



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

EVALUACIÓN Y MEJORA DE UN MOTOR DE RECOMENDACIONES APLICADO A  
UN PROVEEDOR DE CONSUMO MASIVO EN CANAL TRADICIONAL

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

IGNACIO JAVIER DÍAZ ROSALES

PROFESOR GUÍA:  
LUIS ABURTO LAFOURCADE

MIEMBROS DE LA COMISIÓN:  
ALEJANDRA PUENTE CHANDIA  
CLAUDIO PIZARRO TORRES

SANTIAGO DE CHILE  
2017

RESUMEN DE LA MEMORIA PARA OPTAR AL  
TÍTULO DE: Ingeniero Civil Industrial  
POR: Ignacio Javier Díaz Rosales  
FECHA: 24/04/2017  
PROFESOR GUÍA: Luis Aburto Lafourcade

## **EVALUACIÓN Y MEJORA DE UN MOTOR DE RECOMENDACIONES APLICADO A UN PROVEEDOR DE CONSUMO MASIVO EN CANAL TRADICIONAL**

El presente trabajo muestra el estudio y desarrollo de una nueva metodología para el cálculo de las ventas incrementales que genera un motor de recomendaciones utilizado por un proveedor de productos de consumo masivo, con tal de resolver los distintos problemas que presenta la actual metodología. Otra línea del presente trabajo es detectar posibles mejoras en esta herramienta de recomendaciones a través del desarrollo de un nuevo modelo de recomendación y experimentos que permitan confirmar ciertas hipótesis relacionadas principalmente a la distribución de las distintas estrategias que se utilizan.

Los métodos dentro de este trabajo utilizan los datos transaccionales históricos de la empresa fabricante. En el caso del análisis de venta incremental se consideraron datos desde el periodo en que se implementó este motor de recomendaciones, mientras que para el desarrollo de una nueva estrategia de recomendación se utilizaron los datos transaccionales de los últimos 6 meses.

En relación al problema del cálculo de ventas incrementales, se utilizó una metodología basada en el análisis descriptivo de los datos de venta históricos. Donde se estudió a profundidad la evolución que han tenido las ventas incrementales a lo largo del tiempo y como éstas se relacionan con los distintos indicadores utilizados en su cálculo. Lográndose encontrar factores que influyen a estos indicadores de manera tal, que llevan a realizar cálculos erróneos que provocan una subestimación del valor que entrega la herramienta en términos de venta incremental en al menos USD 6.000.0000. Proponiendo seguido de esto una metodología que permita magnificar el valor real de este motor de recomendaciones.

En relación al desarrollo de una nueva estrategia de recomendación se utilizó un modelo de filtros colaborativos, que permite generar recomendaciones más acertadas en función de las preferencias de cada uno de los clientes. Preferencias caracterizadas en función del historial de compra de cada cliente.

Finalmente para validar el impacto del nuevo modelo de recomendación y las hipótesis acerca de las distribuciones de estrategias, se realizan experimentos de campo con grupos de control y experimentales definidos. Teniendo como principales resultados que la nueva estrategia de filtros colaborativos no presenta diferencias significativas respecto a la estrategia a reemplazar, que la posición en la que se encuentra la recomendación dentro de la pantalla de recomendaciones afecta de manera directa a la efectividad que la recomendación tendrá y en última instancia que el número de recomendaciones por estrategia no tiene un efecto significativo sobre la efectividad de esa estrategia.

## **Agradecimientos**

Quiero agradecer a mis padres Raúl y Marcela por su apoyo incondicional en todo aspecto. Siempre han estado en los buenos y malos momentos, por sus consejos y el cariño constante que me brindan día a día. Tampoco puede quedar fuera mi hermano Emilio a quien quiero agradecer siempre su apoyo y la confianza que deposita en mí.

Además quiero agradecer a mi polola Daniela que ha sido un pilar importantísimo en mi vida, sé que siempre puedo contar contigo y créeme que en esta etapa has sido fundamental. Eres mi principal motivación para seguir adelante y para comenzar con mucha energía esta nueva etapa que comienza.

Por otro lado estoy feliz de todos los amigos y personas que me he encontrado en este largo camino, conocer gente tan diversa y llena de alegría definitivamente ha valido la pena. Nunca olvidaré todos los momentos a lo largo de estos años, desde los juegos de taca-taca en primer año hasta los últimos pasajes de la carrera con las reuniones en casas para avanzar en tareas o estudiar para los exámenes que ya se aproximaban.

Por supuesto no puedo dejar fuera a Penta Analytics, empresa que me acogió el último año con un gran equipo de profesionales y sobretodo personas. De verdad que me han enseñado mucho en este tiempo y me motivan cada día a aprender más aún.

Finalmente quiero agradecer a mi profesor guía Luis Aburto, de verdad que el trato siempre tan amable y la paciencia que tuvo conmigo me ayudaron muchísimo en momentos donde no encontraba la motivación y la frustración era un peligro latente. Es definitivamente uno de los mejores profesores que he conocido tanto en la parte humana como académica.

## Tabla de Contenido

1.	Introducción.....	1
1.1	Antecedentes Generales .....	1
1.2	Descripción y justificación del proyecto .....	2
1.3	Objetivos .....	9
1.3.1.	Objetivo General .....	9
1.3.2.	Objetivos Específicos .....	9
1.4	Alcances .....	9
1.5	Resultados Esperados .....	9
2.	Marco Conceptual.....	10
2.1.	Motores de recomendación .....	10
2.2.	Tipos de motores de recomendación .....	10
2.2.1.	Algoritmos de Filtros colaborativos.....	10
3.	Metodología .....	12
3.1.	Venta Incremental .....	12
3.1.1.	Análisis exploratorio.....	13
3.1.2.	Propuesta de cálculo venta incremental.....	13
3.2.	Efectividad de recomendaciones .....	13
3.2.1.	Análisis Descriptivo .....	13
3.2.2.	Selección, limpieza y transformación de datos.....	13
3.2.3.	Generación de modelo.....	13
3.2.4.	Diseño Experimental.....	13
3.2.5.	Evaluación de Resultados.....	14
4.	Desarrollo Metodológico.....	14
4.1.	Venta Incremental .....	14
4.1.1.	Análisis Exploratorio .....	14
4.1.2.	Propuesta Cálculo de Venta Incremental.....	16
4.2.	Efectividad de estrategias.....	21
4.2.1.	Análisis Descriptivo .....	21
4.2.2.	Selección, limpieza y transformación de datos.....	24
4.2.3.	Generación de modelo.....	27
4.2.4.	Diseño Experimental.....	28
4.2.5.	Evaluación de resultados .....	35
5.	Conclusiones .....	43
6.	Recomendaciones y Trabajo Futuro.....	44
7.	Bibliografía.....	45

<b>8.</b>	<b>Anexos.....</b>	<b>47</b>
	<b>Anexo A: Análisis Descriptivo de Clientes Alicorp .....</b>	<b>47</b>
	<b>Anexo B: Análisis Descriptivo de Productos y Vendedores .....</b>	<b>49</b>
	<b>Anexo C: Análisis de Efectividad de Estrategias .....</b>	<b>50</b>
	<b>Anexo D: Metodologías Alternativas .....</b>	<b>52</b>
	<b>Anexo E: Criterios de Similitud .....</b>	<b>54</b>
	<b>Anexo F: Cantidad de Productos Similares .....</b>	<b>55</b>
	<b>Anexo G: Test ANOVA de Grupos Control y Tratamiento .....</b>	<b>56</b>
	<b>Anexo H: Test ANOVA de Experimento 1 .....</b>	<b>56</b>
	<b>Anexo I: Test ANOVA de Experimento 2 .....</b>	<b>58</b>
	<b>Anexo J: Test ANOVA de Experimento 3 .....</b>	<b>59</b>

# 1. Introducción

## 1.1 Antecedentes Generales

Alicorp S.A.A es una empresa fabricante de bienes de consumo masivo que se localiza en Perú, la cual cuenta con participación en gran parte del continente americano, siendo sus principales lugares de exportación Chile, Ecuador, Argentina, Brasil y Bolivia. Es considerada la más grande en su rubro dentro del mercado peruano contando con productos pertenecientes a 16 categorías distintas siendo, en su mayoría, dominadoras del mercado (tabla N°1). Donde sus principales competidores en Perú corresponden a Mondelez, Unilever, Nestlé, Procter & Gamble y Carozzi. A nivel de ventas, es considerada una de las empresas más grandes de Perú, ascendiendo sus ventas totales a US 1.753 MM en 2014.

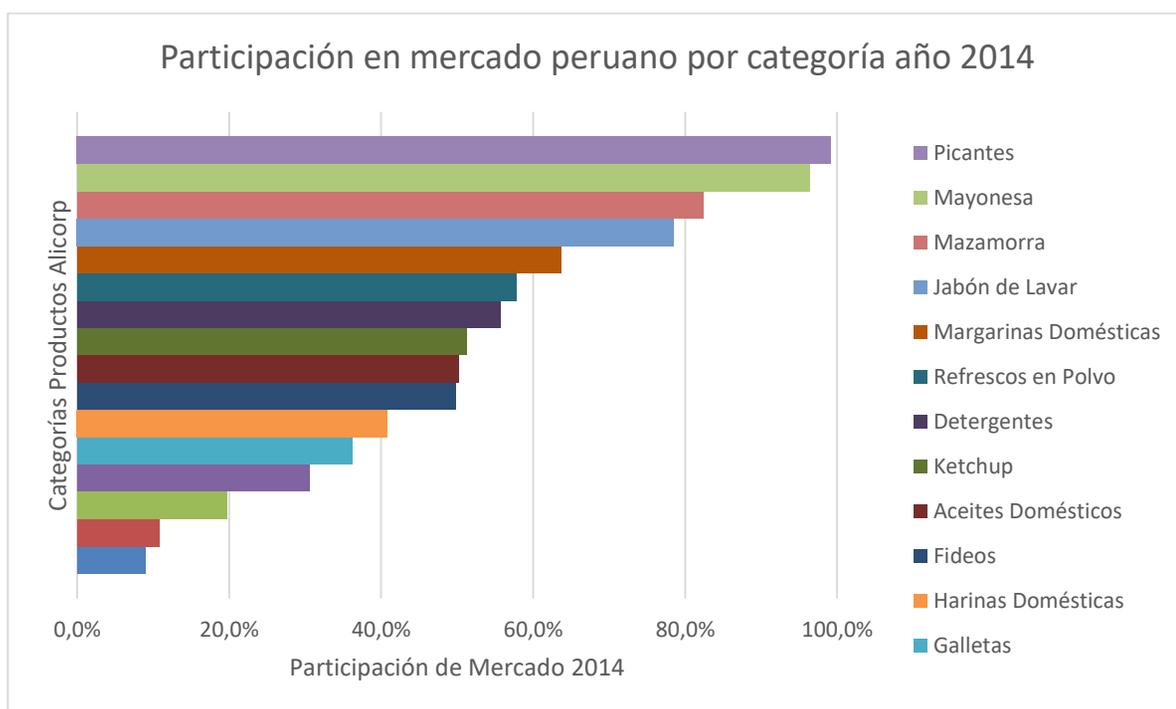


Gráfico 1. Participación de Mercado por categorías de Alicorp en 2014. Elaboración Propia

Al ser una empresa fabricante de bienes de consumo, la industria del Retail es de vital importancia para Alicorp. Siendo principalmente dos los canales de comercialización importantes para este tipo de empresa, primero se encuentra el denominado canal moderno que corresponde a los operadores grandes, es decir, supermercados o centros de distribución mayorista. Y por otro lado, está el canal tradicional donde se encuentran los operadores pequeños como mercados, bodegas, almacenes o mercadillos siendo característico para ellos operar con un canal de distribución directo, es decir, con una fuerza de venta (propia o externa) que llega a ofrecer directamente los productos a este tipo de locales.

La realidad de estos canales a nivel latinoamericano es muy diversa en cuanto a proporción, por ejemplo en el caso de Chile se tiene que el canal moderno corresponde a casi un 80% del mercado chileno y un 20% al tradicional. Sin embargo en Perú, el canal tradicional es el que cuenta con mayor presencia alcanzando casi el 80% del mercado, en desmedro del canal moderno.

Es por esto que el canal tradicional cobra tanta relevancia para Alicorp, siendo este el motivo por el que ha intentado innovar con soluciones que le permitan primero entender el comportamiento de sus clientes en este canal y en base a esto generar acciones que logren impactar de forma directa a sus indicadores críticos de venta. Es en esta etapa que entra Penta Analytics S.A, una empresa chilena con más de 15 años de experiencia en el desarrollo de soluciones analíticas en distintos países de Latinoamérica. Siendo responsables de la implementación y soporte del denominado Pedido Sugerido, un motor de recomendaciones de productos que busca optimizar el proceso de venta entre los vendedores del canal tradicional, dado que existe una problemática relacionada al reducido tiempo de venta efectiva que el vendedor tiene con cada cliente. Lo que dada la gran cantidad de SKU's que se tienen disponible para ofrecer y la alta cantidad de clientes que el vendedor tiene en ruta cada día, no permite aprovechar al máximo la oportunidad de venta.

En este proyecto se busca primero determinar el real aporte que genera, en términos de ventas incremental, el uso de esta herramienta por parte de los vendedores de Alicorp. Y segundo, se busca encontrar mejoras que pudieran implementarse en el Pedido Sugerido para que este sea más efectivo, en términos de aumentar el número de productos recomendados que sean comprados por el cliente.

## **1.2 Descripción y justificación del proyecto**

Al encontrarse en una industria altamente competitiva, las empresas deben encontrar formas de diferenciación. Para Alicorp esta forma es competir en base a precio, innovación y calidad de producto, reconocimiento de marca, eficacia del marketing y capacidad para identificar y satisfacer las necesidades del consumidor. Siendo ésta última estrategia, la responsable de la incorporación del motor de recomendaciones en las operaciones de la empresa.

Actualmente este motor de recomendaciones o Pedido Sugerido, es esencial en las operaciones de la empresa en el canal tradicional. Estando activo para los cerca de 1000 vendedores que la empresa utiliza para garantizar la cobertura de este canal, siendo estos organizados por medio de distintas distribuidoras exclusivas que se encargan de cubrir diferentes zonas asignadas.

El proceso de venta que siguen todos los vendedores, no importando la distribuidora de origen, se realiza de igual manera para todos los tipos de tienda (kiosko, almacenero, etc) y para todos los tipos de clientes. Siendo este descrito a continuación en la siguiente ilustración:

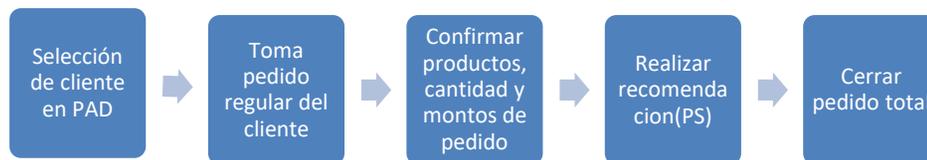


Ilustración 1. Proceso de Venta Canal Tradicional. (Elaboración Propia)

Todo este proceso es operado a través de un dispositivo PALM, ubicado a la izquierda de la ilustración, que permite realizar la carga directa de cada pedido al sistema y a su vez incorporar ciertas etiquetas que facilitan la posterior evaluación del Pedido Sugerido. Se puede apreciar por la ilustración, que el proceso de recomendación se realiza al finalizar el pedido regular del cliente. Acotando que este se realiza en una nueva pantalla dentro del dispositivo y que por construcción sólo recomienda productos que están fuera de ese pedido actual o regular que realizó el cliente.

Básicamente este motor de recomendaciones funciona en base a cuatro tipos de estrategias de recomendación, las cuales tienen un objetivo comercial claro y están disponibles para el vendedor a la hora de tomar el pedido del cliente:

- **Frecuentes:** Se basa en recomendar al cliente productos que están siempre en su canasta habitual, pero que en el pedido actual no está presente. Siendo un recordatorio para el cliente que tiene como objetivo mantener su valor actual, en términos de venta.
- **Asociados:** Son recomendaciones que consideran reglas de asociación entre los productos de la canasta actual con las transacciones conjuntas del perfil. Siendo su objetivo principal incrementar el mix de productos que el cliente compra de manera habitual.
- **Perfil:** Esta estrategia recomienda productos que se encuentran en la canasta base del perfil al cual pertenece el cliente y que actualmente él cliente no éste llevando en su pedido actual. Siendo el objetivo principal de esta estrategia aumentar el mix de productos del cliente y empujarlo a alcanzar el valor máximo del perfil al que está asociado.
- **Zona Caliente:** Son recomendaciones de productos en base a la zona donde el cliente se encuentra, siendo los productos recomendados los que presentan un mejor desempeño comercial (frecuencia, presencia, ventas) en dicha zona. El principal objetivo de esta estrategia es aumentar el valor del cliente.

En la operación actual, la evaluación de este motor se basa en el cálculo de ventas incrementales, siendo el objetivo de este tipo de cálculo determinar la diferencia en ventas entre el período actual y el período anterior. En el caso de Alicorp, lo relevante es determinar cuánto es el incremento en ventas entregado por la herramienta de

recomendaciones, dado que esto permite realizar una evaluación concreta del valor y la rentabilidad que ésta entrega.

Para el cálculo de la venta incremental se utiliza la siguiente fórmula:

$$\text{Venta Incremental} = \text{Venta PS Período anterior} * \text{Crecimiento Real}$$

Dependiendo esta fórmula de los siguientes cálculos previos:

$$\text{Crecimiento Real} = \text{Variación PS} - \text{Variación Mercado}$$

$$\text{Variación Mercado} = \frac{\text{Sum Venta S/PS Actual} - \text{Sum Venta S/PS Anterior}}{\text{Sum Venta S/PS Anterior}}$$

$$\text{Variación PS} = \frac{\text{Sum Venta PS Actual} - \text{Sum Venta PS Anterior}}{\text{Sum Venta PS Anterior}}$$

Dada esta estructura, el cálculo de la venta incremental requiere de la conformación de dos grupos. Uno compuesto por todos los clientes quienes realizaron compras producto de las recomendaciones ofrecidas y otro conformado por los clientes que realizaron compras sin aceptar ninguna de las recomendaciones ofrecidas. Más formalmente esto es:

**Grupo  $PS_i$** : Compro al menos un producto recomendado en el período  $i$

**Grupo  $S/PS_i$** : No compro alguno de los productos recomendados en el período  $i$

En términos simples, lo que la actual metodología intenta hacer es crear un valor de referencia que refleje cuanto sería el aumento en ventas si es que no se hiciera uso de la herramienta de recomendaciones, siendo este valor el indicador denominado Variación Mercado. Esto con el objetivo de que al compararlo con el indicador Variación PS, se pueda aislar el crecimiento inherente al mercado y así lograr obtener el valor de la venta incremental que está asociado netamente al uso del motor de recomendaciones.

Sin embargo, producto de la composición de los grupos y de la fórmula en sí, existen diversos problemas que pueden surgir a nivel de cálculo o de resultados. Los cuales se señalan a continuación:

- A nivel de cálculo pueden existir algunas situaciones que lleven a tener resultados erróneos o que no reflejan el aporte real en las ventas incrementales de las recomendaciones. En la siguiente tabla se indican las con mayor relevancia:

Casos	N° de casos	Probabilidad	Problemas
Caso 1 No existe venta con PS en período anterior	301	6,6%	Al no existir venta anterior no se puede calcular la venta incremental
Caso 2 Crecimiento Real <0 & Monto anterior PS<0	2	0,04%	Se muestra que existe venta incremental cuando realmente esta debiera ser negativa
Caso 3 Variación PS<0 & Variación Mercado<0	141	3,09%	Si $ Variación Mercado  >  Variación PS $ , se obtiene que existe crecimiento real positivo y por ende que hay venta incremental.
Caso 4 Variación PS >0 & Variación Mercado <0	1.070	24%	Se potencia aún más el crecimiento real, cuando esto puede llevar a problemas del grupo de control a considerar.

Tabla 1. Casos con problemas de medición. Elaboración propia.

- Existen problemas debido a que no se consideran factores de heterogeneidad que puedan estar afectando a este modelo, los que surgen esencialmente debido a que esta comparación se realiza con los mismos clientes. En definitiva, esta fórmula de cálculo no permite diferenciar un aumento en las ventas por las recomendaciones que se le hicieron, de un aumento por razones inherentes al comportamiento del cliente.
- Otros comportamientos o efectos que podrían no estar considerándose en esta evaluación, es por ejemplo un aumento en la canasta de compra por parte del cliente debido a razones macroeconómicas. Podría ser que la economía vaya en alza y por ende las personas estén comprando más productos, y por ende a su vez el cliente necesite comprar más productos ante la mayor demanda. En el caso de la evaluación actual esta alza siempre se le atribuye al sistema de recomendación. Otro caso que podría aplicarse en esta línea es que el comportamiento de compra del cliente haya cambiado debido a que tiene más deudas que pagar (educación, salud, etc.) o ha recibido alguna especie de ayuda por parte del gobierno (capitales semillas, bono, etc.), considerando que esta baja o alza en las ventas se deben netamente a efecto de motor de recomendaciones en caso de que se haya aceptado alguna recomendación.

- La evaluación tampoco considera la evolución que pudiera tener el cliente con respecto al negocio, es decir, quizás el comportamiento de compra del cliente cambie debido a que tiene deseos de diversificarse o expandir el negocio, no así porque se le hayan recomendado productos.
- En otra línea las estructuras de los grupos de control podrían estar generando graves problemas de estimación, debido a que es posible que se estén comparando clientes muy distintos entre sí. Esto dado que la distinción que se hace para construir cada grupo considera sólo una variable de comportamiento, no de caracterización de los clientes. Por ejemplo, con la estructura de los grupos actual puede ocurrir el caso extremo donde se comparen los clientes Abarroteros con todos los demás clientes, en el caso de que sólo los clientes Abarroteros acepten recomendaciones en un período dado, teniendo estimaciones muy lejos de la realidad dado que los tipos de productos son totalmente distintos, tanto en frecuencia de compra como valor con respecto a los productos de los demás tipos de clientes.
- Esta evaluación no permite identificar factores humanos que pudieran estar interfiriendo en el impacto que este sistema de recomendación pudiera estar teniendo, en específico como el desempeño del vendedor tiene directa relación con la aceptación de las recomendaciones. Si bien se realiza una pequeña capacitación para utilizar el sistema, esta no se evalúa para ejercer las funciones de vendedor, por ende, puede darse que el vendedor realmente no entienda como funciona el sistema de recomendación (a nivel tecnológico o de estrategias). Otra situación que se puede dar tiene relación con el grado de confianza que el vendedor tiene al motor de recomendaciones, siendo posible que este no realice las recomendaciones de manera correcta debido a que no le cree al sistema o no ha tenido buenas experiencias con la utilización de este.

Debido a estos distintos problemas es que se han estudiado distintas alternativas de solución, con las siguientes características:

Soluciones	Ventajas	Desventajas
Diferencias en diferencias	Permite agregar y controlar por variables a elección	Selección de grupo de control y tratamiento puede llevar a estimaciones erróneas

Modelo de Rubín	Grupos de control mejor definidos dado que agregan más información	<ul style="list-style-type: none"> <li>• Construcción solo considera variables observables</li> <li>• Necesidad conocer historia del experimento</li> </ul>
Experimentos Naturales	Estimación más próxima a la realidad dado que no se tienen supuestos	Alto costo para el cliente

Tabla 2. Trade off distintas soluciones

En primera instancia, se tenía como motivación utilizar alguna de estas metodologías para determinar el efecto real en las ventas de esta herramienta. Sin embargo, se determinó que estas no eran del todo necesarias dadas las características que posee la base transaccional de Alicorp, ya que en ella se puede identificar claramente cuáles son las ventas realizadas a través de la herramienta de recomendaciones. Por lo que el trabajo primero es determinar una nueva metodología de evaluación, más simple y efectiva, y a su vez determinar las razones de porque la evaluación actual presenta problemas, estimando si se está sobreestimando o subestimando la venta realizada por el motor de recomendaciones.

Dejando de lado la evaluación de este motor de recomendaciones, se estudian además las tasas de efectividad de recomendaciones. Definiéndose estas de la siguiente forma:

$$\text{Tasa de efectividad} = \frac{\text{Nº productos recomendados aceptados}}{\text{Nº productos recomendados visualizados}}$$

Encontrándose, a nivel de estrategias, que estas tasas no son altas, es decir, que un bajo porcentaje del total de recomendaciones es aceptado. En específico, se encuentra que la tasa de efectividad promedio de la estrategia asociado es de un 6,5%, la de estrategia frecuente de un 7,5%, la estrategia perfil de un 8,5% y la estrategia zonas calientes de un 1,5%.

Es por esto que se propone generar una estrategia de recomendación que pueda lograr una mejor tasa de aceptación que las demás estrategias, o al menos mejor que la estrategia con menor aceptación. Además, se propone generar experimentos con relación al número de productos recomendados por estrategia y el orden en el que se muestran los productos según las estrategias.

Para la creación de esta estrategia se recurre al estudio de los distintos algoritmos de recomendación existentes actualmente, los cuales se dividen en dos grandes ramas: los algoritmos de filtros colaborativos y los algoritmos en base al contenido.

Para la selección de un algoritmo se deben considerar las limitantes que se encuentran en términos de datos que se disponen. En el caso de Alicorp, en específico del canal minorista, se poseen datos netamente transaccionales, con escasa clasificación a detalle de los productos. Debido a esto es que se opta por seguir un algoritmo en base a filtros colaborativos, dado que utilizan el historial de actividades asociadas a la empresa (datos transaccionales) y que para utilizar algoritmos en base al contenido se necesita de una mayor especificación de las características de los productos.

Ya dentro de los algoritmos de filtros colaborativos, se debe decidir entre los dos tipos que existen siendo estos los filtros colaborativos en base al contenido y los filtros colaborativos en base a un modelo. Siendo el principal trade off entre estos dos la precisión en las recomendaciones y el tiempo de procesamiento para estas. Dado que se quiere obtener una estrategia con mayor aceptación se debe optar por una con una mayor precisión en los resultados, es decir, una que en teoría debiera recomendar productos que pudieran ser de mayor valor para el cliente. Siendo esta la razón de optar por un filtro colaborativo en base al contenido.

Por otra parte, se debe considerar que actualmente las estrategias asociado y perfil corresponden a algoritmos de filtros colaborativos en base a modelos, es específico, a modelos de reglas de asociación y de segmentación. Por tanto, tiene sentido el utilizar un tipo de filtro colaborativo más enfocado a lograr resultados más precisos.

Finalmente, ya pensando en filtros colaborativos basados en la memoria, se tienen dos tipos de modelos:

- Filtros Colaborativos basados en los Usuarios
- Filtros Colaborativos basados en los Productos

Diferenciándose principalmente en que el primero busca recomendar en base a usuarios similares y el segundo busca recomendar en base a productos similares. Seleccionándose en definitiva una estrategia en base a Filtros Colaborativos basados en los Productos, dado que la base transaccional desde la cual se generará el modelo posee una mayor cantidad de usuarios (clientes) que de productos, siendo el tiempo de procesamiento menor si buscamos similitud en estos últimos.

En última instancia se realizarán experimentos de campo, en conjunto con Alicorp, para determinar la efectividad de la utilización de esta nueva estrategia. Para este experimento lo que se medirá será la efectividad o aceptación de la estrategia a implementar, esto en función de los indicadores *precision*. Por otra parte, para tener una correcta conclusión respecto a esta efectividad de la estrategia, debemos tener un grupo de control para poder determinar realmente si esta efectividad o efecto que produjo esta nueva estrategia en el grupo tratado no se debe a factores externos. Esta determinación de los grupos se llevará a cabo primero determinando clientes a los cuales controlar y tratar con perfiles similares, en base a la segmentación que ya posee Alicorp. Pero además se espera determinar clientes para estos grupos que sean propensos a la utilización de recomendaciones, es decir, que hayan utilizado alguna vez el motor de recomendaciones, esto principalmente dado que dado se tiene muy poco

tiempo para realizar pruebas y se necesita que la herramienta sean la mayor cantidad posible.

### **1.3 Objetivos**

#### **1.3.1. Objetivo General**

- Medir el efecto real del motor de recomendaciones sobre las ventas, determinando estrategias de mejora a la efectividad de este.

#### **1.3.2. Objetivos Específicos**

- Definir una metodología para evaluación de motor de recomendaciones.
- Proponer y generar estrategias de mejora para el motor de recomendaciones
- Diseñar experimentos para testear las estrategias de mejora propuestas.
- Evaluar efectividad de las distintas estrategias de mejora en base métricas de evaluación.

### **1.4 Alcances**

El alcance de esta memoria, respecto al problema de la venta incremental, corresponde a un análisis de los factores que afectan el cálculo de ésta con tal de determinar si existe una sobreestimación o subestimación del valor de la venta incremental, y por tanto del motor de recomendaciones en sí. Considerando, para este análisis, toda la información histórica del Pedido Sugerido, es decir, desde la época de implementación de las recomendaciones hasta la actualidad. Además se presentará una metodología alternativa de cálculo, que debiera dar solución a los problemas detectados.

Por otra parte se desarrollará una nueva estrategia de recomendación utilizando filtros colaborativos. Considerando para estas recomendaciones sólo la información transaccional de los clientes en los últimos 6 meses, dado que de ésta manera se asegura que se tomen en consideración sólo productos que actualmente son de preferencia para el cliente. Se debe acotar que este modelo de recomendación no busca ser ejecutable de manera periódica por tanto no se entregará ningún código ni archivo para su ejecución.

Finalmente se definirán distintos grupos de clientes, en función de las efectividades promedio de los últimos 3 meses, con tal de probar las distintas hipótesis que se plantean en relación a la visualización de las estrategias y a la evaluación de la nueva estrategia desarrollada. Siendo esta evaluación también en términos de la efectividad promedio resultante de los experimentos.

### **1.5 Resultados Esperados**

Los resultados que se esperan determinar son:

- Determinación de problemas que presenta la metodología actual de venta incremental.

- Determinar nueva metodología para evaluación de ventas incrementales.
- Modelo complementario a la estrategia de baja efectividad en el motor de recomendaciones.
- Informe de resultados de los experimentos realizados con esta estrategia complementaria.
- Reporte de evaluación de estrategia complementaria.

## 2. Marco Conceptual

### 2.1. Motores de recomendación

Los motores de recomendación tienen como propósito ayudar a los clientes a identificar artículos que sean de su interés, dentro de un amplio espacio de búsqueda. Estas herramientas, han demostrado ser un sistema efectivo e innovador a la hora de encontrar las necesidades de los clientes, y con esto aumentar las ventas de las empresas donde se utilizan.

Existen distintos modelos de recomendación a lo largo de la industria los cuales utilizan distintos modelos o información para la generación de estas. Pudiendo utilizarse datos como el historial de compra, calificaciones de los usuarios y, en el último tiempo, la interacción dentro de los sitios web.

### 2.2. Tipos de motores de recomendación

Como se mencionó anteriormente existen distintos tipos de motores de recomendación, diferenciándose estos en el tipo de información que utilizan. Pudiendo ser esta:

- Datos transaccionales de cada usuario
- Calificación de los productos por parte de los usuarios
- Los metadatos de cada producto o servicio

Por otro lado, los tres grandes tipos de algoritmos que se utilizan en la actualidad corresponden a:

- **Algoritmos de Filtros Colaborativos:** Estos se basan en el historial de interacciones que ha tenido el usuario con la empresa (Calificaciones, compras, etc.)
- **Algoritmos en Base al Contenido:** se construyen en base a los metadatos de los productos o servicios (características del producto, categorías, etc.)
- **Algoritmos Híbridos:** Son básicamente una combinación de los algoritmos mencionados anteriormente.

#### 2.2.1. Algoritmos de Filtros colaborativos

Los filtros colaborativos son usados para asociar clientes o productos similares, pudiendo con esto realizar mejores recomendaciones personalizadas, dado que se dispone de una recomendación más cercana al cliente. Para esto se necesita utilizar toda la información correspondiente a las interacciones del cliente, considerando por ejemplo historial de compra o de recomendaciones aceptadas. Aunque en general se utilizan calificaciones que los mismos clientes han entregado.

Actualmente existen tres técnicas de filtros colaborativos, son las siguientes:

**i. Filtros Colaborativos basados en la memoria:** Usan los datos de los usuarios para calcular similitud entre usuarios o productos y así realizar recomendaciones en función de esas similitudes calculadas.

**ii. Filtros Colaborativos basados en un modelo:** Usan los datos del usuario para estimar o aprender un modelo para hacer predicciones. Pudiendo ser estos modelos algoritmos de data mining y machine learning, como modelos bayesianos, reglas de asociación o redes neuronales.

**iii. Filtros Colaborativos híbridos:** Estos filtros combinan dos o más técnicas de recomendación para hacer predicciones.

A continuación, se muestra una tabla comparativa entre los principales tipos de modelos de Filtros Colaborativos:

Tipo de Filtro Colaborativo	Ventajas	Desventajas
Filtro Colaborativo basado en memoria	<ul style="list-style-type: none"> <li>Más ajustados a la realidad que los basados en modelos</li> <li>Permite actualización de base de datos de forma simple</li> </ul>	<ul style="list-style-type: none"> <li>Cálculo de recomendaciones es lento.</li> <li>Puede no generar recomendaciones para todos los usuarios.</li> </ul>
Filtro Colaborativo basado en modelos	<ul style="list-style-type: none"> <li>Son modelos escalables</li> <li>Es más rápido en generar recomendaciones que los modelos basados en memoria</li> </ul>	<ul style="list-style-type: none"> <li>Son inflexibles, es difícil agregar información.</li> <li>La calidad de los pronósticos no es buena en comparación a los modelos basados en memoria.</li> </ul>

Tabla 2. Comparativa filtros colaborativos basados en memoria y basados en modelos

A continuación, nos centraremos en los Filtros Colaborativos basados en memoria, dado que se quiere lograr una estrategia con una mayor calidad de pronósticos.

### 2.2.1.1. Filtros Colaborativos basados en memoria

Como se mencionó anteriormente los filtros colaborativos basados en la memoria, utilizan toda la base de datos que posea información relevante para poder encontrar recomendaciones de productos para cada cliente. A grandes rasgos, estos se dividen en los filtros colaborativos basados en los usuarios, utilizados para encontrar usuarios similares para cada usuario, y filtros colaborativos basados en productos, que intentan buscar productos similares para cada producto. Centrándonos en estos últimos debido a que la base transaccional de la que se dispone contiene una mayor cantidad de usuarios que de productos, convirtiéndose, por tanto, en la opción con un menor tiempo de procesamiento.

### 2.2.1.2. Filtros Colaborativos basados en los productos

Los filtros colaborativos basados en los productos determinan los productos que son más similares a cada uno, con tal de recomendar productos que estos hayan comprado y él no.

Para determinar el algoritmo que usa esta clase de Filtros Colaborativos primero debemos caracterizar el problema que responden estos. El input del problema es una matriz de interacción  $A = (a_{ij})$  de tamaño  $M \times N$ , asociada con  $M$  usuarios  $C = (c_1, c_2, \dots, c_m)$  y  $N$  productos  $P = (p_1, p_2, \dots, p_n)$ . Enfocándonos en data transaccional, se tiene que  $a_{ij}$  puede tomar valores de 0 o 1, con una representando una transacción observada entre  $c_i$  y  $p_j$ , representado 0 una ausencia de transacción.

Luego, lo primero que realiza el algoritmo es computar una matriz de similitud de productos  $WP = (wp_{st})$ ,  $s, t = 1, 2, \dots, N$ . Este calcula  $wp_{st}$  en base al vector  $A$ . Un alto  $wp_{st}$  indica que el producto  $s$  y  $t$  es similar en que muchos consumidores han comprado ambos  $A \cdot WP$  da los scores potenciales de cada consumidor. En la matriz resultante el elemento en la  $c$ th fila y  $p$ th columna, corresponde a los scores de similitud agregados entre el producto  $p$  y otro producto que el consumidor  $c$  ha comprado.

Se deben definir dos criterios para poder construir este algoritmo:

1. **Similitud de Productos:** Son los tipos de métricas que determinan si un producto es similar a otro, estos dependen de distintos elementos según el criterio que se seleccione.
2. **Cantidad de Productos Similares:** Es necesario definir cuántos productos son considerados como similares para cada uno. Siendo el método de los vecinos más cercanos para cada usuario es la forma más típica de determinar esto.

## 3. Metodología

### 3.1. Venta Incremental

### **3.1.1. Análisis exploratorio**

Se realizará un análisis teniendo como objetivo entender completamente la lógica detrás del cálculo actual de la venta incremental, entendiendo cuáles son sus puntos débiles y en función de estos buscar alternativas que permitan una mejor estimación de esta.

### **3.1.2. Propuesta de cálculo venta incremental**

En esta etapa se tiene como objetivo entregar una fórmula alternativa para el cálculo de la venta incremental del motor de recomendaciones, en base al análisis realizado anteriormente. Se debe mostrar la lógica detrás de esta opción por medio de ejemplos concretos e indicar claramente cuáles son los beneficios respecto a la fórmula actual.

## **3.2. Efectividad de recomendaciones**

### **3.2.1. Análisis Descriptivo**

Se necesita hacer una revisión exhaustiva de los datos disponibles para realizar el modelo de recomendaciones, por lo se debe realizar un análisis para determinar qué información es útil para realizar el modelo y si es posible lograr una mayor precisión y exposición en términos experimentales dado este análisis.

### **3.2.2. Selección, limpieza y transformación de datos**

Es necesario seleccionar, en base al análisis realizado, las variables con los cuales se construirá el modelo. Determinando también la cantidad de datos históricos a considerar para la realización de este. Luego de esto se procede a revisar la calidad de los datos, y de ser necesario, realizar una limpieza para construir el modelo en base a información sólida.

Finalmente realizaremos transformaciones con el fin de poder implementar modelos mediante software para la generación del modelo de recomendación

### **3.2.3. Generación de modelo**

Para esta etapa ya con los datos seleccionados y limpios se procede a construir el modelo de recomendación seleccionado, es decir, el modelo de filtro colaborativo en base a productos. Para esto se realiza su programación en el programa a seleccionar, siendo la primera etapa de esto la determinación de los distintos parámetros de similitud para el modelo. Necesitando para esto último trabajar con la base de datos para encontrar indicadores de precisión que permitan optar por el más conveniente. Finalmente, ya determinado esto se dispone a construir el modelo y realizar pruebas en base a la data transaccional disponible.

### **3.2.4. Diseño Experimental**

En esta etapa primero que todo se necesita determinar cuáles son los grupos de control pertinentes a los criterios anteriormente mencionados, para luego comenzar a realizar

las pruebas en campo. Es fundamental que este proceso se realice de la forma que se planeó por tanto se debe estar en contacto constante con quienes trabajan junto a Alicorp y están encargados del motor de recomendaciones.

### 3.2.5. Evaluación de Resultados

En esta etapa se evaluarán los resultados de los experimentos, en función de indicadores correspondientes a la tasa de aceptación de las recomendaciones entregadas, en específico al indicador de *precision*. Por otro lado, con estos resultados se debe determinar si la estrategia es efectiva para una futura implementación

## 4. Desarrollo Metodológico

### 4.1. Venta Incremental

#### 4.1.1. Análisis Exploratorio

Primero que todo, en esta etapa, es relevante conocer el comportamiento de las ventas logradas por Alicorp para así tener una noción acerca del aporte que puede estar entregando la herramienta de recomendaciones a estas (Pedido Sugerido), por tanto, se realiza una serie de tiempo desde el comienzo de las ventas de Alicorp en 2011 hasta la etapa actual que corresponde a octubre de 2016. Hay que tener claro que la incorporación de esta herramienta comenzó a principios de 2014, siendo esta una implementación paulatina, por ende, el aporte a las ventas no es significativo hasta principios de 2015. En un principio, el valor de estas ventas se encontraba en nuevos soles peruanos, pero se convirtió a dólares con el fin de ayudar a cuantificar de mejor forma la magnitud de estas ventas, siendo la tasa de cambio de 0,29 pesos peruanos por dólar.

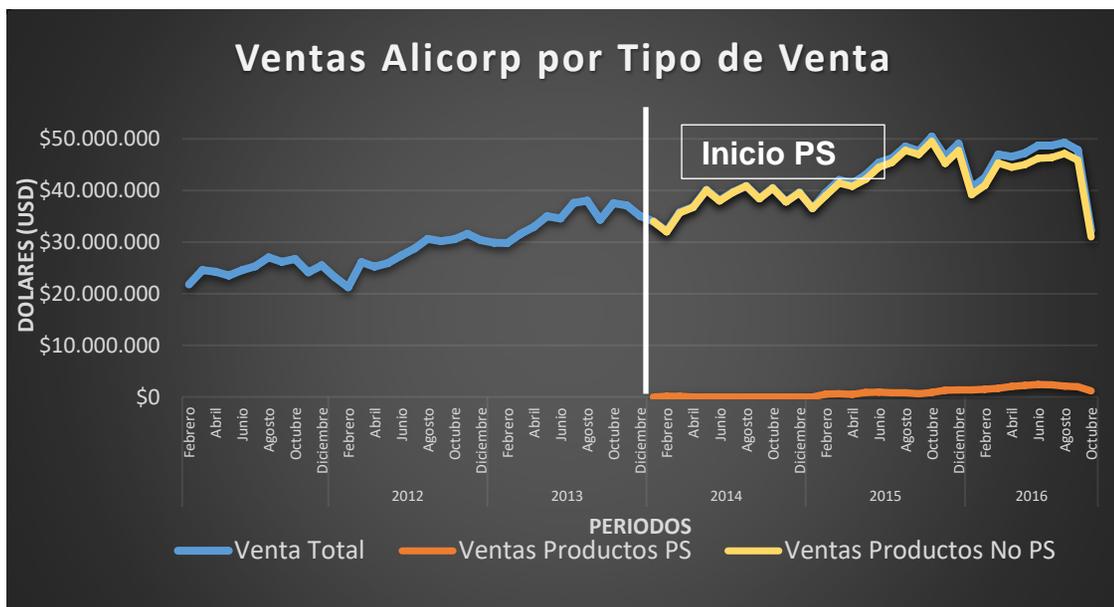


Gráfico 2. Ventas Alicorp por tipo de venta. Elaboración propia

Del Gráfico 2 se puede observar que las ventas por productos recomendados (PS) si son un aporte sustancial a la venta total de Alicorp, sin embargo, este aporte realmente comienza a apreciarse a principios del año 2015 por lo que el análisis a partir de ahora considerará desde este año en adelante.

Para apreciar mejor esto observemos el Gráfico 3, el cual no entrega el aporte mensual del PS en función de la venta total:

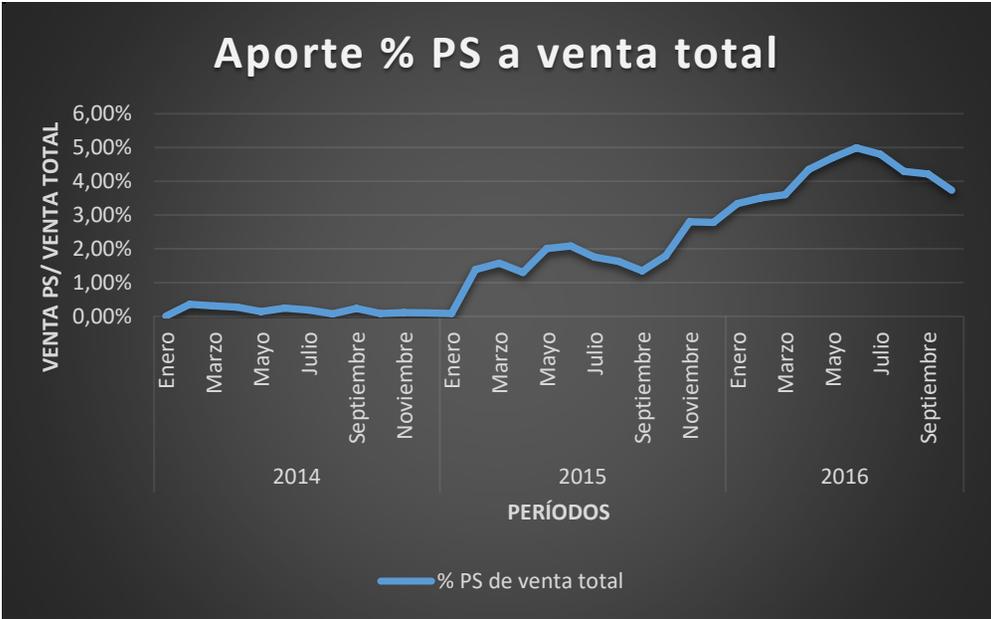


Gráfico 3. Aporte % PS a venta total. Elaboración propia

Se puede apreciar aún más que recién a comienzos del año 2015 la venta por productos recomendados fue un aporte para las ventas totales de Alicorp, por lo que se considerará el análisis de estrategias solo para esos períodos.

Ahora que ya sabemos que la venta producto de las recomendaciones si tiene incidencia en la venta total facturada por Alicorp, debemos determinar realmente si el método que están utilizando actualmente tiene problemas de estimación de este efecto, es decir, si las ventas incrementales que se calculan están siendo sobreestimadas o subestimadas. Para esto observemos el gráfico a continuación, el cual nos muestra los valores de venta incremental que calcula la empresa sobre el total de sus clientes mensualmente y por otro lado muestra el valor de las ventas producto de la compra de productos recomendados por parte del cliente.

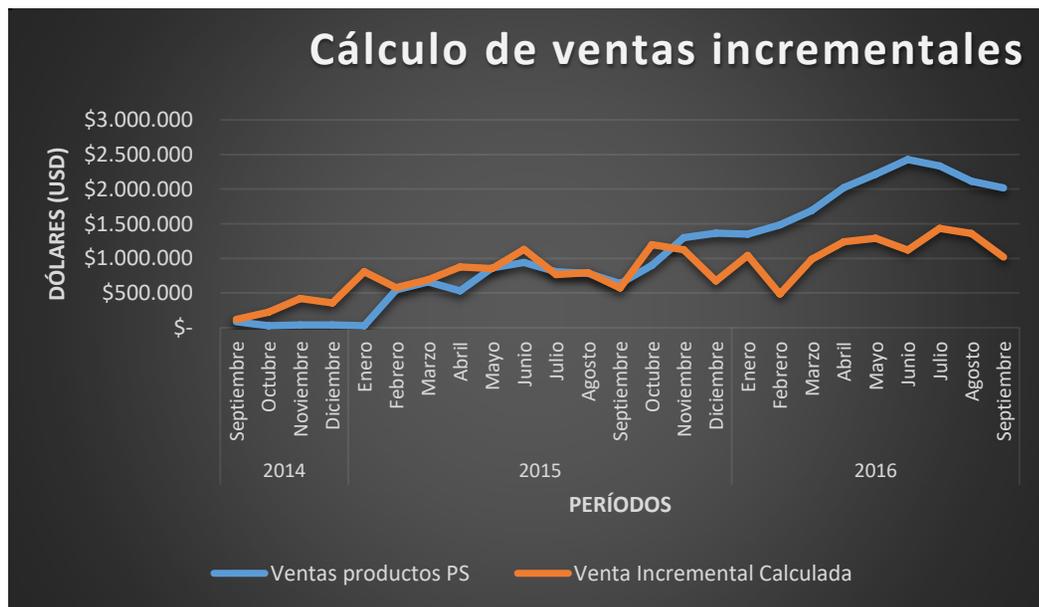


Gráfico 4 Cálculo de ventas incrementales según método. Elaboración propia

En este podemos observar que la subestimación o sobreestimación del cálculo de ventas incrementales ha tenido un comportamiento variable a lo largo del tiempo. Por ejemplo, si consideramos hasta mitad de 2015 (Junio) se aprecia que durante ese período se estuvo sobreestimando la venta incremental que generaban las recomendaciones, sin embargo si consideramos desde Septiembre de 2015 hasta el mismo mes de 2016 podemos concluir que actualmente se está subestimando la venta incremental que entra el PS. Para una mirada más clara, se calculó el área bajo la curva para el período completo del gráfico y además se calculó para el último año móvil, es decir, Septiembre 2015 a Septiembre 2016. Los resultados se presentan en la Tabla 3, a continuación:

Período	Ventas PS (USD)	Venta Incremental (USD)	Diferencias en ventas (USD)
<b>Total</b>	<b>\$27.275.504</b>	<b>\$21.241.377</b>	<b>\$6.034.127</b>
<b>Último Año Móvil</b>	<b>\$21.885.281</b>	<b>\$13.587.316</b>	<b>\$8.297.965</b>

Tabla 3. Diferencias en venta incremental según cálculo. Elaboración propia

Como podemos observar las diferencias son gigantescas, por ende, podemos afirmar que realmente existe un problema en el cálculo de ventas incrementales producto de la herramienta de recomendaciones. Por tanto, el desarrollo a realizar tiene relación con encontrar las razones de esta gran diferencia y determinar si la metodología propuesta realmente se hace cargo de estos problemas.

#### 4.1.2. Propuesta Cálculo de Venta Incremental

Ya determinamos que existe una subvaluación de la venta incremental que entrega la herramienta de recomendaciones y que por tanto la metodología utilizada para su medición tiene problemas. A continuación, se mostrará el desarrollo que se realizó para determinar el porqué de esta subvaluación, utilizando la información histórica

transaccional e información histórica del cálculo de esta venta incremental, además de demostrar porque es factible y necesario simplificar esta metodología de cálculo.

A priori, se determina que esta subvaluación se debe a problemas en el cálculo de la metodología utilizada por la empresa, los cuales se enumeraron en secciones anteriores. Por tanto, para revisar esto, se determinan los factores de crecimiento PS y crecimiento de mercado utilizados por la empresa para calcular la venta incremental. En concreto, se utilizaron los resultados por zona por lo que la posibilidad de casos extremos es nula, centrándose en el caso 3 y caso 4 mostrados en la tabla anexa. Finalmente, si separamos los casos por meses y años tenemos el siguiente gráfico en relación a los casos:

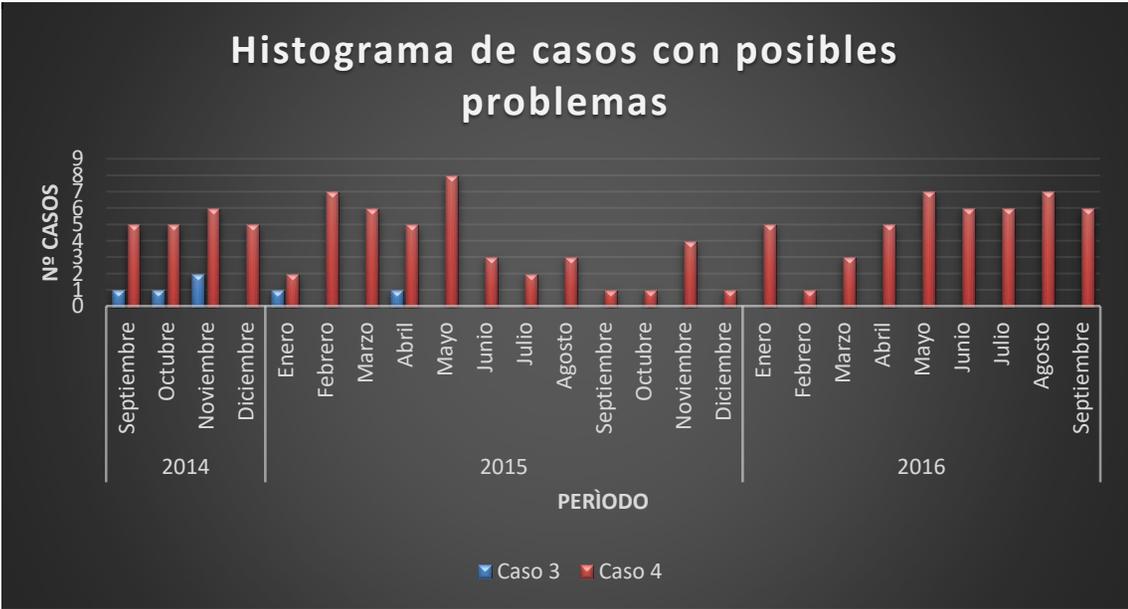


Gráfico 5 Histograma de casos con posibles problemas. Elaboración propia

Es difícil encontrar una relación respecto a los casos encontrados y esta subvaluación de las ventas incrementales, principalmente debido a que los casos encontrados tienden a amplificar este efecto. Pudiendo extraerse principalmente que hay problemas con el grupo de comparación al que está sujeto el grupo de clientes que compran con PS, es decir, con el “mercado” con el que se compara.

Por otra parte, esta subvaluación también tiene relación con los grupos de comparación a los cuales se somete la evaluación, los cuales pueden ser demasiado estrictos y por ende aminorar el efecto real de las recomendaciones. En base a la fórmula existen dos tipos de comparaciones, primero una comparación entre los mismos grupos para determinar el crecimiento de mercado (clientes sin PS) y el crecimiento PS (clientes que han comprado con PS) y en segunda instancia ocurre una comparación entre los clientes sin PS con los que sí han comprado con PS.

Como se menciona en el párrafo anterior para que ocurra una subvaluación de esta venta incremental se necesita que el grupo contra el que se compara tienda a aminorar el efecto en las ventas, lo que en este caso se traduce en que el grupo sin PS tenga un

crecimiento de mercado que vaya en alza. Para esto analizaremos el comportamiento de ventas en el tiempo para cada grupo, obteniendo el siguiente gráfico:



Gráfico 6. Comparación ventas entre clientes con PS y clientes sin PS. Elaboración propia.

Como podemos observar en el gráfico 6 la diferencia de las ventas del grupo compuesto por clientes que, si han comprado por PS, se tiene que los clientes fuera de este grupo tienen un comportamiento constante en cuanto a las compras que realizan. Por tanto, la comparación con este grupo no es un problema en cuanto a la subvaluación dado que no aminoran en el efecto en las ventas incrementales del PS. Sin embargo, que no sean un problema para ese tema en particular no los convierte en un grupo de comparación adecuado. De hecho, en base a la historia de compra de estos clientes podemos afirmar que no son un correcto grupo de comparación ya que al parecer son clientes que no presentan variación alguna y por ende se les puede denominar malos clientes.

A continuación, se estudiará el caso relacionado con la comparación entre los clientes del mismo grupo. Para esta comparación se utilizarán los indicadores de crecimiento de mercado y crecimiento de PS a lo largo del tiempo que el motor de recomendaciones ha sido implementado, siendo relevante que para esto se considerarán los clientes en su totalidad, es decir, sin separar entre clientes nuevos o antiguos. Los resultados se presentan en el siguiente gráfico:

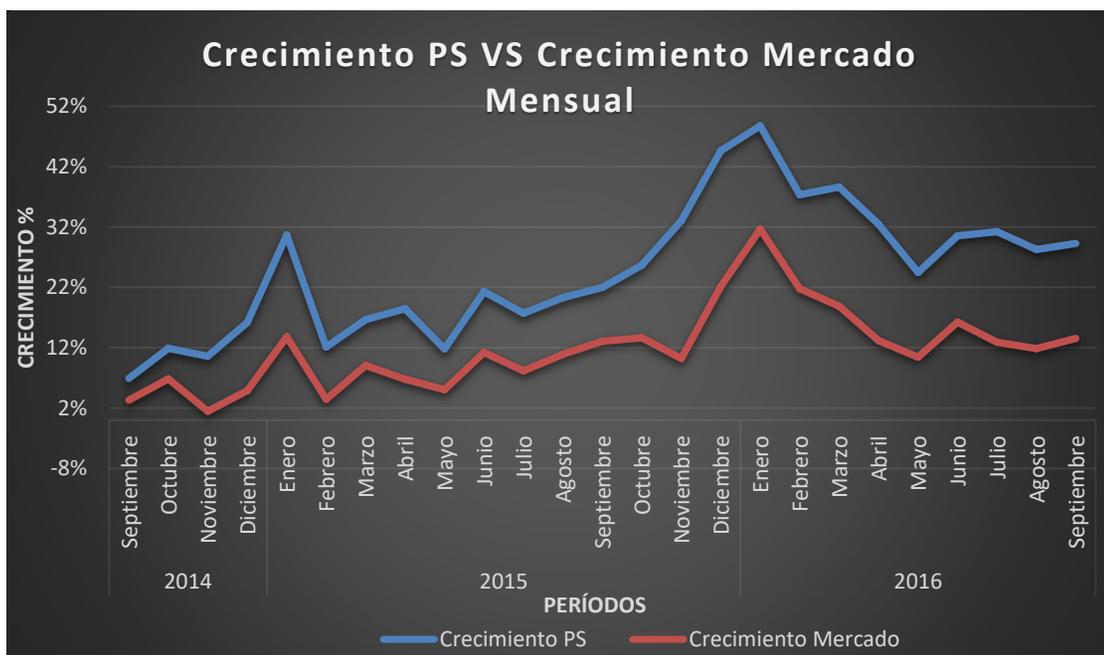


Gráfico 7. Comparación crecimiento PS y crecimiento mercado. Elaboración propia.

Podemos observar que, en el período de interés, el crecimiento PS toma tiene variaciones bastante significativas de un mes a otro. Por ende, con la formulación que existe para el cálculo de la fórmula incremental, se tiene que siempre se está comparando con crecimientos o desviaciones bastante fuertes logrando de esta forma aminorar el efecto y no mostrar que el crecimiento es cada vez más grande en términos de ventas. Siendo esto una de las razones fundamentales de la subestimación de las ventas incrementales entregadas por la herramienta de recomendaciones.

Por otro lado, hay que destacar que este impulso logrado en las ventas del PS, tienen relación no solo con el período de adaptación a esta herramienta, sino que también con la cantidad de clientes que se mantiene utilizando las recomendaciones dadas por esta. Como podemos observar en el siguiente gráfico, la cantidad de clientes antiguos con respecto a los clientes nuevos es bastante superior en el año 2016 por lo que la venta de productos recomendados mantiene la base que tenía en 2015 y a la vez se potencia a través de los nuevos clientes que aceptan este tipo de productos. Además, al mantener clientes antiguos se generan menos problemas de medición con la metodología actual, lo cual también puede ser un factor para que el cálculo de estas ventas sea mayor.

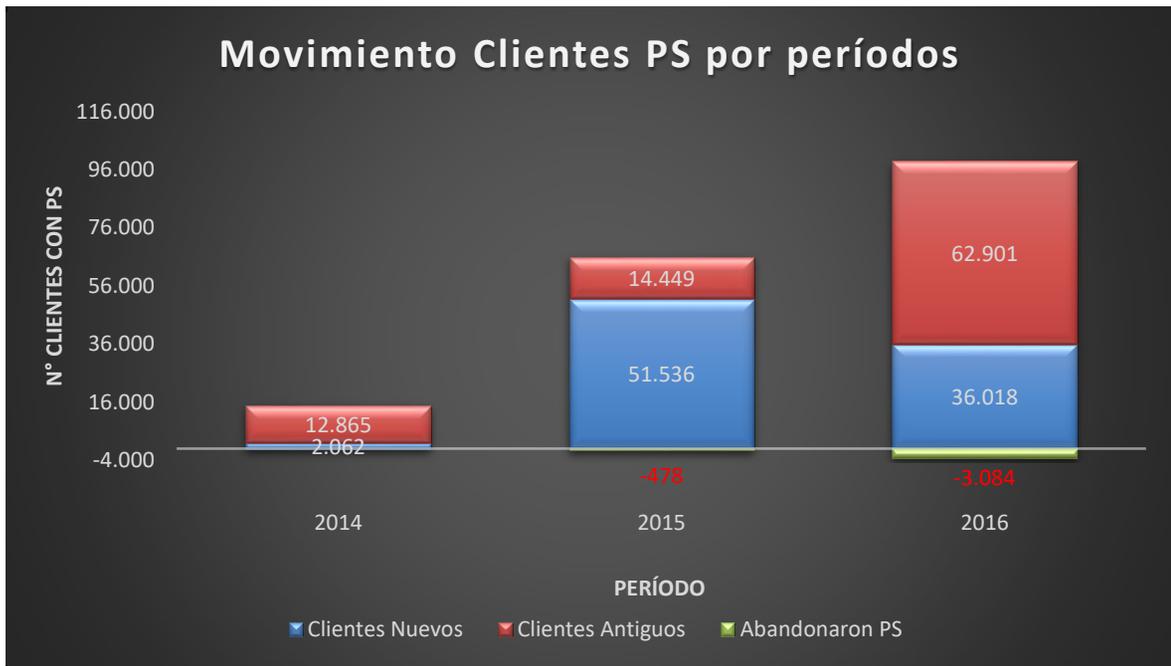


Gráfico 8. Evolución clientes PS anualmente. Elaboración propia

Finalmente dado lo que se demostró, se considera necesario el cambio de la metodología actual dado que es confusa y a veces puede llevar a subvaluación o sobreestimación de las ventas incrementales que entrega la herramienta de recomendaciones. Hay que destacar que esta metodología no es necesaria dado que no se encuentran las mismas restricciones que si se encuentran en la evaluación de promociones, situación para la que originalmente está diseñada esta metodología. A diferencia de las promociones se puede identificar claramente cuanto de las ventas corresponde a los productos recomendados, dado a que se cuenta con un identificador dentro de la base transaccional, y además se puede afirmar que no existe efecto en las compras que realiza el cliente dado que todo el proceso de recomendación ocurre post venta.

En base a esto se propone como metodología a utilizar para determinar la venta incremental, simplemente la suma de las ventas correspondientes a productos recomendados aceptados. Esto ya que de esta forma realmente se está representando el valor agregado que está entregando la herramienta de recomendaciones, es decir, está mostrando las ventas que agrega a un cliente en comparación a que no se hubiera utilizado la herramienta en el cliente. Además, obviamente esto simplifica de sobremanera los cálculos a realizarse y por ende se ahorran recursos de espacio y tiempo tecnológico.

$$Venta Incremental = \sum_i^n Valor \text{ Producto Comprado por } PS_i$$

## 4.2. Efectividad de estrategias

### 4.2.1. Análisis Descriptivo

Cómo se mencionó anteriormente, el año 2014 se considera un período de adaptación e implementación de la herramienta de recomendaciones. Por lo tanto, el cálculo y estudio de cualquier factor relacionado a la efectividad de las estrategias sólo considerará los datos disponibles a contar de principios del año 2015.

La efectividad será considerada como el número de recomendaciones aceptadas respecto al total de recomendaciones realizadas, siendo este indicador muy relevante para determinar cuál estrategia es la que está teniendo mejor respuesta de cara al cliente. De todas formas, hay que considerar que este indicador puede tener un sesgo, dado que no existe alguna etiqueta o una forma de saber si realmente el producto recomendado que fue visualizado en el PAD fue realmente ofrecido por el vendedor al cliente.

Actualmente están operando cuatro estrategias distintas, teniendo distintas preponderancias dentro de la visualización que se le da al vendedor. En la ilustración 1 podemos observar el número de productos que se muestra por cada estrategia, como también el orden en el cual se visualizan estas dentro de la herramienta:



Ilustración 2. Distribución de estrategias en PS

En función de determinar estrategias de mejora se debe observar cuales son las que muestran peor desempeño en función del indicador de efectividad, por lo tanto, se determinó la efectividad promedio de cada una de estas estrategias junto a su desviación estándar. Los resultados se pueden observar en el Gráfico 5 y Tabla 4, a continuación:

Indicador	Asociados	Frecuentes	Perfil	Zonas Calientes
Efectividad (%)	6,47	7,47	8,53	1,47
Desviación Estándar (%)	5,0	5,6	4,4	0,7

Tabla 4. Efectividad promedio y desviación por estrategia. Elaboración propia

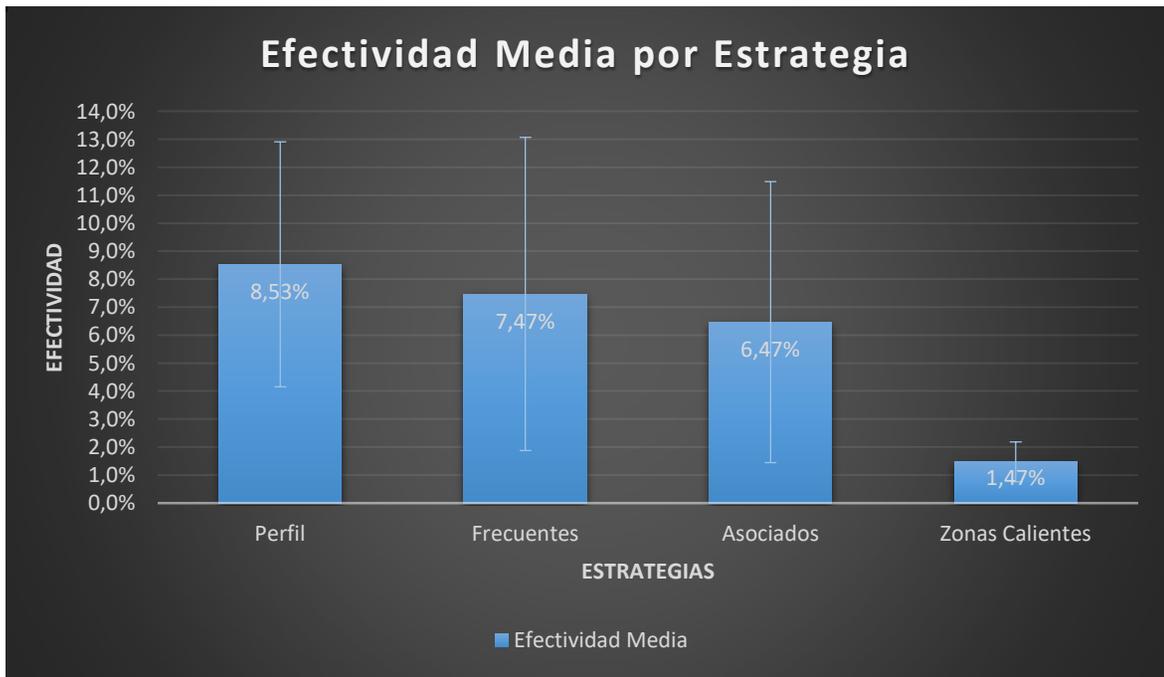


Gráfico 9. Efectividad media por estrategia. Elaboración propia.

Si nos centramos sólo en el indicador de efectividad, se podría pensar que la peor estrategia corresponde a Zonas Calientes. Sin embargo, también es relevante considerar las ventas incrementales que generan cada una de estas estrategias. Esto dado que quizás una estrategia con menor efectividad puede generar mayores ventas incrementales que otra con una mayor efectividad, debido a que la primera recomienda productos que tienen un mayor valor.

Para determinar de manera sencilla qué estrategia genera mayores ventas incrementales, se considera la metodología alternativa propuesta en la sección anterior. Por tanto, la venta incremental se calcula como la suma de las ventas de productos sugeridos que fueron comprados por los clientes en el período 2015-2016, por cada una de las estrategias a evaluar. Los resultados se muestran a continuación en el siguiente gráfico:

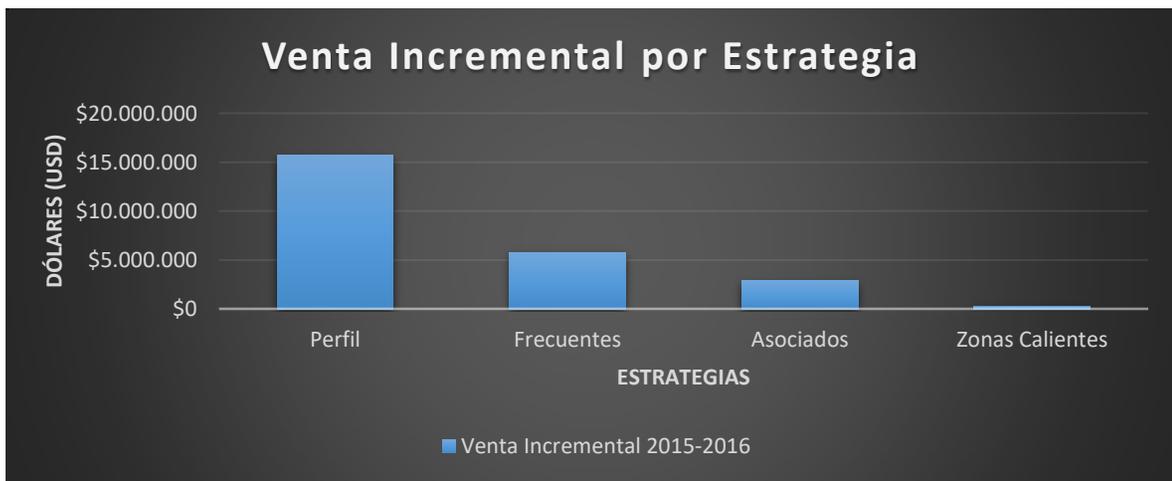


Gráfico 10. Venta incremental por estrategia. Elaboración propia

Del Gráfico 6 se puede observar que el orden en cuanto a las mejores estrategias no cambia, donde la estrategia Zonas Calientes sigue siendo la peor estrategia dentro de las cuatro y la estrategia Perfil la mejor.

Volviendo al análisis de efectividad, se busca determinar causas de porque una estrategia tiene mejor efectividad que otra. Por tanto, se dispuso a analizar las variables que determinan este indicador por cliente, es decir, la variación de productos visualizados y aceptados por cliente anualmente.

Indicador	Asociados	Frecuentes	Perfil	Zonas Calientes
Promedio	12,2	31,5	33,3	12,0
Desviación Estándar	13,9	30,4	28,1	8,7

Tabla 5. Productos Visualizados por cliente anualmente. Elaboración propia.

Indicador	Asociados	Frecuentes	Perfil	Zonas Calientes
Promedio	3,7	4,0	6,2	1,6
Desviación Estándar	4,0	4,8	8,0	1,1

Tabla 6. Productos recomendados aceptados promedio por cliente anualmente. Elaboración propia

Dado estas tablas podemos determinar que al parecer existe una relación entre el grado de efectividad de cada estrategia y el número de productos que se visualizan, esto dado que las dos estrategias con mejor efectividad corresponden a las que visualizan una mayor cantidad de productos recomendados por cliente anualmente. Sin embargo, si nos centramos en las estrategias Asociados y Zonas Calientes podemos observar que esta relación no se cumple dado que aún con una cantidad similar de productos visualizados Zonas Calientes sigue siendo peor en efectividad que la estrategia Asociados.

En base a estos resultados se puede determinar que existen diversos caminos para la mejora del motor de recomendaciones, por un lado, se tiene la creación de una nueva estrategia que pueda reemplazar a la estrategia de Zonas Calientes buscando tener una mejor efectividad que esta. Otra alternativa tiene relación con el orden en que se muestran las recomendaciones, dado que coincidentemente la estrategia con peor efectividad coincide con ser la última estrategia es ser desplegada. Finalmente, otro camino a tomar tiene relación con el número de productos recomendados por estrategia que se le muestran al vendedor y al cliente, donde al parecer la cantidad recomendaciones que se visualizan muestran una mayor probabilidad de que el cliente acepte esta recomendación.

#### 4.2.2. Selección, limpieza y transformación de datos

Dado el alto nivel de procesamiento requerido para realizar una recomendación de filtros colaborativos para la cantidad de clientes y productos en el portafolio de Alicorp, se opta por realizar un modelo acotado a los clientes más valiosos de la cadena y con mayor potencial de ser expuestos a esta nueva estrategia de recomendación, según criterios de venta incremental y número de recomendaciones visualizadas.

Para determinar los perfiles con mayor valor en términos de venta incremental, se continúa utilizando la evaluación propuesta en secciones anteriores, teniendo los siguientes resultados:



Gráfico 11. Venta Incremental por perfil de cliente. Elaboración propia

Según este resultado, por mucho el cliente Abarrotero es el más valioso en términos de la venta incremental que se genera por uso del motor de recomendaciones. Sin embargo, también se debe considerar que este es el perfil que cuenta con más clientes asociados. Por esto es relevante normalizar esta venta incremental que entrega cada perfil en función de su tamaño, obteniéndose el siguiente resultado:



Gráfico 12. Venta Incremental por cada cliente de perfil. Elaboración propia

Coincidiendo con la intuición, al considerar la cantidad de clientes por perfil da como resultado que el tipo de perfil Panaderías es el que entrega mayor venta incremental por cliente.

Ahora consideraremos el potencial de exposición a la nueva estrategia desarrollada, teniendo como indicador la visualización de recomendaciones por perfil. Lo cual podemos observar en el siguiente gráfico:



Gráfico 13. N° de visualizaciones por perfil de cliente. Elaboración propia.

Siendo aún más específicos, dado que la estrategia desarrollada reemplazará en el experimento a la estrategia de Zonas Calientes, consideraremos el número de visualizaciones por perfil que actualmente tiene dicha estrategia con tal de asegurar que el perfil a seleccionar tenga exposición máxima. Obteniéndose el siguiente resultado:



Gráfico 14. N° de Visualizaciones estrategia Zona Caliente por tipo de perfil. Elaboración propia

Confirmando que el perfil con mayor exposición corresponderá al Abarrotero, siendo esto lo más relevante dado que determinará el éxito o fracaso del experimento a realizar. Es claro que si nadie visualiza la nueva estrategia, no hay información para poder evaluar esta.

Finalmente se determina que la estrategia de filtros colaborativos se desarrollará en función del perfil Abarrotero, el cual cuenta con 65.894 clientes y un portafolio de 844 productos distintos, dado que cuenta con una potencial alta exposición a la estrategia desarrollada y un valor significativo en el negocio en términos de venta incremental.

En relación a los datos a utilizar, lo primero a determinar tiene relación con la base de datos a utilizar. En el caso de Alicorp, se dispone de dos tablas con el transaccional histórico, una con la venta realizada al cliente por parte de los vendedores (denominada Preventa) y otra con la venta que efectivamente se le entregó al cliente, dado el stock disponible. Considerando que en este caso en particular, el stock queda fuera del alcance del proyecto y se quiere incrementar la efectividad de la venta del vendedor, se determinó utilizar la base transaccional de preventa.

Otro punto importante a determinar, es la historia transaccional a considerar en el desarrollo de esta nueva estrategia de recomendación. Definiendo finalmente utilizar solo los últimos tres meses de información, dado que se quieren lograr recomendaciones en base a productos que el cliente actualmente está comprando. Siendo específicamente estos meses Octubre, Septiembre y Agosto del año 2016.

La transformación de los datos más importante, tiene relación con la construcción de la matriz para determinar la similitud de productos. Teniendo que estructurar una columna

principal que contenga a todos los clientes considerados y una fila principal con todos los productos que han comprado alguna vez ese tipo de clientes, debiendo completarse con la calificación por cliente para cada producto. Esta calificación puede tomar el valor 1 o 0, siendo 1 cuando el cliente compra el producto dentro del horizonte de tiempo determinado, no importando la cantidad de veces que lo haya comprado, y 0 si es que nunca compro ese producto dentro de este rango.

Cientes ID	Prod 1	Prod 2	....	Prod 1351	Prod 1352	Prod 1353
23	1	1	....	1	1	1
45	0	0	....	1	1	1
41	1	0	...	0	1	0
94	0	0	...	0	1	1
12	1	0	...	0	0	0

....

Tabla 7. Matriz para utilizar filtros colaborativos en base a productos. Elaboración propia.

En relación a esta matriz se realizarán cálculos de similitud entre los distintos productos para generar una lista de productos similares para los distintos productos y así generar recomendaciones.

Además se debieron realizar transformaciones de id, tanto para producto como cliente, para lograr que coincidieran con las utilizadas en la implementación del motor de recomendaciones personalizadas.

#### 4.2.3. Generación de modelo

Para la realización del modelo se utilizó Rapidminer y R, tanto para la estructuración de datos como para el desarrollo del modelo y su posterior evaluación.

Como se mencionó anteriormente, para el desarrollo de filtros colaborativos basados en productos, se debe definir el criterio de similitud y la cantidad de productos similares a considerar. En este caso en particular, se utilizó el criterio de distancia Coseno para determinar la similitud entre productos, dadas las herramientas disponibles. Teniendo el siguiente resultado, considerando que la proporción de datos utilizada como control fue de un 30% y un 70% como entrenamiento.

- ❖ Distancia Coseno
  - Precision = 10,16 %

Por otro lado, para determinar la cantidad de productos similares es posible definir un rango de similitud o una cantidad fija de productos similares. Optando por esta segunda opción dado que de esta forma es posible asegurar obtener recomendaciones para todos los productos a considerar, utilizando el algoritmo de K-Nearest Neighborhood y fijando la cantidad de productos similares a entregar en N=10.

#### **4.2.4. Diseño Experimental**

Un objetivo principal de esta memoria es generar mejoras a esta herramienta de recomendaciones, habiendo como se mencionó anteriormente, tres maneras posibles de realizar esto. Siendo las siguientes:

1. Modificación del número de recomendaciones desplegadas por cada estrategia
2. Modificación en el orden en que se despliegan las distintas estrategias
3. Reemplazo de estrategia con peor efectividad por nueva.

Estas modificaciones van alineadas en aumentar la efectividad para las distintas estrategias y el motor de recomendaciones en general, para lo cual se realizarán experimentos en función de un grupo de control y se determinarán los resultados en función de la efectividad alcanzada para cada experimento.

A continuación se mostrará el diseño realizado para cada uno de estos experimentos, siendo relevantes los grupos determinados para controlar y tratar.

##### **4.2.4.1. Experimento 1: Modificación del número de recomendaciones desplegadas por cada estrategia**

Se cree este es un camino a tomar para la mejora en las recomendaciones, debido a que como se pudo observar en el análisis exploratorio al parecer existe una relación positiva entre el número de recomendaciones que se realizan y el número de productos aceptados. Por tanto, se espera que si se potencia el número de recomendaciones de la estrategia con mejor tasa de efectividad esta aumente aún más.

##### **a. Grupo de Control y Experimentación**

Primero que todo, dado que el objetivo es medir cómo varía la efectividad según este aumento de recomendaciones por estrategia. Se necesita que los clientes a considerar dentro de los grupos de control y tratamiento tengan una efectividad comparable, por tanto se deben conformar en base a este indicador. Dado esto se determina que los clientes dentro de este experimento pertenezcan a un mismo perfil de cliente y no exista una diferencia significativa entre la efectividad promedio de estos dos tipos de grupo.

Debido a que se evaluará el aumento de recomendaciones para dos estrategias, los grupos para el experimento se conformarán de la siguiente forma:



Ilustración 3. Conformación Grupos Experimento 1. Elaboración propia

Según esta estructura, contamos con un grupo de control el cual no se verá expuesto al experimento a realizar (Grupo 1). Luego tenemos al grupo 2 que tendrá una recomendación más de la estrategia perfil y por último el grupo 3 que contará con una recomendación más de tipo asociado. Donde la evaluación de estas dos últimos se realizará en comparación a los resultados obtenidos en el grupo de control.

Estos grupos están compuestos por una cantidad similar de clientes y con efectividades similares. A continuación se muestra en detalle dicha conformación:

i. Cantidad de Clientes

Existen restricciones en el número de clientes a los cuales se les puede realizar este tipo de experimentos por temas de costos e imagen de empresa. Por tanto se determinó llegar a un N suficiente para ser representativo del perfil seleccionado, considerando el universo de estos como aquellos que actualmente están expuestos al motor de recomendaciones. Según esto, se determina la siguiente cantidad de clientes por grupo:



Ilustración 4. Cantidad de clientes por grupo experimento 1. Elaboración propia

ii. Efectividad de Clientes

Dado que el objetivo es determinar las variaciones en la efectividad según los distintos cambios realizados en el experimento, es necesario asegurar que las efectividades promedio sean comparables. Para esto se determina la efectividad promedio de cada grupo y se realiza una prueba estadística que permite determinar que son significativamente iguales.



Ilustración 5. Efectividad promedio por grupo. Elaboración propia.

b. Diseño de Experimentación

Dado que se quiere aumentar el número de recomendaciones de una estrategia y el espacio es limitado para visualizarse en el PAD del vendedor, debemos sacrificar una recomendación de alguna otra estrategia. La estrategia seleccionada corresponderá a la que tiene peor efectividad, siendo reemplazada por la nueva recomendación de la estrategia asignada para cada grupo pero conservando el nombre de la estrategia reemplazada debido a temas técnicos de la herramienta.

Dados los resultados obtenidos en los distintos análisis exploratorios, la estrategia seleccionada será Zonas Calientes. Está siempre se encuentra en última posición, quedando la distribución de la siguiente forma por grupo:

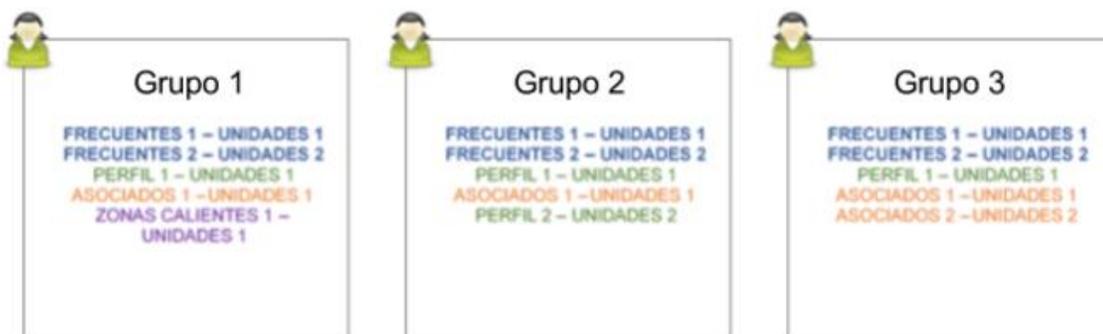


Ilustración 6. Distribución estrategias por grupo experimento 1. Elaboración propia.

#### 4.2.4.2. Experimento 2: Modificación en el orden en que se despliegan las distintas estrategias

La idea de este camino es determinar si realmente el orden en que se despliegan las distintas estrategias tiene injerencia en la efectividad de cada una de estas, teniendo como hipótesis de que efectivamente esto es cierto dado que coincidentemente las estrategias con menor efectividad son las últimas dos en desplegarse. Se cree que esto puede ser debido a que el vendedor sólo recomienda los primeros productos al cliente ya que no quiere sobrecargarlo de recomendaciones o molestarlo con tanta insistencia.

##### a. Grupo de Control y Experimentación

La conformación de los grupos sigue la misma lógica que la del experimento 1, variando solo el número de estos debido a la mayor cantidad de variaciones que se busca evaluar. Conformándose de la siguiente forma:



Ilustración 7. Estructura de grupos experimento 2. Elaboración Propia.

Según esta estructura, contamos con un grupo de control el cual no se verá expuesto al experimento a realizar (Grupo 1). Luego tenemos al grupo 2 que tendrá como primera recomendación la generada para la estrategia Zona caliente, después el grupo 3 será el que tenga asignado como primera recomendación la estrategia Asociado y por último el grupo 4 que contará con una recomendación del tipo Perfil en primera posición.

Estos grupos están compuestos por una cantidad similar de clientes y con efectividades similares. A continuación se muestra en detalle dicha conformación:

##### i. Cantidad de Clientes

Nuevamente al igual que con el experimento 1, existen restricciones en el número de clientes a los cuales se les puede realizar este tipo de experimentos por temas de costos e imagen de empresa. Determinando la siguiente cantidad de clientes por grupo:



Ilustración 8 Cantidad de clientes por tipo de grupo Experimento 2. Elaboración propia

ii. Efectividad de Clientes

Otra vez debemos asegurar que las efectividades promedio sean comparables. Para esto se determina la efectividad promedio de cada grupo y se realiza una prueba estadística que permite determinar que son significativamente iguales.



Ilustración 9. Efectividad promedio por grupo. Elaboración propia.

b. Diseño de Experimentación

En este caso no es necesario el descarte de alguna estrategia en particular, dado que el objetivo es ir rotando todas las estrategias en todas las posiciones posibles. Por tanto lo que se hace es, en función del orden original que está representado por el grupo control, ir desplazando las estrategias una posición más abajo por cada grupo que se conforma.

En este caso se consideran sólo 4 posiciones, ya que las dos recomendaciones de la estrategia Frecuente se consideran como un conjunto que utiliza una sola posición. Quedando la siguiente distribución final por grupos:



Ilustración 10. Distribución estrategias por grupo Experimento 2. Elaboración propia

#### 4.2.4.3. Experimento 3: Reemplazo de estrategia con peor efectividad por nueva

La estrategia que se baraja para reemplazar a la estrategia de Zonas Calientes, actualmente la estrategia con peor efectividad, corresponde a una estrategia que haga uso de filtros colaborativos. En específico, filtros colaborativos que estén centrados en los productos, dada la gran cantidad de clientes con respecto a la cantidad de productos.

##### a. Grupo de Control y Experimentación

La conformación de los grupos continúa con la misma lógica de los experimentos anteriores, siendo el eje central la efectividad promedio entre grupos.

Debido a que se evaluará el cambio de estrategia solo para la con peor efectividad, los grupos para el experimento se conformarán de la siguiente forma:



Ilustración 11. Grupos experimento 3. Elaboración Propia

Según esta estructura, contamos con un grupo de control el cual no se verá expuesto al experimento a realizar (Grupo 1) y por otro lado un grupo 2 que tendrá en lugar de la estrategia de Zona Caliente, la estrategia de Filtros Colaborativos basado en productos

desarrollada. Donde la evaluación se realizará en comparación a los resultados obtenidos en el grupo de control.

Estos grupos están compuestos por una cantidad similar de clientes y con efectividades similares. A continuación se muestra en detalle dicha conformación:

i. Cantidad de Clientes

Al igual que en los experimentos anteriores existen restricciones en el número de clientes a los cuales se les puede realizar este tipo de experimentos por temas de costos e imagen de empresa. Según esto, se determina la siguiente cantidad de clientes por grupo:



Ilustración 12. Cantidad de clientes por grupo experimento 3. Elaboración propia.

ii. Efectividad de Clientes

El objetivo nuevamente es determinar las variaciones en la efectividad según los distintos cambios realizados en el experimento, por lo que es necesario asegurar que las efectividades promedio sean comparables.



Ilustración 13. Efectividad promedio por grupo experimento 3. Elaboración propia.

#### b. Diseño de Experimentación

Dado que se quiere ingresar una nueva recomendación y el espacio de visualización es acotado, se debe reemplazar una estrategia actual. En este caso el objetivo era mejorar con esta estrategia de filtro colaborativo la efectividad del motor de recomendaciones, reemplazando la peor estrategia dentro de este. Por tanto, la estrategia a reemplazar será la de Zona Caliente. Quedando por tanto la nueva estrategia en la última posición, quedando la distribución de la siguiente forma por grupo:



Ilustración 14. Distribución estrategias experimento 3. Elaboración propia

#### 4.2.5. Evaluación de resultados

Se realizaron experimentos durante los meses de Diciembre de 2016 y Enero de 2017, cargándose las recomendaciones utilizadas por semana cada domingo anterior. En específico, dado factores técnicos de Alicorp las semanas válidas para evaluación de los experimentos corresponden a:

- Semana 26 de Diciembre del 2016
- Semana 02 de Enero del 2017
- Semana 09 de Enero del 2017
- Semana 23 de Enero del 2017

Los resultados se determinan en función de los grupos anteriormente definidos, sin embargo existen fluctuaciones entre la participación real de estos en los experimentos. Esto dado que existen semanas donde estos pueden no ser expuestos a las recomendaciones, aun cuando se han cargado estas en la herramienta, debido a decisiones de los vendedores u otro tipo de situaciones que no permiten que se realice el proceso de venta completo. Sin embargo, para el cálculo de cada uno de los resultados se asegura que existen una cantidad suficiente y significativa de clientes a evaluar.

#### 4.2.5.1. Resultados Experimento 1. Variación en el número de productos recomendados por estrategia

- i. Semana 26 de Diciembre:

Grupo	Efectividad Promedio	Desviación Estándar
Control	5,6%	19%
Variación Perfil	5,7%	17,9%
Variación Asociado	4,8%	12,1%

Tabla 8. Resultados Experimento 1 26 de Diciembre. Elaboración Propia

- ii. Semana 02 de Enero del 2017

Grupo	Efectividad Promedio	Desviación Estándar
Control	3,9%	13,5%
Variación Perfil	4,1%	13,7%
Variación Asociado	5,0%	13,8%

Tabla 9. Resultados Experimento 1 02 de Enero. Elaboración Propia

- iii. Semana 09 de Enero del 2017

Grupo	Efectividad Promedio	Desviación Estándar
Control	5,5%	15,8%
Variación Perfil	6%	18,3%
Variación Asociado	5,9%	15,5%

Tabla 10. Resultados Experimento 1 09 de Enero. Elaboración Propia

- iv. Semana 23 de Enero del 2017

Grupo	Efectividad Promedio	Desviación Estándar
Control	5,6%	14,5%

Variación Perfil	4%	14,2%
Variación Asociado	7,3%	17,6%

Tabla 11. Resultados Experimento 1 23 de Enero. Elaboración Propia

A simple vista se puede evidenciar que existe mucha fluctuación entre las efectividades por grupo, dada la alta desviación que se calcula en cada uno de los experimentos. Por otro lado, se observa que no existe un patrón fijo que permita determinar si algún tipo de variación permite generar mejores resultados, dado que en las distintas evaluaciones se ha encontrado que hasta el grupo control presenta una mayor efectividad respecto a las otras variaciones.

Finalmente por el lado estadístico, podemos afirmar que no existen diferencias significativas en la efectividad promedio de ninguno de estos experimentos por lo que no se puede afirmar que la variación pueda mejorar o empeorar el resultado del motor de recomendaciones.

v. Resultado Global

Grupo	Efectividad Promedio	Desviación Estándar
Control	5,1%	15,8%
Variación Perfil	4,9%	16,1%
Variación Asociado	5,7%	14,9%

Tabla 12. Resultado Experimento 1. Elaboración Propia

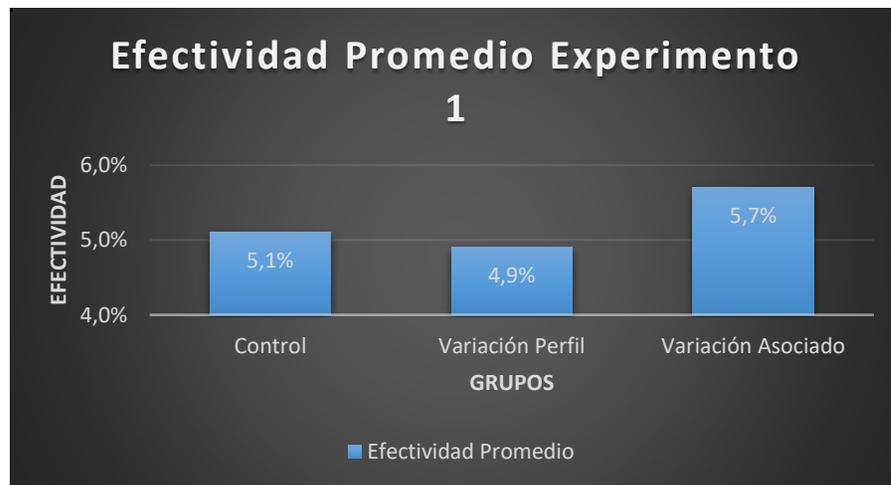


Gráfico 15. Efectividad Promedio Experimento 1. Elaboración Propia

A nivel global, al realizar un test de ANOVA de una vía se determina con un 95% de confianza que no existe relación entre la cantidad de recomendaciones mostradas por estrategia y la efectividad promedio de estas. Por tanto las variaciones de efectividad pueden deberse a factores inherentes a cada estrategia, o a factores situaciones particulares de los grupos del experimento o de la fecha en particular.

De todas formas este análisis no es concluyente debido a que también puede ser afectado por factores de posición en las estrategias desplegadas.

#### 4.2.5.2. Resultados Experimento 2. Variación de Posición de Estrategias

##### i. Resultado Global Estrategia Asociados

Grupo	Efectividad Promedio	Desviación Estándar
Control (Posición 3)	4,5%	14,3%
Grupo 1 (Posición 4)	4,8%	13,8%
Grupo 2 (Posición 1)	5,1%	16,5%
Grupo 3 (Posición 2)	2,3%	10,1%

Tabla 13. Resultados Experimento 2 Estrategia Asociados. Elaboración Propia

Se determinan que existen diferencias significativas en función de los resultados de un test de ANOVA de un factor, dado esto podemos observar que la mejor posición en términos de efectividad promedio para la estrategia asociados corresponde a la primera posición. Sin embargo, existen posiciones que se acercan bastante a estos resultados como la tercera y cuarta posición por tanto no se puede hacer una afirmación definitiva acerca de cuál es la posición óptima para esta estrategia.

En específico, lo que más destaca es que la cuarta posición refleja una alta efectividad para esta estrategia, lo que es totalmente contrario a la intuición que se tiene de que ubicar la estrategia en la última posición tiene una implicancia negativa sobre la efectividad de la misma.

##### ii. Resultado Global Estrategia Perfil

Grupo	Efectividad Promedio	Desviación Estándar
<b>Control (Posición 2)</b>	3,8%	15,0%
<b>Grupo 1 (Posición 3)</b>	4,9%	16,2%
<b>Grupo 2 (Posición 4)</b>	1,7%	10,1%
<b>Grupo 3 (Posición 1)</b>	5,6%	15,3%

Tabla 14. Resultados Experimento 2 Estrategia Perfil. Elaboración Propia

En el caso de la estrategia perfil, también existen diferencias significativas en función de los resultados de un test de ANOVA de un factor. En función de esto podemos afirmar que existe efecto de la posición sobre la efectividad promedio de esta estrategia siendo, en función de los resultados de los experimentos, la primera posición la que entrega una mayor efectividad.

Por otra parte, se puede observar que la segunda y tercera posición son las que le siguen en términos de efectividad. Algo no muy intuitivo dado que se esperaba que la efectividad fuera disminuyendo a la vez que se va desplazando la recomendación a una posición menor.

### iii. Resultado Global Estrategia Frecuentes

Grupo	Efectividad Promedio	Desviación Estándar
Control (Posición 1)	5,1%	16,8%
Grupo 1 (Posición 2)	1,3%	8,7%
Grupo 2 (Posición 3)	4,8%	15,5%
Grupo 3 (Posición 4)	3,3%	12,7%

Tabla 15. Resultados Experimento 2 Estrategia Frecuente. Elaboración Propia

Para la estrategia frecuente, también existen diferencias significativas en función de los resultados de un test de ANOVA de un factor. Por lo que se puede afirmar que existe un efecto de la posición sobre la efectividad de dicha estrategia, siendo nuevamente la primera posición la que presenta una mejor efectividad.

En relación a las otras posiciones, se destaca que la segunda posición es la que presenta menor efectividad en un nivel de magnitud bastante importante. Siendo una hipótesis que esto pueda deberse a que se presenten efectos cruzados entre la posición y la cantidad de recomendaciones que presenta esta estrategia.

### iv. Resultado Global Estrategia Zonas Calientes

Grupo	Efectividad Promedio	Desviación Estándar
Control (Posición 5)	1,7%	11,9%
Grupo 1 (Posición 1)	4,4%	14,8%
Grupo 2 (Posición 2)	3,5%	14,6%
Grupo 3 (Posición 3)	5,8%	17%

Tabla 16. Resultados Experimento 2 Estrategia Zonas Calientes. Elaboración Propia

Se determina que si existen diferencias significativas en la efectividad promedio, según la posición en la cual la estrategia es desplegada en el PAD del vendedor. Esto se condice con la intuición de que la estrategia de Zona Caliente tiene una baja efectividad debido a su posición en el despliegue de estrategias de la herramienta.

Se destaca que, a diferencia de las demás estrategias, la posición con una mejor efectividad promedio no es la primera posición sino que corresponde a la tercera posición. Siendo bastante mayor en magnitud que la estrategia con un segundo mayor resultado (diferencia de 1,4% en efectividad)

A nivel general, se puede observar que todas las estrategias se ven afectadas por el factor posición dentro de la visualización en el PAD. Sin embargo, los resultados no indican un patrón por el cual determinar qué posición es mejor que otra o cómo debieran estar distribuidas estas según el tipo de estrategia. Por tanto, para determinar esto se recurre a realizar una regresión lineal tomando como variable dependiente la efectividad promedio y por como variables independientes la posición y la estrategia detrás de ese resultado. Quedando de la siguiente forma:

$$Efectividad = \beta_0 + \beta_1 X_{Perfil} + \beta_2 X_{Frecuente} + \beta_3 X_{ZonasCalientes} + \beta_4 Z_{1ra\ Posición} + \beta_5 Z_{2da\ Posición} + \beta_6 Z_{3ra\ Posición}$$

Siendo las variables X's, las dummies correspondientes a las estrategias y las variables Z's las dummies que indican la posición. Se debe destacar que el nivel base corresponde a la estrategia Asociados junto a la 4ta posición, cuyo efecto debiera estar indicado por  $\beta_0$ . Siendo los resultados de esta regresión, los indicados en la siguiente tabla:

```
. regress efectividad perfil frecuente zonascalie~s primera segunda tercera
```

Source	SS	df	MS	Number of obs =	15656
Model	1.85191191	6	.308651984	F( 6, 15649) =	15.45
Residual	312.628092	15649	.019977512	Prob > F =	0.0000
				R-squared =	0.0059
				Adj R-squared =	0.0055
				Root MSE =	.14134
Total	314.480004	15655	.020088151		

efectividad	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
perfil	-.0011809	.0031666	-0.37	0.709	-.0073878 .005026
frecuente	-.0051919	.0031215	-1.66	0.096	-.0113104 .0009266
zonascalie~s	-.0093142	.0032408	-2.87	0.004	-.0156665 -.0029619
primera	.019822	.0032113	6.17	0.000	.0135274 .0261166
segunda	-.0041711	.003374	-1.24	0.216	-.0107845 .0024423
tercera	.0167498	.003198	5.24	0.000	.0104814 .0230183
_cons	.0341628	.0029643	11.52	0.000	.0283525 .0399732

Tabla 17. Resultados Regresión Lineal Estrategia y Posición. Programa Stata 11

Podemos observar que existen variables que no son significativas en relación al efecto que tienen sobre la efectividad, siendo estas la variable que representa la estrategia perfil y la variable que indica la segunda posición. Pudiendo deberse esta falta de significancia a la falta de datos suficientes para incorporarlos al modelo.

En base a estos resultados se puede determinar la siguiente tabla con los efectos de las estrategias y posiciones restantes en la efectividad promedio:

	Primera Posición	Tercera Posición	Cuarta Posición
Frecuente	4,8%	4,5%	2,9%
Zonas Calientes	4,4%	4,1%	2,5%
Asociados	5,3%	5,0%	3,4%

Tabla 18. Efecto Posición y Estrategia sobre Efectividad Promedio

En definitiva, según los resultados de la regresión, podemos afirmar que la estrategia con mayor efectividad corresponde a la de Asociado. Sin embargo, esta tabla no permite generar un listado de visualización óptima para la herramienta, dado que por un lado no es un modelo que se ajuste mucho a la situación real (R-squared bajo) y por otro no existen referencias claras acerca del efecto de la segunda posición y la estrategia perfil.

### 4.2.5.3. Resultados Experimento 3: Variación Filtros Colaborativos

i. Semana 26 de Diciembre del 2016

Grupo	Efectividad Promedio	Desviación Estándar
Control	0%	0%
Filtro Colaborativo	0%	0%

Tabla 19.Resultados Experimento 3 26 de Diciembre. Elaboración Propia

ii. Semana 02 de Enero del 2017

Grupo	Efectividad Promedio	Desviación Estándar
Control	0,8%	7,9%
Filtro Colaborativo	0,6%	7,7%

Tabla 20.Resultados Experimento 3 2 de Enero. Elaboración Propia

iii. Semana 09 de Enero del 2017

Grupo	Efectividad Promedio	Desviación Estándar
Control	0,9%	7,9%
Filtro Colaborativo	0,6%	7,6%

Tabla 21.Resultados Experimento 3 9 de Enero. Elaboración Propia

iv. Semana 23 de Enero del 2017

Grupo	Efectividad Promedio	Desviación Estándar
Control	0,5%	4,1%
Filtro Colaborativo	0,6%	5,4%

Tabla 22.Resultados Experimento 2 23 de Enero. Elaboración Propia

En general, para cada uno de los experimentos no se encontraron diferencias significativas entre las efectividades promedio de la estrategia Zonas Calientes y la nueva estrategia de Filtros Colaborativos. A excepción del experimento realizado el 23 de Enero, donde se encuentra que existen diferencias significativas, con un 90% de confianza, en las efectividades de las dos estrategias antes mencionadas.

Se debe señalar, sin embargo, que la realización de estos experimentos presentó diversos problemas debido a la poca efectividad presentada por las estrategias a evaluar. A pesar de que en la construcción de los grupos, se aseguró que tuvieran un nivel de efectividad suficiente en función de los periodos anteriores al experimento.

v. Resultado Global

Grupo	Efectividad Promedio	Desviación Estándar
Control	0,5%	6%

Filtro Colaborativo	0,4%	6%
---------------------	------	----

Tabla 23. Resultados Experimento 3. Elaboración Propia

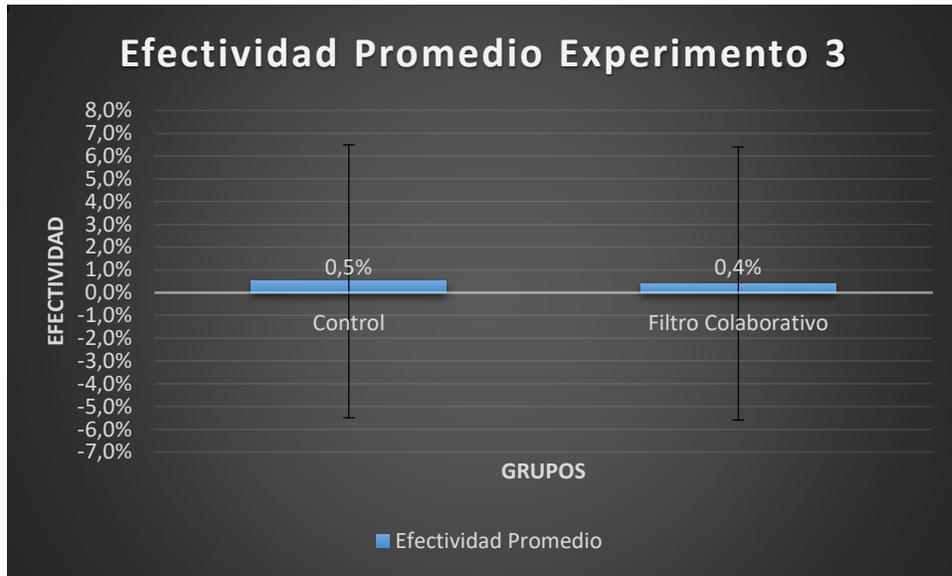


Gráfico 16. Efectividad Promedio Experimento 3. Elaboración Propia

Cómo era de esperar, en función de los resultados individuales, se obtuvo que no existen diferencias significativas entre las efectividades de la estrategia actual (Zonas Calientes) y la nueva estrategia (Filtro Colaborativo). Sin embargo, no se puede afirmar que esto se deba a una mala estructuración de la nueva estrategia o a que las recomendaciones carecen de sentido para el cliente, dado que existen una serie de otros factores cómo la posición y la poca efectividad de los clientes evaluados los que puedan estar afectando a este resultado.

## 5. Conclusiones

El cálculo de venta incremental siempre será un tema difícil de manejar, dado que siempre existirán factores externos que pueden influenciar su valor (por ejemplo el crecimiento de mercado). Siendo esta la razón del nivel de complejidad de la fórmula que utiliza Alicorp actualmente. Sin embargo, es fundamental tener una mirada completa del proceso de venta asociado y ser capaz de ajustar el diseño en base a este proceso. En el caso del Pedido Sugerido, fue innecesario complicar una fórmula de tal forma dado que siempre se realizó un servicio Post Venta, es decir, un proceso extra fuera del proceso de venta regular del vendedor. Ya que debido a esta estructura, se debe considerar toda la venta del Pedido Sugerido como un incremental directo a las ventas regulares de Alicorp.

Una consecuencia importante de esta complejidad es la subestimación que se está realizando del valor que entrega el servicio, dado que con la fórmula propuesta se observa una diferencia de USD 6.034.127 respecto a la fórmula actual. Sin embargo, la implementación de esta nueva evaluación es compleja dado que siempre la parte contratante querrá realizar modificaciones que hagan que esta venta incremental, y por tanto el valor del servicio, sea menor para su conveniencia.

Con respecto al desarrollo del modelo de filtro colaborativo basado en productos, se puede concluir que si bien tenía un índice de *precision* aceptable (10%). Siempre se debe considerar factores externos que pudieran afectar su desempeño. En este caso en particular, se pudo observar que la implementación en reemplazo de la estrategia de Zonas Calientes tuvo mucho que ver en los resultados. Esto dado que a través de otro experimento se determinó que existe un efecto en la efectividad por parte de la posición en que se despliega en la pantalla del vendedor. Por lo tanto era razonable que los resultados fueran parecidos al grupo de control.

En relación al resultado de posición, podemos afirmar que intuitivamente tiene mucho sentido. Dado que es natural que el producto recomendado en primera posición tenga más probabilidades de ser recomendado por el vendedor que el producto ubicado en la última posición, dado que existe un proceso de venta con un tiempo acotado y un cliente al cual no se quiere agotar con una sobreexposición a recomendaciones. Sin embargo, la posición óptima para cada una de las estrategias no es directa dado que existen efectos distintos sobre la efectividad para cada una de ellas. Por lo que es necesario utilizar la información entregada por la regresión para generar un orden óptimo de recomendaciones, en el caso de las variables que son significativas.

Por otro lado, también se determinó que a pesar de la intuición el número de recomendaciones por estrategia no afectan directamente su efectividad. Descartando que la alta efectividad de la estrategia frecuente se deba a que posea dos recomendaciones en la pantalla de recomendación. De todas formas, estos resultados pueden estar afectados por el factor de posición descubierto recientemente o la calidad de las recomendaciones de las otras estrategias.

## 6. Recomendaciones y Trabajo Futuro

Es recomendable que la utilización de la nueva metodología sea implementada en la empresa lo más pronto posible, primero debido a que simplifica el proceso de cálculo de las ventas incrementales y segundo dado que permite clarificar de cara al cliente cual es el real valor de la herramienta, aumentando en consecuencia el poder de negociación ante cualquier clase de modificaciones o requerimientos que puedan surgir por parte del cliente.

En relación a la visualización de las estrategias dentro de la herramienta, se recomienda realizar una mayor cantidad de experimentos con tal de recolectar más datos que permitan aumentar el grado de precisión del modelo de regresión lineal. Buscando así obtener un despliegue de recomendaciones óptimas que potencie la efectividad de las recomendaciones al máximo.

En cuanto al trabajo futuro, se cree que el foco en las mejoras debiera orientarse en términos de mejorar las estrategias que actualmente están funcionando en la herramienta. Esto dado que se observó que muchas de las recomendaciones por estrategia se repetían para un mismo cliente, causando que dado la estructura de visualización algunas de las estrategias quedarán con recomendaciones con una menor probabilidad de compra y en consecuencia una menor efectividad. Por tanto debiera trabajarse en generar estrategias en función de modelos que no se relación de sobremanera.

Otra línea de trabajo pudiera ser complementar el motor de recomendaciones con la incorporación de pricing y recomendaciones de promociones. Esto permitiría, en el caso del pricing, maximizar la venta incremental por oportunidad de compra y, en el caso de las promociones, aumentar esta probabilidad de compra.

## 7. Bibliografía

- [1] SOTOMAYOR, N. CLAUDIA. 2008. Definición de un Sistema de Recomendación para una Empresa de Retail. Memoria de Ingeniero Civil Industrial. Santiago. Universidad de Chile. Facultad de Ciencias Físicas y Matemáticas.
- [2] NALDA, JOSÉ ANTONIO. 2010. Sistema Secuencial de Recomendaciones personalizadas en una Empresa de Home Improvement. Memoria de Ingeniero Civil Industrial. Santiago. Universidad de Chile. Facultad de Ciencias Físicas y Matemáticas.
- [3] FLORÁS, DIEGO. 2015. Modelo de Recomendación de Productos Aplicado a una Empresa de Cupones Online. Memoria de Ingeniero Civil Industrial. Santiago. Universidad de Chile. Facultad de Ciencias Físicas y Matemáticas.
- [4] HERLOCKER, J. L., KONSTAN, J. A., & RIEDL, J. (2000, December). Explaining collaborative filtering recommendations. In Proceedings of the 2000 ACM conference on Computer supported cooperative work (pp. 241-250). ACM.
- [5] LAMBRECHT, A., & TUCKER, C. E. (2015). Field Experiments in Marketing.
- [6] GALLINO, S., & MORENO, A. (2014). Integration of online and offline channels in retail: The impact of sharing reliable inventory availability information. *Management Science*, 60(6), 1434-1451.
- [7] ABRAHAM, M. M., & LODISH, L. M. (1993). An implemented system for improving promotion productivity using store scanner data. *Marketing Science*, 12(3), 248-269.
- [8] WOOLDRIDGE, J. (2007). What's new in econometrics? Lecture 10 difference-in-differences estimation. NBER Summer Institute, available at: [www.nber.org/WNE/Slides7-31-07/slides\\_10\\_diffindiffs.pdf](http://www.nber.org/WNE/Slides7-31-07/slides_10_diffindiffs.pdf), accessed April, 9, 2011.
- [9] GÚZMAN, PAMELA. 2011. Metodología de Evaluación de Promociones en un Supermercado. Memoria de Ingeniero Civil Industrial. Santiago. Universidad de Chile. Facultad de Ciencias Físicas y Matemáticas.
- [10] ANGUITA, FRANCISCO. 2008. Evaluación de Promociones en Tiendas de Conveniencia. Memoria de Ingeniero Civil Industrial. Santiago. Universidad de Chile. Facultad de Ciencias Físicas y Matemáticas.
- [11] MUÑOZ, JAVIER. 2008. Metodología para el Diseño y Evaluación de Promociones Focalizadas para Supermercados. Memoria de Ingeniero Civil Industrial. Santiago. Universidad de Chile. Facultad de Ciencias Físicas y Matemáticas.
- [12] BARBARA, FUENZALIDA. 2012. Pronóstico De Efectividad De Promociones Sobre Clientes, Dadas Sus Características Y Respuestas Pasadas. Memoria de Ingeniero Civil Industrial. Santiago. Universidad de Chile. Facultad de Ciencias Físicas y Matemáticas.

[13] SU, X., & KHOSHGOFTAAR, T. M. (2006, November). Collaborative filtering for multi-class data using belief nets algorithms. In 2006 18th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'06) (pp. 497-504). IEEE.

## 8. Anexos

### Anexo A: Análisis Descriptivo de Clientes Alicorp

#### i. Movimiento de Clientes por períodos

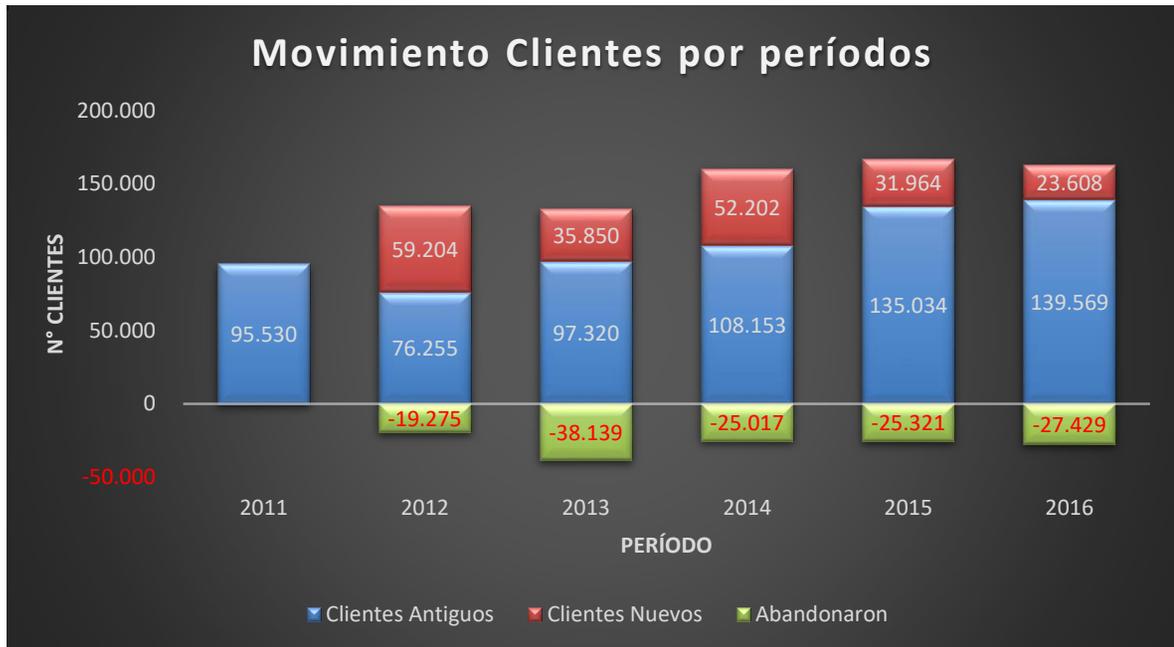


Gráfico 17. Movimiento clientes por períodos. Elaboración propia

#### ii. Número de clientes por perfil año 2016

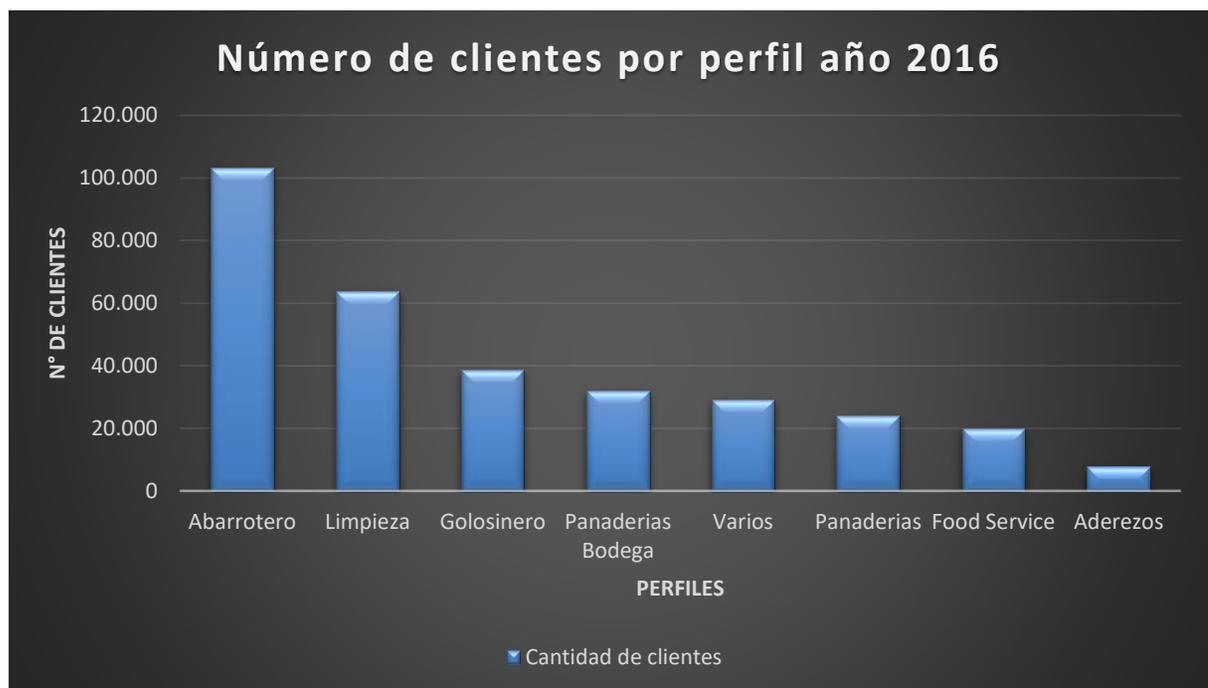


Gráfico 18. Cantidad de clientes por perfil año 2016. Elaboración propia.

iii. Número de productos por perfil de cliente



Gráfico 19. Número de productos por perfil. Elaboración propia

iv. Variación Ventas % por tipo de clientes

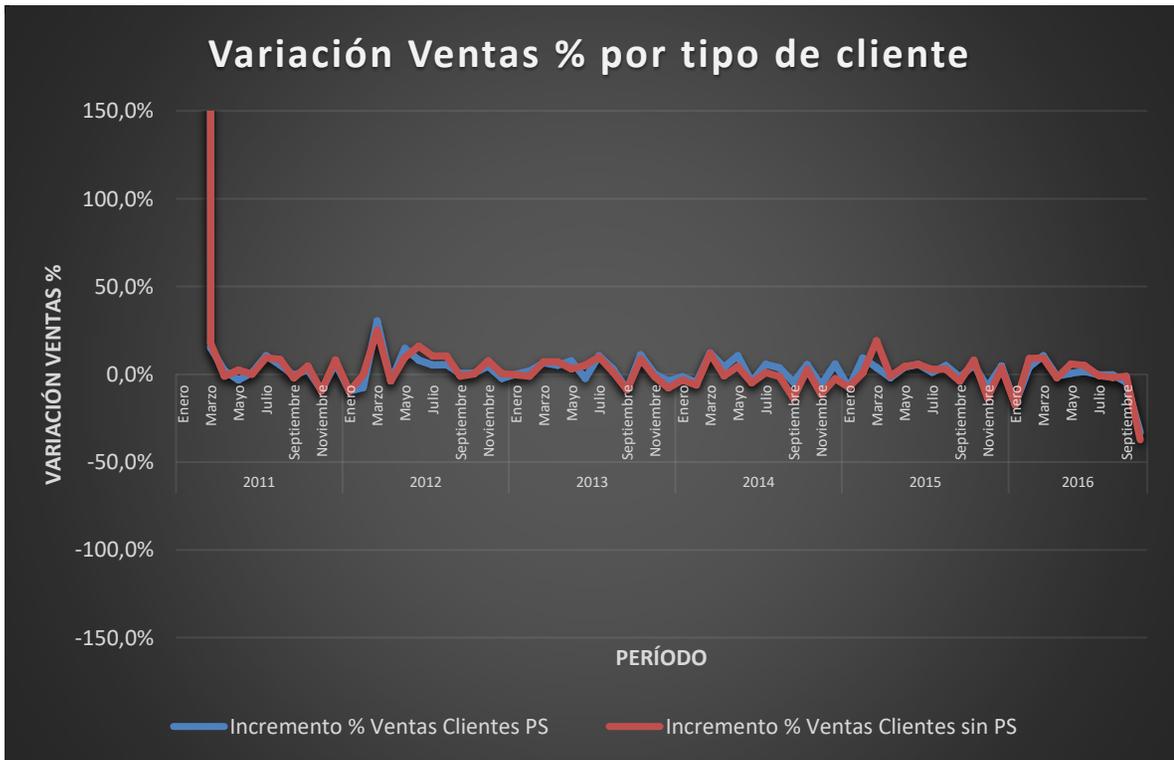


Gráfico 20. Variación de ventas % por tipo de cliente. Elaboración propia.

## Anexo B: Análisis Descriptivo de Productos y Vendedores

### i. Cantidad de vendedores Anuales



Gráfico 21. Cantidad de vendedores anuales. Elaboración propia

### ii. Cantidad de Productos Anuales

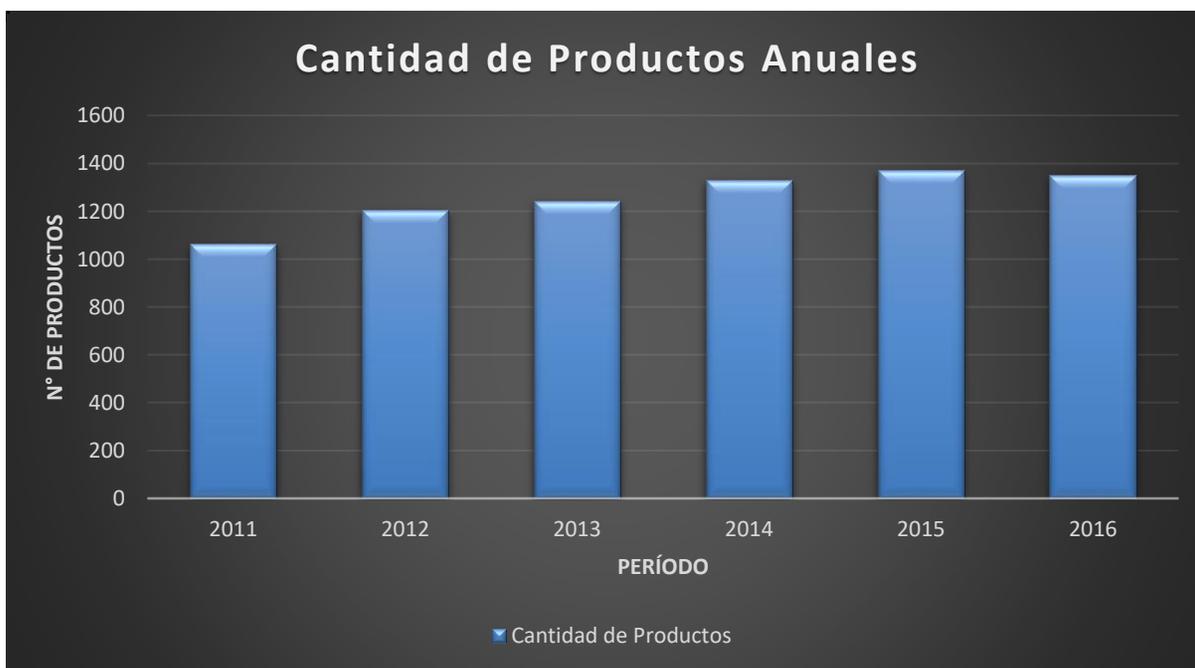


Gráfico 22. Número de productos anuales. Elaboración propia

iii. Incremento de Venta Mensual por Tipo de Producto

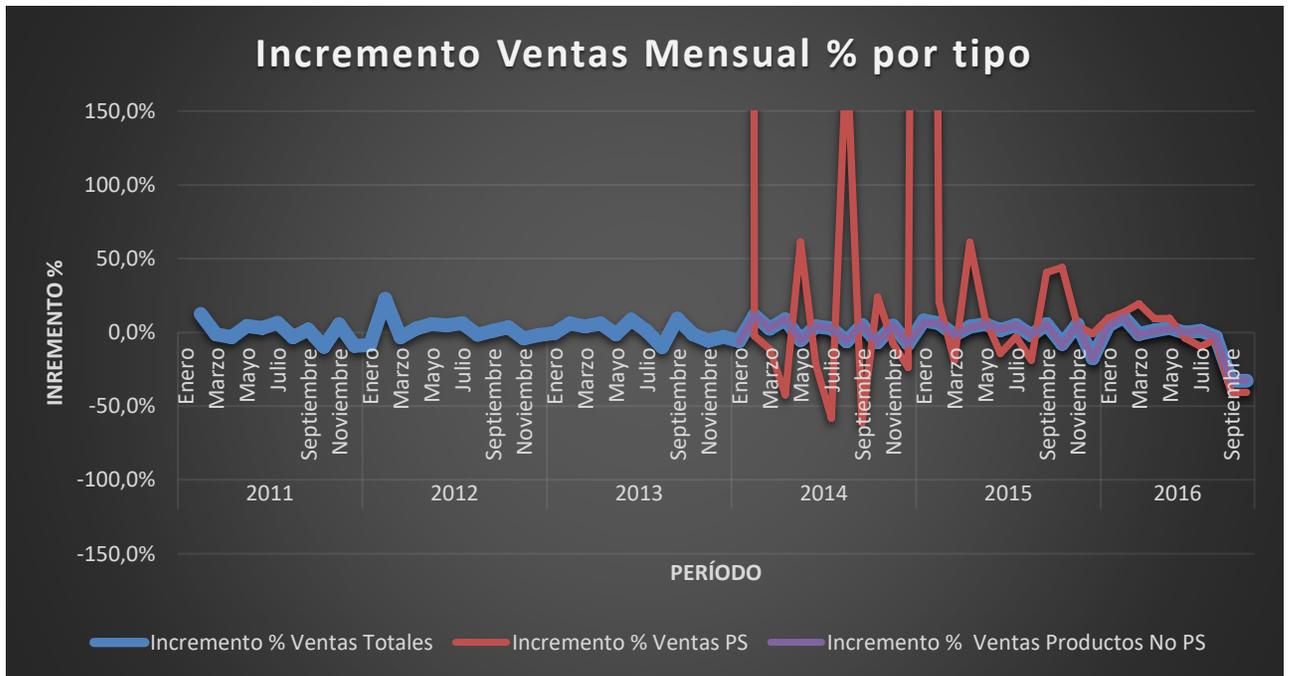


Gráfico 23. Incremento % de ventas mensual por tipo de venta. Elaboración propia.

Anexo C: Análisis de Efectividad de Estrategias

i. Efectividad de Estrategia Asociados

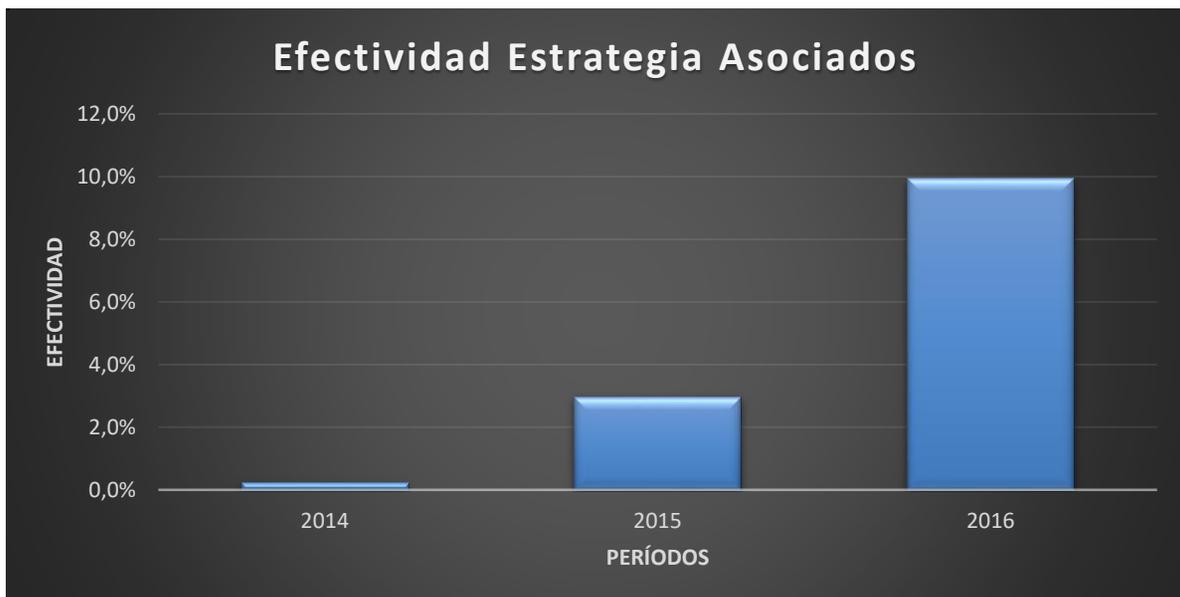


Gráfico 24. Efectividad estrategia asociados. Elaboración propia.

ii. Efectividad de Estrategia Frecuentes

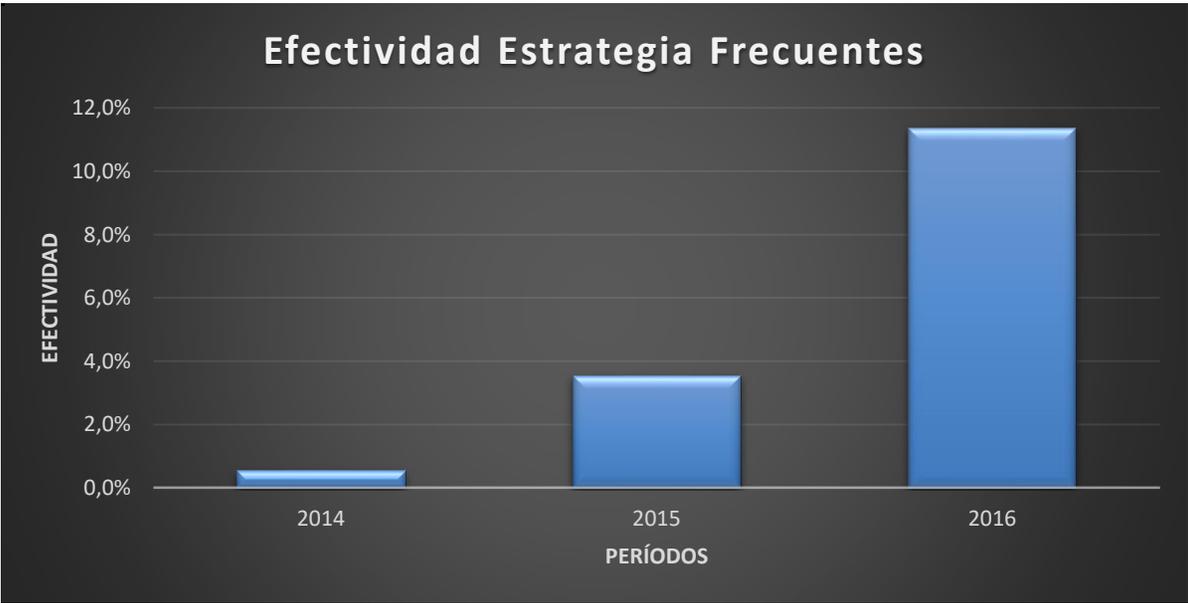


Gráfico 25. Efectividad estrategia frecuentes. Elaboración propia.

iii. Efectividad de Estrategia Perfil



Gráfico 26. Efectividad estrategia perfil. Elaboración propia.

iv. Efectividad de Estrategia Zonas Calientes



Gráfico 27. Efectividad estrategia zonas calientes. Elaboración propia.

## Anexo D: Metodologías Alternativas

### 1. Diferencias en Diferencias

En general, los denominados métodos observacionales se conforman a partir de la distribución de los individuos a evaluar en dos grupos, siendo uno integrado por individuos que han sido expuestos a los efectos de un evento y otro que lo conforman los individuos que no han respondido a este efecto. Teniendo en consideración que todos los individuos son idealmente similares, pero esto en la práctica no siempre es posible, teniendo variables contaminantes no observables. Entrando aquí los estimadores del método “Diferencias en Diferencias”

Lo que hace este método cuasi-experimental es básicamente medir el efecto de un tratamiento o evento dentro de dos períodos de tiempo.

Su configuración más simple considera solo resultados observables para dos grupos en dos periodos distintos. En estos grupos, el primero es expuesto al efecto que se quiere medir durante el segundo periodo, pero no así durante el primer periodo (Grupo tratamiento). Mientras que el segundo grupo no es expuesto durante ninguno de los dos periodos a evaluar (Grupo control). A continuación, en el caso de que se cuente con las mismas unidades observables dentro de un grupo en cada período, se procede a sustraer el beneficio promedio del grupo control de la ganancia promedio del grupo tratamiento. Esto quitara el sesgo en la comparación del segundo periodo entre los grupos de control y de tratamiento, que pueda resultar por la diferencia permanente de estos grupos.

Dado esto, podemos escribir el modelo genérico para cualquier grupo como:

$$y = \beta_0 + \beta_1 dB + \delta_0 d2 + \delta_1 d2 \cdot dB + u$$

Donde  $y$  es el resultado de interés,  $d2$  es una variable dummy para el segundo periodo. La variable dummy  $dB$  captura las posibles diferencias entre el grupo control y el tratado antes de someterse al evento estudiado. La variable dummy de tiempo,  $d2$ , capturar los efectos agregados que puedan causar cambio en  $Y$  aún en ausencia del evento a estudiar. Por otro lado el coeficiente,  $\delta_1$ , multiplica el termino de interacción,  $d2 \cdot dB$ , que representa lo mismo que una variable dummy igual a uno para aquellas observaciones el grupo de tratamiento en el segundo periodo.

Siendo  $\delta_1$  el estimador de diferencias en diferencias:

$$\widehat{\delta}_1 = (\bar{y}_{B,2} - \bar{y}_{B,1}) - (\bar{y}_{A,2} - \bar{y}_{A,1})$$

## 2. Método de Rubin

El Modelo Causal de Rubin, es un modelo que utiliza el concepto de Propensity Score para realizar estimaciones de los efectos que un evento determinado tendrá sobre un resultado.

El denominado Propensity Score corresponde a una variable que resume toda la información que entrega un set de variables caracterizados de un individuo. Por tanto, lo que esta variable logra es relacionar individuos con variables características distintas pero que dado su Propensity Score similar, se puede afirmar que las distribuciones de estas variables características son iguales en todas ellas.

A nivel general, este modelo que intenta determinar efectos causales comparando resultados potenciales de *clientes tratados* y *clientes de control*, a través de la reconstrucción del experimento aleatorio hipotético que dio origen a los datos observacionales con los que se cuenta.

La reconstrucción del experimento aleatorio se realiza de la siguiente forma:

- Se recopila la base de datos de los resultados observados tanto para *clientes tratados*, como para *clientes de control*.
- Se identifican las variables que el tomador de decisión consideró para obtener los datos observacionales.
- Utilizando un modelo predictivo, se calcula la probabilidad de que el tomador de decisión asigne la intervención a cada cliente, independiente de su tipo, a partir de las variables identificadas anteriormente.
- A través de un modelo de matching, se buscan pares de clones *cliente tratado-cliente de control* que posean igual Propensity Score.
- Se calcula la diferencia de resultados entre clones y se genera un ranking de respuestas a la intervención.

Dado este proceso es posible calcular la magnitud del efecto que un evento tiene sobre un conjunto de clientes diverso de manera insesgada.

## Anexo E: Criterios de Similitud

### Similitud de Usuarios

Existen distintas métricas para determinar similitud de usuarios, a continuación, se mostrarán las principales:

i. Distancia Euclidiana:

Esta es conocida por medir la distancia entre dos puntos en un plano mediante la siguiente fórmula:

$$d(X, Y) = \sqrt{\sum_{i=1}^N (x_i - y_i)^2}$$

ii. Distancia Euclidiana Cuadrática:

La Distancia Euclidiana Cuadrática corresponde, como lo dice su nombre, al cuadrado del valor de la Distancia Euclidiana. Su valor es calculado mediante la siguiente fórmula:

$$d(X, Y) = \sum_{i=1}^N (x_i - y_i)^2$$

iii. Distancia Manhattan:

La Distancia Manhattan también es conocida por medir dos puntos en un plano, pero no calculando la distancia diagonal entre estos (Distancia Euclidiana), sino que en distancias horizontales o verticales. Esta distancia es calculada mediante la siguiente fórmula:

$$d(X, Y) = \sum_{i=1}^N |x_i - y_i|^2$$

iv. Distancia Coseno:

Para utilizar la Distancia Coseno, es necesario pensar en cada calificación como un punto en un plano, y trazar un vector desde el origen hasta dicho punto, con el fin de formar un ángulo  $\theta$  entre ambos vectores. Para calcular esta distancia se utiliza la siguiente fórmula:

$$d(X, Y) = \frac{\sum_{i=1}^N x_i y_i}{\sum_{i=1}^N x_i^2 \sum_{i=1}^N y_i^2} = \frac{X \times Y}{\|X\| \cdot \|Y\|}$$

v. Distancia Jaccard:

A diferencia de las distancias mencionadas anteriormente, que capturan la distancia o el ángulo entre los distintos usuarios, el Coeficiente de Tanimoto, o también conocida como la Distancia Jaccard, considera la distancia y el ángulo en el cálculo de la similitud.

Se calcula mediante la siguiente expresión:

$$T = \frac{\sum_{i=1}^N (x_i \wedge y_i)}{\sum_{i=1}^N (x_i \vee y_i)}$$

vi. Coeficiente de Correlación de Pearson:

El Coeficiente de Correlación de Pearson mide la tendencia de dos series de números que se mueven en la misma dirección. En otras palabras, si dos usuarios tienen calificaciones similares o simplemente similares proporcionalmente (al igual que la Distancia Coseno), serán similares según esta métrica.

Para calcular este coeficiente, se utiliza la siguiente expresión:

$$d(X, Y) = \frac{N \sum_{i=1}^N x_i y_i - (\sum_{i=1}^N x_i) (\sum_{i=1}^N y_i)}{\sqrt{(N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2) (N \sum_{i=1}^N y_i^2 - (\sum_{i=1}^N y_i)^2)}}$$

## Anexo F: Cantidad de Productos Similares

Para saber cuántos productos son considerados similares a otro, existen dos tipos de medidas, las cuales son descritas a continuación:

i. K-Vecinos más cercanos:

Se utilizan los K-Vecinos más cercanos, cuando se desea obtener un número fijo K de productos más similares a otro. Es una buena opción si se quiere obtener una cantidad particular de usuarios similares, pero pierde fuerza cuando se necesita que todos los usuarios tengan un cierto grado de similitud.

ii. Vecinos cercanos dependientes de la métrica de similitud:

Esta cantidad de usuarios es utilizada cuando se desea obtener usuarios similares dependiendo de la similitud que se emplee. Es por esto que, dependiendo de la métrica de similitud, se debe elegir cómo determinar la cantidad de usuarios que se considerarán como similares.

Por lo tanto, es conveniente utilizar esta forma de determinar la cantidad de usuarios similares cuando se sabe que existen muchos usuarios similares entre sí, ya que, en caso contrario, puede que para muchos de estos no encuentre ningún vecino más cercano.

## Anexo G: Test ANOVA de Grupos Control y Tratamiento

### i. Test ANOVA Experimento 1

. oneway efectividad tipo\_grupo

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	.002217187	3	.000739062	0.81	0.4883
within groups	1.45621233	1596	.000912414		
Total	1.45842951	1599	.000912089		

Bartlett's test for equal variances:  $\chi^2(3) = 3.1525$  Prob> $\chi^2 = 0.369$

Tabla 24. Test ANOVA Experimento 1. Elaboración Propia

### ii. Test ANOVA Experimento 2

. oneway efectividad tipo\_grupo

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	.001204667	2	.000602333	0.68	0.5048
within groups	1.05423134	1197	.000880728		
Total	1.05543601	1199	.000880264		

Bartlett's test for equal variances:  $\chi^2(2) = 0.7776$  Prob> $\chi^2 = 0.678$

Tabla 25. Test ANOVA Experimento 2. Elaboración Propia

### iii. Test ANOVA Experimento 3

. oneway efectividad tipo\_grupo

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	.000136125	1	.000136125	0.15	0.7009
within groups	.735567795	798	.000921764		
Total	.73570392	799	.000920781		

Tabla 26. Test ANOVA Experimento 3. Elaboración Propia

## Anexo H: Test ANOVA de Experimento 1

### i. Test ANOVA Experimento 1 Global

. oneway efectividad grupo

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	.039282177	2	.019641089	0.81	0.4451
Within groups	79.1472719	3263	.024255983		
Total	79.1865541	3265	.024253156		

Bartlett's test for equal variances:  $\chi^2(2) = 7.6863$  Prob> $\chi^2 = 0.021$

Tabla 27. Test ANOVA Experimento 1 Global. Elaboración Propia

## ii. Test ANOVA Experimento 1 26 de Diciembre

. oneway efectividad grupo

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	.000049523	1	.000049523	0.00	0.9696
Within groups	16.4624687	483	.034083786		
Total	16.4625182	484	.034013467		

Bartlett's test for equal variances:  $\chi^2(1) = 0.9020$  Prob> $\chi^2 = 0.342$

Ilustración 15. Test ANOVA Experimento 1 26 de Diciembre. Elaboración Propia

## iii. Test ANOVA Experimento 1 02 de Enero

. oneway efectividad grupo

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	.02213586	2	.01106793	0.59	0.5529
Within groups	15.414527	826	.018661655		
Total	15.4366629	828	.018643313		

Bartlett's test for equal variances:  $\chi^2(2) = 0.1991$  Prob> $\chi^2 = 0.905$

Ilustración 16. Test ANOVA Experimento 1 02 de Enero. Elaboración Propia

## iv. Test ANOVA Experimento 1 09 de Diciembre

. oneway efectividad grupo

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	.00355708	2	.00177854	0.06	0.9373
Within groups	23.0017956	837	.027481237		
Total	23.0053527	839	.027419967		

Bartlett's test for equal variances:  $\chi^2(2) = 9.6827$  Prob> $\chi^2 = 0.008$

Ilustración 17. Test ANOVA Experimento 1 09 Enero. Elaboración Propia

## v. Test ANOVA Experimento 1 23 de Enero

. oneway efectividad grupo

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	.145131125	2	.072565562	2.99	0.0511
Within groups	19.5651555	805	.024304541		
Total	19.7102866	807	.024424147		

Bartlett's test for equal variances:  $\chi^2(2) = 16.5821$  Prob> $\chi^2 = 0.000$

Ilustración 18. Test ANOVA Experimento 1 23 Enero. Elaboración Propia

## Anexo I: Test ANOVA de Experimento 2

### i. Test ANOVA Experimento 2 Asociados

. oneway efectividad grupo

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	.453255112	3	.151085037	7.62	0.0000
Within groups	81.6701907	4117	.019837306		
Total	82.1234458	4120	.019932875		

Bartlett's test for equal variances:  $\chi^2(3) = 213.6555$  Prob> $\chi^2 = 0.000$

Tabla 28. Test ANOVA Experimento 2 Asociados. Elaboración Propia

### ii. Test ANOVA Experimento 2 Perfil

. oneway efectividad grupo

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	.754237355	3	.251412452	11.77	0.0000
Within groups	82.9450723	3882	.021366582		
Total	83.6993096	3885	.021544224		

Bartlett's test for equal variances:  $\chi^2(3) = 195.3185$  Prob> $\chi^2 = 0.000$

Tabla 29. Test ANOVA Experimento 2 Perfil. Elaboración Propia

### iii. Test ANOVA Experimento 2 Frecuentes

. oneway efectividad grupo

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	.813344069	3	.27111469	13.67	0.0000
Within groups	80.8493419	4077	.019830596		
Total	81.662686	4080	.020015364		

Bartlett's test for equal variances:  $\chi^2(3) = 398.7207$  Prob> $\chi^2 = 0.000$

Tabla 30. Test ANOVA Experimento 2 Frecuentes. Elaboración Propia

### iv. Test ANOVA Experimento 2 Zonas Calientes

. oneway efectividad grupo

Source	Analysis of Variance SS	df	MS	F	Prob > F
Between groups	.476007017	3	.158669006	8.53	0.0000
Within groups	66.3190683	3564	.018608044		
Total	66.7950753	3567	.018725841		

Bartlett's test for equal variances:  $\chi^2(3) = 114.7995$  Prob> $\chi^2 = 0.000$

Tabla 31. Test ANOVA Experimento 2 Zonas Calientes. Elaboración Propia

## Anexo J: Test ANOVA de Experimento 3

### i. Test ANOVA Experimento 3 Global

. oneway efectividad grupo

Source	Analysis of Variance SS	df	MS	F	Prob > F
Between groups	.000403322	1	.000403322	0.11	0.7386
Within groups	5.40809344	1494	.003619875		
Total	5.40849676	1495	.003617724		

Bartlett's test for equal variances:  $\chi^2(1) = 0.0118$  Prob> $\chi^2 = 0.913$

Tabla 32. Test ANOVA Experimento 3 Global. Elaboración Propia

### ii. Test ANOVA Experimento 3 02 Enero

. oneway efectividad grupo

Source	Analysis of Variance SS	df	MS	F	Prob > F
Between groups	.000262294	1	.000262294	0.04	0.8358
Within groups	2.23275401	366	.006100421		
Total	2.2330163	367	.006084513		

Bartlett's test for equal variances:  $\chi^2(1) = 0.2008$  Prob> $\chi^2 = 0.654$

Tabla 33. Test ANOVA Experimento 3 02 Enero. Elaboración Propia

### iii. Test ANOVA Experimento 3 09 Enero

. oneway efectividad grupo

Source	Analysis of Variance SS	df	MS	F	Prob > F
Between groups	.000742862	1	.000742862	0.12	0.7268
Within groups	2.33962464	385	.006076947		
Total	2.34036751	386	.006063128		

Bartlett's test for equal variances:  $\chi^2(1) = 0.3720$  Prob> $\chi^2 = 0.542$

Tabla 34. Test ANOVA Experimento 3 09 Enero. Elaboración Propia

### iv. Test ANOVA Experimento 3 23 Enero

. oneway efectividad grupo

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	.000066282	1	.000066282	0.03	0.8632
Within groups	.8224854	369	.002228958		
Total	.822551682	370	.002223113		

Bartlett's test for equal variances:  $\chi^2(1) = 14.6841$  Prob> $\chi^2 = 0.000$

Tabla 35. Test ANOVA Experimento 3 23 Enero. Elaboración Propia