



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

**ESTUDIO DEL COMPORTAMIENTO DE CLIENTES EN UN AMBIENTE
MULTICANAL**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL INDUSTRIAL

MARÍA FERNANDA PALMA LÓPEZ

**PROFESOR GUÍA:
RICARDO MONTOYA MOREIRA**

**MIEMBROS DE LA COMISIÓN:
MARCEL GOIC FIGUEROA
ALEJANDRA PUENTE CHANDÍA**

**SANTIAGO DE CHILE
2015**

RESUMEN DE LA MEMORIA PARA OPTAR AL
TÍTULO DE: Ingeniero Civil Industrial
POR: María Fernanda Palma López
FECHA: 20/10/2015
PROFESOR GUÍA: Ricardo Montoya

ESTUDIO DEL COMPORTAMIENTO DE CLIENTES EN UN AMBIENTE MULTICANAL

Actualmente los clientes están expuestos a diferentes canales de interacción con la empresa, ya sea como canal de comunicación como canal de venta. Este fenómeno se presenta en el mercado del retail y gestionarlo de forma adecuada puede significar una ventaja competitiva importante para la empresa.

El objetivo de este trabajo es caracterizar el comportamiento de los clientes de una tienda por departamento en un ambiente multicanal. Para esto se definirán y calcularán variables que caractericen este comportamiento multicanal, se caracterizará el comportamiento de compra de estos clientes, se realizará una segmentación para capturar la heterogeneidad y, por último, se analizará la dinámica entre los distintos canales.

Para alcanzar estos objetivos se utilizó información transaccional, de navegación y demográficos de los clientes. Se trabajó con los datos de venta en tienda y a distancia, considerando en este último grupo las ventas en el sitio web, sitio móvil y venta telefónica). Se realizó un análisis descriptivo de las variables demográficas y RFM multicanal de los clientes, obteniéndose como resultado que los clientes con actividad multicanal tienen mayores frecuencias y montos de compra, lo que es consistente con lo que dice la literatura. Se considera como actividad multicanal la compra a distancia y la navegación en el sitio web o móvil.

Luego se realizó una segmentación de los clientes, teniéndose como resultado 3 conglomerados principales: los clientes activos monocanal, activos multicanal e inactivos. Las variables que se utilizaron fueron las 6 variables RFM (cada una de ellas en tienda y a distancia). El grupo de clientes multicanal resultó ser el más pequeño y el con mayores montos y frecuencias de compra. Se estudió la dinámica de comportamiento a través de probabilidades de transición entre estados, teniéndose como resultado que los segmentos más estables son los con menos actividad y actividad monocanal, mientras que el estado multicanal es relativamente inestable, ya que tiene más probabilidades de pasar a un estado monocanal que de permanecer en él.

En el estudio de la interacción entre canales se puede apreciar presencia del efecto ROPO (Research Online Purchase Offline), ya que la cantidad de clientes que compra en tienda el mismo día en que navega es mayor a la cantidad de clientes que compra al día siguiente y este número es decreciente durante la primera semana luego de navegar. Por otro lado, se puede distinguir que los clientes que navegan en el sitio móvil tienen más probabilidades de comprar (en el sitio web y móvil) que los clientes que navegan web.

Se propone estudiar la causalidad entre multicanalidad y aumento en las frecuencias y montos de compra mediante experimentación.

Agradecimientos

No sé si alguien vaya a leer esta parte de la memoria, pero llevo un poco más de un año pensando en qué voy a poner y soñando en que este momento algún día iba a llegar, así que los escribiré de todos modos.

En primer lugar quiero agradecer a mi familia de origen por todo el apoyo durante esta larga etapa, principalmente en esta última parte que se extendió más de lo esperado. Gracias por cuidar al Ignacio y hacerlo sentir en un hogar mientras sus papás trabajan.

Gracias a las Mujeres Bakanes e Hijos por la compañía durante mi primera tesis. Fueron un pilar fundamental para sobrevivir al puerperio que ya está llegando a su fin.

Quiero agradecer al profesor Ricardo Montoya por darme la oportunidad de realizar esta memoria y poder cerrar este ciclo. Gracias también por las palabras de aliento cuando pensaba que iba a morir en el intento.

A Renzo y Andrea por prestarme todo el apoyo y ayuda durante el desarrollo de la memoria. También a Alejandra Puente al profesor Marcel Goic por ayudarme a sacar esto adelante.

A la gente con la que crucé camino durante este paso por la universidad, en especial a la Angie que apareció al final como compañera de travesuras y que sigue presente como una gran amiga.

A mis amigas de siempre, las 45 ever, por estar siempre presentes.

Y finalmente quiero darle las gracias a mi novio Andrés Araneda por acompañarme de forma incondicional en esta última etapa, por ser mi compañero en esta aventura llamada vida y por darme la alegría de convertirme en mamá.

Esto está recién comenzando...

Tabla de Contenido

1. Introducción	1
1.1. Estudios anteriores	1
1.2. Justificación	2
1.3. Ejemplos Internacionales	2
2. Objetivos	4
2.1. Objetivo General	4
2.2. Objetivos Específicos	4
3. Metodología	5
4. Marco Conceptual	7
4.1. Multichannel Customer Management	7
4.2. Modelo RFM	8
4.2.1. Segmentación	9
4.2.2. Matriz de Transición	9
4.3. Test de Hipótesis	10
4.3.1. Test de diferencia de medias	10
4.3.2. Test de diferencia de distribución	10
5. Alcances	12
6. Selección y Procesamiento de datos	13
6.1. Información Disponible	13
6.2. Procesamiento de datos	15
7. Cálculo de Variables	16
8. Análisis Descriptivo	18
8.1. Categorías	18
8.2. Canal de Venta	21
8.3. Variables RFM Multicanal	22
8.4. Conclusiones Análisis Descriptivo	27
9. Segmentación	29
9.1. Segmentación RFM	29
9.2. Segmentación con Visitas y sin Recency	39
9.3. Segmentación Multicanal	45
10. Transición entre Canales	51

11.	Conclusiones y Recomendaciones.....	61
12.	Bibliografía.....	64

1. Introducción

La industria del retail en los últimos años ha migrado de operar en un único canal a disponer de múltiples canales de contacto con los clientes. Si bien este fenómeno existe desde la aparición de los primeros catálogos en Estados Unidos, es en los últimos años, con la aparición de las tiendas online y las redes sociales que se le está dando más importancia a la multicanalidad.

Se entiende por multicanalidad la disposición de más de un canal mediante el que el consumidor realiza su experiencia de compra. Entre los canales generalmente disponibles se encuentran la tienda física, la tienda online, la aplicación móvil y venta telefónica. A esto debe sumarse además, los medios de comunicación masivos que juegan un papel importante en la compra, para conocer e informarse sobre los productos. Sin embargo es difícil recolectar información sobre el impacto que esto tiene en el comportamiento de los clientes.

El consumidor multicanal es aquel que utiliza más de un medio disponible para llevar a cabo el proceso de compra, es decir, interactúa con la empresa a través de distintos canales para cada etapa de este proceso. De esta forma se presenta la oportunidad para la empresa de estar presente en cada uno de los momentos en el que el cliente tome una decisión, ofreciendo alternativas diferentes que se adecúen a cada medio. Según **Neslin et al. (2006)**, este aumento de los canales de interacción con el cliente ha creado el desafío para las firmas de gestionar este ambiente de forma efectiva y le ha presentado a los académicos la oportunidad de generar ideas que ayuden a encaminar este desafío.

1.1. Estudios anteriores

El comportamiento de un cliente está influenciado por un complejo proceso psicológico (consciente e inconsciente) (**Uptal M. et al., 2010**). En este proceso participan componentes propias del individuo, como las emociones, experiencias pasadas, entre otros. Todos estos elementos no son medibles con los datos con los que cuenta la empresa con la que se trabajará, por lo que no serán considerados en el análisis y perfilamiento de los clientes. En **Balasubramanian (2005)**, se describen algunos objetivos que mueven la decisión del canal de compra y muestra cómo estos influyen en la utilidad que el cliente tendrá al finalizar la compra. Estos objetivos son metas económicas, búsqueda de autoafirmación, búsqueda de un significado simbólico asociado con el producto y el proceso de compra, búsqueda por interacción social y su dependencia a esquemas de compra. Estos autores utilizan entrevistas para identificar estas motivaciones. Si bien esto es valioso desde el punto de vista de que se obtiene información que pudo no haberse considerado a priori, no es aplicable a este trabajo ya que sólo se utilizarán datos históricos de transaccionales, demográficos y de navegación de los clientes.

Con un conocimiento acabado del comportamiento de los clientes y de sus necesidades en cada momento se podrían canalizar los recursos de mejor forma, haciendo campañas de marketing más efectivas, que cumplan con el objetivo de informar al cliente lo que quiere saber en el momento y no desperdiciarlo con campañas poco efectivas que puedan

causarle una mala experiencia al cliente (además de desperdiciar recursos). Para esto es necesario comprender su comportamiento, lo que permite hacer una segmentación adecuada. Sin embargo uno de los principales desafíos del estudio del comportamiento de los clientes en un ambiente multicanal es conocer las reales motivaciones que tiene un cliente para utilizar un canal y no otro. Por otro lado, no se puede tener información de algunas actividades que realizan los clientes de forma anónima, como visitar la tienda sin comprar nada, navegar por el sitio web sin iniciar sesión o revisar los productos de la competencia, por ejemplo.

1.2. Justificación

El retail es una industria altamente competitiva, por lo que es importante implementar estrategias de diferenciación. Por otro lado, el fenómeno de multicanalidad se presenta fuertemente en este tipo de empresas, ya que la mayoría de ellas tiene implementado el sistema de tienda online, el envío de correos electrónicos a clientes con ofertas y promociones. Dadas estas condiciones de competitividad y de multicanalidad se presenta la oportunidad de utilizar los distintos canales para contactarse con los clientes de forma óptima.

Al ser este un fenómeno relativamente nuevo en Chile, es poca la investigación que se ha hecho que involucren el efecto de los distintos canales en los clientes. Hacer un estudio de estas características es una oportunidad para conocer aspectos de los clientes que hasta el momento no han sido considerados, lo que trae beneficios tanto para los clientes como para la empresa. Por un lado, los clientes recibirán información de acuerdo a sus necesidades, lo que se traduce en una mayor satisfacción, y por otro lado tenemos el beneficio económico que significa para la empresa ahorrar recursos en marketing.

1.3. Ejemplos Internacionales

En **Adobe (2011)** se presentan tres casos de estudio en los que la gestión multicanal tuvo un gran impacto.

El primer caso es USAA, una empresa que presta servicios financieros a familias de militares. Se dieron cuenta que los clientes cambiaban de canal durante una transacción. Más específicamente, lo hacían del canal web al teléfono. Tomaron la decisión de ayudar a los clientes haciendo más fácil y fluido el proceso en el sitio web. De esta manera aumentaron la satisfacción de los clientes y disminuyeron la cantidad de llamados innecesarios. Utilizaron, además, datos de satisfacción de clientes, atributos demográficos, historial de compra y actividad web y telefónica para segmentar de mejor forma para las campañas de publicidad. Los retornos sobre la inversión de este estudio fueron 12 veces mayores al realizado para segmentar a los clientes de la forma tradicional.

El segundo caso es de la empresa Dollar Thrifty. Esta empresa surge de la fusión de Dollar Rent con Thrifty Car. Los encargados de marketing estaban teniendo problemas para entender el comportamiento de los clientes ya que debido a la cantidad de canales disponibles (reservación vía telefónica gratuita, página web, sucursal, entre otras), era muy difícil obtener datos para analizar. Desarrollaron un sistema que permitía obtener perfiles de clientes basados en el historial de venta, origen de venta, comportamiento de

navegación, transacciones en tienda y otras actividades. Esto les permitió enviar promociones personalizadas de acuerdo al perfil de cada cliente, aumentando así la tasa de conversión en un 152% y optimizando los gastos en publicidad (ahorrando US\$1,7 millones en un año)

El último caso de estudio es Travelocity, empresa líder en servicios de viajes. Analizaron el comportamiento de los clientes haciendo un seguimiento de los productos adquiridos, el canal a través del cual fue hecha la adquisición, atributos de búsqueda, entre otros, de manera de ser más agradable la experiencia de compra. Utilizando una plataforma de análisis multicanal pudieron aumentar en un 15% la tasa de conversión y recuperaron la inversión en tecnología en sólo 9 meses.

De esta manera se puede ver que hace un adecuado análisis del comportamiento de los clientes en forma integral puede significar una reducción importante en los costos de marketing y un aumento en la efectividad de las promociones.

2. Objetivos

2.1. Objetivo General

Caracterizar el comportamiento de clientes de una tienda por departamento en un ambiente multicanal.

2.2. Objetivos Específicos

- Identificar y calcular variables que expliquen el comportamiento multicanal de los clientes .
- Caracterizar el comportamiento de compra de los clientes multicanal y monocal.
- Capturar heterogeneidad en el comportamiento de los clientes en un ambiente multicanal.
- Establecer dinámicas entre canales para clientes multicanal.

3. Metodología

El trabajo estará dividido en 4 etapas, que corresponden al desarrollo de cada uno de los objetivos específicos.

- 1) Identificar y calcular variables que expliquen el comportamiento multicanal de los clientes:

Para identificar las variables que expliquen el comportamiento multicanal de los clientes se hará revisión de la literatura y se conversará con profesores y expertos de la empresa para recoger sus propuestas

El procesamiento de los datos se hará con el *software* SPSS. Para hacer los cálculos se utilizará esta misma herramienta junto con *Microsoft Excel* cuando se necesite hacer cálculos más específicos y haya disminuido la cantidad de datos mediante el muestreo.

- 2) Caracterizar comportamiento de compra

En esta etapa se hará, en primer lugar, un análisis descriptivo del comportamiento de los clientes. Se analizarán variables tanto demográficas como del comportamiento de los clientes para encontrar relaciones entre ellas y que permitan entender cómo se comportan los clientes. En especial se harán comparaciones de las variables RFM multicanal y monocal, junto con otras variables relacionadas con la actividad multicanal (Visitas de navegación, tipos de dispositivos utilizados, etc). Se utilizarán teste de igualdad de medias y de diferencia de proporciones para estudiar el impacto que tiene las características de los clientes en su comportamiento de compra (Frecuencias, Montos, Categorías, etc). Los cálculos se harán con SPSS.

- 3) Capturar heterogeneidad

Para capturar la heterogeneidad se utilizarán analizarán los clientes monocal junto a los multicanal y luego sólo a los multicanal. Para detectar el número óptimo de clusters se comparará el promedio de las distancias intra grupo a medida que se aumenta el número de clusters. Para esto se utilizará la herramienta RStudio, mientras que para el análisis de *K-Means* se utilizará SPSS.

- 4) Detectar dinámicas entre canales

Para detectar la dinámica entre canales se realizarán 2 análisis que permitan describir esta interacción.

En primer lugar se analizará cómo fluyen los clientes de un segmento a otro durante 4 períodos de tiempo. Se construirá una matriz de probabilidades de transición entre estados (considerando cada segmento como un estado diferente) con las frecuencias históricas. Aquí se realizarán dos actividades que permitan detectar cómo interactúan los clientes entre un canal y otro. Estos cálculos se harán utilizando *Microsoft Excel* y *SPSS*.

La segunda etapa es estudiar la el uso de los distintos canales por un mismo cliente. Se analizará cuántos clientes compran a través de cada canal luego de navegar en cada uno de los dispositivos disponibles. Este análisis se extenderá a las categorías para identificar si este comportamiento difiere entre una u otra. Se realizarán test de igualdad de medias

y de diferencia de distribuciones para averiguar si las diferencias que se presenten son significativas o no.

4. Marco Conceptual

4.1. Multichannel Customer Management

Según **Neslin et al. (2006)**, Multichannel Customer Management (Gestión de Clientes Multicanal), MCM, es el diseño, desarrollo, coordinación y evaluación de canales a través de los cuales los clientes y las empresas interactúan, con el objetivo de mejorar el valor del cliente a través de la adquisición efectiva de clientes, retención y desarrollo. Los autores plantean que los clientes deciden utilizar distintos canales para las distintas etapas del proceso de compra (búsqueda de información, compra, servicio de post-venta), por lo que es importante estudiar la forma en que interactúa entre uno y otro.

En un enfoque de gestión de clientes en un ambiente multicanal es necesario considerar cinco actividades principales: Integración de los datos, entendimiento del comportamiento de los clientes, evaluación de los canales, asignación de recursos a través de los canales, coordinación de estrategias de canales. El trabajo de memoria está enfocado en los dos primeros pasos de la gestión de clientes multicanal.

Ya se han hecho estudios que intentan explicar el comportamiento de los clientes en los distintos canales y las variables que influyen en la decisión de escoger uno u otro. Los principales son:

- a) Acciones de marketing: ciertos estímulos enviados a los clientes pueden promover el uso de un canal u otro. Por ejemplo, una oferta enviada por email favorece la compra a través del sitio web si existe un link directo.
- b) Atributos del canal: Los clientes seleccionan un canal de acuerdo a sus necesidades. Por ejemplo, un cliente se divierte revisando un catálogo, entonces preferirá usarlo (lo que no implica que se efectúe la compra).
- c) Integración de los canales: Una buena integración de los canales, puede promover el uso de un canal en específico luego de utilizar otro. Un ejemplo de esto es la opción de retiro en tienda, ya que promueve que los usuarios de la tienda online vayan, además, a la tienda física.
- d) Influencia social: Algunos autores plantean que existe una influencia sobre ciertos clientes que dice que un cierto tipo de personas debe usar un canal determinado.
- e) Factores situacionales: Existen factores que pueden determinar la elección de un canal. Estos son el entorno físico, entorno social, asuntos temporales (hora del día, urgencia de compra, etc.), definición de tareas y estado de ánimo antecedente.
- f) Diferencias individuales: Datos demográficos influyen en la decisión de elección de canal (género, edad, nacionalidad, etc.), además de su estado en el ciclo de vida de cliente.

Sólo algunos de estos elementos son controlables por la empresa para promover el uso de un canal u otro. Sin embargo, un mayor conocimiento de las variables que influyen en la decisión de escoger cierto canal permite que éstas puedan conocer mejor las

necesidades de los clientes y puedan ofrecerles promociones, ofertas, o mensajes adecuados para cada tipo de cliente (que es el objetivo principal del proyecto).

4.2. Modelo RFM

Es común que las empresas quieran calcular el valor de sus clientes. Para esto buscan qué características permiten determinar cuánto vale para la empresa mantener a dicho cliente. Uno de los métodos más conocidos es el modelo RFM (Cheng et al., 2009). Su ventaja es extraer características de un cliente utilizando pocos criterios, disminuyendo así la complejidad de un modelo de valor del cliente.

El modelo analítico RFM es propuesto en (Hughes, 1994) y diferencia a clientes importantes desde una gran base de clientes basándose en 3 atributos:

- **R** (Recency) Recencia de la última compra: Se refiere al intervalo de tiempo entre la última acción de consumo del cliente y el presente. Mientras más corto el intervalo, mayor es el valor de R.
- **F** (Frequency) Frecuencia de las compras: Se refiere al número de transacciones en un periodo particular. Mientras mayor es la frecuencia, más grande es F.
- **M** (Monetary Value) Valor monetario de las compras: Se refiere al monto de dinero consumido en un periodo particular. Mientras mayor sea este monto, mayor será M.

(Jing Wu, 2005) muestra que mientras mayores son los valores de R y F, es más probable que los clientes produzcan una nueva transacción con la empresa. Además, mientras más grande sea M, es más probable que los clientes compren productos o servicios de la empresa nuevamente en el futuro.

Para el estudio de los clientes multicanal se hará una modificación a esta metodología para incluir variables relacionadas con esta condición. Se espera buscar distintos tipos o niveles de clientes multicanal y para esto es necesario medir las intensidades de las variables que lo caracterizan. Las variables a calcular serán las siguientes:

- Recency Online
- Recency Offline
- Frequency Online
- Frequency Offline
- Monetary Online
- Monetary Offline

Las variables Offline incluyen la información de las compras hechas en tienda, mientras que en Online se encuentran las compras hechas a distancia, es decir, por vía telefónica, a través del sitio web o a través del aparato móvil.

Se considerarán sólo datos de compra, ya que la información de navegación es de baja calidad (el 80% de la navegación se hace en forma anónima). Es por esto que se consideró una variable de navegación que tiene valor 1 cuando los clientes han navegado durante el periodo estudiado y 0 cuando no han tenido actividad *online*. Desde noviembre de 2014 se cuenta con un nuevo sistema que permite identificar a los clientes de forma única (cada cliente está asociado a un ID único) y que posibilita recoger la información de navegación previa a una compra, incluso cuando el usuario inicia sesión al momento de pagar. Se propone para trabajos futuros incluir la información de navegación para aumentar las variables que los caracterizan.

Para convertir esta variable de métrica a categórica, se considerarán 2 niveles para cada variable: Alto y Bajo. El valor de corte se calculará ordenando los valores de forma ascendente y se considerará que la mitad es Alto y la otra mitad Bajo, para cada una de las variables.

4.2.1. Segmentación

Con estas variables calculadas para cada uno de los clientes multicanal, se hará una segmentación para encontrar clusters de clientes. Luego se agregarán variables de navegación para hacer un análisis más detallado sólo de los clientes multicanal.

4.2.2. Matriz de Transición

El cálculo de la probabilidad de transición de un estado a otro se hará con la frecuencia empírica durante el período estudiado.

Se considerará como estado cada uno de los segmentos encontrados en la segmentación. Para realizar esto se definirán los estados en los que se encuentran los clientes en 4 periodos de tiempo. Se utilizarán los datos transaccionales de 2 años, por lo tanto, estos sub períodos serán de 6 meses. La cantidad de segmentos depende de lo que arroje el análisis previo. Cada celda de la matriz contendrá la probabilidad de pasar de un estado determinado a otro.

La matriz quedaría de la siguiente forma:

	E_1^t	E_2^t	E_3^t
E_1^{t-1}	$P(E_1^t E_1^{t-1})$	$P(E_2^t E_1^{t-1})$...
E_2^{t-1}	$P(E_1^t E_2^{t-1})$	$P(E_2^t E_2^{t-1})$...
E_3^{t-1}	$P(E_1^t E_3^{t-1})$	$P(E_2^t E_3^{t-1})$...

Tabla 4.1: Matriz de transición entre estados.

4.3. Test de Hipótesis

4.3.1. Test de diferencia de medias

Este test se usa para determinar si las medias de 2 poblaciones son significativamente diferentes basándose en la información obtenida de muestras de cada una de ellas.

Se tiene X_1 y X_2 , las medias de 2 muestras aleatorias independientes, grandes de tamaño n_1 y n_2 .

La distribución muestral del estadístico $X_1 - X_2$ se aproxima a una normal que tiene como media $\mu_1 - \mu_2$ y como desviación estándar $\alpha(X_1 - X_2)$.

Entonces:

$$\alpha(X_1 - X_2) = \sqrt{\left(\frac{\alpha_1^2}{n_1}\right) + \left(\frac{\alpha_2^2}{n_2}\right)}$$

Para muestras de 30 o más elementos se pueden considerar las desviaciones muestrales S_1 y S_2 como estimadores de α_1 y α_2 .

Se tiene que:

$$H_0: \mu_1 = \mu_2$$

Las alternativas pueden ser:

$$H_1: \mu_1 \neq \mu_2$$

$$H_1: \mu_1 > \mu_2$$

$$H_1: \mu_1 < \mu_2$$

Finalmente se prueba la H_0 con el estadístico:

$$z = \frac{X_1 - X_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Se rechaza H_0 si $z > z_{tabla}$, donde z_{tabla} dependerá del nivel de confianza.

4.3.2. Test de diferencia de distribución

Se utiliza el estadístico chi-cuadrado para someter a test de hipótesis referidas a distribuciones de frecuencias. Este test contrasta frecuencias observadas con las frecuencias esperadas de acuerdo con la hipótesis nula.

En este modelo, la hipótesis nula plantea que las variables analizadas son independientes.

Luego se calcula el estadístico de la siguiente manera:

$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i}$$

Donde:

- o_i representa la frecuencia observada
- e_i representa la frecuencia esperada

Luego, no se rechaza la hipótesis nula si el valor del estadístico es menor al valor encontrado en la tabla. Este valor depende del nivel de confianza y de los grados de libertad.

5. Alcances

El trabajo se realizará bajo el marco del proyecto de investigación titulado “Desarrollo de Tecnologías para Mejorar la Experiencia de Compra de Clientes Multicanal” junto a la colaboración de la empresa Falabella. Esta empresa facilitará sus datos históricos de clientes, transacciones y navegación con el fin de llevar a cabo la investigación. La empresa cuenta con 5 canales de venta diferentes. Estos son:

- Tienda física
- Página Web
- Página sitio móvil
- Venta telefónica
- Aplicación móvil

La aplicación está funcionando hace pocos meses, por lo que no se cuenta con suficientes datos para hacer las comparaciones con otros canales. De esta forma, se trabajará con datos de los 4 primeros canales, agrupando las ventas en la página web, en el sitio móvil y las ventas telefónicas como ventas a distancia.

De este modo, queda fuera del análisis otros comportamientos de los clientes que pueden ser interesantes para el estudio, como la actividad de los clientes dentro de la tienda, la comparación que hace de los productos con otras páginas o tiendas, etc.

6. Selección y Procesamiento de datos

6.1. Información Disponible

Las bases de datos que se utilizaron para el cálculo de las variables fueron:

- a) Información de Clientes: Contiene los datos de los clientes que han realizado compras en los últimos 2 años.
 - ID.Cliente
 - Recibe.Email¹
 - Edad
 - Sexo
 - Región
 - Comuna
- b) Navegación: Período de 6 meses (1 de noviembre de 2014 – 30 de abril de 2015). Esta base de datos contiene el historial de navegación de cada cliente por día y por dispositivo.
 - Dia
 - ID.Cliente
 - *Dispositivo*: Los dispositivos son Web (visitas del sitio web a través del computador de escritorio), Mobile (visitas al sitio web a través de un *smartphone*) y Otros (visitas a través de Tablet o *Game Console*)
 - *Visitas*: Contiene el número de páginas vistas por el cliente ese día a través de un dispositivo determinado. El sistema de identificación de clientes que navegan tiene 6 meses de antigüedad. La información de los períodos anteriores es de mala calidad, por lo que se contará sólo con los datos de ese semestre.
 - *Monto*: Si ese día el cliente realiza una compra a través de un dispositivo determinado la variable *Monto* tiene el valor de esa compra. Si las visitas de ese día no se convierten en compra, el valor es 0.
- c) Datos Transaccionales: Período de 2 años (2 de mayo de 2013 – 30 de abril de 2015). Esta tabla contiene la información transaccional de todos los clientes a través de los distintos canales de compra por día y por categoría. Esto quiere decir que si, por ejemplo, un cliente compra 5 productos de una misma categoría en un local determinado, aparecerá como sólo una transacción por el

¹ Los clientes al inscribirse en el sitio web deben decir si desean o no recibir emails de la empresa. Los clientes que desean recibirlos quedan son el grupo de clientes que recibe promociones y otros correos.

total del monto. Por otro lado, si, por ejemplo, compra 1 producto en 4 categorías diferentes en un mismo local, se registrarán 4 transacciones.

- Dia
- ID.Cliente
- *ID.Categoria*: Contiene el código de la categoría en la que se hizo la compra.
- *Nombre.Categoria*: Contiene el nombre de la categoría donde se hizo la compra. Estas categorías son:
 - Accesorio mujer
 - Blanco
 - Calzado
 - Damas
 - Decoración
 - Deportes
 - Dormitorio
 - Electro hogar
 - Gourmet y concesiones
 - Hombres
 - Maneje
 - Mueble
 - Niños
 - Perfumería
 - Regalos
 - Rincón juvenil damas
 - Rincón juvenil hombres
 - Ropa interior
 - Administración
 - Puntos
 - Servicios
- *Monto*: Es el monto total de las compras hechas por un clientes en una misma categoría en un mismo local.

- *ID.Local*: Es el código del local donde se realizó la transacción. Cada tienda tiene un código diferente. Las compras que son hechas a distancia tienen el código 2000 e incluyen las compras a través del sitio web (por mediante cualquier dispositivo) y vía telefónica (Fonocompra). Las compras de los productos que son publicitados a través del canal de televisión y son vendidas a través del teléfono aparecen en esta categoría.

6.2. Procesamiento de datos

Limpieza de datos: En primer lugar, para trabajar la base de datos transaccional, se eliminaron todas las transacciones en las que no tenía *ID.Cliente* o que lo tenían erróneo. Luego, se eliminaron las transacciones de las categorías *Administración*, *Puntos* y *Servicios*, ya que no corresponden a las categorías de productos de la empresa. La base de datos de la información de clientes contenía clientes sin información demográfica, por lo que no se consideran para este trabajo. Luego se cruzó la base de datos transaccional con la de clientes para tener junto a los datos de transacción la información del cliente. Esto se realizó para la base de datos transaccional de cada mes. Para tener sólo transacciones con buena información de clientes y sólo clientes que hayan hecho transacciones en alguna de las 18 categorías de productos. Existen transacciones con montos negativos. Éstas corresponden a las devoluciones.

Selección de clientes: Para obtener la lista definitiva de clientes, se seleccionaron los clientes que hayan realizado transacciones en cada uno de los meses, unificándolos en una sola tabla. Esta tabla final contenía a 2.472.375 clientes.

Selección de muestra: Dado el volumen de clientes de la empresa, se realizó un muestreo aleatorio seleccionando como muestra el 4% del total de clientes. Esto dio como resultado una muestra de 98.895 clientes.

Selección transacciones: El último paso fue seleccionar las transacciones de los clientes de la muestra y unificarlas en una sola tabla. El total de transacciones seleccionadas fue de 1.462.669.

7. Cálculo de Variables

Con la información disponible se construyó un panel con indicadores para cada cliente. Las variables son las siguientes:

a) Variables Navegación:

- a. **Visitas.Total:** Con la información de navegación se obtuvo la cantidad total de páginas visitadas por el cliente durante los 6 meses.
- b. **Visitas.Web:** Total de páginas vistas por el cliente desde un computador de escritorio.
- c. **Visitas.Mobile:** Total de páginas vistas por el cliente desde un *Smartphone*.
- d. **R.Visitas.Total:** *Recency* del total de visitas. Diferencia de días entre la última visita realizada y el último día del período.
- e. **R.Visitas.Web:** *Recency* de las visitas web.
- f. **R.Visitas.Mobile:** *Recency* de las visitas desde dispositivo móvil.

b) Variables Multicanalidad:

- a. **Compra.Distancia:** Tiene el valor 1 si el cliente ha comprado al menos una vez a distancia durante los 2 años.
- b. **Visita:** Esta es una variable que tiene valor 1 si el cliente visitó alguna página durante los 6 meses y 0 si no.
- c. **Multicanal:** Esta variable tiene valor 1 si el cliente navega, 2 si el cliente compra a distancia, 3 si el cliente navega y compra a distancia y 0 si sólo compra en tienda.

c) Variables Compra:

- a. **M.Tienda:** Suma de todas las compras hechas en el período en una tienda física. En esta variable las transacciones de devolución aportan de forma negativa, ya que para efectos del estudio importa el valor total de los montos gastados por cada cliente².
- b. **F.Tienda:** Total de transacciones hechas en el período en una tienda física. Al obtenerse esta información de la tabla de transacciones ya descrita, cada transacción representa el total de transacciones hechas en una misma categoría en un día determinado. Las devoluciones aportan de forma positiva a esta variable, ya que el momento de la devolución es otra oportunidad de interacción entre la empresa y el cliente³.

² Esto es válido también para el cálculo de montos en los otros canales de venta.

³ Esto es válido también para el cálculo de frecuencias de compra en otros canales de venta.

- c. **R.Tienda:** Para calcular la variable *Recency*, se seleccionó para cada cliente el máximo del conjunto de fechas en las que realizó una transacción. Luego se calculó la diferencia de días entre la última compra y el final del período (30 de abril de 2015). Para los clientes que no realizaron ninguna transacción en tienda durante el período de estudio, se reemplazó este valor por 731, es decir, un día más que el total de días que tienen los 2 años.
- d. **M.Distance:** Suma de los montos de todas las transacciones hechas a distancia durante el período⁴.
- e. **F.Distance:** Total de transacciones hechas a distancia durante el período.
- f. **R.Distance:** Diferencia de días entre el final del período y la última compra realizada a distancia.
- g. **M.Total:** Suma de los montos de todas las transacciones hechas en el período.
- h. **F.Total:** Total de transacciones hechas en el período.
- i. **R.Total:** Diferencia de días entre el final del período y la última compra realizada.

⁴ Recordar que ventas a distancia incluye ventas hechas por el canal remoto, es decir, teléfono, computador o celular.

8. Análisis Descriptivo

Se hace una primera observación de las variables calculadas. Este análisis se realiza con los 98895 clientes de la muestra y sus 1.462.669 transacciones.

8.1. Categorías

En primer lugar se hará un análisis de las categorías.

El siguiente gráfico muestra la cantidad de productos comprados en cada categoría dependiendo si la venta fue hecha en tienda o a distancia⁵.

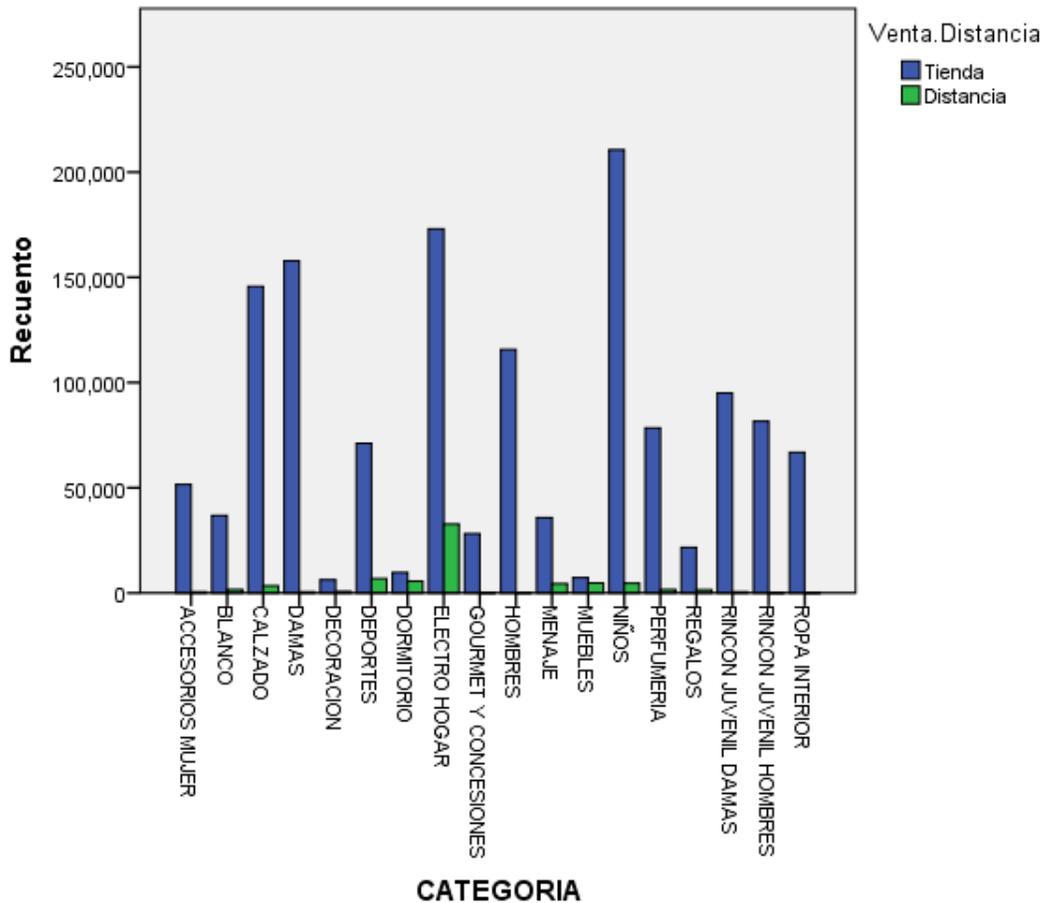


Gráfico 8.1: Cantidad de productos comprados en tienda o distancia.

Se puede ver que la categoría *Electro Hogar* es la que más se adquiere por venta a distancia, mientras que la categoría *Niños* es la más solicitada en tienda. Se hizo un test chi-cuadrado de diferencia de proporciones y arrojó que sí existen diferencias significativas en las distribuciones dependiendo del canal de venta.

⁵ Recordar que las ventas a distancia incluyen las ventas hechas a través del sitio móvil, sitio web o vía telefónica.

El siguiente gráfico muestra la proporción de cada categoría que es vendida en cada canal.

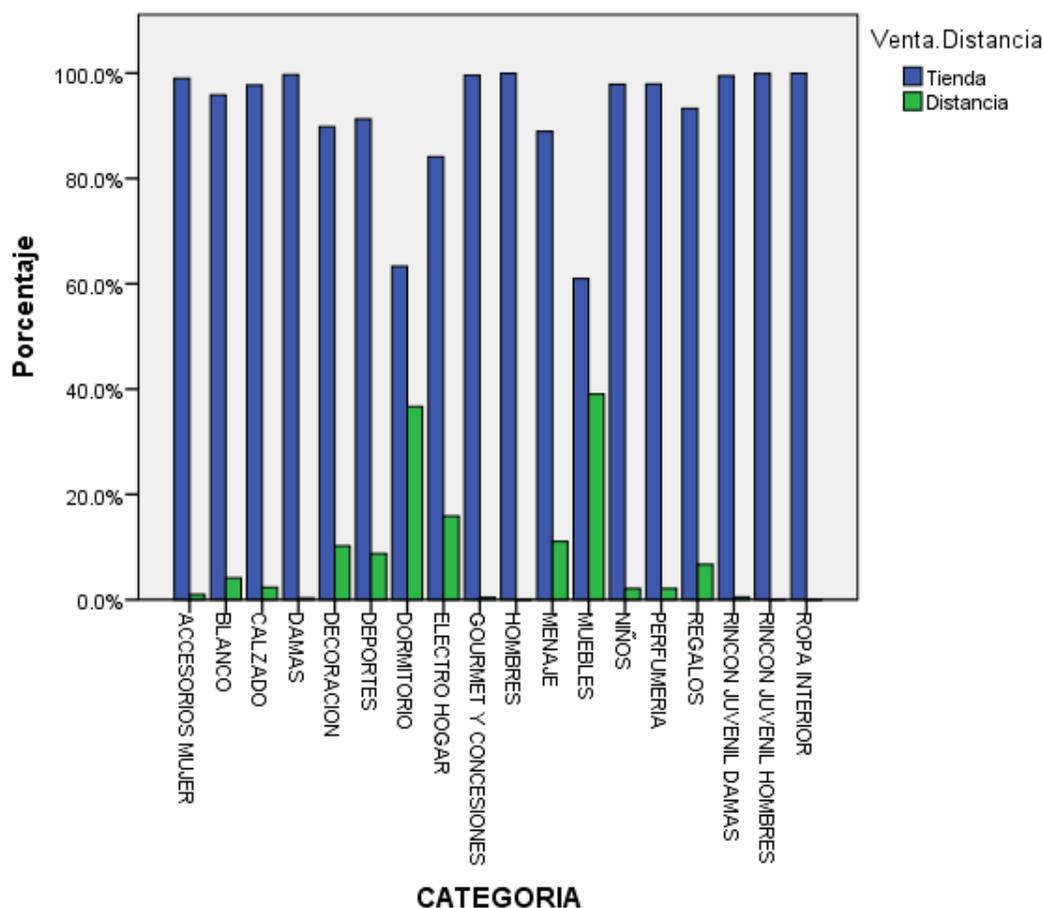


Gráfico 8.2: Porcentaje vendido en tienda o distancia.

Se puede ver que las categorías *Dormitorio* y *Muebles* se venden un gran porcentaje por venta a distancia. Esto puede deberse a que, al tratarse de productos de gran volumen, de todos modos deben pagar el costo de despacho. Así, al hacer la compra desde el computador, teléfono o celular, ahorran tiempo y costo de ir a la tienda. Las categorías que presentan menores ventas a distancia son *Damas*, *Gourmet y concesiones*, *Hombres*, *Rincón Juvenil Damas*, *Rincón Juvenil Hombres* y *Ropa Interior* (este último no se vende a distancia).

Al revisar la distribución entre hombres y mujeres que compran en cada categoría, se obtiene el siguiente gráfico:

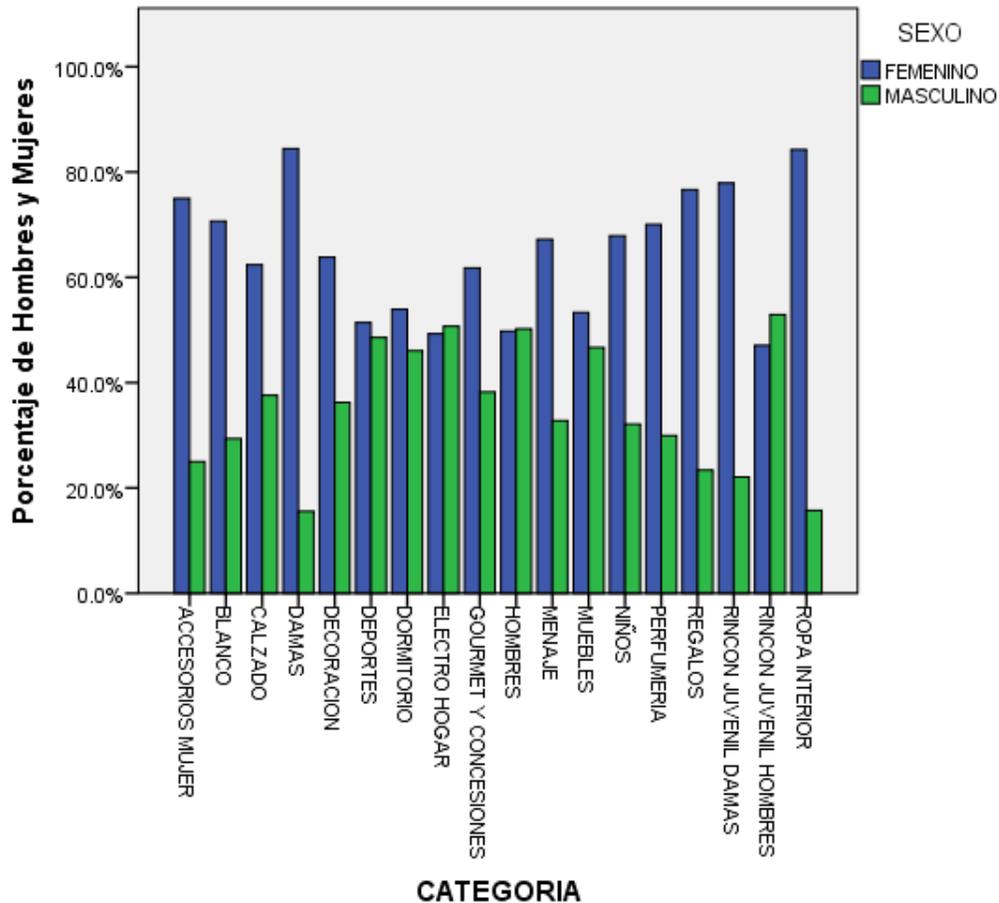


Gráfico 8.3: Porcentaje de compras de hombres y mujeres por categoría.

La mayoría de las categorías son compradas principalmente por mujeres. Sólo *Electro Hogar*, *Hombres* y *Rincón Juvenil Hombres* son compradas mayormente por mujeres. En *Deportes*, *Muebles* y *Dormitorio* las mujeres ocupan menos de un 60% de los clientes.

En el caso de las ventas a distancia, el siguiente gráfico muestra la distribución en porcentajes entre hombres y mujeres para cada categoría.

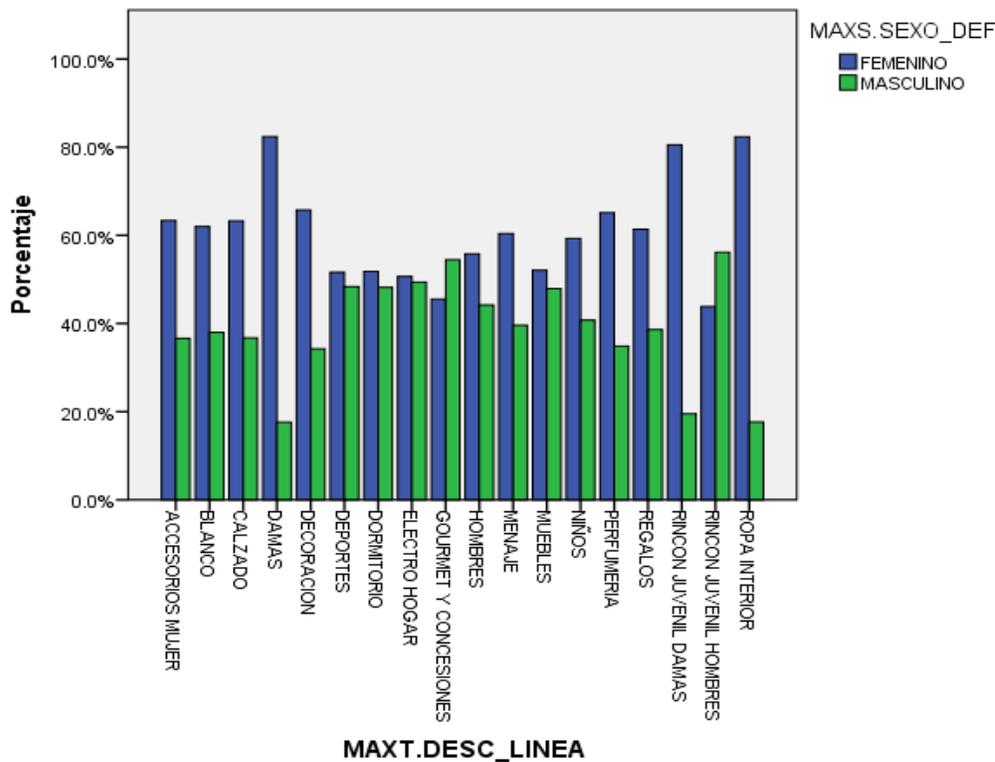


Gráfico 8.4: Porcentaje de compras de hombres y mujeres por categoría a distancia.

Se puede ver que en todas las categorías, excepto *Gourmet y Concesiones* y *Rincón Juvenil Hombres*, las mujeres tienen mayor participación de compra.

8.2. Canal de Venta

La siguiente tabla muestra la diferencia de proporciones de venta en distancia y tienda:

		SEXO	
		FEMENINO	MASCULINO
		% de la fila	% de la fila
Venta.Distance	Tienda	65.1%	34.9%
	Distancia	54.3%	45.7%

Tabla 8.1: Proporciones de venta por sexo a distancia y en tienda.

Se puede ver que las mujeres compran más en ambos canales, sin embargo la brecha disminuye en las ventas a distancia. Se realizó un test chi-cuadrado y el p-valor arrojó que las diferencias son significativas con un 99,9% de confianza.

La siguiente tabla muestra el porcentaje de compras de cada canal que son hechos en Región o en la Región Metropolitana.

		Region	
		Región	Región Metropolitana
		% de la fila	% de la fila
Venta.Distance	Tienda	54.6%	45.4%
	Distancia	38.7%	61.3%

Tabla 8.2: Proporciones de venta por región a distancia y en tienda.

Se puede ver que del total de las ventas en tienda, el 54.6% son hechas desde región, mientras que para las ventas a distancia, el 61,3% son hechas por clientes de la Región Metropolitana.

8.3. Variables RFM Multicanal

Se calcularon las variables RFM *Recency*, *Frecuency* y *Monteray Value* para cada cliente para las compras a distancia, en tienda y en total.

Variable	Media	Desv. Estándar
M.DISTANCIA	68.691,13	234.694,10
M.TIENDA	512.944,01	717.008,64
M.TODO	581.635,14	805.198,68
F.DISTANCIA	0,68	2,21
F.TIENDA	14,12	18,68
F.TODO	14,80	19,23
R.DISTANCIA	260,51	201,62
R.TIENDA	164,83	175,39

Tabla 8.3: Media y desviación estándar para compras a distancia y en tiendas.

Los valores de la tabla anterior corresponden al total de los clientes, es decir, los valores de las frecuencias de distancia de los clientes que sólo compran en tienda fueron reemplazados por 0. Para los valores de *Recency* se calculó el promedio y la desviación estándar sólo para los clientes con actividad en el canal.

A continuación se presentan las medias de estas categorías diferenciadas por sexo y por si recibe o no correos electrónicos.

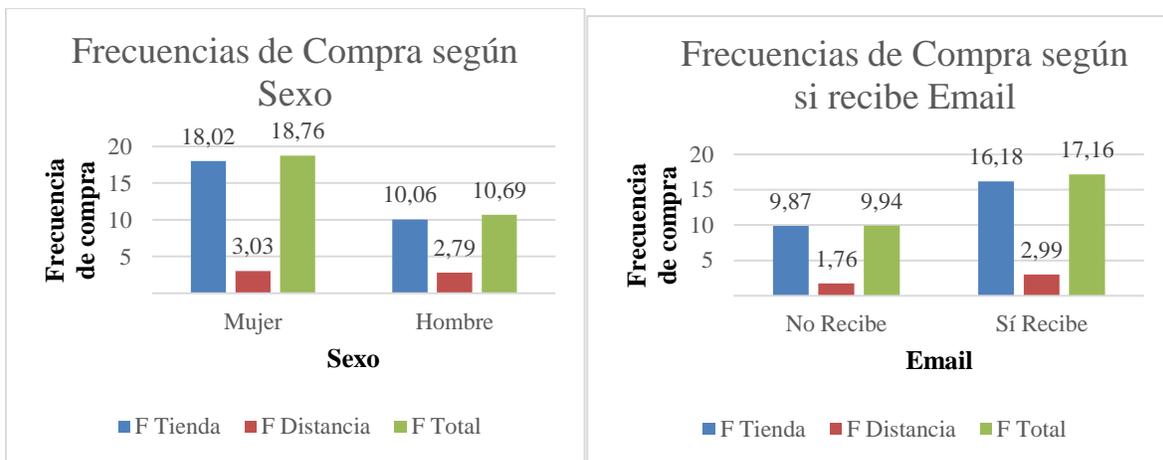


Gráfico 8.5: Frecuencia de compras según sexo y si recibe email.

Como ya se vio anteriormente, las mujeres tienen mayores frecuencias de compra en tienda, a distancia y en total, sin embargo la diferencia disminuye en la frecuencia de compra a distancia. Se hizo el test t-student para comparar las medias para hombres y mujeres y arrojó que con un 99% de confianza las medias son significativamente diferentes.

El segundo gráfico muestra las diferencias entre las medias de las frecuencias para los clientes que reciben email y los que no reciben. Se puede ver que para los tres tipos de frecuencia las medias son mayores para los que sí reciben emails.

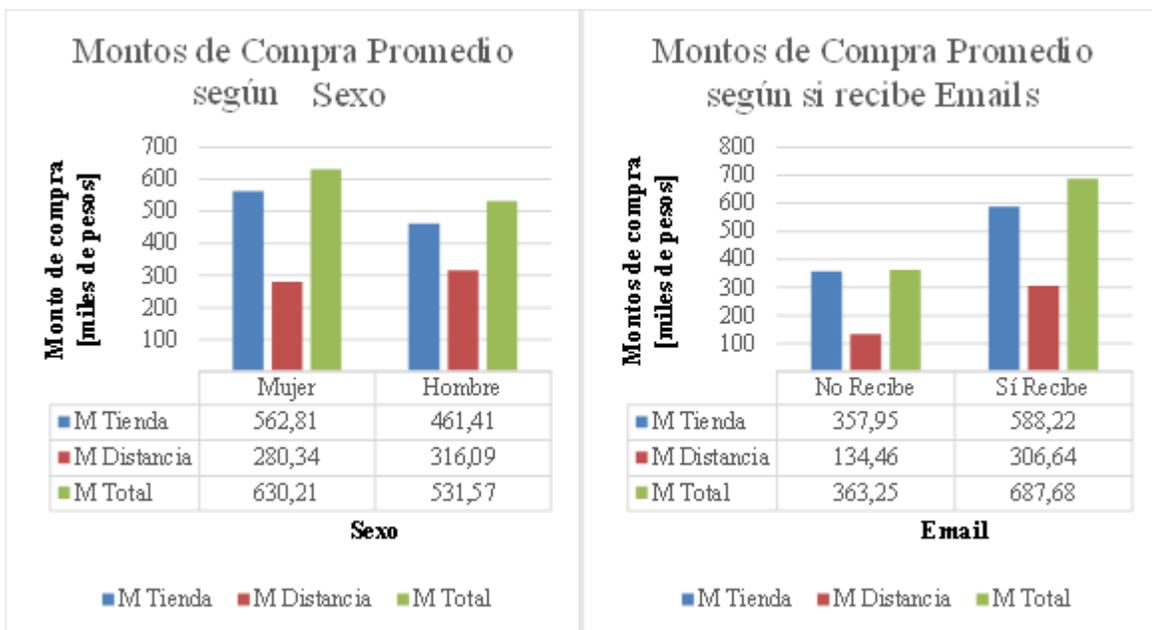


Gráfico 8.6: Montos de compra según sexo y si recibe email.

Se puede ver que las mujeres tienen mayores montos de compra en tienda pero los hombres a distancia. Considerando que la frecuencia de compra de las mujeres es mayor que la de los hombres, se puede concluir que los hombres en promedio hacen compras de montos más grandes que las mujeres, es decir, el ticket promedio a distancia es mayor.

A continuación se analizarán las variables de multicanalidad. Se consideran clientes multicanal los que tengan al menos una actividad multicanal. Estas actividades son navegar, comprar a distancia y recibir *emails*.

		EDAD	MAXS.SEXO_DEF	
			FEMENINO	MASCULINO
		Media	% de la fila	% de la fila
Recibe.Email	No Recibe Email	50.29	51.9%	48.1%
	Recibe Email	40.75	50.5%	49.5%

Tabla 8.4: Recibe email según edad y sexo.

El promedio de edad de los clientes que no recibe email es de 50,29, mientras que los que sí reciben tienen un promedio de 41,75. Se realizó el test t-student para muestras independientes y, con un 95% de confianza arrojó que las medias son significativamente diferentes. Este resultado es esperable, ya que los clientes más jóvenes están más en contacto con la tecnología. Se puede ver, además, que el porcentaje de mujeres es mayor en el grupo de clientes que no recibe emails que el que sí recibe.

La siguiente tabla describe a los clientes de cada categoría multicanal. La variable *Multicanal* se construyó de la siguiente forma: Se le asignó valor 0 a los clientes que sólo tienen actividad en tienda, es decir, a los clientes monocanal⁶; el valor 1 a los clientes que tienen actividad de navegación⁷, valor 2 a los clientes que compran a distancia⁸ y el valor 3 a los clientes que compran a distancia y tienen actividad de navegación.

A continuación se muestra la proporción del total que ocupa cada uno de estos grupos.

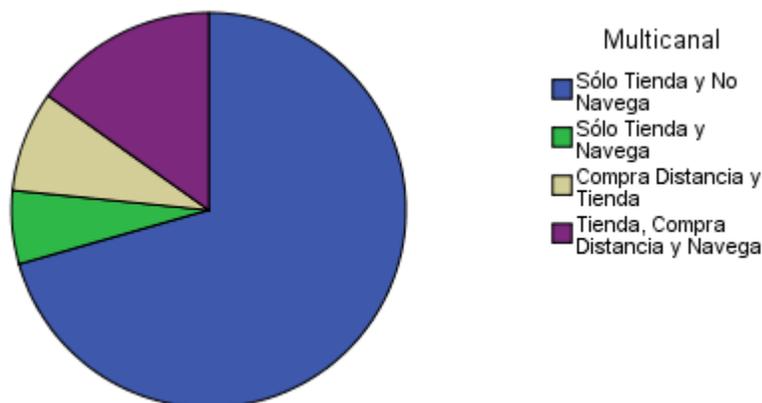


Gráfico 8.7: Proporción según el canal.

⁶ Se considera que todos los clientes tienen actividad en tienda, ya que los clientes que sólo tienen actividad a distancia es de 33, es decir 0,03%.

⁷ Este grupo de clientes no tiene compras a distancia.

⁸ Este grupo de clientes no tiene actividad de navegación, es decir son los clientes que hacen compras a través del Fonocompras.

Multicanal

		Frecuencia	Porcentaje
Válidos	Sólo Tienda y No Navega	69803	70.6
	Sólo Tienda y Navega	5938	6.0
	Compra Distancia y Tienda	8160	8.3
	Tienda, Compra Distancia y Navega	14994	15.2
	Total	98895	100.0

Tabla 8.5: Frecuencia y porcentaje según el canal.

Se puede ver en el siguiente cuadro las características de estos grupos de cliente.

		Recibe.Email		EDAD	MAXS.SEXO_DEF	
		No Recibe Email	Recibe Email	Media	FEMENINO	MASCULINO
		%	%		%	%
Multicanal	Sólo Tienda y No Navega	44.0%	56.0%	45	50.3%	49.7%
	Sólo Tienda y Navega	4.4%	95.6%	38	51.9%	48.1%
	Comra a Distancia y Tienda (No Navega)	7.4%	92.6%	45	53.3%	46.7%
	Tienda, Compra Distancia y Navega	4.5%	95.5%	41	52.8%	47.2%

Tabla 8.6: Características grupo multicanal.

El grupo de clientes que sólo compra en tienda y no navega concentra el mayor porcentaje de clientes que no recibe email, llegando a un 44%. Este porcentaje disminuye drásticamente al incluir actividad de navegación, ya que al iniciar una sesión en la página web los clientes pueden optar a recibir información y promociones por correo electrónico. Los grupos con mayor edad son los de clientes y el de clientes que compran a distancia pero no navegan⁹. El promedio de edad disminuye a 41 en los clientes que compran a distancia y navegan y luego a 38 en los clientes que navegan y sólo compran en tienda. En este grupo se encuentran los clientes más jóvenes con menos poder adquisitivo, por lo que, posiblemente, tampoco cuentan con tarjeta de crédito para comprar vía web.

Se calculó para cada cliente las variables RFM Multicanal y los resultados son los siguientes:

⁹ Recordar que este grupo corresponde a los clientes que compran por vía telefónica.

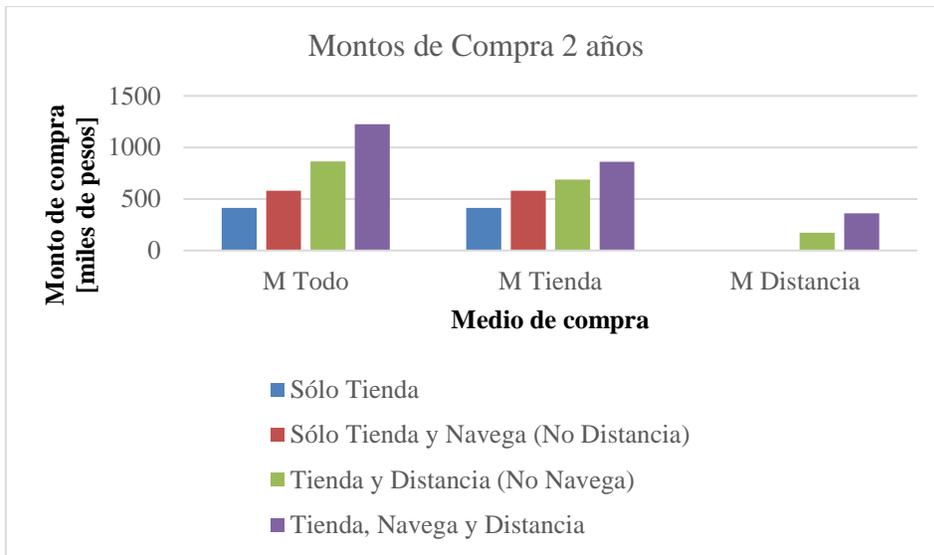


Gráfico 8.8: Montos de compra en 2 años por canal.

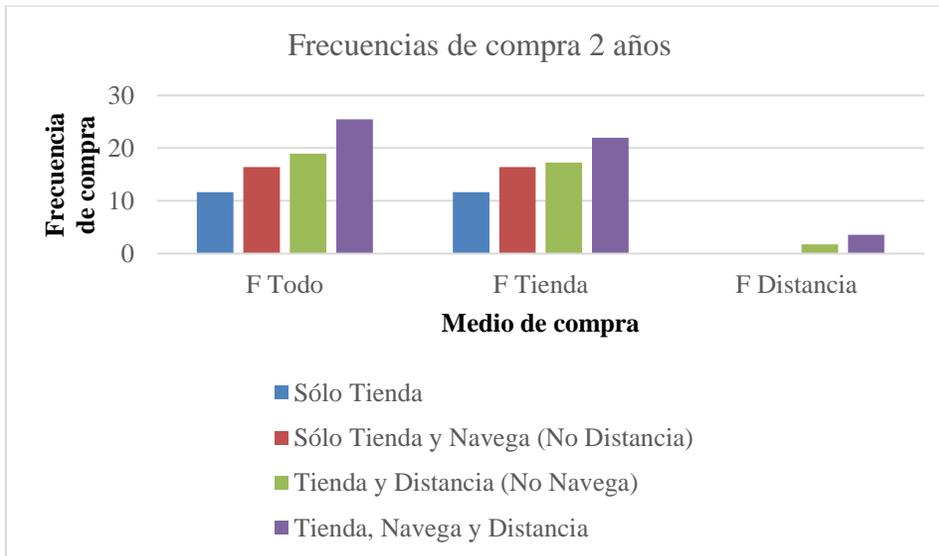


Gráfico 8.9: Frecuencia de compras en 2 años por canal.

Se hizo un test t-student para comparar estas medias y con un 95% de confianza arrojó que son todas diferencias significativas. Se puede ver que los clientes que compran a distancia compran más veces y mayores montos totales que los clientes que compran sólo en tienda. Por otro lado, si se le suma la actividad de navegar, aumenta también las frecuencias y los montos totales promedio de los clientes.

Se hizo luego una matriz de correlaciones para estudiar la relación que tienen entre ellas las variables estudiadas. El resultado es el siguiente:

CORRELACION	M.DIST	M.TIENDA	F.DISTANCIA	F.TIENDA	R.DISTANCIA	R.TIENDA
M.DISTANCIA	1	0.235	0.777	0.173	-0.577	-0.085
M.TIENDA		1	0.195	0.744	-0.216	-0.256
F.DISTANCIA			1	0.191	-0.614	-0.089
F.TIENDA				1	-0.203	-0.273
R.DISTANCIA					1	0.136
R.TIENDA						1

Tabla 8.7: Relación entre las variables estudiadas.

En la tabla aparece el coeficiente de correlación de Pearson, que, en este caso, tiene un 99% de significancia. En la tabla están marcadas con verde las correlaciones positivas (verde oscuro las correlaciones mayores a 0,5 y con verde claro las que son mayores que 0 y menores a 0,5) y de naranja las negativas (naranja claro las que son menores a -0,5 y naranja oscuro las que están entre 0 y -0,5). Las variables de monto y de frecuencia tienen correlación positiva entre sí, mientras que la correlación con las variables de *recency* es negativa. Esto quiere decir que los clientes que compran más veces, compran también montos mayores en total. Se puede ver que las correlaciones más altas son entre las variables de un mismo canal. En el caso de las compras a distancia, los clientes que compran más frecuentemente compran en total montos más altos, lo que es esperable considerando que la variable de monto corresponde a la suma total de las compras hechas durante el período. En ambos canales la correlación es superior a 0,7. Las correlaciones cruzadas (distancia con tienda) son cercanas al 0,2 para las variables de frecuencia y monto. Esto quiere decir, que las personas que compra más y más veces en tienda también lo hace a distancia. Para el caso de *recency*, las correlaciones más grandes (en valor absoluto) se tiene para las ventas a distancia, es decir, a medida que disminuyen los valores de *recency* (más cercana es la última compra a la fecha de referencia) aumentan los montos totales y las frecuencias de compra. También existe una correlación negativa entre *recency* de compra a distancia y los montos y frecuencias en tienda, sin embargo esta correlación es menor en valor absoluto. Las variables de monto y frecuencia y en tienda tienen una correlación similar con las variables de *recency* en tienda y a distancia (alrededor de -0,25).

8.4. Conclusiones Análisis Descriptivo

La distribución de las categorías depende del canal de venta de forma significativa, ya que, por ejemplo, la categoría con mayores ventas a distancia es Electro Hogar mientras que en tienda es Niños. A pesar de que en ambos canales existe mayor presencia femenina, esta diferencia disminuye en las ventas a distancia. Las mujeres presentan, además, mayores montos totales de venta en tienda, sin embargo los hombres tienen mayores montos totales en venta a distancia.

Se puede ver que el promedio de edad es más bajo en los clientes con navegación en la página web. Esto no se ve reflejado directamente en las ventas a distancia, ya que el grupo de clientes monocanal y el de clientes con venta telefónica tienen el mismo promedio de edad y el más alto de los grupos caracterizados.

El grupo de clientes que recibe *emails* corresponde principalmente a clientes con actividad en más de un canal, ya que sólo el 56% de los clientes monocanal tiene activa esta opción, mientras que en los otros grupos es mayor al 90%. Estos clientes tienen una mayor frecuencia de compra y mayores montos totales de compra. Esto puede deberse a que las campañas de *email marketing* tienen efectos positivos en el comportamiento de los clientes, como a que sea una característica propia de los clientes.

Finalmente se puede ver que los clientes que tienen más actividad multicanal (navegación y/o compra a distancia) presentan, en promedio mayores frecuencias y montos de compra en ambos canales. Además, las variables Frecuencia y Monto están correlacionadas positivamente cuando corresponden al mismo canal, mientras que las variables de Frecuencia y Recency presentan una correlación negativa.

9. Segmentación

9.1. Segmentación RFM

Las variables que se consideraran para la segmentación son:

- *R.Tienda* (Recency Tienda): Cantidad de días entre la última compra en tienda en el período y último día de este.
- *M.Tienda* (Monto Tienda): Suma de los montos de todas las compras en tienda hechas en el período.
- *F.Tienda* (Frecuencia Tienda): Cantidad de compras en tienda hechas en el período.
- *R.Distance* (Recency Distancia): Cantidad de días entre la última compra a distancia en el período y último día de este.
- *M.Distance* (Monto Distancia): Suma de los montos de todas las compras a distancia hechas en el período.
- *F.Distance* (Frecuencia Distancia): Cantidad de compras a distancia hechas en el período.

Se dividió el tiempo total de estudio (2 años) en cuatro periodos iguales de seis meses cada uno.

- T1 = 2 de mayo de 2013 hasta 31 de octubre de 2013
- T2 = 1 de noviembre de 2013 hasta 30 de abril de 2014
- T3 = 2 de mayo de 2014 hasta 31 de octubre de 2014
- T4 = 1 de noviembre de 2014 hasta 30 de abril de 2015

Se calcularon las variables para cada cliente para cada semestre. De esta forma, la tabla de datos que se utilizará tiene las siguientes columnas:

- Semestre
- Id.Cliente
- R.Tienda
- F.Tienda
- M.Tienda
- R.Distance
- F.Distance

- M.Distance

El valor de *R.Tienda* y *R.Distance* se calculó como la cantidad de días que transcurrieron desde la última compra del período hasta el final de éste, en cada uno de los canales respectivos, es decir, tomando el Máximo de las fechas en las que el cliente realizó una compra. El máximo de todos los valores es de 182 días, por lo que se reemplazó por 183 los valores de los clientes que no habían realizado ninguna compra durante el semestre.

Los valores de *F.Tienda*, *F.Distance*, *M.Tienda* y *M.Distance* de clientes sin actividad en ese período fueron reemplazados por 0.

Para que todas las variables aportaran de la misma manera en la segmentación, cada variable se normalizó de la siguiente manera para obtener una distribución $N(0,1)$:

$$X.N = \frac{X - \mu_x}{\sigma_x}$$

Es así como las variables que se utilizaron para la segmentación fueron:

R.Tienda.N	F.Tienda. N	M.Tienda. N	R.Distance. N	F.Distance. N	M.Distance. N
------------	----------------	----------------	------------------	------------------	------------------

Para determinar la cantidad de clusters se calculó las distancias intra grupos¹⁰ de 2 a 8 clusters y el resultado es el siguiente.

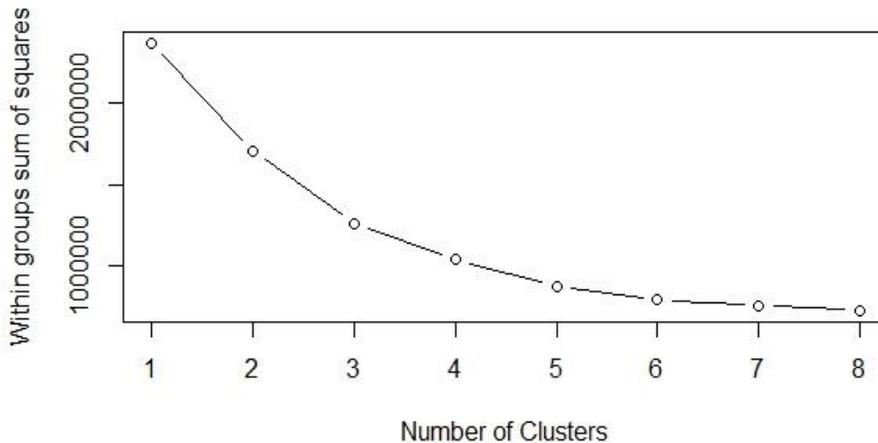


Gráfico 9.1: Distancia intra grupos según número de conglomerados.

¹⁰ Se utilizó la suma de diferencias al cuadrado para calcular las distancias entre los elementos de cada grupo.

La siguiente tabla muestra el cambio porcentual al aumentar la cantidad de conglomerados.

Clusters	Distancia intra grupos	Disminución porcentual
1	2.373.474	
2	1.709.641,7	28.0%
3	1.261.162,4	26.2%
4	1.037.159,8	17.8%
5	876.849,6	15.5%
6	792.658,8	9.6%
7	757.603	4.4%
8	724.360,8	4.4%

Tabla 9.1: Cambio porcentual al aumentar la cantidad de conglomerados.

Utilizando la “regla del codo”, se decide hacer el análisis con 3, 4 y 5 clusters y se evaluará cuál es la mejor alternativa.

4 Clusters

Con el método de k-medias, luego de 34 iteraciones se convergió en los siguientes conglomerados:

Conglomerado	Cientes por Conglomerado	Porcentaje
1	189.210	47,83%
2	176.842	44,70%
3	26.344	6,66%
4	3.184	0,80%

Tabla 9.2: Conglomerados.

Los centros de conglomerados son:

Conglomerado	F Tienda	M Tienda	R Tienda	F Distancia	M Distancia	R Distancia
1	0,69	23.430,75	159,82	0,03	1.496,26	182,35
2	6,11	221.258,58	34,48	0,04	2.203,24	182,12
3	5,87	221.407,82	66,92	1,40	150.617,88	66,01
4	8,43	394.239,39	56,89	4,71	676.067,92	51,38

Tabla 9.3: Centros de conglomerados.

Los siguientes gráficos muestran cómo se comportan las medias para cada una de las variables en los distintos conglomerados.

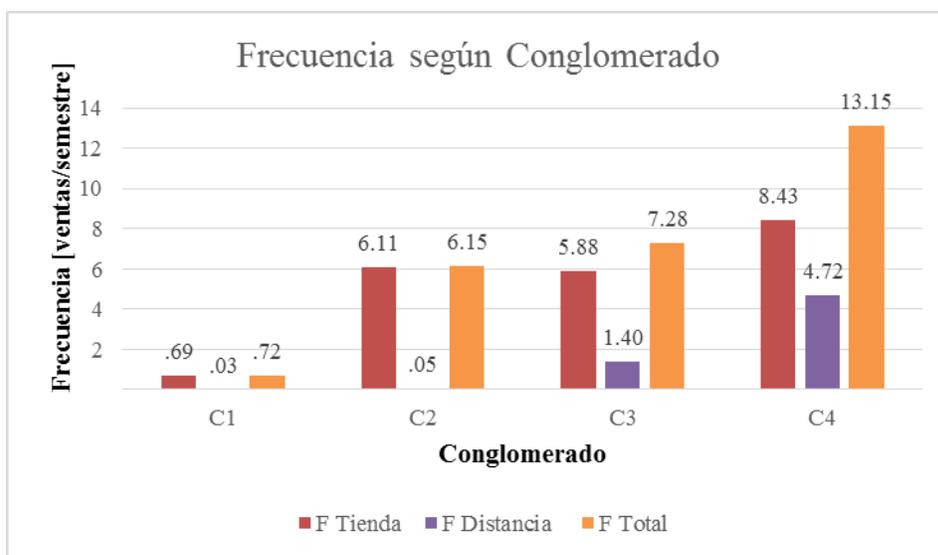


Gráfico 9.2: Frecuencia según conglomerado.

El C1 es el segmento que menos actividad tiene en promedio. Para ambos canales el promedio es menor a 1, es decir realizan menos de una compra al semestre. El C2 contiene a los clientes monocanal, ya que el mayor promedio de frecuencias lo tienen en tienda y la frecuencia a distancia es casi nula. El C3 corresponde a los clientes multicanal con poca actividad. Éste es similar al C4 en el sentido de tener actividad en ambos canales, sin embargo las intensidades son diferentes. El C4 es el segmento de clientes más activo, ya que en promedio compran mucho en tienda y a distancia. Se puede decir, entonces que

- Montos

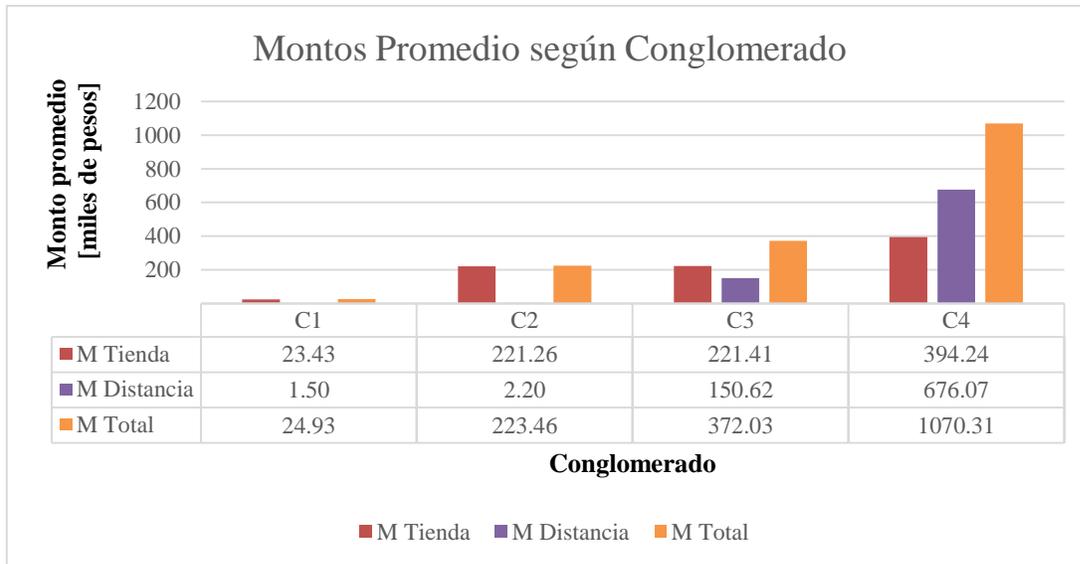


Gráfico 9.3: Montos promedio según conglomerado.

Las variables de monto se comportan similar a las de frecuencia en cada uno de los conglomerados. El C1 corresponde a los clientes que en promedio gastan menos semestralmente. El C2 contiene a los clientes monocanales, ya que las compras que hacen a distancia es mucho menor que el total de compras que hacen en tienda. El C3 desembolsa una cantidad muy similar al C2 en tienda, sin embargo es a distancia donde se diferencia de este segmento. Se puede decir, entonces, que este segmento contiene a clientes multicanal. El C4 también corresponde a los clientes multicanal, pero, a diferencia de los clientes de C3, estos compran en promedio más a distancia que en tienda, es decir, cada vez que compran, gastan montos más grandes a distancia que en tienda. Salvo para el C4, la información que entrega la variable de frecuencia y la de monto es la misma, ya que más compras implican un gasto total mayor, pero en el último segmento, a pesar de comprar más frecuentemente en tienda, se gasta más en el semestre en compras a distancia.

- Recency

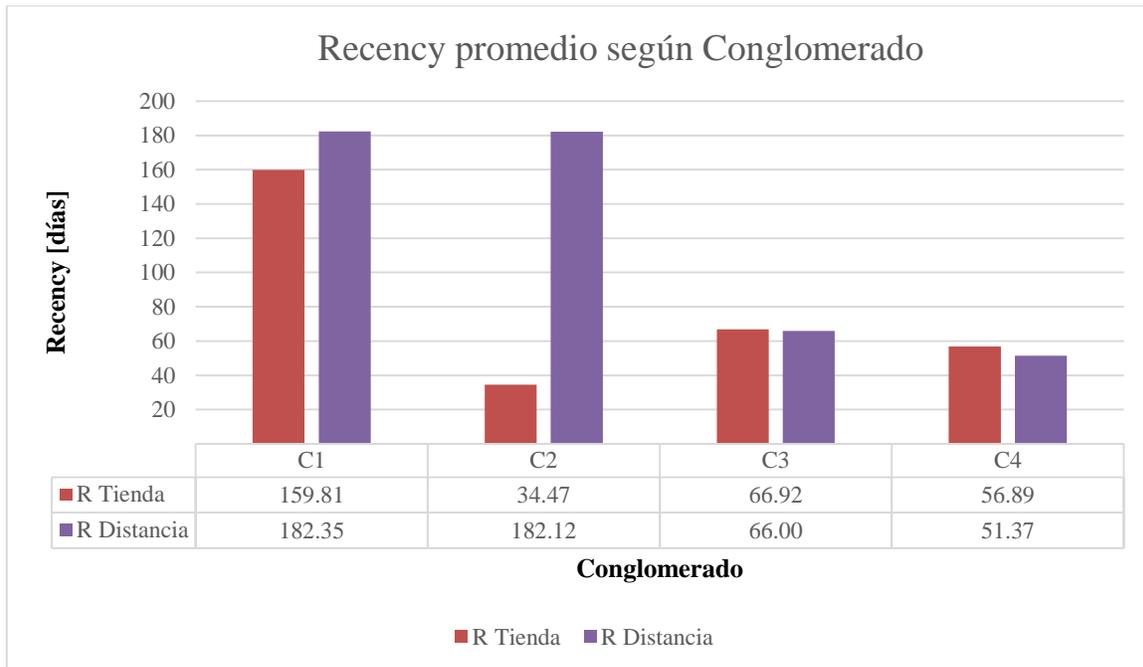


Gráfico 9.4: Recency promedio según conglomerado.

El análisis de Recency entrega resultados similares a los obtenidos de las frecuencias y los montos promedio. El C1 y C2 tienen recency promedio mayor a 182. Considerando que cada semestre tenía como máximo 182 días y se les asignó el valor 183 a los clientes sin actividad durante ese período, se puede decir que estos dos segmentos en promedio no tienen actividad durante el semestre (lo que se concluye también del análisis de las frecuencias).

Se puede ver que los clientes del C4 son los mejores clientes, ya que compran muchas veces, montos altos y son muy activos al tener bajos valores de Recency, sin embargo corresponde sólo al 0,8% del total de clientes. Los siguientes gráficos muestran el aporte neto que hace cada uno de los segmentos.

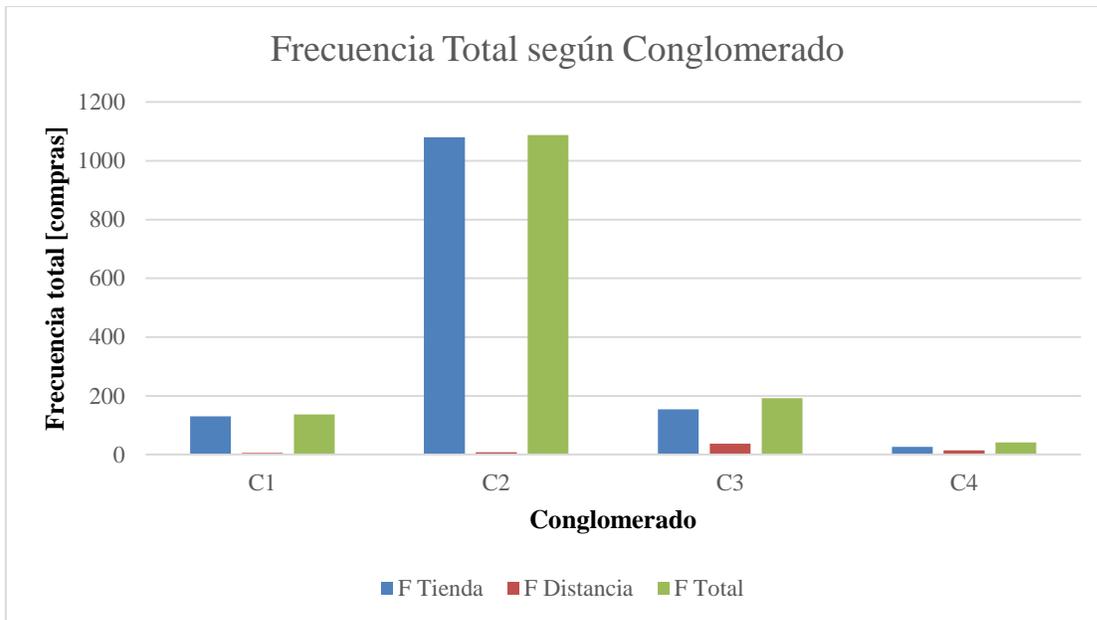


Gráfico 9.5: Frecuencia total según conglomerado.

Se puede ver que el C2 tiene una frecuencia total muy superior al resto de los clusters, a pesar de tener una suma de frecuencias a distancia casi nula. El C3 tiene frecuencias mayores al C1 y se caracteriza por ser el segmento con más compras a distancia. El C4, a pesar de ser el segmento de clientes con más frecuencias de compras, no tiene un gran impacto en el total de las ventas al ser un grupo tan pequeño.

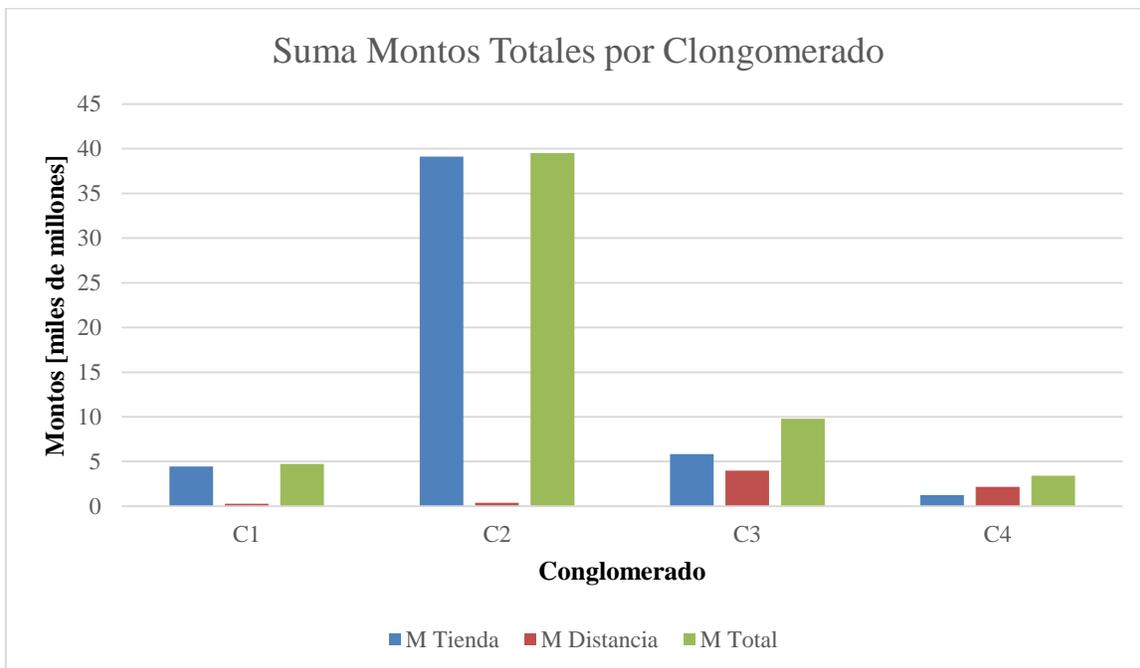


Gráfico 9.6: Montos totales por conglomerado.

Al igual que para las frecuencias, el C2 es el que presenta más altos niveles de compra total, es decir es el segmento más importante ya que realiza el 68,8% de las ventas totales de la empresa. El C3 ocupa el segundo lugar en el nivel de ventas. A pesar de representar

sólo al 6.66% del total de los clientes, tiene un nivel de ventas superior al C1, y cada cliente aporta más a nivel individual. Este es un segmento interesante, ya que es el segmento que más aporta en las ventas a distancia. El C4, a pesar de ser el segmento con mejor calidad de clientes (alta frecuencia y montos totales de venta), al ser un grupo tan pequeño, tiene poco impacto en las ventas totales (5,9% del total de ventas)

Se puede decir, entonces que los conglomerados se pueden caracterizar de la siguiente forma:

C1 - Clientes Inactivos: Son los clientes con poca actividad tanto en tienda como a distancia. Además la suma total de las ventas son bajas para este tipo de cliente y el tener altos valores de *R.Tienda* y *R.Distance* indica que tiene poca actividad dentro del período. Son el segmento más grande, por lo que, a pesar de ser los clientes con menos actividad promedio, la suma neta ocupa el tercer lugar entre los conglomerados.

C2 – Clientes Monocanal: Estos clientes tienen alta frecuencia de compra en tienda y baja en distancia, al igual que los montos totales de compra. El valor de Recency indica lo mismo, ya que en tienda tiene una valor de 34.48 y en distancia, 182,12, es decir prácticamente no tienen actividad a distancia. Este es el segmento de clientes que más aporta a la empresa y aporta un 68.8% de las ventas totales.

C3 – Clientes Multicanal Bajos: Estos clientes tiene actividad a distancia y en tienda, sin embargo sus valores son menores a los del C4. Ocupa el segundo lugar en valor neto de compras, a pesar de tener menos clientes que el C1.

C4 – Clientes Multicanal Altos: Este es el segmento de clientes de mejor calidad. Son clientes multicanales y tienen los indicadores más altos comparados con el resto de los segmentos. Son los clientes más valiosos, ya que interactúan frecuentemente con la empresa a través de los distintos canales. A pesar de esto, es el segmento que menos aporta en la suma total de ventas, ya que es de un tamaño muy reducido.

5 Clusters

Al realizar la segmentación con 5 clusters, uno de ellos contiene sólo al 0,3% de los clientes, por lo que la información que aporta no es relevante en comparación a la segmentación anterior.

3 Clusters

Luego de 73 iteraciones, los conglomerados resultantes son los siguientes:

Conglomerado	Clientes por Conglomerado	Porcentaje
1	27.281	6,9%
2	178.108	45,0%
3	190.191	48,1%

Tabla 9.4: Clientes por conglomerado en 73 iteraciones.

Los centros de conglomerado son:

Conglomerado	F Tienda	M Tienda	R Tienda	F Distancia	M Distancia	R Distancia
1	6,28	247.081,16	64,65	1,83	218.698,53	60,69
2	6,11	221.407,55	34,48	0,05	2.711,60	181,60
3	0,69	23.524,20	159,71	0,03	1.808,38	181,99

Tabla 9.5: Centros de conglomerado.

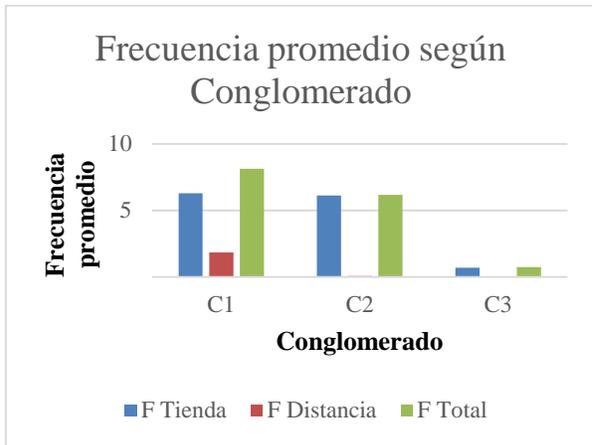


Gráfico 9.7: Frecuencia promedio según conglomerado.

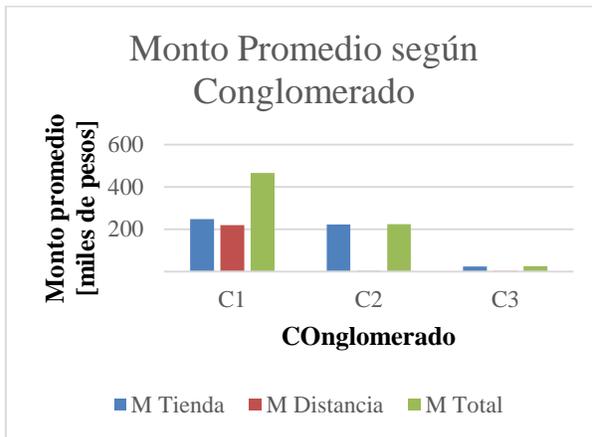


Gráfico 9.8: Monto promedio según conglomerado.

Se puede ver que el C1 corresponde a los clientes multicanal, ya que presentan valores significativos de frecuencia y montos de compras a distancia. Por otro lado, en promedio, son los clientes que más compran también en tienda. Al realizar la prueba t-student para comparar las medias de frecuencia y monto entre el C1 y C2, se obtiene que las diferencias son significativas, es decir, el segmento de clientes C1 tiene mayor frecuencia y mayores montos de venta que el C2, que corresponde al segmento de clientes principalmente monocanal. Por otro lado, el C3 corresponde al segmento de clientes con menos actividad.

Los gráficos a continuación muestran el aporte neto que hace cada uno de los segmentos en las frecuencias y las ventas totales.

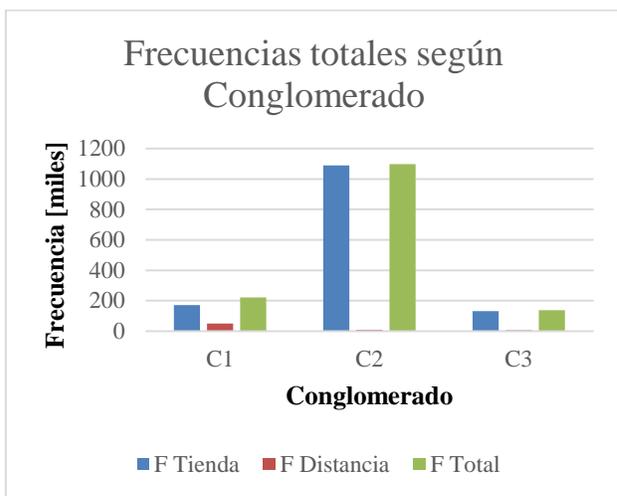


Gráfico 9.9: Frecuencias totales según conglomerado.

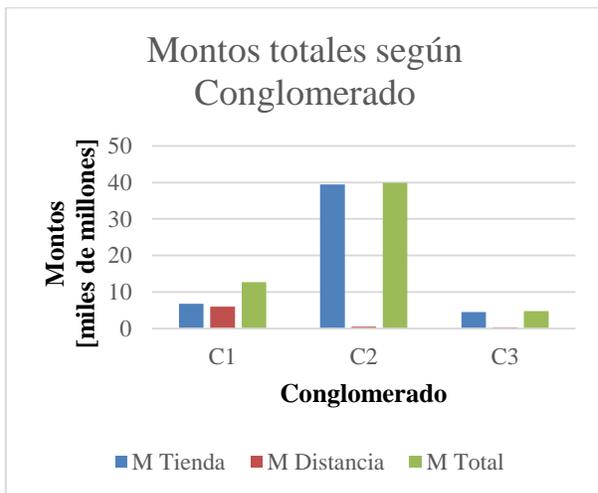


Gráfico 9.10: Montos totales según conglomerado.

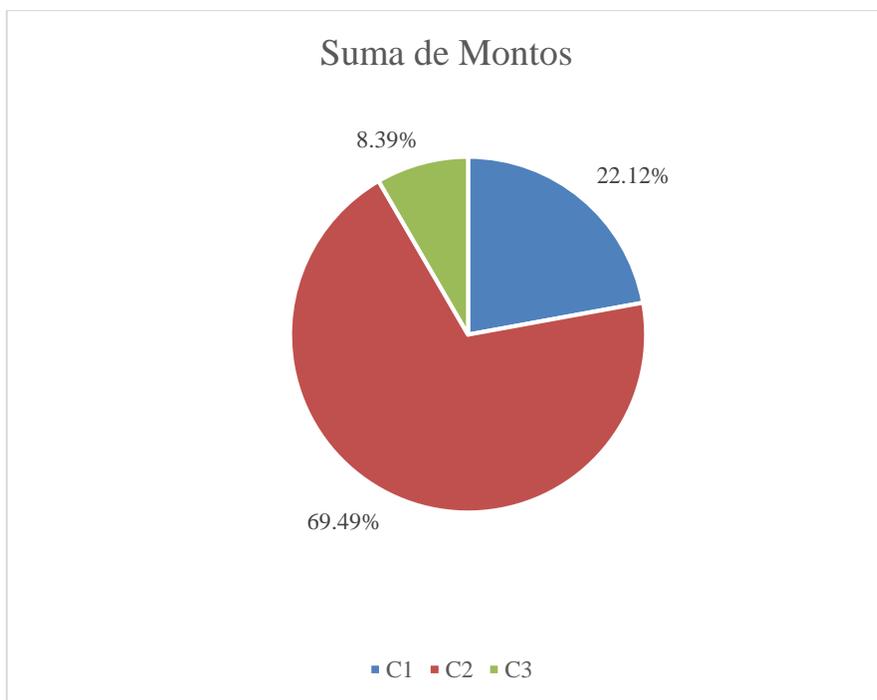


Gráfico 9.11: Suma de montos por conglomerado.

Se puede decir, entonces que los clientes se pueden clasificar de la siguiente forma:

C1 – Clientes Multicanal: Son los clientes con más actividad, tanto en frecuencia promedio por cliente como montos totales por cliente. Es el único segmento con actividad a distancia y tiene, además constante actividad en ambos canales. En términos netos aportan un 22,12% de las ventas y corresponde al 6,9% de los clientes.

C2 – Clientes Monocanal: Tienen actividad principalmente en tienda. En promedio tienen menos frecuencia y menos montos de compra en tienda que el C1. El valor de *R.Tienda* promedio es mayor que el C1, es decir, compran de manera más constante. En términos netos aportan un 69,49% de las ventas totales y está conformado por el 45% de los clientes.

C3 – Clientes Inactivos: Corresponde al segmento de clientes con menos actividad, ya que posee altos valores de *Recency* y bajos valores de frecuencia y montos de compra. Es el segmento más grande (48,1% del total de clientes) y aportan sólo el 8,39% de las ventas.

Al comparar estas dos segmentaciones se puede ver que los clientes están mejor distribuidos en cantidad al separarse en 3 segmentos. Si bien es interesante el comportamiento de los clientes del C4 de la primera segmentación, ya que corresponde a los clientes que gastan altos montos a distancia y en tienda, su tamaño es muy pequeño y en términos netos es bajo el aporte que hace a la empresa.

9.2. Segmentación con Visitas y sin Recency

En esta sección se hace una modificación a las variables utilizadas para segmentar. Al tratarse éste de un estudio del comportamiento multicanal de los clientes, se le agregará la variable Visitas. Esta información está disponible sólo para el último semestre de

estudio, por lo que se utilizaron las variables de frecuencia y monto sólo para el cuarto semestre, es decir, se seleccionaron los datos de la tabla anterior todos los casos correspondientes al último semestre de estudio (entre 1 de noviembre de 2014 y 30 de abril de 2015). En la segmentación anterior no se incluyó esta variable para tener variables de un horizonte más grande de tiempo. Además, se excluyen las variables *Recency.Tienda* y *Recency.Distancia*. Como se dijo anteriormente, para los clientes que no compraron en el período, se consideró la variable de *recency* como el valor máximo más 1, es decir, 183. Esto tiene error, ya que en la práctica no es lo mismo una persona que compró hace realmente 183 días a una que no compró en los 2 años, es decir, en más de 730 días. Es por esto que se decidió eliminar esta variable para evaluar cómo se comportan los clusters que se forman. El siguiente gráfico muestra cómo disminuye la distancia intra grupos a medida que aumenta el número de clusters:

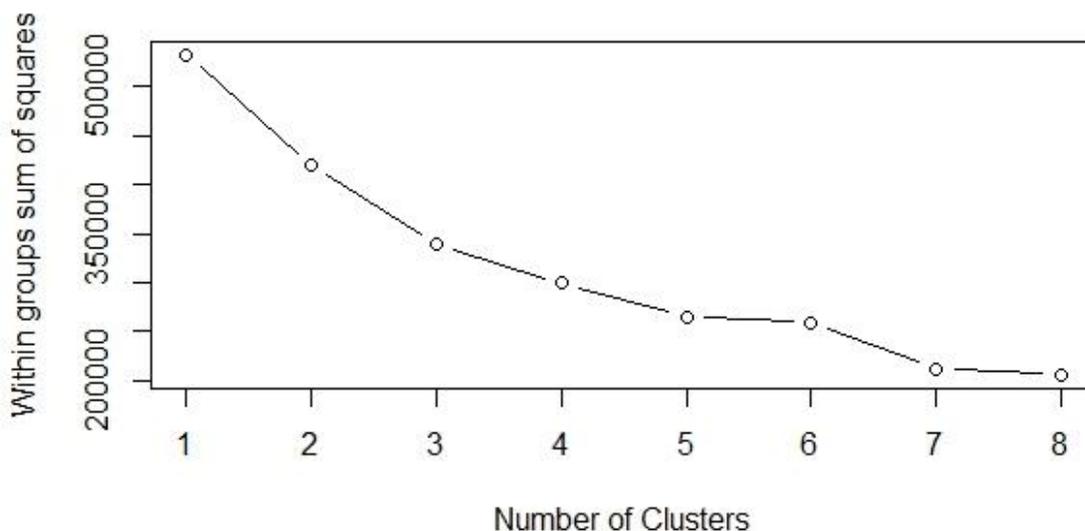


Gráfico 9.12: Distancia intra grupos según número de conglomerados.

Se escogió utilizar 3, 5 y 7 conglomerados.

Al realizar la segmentación con 5 y 7 conglomerados se obtuvieron segmentos que contenían a menos del 0,5% de los clientes, por lo que se descartan estas dos opciones.

Al realizar la segmentación con 3 conglomerados, luego de 50 iteraciones se obtiene la siguiente cantidad de clientes por cluster.

Conglomerado	Cientes por Conglomerado	Porcentaje
1	83.028	83,96%
2	13.371	13,52%
3	2.496	2,52%

Tabla 9.6: Clientes por conglomerado.

Los centros de cada conglomerado se pueden ver en la siguiente tabla.

Conglomerado	Frecuencia Tienda	Monto Tienda	Frecuencia Distancia	Monto Distancia	Vistas
1	1,96	64.963,48	0,10	8.472,58	28,27
2	12,44	517.668,32	0,22	25.679,88	74,13
3	6,43	260.900,30	2,71	402.284,90	1.562,11

Tabla 9.7: Centro de conglomerados.

A continuación se presentan los promedios de las variables para cada conglomerado.

- Frecuencias

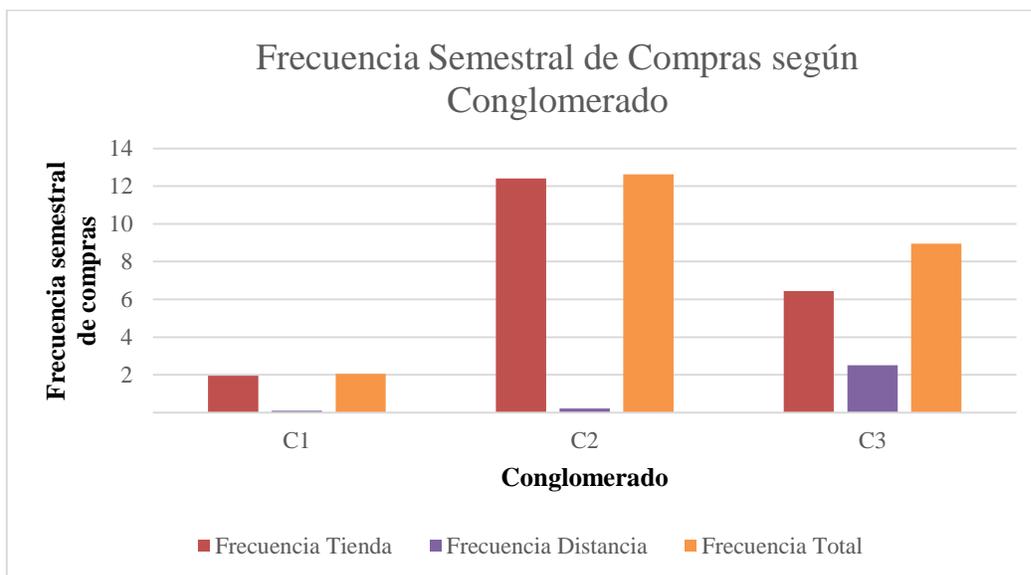


Gráfico 9.13: Frecuencia semestral de compras según conglomerado.

Se puede observar que el C1 es el que contiene a los clientes más inactivos en el período y son, además, principalmente monocanales, ya que sólo compran en tienda. El C2 corresponde a los clientes que compran más frecuentemente y que son monocanales. A pesar de comprar sólo en tienda, son lo que compran más veces en promedio. El C3

corresponde al grupo de clientes multicanal. En promedio son los que más compran a distancia, sin embargo, en promedio, compran menos en tienda que el C2.

- Monto

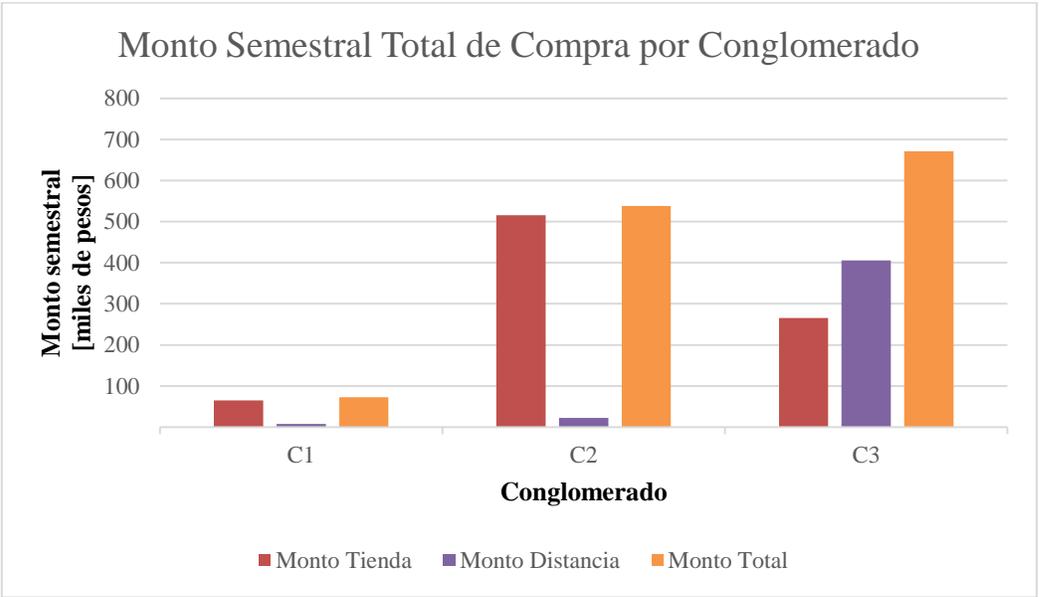


Gráfico 9.14: Monto semestral de compra por conglomerado.

Para el caso de los montos totales promedio de compra del semestre, se puede ver, al igual que en el gráfico anterior, que el C1 corresponde al grupo de clientes con menos actividad. Al C2 pertenecen los clientes monocanal activos, es decir los clientes que compran mucho en tienda y poco a distancia. El C3 corresponde al segmento de clientes multicanal que en promedio gastan más a distancia que en tienda. A diferencia del análisis de frecuencias, este tercer segmento en promedio gasta más en total que el C2. Luego se hará un análisis del aporte total que hace cada uno de estos conglomerados.

- Visitas

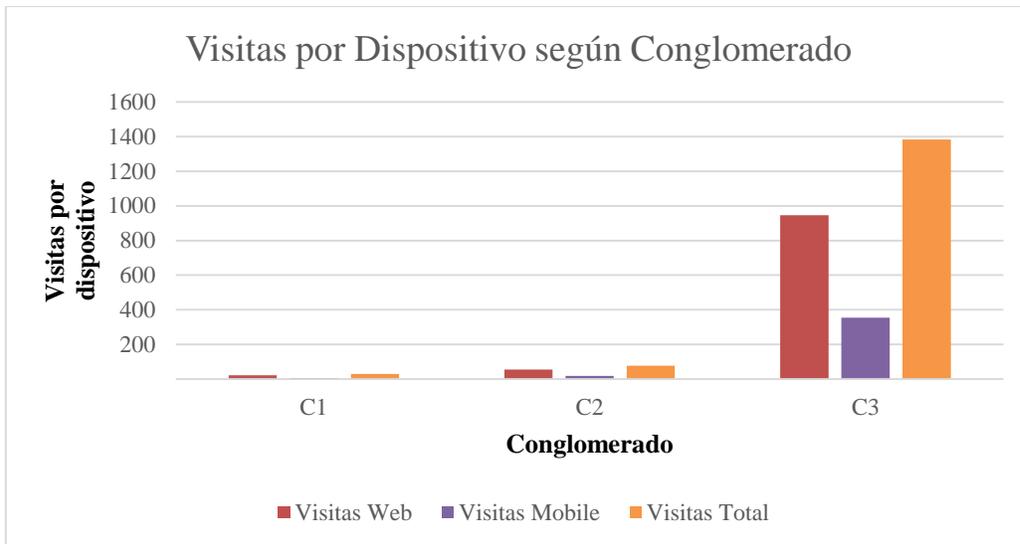


Gráfico 9.15: Visitas por dispositivo según conglomerado.

Como es de esperarse, el C3 es el segmento de clientes que más visita el sitio, ya sea en su modalidad web o mobile.

A continuación se presentan los aportes netos que hacen estos conglomerados tanto para frecuencias de compras como para sumas de montos en el semestre.

- Frecuencia

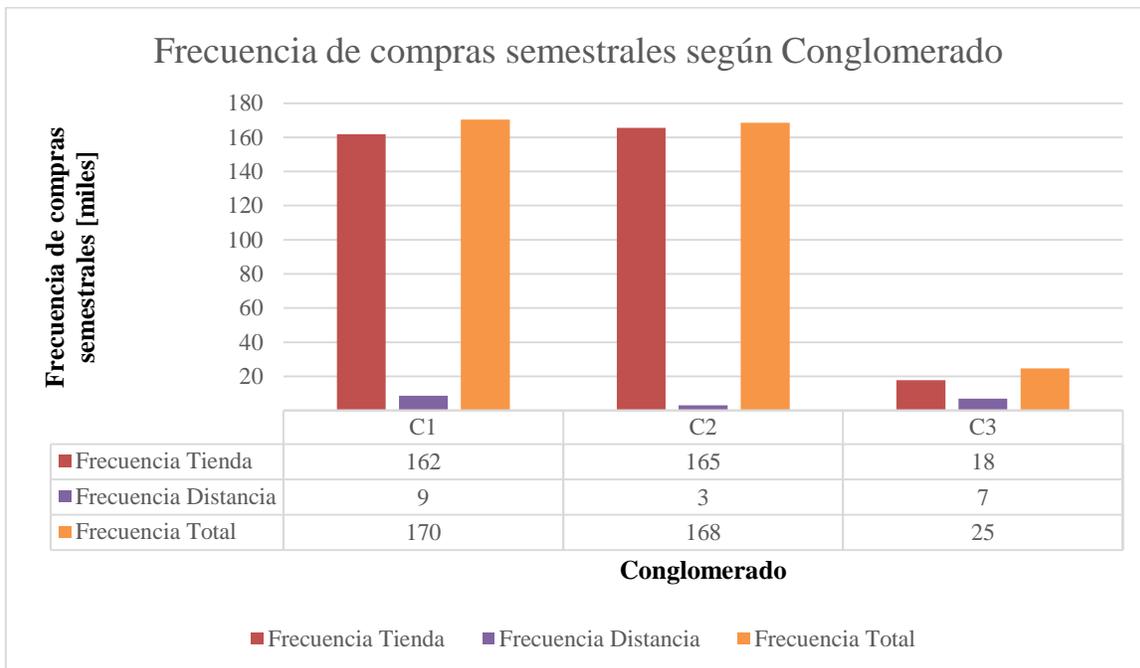


Gráfico 9.16: Frecuencia de compras semestrales según conglomerado.

Se puede observar que el C1 es el que tiene más cantidad de compras totales y a distancia. A pesar de que en promedio tenga sólo 0,1 compras a distancia por cliente al semestre, el hecho de contener al 83.7 del total de clientes hace que la suma total de

compras sea mayor al C3 que tiene un promedio de 2.51 compras a distancia pero sólo el 2,8% de los clientes totales. El C2 es el que tiene mayor frecuencia de compras en tienda y ocupa el segundo lugar en cantidad de compras totales, teniendo sólo un 1,1% menos que el C1. Al ser esta una proporción pequeña, puede cambiar dependiendo del periodo o de la muestra de estudio.

- Monto

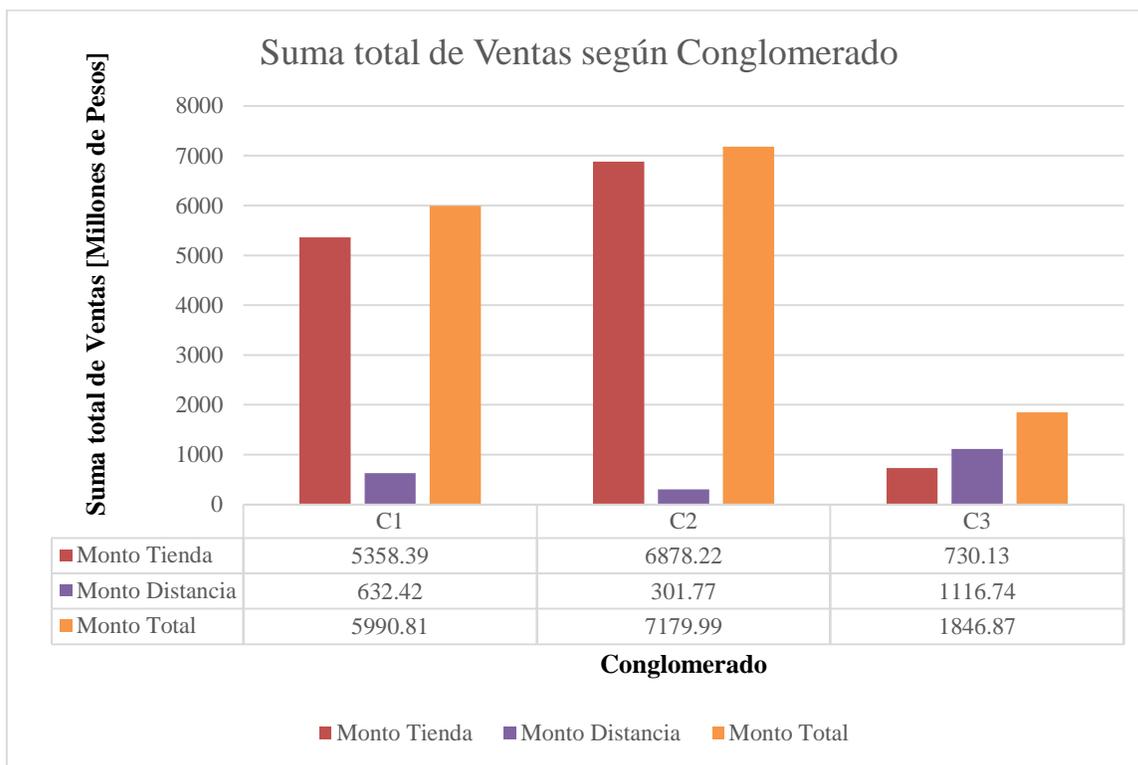


Gráfico 9.17: Suma total de ventas según conglomerado.

En este gráfico se puede ver que el C2 es el conglomerado que aporta mayor cantidad de ventas tanto en tienda como totales. El C3 es el que suma mayor cantidad en ventas a distancia, sin embargo es el segmento que suma menos en ventas totales. El valor del ticket promedio en distancia es mayor para este segmento que el C1, ya que a pesar de realizar menos cantidad de compras, el monto total es mayor.

- Visitas

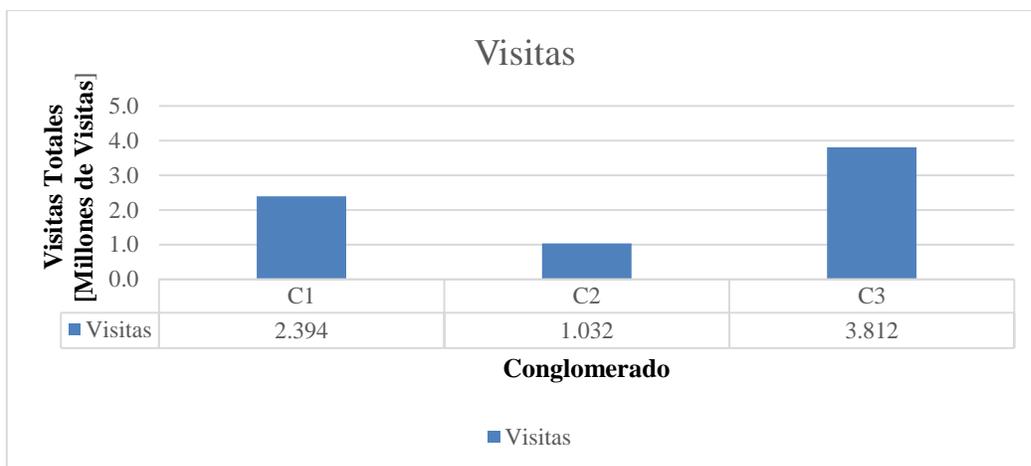


Gráfico 9.18: Visitas.

Para el caso de las visitas podría esperarse que el C1 fuera el que acumulara la mayor cantidad de visitas al sitio, ya que es el segmento que en total tiene más ventas a distancia. Sin embargo es el C3 el conglomerado que acumula la mayor cantidad de éstas. Esto puede deberse a que un mayor número de visitas está asociado a compras de productos más caros a distancia. Otra explicación puede ser que las compras por vía telefónica las realicen principalmente los clientes de este segmento, ya que para las que se realizan por el canal de televisión no hay navegación asociada.

Selección Segmentación

Al comparar ambos tipos de segmentación, se puede ver que coinciden en tener 3 segmentos. A grandes rasgos estos segmentos tienen similares características y son las que se esperaba encontrar a priori: Un segmento con poca actividad, otro segmento con actividad monocanal (clientes que compran principalmente en tienda) y un último segmento más pequeño con actividad multicanal. En la segmentación sin la variable de visitas, el tamaño de los segmentos es más homogéneo y el más pequeño representa al 6.9% del total, diferencia del más pequeño de la última segmentación que representa al 2.5% del total. Por otra parte, se puede ver que en la segmentación sin visitas los segmento monocanal prácticamente no tienen actividad multicanal, a diferencia de los segmentos más grandes de la segmentación final que, a pesar de ser principalmente monocanal, tienen actividad multicanal. Viendo el tamaño de estos grupos se puede ver que parte de los clientes del segmento multicanal de la segmentación sin visitas (que corresponde al 6,9%) en la última segmentación se reparte en los otros segmentos más grandes, lo que se traduce en una disminución de su tamaño y aparición de actividad multicanal en los segmentos principalmente monocanal. Dado esto, se decide seleccionar la segmentación con las variables *recency* y sin la de visitas con 3 conglomerados.

9.3. Segmentación Multicanal

Al tratarse este trabajo de un estudio sobre el comportamiento multicanal de los clientes, y dado que la segmentación hecha del total de los clientes no captura la heterogeneidad de este tipo de clientes (al ser un grupo muy pequeño), se hace un análisis de segmentación incluyendo sólo los clientes con actividad multicanal, es decir, clientes

que tengan *F.Distance* o *Visitas.Total* mayor a cero. En este caso se incluyeron los dispositivos de navegación. Las variables de segmentación fueron las siguientes:

- F.Tienda
- M.Tienda
- F.Distance
- M.Distance
- Visitas.Web
- Visitas.Mobile
- Visitas.Total

Luego de normalizarlas, se calcularon las distancias intra grupo de 2 a 8 conglomerados. El resultado fue el siguiente:

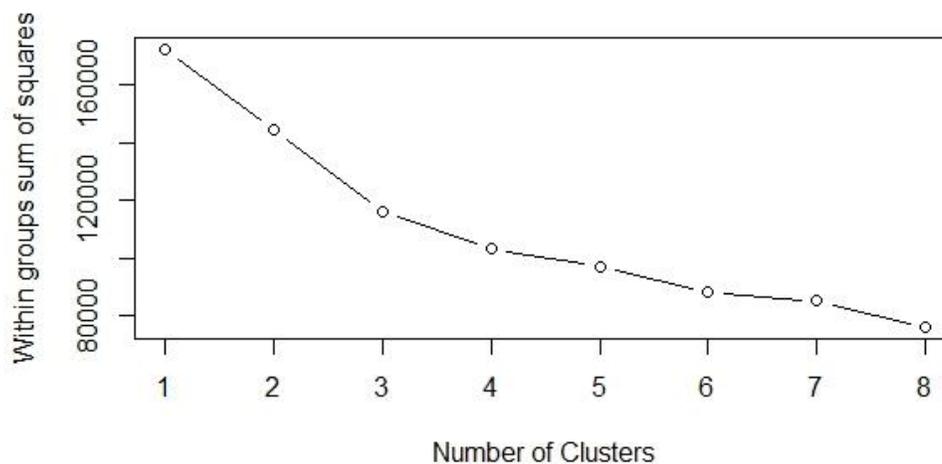


Gráfico 9.19: Distancia intra grupos según conglomerado.

Según la “regla del codo” se decidió hacer una segmentación con 6 conglomerados.

Luego de 25 iteraciones el resultado fue el siguiente:

Conglomerado	Cientes	Porcentaje
1	3.156	12,80%
2	642	2,60%
3	19.241	78,04%
4	31	0,13%
5	1.513	6,14%
6	71	0,29%

Tabla 9.8: Clientes por conglomerado.

Cluster	Frec. Tienda	Monto Tienda	Frec. Distancia	Monto Distancia	Visitas Mobile	Visitas Web	Total Visitas
1	16,55	713959,72	0,70	84.472,27	37,38	143,38	193,75
2	6,77	210483,75	1,19	149.040,65	960,22	2.207,81	3.375,00
3	3,22	112080,51	0,56	43.819,26	24,40	98,54	130,90
4	6,77	239613,97	2,81	249.277,87	10.500,5	1.392,58	12.196,6
5	5,47	233072,97	2,90	526.023,84	69,17	375,19	473,91
6	9,79	338179,16	3,89	294.591,45	967,61	10.119,55	11.903,9

Tabla 9.9: Centro de conglomerado.

Para cada variable, el color rojo más oscuro indica que el valor es más alto, mientras que los más claros son los más bajos, llegando al extremo de ser blanco el mínimo. Se hará un análisis de esta tabla para tener una caracterización integral de estos seis segmentos.

C1 – Clientes Monocanal: Este segmento tiene los valores más altos en *F.Tienda* y *M.Tienda*, mientras que los valores de *F.Distancia* y *M.Distancia* son bajos en comparación al resto de los segmentos (a excepción de C3). Lo mismo ocurre con la cantidad de visitas a la página ya que tiene valores bajos. A pesar de que se están analizando clientes con actividad multicanal, se forma este segmento del 12,8% del total de los clientes, que tiene actividad principalmente monocanal.

C2 – Clientes Multicanal con Actividad Intermedia: Este es el segmento más grande con actividad de navegación (2,6% del total). Tienen más visitas al sitio Web que al sitio Mobile y, si bien tiene actividad de compra multicanal, no es tan intenso como los segmentos C4, C5 y C6.

C3 – Clientes con poca Actividad: Este segmento de clientes tiene los valores más bajos en todas las variables. Tiene más actividad en tienda que a distancia y visita en promedio más el sitio web que el sitio mobile. Es el segmento más grande y corresponde al 78.04% de los clientes.

C4 – Clientes que navegan Mobile: Este es el segmento más pequeño, con sólo un 0,13% del total. Corresponde a los clientes que navegan a través del sitio móvil. El nivel de compra en tienda es similar al del C2. Este tipo de clientes tiene, además un alto nivel de compras a distancia.

C5 – Clientes Multicanal con poca navegación: Este segmento de clientes tiene alto nivel de compras a distancia y bajo nivel de navegación. Estas características indican que puede tratarse del segmento de clientes que compra vía telefónica productos que aparecen en televisión. Representa a un 6,14% de los clientes y es el segmento que más dinero gasta en promedio en ventas a distancia.

C6 – Clientes que navegan Web: Este es el segmento de clientes que tiene un alto nivel de navegación a través del sitio web. Son, además, lo que tienen en promedio una mayor frecuencia de compras a distancia. Si se compara con los segmentos más multicanales C2, C4 y C5, este segmento tiene una mayor actividad y mayores montos promedios de compra en tienda. Representan a un 0,29% de la población.

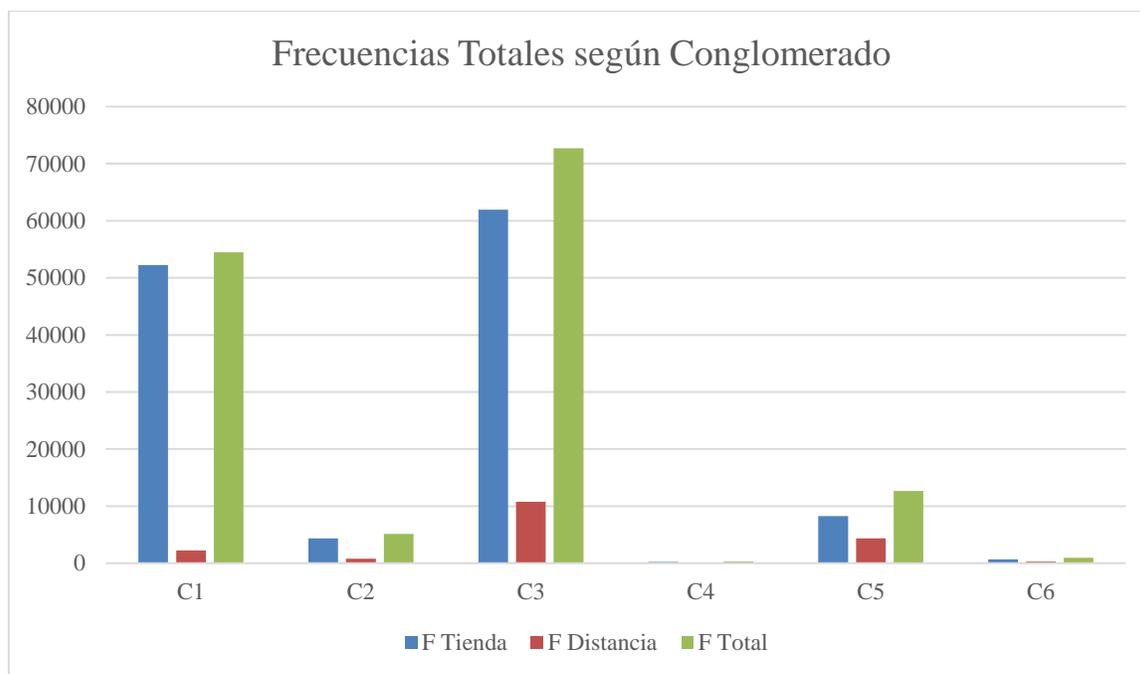


Gráfico 9.20: Frecuencias totales según conglomerado.

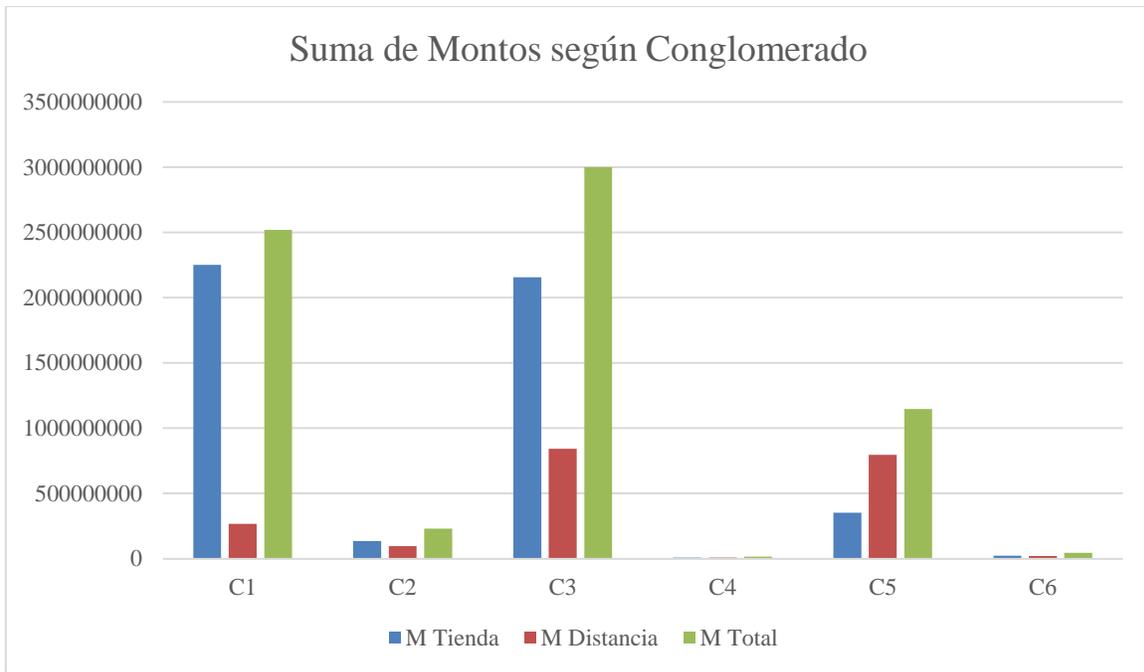


Gráfico 9.21: Montos según conglomerado.

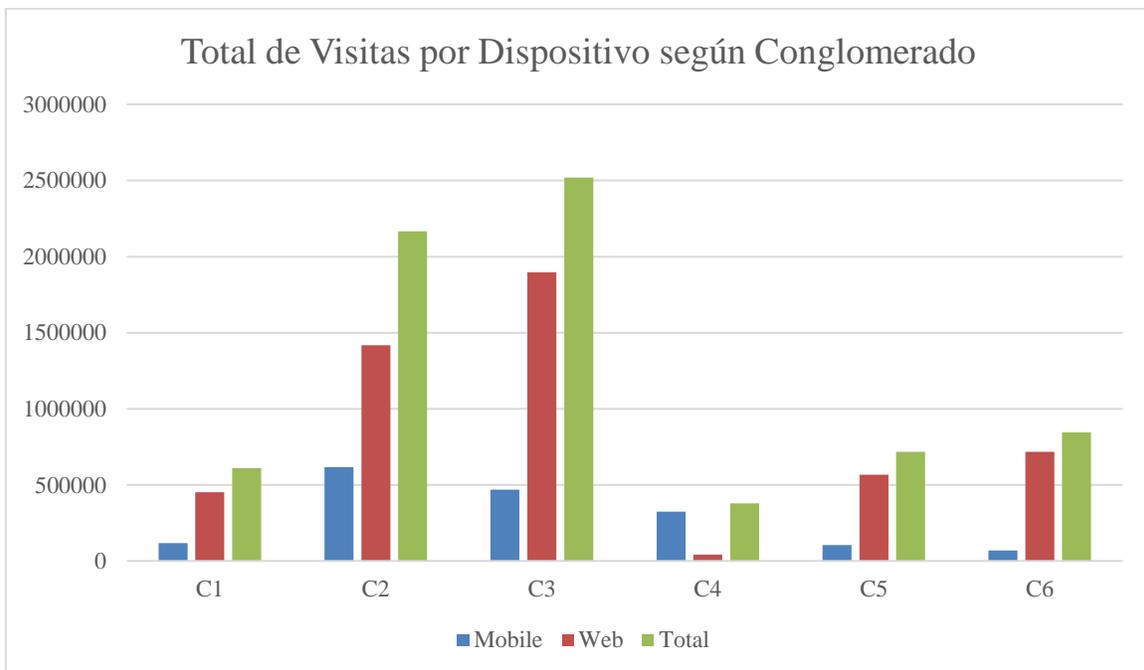


Gráfico 9.22: Visitas por dispositivo según conglomerado.

Se puede concluir que la mayor cantidad de compras están hechas por los clientes con índices más bajos de navegación. Más aún, más de la mitad de las compras a distancia están hechas por clientes que en promedio tienen menos de 1 compra al semestre.

7.1. Matrices de Transición

A continuación se presenta la matriz de transición entre estados.

		T+1		
		C1	C2	C3
T	C1	34.57%	41.11%	24.31%
	C2	7.03%	59.60%	33.37%
	C3	3.45%	30.75%	65.80%

Tabla 9.10. Matriz de transición entre estados

Recordar que:

- C1: Clientes Multicanal
- C2: Clientes Monocanal
- C3: Clientes Inactivos

Esta tabla muestra la transición entre los 3 segmentos de la segmentación que se seleccionó. Los segmentos más estables son los monocanal y los inactivos. Ambos tienen más del 50% de probabilidades de permanecer en el mismo estado en el período siguiente y aproximadamente 30% de probabilidades de pasar de un estado a otro (entre ellos). Dado esto, es poco probable que un cliente inactivo o monocanal se pase a tener actividad multicanal, siendo esta probabilidad menor para los clientes inactivos.

El segmento de clientes multicanal tiene más probabilidades de cambiar al monocanal que de permanecer en el mismo estado. Para este segmento, es más probable permanecer siendo multicanal que no tener actividad al semestre siguiente, ya que esta última opción corresponde al 24.31% (probabilidad que, de todos modos, es más alta que el 3.45% de que ocurra el evento inverso).

Se vio anteriormente que los clientes multicanal son, a nivel individual, mejores clientes que los monocanal, sin embargo es el segmento de clientes más volátil y con menos probabilidades de que nuevos clientes lo integren. Actualmente la empresa realiza descuentos masivos a través del sitio móvil con el fin de aumentar las ventas *online* (Cyberday y Cybermonday). De todos modos, para ellos no está claro que esto aumente la cantidad de clientes multicanal permanentes y no lo sean sólo para esa fecha.

10. Transición entre Canales

El último objetivo específico corresponde al análisis de la transición entre canales. Hay que recordar que los canales que en este estudio se consideran sólo los canales de venta de la empresa, es decir, la tienda física, tienda a través del sitio web, tienda a través del sitio móvil y venta telefónica.

En esta sección se estudiará cómo interactúa un cliente con más de un canal a la vez, más específicamente:

- a) Qué porcentaje de clientes compran el mismo día de navegación, un día después, dos días después y hasta una semana después (en general y en cada uno de los canales de navegación).
- b) Cuántas páginas visitan al día los clientes que compran y no compran el mismo día de navegación, un día después, dos días después y hasta una semana después (en general y en cada uno de los canales de navegación).
- c) Lo mismo de los dos puntos anteriores pero para ventas online cruzadas, es decir, navegación móvil y venta web y viceversa.
- d) Número de páginas visitas promedio de clientes que compran y no compran en tienda el mismo día de la navegación, un día después, dos días después y hasta dos semanas después.

Ventas online

Se utiliza como universo de clientes a todos los que tienen navegación los 6 meses de período de estudio.

Se calculó el promedio de páginas visitadas por los clientes que compraron ese mismo día en el canal de navegación el promedio de páginas visitadas por clientes que no compraron el mismo día de la navegación. Se hizo el cálculo en general y para cada uno de los canales de navegación.

Además se calculó también para las visitas que terminaron en compra al día siguiente, 2 días después, y así hasta 6 días después de la visita.

Las siguientes tablas contienen:

Compra: Número de páginas visitadas promedio al día personas que realizaron compra el mismo día o días después de navegar.

No compra: Número de páginas visitadas promedio al día personas que no realizaron compra el mismo día o días después de navegar.

N: Número de clientes que compra online el mismo día o días después de navegar (si un cliente navega más de un día, se contará la cantidad de días que haya navegado y comprado).

%: Porcentaje de clientes que compran online el mismo día o días después de navegar. El universo de este porcentaje corresponde al total de clientes que navegan (esta información está para cada día y para cada cliente).

General				
Día	Compra	No compra	N	%
0	77.97	22.88	9862	5.47%
1	61.81	25.54	2104	1.17%
2	20.29	26.00	567	0.31%
3	18.95	25.99	387	0.21%
4	17.42	25.99	277	0.15%
5	16.60	25.99	216	0.12%
6	17.24	25.99	185	0.10%

Tabla 10.1. Indicadores navegación General

Mobile				
Día	Compra	No compra	N	%
0	79.08	20.67	473	8.31%
1	36.11	21.18	159	2.79%
2	27.72	21.22	34	0.60%
3	13.93	21.22	21	0.37%
4	7.80	21.22	14	0.25%
5	11.57	21.22	10	0.18%
6	3.25	21.22	6	0.11%

Tabla 10.2. Indicadores Navegación Móvil.

Web				
Día	Compra	No compra	N	%
0	69.83	21.32	9053	6.97%
1	57.17	24.27	1951	1.50%
2	20.56	24.72	533	0.41%
3	20.44	24.71	367	0.28%
4	19.16	24.71	264	0.20%
5	18.28	24.71	207	0.16%
6	18.24	24.71	179	0.14%

Tabla 10.3. Indicadores Navegación Web

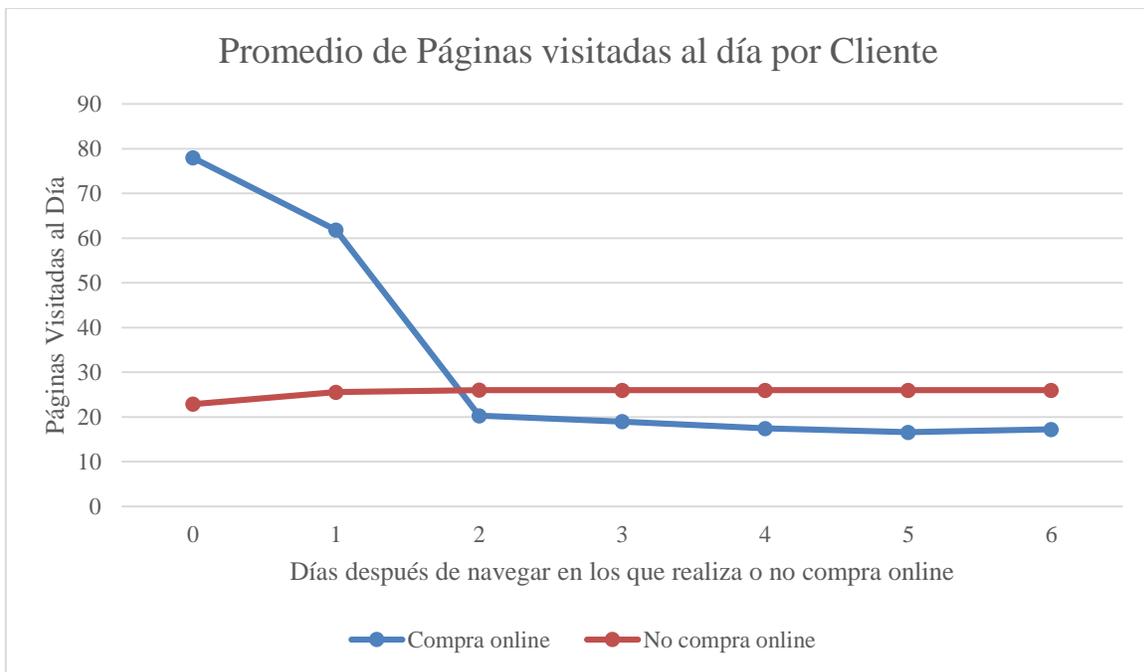


Gráfico 100.1

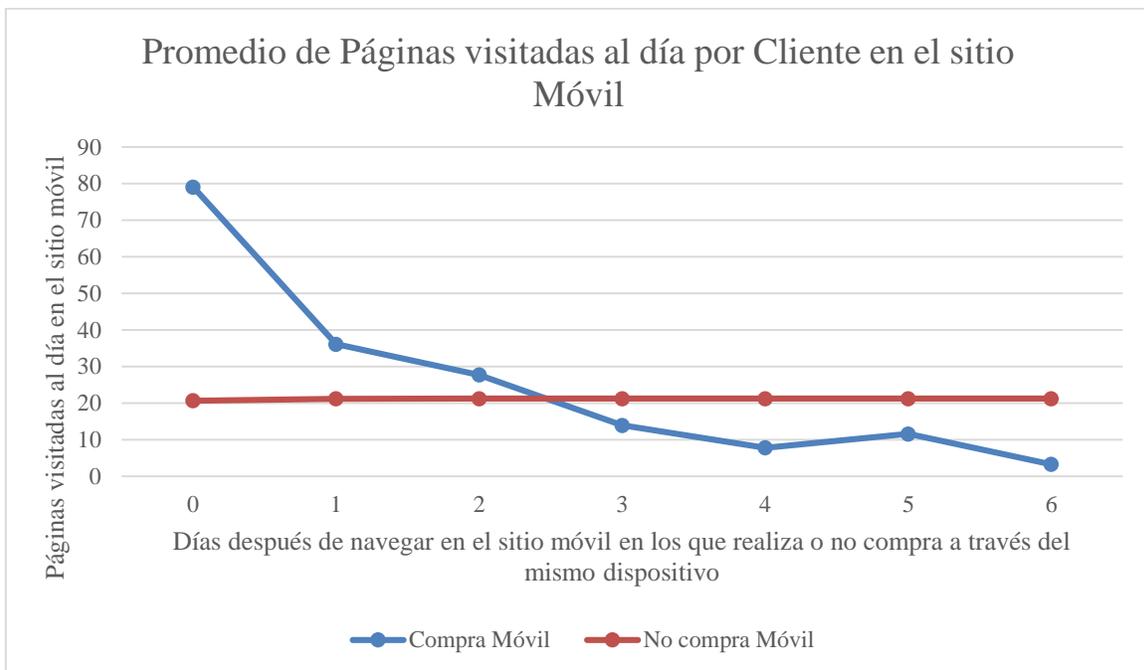


Gráfico 100.2

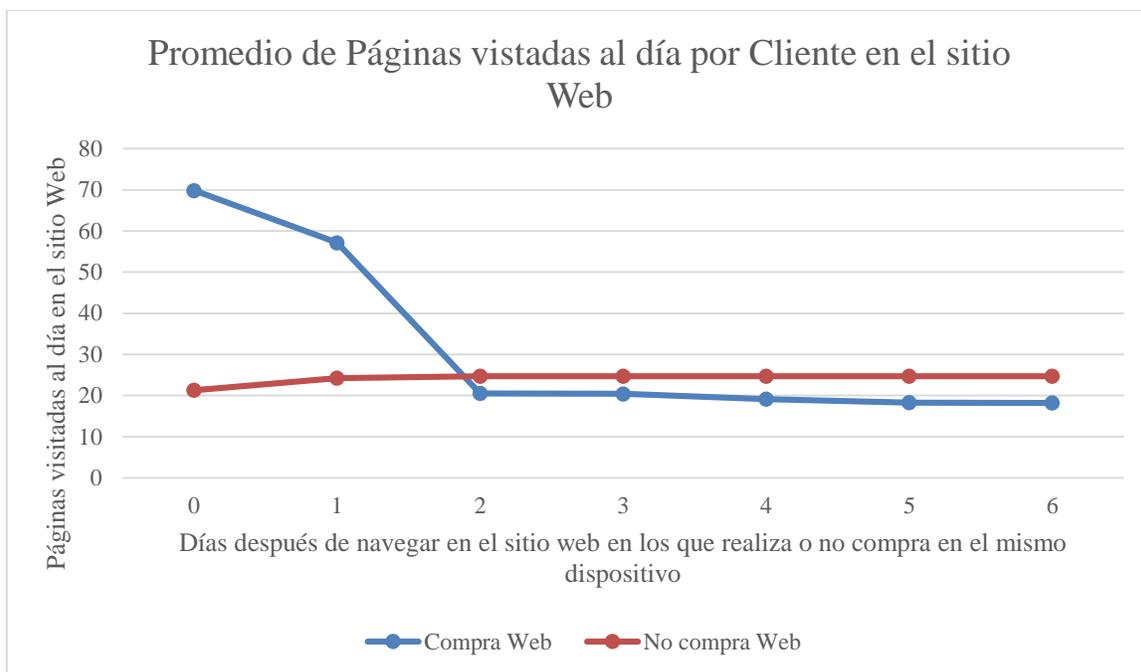


Gráfico 100.3

Se puede ver que los clientes de sitio móvil tienen una mayor cantidad de visitas a la página el día en que realizan una compra por este medio en comparación a los clientes web (cuando realizan compras a través del sitio web). En general y para la navegación web existe un aumento en las visitas cuando se va a realizar una compra el mismo día y 1 día antes de la compra, mientras que para el caso de las visitas el sitio móvil, se puede ver un aumento también para el mismo día, para 1 día antes y 2 días antes. El canal móvil es el que tiene mayor tasa de conversión por día, esto quiere decir que si un cliente visita una página en el sitio móvil es más probable que realice una compra en ese canal que si un cliente visita el sitio web.

En el canal web, si un cliente visita la página para comprar, lo hará más probablemente el mismo día o al día siguiente, mientras que para el sitio móvil, sigue siendo probable que lo haga al tercer día (esta probabilidad es baja, sin embargo, es más alta que para el resto de la semana). Esto se ve reflejado, además, en que se ve un aumento en la cantidad de páginas visitadas al día 2 días antes de la compra (comparado con el promedio de visitas al día de los clientes que no compran al tercer día).

Ventas cruzadas

Luego se hizo el cálculo para las ventas cruzadas de estos dos canales. Aquí se estudia la transición entre el canal web y el canal móvil. Los indicadores fueron los mismos que en la sección anterior, pero considerando las ventas en el otro canal de navegación. Con esto se busca encontrar si existe un canal que sea utilizado más para revisar productos o para comprarlos.

Navega Mobile Compra Web					Navega Web Compra Mobile				
Día	Compra	No Compra	N	%	Día	Compra	No Compra	N	%
0	50.07	20.71	876	15.4%	0	135.86	24.61	107	0.08%
1	35.12	21.10	432	7.6%	1	112.77	24.66	66	0.05%
2	13.36	21.23	89	1.6%	2	16.00	24.70	18	0.01%
3	8.12	21.23	43	0.8%	3	31.67	24.70	6	0.00%

Tabla 10.4

Se puede ver que el canal móvil presenta un 15,4% de conversión cruzada, es decir. Según la información del semestre de estudio, si un cliente inicia actividad de navegación en el sitio móvil, tiene un 15,4% de comprar ese mismo día en el canal web. Si a esa probabilidad se le suma la de venta en el mismo canal, se tiene que el 23,7% de las personas que visitan compran ese mismo día. Ese porcentaje es considerablemente más alto que el 7,05% correspondiente a las personas que compran luego de visitar el sitio web. Para el caso de este último canal, las navegaciones que finalizan en compra en el sitio móvil son sólo del 0.08%.

El canal móvil, entonces, comparándolo con el canal web, se utiliza más con la finalidad de comprar, ya sea por ese mismo medio o por el canal web. De este cálculo no se desprende ninguna causalidad, es decir no se sabe si las personas que ingresan al sitio móvil lo hacen sabiendo previamente que van a comprar, o si deciden hacerlo luego de navegar (pero lo hacen más probablemente en el sitio web por temas de seguridad) o si las personas que navegan por este canal son clientes que compran muy a menudo, a diferencia del canal web que es un medio más popular¹¹.

ROPO

La sigla ROPO (o ROBO) significa en inglés Research Online Purchase Offline. Este fenómeno considera que existe una previa visita a los medio online para luego comprar en la tienda física. En esta memoria se analizará desde dos puntos de vista: (1) considerando como universo a los clientes que navegan online y revisando cuántos de ellos compran en el corto plazo en tienda y (2) considerando a los clientes que compran en tienda como universo y calcular cuántos de ellos visitaron el sitio web o móvil los días previos.

Los gráficos 11.4 y 11.5 muestran el análisis que se hizo en la sección anterior pero considerando las ventas en tienda. Es decir, se calculó el promedio de páginas visitadas cada día por los clientes que realizan y no realizan compras después de navegar. El número de días que aparece en el eje horizontal indica cuántos días después de la navegación se hizo o no se hizo compra.

¹¹ Como se vio en el análisis descriptivo, existe un grupo de clientes que navega web pero sólo realiza compras en tienda.

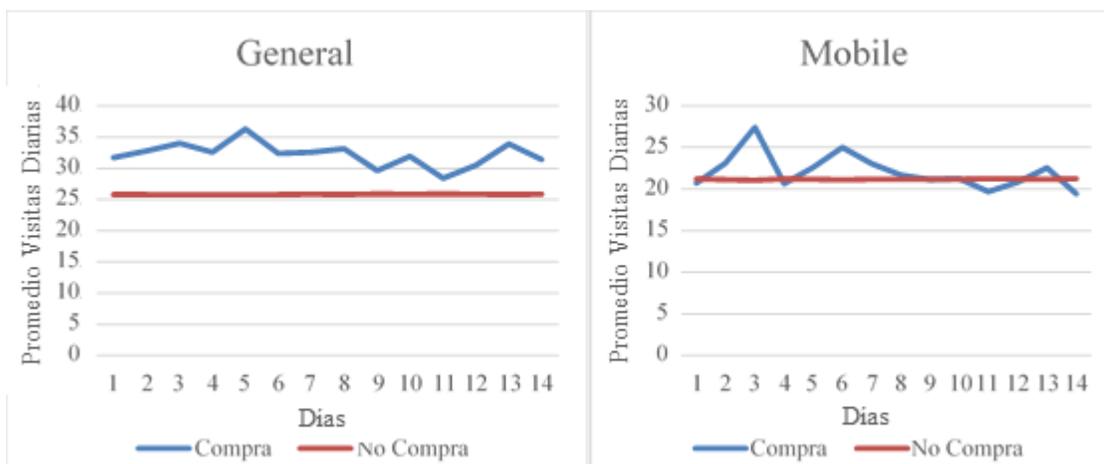


Gráfico 10.4

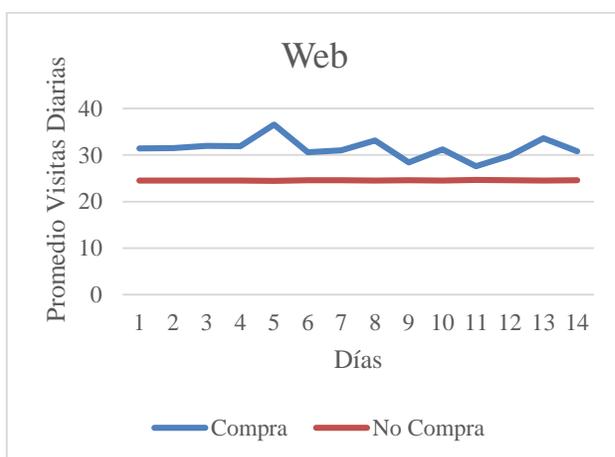


Gráfico 10.5

Como se puede ver en los gráficos 11.4 y 11.5, en el caso general y el caso web se puede ver que el promedio de visitas diario para los clientes que compran luego en tienda (el mismo día, 1 día después, 2 días después, etc.) es mayor al promedio de visitas diarias de las personas que no compran en tienda luego de navegar. Esto puede deberse al efecto ropo, es decir, que buscan productos en internet y luego los compran en tienda. De todos modos, no se ve una tendencia a la baja como podría esperarse (como en el caso de las compras online luego días después de la navegación), por lo que esta diferencia puede deberse a características propias de los clientes y no por causalidad. Es decir, puede ser que los clientes que compran en tienda a menudo, navegan en promedio más que los clientes que no compran en tienda con tanta frecuencia, es decir, que las compras que realizan los clientes luego de navegar podría deberse a la navegación como podría no serlo (por ejemplo clientes que compran con una frecuencia de 1 vez a la semana independiente de la navegación).

El siguiente gráfico muestra cómo evoluciona la cantidad de personas que compran en tienda luego de navegar según el dispositivo.

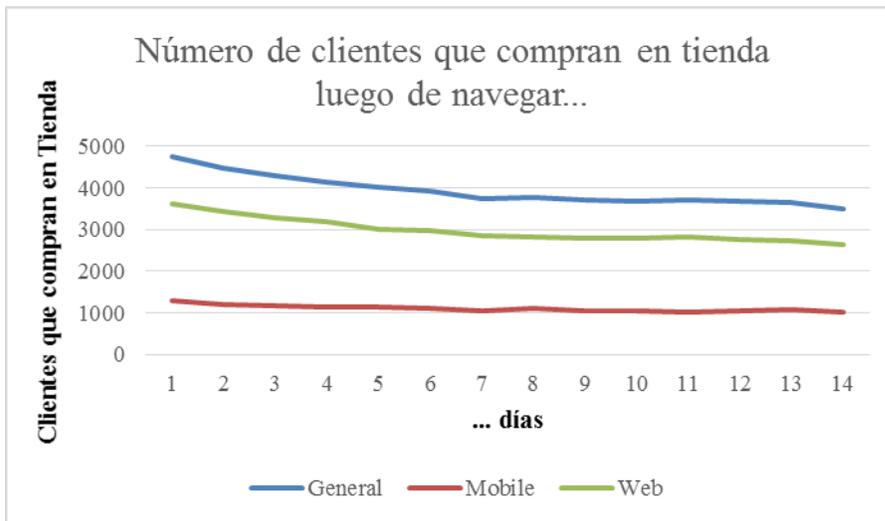


Gráfico 10.6: Número de clientes que compran en tienda luego de navegar.

Para hacer este cálculo, se tomó como universo todos los clientes que realizan navegación y se calculó la cantidad de personas que compran el mismo día, 1 día después de navegar, 2 días después y así consecutivamente hasta las 2 semanas después.

Aquí, en el caso general, caso web y levemente caso móvil puede apreciarse que el número de personas que compran en tienda el mismo día que navegan es mayor al número de clientes que compra al día siguiente de navegar. Este número decrece durante la primera semana y luego se estabiliza. Esta es una forma en la que se presenta el efecto ROPO, ya que del total de clientes que realiza navegación, un porcentaje de ellos compra en tienda el mismo día y otro los días que siguen.

En términos porcentuales la evolución es la siguiente:

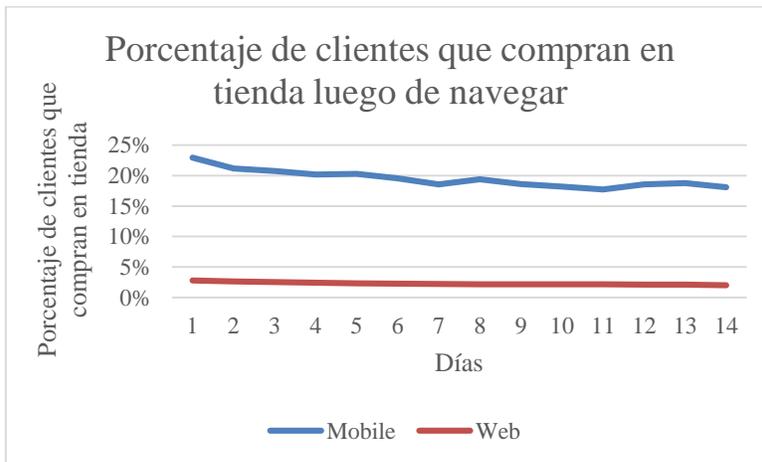


Gráfico 10.7: Porcentaje de clientes que compran en tienda luego de navegar.

Se puede ver que para el caso de la navegación móvil, la probabilidad de que el cliente compre en tienda luego de navegar bordea el 20% y es levemente decreciente a medida que aumenta la cantidad de días. Comparando con los resultados de las ventas *online*, se puede decir que si un cliente navega en el sitio móvil es más probable que realice una compra el mismo día en la tienda física que en el sitio web o móvil. De todos modos, el

porcentaje de clientes móviles que realiza compras el mismo día de navegación es alto comparado con los clientes que navegan a través del sitio web. Esto puede deberse a que los clientes ingresan al sitio móvil sabiendo que luego van a comprar (principalmente en tienda y a través del sitio web, pero en menor cantidad en el mismo sitio móvil) o puede ser debido a características propias de los clientes que utilizan este canal. No se presentan grandes diferencias en el porcentaje de clientes que compran luego de navegar por el sitio web a medida que pasan los días. Si bien de forma neta existen diferencias, al ser pocos clientes comparados con el total de clientes que navegan vía web, esta diferencia se hace despreciable.

Categorías

Para realizar un análisis del comportamiento a nivel de categoría, se debe considerar un nivel de agregación que permita observar patrones de comportamiento bien definidos. Estos grupos se definieron según características de los clientes que los compran.

- **Grupo 1:** *Electro Hogar, Blanco.*
- **Grupo 2:** *Accesorio Mujer, Calzado, Damas, Perfumería, Rincón Juvenil Damas, Ropa Interior, Niños.*
- **Grupo 3:** *Decoración, Maneje, Dormitorio, Muebles, Gourmet y Concesiones, Regalos*
- **Grupo 4:** *Hombres, Deporte, Rincón Juvenil Hombres.*

Para este cálculo se toma como universo el total de compras en tienda realizada por los clientes que tienen navegación durante los 6 meses de estudio. Incluir a todos los clientes que compran en tienda sería agregar más clientes que no van a navegar, por lo que no aportaría en analizar la evolución a través de los días.

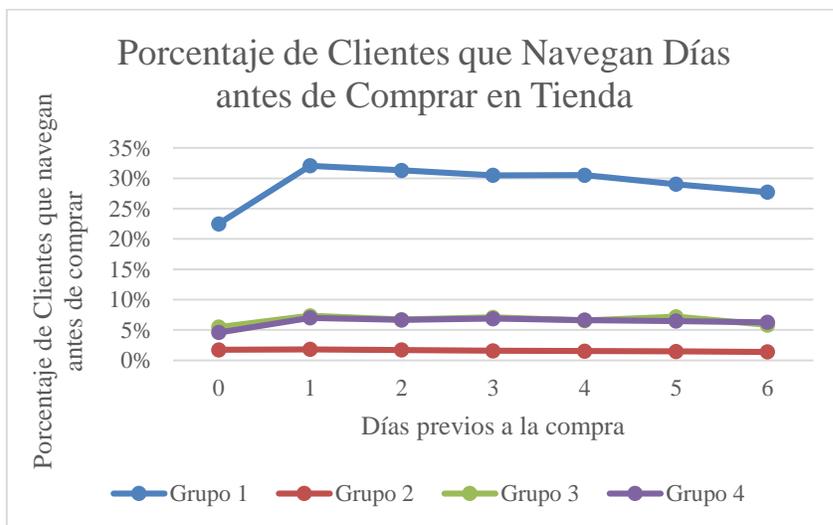


Gráfico 10.8: Porcentaje de clientes que navegan días antes de comprar en Tienda.

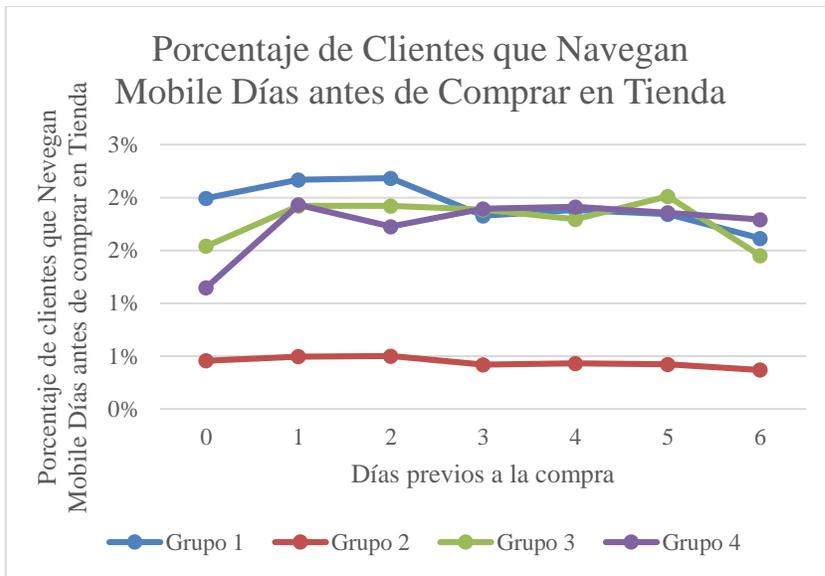


Gráfico 10.9: Porcentaje de clientes que navegan Mobile días antes de comprar en Tienda.

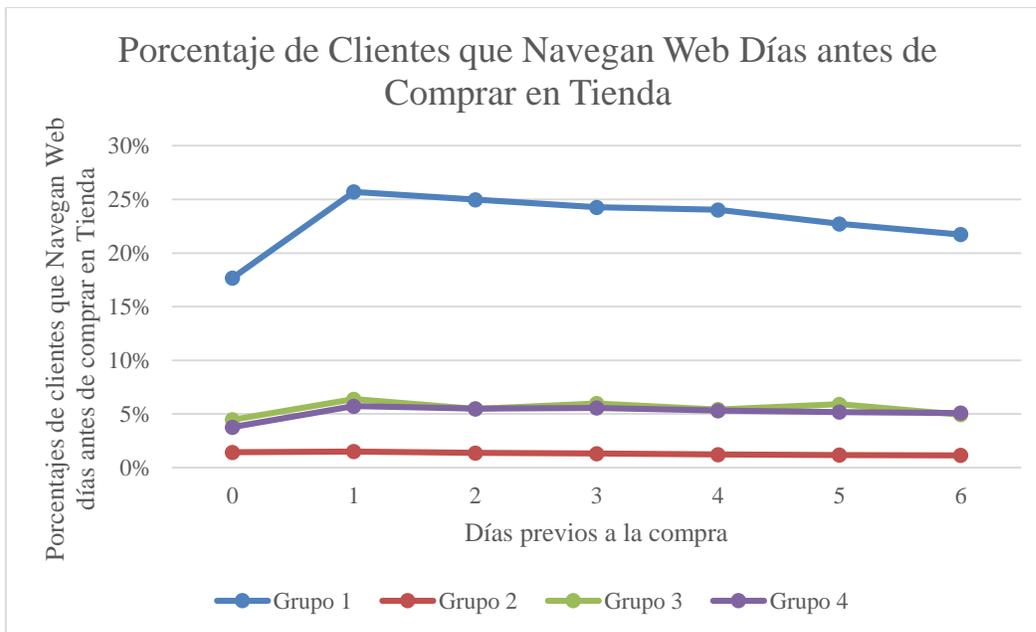


Gráfico 10.10: Porcentaje de clientes que navegan web días antes de comprar en tienda.

Se puede ver que en el Grupo 2 se ve una baja probabilidad de navegar dado que se compró en tienda. Además, esta probabilidad es pareja e independiente de los días hacia atrás de la compra. Esta categoría corresponde a los artículos que compran específicamente las mujeres. Para las otras categorías se puede ver que en el canal web y la suma de ambos canales hay una disminución en la probabilidad de navegar el mismo día de la compra. Es decir, dado que estoy comprando en tienda, la probabilidad de navegar ese mismo día es más baja en comparación a la probabilidad de navegar el día anterior. En el Grupo 1, que corresponde a los artículos de *Blanco* y *Electro Hogar* se puede ver que la probabilidad es mayor. Esto puede deberse a que son productos que más

comúnmente se revisan antes en internet. Una hipótesis es que se dé un fenómeno de sustitución, es decir, por ejemplo, una persona decide comprar una lavadora. Revisa y compara los productos en internet durante la semana y luego va a comprarla el fin de semana. Ese día finaliza la compra, por lo tanto no navega en internet. Se hizo el análisis quitando el período que está 2 días antes y 7 días después de *Cyber Monday* y 1 semana antes y una semana después de navidad. Las curvas tienen la misma forma, es decir, hay una disminución en el primer día, luego un aumento en el segundo día y la curva sigue decreciendo de forma leve.

11. Conclusiones y Recomendaciones

Este trabajo de memoria aborda el desafío de explicar el comportamiento de los clientes en un ambiente multicanal. A continuación se presentan las principales conclusiones de los resultados obtenidos.

Del análisis descriptivo de las variables se obtuvo información tanto del comportamiento de los clientes según sus propias características, como de las características de las transacciones. Del análisis de los clientes se observó que las mujeres cuentan con un mayor número de compras hechas tanto en tienda como a distancia, sin embargo esta diferencia es menor en las ventas a distancia. Las categorías que presentan mayor número de ventas desde el canal remoto son *Electro Hogar*, *Deportes* y *Dormitorio*, y éstas no presentan grandes diferencias entre clientes de sexo femenino o masculino. Por otro lado, las ventas a distancia están hechas principalmente por clientes de la región metropolitana, mientras que las compras en tienda tienen un mayor porcentaje de clientes de región.

Las variables RFM Multicanal presentan correlaciones mayores a 0,5 en términos netos Monto y Frecuencia en tienda (correlación positiva), Monto y Frecuencia a distancia (correlación positiva), Recency y Frecuencia en tienda (correlación negativa) y Recency y Frecuencia a distancia (correlación negativa). Este resultado era de esperarse, ya que al aumentar la frecuencia de compra en un canal, aumenta la suma de los montos de compra, mientras que al realizar compras más frecuentemente, aumenta disminuye el tiempo entre la última compra y el final del período de estudio.

Al realizar la segmentación se obtuvo tres principales segmentos: los clientes monocanal activos, clientes monocanal con poca actividad y los clientes multicanal. El último segmento presenta mayores frecuencias y montos de compra, sin embargo el grupo más pequeño de los tres. El aporte neto que hace a la compañía el grupo de clientes multicanal es bajo si se compara con la suma de los montos de cada uno de los otros segmentos, sin embargo estos clientes son más asequibles para la empresa, ya que se cuenta con más medio de contacto directo y personalizado con ellos (página web, celular, *email*). Actualmente, a pesar de que se pueda tener más influencia en este segmento, es el grupo más inestable de clientes. Si un cliente pertenece a este segmento en un período, es más probable que luego pertenezca a otro con actividad multicanal. Por otro lado, los clientes del grupo de clientes monocanal son los que tienen más probabilidades de pasar al segmento multicanal, por lo que identificarlos para centrar en ellos los esfuerzos de incentivar la multicanalidad podría ser más eficaz que incentivar a todos los clientes por igual.

Para tener un mejor acercamiento a la heterogeneidad de los clientes multicanal se propone incluir otro tipo de variables que no estaban disponibles para este trabajo. Como propone **Neslin et al. (2009)**, es necesario incluir variables no sólo de las compras en cada canal, sino también de uso, preferencia y el nivel de respuesta a cada uno. Para esto es necesario que esta información esté previamente medida en indicadores.

En el estudio del aumento de las vistas de páginas diarias antes de una compra en el mismo canal, se pudo observar en las compras online un significativo aumento en el promedio diario de páginas vistas antes de una compra. Para el caso de la navegación

móvil este aumento se presenta el mismo día de la compra, 1 día antes y 2 días antes. Para el caso general y web el aumento se presenta sólo el mismo día de la compra y 1 día antes. Este resultado podría aplicarse para identificar cuándo un cliente se acerca a una compra ya que se presenta un aumento en las páginas visitadas durante el día. De todos modos este es un análisis general, por lo que identificar la intención de compra debería incluir otras métricas que caractericen mejor a cada cliente en particular.

En el caso de las ventas online cruzadas se puede ver que las navegaciones móviles presentan una mayor tasa de conversión a venta web que en su mismo canal, mientras que para las visitas web no se presenta una probabilidad de compra móvil importante. Se puede concluir que los clientes utilizan el dispositivo móvil más para comprar en web que por el mismo dispositivo y que un cliente que navega a través del sitio móvil tiene más probabilidades de comprar que un cliente que navega a través del sitio web. Si este fenómeno se sigue presentando durante el tiempo con la aplicación móvil (al momento de hacerse este estudio la aplicación había sido recientemente lanzada, por lo que no se contaba con registros de información significativa), podría aceptarse y decidir canalizar los recursos para promover a través de este canal la venta web y considerar este efecto para medir el impacto de la aplicación.

En el estudio del efecto ROPO, se puede ver un aumento en el número de visitas las 2 semanas antes de una compra en tienda, mientras que para los clientes móviles no existe una tendencia clara. Esta característica puede ser propia de los clientes que compran en tienda y navegan, ya que no se presenta una tendencia a la baja durante el tiempo. Queda como trabajo propuesto extender el período de estudio para analizar el efecto real que tenga la venta en tienda en la navegación web. Por otro lado, se puede ver que para los clientes que compran en tienda existe una mayor probabilidad de haber navegado el día anterior que a haber navegado el mismo día en las categorías *Electro Hogar y Blanco*. En las otras categorías no se ve una diferencia pronunciada. En la navegación móvil se puede ver el mismo fenómeno para las categorías de público masculino: *Hombres, Deporte y Rincón Juvenil Hombres*. Esto puede deberse a un efecto de sustitución, ya que los clientes dejan de visitar la página al visitar la tienda física. Para el caso de los clientes que navegan, existe una mayor probabilidad de comprar en tienda el mismo día que al día siguiente. Esta probabilidad disminuye a medida que aumenta el horizonte de días.

Una propuesta de trabajo a futuro es investigar la causalidad de la relación entre la multicanalidad de los clientes y el aumento en los montos y frecuencias de compra. Si bien se pueden hacer esfuerzos para aumentar el número de clientes con actividad multicanal, no se sabe si esto impacta directamente de forma positiva en aumentar significativamente la frecuencia y los montos de venta de forma prolongada en el tiempo. Por otro lado, si bien se conocen los efectos positivos de las campañas de *email marketing*, sería interesante estudiar si aumentar el número de clientes que recibe este tipo de promociones aumenta el número de compras de estos clientes.

Para hacer un estudio más completo del comportamiento multicanal se necesita tener información que actualmente no está al alcance de la empresa, como lo es la información de las visitas en tienda que no finalizan en compra. Esta información se podría obtener mediante el uso de alguna aplicación móvil que ofrezca algún beneficio al registrarse al entrar a la tienda. Otra información que no está disponible actualmente es la relacionada con el efecto ROPO. Cierta número de clientes revisa los productos y luego los compra en la tienda física. Se podría estudiar si este efecto es positivo o negativo para la compañía

midiendo si el cliente realiza otras compras al estar en la tienda (en este caso tendría un efecto positivo, ya que el ticket final es mayor al que si hubiese comprado en tienda) o si hubiese sido preferible que finalizara el proceso de compra desde el canal remoto. De esta forma se podrían buscar mecanismos para incentivar la visita a la tienda luego de la navegación o, en el caso contrario, aumentar las visitas al sitio web que finalizan en compra en el mismo canal. Esto mismo aplica para la transición entre navegación móvil y compra web. Se propone analizar la actividad de navegación de los clientes de forma desagregada para conocer en qué tipos de producto se presenta algún tipo de transición a otro canal.

12. Bibliografía

- Adobe. (2011). *Multichannel Marketing: Creating a Competitive Advantage in Today's Complex Marketing Landscape*.
- Balasubramanian, S., R. R. (2005). *Consumer in an Multichannel Environment: Product Utility, Process Utility and Channel Choice*. Journal of Interactive Marketing, 19 (2), 12 - 30.
- Ching-Hsue Cheng, Y.-S. C. (2009). *Classifying the segmentation of customer value via RFM model and RS theory*. Expert System with Applications .
- Hughes. (1994). *Strategic database marketing*.
- J. M. C. Schijns, G. J. (1996). *Segment selection by relationship strength*. Journal of Direct Marketing.
- Jing Wu, Z. L. (2005). *Research on customer segmentation model by clustering*. Proceedings of the 7th international conference on electronic commerce.
- Neslin, Scott A., D. Grewal, R. Leghorn, V. Shankar, M. L. Teerling, J. S. Thomas, P. C. Verhoef. (2006). *Challenges and Opportunities in Multichannel Management*. Journal of Service Research, 9, 2, 95 - 113.
- Neslin, Scott A., Shankar V. (2009). *Key Issues in Multichannel Customer Management: Current Knowledge and Future Directions*. Journal of Interactive Marketing, 23, 70 - 81.
- Rojas, A. (2014). *Estudio Experimental de Automatización de Email Marketing en un Retail Online, Memoria Ingeniería Civil Industrial, Universidad de Chile, 2014*.
- Uptal M. Dholakia, B. E. (2010). *Consumer Behavior in a Multichannel, Multimedia Retailing Environment*. Journal of Interactive Marketing, 24 (2), 86 - 95.
- Venkatesan R., K. V. (2007). *Multichannel Shopping: Causes and Consequences*. Journal of Marketing, 71 (2), 114 - 132.
- Zahng J., F. P. (2010). *Crafting Integrated Multichannel Retailing Strategies*. Journal of Interactive Marketing, 24 (2), 168 - 180.