evaluar para observar el efecto de la variación sistemática de las variables independientes. Finalmente, el control experimental busca eliminar o neutralizar cualquier fuente de variación que pueda confundir la variable de tratamiento.

El diseño experimental suele tener una planificación estandarizada que consiste en los siguientes pasos:

1. Formulación de hipótesis
2. Selección de variable independiente y dependiente adecuada
3. Control de variables exógenas.
4. Manipulación de las variables independientes y registro de la variable dependiente
5. Análisis estadístico de los datos
6. Inferencia de la relación entre variable dependiente e independiente

6.4 Grupos de control

Para todo proceso experimental, es necesario corroborar que las intervenciones experimentales tengan un efecto (Davis & Smith, 2004).

Las técnicas de control asociadas al diseño se dividen en tres:

1. *Aleatorización*: Consiste en el diseño de grupos con individuos asignados completamente al azar.
2. *Constancia*: Asociado a diseño de grupos apareados y de bloques.
3. *Sujeto como control propio*: Diseño asociado a muestras repetidas de experimentación.

Además de lo anterior, se pueden clasificar los diseños experimentales según la estrategia de comparación:

1. *Diseño entre grupos*
2. *Diseños intra sujetos o de medidas repetidas*
3. *Diseño mixto*

El *diseño experimental entre grupos* tiene por objetivo comparar el comportamiento de la variable dependiente entre distintos grupos de interés.

1. *Diseño entre grupos simple*: Corresponde a diseños de manipulación de una sola variable independiente. Este tipo de diseño puede variar en función del número de grupos y el objetivo de investigación en los siguientes tipos:
 - a. *Diseño de dos grupos*: Una de las situaciones más simples en la investigación experimental donde existe un grupo de control y uno de

- tratamiento que muchas veces se reduce a la presencia o ausencia de alguna variable independiente.
- b. *Diseño multi-grupo*: Corresponde a estructuras de una variable independiente con más de dos valores de dicha variable independiente. A raíz de esto, es posible extraer la relación funcional entre variable independiente y dependiente del experimento.
 - c. *Diseño de bloques de grupos al azar*: El principal objetivo de este tipo de diseños es controlar las diversas fuentes de variación exógena donde cada bloque va controlando diversas fuentes de variación con excepción del control.
2. *Diseño factorial*: Estructura de investigación donde se combinan dos o más diseños simples, es decir, manipulación de dos o más variables independientes.

Adicionalmente a la experimentación entre grupos, existen diseños para mediciones repetidas que corresponde a una extensión al diseño en bloques mencionado anteriormente, donde los sujetos sustituyen los bloques y funcionan como control propio al recibir todos los tratamientos posibles y contrastar la efectividad de tratamiento en sí mismo, aislando fuentes de variación exógena o intrínsecas al individuo tratado. Dentro de este tipo de diseño también se pueden apreciar categorías dependientes de los grupos:

1. *Diseño de grupo simple*:
 - a. *Diseño de un grupo*: Los sujetos se combinan con los distintos tratamientos. La principal ventaja de este tipo de experimentación es que se elimina el error atribuido a las diferencias individuales.
 - b. *Diseño multi-grupo o diseño factorial mixto*: El desarrollo de este tipo de diseño incorpora dos estrategias de inferencia: comparación entre grupos y comparación intra sujetos, pues combina en un mismo experimento el procedimiento de grupos independientes y la utilización de sujetos de control propio. Se utiliza comúnmente en investigaciones donde se encuentren presentes a lo menos dos variables independientes.

En este caso particular, se utilizará una metodología de diseño experimental elaborado entre grupos simples para múltiples grupos, debido a que se realizará la comparación entre grupos de clientes en los que cada uno poseerá un distinto nivel de la variable independiente “experimentación”. De esta forma, un grupo poseerá *experimentación 1*, el segundo grupo tendrá la *experimentación 2* y, finalmente, un grupo tendrá el nivel de *experimentación cero*, que corresponde al grupo de control.

6.5 Evaluación estadística

La evaluación estadística se realiza mediante un test de comparación de medias mediante el estadístico F de Fischer para medir el análisis de varianza, a través de las medias. El test F mide la relación de las medias cuadráticas entre grupos en comparación con las medias cuadráticas intragrupos, como se presenta en la tabla 1.

El test de medias tiene como hipótesis nula $\mu_i = \mu_j$, es decir, las medias de los grupos de tratamiento son iguales.

6.6 Modelos de venta incremental (*uplift*)

Son modelos que buscan caracterizar el comportamiento de los consumidores mediante la respuesta de ellos frente a la intervención de experimentos. Cabe destacar que en este modelo se calcula la respuesta de un cliente en términos relativos, por medio del cambio en la probabilidad de compra condicional a la intervención mediante experimentación o no.

$$P(\text{Compra}) = P(\text{Compra}|\text{Tratamiento}) - P(\text{Compra}|\text{No Tratamiento})$$

Los modelos de venta incremental se basan en modelos estructurados para el cálculo de las probabilidades en cuestión. Dentro estos, en los que se basan los modelos *uplift*, se encuentran:

- Modelos basados en árboles de decisión
 - Los modelos basados en árboles de decisión engloban a los árboles de decisión en sí y a *random forest*. Estos clasifican y ponderan las distintas variables incluidas para poder estimar la probabilidad en cuestión.
- Modelos basados en regresiones
 - Utilizan regresiones, como el modelo *logit*, para poder asignar un peso a las variables y con esto calcular la probabilidad deseada.

Como se ha mencionado, luego de calcular las probabilidades, independiente del modelo subyacente, se estima la diferencia de comportamiento debido a la existencia de intervención y con esto se pueden caracterizar a los clientes que tengan un mayor cambio en su probabilidad de compra.

7 Marco metodológico

La metodología que se desarrollará para poder solucionar el problema planteado y así lograr los objetivos definidos con anterioridad, se divide en nueve etapas:

7.1 Revisión bibliográfica

Antes de comenzar cualquier tipo de desarrollo de proyecto, es necesario informarse respecto de todos los ámbitos relevantes, como por ejemplo, investigaciones existentes en temas similares al que se desea tratar, con el fin de informarse adecuadamente y no volver a realizar estudios ya hechos o generar una innovación en las metodologías ya utilizadas. Además, en este paso se considera la incorporación de información acerca de la metodología y conceptos relevantes para el óptimo desarrollo del proyecto, con el propósito de proponer una solución al problema de negocios. En este caso se revisó bibliografía de tesis o memorias anteriores, documentos académicos sobre metodologías adecuadas, entre otros.

El objetivo de esta sección es poder ampliar el conocimiento y definir objetivos, alcances, metodología, marco conceptual y todos los detalles relevantes para el desarrollo óptimo del proyecto, además de lograr una aproximación a la problemática que se desea abordar y analizar cómo se ha abordado anteriormente.

7.2 Estructura y pre-procesamiento de base de datos

Una vez determinada la estructura metodológica que guiará el proceso, es necesario comenzar el trabajo con los datos que se poseen. Para lo anterior, es necesario realizar un análisis descriptivo inicial de los datos para poder detectar comportamiento de estos y eliminar información que no sea confiable (*outliers*) o que ensucie la base de datos propiamente tal, dado que estas no siempre poseen una buena estructura o calidad de información. Este paso se recoge de la metodología *KDD (Knowledge Discovery database)*.

7.3 Cálculo de probabilidad de compra (*Propensity Score*)

Mediante el uso de regresión logística, se procederá a calcular la probabilidad de compra de los clientes en función de las variables que sean relevantes y presenten significancia adecuada en el modelo en cuestión.

Se utilizará este método (regresión logística), puesto que es de simple comprensión y entrega información respecto de las variables relevantes para poder calcular el *Propensity Score (PS)*. Dentro de las variables que se pretende analizar para poder entender el comportamiento de compra de los clientes, se encuentran: género del cliente, edad, tiempo entre registro en el sitio y primera compra, tiempo entre primera y segunda compra (si el cliente compro dos veces), tiempo entre el día de hoy y su *recency* (último *login*) en el sitio, categoría y subcategoría de producto que compró, nota de evaluación de encuesta de satisfacción, monto de compra e interacción entre variables.

7.4 Segmentación

Se segmentarán los clientes en función de su probabilidad de compra, calculado anteriormente (*Propensity score*). Los segmentos serán tres dependiendo del valor del *PS*: alto, medio y bajo. Lo anterior se realiza con el fin de poder entender los resultados posteriores y establecer si son esperables o no, dada su condición *a priori*.

Además de segmentar en función de su *PS*, se determinarán cuatro grupos experimentales, uno para cada experimento y uno para control. Cada grupo se conformará por individuos asignados aleatoriamente, pero equilibrados según su *PS*.

7.5 Diseño de experimentos

Se deben determinar y diseñar los experimentos que se realizarán a los clientes. Para esto, se debe definir en conjunto con la empresa cuál es el objetivo, atributos, ideas y disponibilidad para poder efectuar el proceso experimental. Definidas las variables que se manejarán y cuáles serán las intervenciones que se realizarán, se deben diseñar los correos que se enviarán a los usuarios con las intervenciones, pues estas serán realizadas por este medio.

Se pretende realizar dos experimentos, donde el primero corresponderá a un sistema de personalización y recomendación simple y el segundo consistirá en un sistema de fidelización que premiará a los clientes que concreten 10 compras válidas.

7.6 Experimentación

Con los experimentos diseñados y con la confección de los correos adecuada, se procederá a implementar dicha experimentación y a intervenir a los clientes según su grupo asignado. A los usuarios se les enviará un mail con la intervención adecuada

una vez por semana, durante 8 semanas, desde el día 26 de julio, donde en cada semana se modificará el contenido del mail dependiendo del estado actual del envío. Se cambiará el *deal* ofrecido para el experimento 1 y se actualizará el estado de compras del cliente para el experimento 2. No se cambiará el tipo de experimento para los clientes asignados a los grupos. El grupo de control no recibirá ningún tipo de intervención.

7.7 Evaluación y análisis de resultados de intervenciones experimentales

Posterior a la realización de las intervenciones experimentales, se procederá a la evaluación de indicadores de interés que estudiará la empresa. Para la estimación del impacto, se medirán las diferencias percibidas en los indicadores entre los distintos niveles de tratamiento.

Los indicadores relevantes para la evaluación son: número de compras y activación de clientes.

La activación de clientes corresponde a la cantidad de consumidores que realizaron compras posteriores a la intervención dentro del total de usuarios del grupo de tratamiento.

Dado el tipo de diseño experimental, es posible evaluar las diferencias de medias, a través de un test de comparación de medias que permite medir la significancia de los efectos principales y las diferencias entre los distintos niveles de tratamiento.

Con el análisis estadístico desarrollado, se procederá a interpretar los resultados sobre la base del comportamiento del consumidor.

7.8 Confección de modelo *uplift*

Con las probabilidades de compra de los clientes, condicionadas por la intervención de los experimentos (según corresponda), es posible desarrollar un modelo de venta incremental que podrá definir los segmentos de clientes que son propensos a modificar su comportamiento debido a la intervención, dicho de otro modo, se tendrá información respecto de los subsegmentos que pueden cambiar su comportamiento para así poder determinar a quienes intervenir en el futuro, con el fin de optimizar la inversión realizada y no intervenir a los clientes que no responderán positivamente frente a estos estímulos. Esto es relevante en la medida en que ayudará a entregar lineamientos futuros a la empresa para instaurar posibles políticas de intervención y retención de clientes.

7.9 Conclusiones y recomendaciones

Finalmente, con las evaluaciones de las intervenciones y la determinación de subsegmentos propensos a comprar, se generará un reporte con el resumen de las intervenciones y variables relevantes en la probabilidad de compra de los clientes, para informar a la empresa y entregar lineamientos para posible implementación de estrategias en el futuro.

Lo anterior se realizará para clientes activos (con al menos una compra después del 01 de noviembre de 2014).

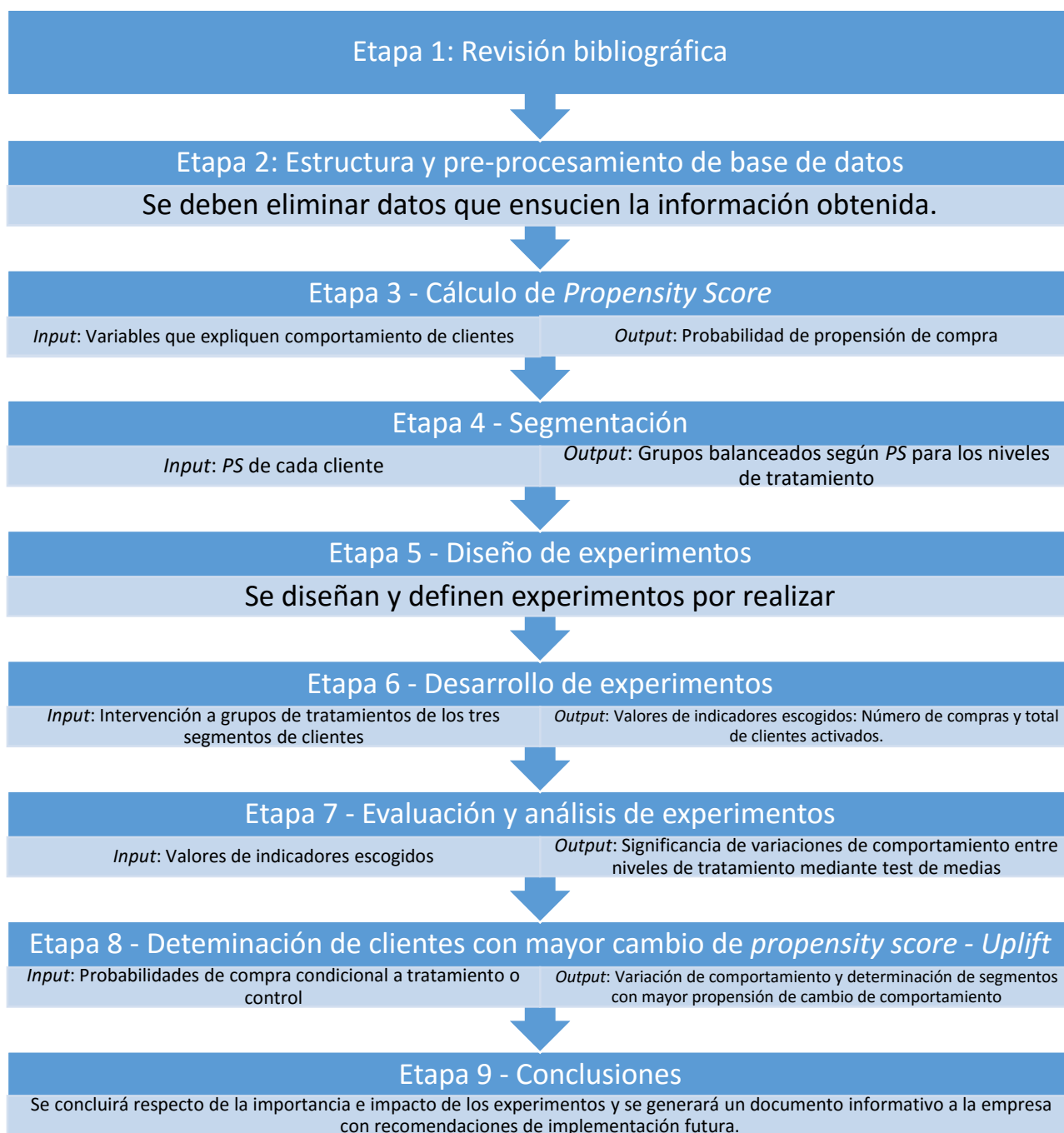


Ilustración 1: Esquema de desarrollo metodológico. Fuente: Elaboración propia

8 Descripción de la situación actual

Como ya se ha señalado anteriormente, la empresa pertenece al rubro del comercio *online*, por lo que se posee una vasta base de datos de información transaccional de los clientes, sin embargo, esta es menos amplia en términos de información demográfica de los usuarios.

8.1 Antecedentes generales

Cuponatic es una empresa que se dedica a la venta de cupones *online*, de modo que ofrece ofertas diarias con descuentos importantes, con la pretensión de brindar a los

clientes las mejores opciones del mercado. Las empresas de este rubro, se aprovechan de las plataformas *online* para poder mejorar su *marketing* aumentando su demanda a través del sitio de descuentos. Finalmente, los clientes son los grandes beneficiados al recibir productos o servicios a precios menores, gracias a los descuentos ofrecidos por el sitio.

En la actualidad, la empresa se encuentra dentro de las cuponeras *online* con mayor participación de mercado junto con *Groupon*.

Cuponatic ofrece 6 categorías principales: Productos, *health & beauty*, gastronomía, entretenimiento, viajes y otros abarcando una amplia gama de ofertas a los clientes y se espera que en el futuro la oferta crezca aún más, convirtiéndose en un *market place* y no solo en una cuponera *online*.

8.2 Análisis descriptivo de clientes

Para el análisis descriptivo, se considerarán los clientes activos de la empresa, entendidos como aquellos que han comprado al menos una vez en los 8 meses previos al día en que se comienza el proceso de limpieza y formato de la base de datos, en este caso, usuarios con compras entre noviembre del año 2014 y junio del año 2015. Se eliminaron los clientes que tuviesen compras en julio, debido a su reciente incorporación al sitio. En total se cuentan 120.326 clientes que cumplen dichas condiciones sobre las que se realizó el análisis descriptivo inicial.

8.2.1 Análisis de clientes por número de compras

Dado que el objetivo principal del estudio corresponde a la retención y fidelización de clientes, resulta pertinente analizar el estado inicial de los clientes en relación con el número de compras. A continuación, se presenta el histograma de los clientes en función del indicador ya mencionado.

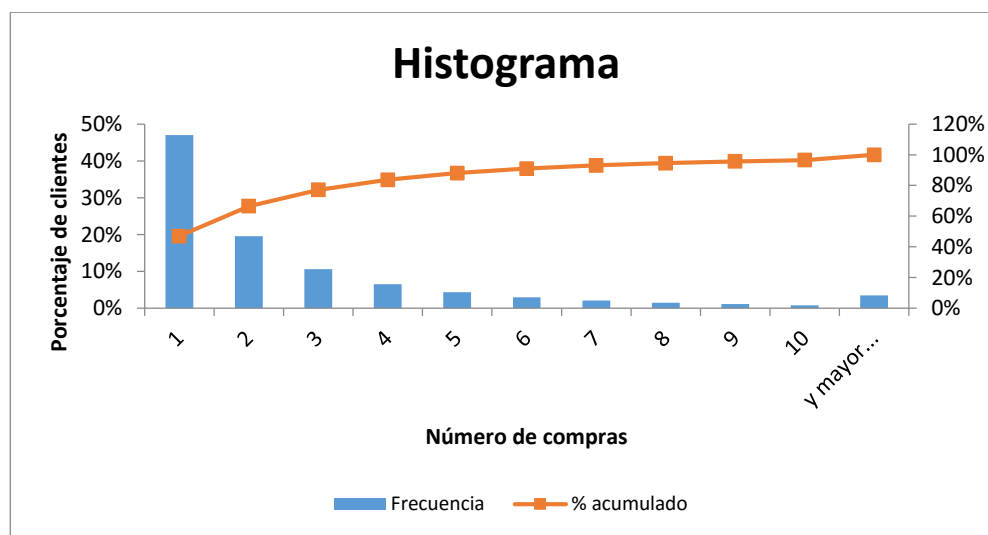


Ilustración 2: Histograma de compra clientes activos. Fuente: Elaboración propia

Dentro del histograma, se destaca que se mantiene la tendencia histórica de comportamiento mencionada en secciones anteriores, donde el 66,55% de los clientes posee menos de tres compras (46,97% de los clientes poseen una compra y 19,58% poseen dos compras históricas).

Es importante señalar que el análisis descriptivo posterior contrastará las distintas variables de interés con la proporción de buenos clientes en cada variable, en cuyo caso se define como buen cliente todo aquel que posea más de dos compras en el sitio.

8.2.2 Análisis de clientes por género

La empresa cuenta con una marcada tendencia de distribución de género, en la que existe un marcado dominio de población femenina dentro de su cartera de clientes, representando al 62,51% del total. Lo anterior se debe principalmente a que el sitio se enfoca en la oferta de servicios de belleza como una de sus dos principales líneas de negocios, lo cual se refleja en el gráfico adjunto.

Adicional a la distribución de la población, dentro del total es posible apreciar la proporción de buenos clientes, con más de dos compras por categoría, donde se destaca que el género femenino posee un comportamiento más deseable que el masculino, en tanto presentan una mayor proporción de buenos clientes. Sin embargo, a la vez se aprecia que esta diferencia es muy pequeña (tan sólo un 0,66% entre ellos).

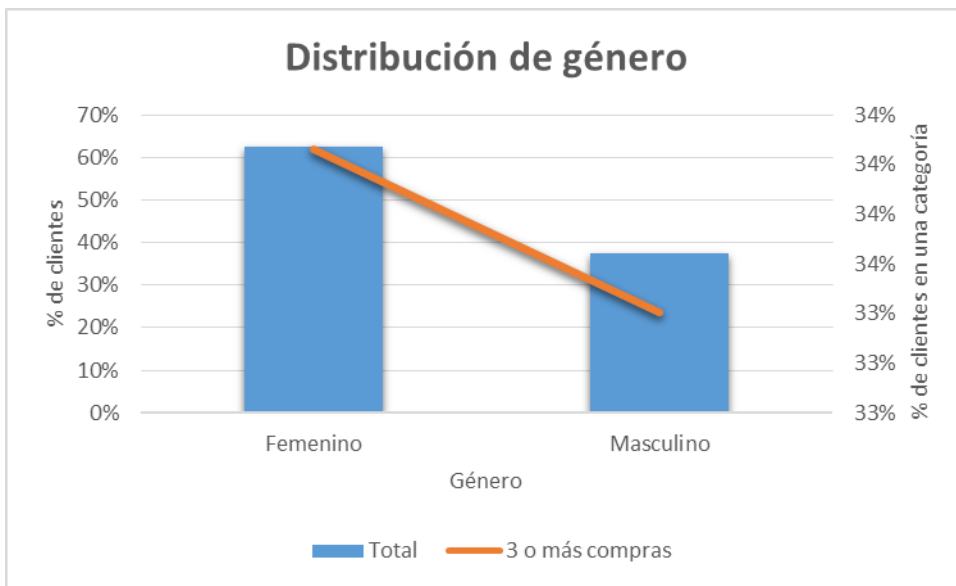


Ilustración 3: Distribución de género en clientes activos. Fuente: Elaboración propia

8.2.3 Análisis de clientes por edad

Para este análisis, se eliminaron todos los clientes que no tuviesen fecha de nacimiento registrada y aquellos que tuvieran fechas de nacimiento irreales, en particular, se consideran clientes con edades entre los 15 y los 85 años.

El público de la empresa posee tendencias a ser adulto joven, dominado por clientes con edad entre 25 y 35 años, que representan al 46% del total.

Dentro de los segmentos etarios, es posible ver qué proporción de buenos clientes es creciente con la edad, es decir, aparentemente los clientes jóvenes realizan compras más impulsivas y aprovechan las buenas ofertas sin tener mayor relación con el oferente, pues su interés es la oferta en sí. Un comportamiento diferente presentan los segmentos de mayor edad, puesto que se podría suponer que ellos usualmente buscan optimizar sus recursos y vuelven al sitio en busca de buenas ofertas, principalmente en productos de gran consumo.

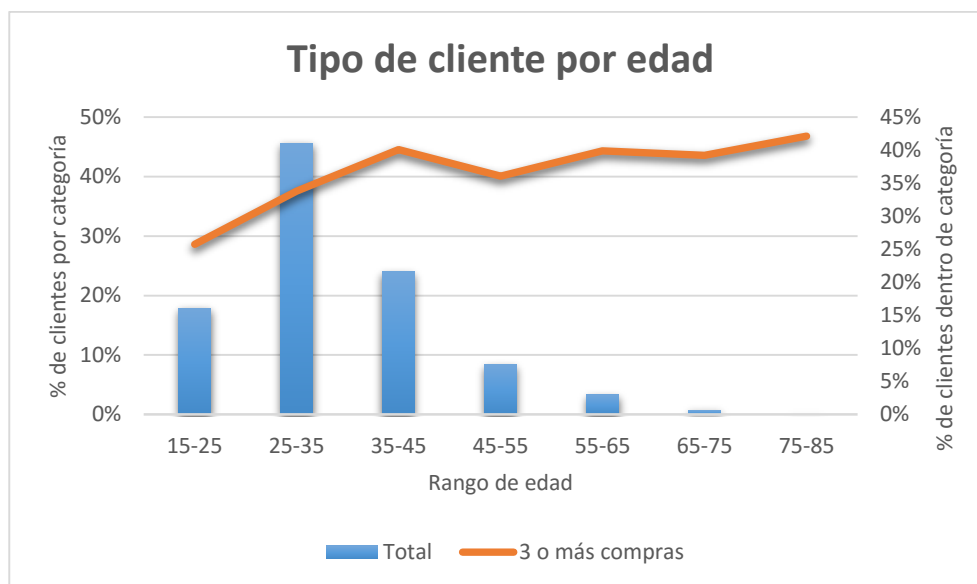


Ilustración 4: Distribución de clientes según tramo de edad. Fuente: Elaboración propia

8.2.4 Análisis de clientes por tiempo de activación

Es posible analizar a los clientes en función de su *tiempo de activación*. Esta noción se define como el tiempo que transcurre entre su registro como cliente en el sitio y la realización de su primera compra.

En función de esta variable, es posible apreciar que existen dos tipos de buenos clientes. El primero corresponde a aquellos que compraron muy cerca de su registro, lo que se puede entender como una masa de clientes que se sintieron atraídos por el sitio desde un comienzo y mantuvieron su relación, porque tuvieron una buena experiencia de compra.

Un segundo comportamiento evidenciado se ubica en el otro extremo, es decir, aquellos clientes que se demoraron más de 6 meses en realizar su primera compra. Lo anterior se traduce en que existe una masa de clientes que ingresaron al sitio, se registraron, pero no observaron nada de interés para consumo inmediato. No obstante, sí se interesaron en las ofertas, por lo que, transcurrido el tiempo, volvieron a ingresar al sitio y decidieron adquirir algún producto o servicio, puesto que sí se logró captar su atención mediante las ofertas.

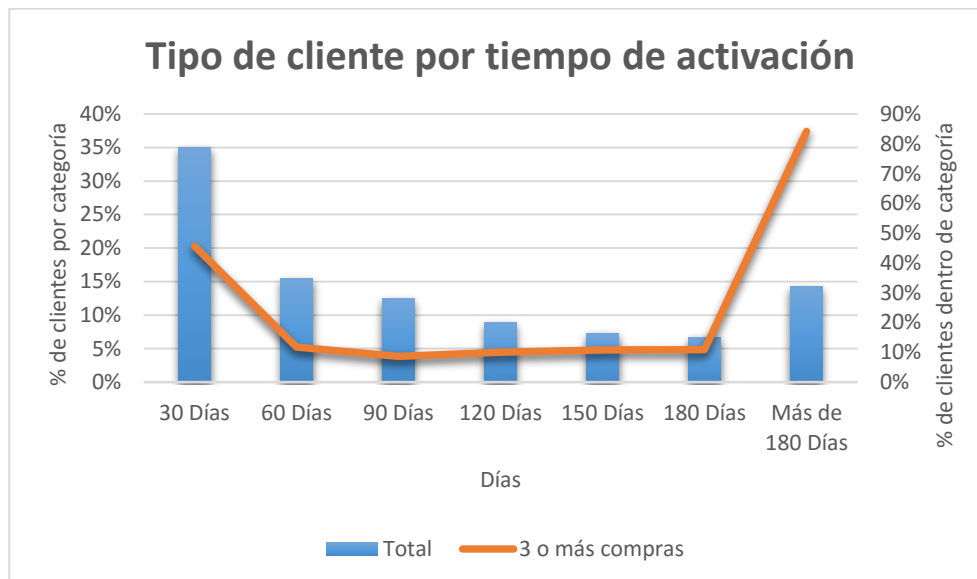


Ilustración 5: Distribución de clientes según tiempo de activación. Fuente: Elaboración propia

8.2.5 Análisis de clientes por tiempo entre compras

En primer lugar, se señala que lo reflejado en este gráfico corresponde al promedio de tiempo entre compras por cliente, sin considerar los que sólo posean una compra, debido a que este es un valor inexistente.

Dados los resultados, es posible advertir que este comportamiento no posee una tendencia lineal, sino más bien ofrece un comportamiento distinto en los extremos, lo que permite interpretar que los clientes que compran con demasiada distancia entre sus transacciones lo hacen por eventos especiales y en busca de algún tipo de descuento especial. Así, no reflejan necesariamente una relación estable con el sitio.

Por otro lado, los clientes que compran con frecuencia muy alta son usuarios que tuvieron una muy buena impresión del sitio, dicho de otro modo, se generó una conexión fuerte con el cliente para que compre con alta frecuencia, aunque no lo suficientemente fuerte como para hacer sostenible en el tiempo esta relación.

El mayor valor de participación se encuentra para los clientes que compran cada dos o tres meses, lo que significa que estos visitan frecuentemente el sitio y realizan compras periódicamente y no sólo por oportunidad o impulsividad.

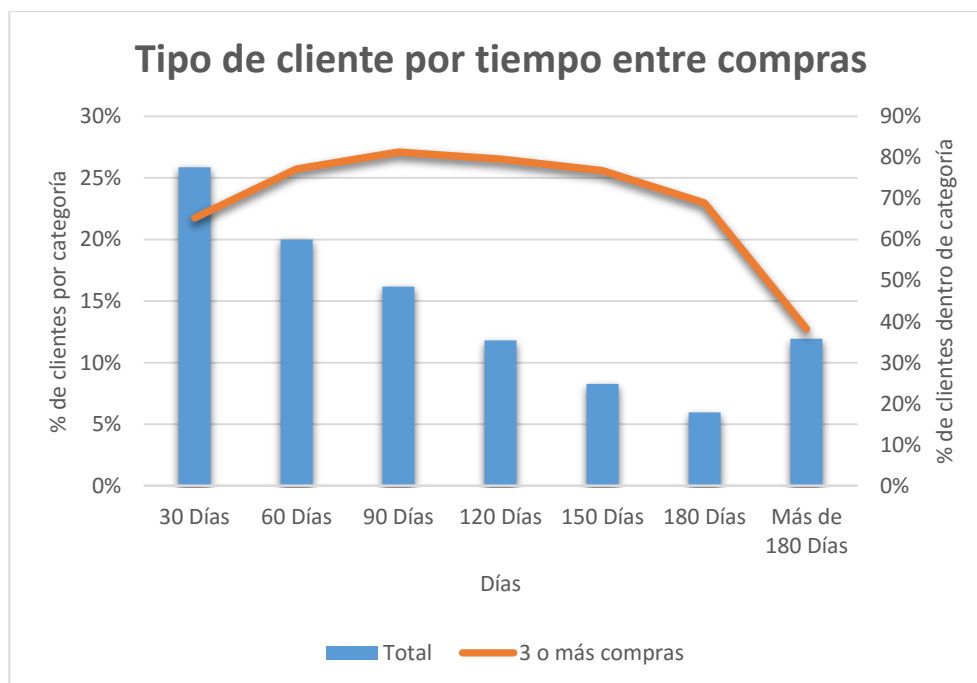


Ilustración 6: Distribución de clientes según tiempo entre compras. Fuente: Elaboración propia

8.2.6 Análisis de clientes por categoría de primera compra

Las compras realizadas en el sitio se encuentran catalogadas por dos niveles de agregación: *categoría* y *subcategoría*. La primera engloba a las seis principales líneas de negocios del sitio. Por otro lado se cuenta con 47 subcategorías distintas que evidencian un nivel mayor de especificidad de las compras realizadas por los clientes. En la ilustración 7 se presenta la estructura de categorías y subcategorías del sitio.

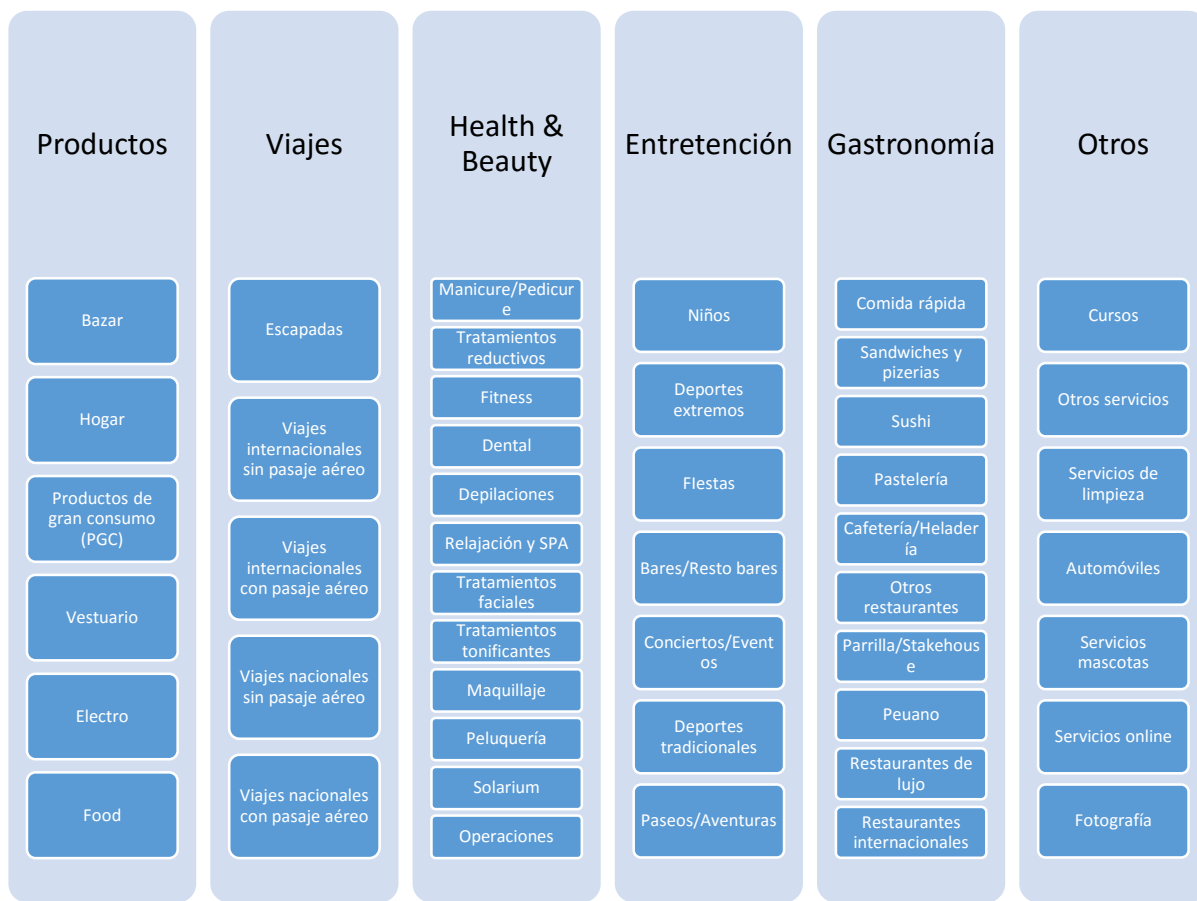


Ilustración 7: Benchmark de categorías y subcategorías del sitio.

Las categorías de productos y *Health & beauty* son las dos principales categorías de negocio del sitio que en conjunto reúnen al 84% de las compras realizadas en la primera oportunidad del cliente. Es por esto que se considera que este volumen representa un comportamiento importante en la población. Además de lo anterior, existe una tercera categoría que representa un comportamiento inusual, que corresponde a gastronomía. Dicha clase posee la mayor proporción de buenos clientes, pero su peso relativo es bajo, solo 11% del total, por lo que es necesario considerar la evaluación del impacto de la categoría en la capacidad de generar relación entre cliente y empresa.



Ilustración 8: Distribución de clientes según categoría de primera compra. Fuente: Elaboración propia

8.2.7 Análisis de clientes por subcategoría de primera compra

Dentro del análisis de la subcategoría relacionada con la primera compra del cliente, es posible evidenciar que existen cinco subcategorías que abarcan una mayor proporción de preferencias, de las cuales una se destaca considerablemente: “Productos de gran consumo”.

A pesar de lo anterior es posible ver que existen tres subcategorías que contienen la mayor proporción de buenos clientes: “Sushi”, “Comida rápida” y “Sandwich y pizzerías”.

Es posible relacionar estos tres *peaks* de concentración con la realidad evidenciada por categoría, pues estas están dentro de la categoría “Gastronomía”, por lo que el análisis realizado previamente resulta consecuente.

Los tres *peaks* tienen sentido, en tanto al ser un servicio que es externo al sitio, por lo general, se asocia a la reputación del emisor del cupón (empresa oferente) y los clientes podrían aprovechar de obtener un descuento por un servicio ya conocido y sin riesgo.

Lo interesante de analizar es el peso relativo de las principales categorías en la fidelidad del cliente (las cinco subcategorías principales), ya que existe un alto porcentaje de la población que llega al sitio por este tipo de ofertas.

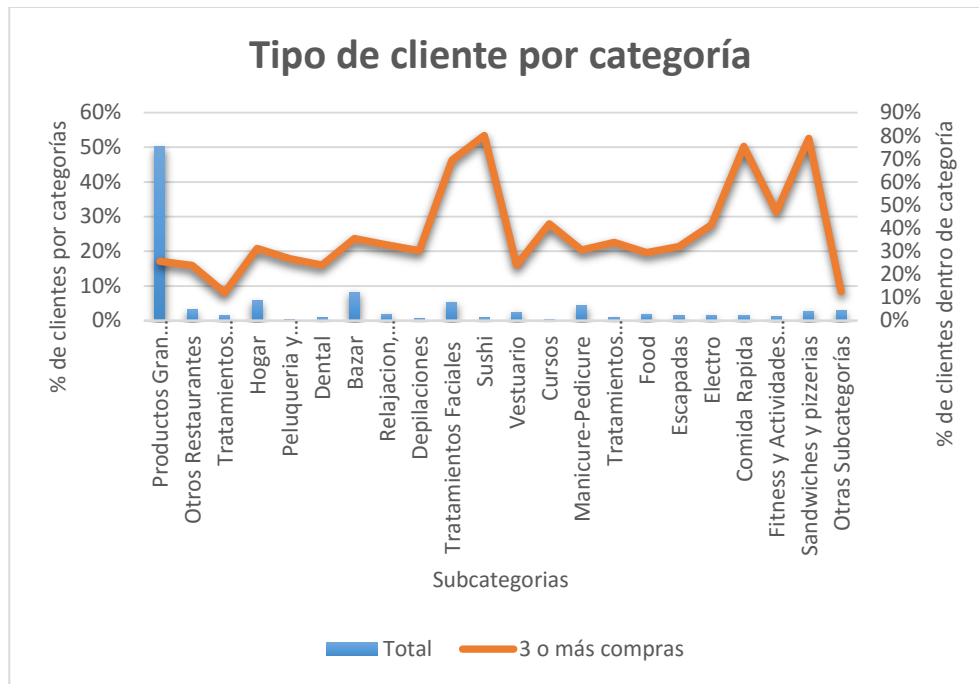


Ilustración 9: Distribución de clientes según categoría de primera compra. Fuente: Elaboración propia

9 Desarrollo de modelos

9.1 Estructura y pre-procesamiento de base de datos

Para el análisis descriptivo anterior, se consideraron los clientes activos de la empresa donde se define activo a un cliente que ha comprado al menos una vez en los 8 meses previos a partir del día de realización del comienzo del proceso de limpieza y formato de la base de datos, en este caso clientes con compras entre noviembre del año 2014 y junio del año 2015. Se eliminaron los clientes que tuviesen compras en julio debido a su reciente incorporación al sitio como cliente. En total se cuentan 120.326 clientes que cumplen dichas condiciones sobre las que se realizó el análisis descriptivo inicial.

De esta base de clientes, se eliminaron 2.210 registros que no tenían información de género, además de 12.869 clientes que tenían más de 5 compras. Este filtro se realiza, dado que no se desea intervenir a buenos clientes, solo aquellos que tengan más de dos compras para ser incluidos en la lista de usuarios que serán parte del experimento 2, dados los requerimientos de la empresa. Adicionalmente a lo anterior, se eliminan 6.221 clientes que no poseen fecha de registro en el sitio y, finalmente, se prescinde de 37.720 consumidores que poseen edades poco creíbles (menos de 15 años o más de 85 años), lo cual se debe a problemas de implementación del registro de clientes en los primeros años de la empresa.

Los filtros y el formateo de la base de datos se justifica en la medida en que se requiere información de estas variables que pueda ser incorporada a la regresión logística *a priori*, de modo de no tener celdas vacías.

Lo anterior da como resultado una base de datos de 61.306 clientes con información válida.

Posterior a la limpieza de la base de datos, se procede a reestructurar dicha base para poder incluirse al modelo de regresión lineal. Esto se realiza transformando la información de categoría y subcategoría de la primera compra realizada por los clientes a variables binarias, al igual que el género, que adquiere valor 1 cuando el cliente es de género femenino y 0 si es masculino.

9.2 Cálculo de *propensity score*

Para desarrollar el cálculo del *propensity score* se incluye la información de toda la base de datos mencionada anteriormente ya limpia y con el formato adecuado.

Para el correcto desarrollo del modelo, se incorporan todas las variables consideradas en el análisis descriptivo, pues representan las de mayor interés. Con estas, se desarrollaron 8 modelos que incorporan más variables cada vez y se comparan los resultados para determinar el que representa de mejor manera el comportamiento de los clientes. Se comienza con un modelo simple que incluye las principales categorías como variables independientes, las categorías son “Productos”, “*health & beauty*” y “gastronomía”, posteriormente se agregan las principales subcategorías asociadas a las anteriores, como “hogar”, “productos de gran consumo”, “otros restaurantes” y “tratamientos reductivos”, de las cuales las dos primeras están relacionadas con la categoría “productos”, la siguiente se asocia con gastronomía y la última subcategoría, se desprende de *health & beauty*.

En los siguientes modelos, se incorporan variables demográficas de los clientes como edad y género. El cuarto modelo incorpora, además, el monto desembolsado en la primera compra, que deja de considerarse en el quinto por su valor y significancia, al igual que la subcategoría “tratamientos reductivos”. Para el siguiente modelo, se incorpora la frecuencia de compras y se genera la interacción entre la variable binaria género (1 cuando el género del cliente es femenino) y la subcategoría “tratamientos reductivos”. El sexto modelo incorpora la variable activación, correspondiente al tiempo entre el registro del cliente en el sitio y su primera compra; y se cambia la interacción de género con subcategoría por la interacción entre el género femenino y la categoría “*health & beauty*”. Posteriormente, se agrega la información de *recency*, que corresponde al tiempo transcurrido desde el último ingreso al sitio. Finalmente, se elimina la variable género y se transforma la variable edad en el logaritmo de dicha variable, por el comportamiento observado de esta en el análisis descriptivo.

Para el cálculo de la regresión, se considera que la variable dependiente, o variable que se buscará predecir, es la probabilidad de que un cliente posea más de 2 compras en su historia en función de las variables adecuadas, es decir, la variable independiente de la regresión es 1 si el cliente posee más de 2 compras y 0 en caso contrario.

El resumen de los modelos se presenta en las Tablas 1 y 2, en la que se destaca la comparación de los diferentes modelos mediante dos ámbitos principales: significancia individual de las variables y valor de *AIC* (*Akaike information criterion*). Para el primer ámbito, es deseable que todas las variables sean

significativas, al menos al 95% de confianza, lo cual se encuentra representado en las Tablas 1 y 2, en la columna *Sig.* de cada modelo. De esto es posible desprender que los modelos 1, 5 y 8 coinciden en que todas sus variables son significativas a más de un 95% de confianza. Para el segundo ámbito, es deseable que el mejor modelo presente el menor *AIC*, pues este indicador considera el *trade-off* entre la bondad de ajuste del modelo y la complejidad del mismo, mediante la fórmula $AIC = 2k - 2\ln(L)$, donde k representa el número de parámetros del modelo y L , la verosimilitud, es decir, a medida que se incorporan variables, el término *AIC* debería crecer, penalizando la complejidad para evitar el sobreajuste del modelo representado por la verosimilitud (L). Sobre la base de esta última información, es posible ver que cada modelo desarrollado va mejorando con este índice, con excepción del modelo 3 y 4, que mantienen el valor de *AIC*, vale decir, cada variable incorporada agrega información para la estimación realizada.

Por último, es posible ver que, si se consideran los tres modelos (modelo 1, modelo 5 y modelo 8) con variables significativas individualmente y se comparan mediante el criterio de *AIC*, es evidente que el modelo 8 posee el mejor ajuste, por lo tanto, será el adecuado para la estimación de *propensity score*.

9.2.1 Nomenclatura de variables en modelo

Se utilizará la siguiente nomenclatura para las variables que serán incorporadas al modelo *logit*, con el cual se calculará el *propensity score*.

La variable dependiente corresponde a la variable binaria 1 si el cliente posee más de 2 compras y 0 en caso contrario (posee 1 o 2 compras).

- Género: Variable binaria que vale 1 si el género del cliente es femenino y 0 si es masculino.
- Edad: Variable continua que presenta la edad de los clientes.
- Monto 1 OC: Monto de la primera compra realizada por el cliente.
- *Recency*: Variable que representa el tiempo transcurrido desde la última vez que el cliente se registró en el sitio.
- Frecuencia: Promedio en días entre las compras realizadas por el cliente.
- Activación: Tiempo transcurrido desde el registro del cliente en el sitio y su primera compra.
- *C_Productos*: Variable binaria que toma valor 1 si el cliente, en su primera compra, adquirió un producto de la categoría “Productos”.
- *C_H&B*: Variable binaria que toma valor 1 si el cliente, en su primera compra, adquirió un servicio de la categoría “*Health & beauty*”.
- *C_Gastronomía*: Variable binaria que toma valor 1 si el cliente, en su primera compra, adquirió un servicio de la categoría “Gastronomía”.
- *S_Hogar*: Variable binaria que toma valor 1 si el cliente, en su primera compra, adquirió un producto de la subcategoría “Hogar”.
- *S_Otros restaurantes*: Variable binaria que toma valor 1 si el cliente, en su primera compra, adquirió un servicio de la subcategoría “Otros restaurantes”.
- *S_PGC*: Variable binaria que toma valor 1 si el cliente, en su primera compra, adquirió un producto de la subcategoría “Productos de gran consumo”.

- *S_Tratamientos reductivos*: Variable binaria que toma valor 1 si el cliente, en su primera compra, adquirió un servicio de la subcategoría “Tratamientos reductivos”.
- *Género x Tratamientos reductivos*: Variable binaria que toma valor 1 si el cliente es de género femenino y si, en su primera compra, adquirió un bien de la subcategoría “Tratamientos reductivos”.
- *Género x H&B*: Variable binaria que toma valor 1 si el cliente es de género femenino y si, en su primera compra, adquirió un bien de la categoría “Health & beauty”.
- *Log(edad)*: Variable que evidencia transformación no-lineal de la edad mediante el logaritmo natural de esta última. Se utiliza el logaritmo con el fin de quitar peso a los valores extremos superiores y aumentarlo a los valores inferiores.

Variable	Modelo 1			Modelo 2			Modelo 3			Modelo 4		
	Coef	Err Std	Sig	Coef	Err Std	Sig	Coef	Err Std	Sig	Coef	Err Std	Sig
Intercepto	-1,156	0,045	***	-1,165	0,045	***	-1,812	0,056	***	-1,780	0,059	***
Género							0,110	0,019	***	0,110	0,019	***
Edad							0,017	<0,001	***	0,017	0,001	***
Monto 1 OC										0,000	<0,001	***
Recency												
Frecuencia												
Activación												
C_Productos	0,592	0,046	***	2,678	0,063	***	2,653	0,063	***	2,630	0,065	***
C_H&B	0,392	0,049	***	0,309	0,050	***	0,293	0,050	***	0,268	0,053	***
C_Gastronomia	0,648	0,051	***	0,801	0,066	***	0,831	0,066	***	0,804	0,068	***
S_Hogar				-0,551	0,060	***	-0,581	0,060	***	-0,580	0,060	***
S_Otros Restaurant				-0,208	0,056	***	-0,207	0,056	***	-0,208	0,057	***
S_PGC				-2,513	0,045	***	-2,519	0,046	***	-2,520	0,046	***
S_Tratamientos Reductivos				0,073	0,042	.	0,062	0,042		0,006	0,042	
Género x S_Tratamientos Reductivos												
Género x C_H&B												
Log(Edad)												
AIC	75,232			69,386			69,014			69,013		

Tabla 1: Resumen coeficientes regresiones. Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. Fuente: Elaboración propia con software R.

Variable	Modelo 5			Modelo 6			Modelo 7			Modelo 8		
	Coef	Err Std	Sig	Coef	Err Std	Sig	Coef	Err Std	Sig	Coef	Err Std	Sig
Intercepto	-1,740	0,056	***	-1,770	0,046	***	-1,380	0,047	***	-3,180	0,127	***
Género				0,050	0,022	*	0,041	0,022	.			
Edad	0,017	0,001	***	0,016	0,001	***	0,017	0,001	***			
Montro 1 OC												
Recency							-0,005	0,000	***	-0,005	0,000	***
Frecuencia	0,000	<0,001	***	-0,001	<0,001	***	-0,001	<0,001	***	-0,001	<0,001	***
Activación				0,000	<0,001	***	0,001	<0,001	***	0,001	<0,001	***
C_Productos	2,660	0,063	***	2,530	0,055	***	2,610	0,056	***	2,620	0,056	***
C_H&B	0,319	0,050	***									
C_Gastronomía	0,834	0,067	***	0,721	0,059	***	0,748	0,060	***	0,766	0,060	***
S_Hogar	-0,573	0,060	***	-0,556	0,061	***	-0,523	0,062	***	-0,520	0,062	***
S_Otros Restaurant	-0,202	0,057	***	-0,200	0,058	***	-0,256	0,059	***	-0,253	0,059	***
S_PGC	-2,510	0,046	***	-2,470	0,047	***	-2,550	0,048	***	-2,540	0,048	***
S_Tratamientos Reductivos				0,064	0,042		0,172	0,048	***	0,165	0,043	***
Género x S_Tratamientos Reductivos	0,127	0,047	**									
Género x C_H&B				0,267	0,041	***	0,263	0,042	***	0,300	0,038	***
Log(Edad)										0,685	0,035	***
AIC	68,940			67,232			65,902			65,845		

Tabla 2: Resumen coeficientes regresiones. Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. Fuente: Elaboración propia con software R.

9.2.2 Interpretación de coeficientes

Teniendo los coeficientes estimados por el modelo definido anteriormente, es necesario realizar una primera comprobación del sentido y magnitud de dichos coeficientes, a través de la interpretación de estos.

En primer lugar, se aprecia que el *intercepto* es negativo. Este punto representa la probabilidad intrínseca de cada cliente de llegar a ser un buen cliente, y se hace esperable que tenga un valor negativo dadas las características de los usuarios del sitio, que justifican el problema de negocios que se enfrenta en este proyecto.

El *recency* y frecuencia son negativos, pero con un valor muy bajo, es decir, afectan negativamente la probabilidad de ser un buen cliente, puesto que si un cliente no ha tenido interacción con el sitio es poco probable que sea un buen cliente y, además, si este compra con mucha distancia temporal quiere decir que no posee una relación cercana con el sitio, disminuyendo así su probabilidad de ser un buen cliente. La magnitud es esperable, pues ambas mediciones son en días, por lo que el impacto crece sustancialmente a medida que el tiempo avanza. La activación posee una magnitud similar a las variables anteriores, pero con signo positivo. Esto es esperable, pues en la medida que un cliente se demora más en activarse, quiere decir que este no es oportunista; sino que revisa el sitio hasta encontrar su mejor opción y, al igual que las variables anteriores, al ser medida en días, se espera que la magnitud del coeficiente sea baja.

Posteriormente, las variables de categoría y subcategoría hablan de las preferencias de los clientes por ciertas líneas de negocios y entregan mayor valoración a los productos, pues estos pueden ser de bajo precio y de fácil adquisición, ya que que no requieren un involucramiento superior, a diferencia de otras líneas de negocio. Además, los productos de hogar y de gran consumo poseen coeficientes negativos, pues pueden corresponder a compras más impulsivas. El caso de “*health & beauty*” es esperable también, debido a que posee un valor mayor para mujeres que para hombres, dadas sus características.

9.2.3 Evaluación del modelo

Además de los valores estimados del modelo y del ajuste del mismo, es necesario evaluar si es o no un buen modelo, es decir, si su capacidad de predicción es adecuada. Para lo anterior, se generan dos elementos que ayudan a evaluar la capacidad predictiva del modelo: *Matriz de confusión* y *curva ROC*.

Para la realización de esta evaluación, se efectuó una prueba de predicción considerando los clientes y sus compras hasta el 31 de enero de 2015 (adelantando el proceso 5 meses) y se contrastó con las compras realizadas por los mismos por dos meses, hasta el 30 de abril de 2015. Posteriormente, se corroboró la capacidad de predicción del modelo, para lo cual se generó una matriz de confusión en la que se contabilizan la cantidad de aciertos y errores del modelo sobre las reales realizaciones. La matriz de confusión generada se presenta en la tabla 3, donde son apreciables los contrastes entre valores predichos y valores reales, tomando como valor umbral el de 30%. Dada la distribución de probabilidades obtenidas hasta ese punto, se explica más del 80% de los valores obtenidos.

		Predicted		Total
		Buen cliente	Mal cliente	
Real	Buen cliente	VP=16.015	FN=4.399	P=20.414
	Mal cliente	FP=6.620	VN=21.422	N=28.042
Total		p=22.635	n=25.821	

Tabla 3: Matriz de confusión, umbral 30%. Fuente: Elaboración propia.

Dentro de la tabla, se mencionan los valores VP, FP, FN y VN que corresponden a *verdaderos positivos*, *falsos positivos*, *falsos negativos* y *verdaderos negativos*, respectivamente, dicho de otro modo, corresponden a la cantidad de veces que se observó un valor positivo y este fue predicho como tal, si un cliente se observó como mal cliente siendo que se predijo que sería bueno, si un cliente resultó ser uno bueno siendo que se predijo que sería malo y, finalmente, si un cliente se predijo como mal cliente y efectivamente se observó esto.

Dados los valores obtenidos, se calculan 4 indicadores de la matriz de confusión para poder corroborar el poder de predicción: *precisión*, *predictivos positivos*, *predictivos negativos* y *sensibilidad*.

Se define *precisión* como $(VP + VN)/(P + N)$ y busca medir la capacidad de acercarse a los valores reales midiendo los verdaderos positivos (VP) y verdaderos negativos (VN) sobre el total de valores, es decir, mide el total de aciertos sobre el total. Este valor para la matriz de confusión corresponde a 77,26%, con lo que el modelo acierta en su predicción a más de 77% de los clientes.

Predictivo positivo mide la real capacidad de eficacia diagnóstica y se calcula con la fórmula $VP/(VP + FP)$, que busca medir la cantidad de buenos aciertos sobre el total de estimaciones positivas, es decir, cuántas veces el modelo predice y en la realidad se cumple dicha predicción sobre el total de casos positivos predichos, del mismo modo para el caso de predicciones negativas, donde la fórmula es análoga pero sobre los valores negativos ($VN/(VN + FN)$). Para el modelo testeado, se obtiene un 70,75% de predicciones positivas y un 82,96% de predicciones negativas, es decir, en más del 70% de los valores predichos positivos se cumple la predicción y esto sucede de forma análoga para los casos negativos.

Finalmente se mide la sensibilidad que busca explicar la cantidad de verdaderos positivos predichos por sobre los resultados reales, este indicador se define como $VP/(VN + FN)$ y para esta matriz de confusión se tiene que la sensibilidad corresponde al 78,45%.

Dados los indicadores de la matriz de confusión, se puede concluir que el modelo presenta una buena capacidad predictiva, ya que posee una buena capacidad de predicción positiva y negativa además de tener precisión y sensibilidad sobre el 75%.

Por otro lado, se puede evaluar gráficamente lo anterior mediante una curva *ROC* (*Receiver Operating Characteristic*) que representa la sensibilidad versus 1-especificidad, donde la especificidad se define como $VN/(VN + FP)$ y busca ver la tasa verdaderos negativos obtenidos.

Esta curva representa la clasificación binaria de *verdaderos positivos* frente a los *falsos positivos* que permite advertir la capacidad del modelo de predecir correctamente. Además, se contrasta con una línea diagonal que representa la probabilidad de acierto aleatorio del modelo, en otras palabras, la asignación con probabilidad de 50% para caracterizar a los clientes sin modelo. La curva desarrollada para el modelo se aprecia en Ilustración 10.

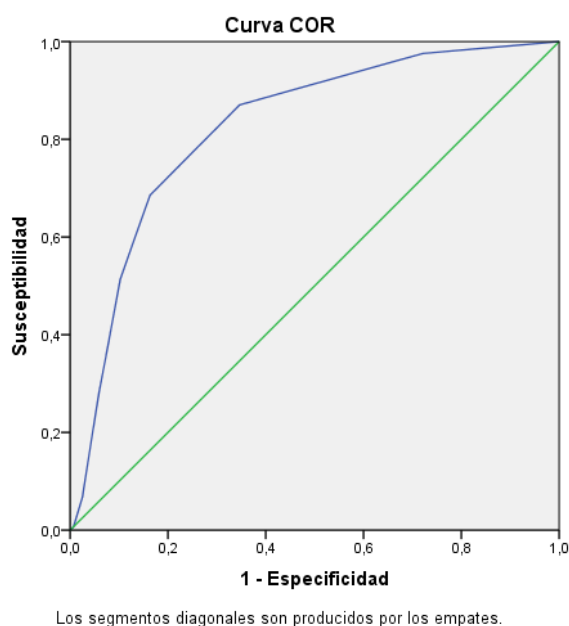


Ilustración 10: Curva ROC. Fuente: Elaboración propia en software

A partir de la curva *ROC*, es posible ver que el modelo presenta un beneficio de información por sobre una asignación aleatoria, debido a que la curva del modelo se presenta por sobre la línea diagonal y el área entre las curvas representa la ganancia de información debido al mismo, dicha área corresponde a un 82,4% (ver tabla 4), por lo que se concluye que este posee un buen poder predictivo.

Área bajo la curva

VARIABLES RESULTADO DE CONTRASTE: u

Área	Error típ. ^a	Sig. asintótica ^b	Intervalo de confianza asintótico al 95%	
			Límite inferior	Límite superior
,824	,002	,000	,820	,828

Tabla 4: Área bajo la curva ROC. Fuente: Elaboración propia

En conclusión, el modelo seleccionado posee la cualidad de ser el mejor modelo, estadísticamente hablando, y se corrobora con la prueba de predicción mediante su matriz de confusión y curva ROC, donde se aprecian buenos indicadores de rendimiento y la curva de información del modelo presenta claras ventajas por sobre la asignación aleatoria.

9.3 Segmentación

Con los resultados del modelo de propensión, se calcularon las probabilidades individuales para los clientes dadas sus características individuales. La distribución de las probabilidades individuales se presentan graficadas en el siguiente histograma (ver ilustración 11).

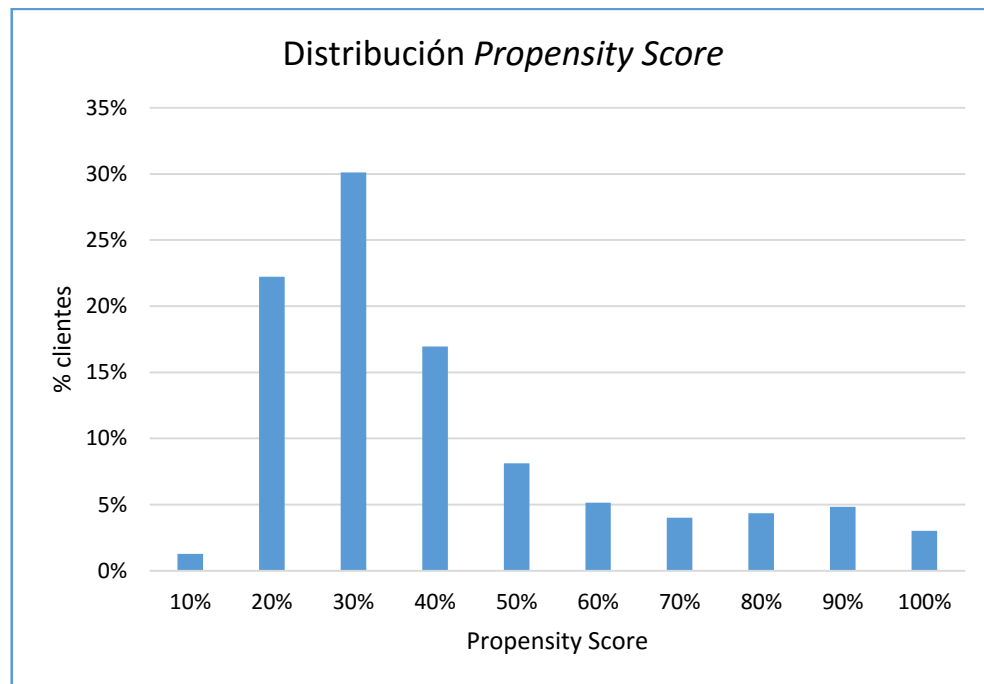


Ilustración 11: Histograma de distribución propensity score. Fuente: Elaboración propia.

De la distribución de *propensity score* es posible ver que existe una fuerte tendencia a bajas probabilidades de compra lo que es consecuente con la realidad de la empresa y sus clientes.

A partir de estos resultados, se generaron tres segmentos de clientes en función de estos valores: bajo *propensity score* (*BPS*), medio *propensity score* (*MPS*) y alto *propensity score* (*APS*), los cuales se caracterizan según se observa en la Tabla 5.

	Bajo <i>propensity score</i> (<i>BPS</i>)	Medio <i>propensity score</i> (<i>MPS</i>)	Alto <i>propensity score</i> (<i>APS</i>)
Intervalo de <i>propensity score</i>	[0%,25%]]25%,50%]]50%,100%]
Cantidad de clientes	38,9%	39,8%	21,3%

Tabla 5: Distribución de *propensity score* para segmentos de clientes. Fuente: Elaboración propia.

Las asignaciones se realizan para poder distribuir y formar grupos con tamaños robustos y no tener un desbalanceo del tamaño de los segmentos.

9.4 Diseño de experimentos

9.4.1 Grupos de experimentación

El diseño experimental que se aplicará consiste en un diseño con un factor y múltiples tratamientos, donde los niveles de tratamiento son tres, asociados a las dos distintas intervenciones que se realizarán y un nivel de intervención cero (grupo de control). De acuerdo con esto, el nivel 1 corresponde a tratamiento de personalización y recomendación simple, el nivel 2 se refiere a carrera por premios (club de fidelización) y el nivel 3, a grupo de control que no posee intervención.

Para cada nivel de tratamiento se selecciona una muestra aleatoria de cada segmento de *propensity score*, balanceando la muestra con un 40% de la muestra perteneciente a *BPS*, 40% de *MPS* y 20% de *APS*.

Las características de los grupos que serán objeto de experimentación, se presenta en la Tabla 7, en la que se destaca que existe un balance en pertenencia a segmento de clientes, los grupos de experimentación tienen tamaños similares y corresponde cada uno al 40% del grupo de experimentación y el control tiene un menor tamaño, correspondiente al 20% del grupo de experimentación.

Además, los tamaños muestrales se determinaron a partir de la potencia muestral deseada. Lo anterior se realizó considerando los dos primeros niveles de tratamiento juntos, con el fin de contrastar sobre el control deseando aumentar la tasa de respuesta de un 2% a un 3%, lo que nace a partir de la tasa de conversión y activación de clientes suscritos mensualmente. Esto quiere decir que solo un 2% de los clientes que se incorporan al sitio para recibir el *newsletter* generan una compra efectiva, por lo que es de especial interés para la empresa poder mejorar la activación de suscritos a través futuras intervenciones. Así, si se logra modificar en un 50%, esa tasa ya se considera interesante para la empresa. Dado lo anterior, si la potencia muestral fuese un 85%, se necesitarían 10.939 clientes intervenidos y 2.735 clientes de control, asumiendo una proporción de 80% intervenidos sobre un 20% de control para poder intervenir a la mayor cantidad de usuarios posible.

	# Clientes	# Clientes BPS	# Clientes MPS	# Clientes APS
Nivel 1	5.361	2.196	2.144	1.021
Nivel 2	5.352	2.107	2.137	1.108
Nivel 3	2.695	1.053	1.117	525
Total	13.408	5.356	5.398	2.654

Tabla 6: Distribución de número de clientes para los grupos de experimentación según nivel de tratamiento y segmento de propensity score. Fuente: Elaboración propia.

	% Clientes	% Clientes BPS	% Clientes MPS	% Clientes APS
Nivel 1	40,0%	41,0%	40,0%	19,0%
Nivel 2	39,9%	39,4%	39,9%	20,7%
Nivel 3	20,1%	39,1%	41,4%	19,5%
Total	100,0%	39,9%	40,3%	19,8%

Tabla 7: Distribución de proporción de clientes para los grupos de experimentación según nivel de tratamiento y segmento de propensity score. Fuente: Elaboración propia.

9.4.2 Experimentos que se realizarán

Como se ha mencionado en secciones anteriores, el proceso experimental consta de tres niveles de tratamientos, donde estos apuntan a testear si los distintos niveles experimentales aumentan el número de compras realizadas y cantidad de clientes activados.

Los niveles experimentales mencionados anteriormente se detallan a continuación.

9.4.2.1 Nivel 1: Personalización y recomendación simple

El nivel 1 de experimentación consiste en el envío de un correo electrónico a los clientes del grupo de experimentación, con un asunto de correo (*subject*) personalizado, que demuestre mayor involucramiento y cercanía con el cliente. Dentro del correo, se ofrecerá la oferta del sitio que presente mayor conversión (número de compras/visitas), relacionada con la historia de compras del cliente. Particularmente, se enviará la mejor oferta que pertenezca a la subcategoría de la primera compra realizada por el cliente, es decir, se considera la subcategoría de la primera compra realizada por el cliente y se ofrece lo mejor del sitio para cada cliente que pertenezca a la subcategoría en cuestión.

9.4.2.1.1 Hipótesis a testear – Nivel 1: Personalización y recomendación simple

La hipótesis de comportamiento que se pretende probar mediante la experimentación, consiste en que muchos clientes no ven las ofertas adecuadas para poder comprar en el sitio y les parece poco atractiva la forma en que estas se muestran, por lo que basta con hacer más llamativo el anuncio y ofrecer lo que le gusta para que compren más.

9.4.2.2 Nivel 2: Carrera por premio (Club de fidelización)

El nivel de intervención número 2 corresponde a un club de fidelización del sitio, que premiará a aquellos clientes que concreten 10 compras durante el año 2015. El premio consiste en una *gift card* libre de uso por el monto promedio de las compras realizadas durante el año. Se destaca que existen clientes que pueden haber comenzado la carrera con algunas compras ya hechas, pero no pueden ser más de 5 compras antes de la experimentación, en otras palabras, un cliente tiene que comprar al menos 5 veces desde el inicio del proceso experimental.

9.4.2.2.1 Hipótesis a testear – Nivel 2: Carrera por premio (Club de fidelización)

La hipótesis de comportamiento que se testeará mediante este nivel de experimentación consiste en que los clientes aceleran su comportamiento de compra a medida que se acercan a lograr un premio.

9.4.2.3 Nivel 3: Control

El tercer nivel experimental consiste en el grupo de control que no posee ningún tipo de intervención adicional a los clientes. Se utiliza este grupo para contrastar posteriormente el efecto de las intervenciones en el comportamiento de compras de los consumidores.

Se destaca que los niveles experimentales fueron escogidos en conjunto con la empresa buscando incentivar a los clientes mediante dos grados distintos de involucramiento con el cliente donde el primer nivel genera un bajo nivel de involucramiento y el segundo nivel se espera que genera un involucramiento mayor entre el cliente y la empresa.

9.5 Desarrollo de experimentos

9.5.1 Nivel 1: Personalización y recomendación simple.

El proceso experimental tendrá como duración dos meses, durante los cuales se enviarán recordatorios semanales a los clientes para cada nivel de tratamiento. El proceso de experimentación comienza el día 26 de julio y se extiende por ocho semanas. Para el nivel 1, se enviará cada semana el mismo *mail*, con el mismo asunto personalizado pero cambiará en contenido, pues la oferta cambia dinámicamente en el sitio, por lo que semana a semana se cambiará el contenido de la oferta al cliente.

El correo enviado a cada cliente se presenta a continuación (ver ilustración 12). El asunto del correo es “[Nombre], esta oferta te podría interesar!” donde la variable [Nombre] corresponde al nombre de pila del cliente.



Ilustración 12: Correo enviado a clientes nivel 1 de tratamiento: Personalización y recomendación simple.
Fuente: Elaboración propia.

9.5.2 Nivel 2: Carrera por premio (Club de fidelización)

Para el nivel 2 se enviará una actualización en su estado de avance de la carrera y a los ganadores se enviará un correo especial, entregando el código de la *gift card* de premio y el monto asociado a esta. Además, en el *footer* del correo, se enviarán tres ofertas a los clientes, las cuales cambiarán semana a semana por el mismo motivo del cambio del nivel 1. Las tres ofertas corresponden a las destacadas para hombres y mujeres, y a la que posea mayor número de visitas en los últimos tres días (sobre la base de *google analytics*).

En el primer envío de experimentos, se remite a los clientes un *mail* informando que son parte del club de fidelización llamado “Club cuponatic”, destacando las reglas de la intervención y su estado actual en el proceso (número de compras y premio al cual están accediendo hasta el momento). Un ejemplo del mail de invitación se aprecia a continuación (ver ilustración 13).



Ilustración 13: Mail invitación experimento nivel 2: Carrera por compras (club de fidelización). Fuente: Elaboración propia.

En los siguientes envíos, se emite un correo similar al anterior, pero cambiando el mensaje, demostrando así que se está enviando un resumen semanal de estados de cuenta, como se puede apreciar en ilustración 14.

Finalmente, si algún cliente completa la meta de las 10 compras, será premiado con la *gift card* y se le anunciará el hecho mediante un correo electrónico que contenga la información necesaria para tal propósito. El correo en cuestión se puede apreciar en la ilustración 14.

A continuación, se presenta una tabla comparativa con las características principales de los distintos niveles de experimentación:

	Nivel 1	Nivel 2	Nivel 3
Envío de mail	Sí	Sí	No
Personalización asunto	Sí	Sí	No
Personalización de ofertas	Sí	No	No
Estado de avance	No	Sí	No

Tabla 8: Caracterización de niveles experimentales. Fuente: Elaboración propia

CUPONATIC
¡Vive y Sorpréndete!

CLUB CUPONATIC Resumen semanal

¡Sigue Así! Falta menos...

RESUMEN AL 16 DE SEPTIEMBRE 2015:
Junta 9 cupones más antes del 31 de diciembre del 2015 y te daremos una compra GRATIS por el valor promedio de tus cupones acumulados.

TIENES 1/10 CUPONES. VALOR PROMEDIO \$2.990

¡CONSIGUE MÁS!

OFERTAS DE HOY QUE TE PODRÍAN INTERESAR

<p>Sesión Ácido Hialurónico con Microaguja DERMAPEN \$39.990 (Mejor precio)</p>	<p>40% Set de 4 Mancuernas \$2.990 \$5.000</p>	<p>Rack Tv Home Theater \$45.990 \$139.990</p>
--	---	---

¿Tienes alguna sugerencia, duda o consulta sobre nuestro servicio? Estaremos felices de ayudarte. Por favor escribemos a contacto@cuponatic.com o llámanos al [+562 23071900](tel:+56223071900).

¡Que tengas un día Cuponatic!

CUPONATIC
¡Vive y Sorpréndete!

CLUB CUPONATIC ¡Ganaste!

¡Felicitaciones Marisol! Has ganado.

Has llegado a la meta completando 10 cupones antes del 31 de diciembre del 2015, como premio a tu fidelidad te daremos una compra GRATIS por el valor promedio de tus cupones acumulados.

Has ganado una Gift Card por el monto de: \$6.290, el código es: MarisolWinner

¡VEN A CELEBRAR!

OFERTAS DE HOY QUE TE PODRÍAN INTERESAR

<p>88% 20 Ses. Trat. Tonicificante Abdomen y Glúteos / Stgo. Centro \$7.490 \$54.990</p>	<p>\$2.990 por \$21.000 en créditos para Uber nuevos usuarios \$2.990 \$21.000</p>	<p>89% Spa completo para 1 o 2 personas en Stgo Centro \$9.990 \$75.000</p>
---	---	--

¿Tienes alguna sugerencia, duda o consulta sobre nuestro servicio? Estaremos felices de ayudarte. Por favor escribemos a contacto@cuponatic.com o llámanos al [+562 23071900](tel:+56223071900).

¡Que tengas un día Cuponatic!

Ilustración 14: Correos nivel 2: Carrera por compras (club de fidelización). Izquierda: Correo semanal de estado de avance. Derecha: Correo de ganadores que completaron la meta estipulada. Fuente: Elaboración propia.

9.6 Evaluación y análisis de experimentos

Una vez transcurrido el periodo de experimentación, se procede a la obtención de los resultados del proceso de experimentación. En particular, se obtienen datos de: cantidad de clientes que compraron (activados) dentro de los segmentos y número de compras realizadas por los clientes. La evaluación estadística de comparación de medias se realiza mediante el *software SPSS*.

9.7 Indicadores de correo (Open rate y Click through rate)

Dado que el proceso de experimentación se basa en el envío de correos electrónicos es pertinente evidenciar los principales indicadores de envío de correos electrónico como lo son Open rate y click through rate (CTR) donde el primero muestra la cantidad de aperturas realizadas por clientes distintos al correo electrónico y el segundo evidencia la proporción de clicks realizados dentro del correo. Los indicadores en cuestión se pueden apreciar en la tabla 9.

Se destaca que se aprecian 8 indicadores pues se realizaron 8 envíos (dos meses de experimentación con recordatorios semanales).

Envío	Nivel 1 – Recomendación simple		Nivel 2 – Carrera por premio	
	Open rate	CTR	Open rate	CTR
Recordatorio 1	32,75 %	1,9 %	31,11 %	8,33 %
Recordatorio 2	29,83 %	1,24 %	29,85 %	9,92v %
Recordatorio 3	30,57 %	1,64 %	26,23 %	7,9 %
Recordatorio 4	28,43 %	1,84 %	23,95 %	6,53 %
Recordatorio 5	27,57 %	1,12 %	22,92 %	7,27 %
Recordatorio 6	27,56 %	1,12 %	23,66 %	5,19 %
Recordatorio 7	26,85 %	1,07 %	22,57 %	4,86 %
Recordatorio 8	26,02 %	1,05 %	21,89 %	5,2 %

Tabla 9: Indicadores de rendimiento para correo electrónico, Open rate y Click through rate (CTR). Fuente: Elaboración propia en base a información de Sendgrid

9.7.1 Activación de clientes

Es necesario entender que se define como activado a un cliente, toda vez que este genera al menos una compra durante el periodo de experimentación.

A continuación, se presentan los datos obtenidos para la variable dependiente “activación”.

Sobre la base de esta definición, es posible ver en la tabla 10 que el máximo nivel de activación lo obtiene el nivel 2 de tratamiento, con un 10,1%, que corresponde al 0,5% más que los otros dos niveles de intervención. A pesar de esta diferencia de medias inicial, es pertinente comparar estadísticamente esta diferencia mediante un test de comparación de medias.

Estimaciones

Variable dependiente: activacion

Tratamiento	Media	Error tip.	Intervalo de confianza 95%	
			Límite inferior	Límite superior
Nivel 1	,096	,004	,088	,104
Nivel 2	,101	,004	,093	,109
Nivel 3	,096	,006	,085	,107

Tabla 10: Estimaciones de medias para niveles de tratamiento. Variable dependiente: Activación. Efectos fijos: Nivel de tratamiento. Fuente: Elaboración propia.

A partir de la tabla 11, es posible advertir que el valor del test F de la prueba de comparación de medias es 0,413, lo que implica que este valor no es significativamente distinto a 1, por lo que se acepta la hipótesis nula de igualdad de medias para el test, lo que, en términos simples, significa que las medias para los distintos niveles de tratamiento son estadísticamente iguales y no existe un cambio de comportamiento significativo en la variable activación provocado por el efecto del tratamiento.

Pruebas de los efectos inter-sujetos

Variable dependiente: activacion

Origen	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
Modelo corregido	,073 ^a	2	,036	,413	,662
Intersección	114,933	1	114,933	1301,749	,000
Tratamiento	,073	2	,036	,413	,662
Error	1183,545	13405	,088		
Total	1312,000	13408			
Total corregida	1183,618	13407			

a. R cuadrado = ,000 (R cuadrado corregida = ,000)

Tabla 11: Prueba de efectos inter sujetos. Variable dependiente: Activación. Efectos fijos: Nivel de tratamiento. Fuente: Elaboración propia.

Sumado a lo anterior, se evalúa el efecto que generan los distintos tratamientos comparándolos entre sí. Esto se aprecia en la tabla 12, en la que se comparan las medias de las distintas interacciones de tratamientos posibles. En dicha tabla, se aprecia que no existe ninguna interacción de comparación de medias que represente un cambio significativo de comportamiento, lo que se traduce –en términos simples– en que el efecto tratamiento no genera un cambio en el comportamiento de los clientes independiente del nivel de tratamiento y su comparación.

Comparaciones por pares

Variable dependiente: activacion

(I) Tratamiento	(J) Tratamiento	Diferencia de medias (I-J)	Error típ.	Sig. ^a	Intervalo de confianza al 95 % para la diferencia ^a	
					Límite inferior	Límite superior
Nivel 1	Nivel 2	-,005	,006	,418	-,016	,007
	Nivel 3	,000	,007	,962	-,013	,014
Nivel 2	Nivel 1	,005	,006	,418	-,007	,016
	Nivel 3	,005	,007	,478	-,009	,019
Nivel 3	Nivel 1	,000	,007	,962	-,014	,013
	Nivel 2	-,005	,007	,478	-,019	,009

Basadas en las medias marginales estimadas.

a. Ajuste para comparaciones múltiples: Diferencia menos significativa (equivalente a la ausencia de ajuste).

Tabla 12: Comparación de pares entre niveles de tratamiento. Variable dependiente: Activación. Efectos fijos: Nivel de tratamiento. Fuente: Elaboración propia.

Lo anterior es corroborado por un test de proporciones de Chi-cuadrado, donde en la tabla 13 se aprecia el resumen de casos validos donde se destaca que el valor 0 para la variable activación corresponde a un cliente que no fue activado y si el valor es 1 corresponde a un cliente activado (i.e. generó una compra en el periodo).

Tabla de contingencia activacion ^ Tratamiento

Recuento		Tratamiento			Total
		Nivel 1	Nivel 2	Nivel 3	
activacion	0	4846	4813	2437	12096
	1	515	539	258	1312
Total		5361	5352	2695	13408

Tabla 13: Tabla de contingencia para variable activación en los niveles de tratamiento. Fuente: Elaboración propia

Además, la prueba de Chi-cuadrado entrega una significancia asintótica de igual valor al test de medias calculado con anterioridad lo que corrobora el resultado planteado con anterioridad. Se destaca que la prueba de Chi-cuadrado se utiliza para contrastar proporciones lo que es más adecuado para medir estas diferencias en variables binarias como la activación.

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	,826 ^a	2	,662
Razón de verosimilitudes	,824	2	,662
N de casos válidos	13408		

a. 0 casillas (0,0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 263,71.

Tabla 14: Prueba Chi-cuadrado para variable activación en niveles de tratamiento. Fuente: Elaboración propia

En segundo lugar, se medirá estadísticamente las diferencias de activación dependiendo del segmento de cliente al que pertenece.

Se aprecia en la tabla 15 las medias de activación para cada segmento de clientes, de la cual se desprende que el segmento de alto *propensity score* tiene una activación media de 12,5% clientes, equivalente a un 6,3% superior al segmento con menor proporción de activación (segmento de bajo *propensity score*).

Estimaciones

Variable dependiente: activacion

segmento	Media	Error típ.	Intervalo de confianza 95%	
			Límite inferior	Límite superior
APS	,125	,006	,114	,137
BPS	,062	,004	,054	,070
MPS	,120	,004	,112	,128

Tabla 15: Estimaciones de medias para segmento de clientes. Variable dependiente: Activación. Efectos fijos: Segmento de clientes Fuente: Elaboración propia.

La diferencia antes mencionada se compara estadísticamente, y el resultado de esta comparación se aprecia en la tabla 16, donde el valor del estadístico F es 67,27, el cual es significativamente distinto de 1 a un 95% de confianza. Esto genera la conclusión de que el nivel de tratamiento efectivamente genera un cambio de comportamiento en los clientes –estadísticamente hablado–, pues se rechaza la hipótesis nula de igualdad de medias.

Pruebas de los efectos inter-sujetos

Variable dependiente: activacion

Origen	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
Modelo corregido	11,761 ^a	2	5,881	67,270	,000
Intersección segmento	126,133	1	126,133	1442,845	,000
Error	1171,857	13405	,087	67,270	,000
Total	1312,000	13408			
Total corregida	1183,618	13407			

a. R cuadrado = ,010 (R cuadrado corregida = ,010)

Tabla 16: Prueba de efectos inter sujetos. Variable dependiente: Activación. Efectos fijos: Segmento de clientes. Fuente: Elaboración propia.

A pesar de lo anterior, se vuelve importante evidenciar en qué segmentos de clientes existe una mayor diferencia (si la hay). Lo anterior es posible apreciarlo a partir de la tabla 17, donde se comparan las medias evidenciadas en la tabla 16. A partir de los valores de la tabla 15, es posible ver que los segmentos de alto y medio *propensity score* presentan un comportamiento significativamente distinto al 95% de confianza, es decir, estos segmentos de clientes efectivamente presentan un mejor comportamiento en términos de activación de cliente respecto del segmento de bajo *propensity score*, por lo que, independiente del nivel de tratamiento, los segmentos generados presentan un comportamiento diferente y se reconoce lo anterior en términos estadísticos.

Comparaciones por pares

Variable dependiente: activacion

(I)segmento	(J)segmento	Diferencia de medias (I-J)	Error típ.	Sig. ^b	Intervalo de confianza al 95 % para la diferencia ^b	
					Límite inferior	Límite superior
APS	BPS	,064*	,007	,000	,050	,078
	MPS	,005	,007	,455	-,008	,019
BPS	APS	-,064*	,007	,000	-,078	-,050
	MPS	-,059*	,006	,000	-,070	-,047
MPS	APS	-,005	,007	,455	-,019	,008
	BPS	,059*	,006	,000	,047	,070

Basadas en las medias marginales estimadas.

*. La diferencia de medias es significativa al nivel ,05.

b. Ajuste para comparaciones múltiples: Diferencia menos significativa (equivalente a la ausencia de ajuste).

Tabla 17: Comparación de pares entre segmento de clientes. Variable dependiente: Activación. Efectos fijos: Segmento de clientes. Fuente: Elaboración propia.

Finalmente, se generará la comparación de medias para la interacción entre nivel de tratamiento y segmento de clientes, dicho de otro modo, el efecto del nivel de tratamiento para los diversos segmentos de clientes y viceversa, o bien, el efecto del segmento de clientes para los diversos niveles de tratamiento.

En primer lugar, se medirá el efecto de la segmentación dentro de cada nivel de tratamiento. Las estimaciones de medias para cada dupla de interés se presentan en la tabla 18. A partir de esta tabla es posible distinguir que, para el nivel 1 de

tratamiento, la diferencia entre el segmento con mayor proporción de activos y el menor (MPS versus BPS) posee una diferencia de 6,2%, mientras que en el nivel 2 dicha diferencia, de 6,7%, se genera entre el segmento de alto *propensity score* y el segmento de bajo *propensity score* (APS vs BPS) y para el nivel 3 la diferencia se genera nuevamente entre el segmento de alto y bajo *propensity score*, con un 6,7% de diferencia.

Estimaciones

Variable dependiente: activacion

Tratamiento	segmento	Media	Error tip.	Intervalo de confianza 95%	
				Límite inferior	Límite superior
Nivel 1	APS	,119	,009	,100	,137
	BPS	,060	,006	,048	,072
	MPS	,122	,006	,110	,135
Nivel 2	APS	,128	,009	,111	,146
	BPS	,061	,006	,049	,074
	MPS	,125	,006	,113	,138
Nivel 3	APS	,133	,013	,108	,159
	BPS	,066	,009	,048	,083
	MPS	,107	,009	,089	,124

Tabla 18: Estimaciones de medias para segmento de clientes dentro de cada nivel de tratamiento. Variable dependiente: Activación. Efectos fijos: Nivel de tratamiento y Segmento de clientes Fuente: Elaboración propia.

Las diferencias antes mencionadas se evaluaron estadísticamente mediante un test de comparación de medias. Esta interacción de factores fijos generó un valor del estadístico F de 0,938, que no es significativamente distinto de 1 al 95% de confianza, es decir, dentro de los niveles de tratamiento, no existe un cambio significativo de comportamiento para los distintos segmentos de cliente. Lo anterior es posible verlo en la tabla 19.

Pruebas de los efectos inter-sujetos

Variable dependiente: activacion

Origen	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
Modelo corregido	12,148 ^a	8	1,518	17,368	,000
Intersección	113,162	1	113,162	1294,319	,000
Tratamiento	,053	2	,027	,306	,737
segmento	10,036	2	5,018	57,394	,000
Tratamiento * segmento	,328	4	,082	,938	,440
Error	1171,470	13399	,087		
Total	1312,000	13408			
Total corregida	1183,618	13407			

a. R cuadrado = ,010 (R cuadrado corregida = ,010)

Tabla 19: Prueba de efectos inter sujetos. Variable dependiente: Activación. Efectos fijos: Nivel de tratamiento y Segmento de clientes. Fuente: Elaboración propia.

Adicional a lo anterior, se estudia la diferencia entre los distintos segmentos para los múltiples niveles de tratamiento que se evidencia en la tabla 20. En esta tabla se aprecia que nuevamente los segmentos de medio y alto *propensity score* poseen un comportamiento significativamente distinto al segmento de bajo *propensity score* para todos los niveles de tratamiento, corroborando así que el nivel de tratamiento

aparentemente no genera un cambio concreto de comportamiento para la variable “activación”; sino que el cambio se evidencia para la segmentación, ya que en el nivel agregado no se corrobora un cambio de comportamiento significativo.

Comparaciones por pares

Variable dependiente: activacion

Tratamiento	(I)segmento	(J)segmento	Diferencia de medias (I-J)	Error típ.	Sig. ^b	Intervalo de confianza al 95 % para la diferencia ^b	
						Límite inferior	Límite superior
Nivel 1	APS	BPS	,058*	,011	,000	,036	,080
		MPS	-,004	,011	,743	-,026	,018
	BPS	APS	-,058*	,011	,000	-,080	-,036
		MPS	-,062*	,009	,000	-,080	-,044
MPS	APS	,004	,011	,743	-,018	,026	
	BPS	,062*	,009	,000	,044	,080	
Nivel 2	APS	BPS	,067*	,011	,000	,045	,088
		MPS	,003	,011	,802	-,019	,024
	BPS	APS	-,067*	,011	,000	-,088	-,045
		MPS	-,064*	,009	,000	-,082	-,046
MPS	APS	-,003	,011	,802	-,024	,019	
	BPS	,064*	,009	,000	,046	,082	
Nivel 3	APS	BPS	,068*	,016	,000	,037	,099
		MPS	,027	,016	,087	-,004	,057
	BPS	APS	-,068*	,016	,000	-,099	-,037
		MPS	-,041*	,013	,001	-,066	-,016
MPS	APS	-,027	,016	,087	-,057	,004	
	BPS	,041*	,013	,001	,016	,066	

Basadas en las medias marginales estimadas.

*. La diferencia de medias es significativa al nivel ,05.

b. Ajuste para comparaciones múltiples: Diferencia menos significativa (equivalente a la ausencia de ajuste).

Tabla 20: Comparación de pares entre segmento de clientes para los distintos segmentos de clientes. Variable dependiente: Activación. Efectos fijos: Nivel de tratamiento y Segmento de clientes. Fuente: Elaboración propia.

De manera análoga, se genera el análisis estadístico para el efecto del nivel de tratamiento dentro de los distintos segmentos de clientes; pero este análisis se realiza solo para el nivel desagregado, puesto que en el nivel agregado el resultado es igual al presentado en la tabla 19.

En primer lugar, se aprecia que para el segmento de alto *propensity score* el nivel 3 de tratamiento presenta un mejor promedio de activación con un 13,3%, un 1,4% más alto que el nivel con peor impacto que es el nivel 1. Del mismo modo, se aprecia que el nivel 3 es el tratamiento con mejor resultado en términos de activación para el segmento de bajo *propensity score*, con una diferencia de 0,6% respecto del nivel 1, que presentó el peor nivel de aprobación. Para el segmento de medio *propensity score*, se ve que el nivel 2 de tratamiento presenta el mejor incremento de activación, con un 12,5%, 1,8% superior que el nivel 3 de tratamiento, que presenta el peor nivel de activación según la tabla 21.

Estimaciones

Variable dependiente: activacion

segmento	Tratamiento	Media	Error típ.	Intervalo de confianza 95%	
				Límite inferior	Límite superior
APS	Nivel 1	,119	,009	,100	,137
	Nivel 2	,128	,009	,111	,146
	Nivel 3	,133	,013	,108	,159
BPS	Nivel 1	,060	,006	,048	,072
	Nivel 2	,061	,006	,049	,074
	Nivel 3	,066	,009	,048	,083
MPS	Nivel 1	,122	,006	,110	,135
	Nivel 2	,125	,006	,113	,138
	Nivel 3	,107	,009	,089	,124

Tabla 21: Estimaciones de medias para nivel de tratamiento dentro de segmento de clientes. Variable dependiente: Activación. Efectos fijos: Segmento de clientes y Nivel de tratamiento. Fuente: Elaboración propia.

Finalmente, se aprecia en la tabla 22 que dentro de cada segmento de cliente no existe un tratamiento que sea estadísticamente superior, corroborando que no hay un cambio de comportamiento significativamente distinto dentro de los segmentos de cliente.

Comparaciones por pares

Variable dependiente: activacion

segmento	(I)Tratamiento	(J)Tratamiento	Diferencia de medias (I-J)	Error típ.	Sig. ^a	Intervalo de confianza al 95 % para la diferencia ^a	
						Límite inferior	Límite superior
APS	Nivel 1	Nivel 2	-,010	,013	,452	-,035	,015
		Nivel 3	-,015	,016	,351	-,046	,016
	Nivel 2	Nivel 1	,010	,013	,452	-,015	,035
		Nivel 3	-,005	,016	,741	-,036	,026
	Nivel 3	Nivel 1	,015	,016	,351	-,016	,046
		Nivel 2	,005	,016	,741	-,026	,036
BPS	Nivel 1	Nivel 2	-,001	,009	,902	-,019	,017
		Nivel 3	-,005	,011	,625	-,027	,016
	Nivel 2	Nivel 1	,001	,009	,902	-,017	,019
		Nivel 3	-,004	,011	,700	-,026	,018
	Nivel 3	Nivel 1	,005	,011	,625	-,016	,027
		Nivel 2	,004	,011	,700	-,018	,026
MPS	Nivel 1	Nivel 2	-,003	,009	,723	-,021	,015
		Nivel 3	,016	,011	,151	-,006	,037
	Nivel 2	Nivel 1	,003	,009	,723	-,015	,021
		Nivel 3	,019	,011	,084	-,003	,040
	Nivel 3	Nivel 1	-,016	,011	,151	-,037	,006
		Nivel 2	-,019	,011	,084	-,040	,003

Basadas en las medias marginales estimadas.

a. Ajuste para comparaciones múltiples: Diferencia menos significativa (equivalente a la ausencia de ajuste).

Tabla 22: Comparación de pares entre nivel de tratamiento dentro de cada segmento de clientes. Variable dependiente: Activación. Efectos fijos: Segmento de clientes y Nivel de tratamiento. Fuente: Elaboración propia.

9.7.2 Impacto en número de compras

Es de gran interés saber el promedio de compras de los clientes, y no sólo su activación, pues lo relevante es poder generar una relación de fidelidad con el cliente y que este genere más órdenes de compra.

Para la presente sección, se generará el análisis para la variable dependiente: número de compras, consistente en la cantidad total de compras generadas por los clientes durante el proceso de intervención.

Al igual que en la sección anterior, se generará el análisis de medias para el número de órdenes de compras para el nivel de tratamiento, segmento de clientes y la interacción de estos efectos fijos.

En la tabla 23, se aprecian las medias de órdenes de compra para los distintos niveles de tratamiento. Se aprecia, en primer lugar, que el nivel 2 de tratamiento presenta el mejor promedio de órdenes de compras generadas con 0,15 órdenes de compra, que resulta ser superior en 0,01 órdenes de compra al nivel con peor promedio (nivel 1).

Estimaciones

Variable dependiente: numero_compras

Tratamiento	Media	Error típ.	Intervalo de confianza 95%	
			Límite inferior	Límite superior
Nivel 1	,140	,007	,126	,155
Nivel 2	,159	,007	,144	,173
Nivel 3	,141	,011	,121	,162

Tabla 23: Estimaciones de medias para niveles de tratamiento. Variable dependiente: Número de compras. Efectos fijos: Nivel de tratamiento. Fuente: Elaboración propia.

A pesar de esta superioridad, considerando la media de número de compras, se comparan estas diferencias en términos estadísticos. Los resultados de este contraste se presentan en la tabla 24. Los resultados del contraste indican que el valor del estadístico F es 1,743, que no es significativamente distinto de cero, al 95% de confianza, es decir, no existe una diferencia significativa para el número de compras de los clientes dependiendo del nivel de tratamiento al que han sido expuestos.

Pruebas de los efectos inter-sujetos

Variable dependiente: numero_compras

Origen	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
Modelo corregido	1,043 ^a	2	,522	1,743	,175
Intersección	260,391	1	260,391	870,031	,000
Tratamiento	1,043	2	,522	1,743	,175
Error	4011,973	13405	,299		
Total	4306,000	13408			
Total corregida	4013,016	13407			

a. R cuadrado = ,000 (R cuadrado corregida = ,000)

Tabla 24: Prueba de efectos inter sujetos. Variable dependiente: Número de compras. Efectos fijos: Nivel de tratamiento. Fuente: Elaboración propia.

A pesar de la falta de significancia global, se estudia la comparación entre los distintos niveles de tratamiento para el número de compras. Dicha comparación se presenta en la tabla 25, donde se aprecia que no existen diferencias significativas en la comparación entre los distintos niveles de tratamiento, en consecuencia, no existe un efecto significativo de cambio de comportamiento para el número de compras generadas por los clientes.

Comparaciones por pares

Variable dependiente: numero_compras

(I) Tratamiento	(J) Tratamiento	Diferencia de medias (I-J)	Error típ.	Sig. ^a	Intervalo de confianza al 95 % para la diferencia ^a	
					Límite inferior	Límite superior
Nivel 1	Nivel 2	-,018	,011	,082	-,039	,002
	Nivel 3	-,001	,013	,932	-,026	,024
Nivel 2	Nivel 1	,018	,011	,082	-,002	,039
	Nivel 3	,017	,013	,182	-,008	,043
Nivel 3	Nivel 1	,001	,013	,932	-,024	,026
	Nivel 2	-,017	,013	,182	-,043	,008

Basadas en las medias marginales estimadas.

a. Ajuste para comparaciones múltiples: Diferencia menos significativa (equivalente a la ausencia de ajuste).

Tabla 25: Comparación de pares para nivel de tratamiento. Variable dependiente: Número de compras. Efectos fijos: Nivel de tratamiento. Fuente: Elaboración propia.

En la tabla 26, se aprecian las medias de número de compras para los distintos segmentos de clientes establecidos, en esta tabla se aprecia que el segmento de alto *propensity score* posee un promedio de 0,18 compras realizadas y el segmento con peor rendimiento es el de bajo *propensity score*, con 0,09 compras generadas.

Estimaciones

Variable dependiente: numero_compras

segmento	Media	Error típ.	Intervalo de confianza 95%	
			Límite inferior	Límite superior
APS	,187	,011	,166	,207
BPS	,090	,007	,075	,104
MPS	,186	,007	,172	,201

Tabla 26: Estimaciones de medias para Segmento de clientes. Variable dependiente: Número de compras. Efectos fijos: Segmento de clientes Fuente: Elaboración propia.

Las diferencias enunciadas presentan un comportamiento estadísticamente diferente, debido a que el valor del estadístico F para este contraste es de 50,51 el cual es significativamente distinto al 95% de confianza, de acuerdo con la tabla 27. Lo anterior implica que, efectivamente, la segmentación evidencia un cambio de comportamiento de los clientes de estos segmentos para el número de compras.

Pruebas de los efectos inter-sujetos

Variable dependiente: numero_compras

Origen	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
Modelo corregido	30,019 ^a	2	15,010	50,516	,000
Intersección segmento	285,909	1	285,909	962,244	,000
Error	3982,997	13405	,297		
Total	4306,000	13408			
Total corregida	4013,016	13407			

a. R cuadrado = ,007 (R cuadrado corregida = ,007)

Tabla 27: Prueba de efectos inter sujetos. Variable dependiente: Número de compras. Efectos fijos: Segmento de clientes. Fuente: Elaboración propia.

A partir del resultado anterior, se analizan las comparaciones entre los distintos segmentos de clientes para ver cuál es el mejor. Lo anterior se evidencia en la tabla 28, en la que se aprecia que la diferencia es significativa para los segmentos de alto y medio *propensity score* cuando se comparan con el segmento de bajo *propensity score*, es decir, el segmento de bajo *propensity score* es estadísticamente inferior a los demás para el número de compras, pero no existen diferencias significativas entre los de medio y alto *propensity score*.

Comparaciones por pares

Variable dependiente: numero_compras

(I)segmento	(J)segmento	Diferencia de medias (I-J)	Error típ.	Sig. ^b	Intervalo de confianza al 95 % para la diferencia ^b	
					Límite inferior	Límite superior
APS	BPS	,097 [*]	,013	,000	,071	,122
	MPS	,000	,013	,991	-,025	,025
BPS	APS	-,097 [*]	,013	,000	-,122	-,071
	MPS	-,097 [*]	,011	,000	-,117	-,076
MPS	APS	,000	,013	,991	-,025	,025
	BPS	,097 [*]	,011	,000	,076	,117

Basadas en las medias marginales estimadas.

*. La diferencia de medias es significativa al nivel ,05.

b. Ajuste para comparaciones múltiples: Diferencia menos significativa (equivalente a la ausencia de ajuste).

Tabla 28: Comparación de pares para segmentos de cliente. Variable dependiente: Número de compras. Efectos fijos: Segmento de clientes. Fuente: Elaboración propia.

Para finalizar, se comparará la interacción entre los factores fijos: Segmento de clientes y nivel de tratamiento. Al igual que en la sección anterior, en primer lugar, se analizará el efecto de la segmentación para cada nivel de tratamiento. Esta comparación se presenta en la tabla 29, donde se advierte que para el nivel 1 de tratamiento, el segmento de medio *propensity score* presenta el mejor promedio de compras (0,18), con un incremento de 0,1 órdenes en promedio en comparación con el segmento con peor promedio de órdenes de compra, que pertenece al segmento con bajo *propensity score* (0,08 órdenes de compra).

Estimaciones

Variable dependiente: numero_compras

Tratamiento	segmento	Media	Error típ.	Intervalo de confianza 95%	
				Límite inferior	Límite superior
Nivel 1	APS	,177	,017	,144	,211
	BPS	,084	,012	,061	,107
	MPS	,180	,012	,157	,203
Nivel 2	APS	,198	,016	,166	,230
	BPS	,091	,012	,068	,114
	MPS	,205	,012	,182	,228
Nivel 3	APS	,181	,024	,134	,228
	BPS	,099	,017	,066	,132
	MPS	,163	,016	,131	,195

Tabla 29: Estimaciones de medias para segmento de clientes dentro de cada nivel de tratamiento. Variable dependiente: Número de compras. Efectos fijos: Nivel de tratamiento y Segmento de clientes. Fuente: Elaboración propia.

Las diferencias de medias presentadas anteriormente, a pesar de ser aparentemente altas, no presentan una diferencia significativamente distinta a un 95% de confianza, pues el valor del estadístico F es de 0,75 y resulta no ser significativamente distinto de 1, por lo que se acepta la hipótesis nula de igualdad de medias para el conjunto testeado, según se aprecia en la tabla 30.

Pruebas de los efectos inter-sujetos

Variable dependiente: numero_compras

Origen	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
Modelo corregido	31,854 ^a	8	3,982	13,401	,000
Intersección	253,302	1	253,302	852,514	,000
Tratamiento	,864	2	,432	1,454	,234
segmento	24,549	2	12,274	41,311	,000
Tratamiento * segmento	,898	4	,224	,755	,554
Error	3981,162	13399	,297		
Total	4306,000	13408			
Total corregida	4013,016	13407			

a. R cuadrado = ,008 (R cuadrado corregida = ,007)

Tabla 30: Prueba de efectos inter sujetos. Variable dependiente: Número de compras. Efectos fijos: Nivel de tratamiento y Segmento de clientes. Fuente: Elaboración propia.

Sin embargo, a pesar de que esta diferencia no es globalmente significativa, sí es posible observar que, al igual que en la activación, existen diferencias importantes para los niveles de segmento, independiente del nivel de tratamiento, y dichos segmentos son los mismos que los apreciados con anterioridad, en cuyo caso el segmento con bajo *propensity score* es el más afectado de los tres, pues posee el menor promedio de número de compras y es significativamente menor en comparación con los otros dos segmentos (ver tabla 31).

Comparaciones por pares

Variable dependiente: numero_compras

Tratamiento	(I)segmento	(J)segmento	Diferencia de medias (I-J)	Error tip.	Sig. ^b	Intervalo de confianza al 95 % para la diferencia ^b	
						Límite inferior	Límite superior
Nivel 1	APS	BPS	,093*	,021	,000	,053	,134
		MPS	-,003	,021	,894	-,043	,038
	BPS	APS	-,093*	,021	,000	-,134	-,053
		MPS	-,096*	,017	,000	-,128	-,063
	MPS	APS	,003	,021	,894	-,038	,043
		BPS	,096*	,017	,000	,063	,128
Nivel 2	APS	BPS	,107*	,020	,000	,067	,146
		MPS	-,007	,020	,717	-,047	,032
	BPS	APS	-,107*	,020	,000	-,146	-,067
		MPS	-,114*	,017	,000	-,147	-,081
	MPS	APS	,007	,020	,717	-,032	,047
		BPS	,114*	,017	,000	,081	,147
Nivel 3	APS	BPS	,082*	,029	,005	,025	,139
		MPS	,018	,029	,532	-,039	,075
	BPS	APS	-,082*	,029	,005	-,139	-,025
		MPS	-,064*	,023	,006	-,110	-,018
	MPS	APS	-,018	,029	,532	-,075	,039
		BPS	,064*	,023	,006	,018	,110

Basadas en las medias marginales estimadas.

*. La diferencia de medias es significativa al nivel ,05.

b. Ajuste para comparaciones múltiples: Diferencia menos significativa (equivalente a la ausencia de ajuste).

Tabla 31: Comparación de pares entre segmento de clientes para los distintos segmentos de clientes. Variable dependiente: Número de compras. Efectos fijos: Nivel de tratamiento y Segmento de clientes. Fuente: Elaboración propia.

Si se invierte el análisis y se evalúa el tratamiento dentro de cada segmento de clientes, se logra apreciar que para el segmento de alto *propensity score* el nivel de tratamiento 2 posee el mejor promedio de número de compras, que corresponde a 0,02 órdenes de compras. Esto es superior al nivel 1 de tratamiento, que posee el peor promedio de órdenes de compras. Por otro lado, para el segmento con medio *propensity score*, el nivel 2 de tratamiento vuelve a ser el que presenta el mejor promedio de órdenes de compra (0,2 órdenes de compra) que es 0,04 órdenes de compras superior al peor tratamiento (Nivel 3 con 0,16 órdenes de compras en promedio). No obstante lo anterior, para el segmento de bajo *propensity score* se evidencia que el nivel de tratamiento que posee mejor promedio de órdenes de compra es el nivel de tratamiento 3, con 0,09 órdenes de compra, con 0,01 órdenes de compra superior respecto del nivel de tratamiento 1, que posee el peor valor (0,08, según la tabla 32).

Estimaciones

Variable dependiente: numero_compras

segmento	Tratamiento	Media	Error típ.	Intervalo de confianza 95%	
				Límite inferior	Límite superior
APS	Nivel 1	,177	,017	,144	,211
	Nivel 2	,198	,016	,166	,230
	Nivel 3	,181	,024	,134	,228
BPS	Nivel 1	,084	,012	,061	,107
	Nivel 2	,091	,012	,068	,114
	Nivel 3	,099	,017	,066	,132
MPS	Nivel 1	,180	,012	,157	,203
	Nivel 2	,205	,012	,182	,228
	Nivel 3	,163	,016	,131	,195

Tabla 32: Estimaciones de medias para nivel de tratamiento dentro de segmento de clientes. Variable dependiente: Número de compras. Efectos fijos: Segmento de clientes y Nivel de tratamiento. Fuente: Elaboración propia.

Las diferencias antes mencionadas se analizaron para cada nivel de tratamiento realizado en cada segmento de clientes, al igual que para la activación. En este caso, se destaca que para el segmento de medio *propensity score* sí existe un comportamiento significativamente diferente entre el tratamiento nivel 2 y nivel 3, al 95% de confianza. Esto se traduce en que, si bien no existe una activación superior en este subsegmento de clientes, estos sí generan un número de compras superior y mayor que el tratamiento de nivel 3, abriéndose así una opción de ganancia para la empresa, puesto que en este subsegmento sí hay un comportamiento favorable y el nivel de tratamiento 2 sí generó un cambio de comportamiento significativo para estos clientes, lo que se corrobora en la tabla 33.

Comparaciones por pares

Variable dependiente: numero_compras

segmento	(I) Tratamiento	(J) Tratamiento	Diferencia de medias (I-J)	Error típ.	Sig. ^b	Intervalo de confianza al 95 % para la diferencia ^b	
						Límite inferior	Límite superior
APS	Nivel 1	Nivel 2	-,020	,024	,389	-,067	,026
		Nivel 3	-,004	,029	,900	-,061	,054
	Nivel 2	Nivel 1	,020	,024	,389	-,026	,067
		Nivel 3	,017	,029	,563	-,040	,073
	Nivel 3	Nivel 1	,004	,029	,900	-,054	,061
		Nivel 2	-,017	,029	,563	-,073	,040
BPS	Nivel 1	Nivel 2	-,007	,017	,679	-,039	,026
		Nivel 3	-,015	,020	,477	-,055	,026
	Nivel 2	Nivel 1	,007	,017	,679	-,026	,039
		Nivel 3	-,008	,021	,710	-,048	,033
	Nivel 3	Nivel 1	,015	,020	,477	-,026	,055
		Nivel 2	,008	,021	,710	-,033	,048
MPS	Nivel 1	Nivel 2	-,025	,017	,135	-,058	,008
		Nivel 3	,017	,020	,395	-,022	,057
	Nivel 2	Nivel 1	,025	,017	,135	-,008	,058
		Nivel 3	,042 [*]	,020	,037	,003	,081
	Nivel 3	Nivel 1	-,017	,020	,395	-,057	,022
		Nivel 2	-,042 [*]	,020	,037	-,081	-,003

Basadas en las medias marginales estimadas.

*. La diferencia de medias es significativa al nivel ,05.

b. Ajuste para comparaciones múltiples: Diferencia menos significativa (equivalente a la ausencia de ajuste).

Tabla 33: Comparación de pares entre nivel de tratamiento dentro de cada segmento de clientes. Variable dependiente: Número de compras. Efectos fijos: Segmento de clientes y Nivel de tratamiento. Fuente: Elaboración propia.

Lo anterior puede apreciarse gráficamente en la ilustración 15, que representa el promedio de compras realizadas por los clientes durante el proceso de experimentación agrupado por nivel de tratamiento y por segmento de pertenencia; y se aprecia que efectivamente el segmento de bajo *propensity score* (BPS) posee una media menor que los otros dos segmentos de clientes.

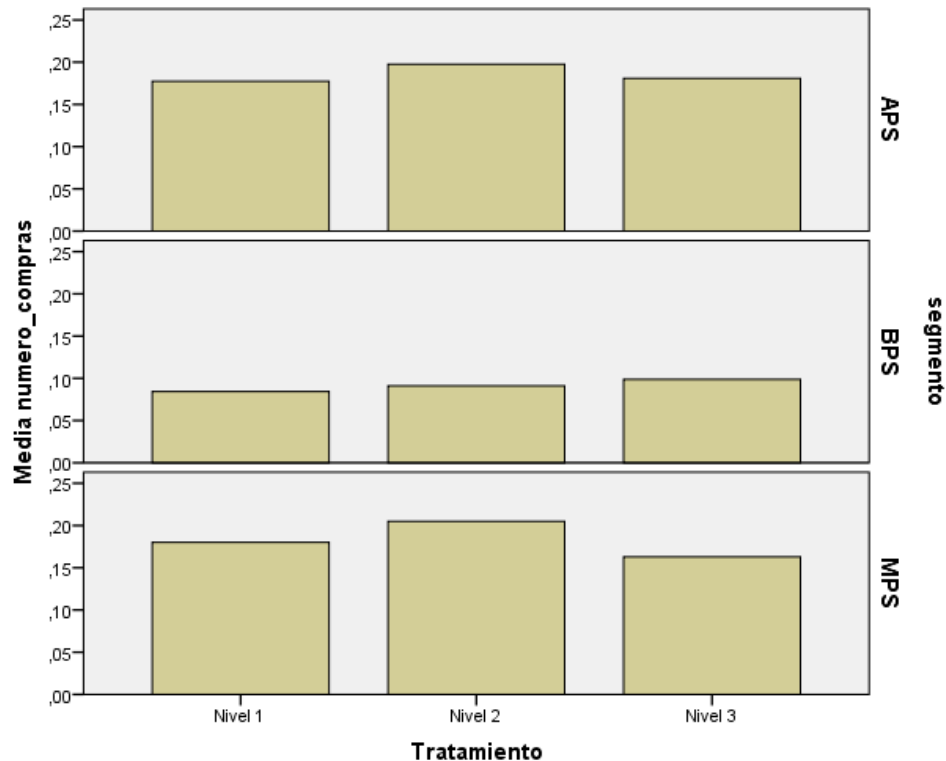


Ilustración 15: Gráfico de media de número de compras realizada en el periodo de experimentación segmento por segmento de clientes. APS: Alto propensity score, MPS: Medio propensity score, BPS: Bajo propensity score. Fuente: Elaboración propia en software IBM SPSS

A partir de los resultados obtenidos, es posible apreciar que, en términos generales, los distintos niveles de experimentación no generaron un cambio de comportamiento en los clientes y que, si bien existen diferencias en la activación y el número de compras, estas diferencias no pueden ser atribuidas a la existencia de la experimentación.

Es importante destacar que se apreció un alto índice de apertura de correos (28,7% para nivel 1 y 25,3% para nivel 2 de experimentación) por lo que no se puede asignar los malos resultados a esta variable sino que éstos se deben a que los clientes poseen un comportamiento particular y que es poco rentable para la empresa o bien que los incentivos no fueron los adecuados para el proceso de experimentación en función de los objetivos.

Además es posible que la hora en que fueron enviadas las intervenciones no fuera la mejor y queda propuesto experimentar utilizando la hora de envío como variable dentro de las posibles experimentaciones futuras.

A pesar de lo anterior, sí es posible establecer que, en términos generales, la segmentación realizada mediante el cálculo de *propensity score* para los clientes explica la diferencia de activación y del número de compras realizadas por los clientes y que esta diferencia es relevante cuando se compara el segmento de bajo *propensity score* con los otros dos segmentos, además los segmentos de medio y alto *propensity score* no presentan diferencias por lo que se podrían agrupar y diferenciar del tercer segmento antes mencionado. Además, esta diferencia se mantiene estable en los distintos niveles de tratamiento, por lo que se recomendaría en este caso segmentar a los clientes y enfocarse en esta segmentación para aumentar la activación y el número de compras, pero las intervenciones no es necesaria hacerla para todos.

Un caso particular detectado al analizar los clientes del segmento de medio *propensity score*, pues en esta población se aprecia que el nivel de tratamiento 2 sí genera un cambio de comportamiento cuando se compara con el nivel de tratamiento 3, en términos del número de compras realizadas, por lo que es posible concluir que en este segmento de clientes el tratamiento de nivel 2 sí es efectivo y sí debería desarrollarse para ellos, puesto que es el único subsegmento de clientes que presenta un cambio de comportamiento estadísticamente significativo.

A partir de este análisis, es posible interpretar que los clientes del sitio *web* presentan un comportamiento distinto al común de los clientes pues se destacan por realizar compras impulsivas y ocasionales, donde a pesar de personalizar el principal medio de comunicación y ofrecerles un descuento que podría ser más interesante para él, debido a la relación con sus preferencias pasadas y ,a pesar de ofrecer un premio si se concreta una prueba, estos no cambian su comportamiento y no se ven afectados por estas intervenciones, ya que no presentan un cambio si se comparan con los clientes a quienes no se les envió ningún correo adicional ni se les ofreció un premio. Lo anterior permite establecer que las intervenciones realizadas no fueron las mejores para fidelizar a los clientes al no verse incentivados por las intervenciones.

Por otro lado, se comprobó que la segmentación fue adecuada, por lo que se deberían utilizar estos segmentos para enfocar los esfuerzos de *marketing* pues éstos clientes son más propensos a ser buenos clientes y se deberían potenciar en la empresa. En particular, esto se puede realizar considerando las variables que son relevantes para calcular el *propensity score* que posteriormente da lugar a la segmentación y ejecutar acciones cuando los clientes cumplan con las variables relevantes. De esta forma, la distribución de buenos y malos clientes se puede modificar.

Finalmente, es posible caracterizar a los clientes que alteran el promedio de compras del nivel 2, es decir, caracterizar a los clientes que generan un alto número de compras en este nivel de tratamiento para el segmento de medio *propensity score* para encontrar clientes que cumplan el perfil y así poder mejorar los resultados de la empresa. Adicional a este segmento, se hace necesario entender a los clientes que lograron completar la meta de este nivel de tratamiento, pues existen 7 clientes que concretaron la meta de las 10 compras. Por el motivo anterior, se hace fundamental buscar el perfil de estos clientes para entender lo que hace que sean buenos clientes.

Más específicamente aún, hay un cliente que pertenece al segmento de bajo *propensity score* que logró la meta, con lo que se vuelve fundamental caracterizar este perfil para hacer más exitoso el proceso de fidelización de clientes que pertenecen a este segmento.

9.8 Determinación de clientes con mayor cambio de comportamiento debido a intervenciones.

Para la determinación de clientes con mayor cambio en su comportamiento condicional a la existencia de intervenciones, se utiliza un modelo de venta incremental. Dicho modelo busca contrastar el cambio de la probabilidad de ser un buen cliente sujeto a la existencia de experimentación.

9.8.1 Modelo *Uplift*

Para el cálculo de la venta incremental, se contrastaron dos modelos distintos, uno basado en un modelo *Logit*, el mismo utilizado para el cálculo del *propensity score*, y un modelo basado en *random forest*.

Lo anterior se realizó, puesto que en la literatura usualmente comparan estos dos modelos y se ha determinado que los modelos *uplift* basados en árboles de decisión tienden a representar una mejor ganancia en comparación con los modelos basados en regresiones, por lo que se hace pertinente comparar el rendimiento de estos (Radcliffe & Radcliffe, n.d.).

9.8.1.1 Comparación de modelos.

Como se mencionó previamente, se desarrollaron dos modelos distintos para el cálculo del modelo de venta incremental.

La comparación de los modelos se realiza mediante la diferencia de ganancia de información obtenida por ambos modelos, lo cual se mide mediante el coeficiente *qini*. El coeficiente *qini*, mide el área entre la curva de ganancia de información y la línea diagonal que representa una selección al azar. Dado lo anterior, un modelo que presente un mayor valor del coeficiente en cuestión representará el mejor modelo.

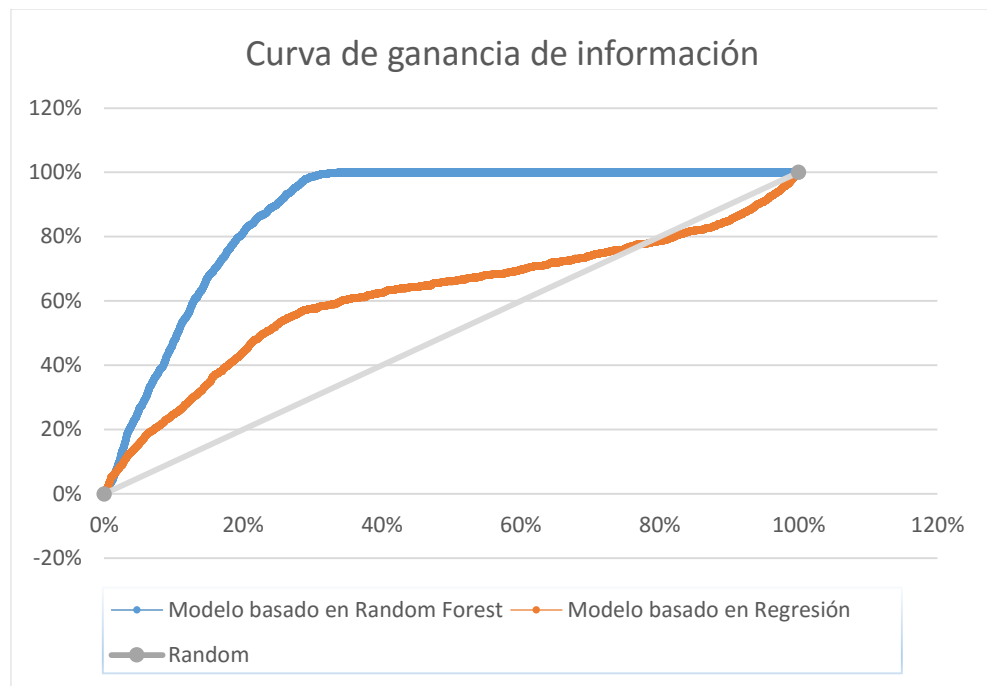


Ilustración 16: Curva de ganancia de información para modelos Uplift. Q de Qini para modelo basado en Random Forest: 2,76%. Q de Qini para modelo basado en Regresión: 0,07%. Fuente: Elaboración propia en software R.

Lo anterior se ve representado gráficamente en la ilustración 16, donde es posible observar que el modelo basado en *Random Forest* representa el mejor estimador para poder generar el modelo de venta incremental.

Se destaca que el modelo de *Random Forest*, que evidenció el mejor rendimiento basado en su criterio de separación, es el que utiliza la divergencia de Kullback-Leibler, la cual representa una medida no simétrica de similitud entre dos funciones de distribución de probabilidad. El mejor rendimiento de criterio de separación se representa cuantitativamente en la tabla 34, donde se ve que el método anterior posee el mejor valor para el valor de Q de qini.

	Divergencia Kullback-Leibler	Método de interacción	Divergencia Chi-Cuadrado	Distancia Euclidiana
Q de Qini	2,76%	2,19%	2,59%	2,42%

Tabla 34: Q de Qini para distintos método de separación para Random Forest. Fuente: Elaboración propia en software R.

9.8.1.2 Resultados de modelo

El modelo *uplift* se desarrolla ordenando a los clientes por el *lift* generado, debido a la existencia de tratamiento, a la diferencia entre probabilidad condicional a tratamiento y la probabilidad condicional al no-tratamiento, para luego ser agrupados en deciles. Con los deciles ya formados, se procede a calcular el porcentaje de órdenes de compras generadas por los clientes tratados y el porcentaje de órdenes de compras generadas por los clientes no tratados (control) y se calcula la diferencia entre estos dos valores. A este último valor se le denomina *Uplift* del decil y representa el incremento de órdenes de compras generadas, gracias al tratamiento (en porcentaje). El resultado del proceso anterior para los clientes intervenidos se presenta en la tabla 35 y gráficamente en la ilustración 18.

Decil	# Clientes tratados	# Clientes no-tratados	# Órdenes de compras - Tratamiento	# Órdenes de compra - Control	% Órdenes de compras - Tratamiento	% Órdenes de compra - Control	Uplift
1	1.071	270	307	19	28,43%	7,28%	21,1%
2	1.043	298	216	31	20,93%	10,03%	10,9%
3	1.063	279	154	27	14,60%	9,06%	5,5%
4	1.069	270	53	10	5,10%	3,46%	1,6%
5	1.061	280	39	8	3,65%	2,93%	0,7%
6	1.082	259	40	7	3,63%	2,93%	0,7%
7	1.056	284	43	12	4,04%	4,36%	-0,3%
8	1.088	253	49	16	4,50%	6,37%	-1,9%
9	1.097	244	52	34	4,72%	14,17%	-9,4%
10	1.083	258	101	94	9,34%	36,15%	-26,8%

Tabla 35: Tabla resumen de resultado modelo Uplift basado en Random Forest ordenado por deciles. Fuente: Elaboración propia en software R

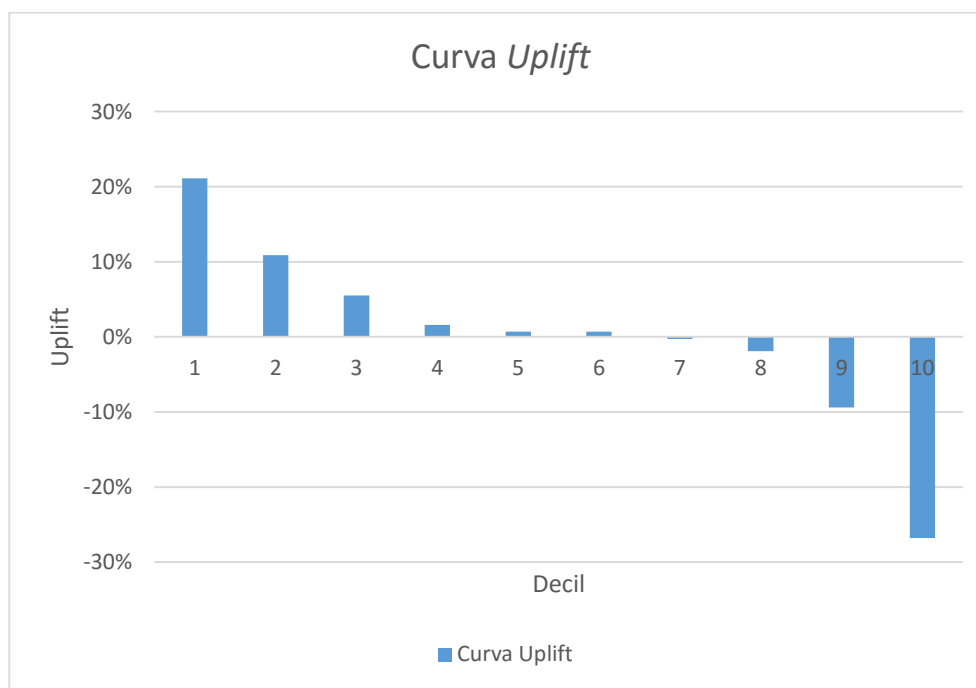


Ilustración 17: Curva uplift, modelo basado en Random Forest ordenado por deciles. Fuente: Elaboración propia

De los resultados anteriores, es posible apreciar que existen claras diferencias de comportamiento entre los distintos deciles, donde se destacan los dos primeros deciles que poseen un alto incremento en su proporción de venta (superior al 10%) y, por otro lado, existen dos deciles que poseen un evidente decaimiento en este valor que se refleja en su *uplift* negativo. Además, se aprecian 4 deciles con incremento bajo y dos deciles con decrecimiento marginal.

Los resultados anteriores pueden ser interpretados de la siguiente manera:

Los dos deciles con un incremento mayor se pueden clasificar como deciles de clientes persuadibles por intervenciones experimentales, pues la existencia de este

elemento provoca un considerable incremento en su porcentaje de venta incremental y a estos clientes sí se les recomienda intervenir experimentalmente, dado que responden positivamente a estos estímulos.

Por otro lado, los dos deciles que poseen el mayor decaimiento en su *uplift*, los dos deciles con *uplift* más negativos, se pueden clasificar como clientes “*boomerang*”, que son clientes con alta probabilidad de comprar si no son intervenidos; son clientes a los cuales el sobreenvío de correos puede ser molesto y conviene no molestarlos, por lo que se recomienda no intervenir.

En tercer lugar, existen cuatro deciles con incrementos en su *uplift* marginalmente positivo que representan a clientes que probablemente comprarían sin necesidad de una intervención adicional.

Finalmente, existen dos deciles con *uplift* marginalmente negativos, los cuales representan a causas perdidas, es decir, clientes que, independiente de la existencia o inexistencia de una intervención no comprarían.

Para los últimos dos casos, se recomienda no generar intervenciones, pues la existencia de intervenciones no generan cambios significativos en su comportamiento y la existencia de intervenciones solo generan un gasto mal realizado por parte de la compañía al no generar los beneficios esperados.

9.8.1.3 Caracterización de clientes

Con el fin de poder determinar un perfil de clientes que sean atractivos para poder intervenir en el futuro, se procederá a caracterizar a los clientes que pertenecen a los dos primeros deciles (clientes influenciables) y a los últimos dos deciles (clientes *boomerang*), para así poder ver qué características son distintivas entre estos grupos y así generar el perfil adecuado para la compañía.

9.8.1.3.1 Edad

En primer lugar se procederá a caracterizar la edad de los clientes de los deciles.

A continuación, se presenta el histograma de edad para los dos grupos de deciles de interés (Ver ilustración 18). Se puede apreciar que existe una tendencia clara en que los clientes de los dos últimos deciles posean una edad menor en comparación con los clientes de los dos primeros deciles que poseen una distribución más cargada a la derecha del histograma.

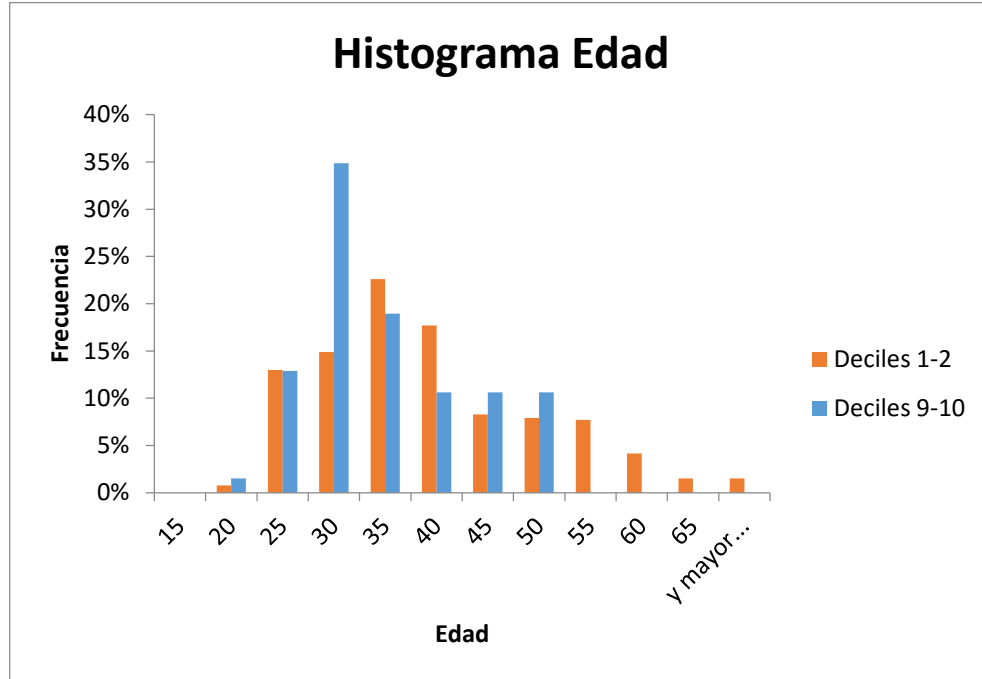


Ilustración 18: Histograma de edades para deciles de interés. Deciles 1, 2, 9 y 10. Fuente: Elaboración propia

Estadísticamente, lo anterior se refleja en que los primeros deciles poseen un promedio de edad de 36,94 años y los últimos dos tienen un promedio de 32,06 años (ver tabla 36) y se parecían mayores rangos de edad para los primeros deciles, pues el máximo valor es 71,9 en comparación con 46,27 para el cliente de mayor edad de los últimos deciles.

Estadísticamente, se realizó un test de medias para estos dos valores, el cual tiene como hipótesis nula que las medias de las dos muestras son iguales. Con los datos de las muestras se obtiene que la diferencia de medias es de 4,9 y poseen una diferencia significativa al 95% (ver anexos 13.9), lo que se traduce en que se rechaza la hipótesis nula de comportamiento y, por lo tanto, se puede asumir que las medias de las dos muestras no son iguales estadísticamente para la variable edad entre los dos grupos de clientes analizados.

	Deciles 1-2	Deciles 9-10
Promedio	36,94 años	32,06 años
Desviación estándar	10,74	7,68
Mínimo	18,57 años	17,46 años
Máximo	71,9 años	46,27 años

Tabla 36: Resumen de estadísticos descriptivos para edad. Fuente: Elaboración propia

9.8.1.3.2 Frecuencia de compras

En segundo lugar, se analizará la variable frecuencia de compra para los dos grupos de clientes. A partir del histograma de frecuencia de compra para los dos grupos de clientes (ver ilustración 19), se puede apreciar que los primeros deciles poseen una distribución con tendencia a un promedio de días menor en comparación con los últimos dos deciles que poseen una frecuencia de compra mayor, con un *peak* del histograma en una frecuencia mayor a los 150 días.

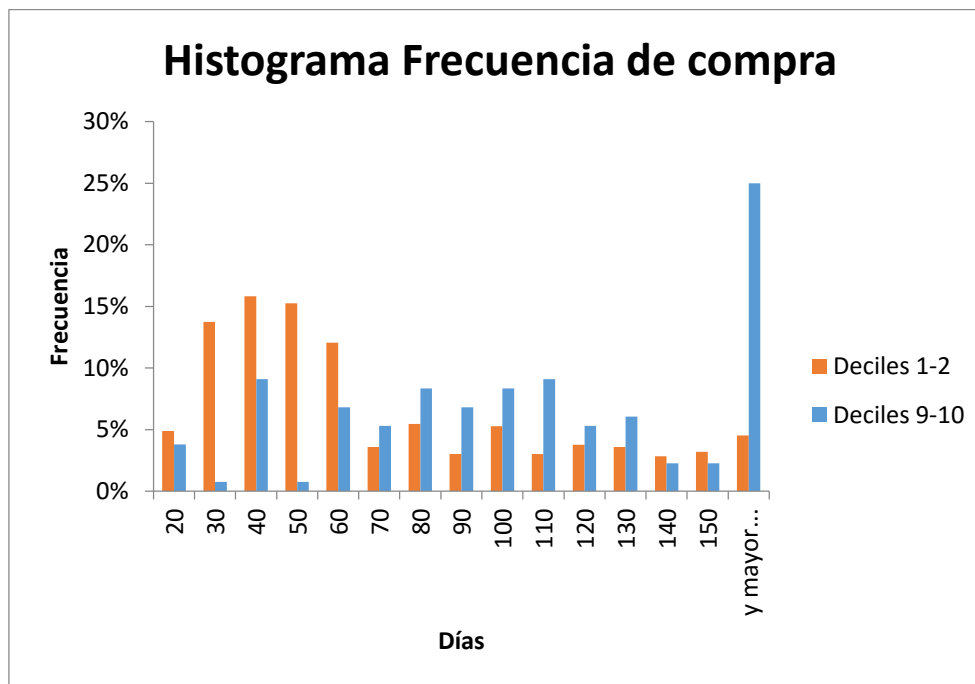


Ilustración 19: Histograma de frecuencia de compras para deciles 1, 2, 9 y 10. Fuente: Elaboración propia.

En términos objetivos, se aprecia que el promedio de días de frecuencia de compra en los primeros deciles es de 65,96 días, representado por un 60,8% del promedio de días de los últimos deciles, cuyo promedio es 108,36 días (ver tabla 37).

Para evaluar la diferencia de estos valores de manera estadística, se realizó un test de medias donde la hipótesis nula de comportamiento es que las medias de las poblaciones son iguales, pero a partir de los datos se obtiene que la diferencia de medias es de -42,5 y es estadísticamente significativo al 95% (ver anexos 13.10), es decir, existe un comportamiento estadísticamente distinto entre los dos grupos de clientes.

	Deciles 1-2	Deciles 9-10
Promedio	65,96 días	108,36 días
Desviación estándar	44,96	57,95
Mínimo	0 días	0 días
Máximo	284,3 días	315,06 días

Tabla 37: Estadísticos descriptivos para variable Frecuencia de compra. Fuente: Elaboración propia.

9.8.1.3.3 Activación

En una primera instancia, se puede apreciar que las distribuciones de esta variable para los dos grupos de interés siguen un mismo patrón y no se puede determinar un comportamiento diferente solo con el gráfico presentado. (Ver ilustración 20).

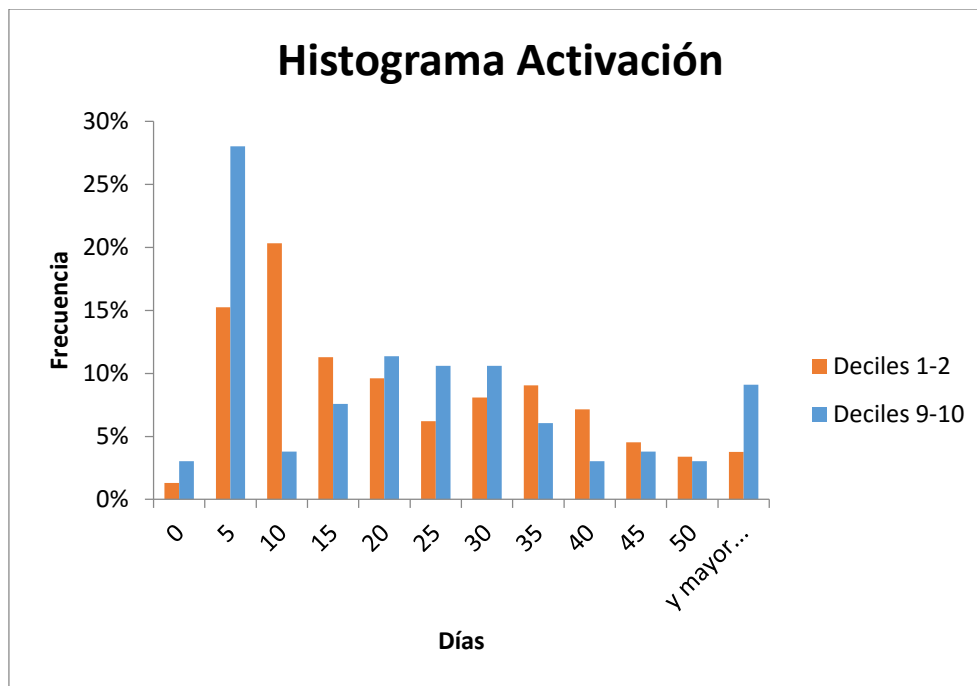


Ilustración 20: Histograma de frecuencia de compras para deciles 1, 2, 9 y 10. Fuente: Elaboración propia.

En términos concretos, se aprecia que los primeros deciles poseen un mayor tiempo de activación; pero la diferencia es de aproximadamente 0,5 días (ver tabla 38), lo que hace suponer –en un comienzo– que la diferencia no es lo suficientemente grande como para asignar un comportamiento distinto entre los grupos.

Estadísticamente se realizó un test de medias para poder evaluar la diferencia entre estos dos grupos, la hipótesis nula es igualdad de medias para las dos muestras, y se obtiene que la de medias es de 0,5; pero su valor estadístico (p-valor) es de $0,745 > 0,05$ (ver anexos 13.11), por lo que no se puede rechazar la hipótesis nula, lo que implica que la diferencia no es significativa por lo que se puede concluir que la variable activación no representa una variable significativa de diferencia de comportamiento para estos dos grupos.

	Deciles 1-2	Deciles 9-10
Promedio	19,43 días	18,92 días
Desviación estándar	15,22	16,82
Mínimo	0 días	0 días
Máximo	60,49 días	1.115,95 días

Tabla 38: Estadísticos descriptivos para variable Tiempo de activación. Fuente: Elaboración propia.

9.8.1.3.4 Categoría de compra

Para esta sección, se analizarán las compras realizadas por los clientes durante el proceso de experimentación y se considerarán las tres categorías más importantes: “Productos”, “gastronomía” y “Health & beauty”. El resto de las categorías se agruparán en una variable nueva denominada “otras”.

A partir de la proporción de compras generadas por categorías realizadas por los clientes de los primeros deciles durante el proceso de experimentación, se puede observar que la categoría “Productos” es la preferida de estos clientes, abarcando el

71% de las compras realizadas (ver ilustración 21). Además, se ve una tendencia similar para los últimos deciles pues la categoría “Productos” también es predominante en este segmento de clientes, pero en menor cuantía comparado con los primeros deciles, lo que produce una diferencia entre ellos de un 10% (ver ilustraciones 21 y 22). Dicha diferencia se distribuye principalmente a la categoría “Health & Beauty” en los últimos deciles.

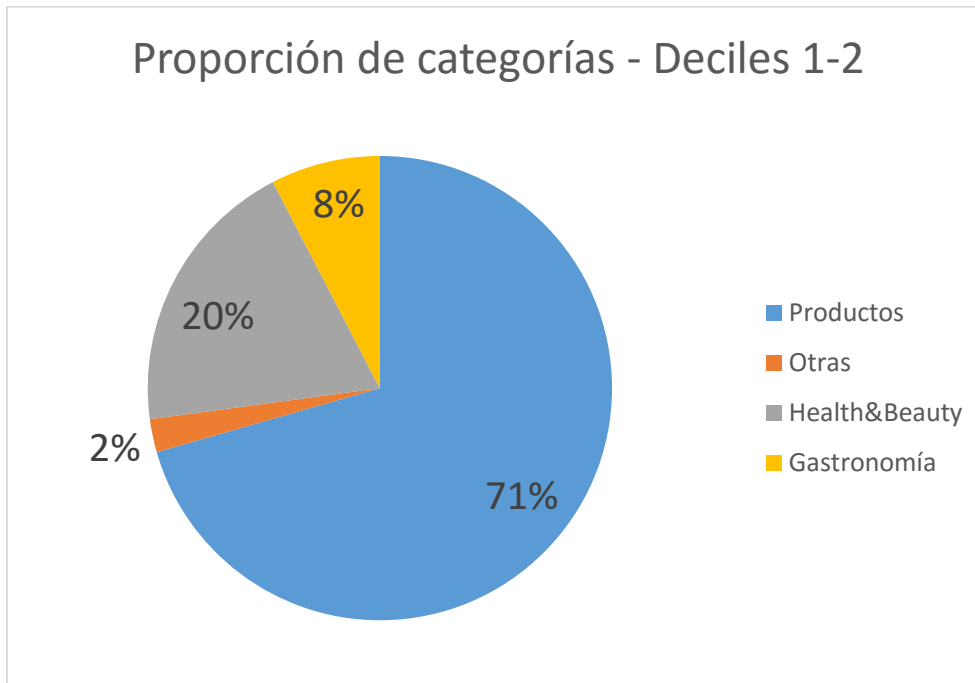


Ilustración 21: Distribución de categorías de órdenes de compras generadas por los deciles 1 y 2 durante proceso de experimentación. Fuente: Elaboración propia

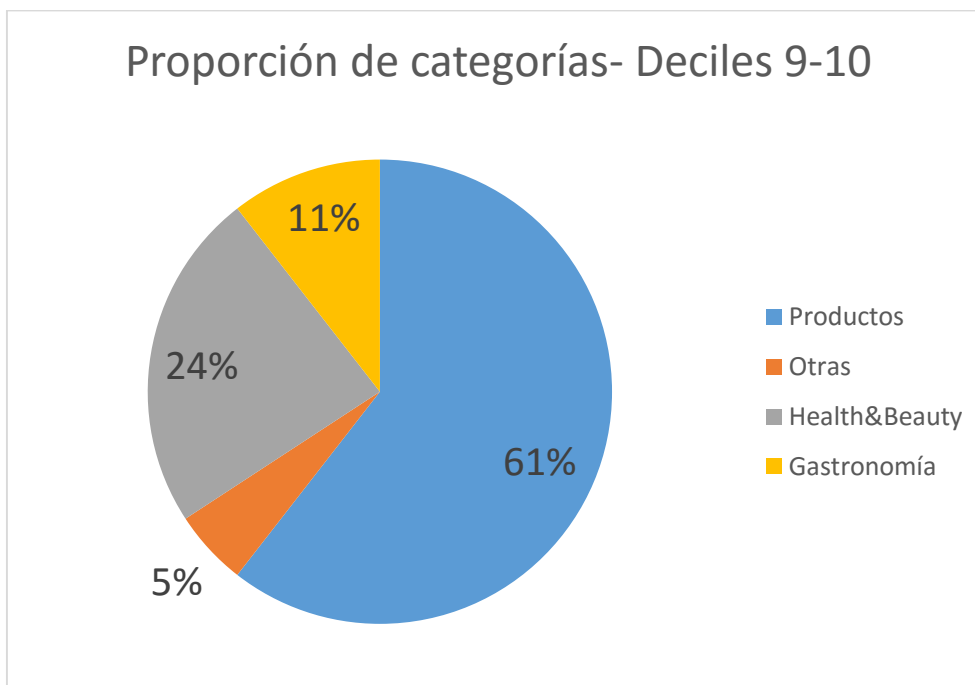


Ilustración 22: Distribución de categorías de órdenes de compras generadas por los deciles 9 y 10 durante proceso de experimentación. Fuente: Elaboración propia

En términos estadísticos concretos, se calculó un test de proporciones para dos muestras por medio de una tabla de contingencia (ver anexos 13.12) para comparar la diferencia de proporción de compras asociadas a la categoría Productos, que es la que posee la mayor proporción. Este test posee una hipótesis nula que consiste en que las dos proporciones son iguales, es decir, que ambos grupos de clientes compran la misma cantidad de Productos. De los datos obtenidos, se desprende un valor del estadístico *Chi-cuadrado* de 2,3, lo que implica una significancia de $0,129 > 0,05$ (ver anexos 13.13) y se traduce en que la diferencia no es significativa.

Debido a esto, es posible asumir que no existe una diferencia de comportamiento en la distribución de compras realizadas entre los deciles.

	# Órdenes de compras – Deciles 1-2	% Órdenes de compras – Deciles 1-2	# Órdenes de compras – Deciles 9-10	% Órdenes de compras – Deciles 9-10
Productos	611	71%	46	61%
Health & Beauty	169	20%	18	24%
Gastronomía	66	8%	8	11%
Otras	51	2%	4	5%

Tabla 39: Distribución de órdenes de compra según categoría para deciles 1, 2, 9 y 10. Fuente: Elaboración propia

9.8.1.3.5 Subcategoría de compra

En esta sección, se analizará la distribución de las compras realizadas según la subcategoría. Se destaca que la subcategoría corresponde a un nivel mayor de detalle de la caracterización de los bienes adquiridos; es un nivel mayor que la categoría (elemento revisado con anterioridad). En particular, se analizarán las cuatro subcategorías principales: “Hogar”, “otros restaurantes”, “productos de gran consumo” (PGC) y “tratamientos reductivos”. El resto de las subcategorías se agruparán en una nueva, denominada “otros”. Se destaca, además, que las subcategorías “hogar” y “productos de gran consumo” se asocian a la categoría “productos”, la subcategoría “otros restaurantes” se asocia a la categoría “gastronomía” y la subcategoría “tratamientos reductivos” se asocia a la categoría “*health & beauty*”

A partir de los gráficos presentados (ver ilustración 23 y 24), se aprecia una predominancia en la subcategoría “productos de gran consumo”, lo cual es esperable debido a la tendencia de compra de productos de los clientes y la diferencia entre los dos grupos es de 5,7% (67,7% de predominancia en los primeros deciles versus 62% de los últimos deciles de clientes); pero la distribución de las demás subcategorías cambia, pues en los últimos deciles aumenta la proporción de “otros”.

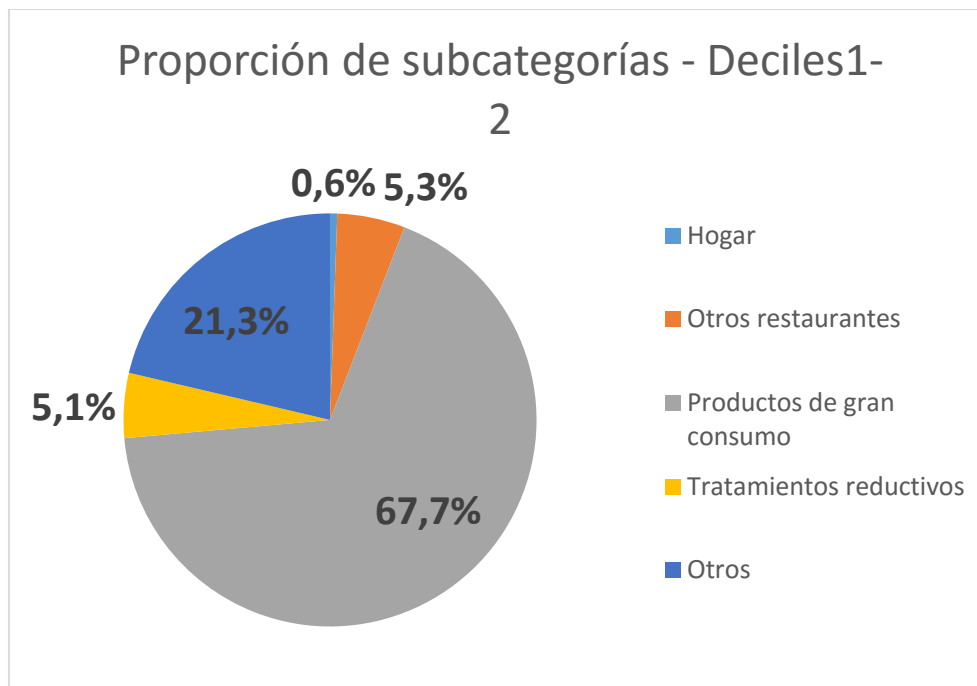


Ilustración 23: Distribución de subcategorías de órdenes de compras generadas por los deciles 1 y 2 durante proceso de experimentación. Fuente: Elaboración propia

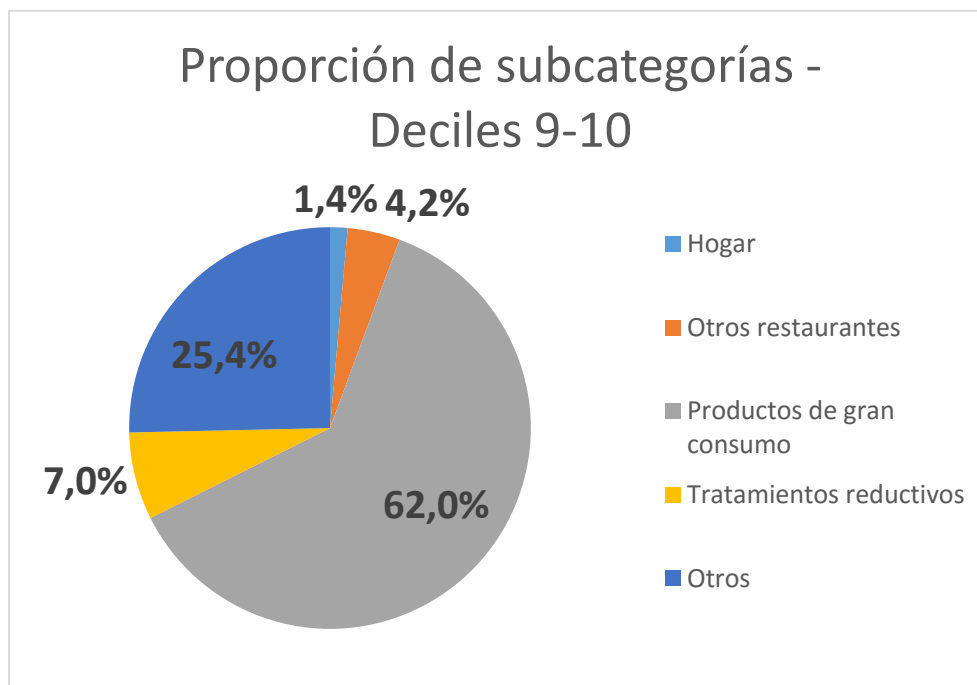


Ilustración 24: Distribución de subcategorías de órdenes de compras generadas por los deciles 9 y 10 durante proceso de experimentación. Fuente: Elaboración propia

Nuevamente se realizó un test de proporciones para la subcategoría principal de compra de los segmentos de clientes, donde la hipótesis nula indica igualdad de proporción en las poblaciones de estudio. Sobre la base de los datos obtenidos, se obtiene un valor de *Chi-cuadrado* de 3,054 que representa una significancia (p-valor) de $0,081 > 0,05$ (ver anexos 13.15), por lo que se rechaza la hipótesis nula. Esto

implica que no existen diferencias significativas en la proporción de subcategorías de las compras realizadas por los clientes.

	# Órdenes de compras – Deciles 1-2	% Órdenes de compras – Deciles 1-2	# Órdenes de compras – Deciles 9-10	% Órdenes de compras – Deciles 9-10
Productos de gran consumo (PGC)	600	67,7%	44	62,0%
Tratamientos reductivos	45	5,1%	5	7,0%
Otros restaurantes	47	5,3%	3	4,2%
Hogar	5	0,6%	1	1,4%
Otras	189	21,3%	18	25,4%

Tabla 40: Distribución de órdenes de compra según subcategoría para deciles 1, 2, 9 y 10. Fuente: Elaboración propia

10 Recomendaciones y trabajo futuro

10.1 Recomendaciones

A partir de lo evidenciado, es posible recomendar a la empresa no realizar estas intervenciones experimentales a todos sus clientes si sólo buscan optimizar sus recursos en términos contables netos, ya que la inversión en implementación de estas actividades pueden ser mayores a los retornos esperados. Sin embargo, si la empresa busca mejorar otras variables no cuantificables, como la cercanía con clientes y felicidad, entonces el nivel de experimentación 2 (club de fidelización) podría ser una buena iniciativa, pues genera una conexión concreta con los clientes que sí se vieron incentivados. En particular, esta iniciativa se podría implementar solo para un grupo de clientes, como los segmentos con *propensity score* medio y/o alto, pues ellos sí compran en el sitio y con esta iniciativa tendrían un motivo adicional para escoger este sitio particular y no la competencia. Además, actividades como estas pueden mejorar la percepción de clientes y mejorar las recomendaciones, mejorando el *NPS (net promoter score)* de la empresa, lo que resulta relevante, pues la misión y visión de la empresa van alineadas con este indicador. Por consiguiente, la inversión podría compensarse en la mejora de este indicador, aunque no lo sea en términos contables.

Finalmente, se recomienda analizar actividades e iniciativas tipo *trigger* a los clientes que cumplan el perfil de clientes con 37 años y frecuencia de compra de 66 días, pues se determinó que los clientes con estas características responden estadísticamente mejor a las intervenciones. De este modo, enviarle intervenciones podría mejorar la calidad de clientes de estos perfiles.

10.2 Trabajo futuro

Si bien, estos dos niveles de intervenciones experimentales no conllevaron un cambio significativo de comportamiento, es posible que otras acciones sí produzcan efectos en los clientes, por lo que se recomienda seguir buscando intervenciones para encontrar la que mejore la calidad de los clientes de la empresa. Un ejemplo de intervención futura posible de testear consiste en la entrega de despacho gratuito, pues se ha determinado que más del 50% de los clientes de comercio *online* se han declarado sensibles al costo de despacho. Así, esta iniciativa podría mejorar y

cambiar el comportamiento de los usuarios, especialmente en Chile, donde no existen sitios que posean esta característica.

Además se recomienda manipular otras variables que puedan afectar el comportamiento de los clientes como la hora de envío, la estructura del newsletter o correo que se desee enviar con el fin de encontrar la comunicación que genere el mejor rendimiento a las comunicaciones realizadas independiente del contenido de éste (que pueden ser incentivos o no para los clientes).

11 Conclusiones

Una vez finalizado el proyecto, y luego de la obtención y análisis de los resultados, es posible obtener conocimiento y conclusiones del proceso experimental realizado.

En primer lugar, se puede apreciar que la segmentación realizada en la primera etapa del proyecto fue capaz de caracterizar e identificar a los potenciales buenos clientes de la empresa, con el cual se identificaron las principales variables que afectan el comportamiento de compra en el sitio *web*. Este punto es importante, pues con esta información se pueden estructurar de mejor manera las acciones que se deben realizar por la compañía, enfocando sus acciones de *marketing* de captación y retención de clientes a las personas que posean un *Propensity Score* dentro de los determinados que sí poseen un comportamiento favorable. En particular, la empresa podría instaurar el cálculo de este valor a todos sus clientes, con el fin de realizar las acciones antes mencionadas.

En segundo lugar, se puede concluir que los clientes no son influenciados por las intervenciones realizadas, pues ninguno de los dos niveles de experimentación logró generar un cambio de comportamiento estadísticamente significativo, inclusive, las diferencias percibidas puede que se deban a ruido generado en el proceso experimental o, quizás, al simple hecho de que los clientes intervenidos tuvieron un correo más en su casilla personal y no necesariamente por la intervención. Aún así, estas intervenciones no fueron fructíferas en términos de órdenes de compras generadas y en la variable cantidad de clientes activados. En términos de las hipótesis de comportamiento planteadas en un comienzo es posible ver que, con respecto a la hipótesis 1, los clientes efectivamente no ven las ofertas, no abren el correo electrónico y por lo tanto, a pesar que se le puedan ofrecer ofertas más personalizadas, estas no son efectivas pues el problema está en el primer paso que es abrir el correo y el cambio en el asunto del correo no es efectivo. Por otro lado, con respecto a la hipótesis 2, se verificó que los clientes de comercio online como las cuponeras, no responden efectivamente a incentivos de aceleración de compra con un premio cuando se cumple la meta estipulada.

En tercer lugar, a pesar de que las intervenciones no generan cambio de comportamiento significativo para toda la población, sí se pudieron identificar subsegmentos de clientes que sí responden positivamente frente a intervenciones experimentales mediante el modelo de venta incremental y, de forma análoga, se determinaron segmentos de clientes que responden muy negativamente a interacciones adicionales. Con este descubrimiento, es posible establecer un perfil de clientes que responden positivamente para poder mejorar aún más las estrategias de *marketing* de la empresa. El perfil de cliente corresponde a aquellos con 37 años, con un promedio de frecuencia de compra de 66 días, ya que sólo estas variables son estadísticamente significativas en comparación con los dos peores deciles de clientes. Por lo tanto, con este perfil se pueden ejecutar acciones en base a *triggers* que se ejecuten cuando se detecten clientes de este estilo.

En resumen, se puede concluir que los segmentos de alto y medio *propensity score* sí tienen un comportamiento distinto y que los clientes que cumplen con el perfil descrito deben ser intervenidos experimentalmente. Sin embargo, no es necesario realizar esto para toda la población, pues las intervenciones no son capaces de modificar el comportamiento de los clientes en términos estadísticos.

12 Bibliografía

- Araya, M. (2011). *Estimación de la efectividad promocional en un supermercado mayorista*. Universidad de Chile.
- Davis, S., & Smith, R. (2004). *An introduction to statistics and research methods*. Prentice Hall.
- Fuenzalida, B. (2012). *Pronóstico de efectividad de promociones sobre clientes, dadas sus características y respuestas pasadas*. Universidad de Chile.
- Hill, P. (2014). US\$ 2.000 millones de ventas por internet. Retrieved from <http://www.lun.com/lunmobile//Pages/NewsDetailMobile.aspx?dt=2014-05-10&PaginaId=24&SupplementId=0&bodyid=0&IsNPHR=1>
- Imhmed, Mohmed; Binti, Azizan; Zalisham, J. (2013). *The impact of trust and past experience on intention to purchase in E-Commerce*.
- Marchant, I. (2010). No Title. Retrieved from <http://colombia.ecommercelatam.com/wp-content/uploads/2010/12/Marchant.pdf>
- Mohd, N., & Ismail, S. (n.d.). *Antecedents of customer loyalty in an e-commerce setting: an empirical study*. Doha, Qatar.
- Muzumdar, A. (2010). *Predicting customer purchase in an online retail business, a Data Mining approach*. Rourkela.
- Qureshi, I., Yulin, F., Ramsey, E., McCole, P., Ibbotson, P., & Compeau, D. (2009). Understanding online customer repurchasing intention and the mediating role of trust - An empirical investigation in two developed countries. *European Journal of Information Systems*, 205–222.
- Radcliffe, N. J., & Radcliffe, N. J. (n.d.). Using Control Groups to Target on Predicted Lift: *Methods*, (3).
- Software, P. (2006). Optimal Targeting through Uplift Modeling : Generating higher demand and increasing customer retention while reducing marketing costs . Uplift Modeling : An Overview, (December).
- Wansink, B. (2003). Developing a cost effective brand loyalty program. *Journal of Advertising Research*, 43(03), 301–309.
- Wen, C., Prybutok, V., & Xu, C. (2011). An integrated model for customer online repurchase intention. *Journal of Computer Information Systems*.
- Zhang, Y., Fang, Y., Wei, K.-K., Ramsey, E., McCole, P., & Chen, H. (2010). *Repurchase intention in B2C e-commerce - A relationship quality perspective*.

13 Anexos

13.1 Anexo 1: Histograma de compras

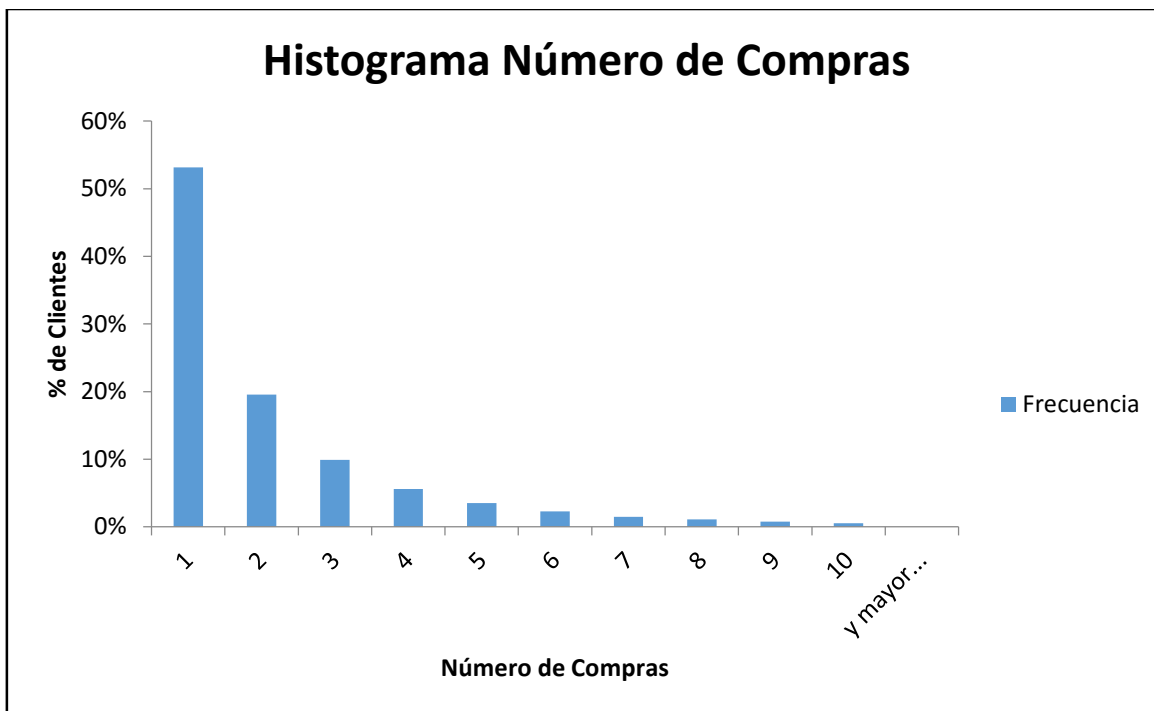


Ilustración 25: Distribución de número de compras históricas en la empresa. Fuente: Elaboración propia.

13.2 Anexos 2: Resultados regresión logística 1

Call:

```
glm2(formula = indep ~ Productos + HealthBeauty + Gastronomía,  
      family = binomial(), data = new_compras2, control = list(maxit = 50),  
      singular.ok = TRUE)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-0.967	-0.946	-0.849	1.428	1.695

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.1656	0.0453	-25.75	< 2e-16 ***
Productos	0.5928	0.0466	12.72	< 2e-16 ***
HealthBeauty	0.3296	0.0490	6.73	1.7e-11 ***
Gastronomía	0.6482	0.0518	12.52	< 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 75543 on 58742 degrees of freedom
Residual deviance: 75224 on 58739 degrees of freedom
AIC: 75232

Number of Fisher Scoring iterations: 4

Ilustración 26: Resultado modelo 1 – Categorías: producto, health & beauty y gastronomía

13.3 Anexos 3: Resultados regresión logística 2

Call:

```
glm2(formula = indep ~ Genero + edad + Productos + HealthBeauty +  
      Gastronomía + Hogar + OtrosRestaurantes + PGC + TratamientosReductivos,  
      family = binomial(), data = new_compras2, control = list(maxit = 50),  
      singular.ok = TRUE)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.155	-0.835	-0.756	1.306	1.821

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-1.812378	0.056521	-32.07	< 2e-16	***
Genero	0.110544	0.019397	5.70	1.2e-08	***
edad	0.017366	0.000925	18.78	< 2e-16	***
Productos	2.653837	0.063274	41.94	< 2e-16	***
HealthBeauty	0.293332	0.050855	5.77	8.0e-09	***
Gastronomía	0.831640	0.066657	12.48	< 2e-16	***
Hogar	-0.581987	0.060342	-9.64	< 2e-16	***
OtrosRestaurantes	-0.207945	0.056954	-3.65	0.00026	***
PGC	-2.519356	0.046057	-54.70	< 2e-16	***
TratamientosReductivos	0.062748	0.042126	1.49	0.13635	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 75543 on 58742 degrees of freedom

Residual deviance: 68994 on 58733 degrees of freedom

AIC: 69014

Number of Fisher Scoring iterations: 4

Ilustración 28: Resultado modelo 3: Género, edad, categorías: productos, health & beauty, gastronomía y subcategorías: hogar, otros restaurantes, productos de gran consumo y tratamientos reductivos

13.4 Anexos 4: Resultados regresión logística 3

Call:

```
glm2(formula = indep ~ monto_loc + Genero + edad + Productos +  
      HealthBeauty + Gastronomía + Hogar + OtrosRestaurantes +  
      PGC + TratamientosReductivos, family = binomial(), data = new_compras2,  
      control = list(maxit = 50), singular.ok = TRUE)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.157	-0.835	-0.754	1.305	1.939

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-1.78e+00	5.87e-02	-30.35	< 2e-16	***
monto_loc	-4.47e-07	2.44e-07	-1.83	0.06676	.
Genero	1.10e-01	1.94e-02	5.67	1.4e-08	***
edad	1.74e-02	9.25e-04	18.81	< 2e-16	***
Productos	2.63e+00	6.47e-02	40.65	< 2e-16	***
HealthBeauty	2.68e-01	5.25e-02	5.10	3.3e-07	***
Gastronomía	8.04e-01	6.82e-02	11.78	< 2e-16	***
Hogar	-5.80e-01	6.04e-02	-9.60	< 2e-16	***
OtrosRestaurantes	-2.08e-01	5.70e-02	-3.65	0.00027	***
PGC	-2.52e+00	4.61e-02	-54.66	< 2e-16	***
TratamientosReductivos	6.19e-02	4.21e-02	1.47	0.14144	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 75543 on 58742 degrees of freedom
Residual deviance: 68991 on 58732 degrees of freedom
AIC: 69013

Number of Fisher Scoring iterations: 4

Ilustración 29: Resultado modelo 4 – Monto de primera compra, género, edad, categorías: productos, health & beauty, gastronomía y subcategorías: hogar, otros restaurantes, productos de gran consumo y tratamientos reductivos

13.5 Anexos 5: Resultados regresión logística 4

Call:

```
glm2(formula = indep ~ frecuencia + edad + Productos + HealthBeauty +  
      Gastronomía + Hogar + OtrosRestaurantes + PGC + Genero:TratamientosReductivos,  
      family = binomial(), data = new_compras2, control = list(maxit = 50),  
      singular.ok = TRUE)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.135	-0.836	-0.754	1.316	1.835

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-1.74e+00	5.57e-02	-31.23	< 2e-16	***
frecuencia	-1.81e-04	5.22e-05	-3.47	0.00052	***
edad	1.72e-02	9.24e-04	18.57	< 2e-16	***
Productos	2.66e+00	6.33e-02	42.03	< 2e-16	***
HealthBeauty	3.19e-01	5.01e-02	6.36	2e-10	***
Gastronomía	8.34e-01	6.67e-02	12.51	< 2e-16	***
Hogar	-5.73e-01	6.03e-02	-9.50	< 2e-16	***
OtrosRestaurantes	-2.02e-01	5.70e-02	-3.55	0.00039	***
PGC	-2.51e+00	4.60e-02	-54.59	< 2e-16	***
Genero:TratamientosReductivos	1.27e-01	4.66e-02	2.72	0.00662	**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 75442 on 58639 degrees of freedom
Residual deviance: 68920 on 58630 degrees of freedom
(103 observations deleted due to missingness)
AIC: 68940

Number of Fisher Scoring iterations: 4

Ilustración 30: Resultado modelo 5 - Frecuencia, edad, categorías: productos, health & beauty, gastronomía y subcategorías: hogar, otros restaurantes, productos de gran consumo y tratamientos reductivos sólo si género es femenino.

13.6 Anexos 6: Resultados regresión logística 5

Call:

```
glm2(formula = indep ~ activacion + frecuencia + edad + Productos +  
      Genero:HealthBeauty + Gastronomía + Hogar + OtrosRestaurantes +  
      PGC + Genero + TratamientosReductivos, family = binomial(),  
      data = new_compras2, control = list(maxit = 50), singular.ok = TRUE)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.335	-0.817	-0.723	1.109	1.883

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-1.77e+00	4.56e-02	-38.80	< 2e-16	***
activacion	9.69e-04	3.35e-05	28.93	< 2e-16	***
frecuencia	-1.06e-03	6.02e-05	-17.57	< 2e-16	***
edad	1.64e-02	9.40e-04	17.48	< 2e-16	***
Productos	2.53e+00	5.50e-02	46.04	< 2e-16	***
Gastronomía	7.21e-01	5.90e-02	12.22	< 2e-16	***
Hogar	-5.56e-01	6.12e-02	-9.08	< 2e-16	***
OtrosRestaurantes	-2.00e-01	5.79e-02	-3.45	0.00056	***
PGC	-2.47e+00	4.67e-02	-52.93	< 2e-16	***
Genero	5.09e-02	2.15e-02	2.36	0.01819	*
TratamientosReductivos	6.46e-02	4.20e-02	1.54	0.12466	
Genero:HealthBeauty	2.67e-01	4.14e-02	6.44	1.2e-10	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 74494 on 58088 degrees of freedom

Residual deviance: 67208 on 58077 degrees of freedom

(654 observations deleted due to missingness)

AIC: 67232

Number of Fisher Scoring iterations: 4

Ilustración 31: Resultado modelo 6 – Activación, frecuencia, edad, categorías: productos, health & beauty sólo si género es femenino, gastronomía y subcategorías: hogar, otros restaurantes, productos de gran consumo y tratamientos reductivos

13.7 Anexos 7: Resultados regresión logística 6

Call:

```
glm2(formula = indep ~ recency +activacion + frecuencia +  
      edad + Productos + Genero:HealthBeauty + Gastronomía + Hogar +  
      OtrosRestaurantes + PGC + Genero + TratamientosReductivos,  
      family = binomial(), data = new_compras2, control = list(maxit = 50),  
      singular.ok = TRUE)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.532	-0.836	-0.684	1.062	2.219

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-1.38e+00	4.70e-02	-29.43	< 2e-16	***
recency	-5.17e-03	1.46e-04	-35.43	< 2e-16	***
activacion	1.08e-03	3.43e-05	31.53	< 2e-16	***
frecuencia	-1.33e-03	6.14e-05	-21.74	< 2e-16	***
edad	1.72e-02	9.50e-04	18.11	< 2e-16	***
Productos	2.61e+00	5.59e-02	46.74	< 2e-16	***
Gastronomía	7.48e-01	5.99e-02	12.49	< 2e-16	***
Hogar	-5.23e-01	6.23e-02	-8.39	< 2e-16	***
OtrosRestaurantes	-2.56e-01	5.89e-02	-4.35	1.4e-05	***
PGC	-2.55e+00	4.76e-02	-53.58	< 2e-16	***
Genero	4.11e-02	2.18e-02	1.89	0.059	.
TratamientosReductivos	1.72e-01	4.27e-02	4.04	5.4e-05	***
Genero:HealthBeauty	2.63e-01	4.19e-02	6.28	3.4e-10	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 74494 on 58088 degrees of freedom

Residual deviance: 65876 on 58076 degrees of freedom

(654 observations deleted due to missingness)

AIC: 65902

Number of Fisher Scoring iterations: 4

Ilustración 32: Resultado modelo 7 – Recency, activación, frecuencia, edad, categorías: productos, health & beauty sólo si género es femenino, gastronomía y subcategorías: hogar, otros restaurantes, productos de gran consumo y tratamientos reductivos

13.8 Anexos 8: Resultados regresión logística 7

Call:

```
glm2(formula = indep ~ recency + activacion + frecuencia +  
      log(edad) + Productos + Genero:HealthBeauty + Gastronomía +  
      Hogar + OtrosRestaurantes + PGC + TratamientosReductivos,  
      family = binomial(), data = new_compras2, control = list(maxit = 50),  
      singular.ok = TRUE)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.524	-0.839	-0.682	1.057	2.232

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-3.18e+00	1.27e-01	-25.07	< 2e-16	***
recency	-5.19e-03	1.46e-04	-35.54	< 2e-16	***
activacion	1.07e-03	3.43e-05	31.33	< 2e-16	***
frecuencia	-1.33e-03	6.14e-05	-21.71	< 2e-16	***
log(edad)	6.85e-01	3.50e-02	19.54	< 2e-16	***
Productos	2.62e+00	5.56e-02	47.15	< 2e-16	***
Gastronomía	7.66e-01	5.97e-02	12.84	< 2e-16	***
Hogar	-5.20e-01	6.23e-02	-8.35	< 2e-16	***
OtrosRestaurantes	-2.53e-01	5.89e-02	-4.29	1.8e-05	***
PGC	-2.54e+00	4.75e-02	-53.47	< 2e-16	***
TratamientosReductivos	1.65e-01	4.27e-02	3.88	0.00011	***
Genero:HealthBeauty	3.00e-01	3.81e-02	7.87	3.7e-15	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 74494 on 58088 degrees of freedom

Residual deviance: 65821 on 58077 degrees of freedom

(654 observations deleted due to missingness)

AIC: 65845

Number of Fisher Scoring iterations: 4

Ilustración 33: Resultado modelo 8 – Recency, activación, frecuencia, logaritmo de edad, categorías: productos, health & beauty sólo cuando género es femenino, gastronomía y subcategorías: hogar, otros restaurantes, productos de gran consumo y tratamientos reductivos

13.9 Anexos 9: Resultados test de medias. Caracterización clientes deciles 1, 2, 9 y 10. Variable Edad

Prueba de muestras independientes

		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia	
									Inferior	Superior
Edad	Se han asumido varianzas iguales	12,649	,000	4,905	661	,000	4,8731908	,9935204	2,9223545	6,8240271
	No se han asumido varianzas iguales			5,977	273,530	,000	4,8731908	,8153338	3,2680638	6,4783178

Ilustración 34: Test de medias para variable: Edad para deciles 1, 2, 9 y 10. Fuente: Elaboración propia en software IBM SPSS

13.10 Anexos 10: Resultados test de medias. Caracterización clientes deciles 1, 2, 9 y 10. Variable Frecuencia de compras

Prueba de muestras independientes

		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia	
									Inferior	Superior
Frecuencia	Se han asumido varianzas iguales	13,547	,000	-9,118	661	,000	-42,4047282	4,6504736	-51,5362092	-33,2732472
	No se han asumido varianzas iguales			-7,841	172,183	,000	-42,4047282	5,4081911	-53,0796175	-31,7298389

Ilustración 35: Test de medias para variable: Frecuencia de compras para deciles 1, 2, 9 y 10. Fuente: Elaboración propia en software IBM SPSS

13.11 Anexos 11: Resultados test de medias. Caracterización clientes deciles 1, 2, 9 y 10. Variable Tiempo de activación

Prueba de muestras independientes

		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia	
									Inferior	Superior
Activacion	Se han asumido varianzas iguales	1,122	,290	-,326	661	,745	-,4928886	1,5119972	-3,4617848	2,4760076
	No se han asumido varianzas iguales			-,307	187,813	,759	-,4928886	1,6058458	-3,6607012	2,6749239

Ilustración 36: Test de medias para variable: Tiempo de activación para deciles 1, 2, 9 y 10. Fuente: Elaboración propia en software IBM SPSS

13.12 Anexos 12: Tabla de contingencia Categoría productos.

Tabla de contingencia Categoría ^ Tratamiento

Recuento		Tratamiento		Total
		1-2	9-10	
Categoría	Otros	275	30	305
	Productos	611	46	657
Total		886	76	962

Ilustración 37: Tabla de contingencia: Categoría "Productos" para deciles 1, 2, 9 y 10. Fuente: Elaboración propia en software IBM SPSS

13.13 Anexos 13: Test de proporciones categoría Productos

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)	Sig. exacta (bilateral)	Sig. exacta (unilateral)
Chi-cuadrado de Pearson	2,300 ^a	1	,129		
Corrección por continuidad ^b	1,927	1	,165		
Razón de verosimilitudes	2,222	1	,136		
Estadístico exacto de Fisher				,157	,084
N de casos válidos	962				

a. 0 casillas (0,0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 24,10.

b. Calculado sólo para una tabla de 2x2.

Ilustración 38: Prueba Chi-Cuadrado de proporciones para dos muestras independientes. Categoría "Productos" para deciles 1, 2, 9 y 10. Fuente: Elaboración propia en software IBM SPSS

13.14 Anexos 14: Tabla de contingencia, Subcategoría Productos de gran consumo

Tabla de contingencia Subcategoría * Tratamiento

Recuento		Tratamiento		Total
		1-2	9-10	
Subcategoría	Otros	286	32	318
	Productos gran consumo	600	44	644
Total		886	76	962

Ilustración 39: Tabla de contingencia: Subcategoría "Productos gran consumo" para deciles 1, 2, 9 y 10. Fuente: Elaboración propia en software IBM SPSS

13.15 Anexos 15: Test de proporciones, Subcategoría Productos de gran consumo.

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)	Sig. exacta (bilateral)	Sig. exacta (unilateral)
Chi-cuadrado de Pearson	3,054 ^a	1	,081		
Corrección por continuidad ^b	2,626	1	,105		
Razón de verosimilitudes	2,948	1	,086		
Estadístico exacto de Fisher				,098	,054
N de casos válidos	962				

a. 0 casillas (0,0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 25,12.

b. Calculado sólo para una tabla de 2x2.

Ilustración 40: Prueba Chi-Cuadrado de proporciones para dos muestras independientes. Subcategoría "Productos gran consumo" para deciles 1, 2, 9 y 10. Fuente: Elaboración propia en software IBM SPSS