



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

**ESTIMACIÓN DE DEMANDA DE PREMIOS, PARA UN CLUB DE FIDELIZACIÓN
DE ACUMULACIÓN DE PUNTOS**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

IVÁN MANUEL ESPINOZA MELÉNDEZ

**PROFESOR GUÍA:
MÁXIMO BOSCH PASSALACQUA**

**MIEMBROS DE LA COMICIÓN:
LUIS ABURTO LAFOURCADE
ALEJANDRA PUENTE CHANDIA**

SANTIAGO DE CHILE

OCTUBRE 2008

“Estimación de Demanda de Premios, para un club de Fidelización de Acumulación de Puntos”

El objetivo general del presente trabajo de título es desarrollar un modelo de estimación que permita predecir la demanda de premios que se van a canjear en una determinada semana, con el fin de aportar información relevante para que la empresa tome decisiones, en relación al stock adecuado de productos que se deben mantener en las tiendas.

La tienda por departamentos en donde se desarrolló el trabajo posee un club de fidelización, en el cual los clientes acumulan puntos que se pueden canjear por los productos que deseen dentro las posibilidades existentes en un catálogo y al nivel de que logren acceder. Existen 8 niveles y a medida que se acumulan más puntos se va accediendo a los niveles superiores y por lo tanto a mejores productos.

En el trabajo se estimó que la fuga de clientes por insatisfacción del programa de puntos provoca una pérdida mínima de MM\$1.200 anuales en ventas en la multitienda. Una parte de esto, se deberá a los faltantes de premios en las tiendas. Se estima también en el trabajo, un costo de MM\$648 anuales de sobrantes de premios al final de la temporada. Por lo tanto, es necesario un buen control de inventarios y una buena estimación de demanda de canjes.

La estimación de demanda de canje semanal por premios se realizó en tres niveles, estos son, el total de canje, la distribución de canje por nivel de puntaje y dentro de cada nivel de puntaje las categorías genéricas de productos que se canjean. Para estimar los canjes totales, se utilizaron las series de tiempo ARIMAX y suavizado exponencial. La estimación por nivel de puntos se realizó mediante series de tiempo ARIMA, suavizado exponencial y un modelo de participación de mercado autorregresivo. Para las categorías genéricas de productos, dada la discontinuidad de los datos en el tiempo, no se pudo realizar un pronóstico por series de tiempo y se utilizó un modelo de participación de mercado autorregresivo.

Los resultados para el total de canje muestran que el modelo que mejor predijo, con un MAPE de 10,5%, es un modelo ARIMAX(1,0,2), en donde se incorporaron variables externas como algunos fines de mes, la semana de navidad y puntos acumulados. Para la distribución de canje por nivel de puntaje, el modelo que mejor predijo, para los primeros 4 niveles, es el de participación de mercado autorregresivo con un MAPE que aumenta desde un 8% en el nivel 1 a un 17% en el nivel 4. Para los 4 niveles superiores se recomienda utilizar un modelo suavizado exponencial que arrojó errores menores al modelo anterior y resultados similares al modelo ARIMA. Para la estimación de las categorías genéricas, los resultados fueron aumentando a medida que aumentaba el nivel y disminuía el número de productos de la categoría, el MAPE para las categorías varía entre un 9% el menor y un 44% el mayor.

Como conclusión. Para los canjes totales el mejor modelo a utilizar es un modelo ARIMAX. Sin embargo, este modelo no es el mejor para toda las etapas. El modelo de participación autorregresivo es mejor para los niveles 1 al 4 y un modelo simple como el suavizado exponencial es el más adecuado para los niveles altos y las categorías. Se debe tener en cuenta que no sólo un buen sistema de estimación de demanda es suficiente para brindar satisfacción a los clientes, sino también, una buena y gestión de la cadena de abastecimiento hasta el punto de entrega en las tiendas. Para mejorar la estimación de demanda se sugiere incorporar en los modelos variables de acciones de marketing y usar la función de utilidad determinada en el modelo de demanda como apoyo para la confección de catálogos.

Índice de Contenidos

1.	Introducción	1
1.1	Antecedentes Generales	1
1.2	CRM (Clubes de fidelización)	2
1.3	Planteamiento del problema y justificación	3
1.4	Objetivos y alcances.....	4
1.4.1	Objetivo General.....	4
1.4.2	Objetivos Específicos	4
1.4.3	Alcance	4
1.5	Resultados Esperados	5
2.	Marco Conceptual.....	6
2.1	Series de Tiempo.....	6
2.1.1	Modelo autorregresivo (AR).....	6
2.1.2	Modelo de medias móviles (MA)	7
2.1.3	Modelo ARMA(p,q)	8
2.1.4	Modelo ARIMA(p,d,q).....	8
2.1.5	Modelo SARIMA (p,d,q)(sp,sd,sq)	9
2.1.6	Modelo SARIMAX (p,d,q)(sp,sd,sq)X.....	9
2.1.7	Modelación según Box Jenkins [8].....	10
2.2	Modelos de participación de mercado.....	11
2.2.1	Modelo de interacción Competitiva Multiplicativa (MCI):	11
2.2.2	Modelo Logit Multinominales (MNL):	11
2.2.3	Modelo propuesto. Participación de mercado autorregresivo.	12
2.3	Medidas de error	13
2.3.1	Raíz del error cuadrático medio (RMSE)	14
2.3.2	Error porcentual absoluto medio (MAPE).....	14
2.3.3	Máximo error absoluto porcentual (MAX MAPE).....	14
2.3.4	Error cuadrático medio normalizado (NMSE):.....	15
2.3.5	Sin Stock Medio (SSM, Elaboración propia):	15
2.3.6	Sin Stock Medio Porcentual (SSMP, Elaboración propia):	15
3.	Metodología	16
3.1	Estudio de la Situación Actual	16
3.2	Procesamiento de Datos y Validación de Hipótesis	16
3.3	Estimación de demanda de globales.....	16
3.4	Estimación de demanda por cada nivel de puntos.....	16
3.5	Estimación de demanda por productos por cada nivel de puntos.....	16

4.	Desarrollo Metodológico.....	18
4.1	Estudio de la Situación Actual	18
4.1.1	Descripción del programa de puntos	18
4.1.2	Situación Actual de Canjes.....	19
4.1.3	Generación de hipótesis.....	23
4.1.4	Estimación de demanda existente.....	23
4.2	Procesamiento de Datos y Validación de Hipótesis	24
4.2.1	Limpieza de Datos.....	24
4.2.2	Procesamiento de Datos.....	25
4.2.3	Validación de Hipótesis.....	25
4.3	Estimación de demandas Globales	31
4.3.1	Desarrollo del Modelo	32
4.3.1.1	Estimación de orden del modelo _____	32
4.3.1.2	Estimación de parámetros del modelo _____	35
4.3.2	Errores del Modelo	36
4.3.2.1	Ajuste del modelo	36
4.3.2.2	Predicción del Modelo _____	38
4.4	Estimación de demanda por nivel de puntos	41
4.4.1	Desarrollo del Modelo	41
4.4.2	Errores del Modelo	42
4.4.2.1	Ajuste del modelo _____	42
4.4.2.2	Predicción del Modelo _____	43
4.5	Estimación de demanda por tipo de productos	46
4.5.1	Desarrollo del Modelo	46
4.5.2	Errores del Modelo	48
4.5.2.1	Ajuste del modelo _____	48
4.5.2.2	Predicción del Modelo _____	50
5.	Resumen de resultados	54
6.	Conclusiones	58
7.	Estudios Futuros y propuestas de mejora	60
	Carta Gantt	62
	Bibliografía.....	63
	Anexos	64

ÍNDICE DE TABLAS

Tabla 1: Nivel de Puntos.....	18
Tabla 2: Categorías Empresa Principal	22
Tabla 3: Categorías Empresa Holding.....	23
Tabla 4: Error de la Estimación Actual	24
Tabla 5: Correlaciones entre Número de canjes y antigüedad del catálogo	31
Tabla 6: Test Anova Diferencia de Medias	33
Tabla 7: Estimación de Parámetros.....	36
Tabla 8: Ajuste del Modelo	36
Tabla 9: Tabla de autocorrelaciones del Modelo.....	37
Tabla 10: Canjes estimados v/s Canjes Reales	39
Tabla 11: Errores de Estimación.....	39
Tabla 12: Ajuste de la Estimación.....	39
Tabla 13: Errores de calibración del modelo.....	40
Tabla 14: Coeficientes Modelo suavizado Exponencial	40
Tabla 15: Errores Modelo Suavizado Exponencial	40
Tabla 16: Parámetros del Modelo.....	42
Tabla 17: Ajuste del modelo.....	42
Tabla 18: Significancia de las variables del modelo	43
Tabla 19: Predicción de canjes semanales	43
Tabla 20: Error porcentual Semanal.....	44
Tabla 21: Indicadores de error	44
Tabla 22: Indicadores de error	44
Tabla 23: Ajuste del modelo (periodo de validación)	45
Tabla 24: Estimación Parámetros Suavizado Exponencial	45
Tabla 25: Errores Suavizado	45
Tabla 26: Errores ARIMA	46
Tabla 27: Categorías Nivel 1	47
Tabla 28: Coeficientes del modelo (Categorías Empresa Principal).....	47
Tabla 29: Coeficientes del modelo (Categorías Empresa Holding)	48
Tabla 30: Error de Ajuste de los modelos	49
Tabla 31: Significancia de las variables del modelo (Administración)	49
Tabla 32: Significancia de las variables del modelo (Básicos Decohogar).....	49
Tabla 33: Significancia de las variables del modelo (Blanco)	49
Tabla 34: Predicción de canjes semanales (Empresa Principal).....	50
Tabla 35: Predicción de canjes semanales (Empresa Holding)	50
Tabla 36: Error porcentual Semanal (Empresa Principal)	51
Tabla 37: Error porcentual Semanal (Empresa Holding).....	51
Tabla 38: Ajuste del modelo (periodo Validación).....	52
Tabla 39: Indicadores de error	52
Tabla 40: Indicadores de Error	53
Tabla 41: Errores de Estimación (ARIMAX)	54

Tabla 42: Errores Modelo Suavizado Exponencial	54
Tabla 43: Errores Participación de mercado autorregresivo	54
Tabla 44: Errores Suavizado	55
Tabla 45: Errores ARIMA	55
Tabla 46: Errores Categorías Nivel 1	56
Tabla 47: Errores Categorías Nivel 2	56
Tabla 48: Errores Categorías Nivel 3	56
Tabla 49: Errores Categorías Nivel 4	57
Tabla 50: Significancia de las variables del modelo (nivel 2)	68
Tabla 51: Significancia de las variables del modelo (nivel 3)	68
Tabla 52: Significancia de las variables del modelo (nivel 4)	68
Tabla 53: Significancia de las variables del modelo (nivel 5)	68
Tabla 54: Significancia de las variables del modelo (nivel 6)	68
Tabla 55: Significancia de las variables del modelo (nivel 7)	69
Tabla 56: Significancia de las variables del modelo (nivel 8)	69
Tabla 57: Significancia de las variables del modelo (Deportes)	70
Tabla 58: Significancia de las variables del modelo (Electrohogar)	70
Tabla 59: Significancia de las variables del modelo (Comunicación, Electrónica)	70
Tabla 60: Significancia de las variables del modelo (Electricidad)	70
Tabla 61: Significancia de las variables del modelo (Electricidad)	70
Tabla 62: Significancia de las variables del modelo (Menaje).....	71
Tabla 63: Significancia de las variables del modelo (Muebles y organización).....	71
Tabla 64: Categorías Nivel 2	71
Tabla 65: Error de Ajuste de los modelos Nivel 2 (periodo de Validación).....	71
Tabla 66: Error de Ajuste de los modelos Nivel 2 (periodo de Validación).....	72
Tabla 67: Significancia de las variables del modelo (Básicos Decohogar).....	72
Tabla 68: Significancia de las variables del modelo (Blanco)	72
Tabla 69: Significancia de las variables del modelo (Electrohogar)	72
Tabla 70: Significancia de las variables del modelo (Accesorios Automóviles)	73
Tabla 71: Significancia de las variables del modelo (Comunicación, Electrónica)	73
Tabla 72: Significancia de las variables del modelo (Electricidad)	73
Tabla 73: Significancia de las variables del modelo (Menaje).....	73
Tabla 74: Significancia de las variables del modelo (Muebles y Organización)	73
Tabla 75: Indicadores de error	74
Tabla 76: Categorías Nivel 3	74
Tabla 77: Error de Ajuste de los modelos Nivel 3	74
Tabla 78: Error de Ajuste de los modelos Nivel 3 (periodo de validación).....	75
Tabla 79: Significancia de las variables del modelo (Administración)	75
Tabla 80: Significancia de las variables del modelo (Blanco)	75
Tabla 81: Significancia de las variables del modelo (Electrohogar)	75
Tabla 82: Significancia de las variables del modelo (Accesorios Automóviles)	76
Tabla 83: Significancia de las variables del modelo (Agro)	76
Tabla 84: Significancia de las variables del modelo (Comunicación, Electrónica)	76

Tabla 85: Significancia de las variables del modelo (Ferretería).....	76
Tabla 86: Significancia de las variables del modelo (Menaje).....	76
Tabla 87: Indicadores de error	77
Tabla 88: Categorías Nivel 4.....	77
Tabla 89: Error de Ajuste de los modelos Nivel 4	77
Tabla 90: Error de Ajuste de los modelos Nivel 4 (periodo de validación)	78
Tabla 91: Significancia de las variables del modelo (Administración)	78
Tabla 92: Significancia de las variables del modelo (Básicos Decohogar).....	78
Tabla 93: Significancia de las variables del modelo (Electrohogar)	78
Tabla 94: Significancia de las variables del modelo (Niños).....	79
Tabla 95: Significancia de las variables del modelo (Comunicación Electrónica)	79
Tabla 96: Significancia de las variables del modelo (Menaje).....	79
Tabla 97: Significancia de las variables del modelo (Muebles y Organización)	79
Tabla 98: Significancia de las variables del modelo (Tabiquería/Techumbre).....	79
Tabla 99: Indicadores de Error	80

ÍNDICE DE GRÁFICOS Y FIGURAS

Figura 1: Pasos Metodológicos.....	17
Gráfico 1: Top 16 de ventas en Retail año 2007	1
Gráfico 2: Nivel de puntos y Dinero necesario para acceder a cada nivel.....	19
Gráfico 3: Canjes Totales.....	19
Gráfico 4: Share de canjes por Empresa	20
Gráfico 5: Shares de canjes por Nivel (Nivel 1-Nivel 4).....	20
Gráfico 6: Shares de canjes por Nivel (Nivel 5-Nivel 8).....	21
Gráfico 7. Shares de canjes por Categorías.....	21
Gráfico 8: Shares de canjes por Categorías.....	22
Gráfico 9: Correlación Cruzada, Puntos Entregados y Productos Canjeados	26
Gráfico 10: Canjes Mensuales	27
Gráfico 11: Monto de Compra y Número de Alianzas	28
Gráfico 12: Medias de los canjes semanales por mes.....	28
Gráfico 13: Canjes Semanales (Febrero).....	29
Gráfico 14: Canjes Semanales (Marzo)	29
Gráfico 15: Canjes Semanales (Septiembre).....	30
Gráfico 16: Canjes Semanales (Octubre).....	30
Gráfico 17: Canjes Semanales (Diciembre)	31
Gráfico 18: Serie de Canjes	32
Gráfico 19: Medias Trimestrales	32
Gráfico 20: Autocorrelación Serie de Canjes.....	33
Gráfico 21: Autocorrelación Parcial Serie de Canjes.....	34
Gráfico 22: Análisis Espectral Serie de Canjes.....	35
Gráfico 23: Autocorrelación Errores del modelo	37
Gráfico 24: Ajuste del Modelo	38
Gráfico 25: Predicción del Modelo.....	39
Gráfico 26: Shares de canjes por Nivel (Nivel 1-Nivel 4).....	41
Gráfico 27: Shares de canjes por Nivel (Nivel 5-Nivel 8).....	41
Gráfico 28: Canjes Semanales (Enero)	64
Gráfico 29: Canjes Semanales (Abril)	64
Gráfico 30: Canjes Semanales (Mayo).....	65
Gráfico 31: Canjes Semanales (Junio)	65
Gráfico 32: Canjes Semanales (Julio).....	66
Gráfico 33: Canjes Semanales (Agosto).....	66
Gráfico 34: Canjes Semanales (Noviembre).....	67

1. Introducción

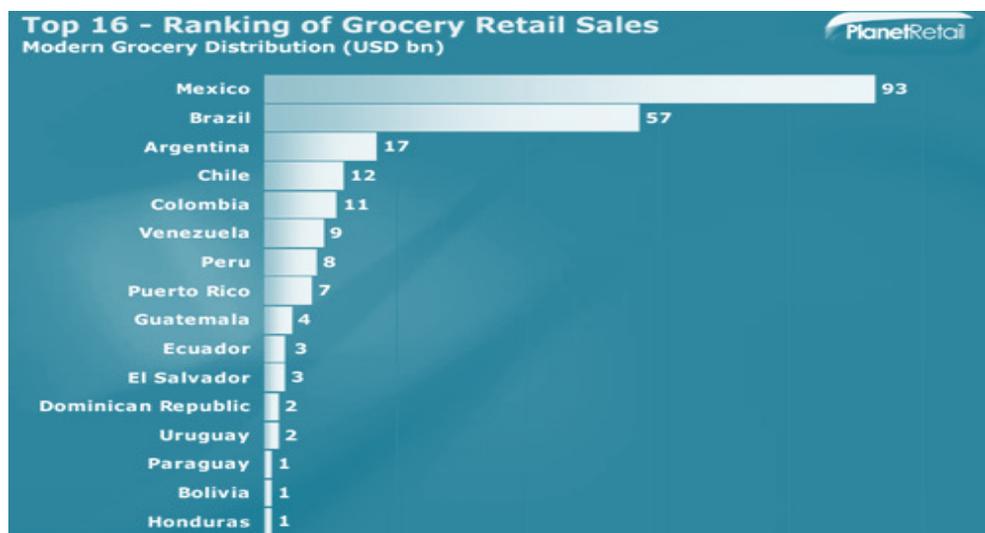
1.1 Antecedentes Generales

Las tiendas por departamentos son establecimientos pertenecientes al mundo del Retail, los cuales ofrecen una gran variedad de productos. Estos establecimientos pretenden satisfacer una amplia gama de necesidades de los clientes. Normalmente, poseen enormes dimensiones y se encuentran principalmente ubicados en lugares céntricos de la ciudad.

En Chile, la industria del Retail ha sido muy exitosa y el mercado se encuentra prácticamente saturado. Es por esto que, cada vez se hace más valioso el generar nuevas técnicas o estrategias para crear valor al cliente, y lograr una mayor participación en la canasta de éstos. Por lo tanto, muchas de estas empresas han estado en constante movimiento, incluso han llegado a ofrecer sistemas financieros. Por otro lado, han decidido expandirse dentro de Latinoamérica, principalmente en Perú, Argentina y Colombia.

En el año 2007 Chile se encuentra en el cuarto lugar¹ en Latinoamérica en ingresos provenientes del Retail, con alrededor de 12 mil millones de USD\$ en ventas. El siguiente gráfico expone lo antes mencionado.

Gráfico 1: Top 16 de ventas en Retail año 2007



Cada una de las tiendas por departamentos, para lograr retener a sus clientes, han desarrollado sistemas de fidelización distintos. En particular, la tienda por departamentos en la cual se realiza el estudio, posee una tarjeta de fidelización y de crédito, la cual permite a los clientes poder realizar sus compras en cuotas y por lo tanto, facilita el acceso a productos los cuales no podrían ser adquiridos “al contado”.

Otro de los beneficios que posee esta tarjeta, es un sistema de acumulación de puntos, que consiste en que cada vez que el cliente realiza sus compras pagando con la tarjeta, ya sea en la tienda o en alguna de sus alianzas, acumula una cierta cantidad de puntos, los cuales puede canjear al llegar a cierto nivel. Existen 8 niveles de puntos, los cuales se diferencian en la

¹ Según PanetRetail

cantidad de puntos acumulados. En cada uno de estos niveles, el cliente puede canjear distintos tipos de productos, los cuales se encuentran en un catálogo, y por lo tanto, se debe mantener un stock de éstos.

Uno de los principales problemas de este tipo de sistema, es que existe un costo de inventario de los productos que se deben mantener en stock. Por lo tanto, es imprescindible para la empresa lograr determinar una cantidad en stock que permita reducir estos costos y que a la vez, satisfagan la demanda existente.

La temática de este trabajo, es desarrollar un modelo que permita predecir el canje de productos para cada nivel de puntos y por tipo de productos.

1.2 CRM (Clubes de fidelización)

El CRM (Customer Relationship Management), es una estrategia de negocio basada en el cliente, optimizando su valor en el largo plazo [1]. No hay que limitar este concepto sólo en la compra y venta de bienes o servicios, si no que éste adquiere sentido en toda la gestión del cliente, como por ejemplo, dudas, sugerencias, reclamos, etc.

Los programas CRM ofrecen:

- a) Mayor conocimiento del cliente y personalización del trato [2]:
 - Identificar y conocer mejor a los clientes de la empresa
 - Personalizar de forma más precisa
 - Tener centralizada toda la información
 - Accesibilidad a la información de todos los empleados
 - Disponer de un perfil dinámico de clientes
 - Conocimiento de mercado para las direcciones de ventas y marketing

- b) Aumento de la satisfacción y lealtad de los clientes:
 - Aumento del grado de satisfacción de los clientes
 - Reducción de los costes de las campañas de promoción de nuevos clientes
 - Comunidad de clientes

- c) Aumento de las ventas:
 - Up-selling
 - Cross-selling
 - Precios premium

- d) Reducción de los costos de servicio:
 - Menores costes de atención al cliente
 - Menores costes de rotación de empleados

En el presente trabajo se debe entender el comportamiento de los clientes frente al canje de productos, para poder generar un modelo de predicción de demanda de éstos. Por lo anterior, es importante conocer el concepto CRM.

En el siguiente capítulo se presenta con más detalle el problema a abordar y su respectiva justificación.

1.3 Planteamiento del problema y justificación

En el último tiempo, las grandes empresas de Retail han tratado de retener a sus clientes mediante “tarjetas de fidelización”. Estas empresas desean premiar a sus clientes más leales con descuentos y puntos intercambiables por productos a medida que van realizando una mayor cantidad de compras con la tarjeta. Sin embargo, los primeros resultados de estos programas no han sido satisfactorios, demostrando que la fidelización de clientes debe ir más allá para ser efectiva [3].

Un programa de marketing relacional debe partir de una decisión estratégica consciente por parte de la alta dirección de los retailers. Se debe aspirar a crear y desarrollar relaciones más estrechas y valiosas tanto para el Retail como para el cliente. Es decir, se pretende influir positivamente en los clientes para que actúen de una determinada manera que les beneficie tanto a ellos como a la propia empresa. Una tarjeta de fidelización, por lo general, no es suficiente para lograr lo anterior. Hay que tratar de sacar el máximo potencial a las tecnologías existentes y, en particular, del CRM [4].

En particular la tienda en estudio está en un rápido crecimiento y con grandes aspiraciones de posicionarse como una de las empresas de Retail más importantes de Latinoamérica, su modelo de tarjeta de fidelización es parte fundamental para la empresa. El problema radica, en que no existe un modelo de predicción de demanda de canje de productos adecuado, que permita disminuir los grandes quiebres de stock existentes que hoy alcanzan un 40%². Estos quiebres, podrían provocar una molestia en el cliente y por lo tanto, una posible fuga de la tienda hacia la competencia.

Uno de los objetivos a desarrollar en el trabajo es crear un modelo que permita determinar la cantidad de productos que se demandará para cada nivel de puntos y categoría pertenecientes al catálogo.

Con respecto a las temáticas a desarrollar durante el proyecto, se vislumbran distintas disciplinas, tales como:

Gestión de operaciones: Desarrollar un modelo de estimación de demanda es parte de esta disciplina.

Marketing: Se debe conocer a los clientes, y determinar aquellas variables importantes que pueden influir en su decisión de canje de productos o de uso de la tarjeta a nivel macro.

Con respecto a los costos que se deben disminuir se encuentran los costos directos o medibles y los costos indirectos o no medibles.

Los costos medibles, son los costos que provocan aquellos productos que no se canjean durante la duración del catálogo y son liquidados a un precio menor que al que serian vendidos al momento de la adquisición. Actualmente, el 30% de los productos no se canjean en el periodo de duración del catalogo. Si se considera que estos productos son liquidados a un 80%³ del precio

² Dato entregado por el Gerente de Marketing

³ Datos de liquidación de productos y de productos que no se canjean entregado por el Gerente de Fidelización

luego de haber terminado el catálogo (el cual tiene una duración aproximada de 3 meses), implica dejar de ganar \$54.000.000 de pesos mensuales (\$648.000.000 anuales).

Con respecto a los costos no medibles se encuentra la decepción que provocaría en el cliente la no existencia de los productos que desean canjear que podrán provocar eventualmente su fuga del programa hacia la competencia [5]. Esta decepción se agravaría si el cliente apresuró sus compras para poder acceder a un nivel en donde existe un producto que éste desea. Suponiendo una fuga del 1% de los clientes insatisfechos se estima un costo mínimo de MM\$1.200 anuales (El detalle se encuentra en trabajos futuros).

Finalmente, entender aquellas variables que influyen en los canjes, es importante para una buena planificación de campañas y catálogos. Por ejemplo, el conocer que en ciertas semanas o meses los canjes disminuirán, se podrían generar campañas para aumentar los canjes en éstas y disminuirlos en aquellas donde son mucho mayores, suavizando así un poco la demanda y por lo tanto disminuir los costos operacionales en los meses o semanas que la demanda crece enormemente.

1.4 Objetivos y alcances

1.4.1 Objetivo General

Desarrollar un modelo de estimación de demanda que permita estimar la demanda de productos que se van a canjear en una determinada semana. Para ayudar a la toma de decisiones relacionadas a mantener el stock adecuado de productos.

1.4.2 Objetivos Específicos

- Estimar la demanda global de puntos que se van a canjear en una semana, que permitirá obtener un stock global de productos adecuado para la empresa.
- Desarrollar un modelo de predicción de demanda de puntos que se van a canjear por niveles de puntos, para poder determinar la estimación de demanda para la siguiente etapa y poder repartir el presupuesto de canje de manera adecuada en los distintos niveles.
- Predecir la demanda a nivel de tipo de producto para los niveles 1 al 4 en la semana siguiente, así se podrá tener un stock de productos adecuados al nivel y poder mantener una relación más cercana con los proveedores de éstos.
- Realizar un análisis cualitativo de las variables predictivas del modelo.
- Determinar una cuantificación de los errores de pronósticos y la capacidad predictiva de los modelos.

1.4.3 Alcance

El alcance del trabajo realizado contempla realizar un modelo de demanda que permita predecir el canje de puntos a nivel global, niveles de punto, tipo de producto y productos nuevos de catálogo.

El modelo a desarrollar contemplará los canjes globales de todos los productos dentro del territorio chileno y no se estudiará en este trabajo los canjes a nivel de tienda.

Se deberá entregar un pronóstico para las categorías más importantes, es decir, aquellas que en conjunto posean más del 80% de los canjes por nivel.

El modelo debe permitir a la empresa estimar semanalmente la cantidad de productos por categorías que se demandarán en la semana siguiente y poder responder las necesidades de canjes de sus clientes.

El modelo no se hará cargo de la entrada de nuevas categorías de productos.

Se llegará a un nivel de prototipo de los modelos, entregando un manual a la empresa de cómo implementarlos.

1.5 Resultados Esperados

Al finalizar el proyecto, se espera obtener como producto final los siguientes resultados:

1. Un modelo que permita estimar la demanda de productos pertenecientes a cada nivel de puntos. El sistema recibirá como parámetros:

- El resultado de canjes por cada nivel de puntos durante el período de estudio.
- El resultado de canjes por cada categoría en el periodo de estudio.
- Los puntos entregados en el periodo de evaluación.
- El número de productos disponibles en la categoría.

Los outputs del sistema contemplarán los siguientes elementos:

- El pronóstico de canjes para cada tipo de productos, para la semana siguiente, para los niveles 1 al 4.
- Una estimación de la cantidad de canjes por cada nivel de puntos.
- El pronóstico de canje a nivel global.

2. Una comparación de las ventajas y desventajas de la utilización de esta metodología respecto de los enfoques actualmente utilizados en la empresa.

3. Finalmente, se entregará una descripción de la metodología que permitirá afinar el modelo en el futuro.

2. Marco Conceptual

A continuación se expondrán los métodos de estimación de demanda que se necesitan para el desarrollo de la memoria. Se explicarán los modelos de estimación de Serie de Tiempo ARIMA y de participación de mercado.

2.1 Series de Tiempo

Para abordar el problema de estimación de demanda, se han desarrollado una gran cantidad de metodologías o modelos, como por ejemplo los modelos de regresión simple, métodos de suavización exponencial, series de tiempo, redes neuronales, modelos de atributos, etc. El utilizar uno u otro modelo depende del problema que se esté enfrentado y la disponibilidad de datos con la que se cuente. En particular, entre los modelos de pronóstico de series de tiempo existentes se estudiarán aquellos propuestos por Box y Jenkins [6], denominados como modelos ARIMA.

Un proceso estocástico es una sucesión de variables aleatorias Y_t ordenadas, pudiendo tomar t (tiempo o periodo) cualquier valor entre $-\infty$ y ∞ . Por ejemplo, la siguiente sucesión de variables aleatorias puede ser considerada como proceso estocástico:

$$Y_{-5}, Y_{-4}, Y_{-3}, Y_{-2}, \dots, Y_3, Y_4$$

Definimos un modelo como autorregresivo si la variable endógena de un período t es explicada por las observaciones de ella misma correspondientes a períodos anteriores añadiéndose, como en los modelos estructurales, un término de error. En el caso de procesos estacionarios con distribución normal, la teoría estadística de los procesos estocásticos dice que, bajo determinadas condiciones previas, toda Y_t puede expresarse como una combinación lineal de sus valores pasados (parte sistemática) más un término de error (innovación).

A continuación se exponen los conceptos más utilizados en los siguientes capítulos:

Y: Serie de tiempo de n observaciones $\{Y_1 \dots Y_n\}$, con Y_i = registro de Y en el periodo i .

a: Serie de tiempo de n observaciones $\{a_1 \dots a_n\}$ “ruido blanco”, con media cero y varianza σ_a^2 .

L: Operador de retraso (Lag), tal que $L^p Y_t = Y_{t-p}$

2.1.1 Modelo autorregresivo (AR)

Los modelos autorregresivos se abrevian con la palabra AR tras la que se indica el orden del modelo: AR(1), AR(2),...etc. El orden del modelo expresa el número de observaciones retasadas de la serie temporal analizada que intervienen en la ecuación. Así, por ejemplo, un modelo AR(1) tendría la siguiente expresión:

$$Y_t = C_0 + C_1 * Y_{t-1} + a_t$$

Ecuación 1

El término de error (a_t) de los modelos de este tipo se denomina generalmente ruido blanco cuando cumple las tres hipótesis básicas siguientes:

- media nula
- varianza constante
- covarianza nula entre errores correspondientes a observaciones diferentes

La expresión genérica de un modelo autorregresivo, no ya de un AR(1) sino de un AR(p) sería la siguiente:

$$Y_t = C_0 + C_1 * Y_{t-1} + C_2 * Y_{t-2} + \dots + C_p * Y_{t-p} + a_t$$

Ecuación 2

Pudiéndose escribir de forma abreviada como:

$$\phi_p(L) * Y_t = (1 - \phi_1 * L - \phi_2 * L^2 - \dots - \phi_p * L^p) * Y_t$$

Ecuación 3

$$\phi_p(L) * Y_t = Y_t - \phi_1 * Y_{t-1} - \phi_2 * Y_{t-2} - \dots - \phi_p * Y_{t-p}$$

Ecuación 4

En otras palabras, la variable a pronosticar “ Y_{t+1} ” se construye mediante una combinación lineal p observaciones de esa variable: Y_t, Y_{t-p} .

2.1.2 Modelo de medias móviles (MA)

Un modelo de los denominados de medias móviles es aquel que explica el valor de una determinada variable en un período t en función de un término independiente y una sucesión de errores correspondientes a períodos precedentes, ponderados convenientemente. Estos modelos se denotan normalmente con las siglas MA, seguidos, como en el caso de los modelos autorregresivos, del orden entre paréntesis. Así, un modelo con q términos de error MA(q) respondería a la siguiente expresión:

$$Y_t = K_0 + a_t + K_1 * a_{t-1} + K_2 * a_{t-2} + \dots + K_p * a_{t-q}$$

Ecuación 5

Nuevamente puede abreviarse utilizando el polinomio de retardos (como en el caso de los modelos AR):

$$\theta_q(L) * a_t = a_t - \vartheta_1 * a_{t-1} - \vartheta_2 * a_{t-2} - \dots - \vartheta_p * a_{t-q}$$

Ecuación 6

Al igual que en el caso de los modelos autorregresivos, el orden de los modelos de medias móviles suele ser bajo MA(1), MA(2) que corresponde con la periodicidad de los datos analizados.

2.1.3 Modelo ARMA(p,q)

El modelo ARMA(p,q) combina las aproximaciones realizadas por los modelos AR(p) y MA(q). Es más general que los anteriores, pero tiene como requisito primordial que la serie a modelar debe ser estacionaria, es decir, que cada observación perteneciente a dicha serie debe poseer igual distribución de probabilidad. Esto implica, según Box y Jenkins, que la media y la varianza son iguales para cualquier observación de la serie.

La respuesta de este modelo son dos vectores de coeficientes C y K, los cuales se obtienen mediante estimaciones de máxima verosimilitud o de mínimos cuadrados.

$$\phi_p(L) * (Y_t - \mu) = \theta_q(L) * a_t$$

Ecuación 7

Con $\mu = \text{constante}$.

2.1.4 Modelo ARIMA(p,d,q)

En caso de que la serie no sea estacionaria, se puede modelar mediante el modelo ARIMA(p,d,q). Este modelo realiza transformaciones a la serie de manera de convertirla en una serie estacionaria, mediante la diferenciación de la serie. De esta manera, se define el operador ∇ como “operador diferencial”, tal que: $\nabla X_t = (X_t - X_{t-1})$.

En general, podemos definir $\nabla^d X_t = (X_t - X_{t-d})$

Con lo que la expresión general del modelo sería:

$$\phi_p(L) * (\nabla^d Y_t - \mu) = \theta_q(L) * a_t$$

Ecuación 8

La utilización de modelos ARIMA(p,d,q) es muy frecuente para casos en que la serie, además de no ser estacionaria, presenta tendencia, ya sea creciente o decreciente. En estos casos, se recomienda utilizar un valor de d=1 cuando la serie presenta tendencia lineal y d=2 si la tendencia es cuadrática [7].

2.1.5 Modelo SARIMA (p,d,q)(sp,sd,sq)

El modelo SARIMA es una extensión del modelo anterior. Si la serie de tiempo a modelar presenta una estacionalidad s, este modelo recobra los términos que reflejen dicho comportamiento. Por lo anterior, la variable Y_t está es función de los valores que toma la serie en el periodo t-s y t-k*s, logrando que el modelo sea más completo al incorporar mayor información a éste. Tal como se definió en el modelo anterior el operador diferencial, al existir estacionalidad, se define el operador ∇_s como “operador diferencial estacional”, tal que $\nabla_s Y = Y_t - Y_{t-s}$, siendo s la estacionalidad de Y.

Se crean así 3 nuevos términos, mediante los cuales se incorpora dicha estacionalidad:

sp: término estacional autorregresivo

sd: término estacional diferencial

sq: término estacional de medias móviles

El modelo SARIMA (ARIMA Estacional) queda definido por la siguiente ecuación:

$$\phi_p(L) * \Phi_{sp}(L) * (\nabla^d \nabla_s^{sd} Y_t - \mu) = \theta_q(L) * \Theta_{sq}(L) * a_t$$

Ecuación 9

Donde:

$\Phi_{sp}(L) = 1 - \Phi_1 * L - \Phi_2 * L^2 - \dots - \Phi_{sp} * L^{sp}$: Polinomio autorregresivo estacional de orden sp.

$\Theta_{sq}(L) = 1 - \Theta_1 * L - \Theta_2 * L^2 - \dots - \Theta_{sq} * L^{sq}$: Polinomio de medias móviles estacional de orden sq.

2.1.6 Modelo SARIMAX (p,d,q)(sp,sd,sq)X

Finalmente, otra agregación al modelo ARIMA es cuando se incorporan otras variables que puedan influir en el comportamiento de la serie de tiempo que se va a modelar. Estas variables, permiten incrementar la precisión en el pronóstico. De esta manera, se puede relacionar a la variable Y_t con un conjunto de r variables X, externas al proceso.

$$\phi_p(L)\Phi_{sp}(L)\left[\nabla^d\nabla_s^{sd}(Y_t - \sum_{i=1}^r c_i X_i) - \mu\right] = \theta_q(L)\Theta_{sq}(L)A_t$$

Ecuación 10

Con c_i = coeficientes de las variables externas Y (Y_i).

2.1.7 Modelación según Box Jenkins [8]

Para el desarrollo de los modelos antes mencionados, Box Jenkins ha propuesto una metodología que se basa en una serie de pasos, los cuales permiten incorporar todos los supuestos que son necesarios para determinar si el modelo es adecuado.

Los pasos que rigen esta metodología son los siguientes:

- Identificación tentativa de la serie (Visualización de la serie, gráficos de autocorrelación y autocorrelación parcial):

El gráfico de autocorrelación, permite observar la relación que existe entre un valor de la serie y los valores de ésta pero “k” periodos previos. Formalmente, la autocorrelación de un valor de la serie con el k-ésimo retardo r_k se define como:

$$r_k = \frac{\sum_{i=1}^r (x_i - \bar{x}) * (x_{i+k} - \bar{x})}{\sum_{i=1}^r (x_i - \bar{x})^2}$$

Ecuación 11

Donde \bar{x} es el promedio de las r observaciones x de la serie de tiempo. Los gráficos de autocorrelación parcial, miden la correlación entre un intervalo t y k retardos, sin incluir los efectos de las observaciones entre los valores originales. Formalmente, la autocorrelación parcial con la k-ésima muestra de retardo L_{kk} se define como:

$$L_{kk} = \frac{(r_k - \sum_{j=1}^{k-1} L_{k-1,j} * r_{k-j})}{(1 - \sum_{j=1}^{k-1} L_{k-1,j} * r_j)}$$

Ecuación 12

Al graficar las autocorrelaciones se obtienen los denominados correlogramas, los cuales entregan información sobre cuál es el orden del modelo ARIMA que más se ajustaría a la serie de datos.

- Determinación del orden del modelo (definición de los parámetros p,d y q)
- Estimación de los coeficientes
- Chequeo y Diagnóstico

- Volver al primer paso en caso de resultados insatisfactorios

2.2 Modelos de participación de mercado

A continuación, se explicarán los modelos de participación de mercado existentes en la literatura de interacción competitiva multiplicativa y modelo logit multinomiales [9]. A continuación el modelo propuesto para esta memoria. Antes de explicar estos modelos, se explicarán las componentes presentes en éstos:

A_i	=	atracción o utilidad del sku i
S_i	=	participación de mercado del sku i
X_{ki}	=	valor de la k-ésima variable explicativa del sku i
K	=	número de variables explicativas
β_k	=	parámetro a estimar
α_i	=	parámetro que explica influencia constante del sku i

2.2.1 Modelo de interacción Competitiva Multiplicativa (MCI):

$$A_i = \exp(\alpha_i) \prod_{k=1}^K X_{ki}^{\beta_k} \varepsilon_i$$

Ecuación 13

$$S_i = \frac{A_i}{\sum_{j=1}^m A_j}$$

Ecuación 14

2.2.2 Modelo Logit Multinomiales (MNL):

$$A_i = \exp\left(\alpha_i + \sum_{k=1}^K \beta_k X_{ki} + \varepsilon_i\right)$$

Ecuación 15

$$S_i = \frac{A_i}{\sum_{j=1}^m A_j}$$

Ecuación 16

Linealizando la diferencia de shares⁴ de ambos modelos se obtiene las siguientes expresiones [10]:

$$\log\left(\frac{S_i}{\tilde{S}}\right) = \alpha_i^* + \sum_{k=1}^K \beta_k \log\left(\frac{X_{ki}}{\tilde{X}_k}\right) + \varepsilon_i^* \quad (\text{MCI})$$

Ecuación 17

$$\log\left(\frac{S_i}{\tilde{S}}\right) = \alpha_i^* + \sum_{k=1}^K \beta_k \left(X_{ki} - \bar{X}_k\right) + \varepsilon_i^* \quad (\text{MNL})$$

Ecuación 18

Donde:

\tilde{S} y \tilde{X}_{ki} : Medias geométricas

\bar{S} y \bar{X}_{ki} : Medias aritméticas

$\alpha_i^* = (\alpha_i - \bar{\alpha})$ y $\varepsilon_i^* = (\varepsilon_i - \bar{\varepsilon})$

2.2.3 Modelo propuesto. Participación de mercado autorregresivo.

Se parte de la base de un modelo MNL, que describe la participación de mercado de un producto⁵ en un período t (S_{jt}) en términos de la utilidad o atracción que genera cada uno de ellos en ese periodo (A_{jt}).

$$s_{jt} = \frac{\exp(A_{jt})}{\sum_k \exp(A_{kt})}$$

Ecuación 19

⁴ Con respecto a un bien de referencia (outside good)

⁵ Puede ser SKU, Categoría, etc.

Se define la Atracción de cada producto j en el tiempo t en función de la Atracción de q periodos anteriores, la diferencia entre la Atracciones del periodo anterior y el anterior a éste y una combinación lineal de variables exógenas a la serie de tiempo. La siguiente formula ilustra lo antes mencionado.

$$A_{jt} = \sum_{T=t-1}^{T=t_0} \alpha_T * A_{jT} + \gamma_T * (A_{j(t-1)} - A_{j(t-2)}) + \sum_{q=1}^{q=Q} \beta_q * x_{jq} + C_j$$

Ecuación 20 (Elaboración Propia)

La estimación de parámetros se ve facilitada mediante la utilización de una transformación lineal [11]:

$$\ln(s_{jt}) = A_{jt}$$

Ecuación 21

Los parámetros de A_{jt} son estimables mediante una regresión de mínimos cuadrados ordinarios.

Una vez realizada la estimación se cuenta con la siguiente información del modelo.

- 1) Estimaciones de los parámetros del modelo.
- 2) Error de la estimación de ajuste del modelo.
- 3) Market Share de cada producto en los distintos periodos.
- 4) Estimación del Market share para el siguiente periodo.
- 5) Error de la estimación para cada producto.

Para llevar el pronóstico desde participaciones de mercado a unidades, se multiplica las participaciones de mercado de cada producto por el pronóstico realizado por serie de tiempo (ARIMA) del tamaño del mercado potencial de cada nivel de puntos.

2.3 Medidas de error

Para evaluar las estimaciones se utilizan distintas medidas de error dependiendo del modelo utilizado. Para las estimaciones mediante análisis de regresión se utiliza el R^2 y R^2 -ajustado. El R^2 mide el porcentaje de variación explicada por la regresión, mientras que el R^2 -ajustado es la variación explicada, ajustada por los grados de libertad del modelo [12].

Para evaluar el desempeño del pronóstico (\hat{X}), se utilizarán medidas que comparan la predicción con el valor real. Estas medidas también pueden ser utilizadas para la calibración de los datos. Para una serie $(X_k)_{k=1}^n$ de media \bar{X} y desviación σ , se tienen las siguientes medidas de error:

2.3.1 Raíz del error cuadrático medio (RMSE)

$$RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^n (X_k - \hat{X}_k)^2}$$

Ecuación 22

Este indicador entrega un error promedio con respecto a los errores cuadráticos obtenidos en el pronóstico de toda la serie de tiempo.

2.3.2 Error porcentual absoluto medio (MAPE)

$$MAPE = \frac{1}{n} \sum_{k=1}^n \left| \frac{(X_k - \hat{X}_k)}{X_k} \right|$$

Ecuación 23

Este indicador entrega en promedio el error absoluto obtenido en el pronóstico de toda la serie, como porcentaje de la serie original.

2.3.3 Máximo error absoluto porcentual (MAX MAPE)

$$MAXP = \underset{k=1, \dots, n}{MAX} \left| \frac{(X_k - \hat{X}_k)}{X_k} \right|$$

Ecuación 24

Esta medida encuentra el valor más alejado de la realidad, en términos absolutos y porcentuales con respecto a la serie original, durante el período de pronóstico.

2.3.4 Error cuadrático medio normalizado (NMSE):

$$NMSE = \frac{\sum_{k=1}^N (\hat{X}_k - X_k)^2}{\sum_{k=1}^N \hat{X}_k * \sum_{k=1}^N X_k} * N$$

Ecuación 25

De esta medida se recupera la discrepancia normalizada entre los valores observados y los estimados.

2.3.5 Sin Stock Medio (SSM, Elaboración propia):

$$SSP = \begin{cases} \frac{|\hat{X}_k - X_k|}{N}, & \hat{X}_k \leq X_k \\ 0, & \text{En otro caso} \end{cases}$$

Ecuación 26

Esta medida entrega el número de unidades que hubiesen faltado (quiebre de stock), al utilizar el modelo, suponiendo una eficiente distribución de inventario.

2.3.6 Sin Stock Medio Porcentual (SSMP, Elaboración propia):

$$SSP = \begin{cases} \frac{|\hat{X}_k - X_k|}{\hat{X}_k}, & \hat{X}_k \leq X_k \\ 0, & \text{En otro caso} \end{cases}$$

Ecuación 27

Entrega el porcentaje de error promedio de stock en el cual se incurre.

3. Metodología

En el presente capítulo se expondrá cómo aplicar la teoría antes mencionada al problema específico de la empresa. A continuación se exponen en detalle cada uno de los pasos metodológicos a considerar en el proyecto.

3.1 Estudio de la Situación Actual

En esta primera etapa, se describirá en el programa de puntos en el cual se desarrollará el proyecto. Además, se ilustrarán las series de canjes actuales, que son las que se pretenden modelar. Finalmente, se generarán hipótesis que tienen relación a las variables que pueden influir en el canje de productos.

3.2 Procesamiento de Datos y Validación de Hipótesis

En esta fase de la metodología se realizará una limpieza de aquellos valores que son errores de la base de datos, como por ejemplo, valores perdidos, incoherencias como canjes negativos, números de canjes decimales, etc. A continuación se procesarán los datos, de manera de dejar las bases de forma amigable para el futuro desarrollo de los modelos. Finalmente se procederá a validar las hipótesis generadas en la etapa anterior.

3.3 Estimación de demanda de globales

En este paso se utilizará un modelo de estimación de demanda basado en series de tiempo. El modelo escogido es un modelo SARIMAX el cual permite incorporar las estacionalidades y variables externas que influyen en el comportamiento de canje de los clientes. Una vez obtenido un modelo se realizará un estudio de los errores de ajuste que ha incurrido el modelo, para asegurarse que no existen otros factores que podrían explicar los canjes (Ruido Blanco).

3.4 Estimación de demanda por cada nivel de puntos

En esta etapa se utilizará un modelo de participación de mercado autorregresivo. Esta estimación permitirá estimar los costos que se incurrirán por cada nivel de puntos y además son base de la estimación de la siguiente etapa. Como en la etapa anterior, una vez obtenido un modelo se realizará un estudio de los errores de éste, para asegurarse que no existen otros factores que podrían explicar los canjes.

3.5 Estimación de demanda por productos por cada nivel de puntos

En esta etapa se utilizará un modelo de participación de mercado autorregresivo. Esta estimación permitirá a la empresa tener una visión más clara de aquellas categorías que a los

clientes les otorgan una mayor utilidad y por lo tanto al momento de decidir qué productos incorporar en un determinado catálogo aumentar el valor agregado que estos pueden traspasar a los clientes. Además, se puede reaccionar ante algún cambio en la tendencia de las categorías preferidas y así la empresa puede tener menor incertidumbre al decidir reabastecerse de algún tipo de producto.

El siguiente esquema resume los pasos metodológicos a seguir.

Figura 1: Pasos Metodológicos



4. Desarrollo Metodológico

En este capítulo, se expone el desarrollo de la metodología antes mencionada y como se abordaron cada una de estas etapas.

4.1 Estudio de la Situación Actual

4.1.1 Descripción del programa de puntos

El programa de puntos permite a las personas que formen parte de él, acumular Puntos que en definitiva permitirán canjear puntos en las tiendas de la empresa en estudio o adheridas. Cada compra que un socio participante realice en las tiendas de la tienda de estudio o tiendas adheridas a esta promoción, en que el medio de pago sea la tarjeta, generará una cierta cantidad de puntos que se acumularán en una cuenta especial que se creará al efecto para cada socio participante. Los puntos a los que accedan los clientes que posean tarjetas adicionales, serán acumulados en la cuenta del titular de la tarjeta [13].

Cada \$120 de compra de productos o servicios en que el medio de pago utilizado corresponde a la Tarjeta da derecho al socio de acumular 1 punto. Cada \$950 de Avance en efectivo contra el crédito el socio participante acumula 1 punto. Cuando el socio solicite emisión de facturas como respaldo de compra del bien o servicio, cada \$360 de dichas compras, dará derecho a acumular 1 punto.

Existen 8 niveles de puntos, el cliente puede acceder a estos niveles a medida que va acumulando mayor cantidad de puntos, mientras mayor es el nivel el cliente tendrá acceso a productos, cuyo precio en el mercado son mayores. Si un cliente se encuentra entre dos niveles de puntos, solo podrá acceder a los productos de los niveles inferiores a los puntos que actualmente posee.

El detalle de los niveles de puntos y los puntos necesarios para acceder a éstos, se señalan a continuación:

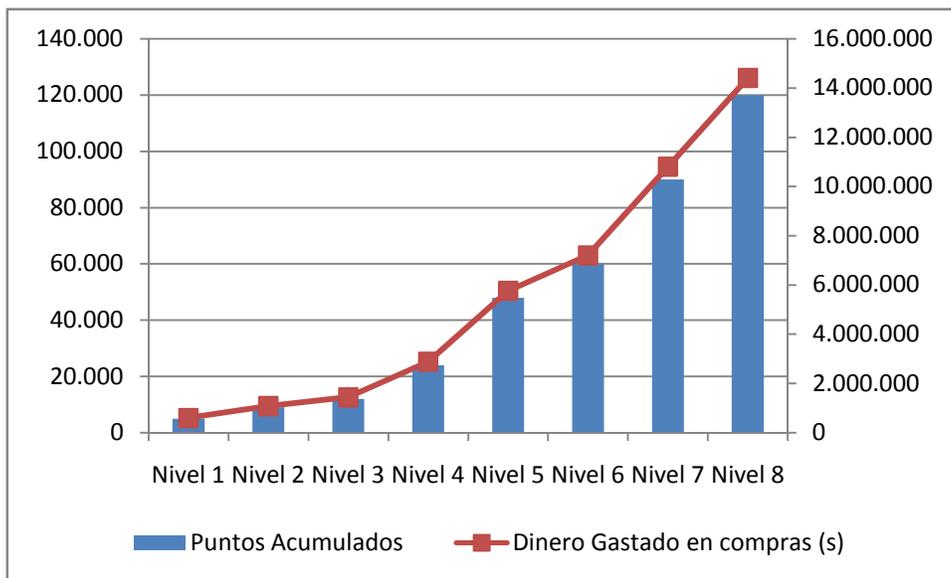
Tabla 1: Nivel de Puntos

Nivel De Puntos	Puntos Acumulados	Costo Promedio \$	Precio Promedio \$
Nivel 1	5.000	3.645	8.954
Nivel 2	9.000	6.864	14.565
Nivel 3	12.000	10.627	23.437
Nivel 4	24.000	27.883	52.564
Nivel 5	48.000	55.602	103.726
Nivel 6	60.000	131.383	214.394
Nivel 7	90.000	284.683	425.648
Nivel 8	120.000	537.674	1.350.234

El programa entró en vigencia el 1 de Marzo de 2003 y se extenderá indefinidamente. Los puntos obtenidos en cada compra realizada durante la duración del programa, tendrá vigencia de 12 meses. Los puntos no canjeados al finalizar dicho periodo, serán descontados de la cuenta.

De todo lo anterior, se puede determinar cuánto dinero un cliente debe comprar con la tarjeta para poder acceder a cada nivel de puntos. El siguiente gráfico ilustra lo antes mencionado:

Gráfico 2: Nivel de puntos y Dinero necesario para acceder a cada nivel

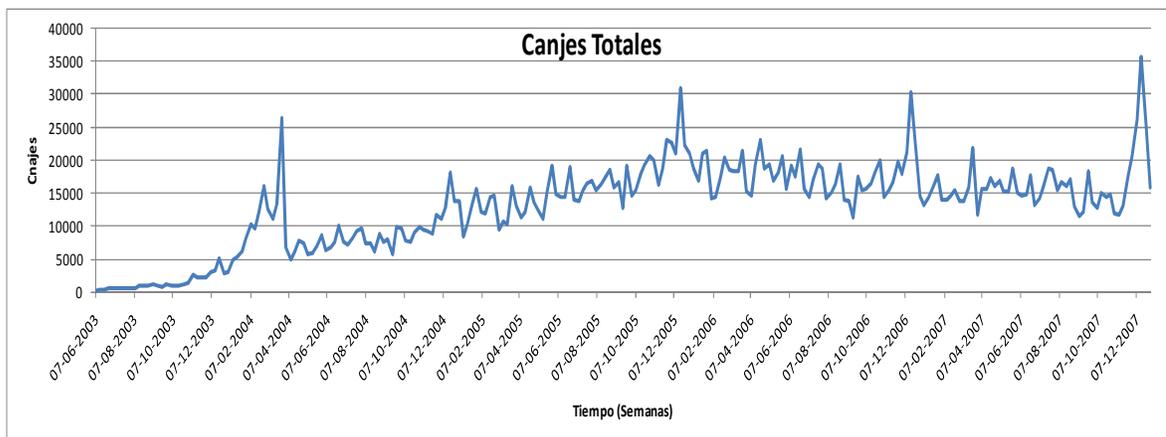


4.1.2 Situación Actual de Canjes

A continuación se presentan los canjes de productos, desde que comenzó el programa de puntos, en los distintos tipos de agregación, en primer lugar se presenta los canjes totales, luego por empresa (empresa de estudio o Holding), por nivel de puntos y por tipo de productos.

El siguiente gráfico representa la evolución de los canjes totales en el tiempo.

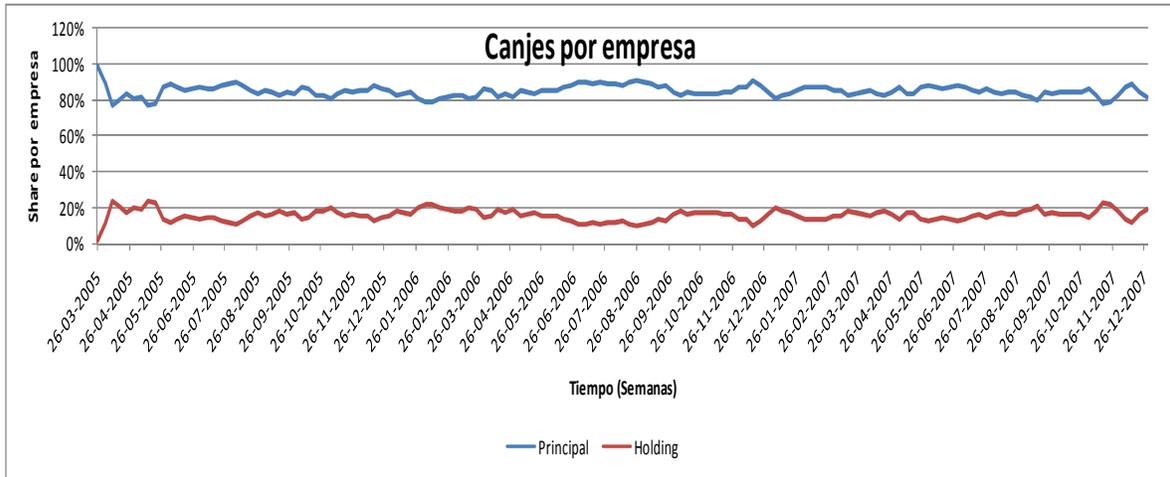
Gráfico 3: Canjes Totales



Del gráfico anterior se desprende, que los canjes fueron en aumento en los primeros dos años y medio y luego se mantienen relativamente constantes, además, existen estacionalidades marcadas en Diciembre de cada año variable a considerar en los modelos. Lo anterior implica que el modelo debería considerarse desde Junio del año 2005 donde la serie se comporta de manera estacional.

A continuación se presenta la historia de canjes por tipo de empresa, ya sea en la tienda en estudio o en el Holding (esto es dónde se canjea el producto).

Gráfico 4: Share de canjes por Empresa



Se observa que el Holding comienza a participar en el programa en Marzo de 2005, y que la participación de los canjes de ésta ha crecido en los últimos 6 meses.

Los canjes por nivel de puntos son los siguientes.

Gráfico 5: Shares de canjes por Nivel (Nivel 1-Nivel 4)

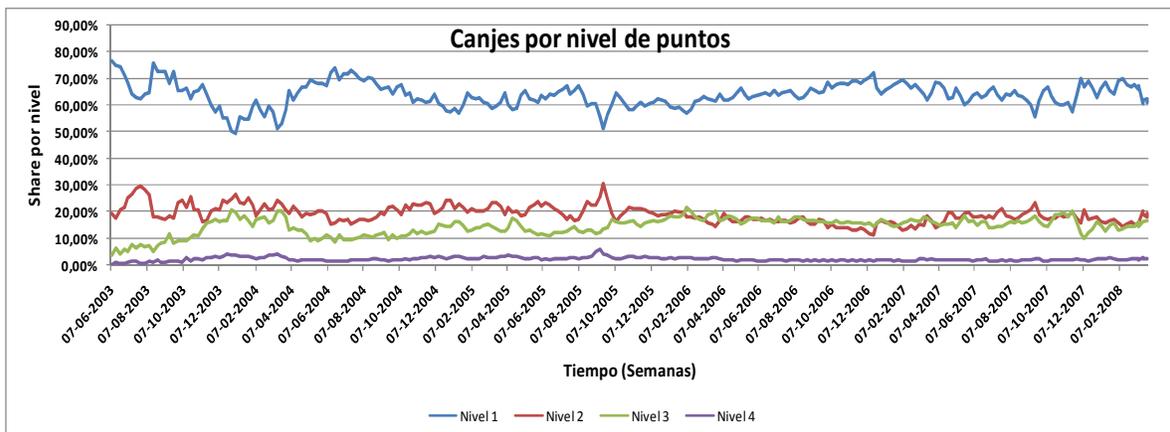
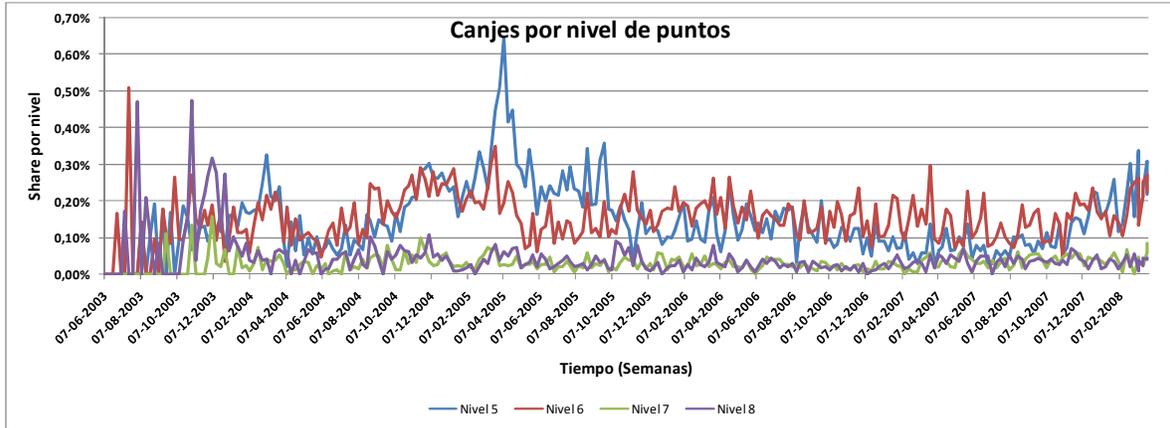


Gráfico 6: Shares de canjes por Nivel (Nivel 5-Nivel 8)



Se observa que a medida que aumentan los niveles disminuye el nivel de canjes y aumenta la volatilidad porcentual de éstos y que los primeros 4 niveles concentran más del 97% de los canjes.

Finalmente, se exponen los canjes por tipo de productos, para las principales categorías (mayor share de canjes).

Gráfico 7. Shares de canjes por Categorías

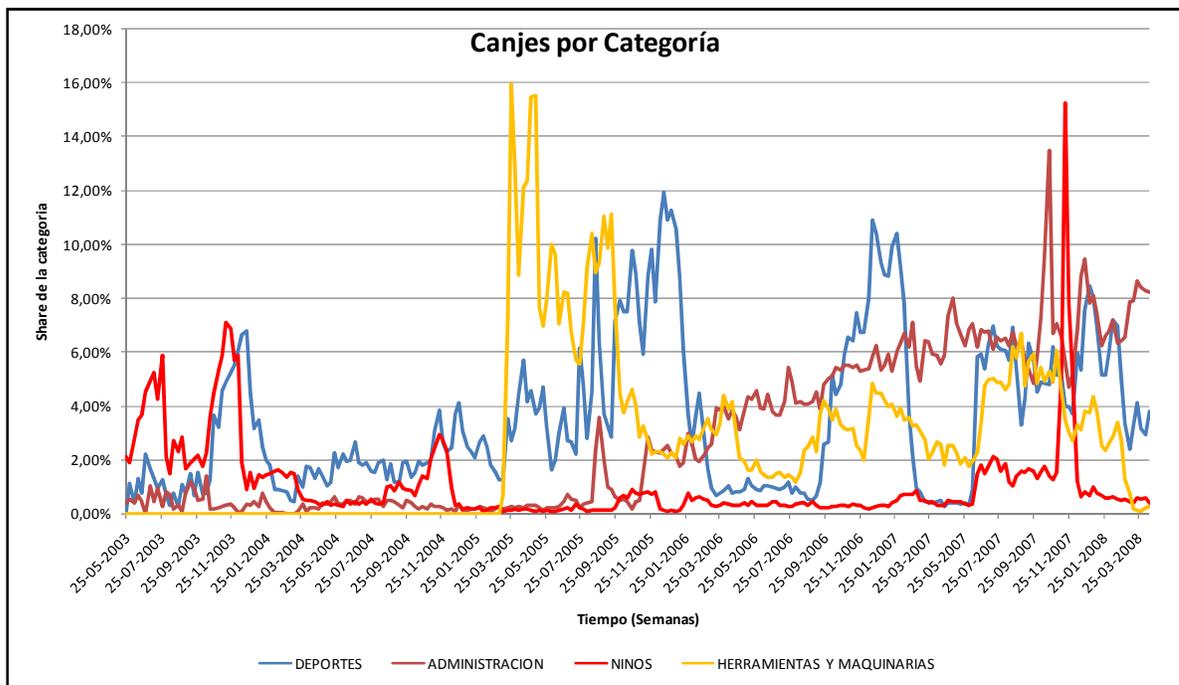
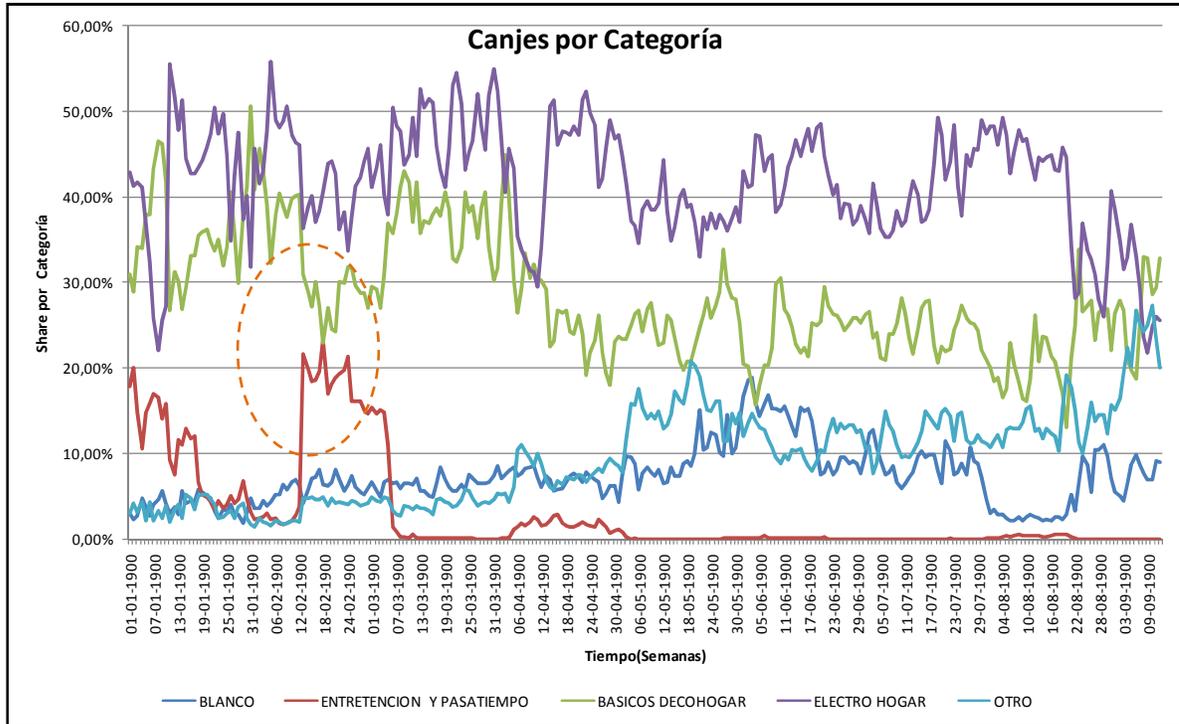


Gráfico 8: Shares de canjes por Categorías



Se observa una gran aleatoriedad entre las distintas categorías, lo que dificulta la estimación. Además se desprende, que cuando entretenimiento y pasatiempo sube su participación, Básicos decohogar la disminuye, lo que sugiere que existe algún efecto de sustitución.

Los tipos de productos o líneas totales que han estado alguna vez en al menos en un catálogo para la empresa principal son las siguientes:

Tabla 2: Categorías Empresa Principal

Categoría	Share Promedio	Número de sku Promedio (por catálogo)
Electro Hogar	41%	300
Básicos Decohogar	27%	172
Blanco	8%	67
Deportes	4%	37
Administración	4%	34
Decoración	1%	32
Entretenimiento y Pasatiempo	1%	18
Niños	1%	26
Electricidad	1%	3
Perfumería	1%	16
Estacional	0%	1
Accesorios Mujer	0%	3

Las líneas presentes en la empresa del Holding son las siguientes:

Tabla 3: Categorías Empresa Holding

Categoría	Share Promedio	Número de sku Promedio (por catálogo)
Menaje	10%	37
Herramientas y maquinarias	5%	44
Comunicación, Electrónica	2%	8
Decoración	2%	32
Electricidad	2%	3
Línea blanca y climatización	2%	4
Muebles y organización	2%	23

4.1.3 Generación de hipótesis

Para el desarrollo de los modelos de predicción de demanda, se han generado hipótesis las cuales deben ser comprobadas al tener acceso a la base de datos de la empresa. Las hipótesis generadas son respecto a las variables externas que pueden influir en los puntos que se van a canjear en un determinado periodo.

A nivel de mercado las variables que pueden afectar la demanda de productos a través del canje de puntos son las siguientes:

- Puntos Entregados: Los puntos entregados en una determinada semana (semanas anteriores) aumenta la probabilidad de canjes, ya que al haber más ventas aumentan los puntos y por lo tanto los clientes podrían llegar a acceder a nuevos productos (alcanzan nuevos niveles).
- Puntos a expirar: Al existir una mayor cantidad de puntos a expirar en un periodo determinado, el número de puntos canjeados en ese periodo aumentan proporcionalmente
- Mes: Existen meses en los cuales los canjes de puntos se concentran en mayor cantidad que el resto del año.
- Promociones doble puntaje: Aumentan las ventas y la acumulación de puntos, por lo tanto, afecta directamente el nivel de canje de estos, dado que los clientes pueden acceder a los niveles de puntos de forma más rápida.
- Ingreso Alianzas: Al ingresar nuevas alianzas al programa, la acumulación de puntos será mayor y por lo tanto, se canjearán una mayor cantidad de puntos.
- Fin de mes: Existe un mayor nivel de canje los fines de mes, al tener un mayor flujo de personas en la tienda.
- Antigüedad del Catálogo: el tiempo que lleve un catálogo en vigencia, podría influir en los canjes de productos.

4.1.4 Estimación de demanda existente

Actualmente, existe un modelo de estimación de demanda lineal **Mensual**, que consiste en multiplicar la demanda de puntos del mismo mes pero del año anterior por una variable que es la tasa de aumento o disminución de las ventas con la tarjeta del año anterior.

A continuación, se expone el MAPE de la estimación existente a nivel global y por nivel de puntos, en los últimos 16 meses.

Tabla 4: Error de la Estimación Actual

Nivel	Mape
Total Canjes	30%
Nivel 1	11%
Nivel 2	14%
Nivel 3	18%
Nivel 4	17%
Nivel 5	47%
Nivel 6	25%
Nivel 7	39%
Nivel 8	40%

De la tabla anterior se desprende que existe un error mayor al 16% incluso en los niveles más altos este error aumenta. Además, no existe una estimación por tipo de producto, actualmente esta estimación se hace con métodos informales poco estructurados o criterio de expertos provocando que exista muchos quiebres de stock en las tiendas⁶.

4.2 Procesamiento de Datos y Validación de Hipótesis

4.2.1 Limpieza de Datos

Entre los procesos que requieren mayor tiempo se encuentra el de limpieza de datos. Uno de los principales “errores” de la base de datos, fue obtener datos negativos para algunos canjes de productos. Antes de decidir eliminar estos datos, se estudiaron en detalle, ya que a veces el que existan canjes negativos significa que un cliente realizó una devolución del producto canjeado. Para esto, se analizaron los clientes en forma independiente para aquellos casos que habían canjes negativos y se hizo un match entre los skus canjeados por el cliente y los devueltos por éste. Hubo 3.450 canjes negativos de los cuales el 98% correspondían a devoluciones y por lo tanto se eliminaron el 2% de los datos por no corresponder a devoluciones. Para el caso de las devoluciones, éstas se anulan con los canjes de aquellos productos que se devuelven, por lo tanto, no fue necesario eliminar estos datos negativos.

⁶ Según Gerente de Marketing, el quiebre alcanza el 40%

4.2.2 Procesamiento de Datos

La calidad de la información que se posee para realizar el pronóstico es primordial para el desarrollo de un modelo que prediga de manera precisa y confiable. Es por esta razón que no sólo una buena limpieza de datos es necesaria, sino también, una buena estructura de éstos que permita desarrollar los modelos de manera un poco más sencilla y amigable. La eficacia del modelo depende en gran manera al buen tratamiento previo que se realice de la información con que se cuenta.

Para el estudio de los datos y el desarrollo del modelo se necesitaban que éstos tuviesen la siguiente estructura.

Fecha (Semana)	Número de Canjes	Puntos Entregados
----------------	------------------	-------------------

SKU	Fecha (Semana)	Categoría	Nivel	Empresa	Número de Canjes
-----	----------------	-----------	-------	---------	------------------

Sin embargo, la estructura de la base de datos era la siguiente:

SKU	LINEA	Empresa
-----	-------	---------

Cliente	Fecha (Día)	SKU	Puntos Entregados
---------	-------------	-----	-------------------

Al encontrarse los datos a nivel diario, se cuenta con alrededor de 300 millones de datos lo que provoca que el procesamiento sea lento dada las limitaciones del procesador. Este proceso fue el más lento de todo el proyecto.

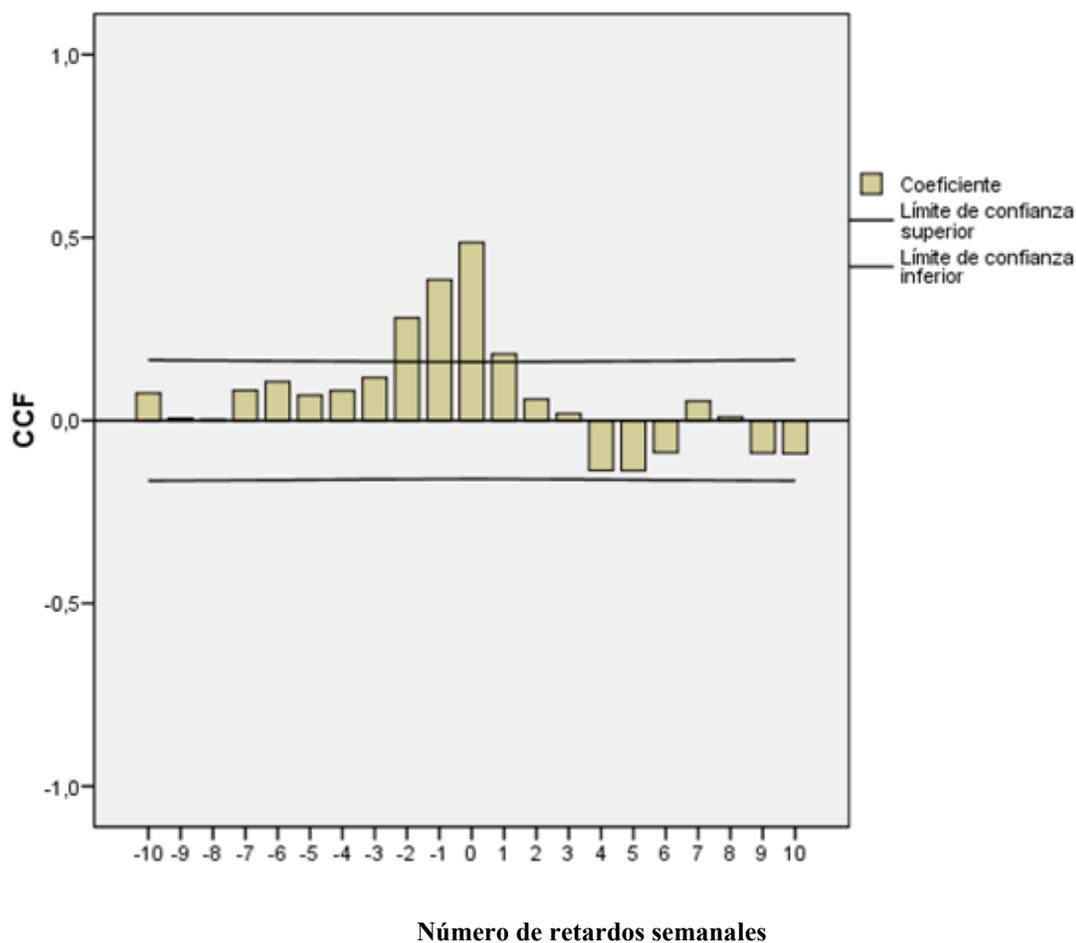
4.2.3 Validación de Hipótesis

Una vez obtenida la estructura de los datos, se procedió a validar las hipótesis antes mencionadas, estas hipótesis tienen relación a las variables que podrían afectar la cantidad de canje en alguna semana. A continuación se exponen como se validaron cada una de ellas.

- Puntos Entregados: Esta variable tiene directa relación con las ventas de la empresa y sus alianzas. Para validar que los puntos acumulados se calcula la **correlación** cruzada entre esta variable y los canjes de productos para las últimas 10 semanas.

A continuación se presenta el gráfico de correlación cruzada antes mencionado.

Gráfico 9: Correlación Cruzada, Puntos Entregados y Productos Canjeados

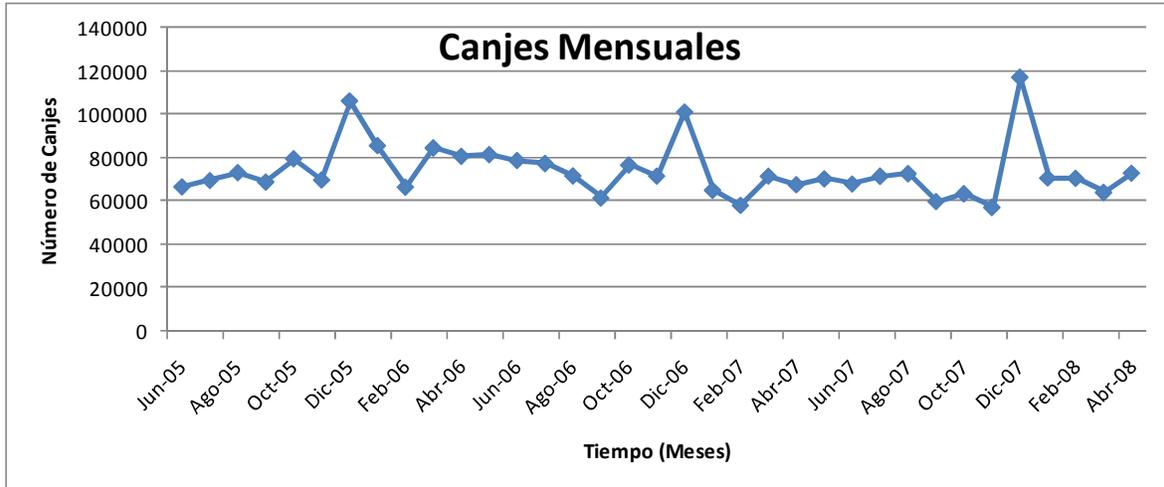


Del gráfico anterior se desprende que existe una directa relación entre los puntos entregados a los clientes y los canjes de productos. Esta correlación se encuentra no solamente la misma semana si no también para un desfase de hasta 2 semanas (superan el límite de confianza superior). Es importante destacar, que la información de puntos entregados de la misma semana que se desea predecir no es relevante, ya que no se poseerá al momento de realizar el pronóstico.

- Puntos a expirar: Lamentablemente la información de puntos a expirar es una variable que se va pisando en la base de datos, por la tanto, no existe la información histórica de ésta. Se sugiere a la empresa, guardar esta variable en el futuro para poder incorporarla a los modelos de estimación de canje. Calcularla tampoco fue posible, debido a que al procesar los datos, se encontró que el 30% de los clientes tenían un acumulación de canjes negativas, por lo que no era de fiar la base de datos por rut del cliente.

- Mes: A continuación se exponen las series de canjes mensuales a nivel global.

Gráfico 10: Canjes Mensuales



Mirando las series de canjes, se observa que existe un claro aumento de éstos en el mes de Diciembre, este aumento se debe a que en la semana de navidad aumentan drásticamente las ventas y por lo tanto es el periodo en donde las personas acumulan mayor cantidad de puntos. Como se vio anteriormente la cantidad de puntos acumulados es una variable que afecta directamente los canjes.

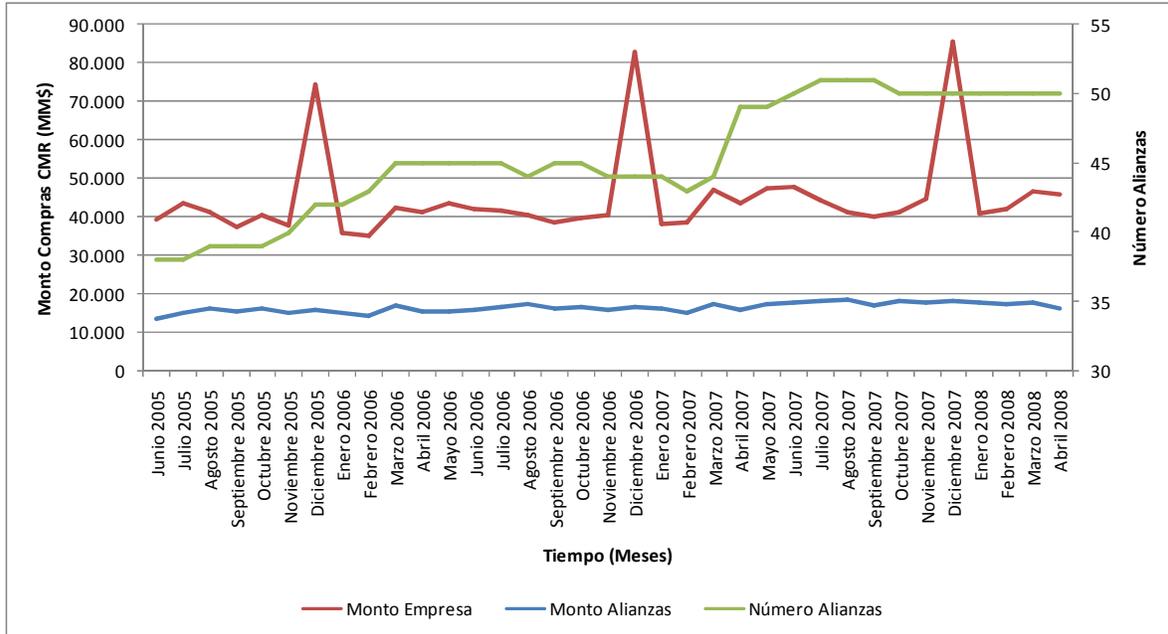
Además de Diciembre, se observa que existen meses en que la cantidad de canjes aumenta en relación al resto de los meses, estos son Abril, Julio y Octubre. Además, Febrero, Mayo, Septiembre y Noviembre la cantidad de canjes disminuye de forma drástica. Más adelante, se realizará un análisis de las semanas más fuertes en canjes por mes, para ver si el efecto es solo en alguna semana o se extiende en el mes completo.

- Promociones doble puntaje: No se contó con esta información, no tienen registradas las fechas de las promociones en forma detallada. Poseen solamente algunas fechas y en forma mensual. Al no encontrarse la información de promociones en forma semanal es difícil aislar el efecto que ésta produjo en la semana que se realizó o en las semanas posteriores.

- Ingreso Alianzas: A continuación se ilustra el gráfico del número de alianzas principales⁷ [14] y el monto de compra con la tarjeta de crédito.

⁷ Participación mayor a un 0,05% mensual en las ventas de la empresa

Gráfico 11: Monto de Compra y Número de Alianzas

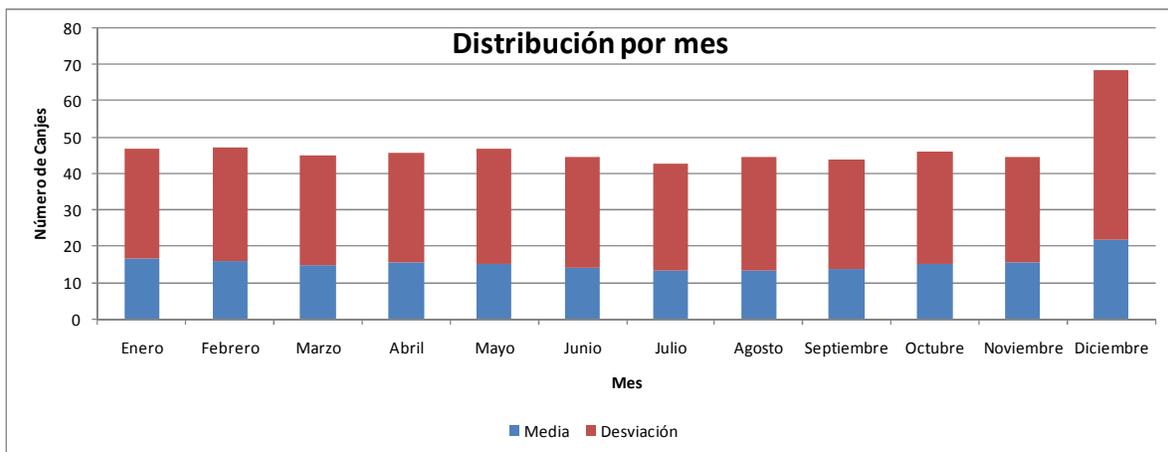


Del gráfico anterior se desprende que a pesar que el número de alianzas principales creció de 37 a 50, sin embargo, el monto de éstas en conjunto se mantiene en el rango de los MM\$17.500 y MM\$18.300. De lo anterior se concluye, que no existe un efecto de aumento drástico en la acumulación de puntos, así que dada la complejidad del análisis de alianzas, no se considerarán en el análisis.

- Fin de Mes: En el siguiente gráfico se muestran las medias semanales de canjes por mes, junto con sus respectivas desviaciones.

-

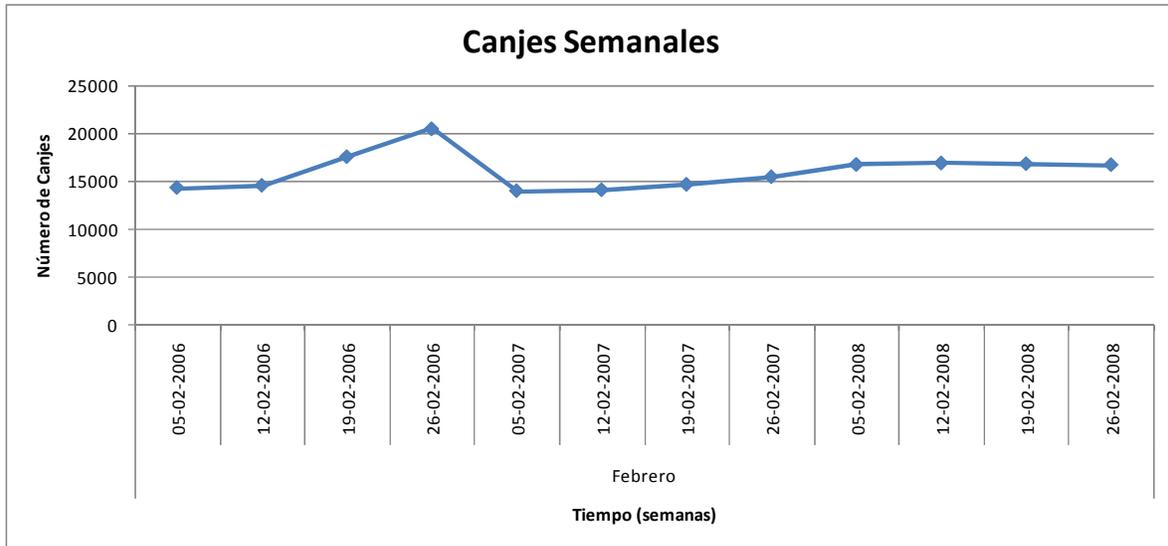
Gráfico 12: Medias de los canjes semanales por mes



Se desprende del gráfico anterior, que existe una gran desviación de los canjes, es decir, éstos no se reparten de manera pareja durante el mes. Por lo tanto, es necesario estudiar cómo se

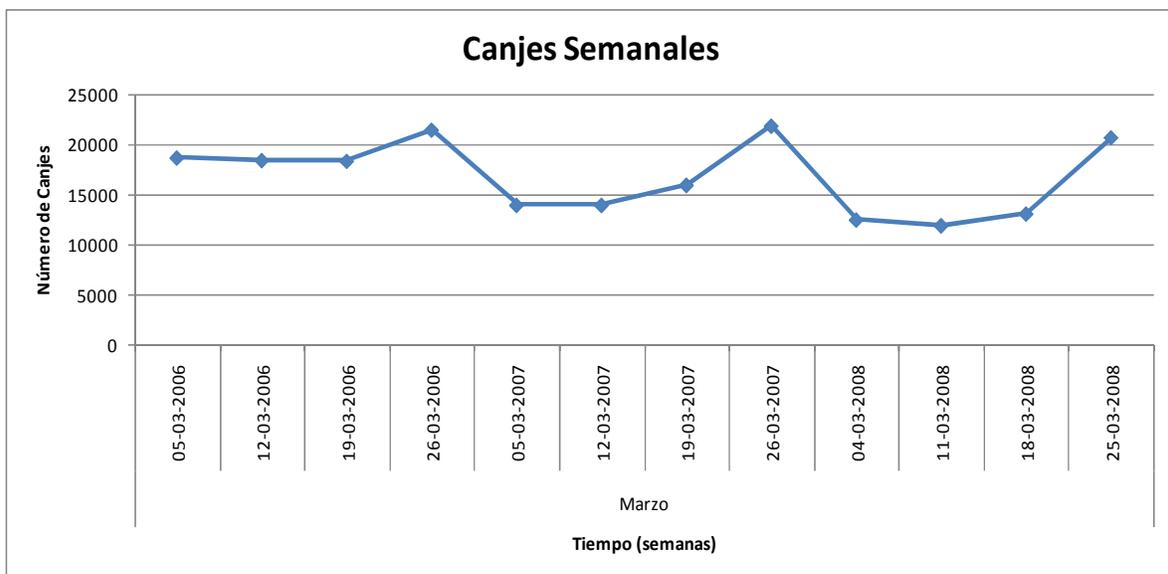
distribuyen los canjes en un determinado mes, para encontrar aquellas concentraciones que se pudiesen dar en una determinada semana. A continuación se muestran las series semanales de canjes en para los meses de Febrero, Marzo, Septiembre, Octubre y Diciembre. Es resto de los meses se encuentran en Anexo A: Canjes semanales por meses.

Gráfico 13: Canjes Semanales (Febrero)



Salvo el primer año, no existe un aumento significativo en alguna semana en los siguientes años. A pesar de lo anterior, hay que considerar que existe una disminución de los canjes en este mes (Ver **¡Error! No se encuentra el origen de la referencia.**).

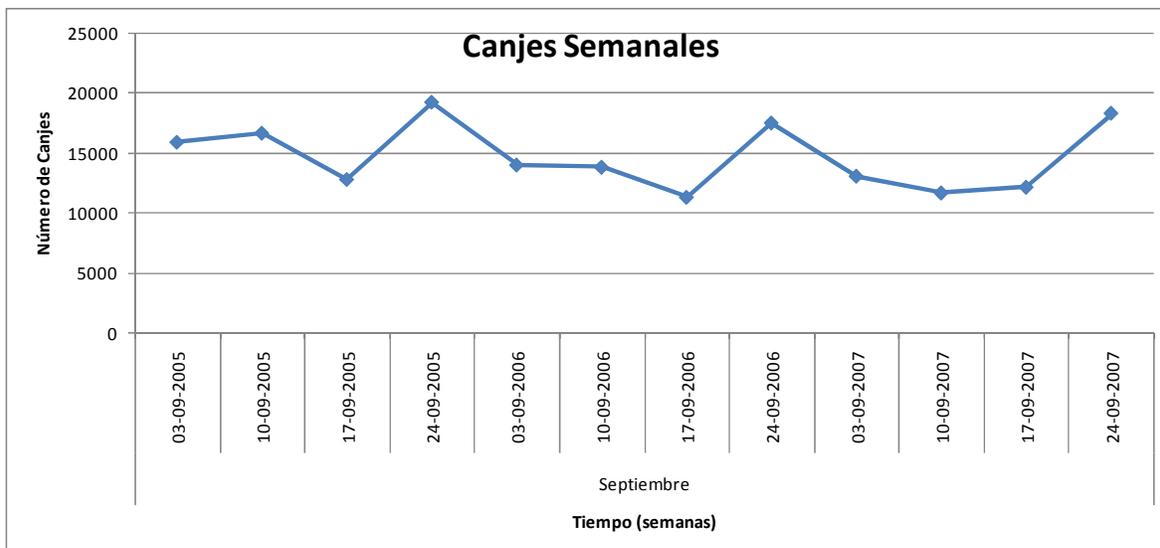
Gráfico 14: Canjes Semanales (Marzo)



Se observa un claro aumento de los canjes en la última semana de este mes. Se hubiese esperado un aumento de los canjes al comienzo del mes debido a las compras para la época escolar y no al final del periodo. Lo anterior se podría deber simplemente porque los puntos no

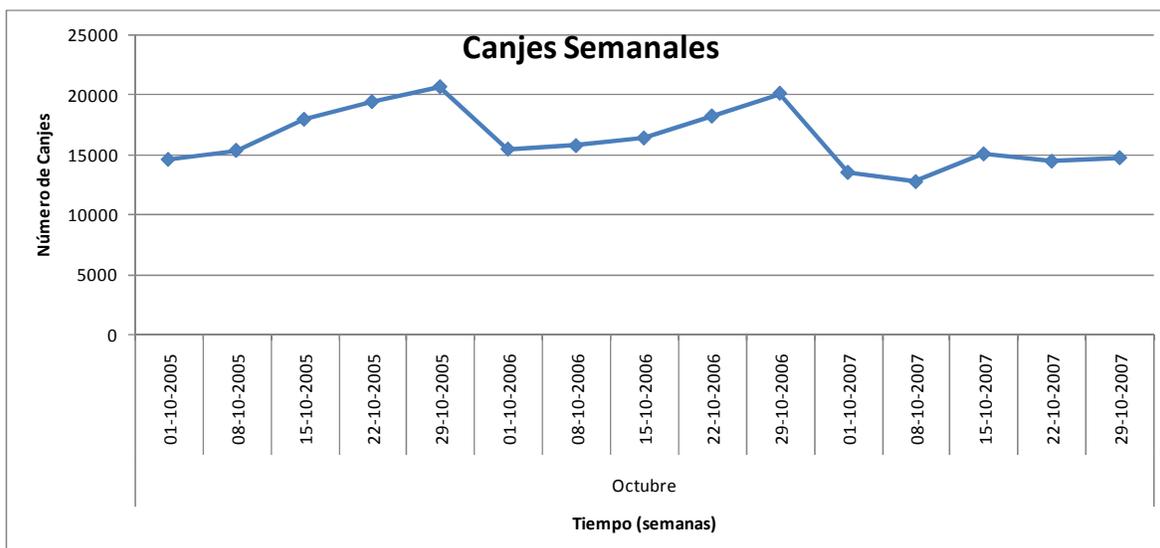
duran exactamente un año calendario sino que vencen a final de cada mes lo que podría provocar que las personas canjeen en la última semana antes que venzan los puntos del año anterior.

Gráfico 15: Canjes Semanales (Septiembre)



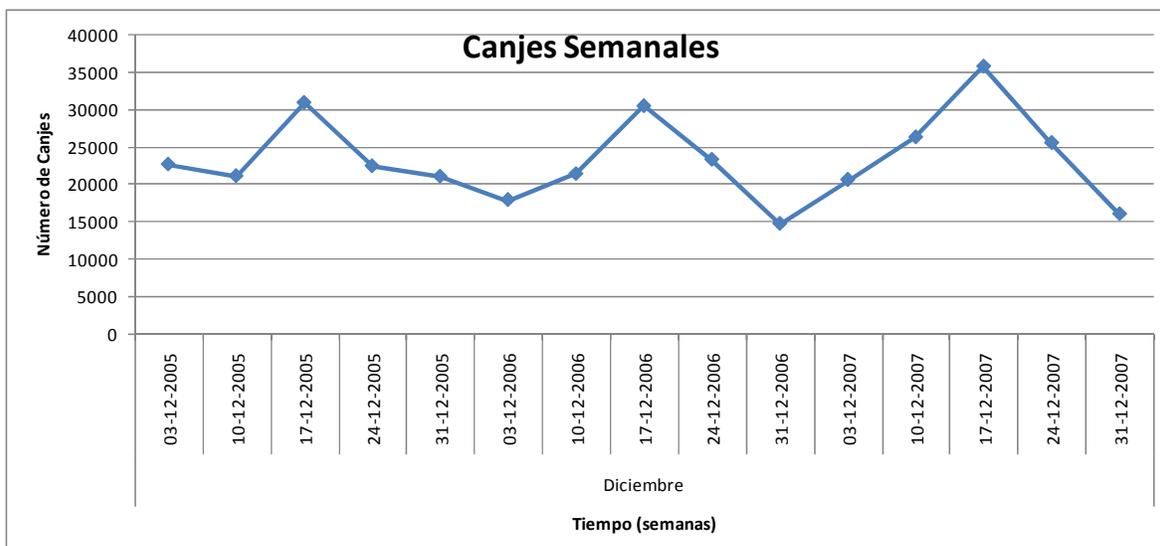
Se observa un aumento en la última semana del mes. Además, existe una disminución de los canjes en la semana de fiestas patrias y la semana previa a esta festividad.

Gráfico 16: Canjes Semanales (Octubre)



Se observa un aumento de los canjes en los primeros años y un pequeño aumento en el año 2007 y no el gran salto que se esperaba como los años anteriores.

Gráfico 17: Canjes Semanales (Diciembre)



Como se mencionó anteriormente, existe un aumento significativo de los canjes en la semana previa a navidad. Lo anterior se debe a la gran cantidad de ventas que se realizan en esta fecha y por lo tanto aumentan drásticamente los puntos entregados en este periodo.

- Antigüedad del Catálogo: En la siguiente tabla se encuentran las correlaciones entre la antigüedad del catálogo (semanas que lleva el catálogo en vigencia) y Número de canjes semanales. Además, se encuentra la correlación en el Número de canjes semanales con la antigüedad del catálogo al cuadrado (Suponiendo que la gente se demora un tiempo “cuadrático” en integrar el catalogo), con el logaritmo natural de la antigüedad del catálogo (Suponiendo que la gente acelera sus canjes al cambiar el catálogo) y la exponencial de la antigüedad al catálogo (Suponiendo que la gente se demora un tiempo “exponencial” en integrar el catalogo).

Tabla 5: Correlaciones entre Número de canjes y antigüedad del catálogo

Coefficiente Correlación	Antigüedad Catalogo	AntCatcuadrado	Ln(AntCat)	Exp(AntCat)
Productos Canjeados	0,026	-0,028	0,134	0,139

De la tabla anterior se desprende que no existe relación entre la antigüedad del catálogo y los canjes semanales de productos.

4.3 Estimación de demandas Globales

En los siguientes capítulos, se explicará cómo se estimará la demanda a nivel global de puntos y se desarrollará un modelo de estimación de canjes, con su respectiva validación.

4.3.1 Desarrollo del Modelo

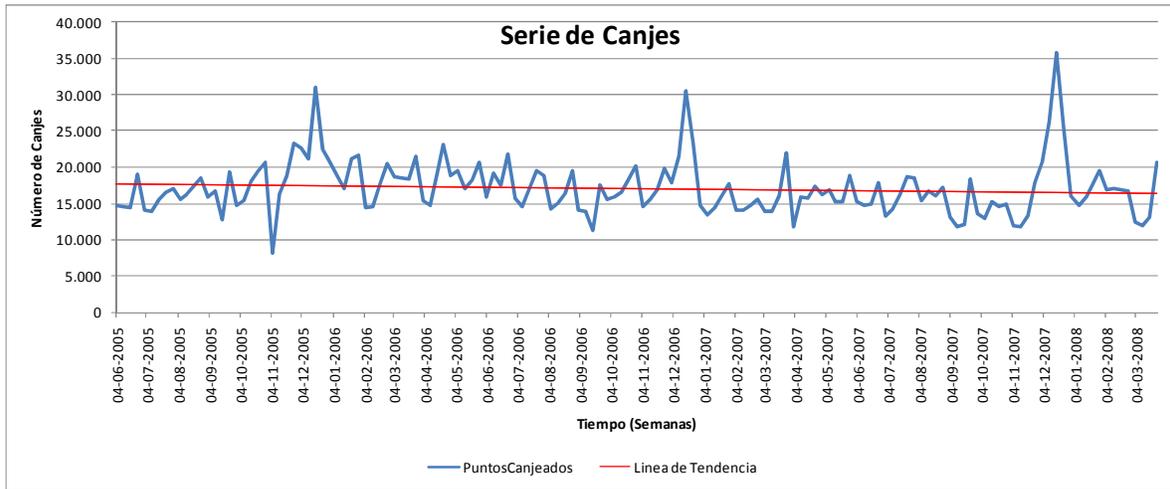
4.3.1.1 Estimación de orden del modelo

Una vez validadas las hipótesis, se procede a desarrollar un modelo de predicción de demanda mediante los pasos metodológicos propuestos por Box Jenkins.

El modelo de estimación de demanda a nivel global almacenará como datos de entrada los valores correspondientes al periodo comprendido entre las fechas 01 de Junio de 2005 y 01 de Abril de 2008.

A continuación se expone la serie de datos correspondiente al periodo antes mencionado.

Gráfico 18: Serie de Canjes



El siguiente gráfico muestra la media de canjes semanales para cada uno de los 10 trimestres de estudio.

Gráfico 19: Medias Trimestrales

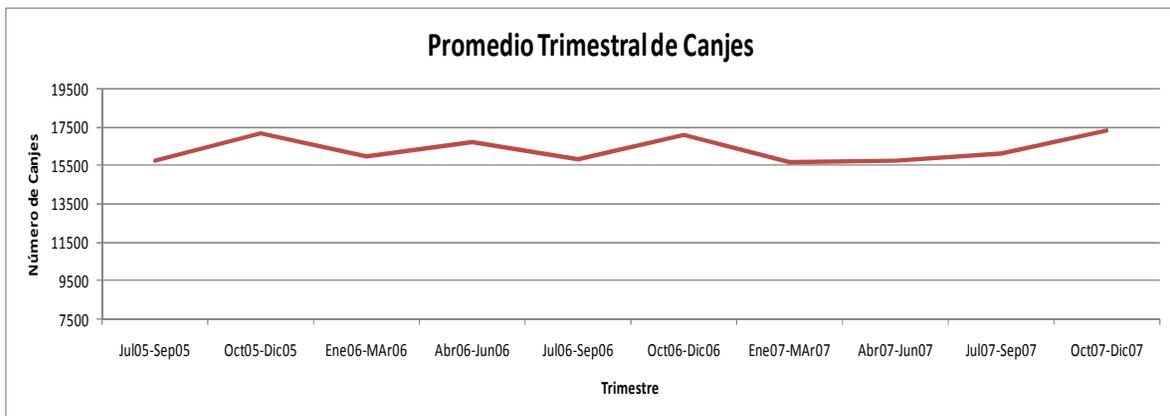
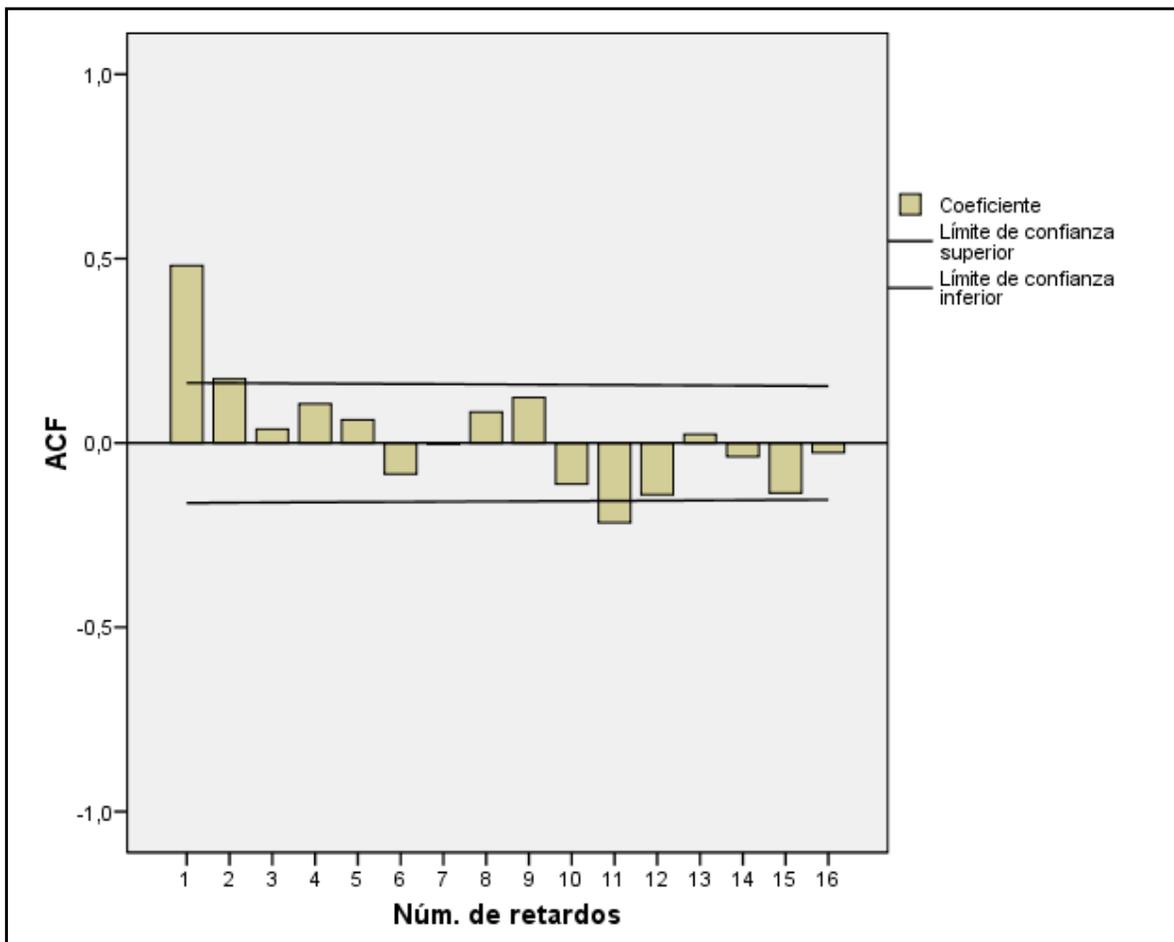


Tabla 6: Test Anova Diferencia de Medias

	Gl	F	Sig.
Inter-Grupos	9	0,79	0,318
Intra-Grupos	260	0,89	0,125
Total	269		

Luego de analizar la serie y la media trimestral se desprende que la serie es estacionaria, y que por lo tanto no necesitará ser diferenciada ya que no existe una diferencia de medias significativas. A continuación se exponen los gráficos de Autocorrelación y Autocorrelación Parcial.

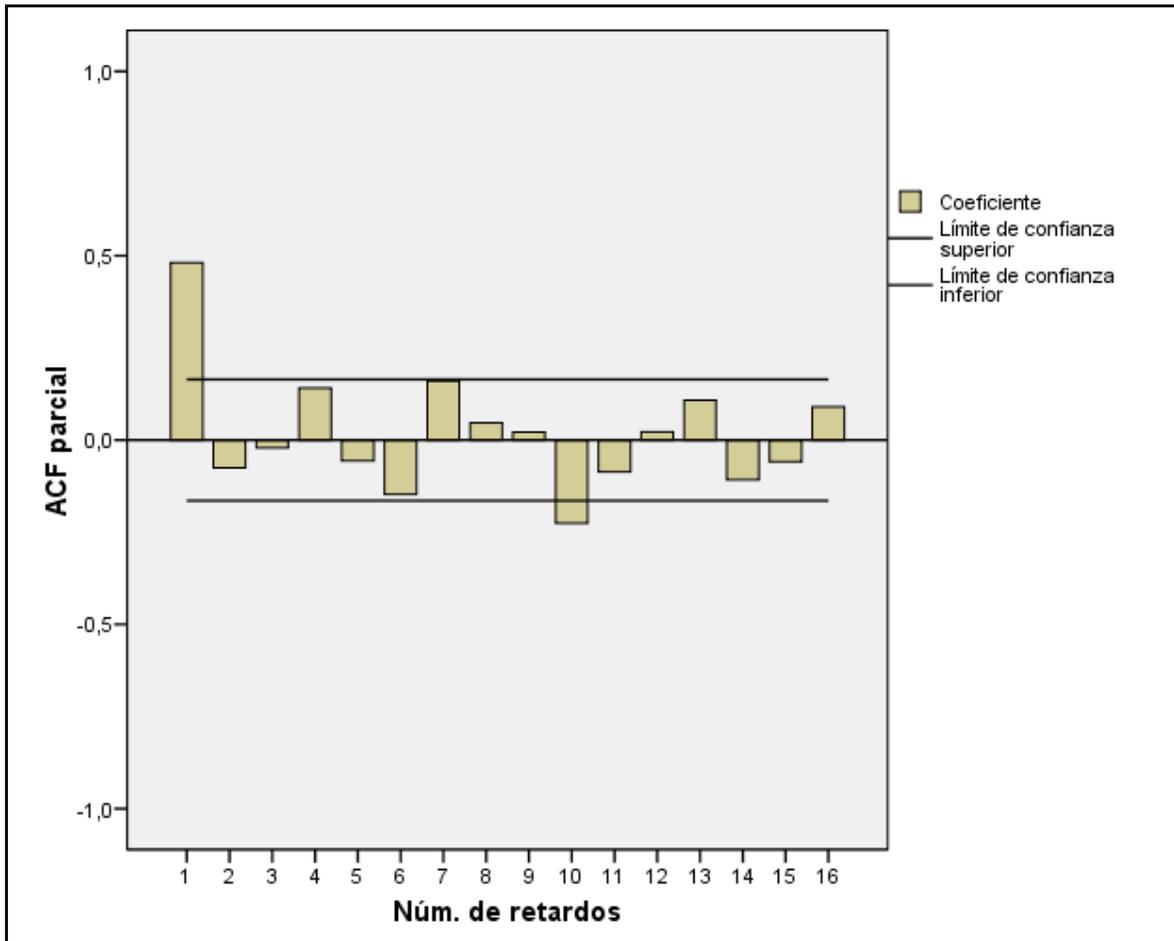
Gráfico 20: Autocorrelación Serie de Canjes



El gráfico anterior muestra la autocorrelación de los rezagos de la serie. Se observa que existe una autocorrelación en los retardos 1 y 2 lo que sugiere que la componente MA es de nivel 2 (MA(2)), es decir, que la serie está explicada por los errores de las dos semanas anteriores.

Para estudiar la componente autorregresiva de la serie se presenta el gráfico de autocorrelación de ésta.

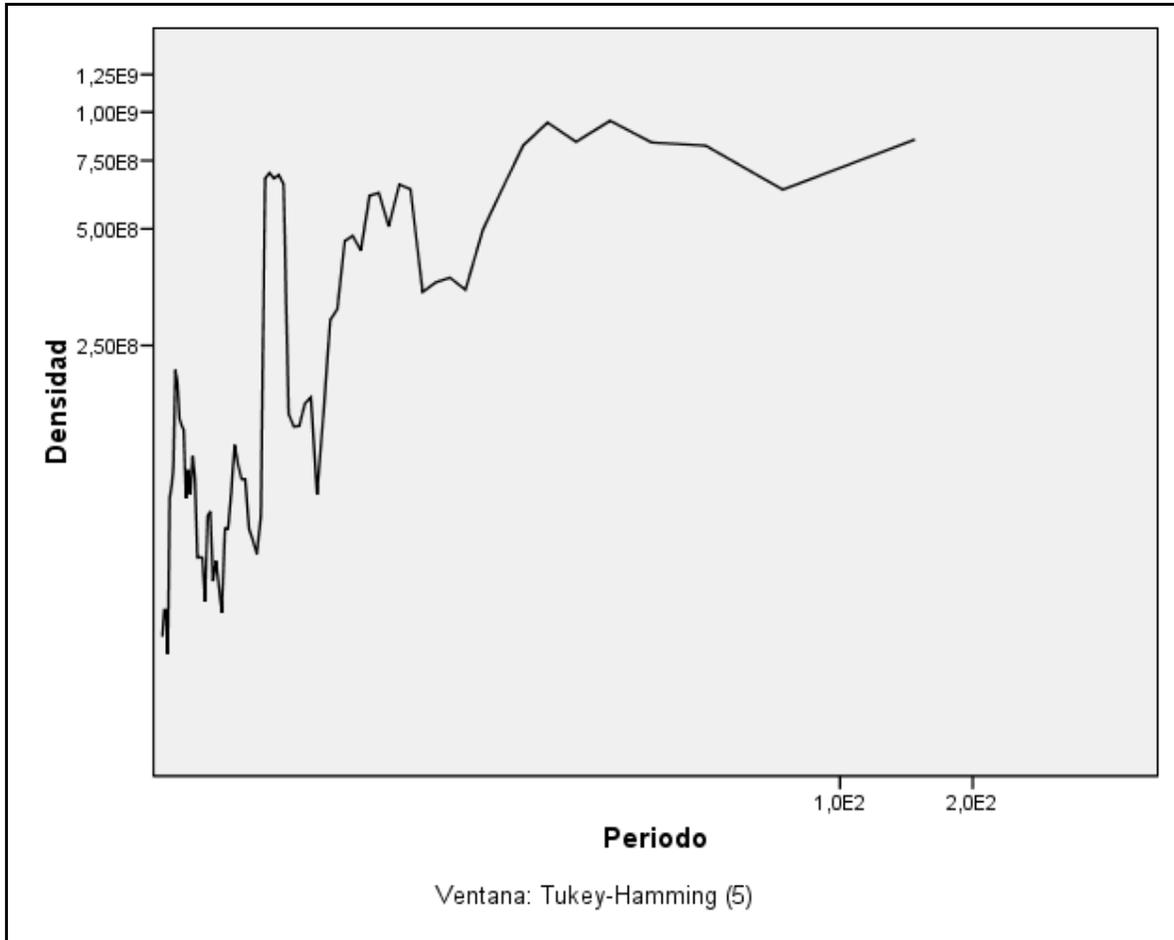
Gráfico 21: Autocorrelación Parcial Serie de Canjes



Se observa que existe una autocorrelación en el retardo 1 lo que sugiere que la componente AR es de nivel 1 (AR(1)). Es decir, la serie de canjes de productos está explicada por el valor de ésta en la semana inmediatamente anterior.

Finalmente, para estudiar si existe algún tipo de estacionalidad, se presenta el gráfico espectral.

Gráfico 22: Análisis Espectral Serie de Canjes



Al observar el gráfico anterior, se observa que de existir una periodicidad o estacionalidad, ésta es menor que 1 semana, por lo tanto, no será necesario considerarla en el modelo. Anualmente existe una periodicidad que se ve reflejada en el aumento de los canjes en las semanas de navidad, para esto se incorporará una variable al modelo que corresponderá a la semana de navidad.

Por todo lo anterior, se sugiere un modelo ARMA (1,2) que formalmente sería un modelo ARIMA (1,0,2)(0,0,0).

4.3.1.2 Estimación de parámetros del modelo

Una vez obtenido el orden del modelo se procede a estimar los parámetros del modelo, para esto se incorporan como variables externas al modelo todas aquellas variables que fueron validadas en la etapa validación de hipótesis. Sin embargo, no todas estas variables serán necesariamente significativas para el modelo, y se excluirán de éste aquellas que tengan una significancia menor a 0,05. Lo anterior significa que al incorporar una variable x al modelo se tendrá una probabilidad menor al 5% de equivocarse al rechazar la hipótesis nula de que la variable x no es significativa para el modelo.

El siguiente gráfico ilustra aquellas variables que resultaron significativas para el modelo con su respectiva significancia y el valor de su coeficiente.

Tabla 7: Estimación de Parámetros

Estimaciones de los parámetros				
Retardos no estacionales	Estimaciones	Error típico	t	Sig. aprox.
AR1	0,9994	0,0010	1012,45	0,00
MA1	0,5104	0,0833	6,13	0,00
MA2	0,3509	0,0821	4,27	0,00
Coefficientes de regresión				
Puntos Acumulados semana anterior	1,9E-05	6,6E-06	2,85	0,01
Fin de mes Enero	2,9E+03	1,2E+03	2,44	0,02
Fin de mes Marzo	5,7E+03	1,0E+03	5,56	0,00
Fin de mes May	3,6E+03	1,4E+03	2,49	0,01
Fin de mes Junio	4,3E+03	1,2E+03	3,64	0,00
Fin de mes Septiembre	4,3E+03	1,2E+03	3,66	0,00
Fin de mes Diciembre	4,1E+03	1,6E+03	2,61	0,01
Navidad	1,3E+04	1,3E+03	9,92	0,00

Las variables que resultaron ser significativas para el modelo son:

- Fines de mes de Enero, Marzo, mayo, Junio, Septiembre y Diciembre.
- La semana de navidad.
- Puntos Acumulados en la semana anterior.
- Además de las propias del modelo AR(1), MA(1), MA(2).

En el siguiente capítulo se discute sobre el ajuste del modelo.

4.3.2 Errores del Modelo

4.3.2.1 Ajuste del modelo

La siguiente tabla presenta el ajuste del modelo a los datos.

Tabla 8: Ajuste del Modelo

R	R^2	R^2 Ajustado	RMSE	MAPE	MAXP
0,93	0,86	0,84	1175	10,2%	15,1%

La tabla anterior ilustra el porcentaje de varianza explicada por el modelo, es decir, el modelo expuesto anteriormente explica un 84% de la varianza de los datos. Para determinar si existe algún otro factor significativo que no esté siendo considerado en el modelo y que pudiese explicar la varianza que no explicada por el modelo en sí, se estudiarán los errores de éste.

La siguiente tabla ilustra la autocorrelación existente entre los errores del modelo.

Gráfico 23: Autocorrelación Errores del modelo

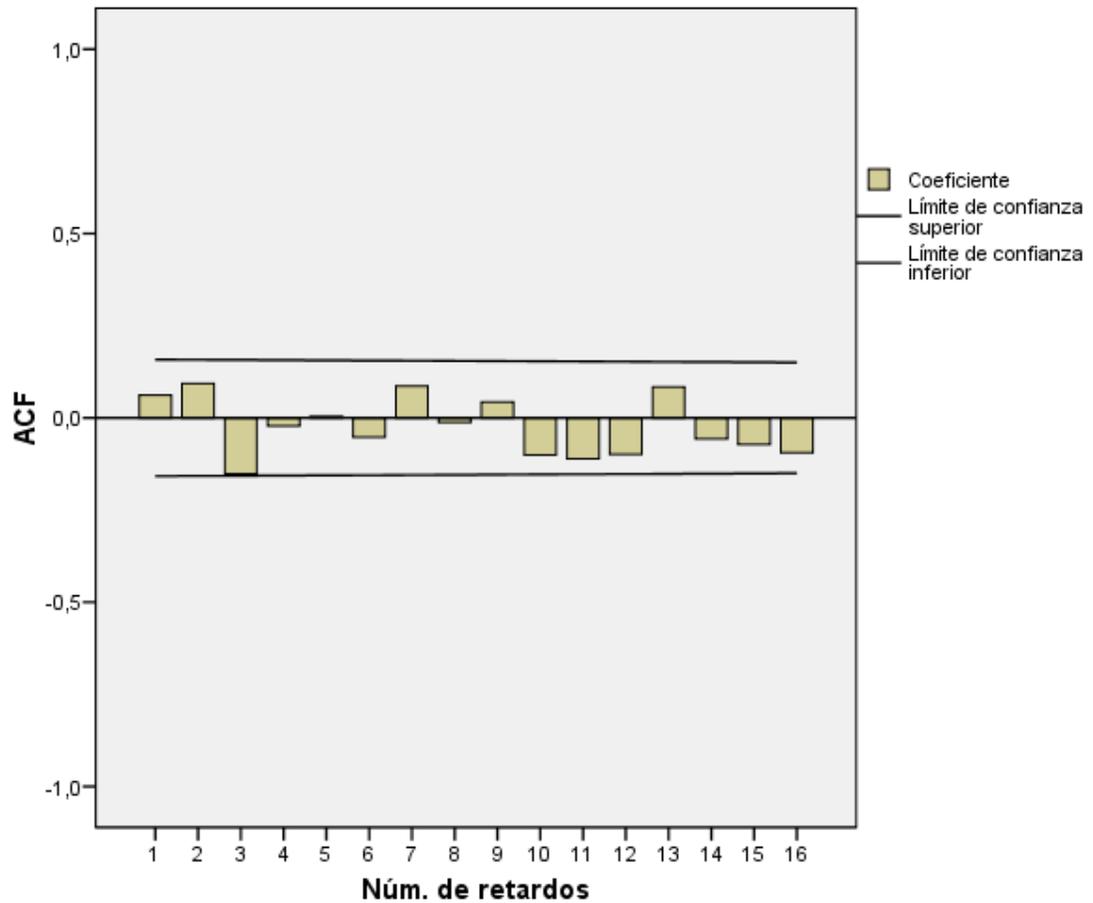


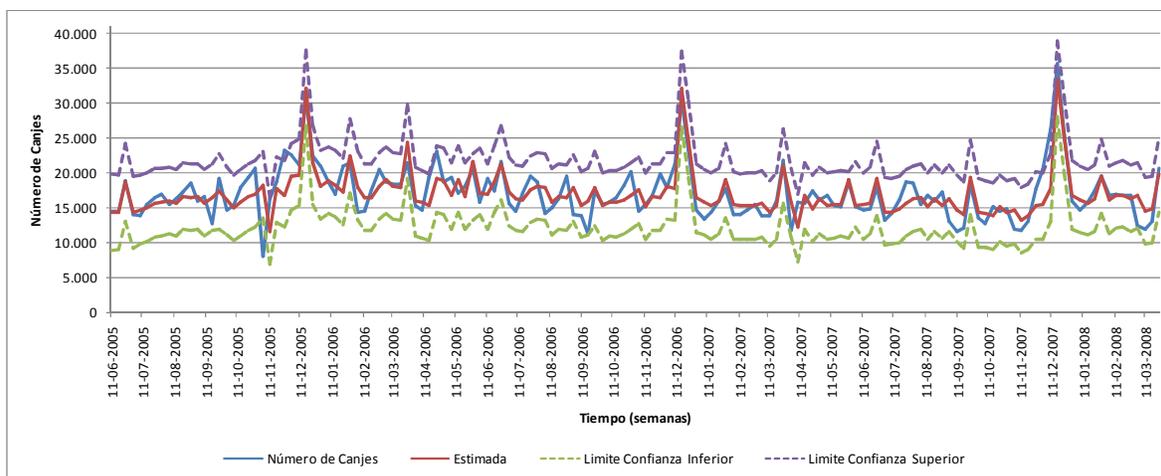
Tabla 9: Tabla de autocorrelaciones del Modelo

Autocorrelaciones					
Retardo	Autocorrelación	Error típico(a)	Estadístico de Box-Ljung		
	Valor	GI	Sig.(b)	Valor	GI
1	0,06	0,08	0,61	1	0,44
2	0,09	0,08	2,01	2	0,37
3	-0,15	0,08	5,76	3	0,12
4	-0,02	0,08	5,84	4	0,21
5	0,00	0,08	5,84	5	0,32
6	-0,05	0,08	6,29	6	0,39
7	0,09	0,08	7,53	7	0,38
8	-0,01	0,08	7,56	8	0,48
9	0,04	0,08	7,87	9	0,55
10	-0,10	0,08	9,59	10	0,48
A	El proceso subyacente asumido es la independencia (ruido blanco).				
B	Basado en la aproximación chi cuadrado asintótica.				

Se observa que no existe autocorrelación significativa en los errores del modelo. Para corroborar lo anterior se presenta la tabla de autocorrelaciones, que tiene como hipótesis nula la independencia del modelo, es decir, que el modelo no deja de lado algún factor relevante y que por lo tanto, los errores están siendo explicados solamente por el ruido blanco.

A continuación se presenta el gráfico de la serie de tiempo con el ajuste realizado por el modelo, junto con los límites de confianza al 95% superiores e inferiores.

Gráfico 24: Ajuste del Modelo



Se desprende del gráfico anterior que la serie de datos se encuentra siempre dentro del rango de límites de confianza.

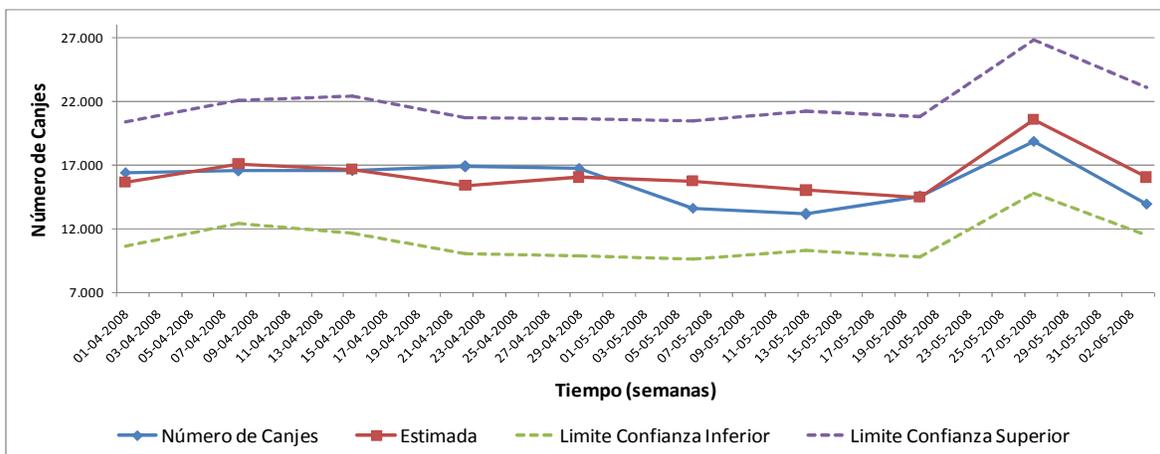
Una vez analizado los errores del modelo, se estudiará la eficacia en la predicción de éste.

4.3.2.2 Predicción del Modelo

Para validar la calidad de predicción del modelo, se pronosticarán los canjes para el periodo comprendido entre las fechas 01 de Abril de 2008 al 10 de Junio de 2008.

A continuación se presenta el gráfico de la serie junto con la predicción y los límites de confianza superiores e inferiores.

Gráfico 25: Predicción del Modelo



La siguiente tabla presenta los canjes estimados y los canjes reales para las 10 semanas en estudio.

Tabla 10: Canjes estimados v/s Canjes Reales

Semana	Número de Canjes	Estimada	Error	Error %
1	16.407	15.685	749	4,8%
2	16.546	17.110	-564	3,3%
3	16.588	16.644	-56	0,3%
4	16.874	15.385	1.489	9,7%
5	16.768	16.025	743	4,6%
6	13.593	15.716	-2.123	13,5%
7	13.138	15.007	-1.869	12,5%
8	14.487	14.442	45	0,3%
9	18.892	20.584	-1.692	9,0%
10	13.960	16.013	-2.053	14,7%

En primera instancia se desprende del gráfico y tabla anteriores que el modelo ha predicho bien. En la siguiente tabla se presentan los indicadores de error definidos anteriormente.

Tabla 11: Errores de Estimación

RMSE	MAPE	MAXP	NMSE	SSMP	SSM
1369	10,5%	15,6%	0,007	1,65%	275

Tabla 12: Ajuste de la Estimación

R	R^2	R^2 Ajustado
0,90	0,82	0,82

Tabla 13: Errores de calibración del modelo

R	R ²	R ² Ajustado	RMSE	MAPE	MAXP
0,93	0,86	0,84	1175	10,2%	15,1%

Se observa que el error de estimación semanal es de un 10,5% que representa en promedio alrededor de 1369 unidades. Además se tiene un máximo error de pronóstico de un 15,6% que se encuentra en el mes de Mayo donde la estimación es superior a los canjes de la semana. Comparando los errores del modelo con respecto al actual sistema de estimación se observa que el modelo estima a nivel **semanal** un error mucho menor que el modelo **mensual** de la empresa, que como vimos anteriormente posee un MAPE de un 30%. Además, no existe un sobre ajuste de los datos en el periodo de calibración de los parámetros, esto se desprende al observar el indicador de ajuste (R^2) que no varía significativamente entre el periodo de prueba y el de validación. Además, no existe mucha diferencia entre los distintos errores entre ambos periodos (tablas 11 y 13).

Muchas veces se incurre en el error de ocupar un modelo complicado para lograr una estimación mejor, cuando a veces un modelo más simple podría predecir de forma parecida pero su implementación es más rápida y simple. A pesar de que, el modelo no es complicado de implementar y que las variables introducidas son directas y no requieren mucho tiempo en su cálculo, para no incurrir en el error antes mencionado se comparará el modelo con uno de suavizado exponencial. A continuación se exponen los resultados del modelo antes descrito.

Tabla 14: Coeficientes Modelo suavizado Exponencial

Serie	Alpha (Nivel)	Gamma (Tendencia)
Canjes	0,4	0

Tabla 15: Errores Modelo Suavizado Exponencial

R	R ²	R ² Ajustado	RMSE	MAPE	MAXP
0,84	0,71	0,69	2130	15%	27,5%

Se observa que con un modelo de suavizado exponencial el modelo se ajusta de menor medida a la serie (R^2 menor al modelo ARIMA). Además, el MAPE de este modelo es mayor al Máximo MAPE del modelo ARIMA, por lo que sugiere utilizar el modelo ARIMA para las futuras estimaciones. Sin embargo, este último modelo sigue siendo mucho mejor al modelo existente en la empresa.

El Siguiete capítulo se centrará en cómo se dividirá esta estimación en los 8 niveles de puntos existentes en la empresa.

4.4 Estimación de demanda por nivel de puntos

4.4.1 Desarrollo del Modelo

Como se mencionó anteriormente, para esta etapa del proceso se utilizará un modelo de estimación de demanda basada en participación de mercados auto regresiva.

En los siguientes dos gráficos, se muestran las series de canjes por nivel de puntos por porcentajes (Market Share).

Gráfico 26: Shares de canjes por Nivel (Nivel 1-Nivel 4)

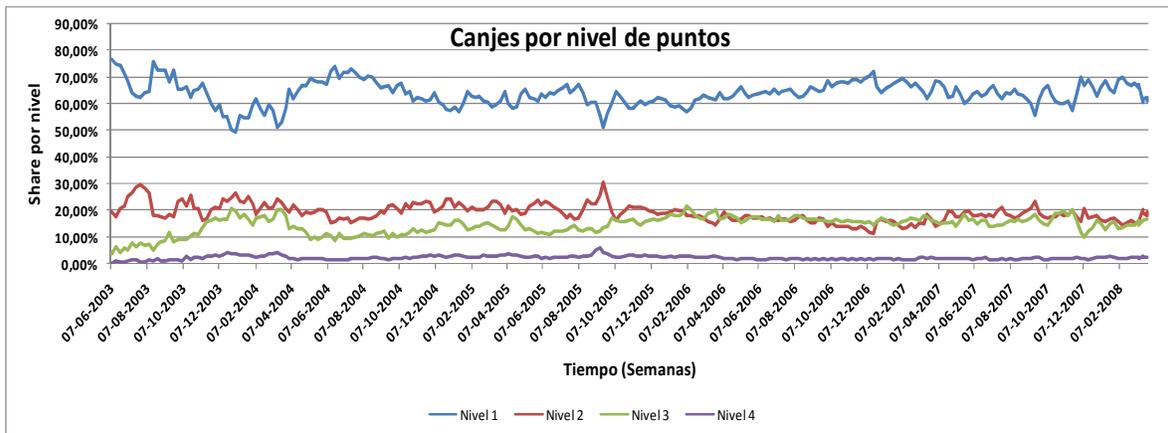
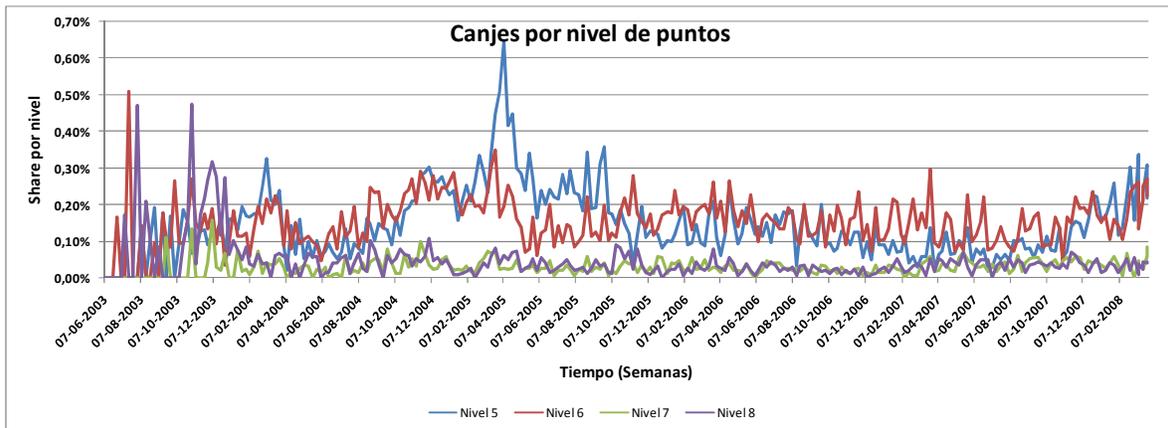


Gráfico 27: Shares de canjes por Nivel (Nivel 5-Nivel 8)



Como se mencionó en el capítulo de estudio de la situación actual, los primeros 4 niveles de puntos concentran más del 97% de los canjes por semana y que existe una gran volatilidad en los niveles más altos.

El modelo de estimación de demanda basado en participación de demanda autorregresivo, se utilizará para estimar la demanda en esta etapa.

$$A_{jt} = \sum_{T=t-1}^{T=t_0} \alpha_T * A_{jT} + \gamma_T * (A_{j(t-1)} - A_{j(t-2)}) + \sum_{q=1}^{q=Q} \beta_q * x_{jq} + C_j$$

Modelo basado en participación de mercado autorregresivo

A continuación se exponen los parámetros del modelo anterior para estimar la atracción para cada nivel de puntos.

Tabla 16: Parámetros del Modelo

	Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Nivel 6	Nivel 7	Nivel 8
Constante	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Lag 1	0,90	0,83	0,87	1,00	1,00	1,00	0,55	0,49
Lag 2	-0,23	0,00	0,00	0,00	0,00	0,00	0,44	0,51
Lag 3	0,32	0,17	0,13	0,00	0,00	0,00	0,00	0,00
Diferencia	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

De la tabla anterior se desprende que la atracción de los productos de un determinado nivel es una combinación lineal de la atracción que tuvo el nivel en el periodo o semana anterior. Como se expuso anteriormente, la atracción real del nivel en un determinado periodo, viene dado por el logaritmo del market share del nivel en ese periodo (ver ecuación anterior).

4.4.2 Errores del Modelo

En este capítulo se presenta el análisis de los errores del modelo de predicción de canjes por nivel de puntos. Estos errores serán los errores de ajuste del modelo y de la capacidad de predicción de éste. A continuación se presentará el análisis para el nivel 1, los siguientes niveles se encuentran en el Anexo B: Ajuste del modelo por nivel.

4.4.2.1 Ajuste del modelo

El ajuste de los parámetros del modelo de predicción por nivel de puntos se realizó para la serie de canjes comprendida en el mismo periodo de ajuste del modelo de predicción de demanda global antes descrito, es decir, entre el 01 de Junio de 2005 y el 01 de Abril de 2008.

La siguiente tabla presenta la calidad de ajuste del modelo para cada nivel.

Tabla 17: Ajuste del modelo

Indicador	Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Nivel 6	Nivel 7	Nivel 8
R	0,97	0,95	0,96	0,92	0,88	0,91	0,89	0,89
R^2	0,94	0,91	0,92	0,84	0,78	0,82	0,80	0,80
R^2 Ajustado	0,94	0,90	0,92	0,83	0,77	0,82	0,79	0,80

Se observa de la tabla anterior que existe un 94% de la varianza explicada por el modelo para el nivel 1. Una vez estudiado el ajuste del modelo, se procede a analizar la significancia de los parámetros de éste. Para lo anterior se presenta la siguiente tabla.

Tabla 18: Significancia de las variables del modelo

Variables Explicativas	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	T	Sig.
LAGS(A,1)	0,90	0,08	0,91	11,64	0,00
LAGS(A,2)	-0,23	0,11	-0,23	-2,16	0,03
LAGS(A,3)	0,32	0,08	0,32	4,10	0,00

La atracción correspondiente al nivel 1 en la semana t, es una combinación lineal del valor que toma ésta en las 3 semanas anteriores. Además, se puede observar de la tabla anterior que todas las variables autorregresivas son significativas al 5%, es decir, que existe una probabilidad menor al 5% para cada variable de equivocarse al rechazar la hipótesis nula de no significancia de las variables.

4.4.2.2 Predicción del Modelo

Una vez realizado el estudio de los errores de ajuste del modelo, es necesario estudiar la capacidad predictiva de éste. Para lo anterior, se ha realizado una estimación para el periodo comprendido entre el 01-04-2008 y el 10-06-2008 (10 semanas).

Como el modelo predice la utilidad del nivel y no los canjes de éste se calcula el market share de cada nivel a través de la Ecuación 19. Una vez obtenidos los market share se procede a multiplicar éstos por la cantidad estimada a través del modelo de predicción de demanda global

La siguiente tabla presenta la predicción del modelo para las 10 semanas en estudio.

Tabla 19: Predicción de canjes semanales

Semana	Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Nivel 6	Nivel 7	Nivel 8
1	9141	3312	2691	298	29	29	6	6
2	9721	3725	3002	421	37	28	11	6
3	9819	3214	2697	335	64	42	12	5
4	8978	3254	2643	292	28	28	11	5
5	9104	3489	2811	395	34	27	10	6
6	9272	3035	2547	316	61	40	11	5
7	9047	2946	1995	284	19	22	6	4
8	7965	3432	2451	300	27	26	4	5
9	11564	4558	3628	471	46	32	8	9
10	9496	3319	2611	302	37	30	12	10

El error porcentual por semana para cada nivel de puntos se expone en la siguiente tabla.

Tabla 20: Error porcentual Semanal

Semana	Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Nivel 6	Nivel 7	Nivel 8
1	6%	1%	5%	40%	32%	13%	45%	28%
2	4%	27%	15%	3%	52%	45%	28%	26%
3	2%	1%	2%	13%	79%	14%	18%	23%
4	9%	5%	8%	42%	35%	16%	4%	41%
5	11%	17%	6%	5%	56%	49%	34%	16%
6	8%	24%	50%	1%	204%	67%	172%	27%
7	23%	2%	6%	13%	36%	22%	56%	45%
8	2%	16%	0%	25%	28%	4%	58%	30%
9	3%	28%	24%	10%	10%	24%	62%	55%
10	11%	18%	39%	17%	36%	14%	193%	38%

De lo anterior se obtiene los indicadores de error de estimación son los siguientes.

Tabla 21: Indicadores de error

Indicador	Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Nivel 6	Nivel 7	Nivel 8
MAPE	8%	14%	16%	17%	57%	27%	67%	33%
MAXP	23%	27%	50%	42%	204%	67%	193%	55%
RMSE	836	518	450	102	26	13	6	4
NMSE	0,008	0,025	0,030	0,077	0,403	0,154	0,385	0,349
SSMP	3%	1%	2%	14%	25%	17%	25%	22%
SSM	341	25	54	62	14	8	4	2

Para los cuatro primeros niveles, que son los que concentran más del 95% de los canjes, se tiene un MAPE menor al 20% con un MAXP menor al 50%. Para los niveles más altos se tiene un error mayor pero si se observa el RMSE de cada nivel el error no supera los 26 productos para los niveles más altos en comparación a los niveles más bajos que este indicador es mucho más alto.

Considerando ahora que existen dos semanas (6 y 10) en que los canjes totales disminuyen drásticamente y que son los máximos errores de la estimación global de canjes, este error es arrastrado a los niveles. Eliminando estas dos semanas, los indicadores serían los siguientes.

Tabla 22: Indicadores de error

Indicador	Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Nivel 6	Nivel 7	Nivel 8
MAPE	7%	11%	8%	18%	41%	23%	38%	33%
MAXP	22%	27%	23%	42%	79%	48%	62%	55%
RMSE	837	510	309	113	25	13	6	4

Tabla 23: Ajuste del modelo (periodo de validación)

Indicador	Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Nivel 6	Nivel 7	Nivel 8
R	0,96	0,93	0,94	0,89	0,85	0,88	0,85	0,86
R ²	0,92	0,87	0,89	0,79	0,73	0,78	0,73	0,74
R ² Ajustado	0,92	0,87	0,89	0,78	0,73	0,77	0,72	0,73

En general el error promedio disminuye en especial en los niveles en donde éste era mayor, esto se debe a que se eliminaron aquellas semanas en donde el error global era mayor. Además, no existe una disminución de los indicadores de ajuste del modelo, por lo que se desprende que el modelo no está sobre ajustado a los datos del periodo de calibración.

Nuevamente se comparará el resultado del modelo propuesto con uno de suavizado exponencial. A continuación se exponen los resultados de este modelo.

Tabla 24: Estimación Parámetros Suavizado Exponencial

Serie	Alpha (Nivel)	Gamma (Tendencia)
Nivel 1	0,7	0
Nivel 2	0,5	0
Nivel 3	0,3	0
Nivel 4	0,4	0
Nivel 5	0,2	0
Nivel 6	0,2	0
Nivel 7	0,1	0
Nivel 8	0,1	0

Tabla 25: Errores Suavizado

Indicador	Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Nivel 6	Nivel 7	Nivel 8
MAPE	22%	18%	27%	21%	33%	24%	57%	26%
MAXP	53%	43%	56%	50%	96%	49%	87%	71%
RMSE	2197	744	428	178	20	10	7	5
NMSE	0,056	0,148	0,132	0,167	0,305	0,103	0,956	0,610
SSMP	2,5%	6,1%	11,2%	7,8%	10,1%	12,5%	28,6%	17,5%
SSM	201	145	265	32	7	5	4	2

A continuación se exponen los resultados para un modelo de serie de tiempo Arima en todos los niveles de puntos.

Tabla 26: Errores ARIMA

Indicador	Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Nivel 6	Nivel 7	Nivel 8
MAPE	20%	18%	15%	16%	30%	20%	52%	29%
MAXP	40%	29%	40%	35%	83%	41%	71%	75%
RMSE	1915	377	413	74	20	11	8	5
NMSE	0,036	0,014	0,029	0,038	0,236	0,106	0,856	0,678
SSMP	1,3%	2,1%	7,9%	10,1%	15,9%	14,2%	38,6%	27,5%
SSM	147	74	219	46	11	7	6	3

De la tabla anterior se desprende que para los niveles bajos (1-4) el modelo propuesto predice mejor, pero para los niveles más altos se sugiere usar un modelo de suavizado exponencial que disminuye el error promedio y máximo y es más simple de implementar que un modelo Arima en donde además, el indicador SSM es mayor.

4.5 Estimación de demanda por tipo de productos

4.5.1 Desarrollo del Modelo

Una vez obtenido un modelo de estimación de demanda para los canjes globales y por nivel de puntos, se desarrollará un modelo para los tipos de productos o categorías pertenecientes a los primeros 4 niveles⁸.

Solamente el nivel 1 se expondrá en detalle, el resto de los niveles es una réplica del procedimiento y los detalles se encuentran en Anexo C: Categoría de productos por nivel.

Para el desarrollo de estos modelos, se utilizará el modelo de predicción de demanda basado en participación de mercado autorregresivo, debido a que las series no son continuas y por lo tanto no se cumplen las condiciones para un modelo de serie de tiempo.

$$A_{jt} = \sum_{T=t-1}^{T=t_0} \alpha_T * A_{jT} + \gamma_T * (A_{j(t-1)} - A_{j(t-2)}) + \sum_{q=1}^{q=Q} \beta_q * x_{jq} + C_j$$

Modelo de Participación Autorregresivo

Las variables exógenas que se incorporarán al modelo son aquellas que tienen que ver con la cantidad de productos de la categoría en el nivel. Las variables consideradas para determinar las utilidades de cada categoría en cada mes son:

N_{jt} = Número de productos de la categoría en el periodo t.

⁸ Debido a que representan más del 97% de los canjes

SN_{jt} = Share de la cantidad de productos de la categoría con respecto al total del nivel en el periodo t (se mantiene durante el catálogo).

$\ln(SN_{jt})$ = Logaritmo natural de la variable anterior.

$SN_{jt} - SN_{jt-1}$ = Es la diferencia que existe entre el share de la cantidad de productos de una categoría en el nivel con respecto a un periodo a otro.

$\ln(SN_{jt}) - \ln(SN_{jt-1})$ = Es la diferencia que existe entre los logaritmo natural del share de la cantidad de productos de una categoría en el nivel con respecto a un periodo a otro.

Una vez definidas las variables incorporadas en el modelo, se procede a determinar los coeficientes que acompañan a éstas. Cabe destacar que para cada categoría no se debe tener más de una variable relacionada con el número de productos, ya que son directamente proporcionales y no estarían agregando información adicional al modelo solo aumentarían su complejidad, a no ser que esta variable sea la diferencia del número de productos.

A continuación se exponen las categorías a estudiar en el nivel 1.

Tabla 27: Categorías Nivel 1

Administración	Básicos Decohogar	Blanco	Deportes
Electrohogar	Comunicación, Electrónica	Electricidad	Herramientas y maquinarias
Menaje	Muebles y organización		

A continuación se presenta los coeficientes de cada variable para determinar la utilidad en el periodo t de cada una de las categorías antes mencionadas divididas por empresa. En primer lugar se presentarán las categorías pertenecientes para la empresa principal o sostenedora y a continuación a aquellas que pertenecen a la empresa perteneciente al Holding. Solo se presentan aquellas variables significativas, la significancia de éstas se exponen en el Anexo 3.

Tabla 28: Coeficientes del modelo (Categorías Empresa Principal)

Coefficientes	UAdministración	UBásicos Decohogar	UBlanco	UDeportes	UElectro hogar
Constante	0	0	0	0	0
Ln(S)	0	0	0	0	0
N	0	0	0	0	0
SN	0	0	0	0	0,704
Ln(SN)	0,091	0	0	0	0
lag1(A)	0,916	0,861	0,862	1,056	0,883
lag2(A)	0	0	0	-0,155	0
lag3(A)	0	0	0	0	0
diff(A)	0	0	0	0	0
diff(SN)	0	0	0	0	0
diff(Ln(SN))	0	0	0	0	0

En la tabla anterior se observa que todas las atracciones son explicadas principalmente por el valor que estas tuvieron la semana anterior. Solamente dos categorías son explicadas además por el número de productos que éstas posean.

A continuación se exponen las categorías para la empresa del Holding.

Tabla 29: Coeficientes del modelo (Categorías Empresa Holding)

Coefficientes	UComunicación. electronica	UElectricidad	UHerramientas y maquinarias	UMenaje	UMuebles y organización
Constante	0	-0,697	-0,29	0	0
Ln(S)	0	0	0	0	0
N	0	0	0	0	0
SN	0	0	0	0	0
Ln(SN)	0	0,151	0	0,063	0
lag1(A)	0,8	0,366	0,931	0,931	,998
lag2(A)	0,184	0,407	0	0	0
lag3(A)	0	0	0	0	0
diff(A)	0	0	0	0	0
diff(SN)	0	0	0	0	0
diff(Ln(SN))	0	0,634	0,511	0	0

Se desprende de la tabla anterior que el número de productos de la categoría explica mejor las atracciones de las categorías pertenecientes a la empresa del Holding que a la empresa principal. Además, dependen en general del valor de la atracción de una semana anterior.

4.5.2 Errores del Modelo

En este capítulo se presenta el análisis de los errores del modelo de predicción de canjes por categoría de productos. Estos errores serán los errores de ajuste del modelo y de la capacidad de predicción de éste.

4.5.2.1 Ajuste del modelo

El ajuste de los parámetros del modelo de predicción por nivel de puntos se realizó para las series de canjes comprendidas, entre el 01 de Junio de 2005 y el 01 de Abril de 2008.

Tabla 30: Error de Ajuste de los modelos

Categoría	R	R ²	R ² Ajustado	Error Tip. De la Estimación
Administración	0,95	0,89	0,89	0,37
Básicos Decohogar	0,95	0,90	0,90	0,14
Blanco	0,86	0,73	0,73	2,11
Deportes	0,90	0,82	0,81	1,16
Electrohogar	0,94	0,89	0,89	0,11
Comunicación, Electrónica	0,93	0,86	0,86	2,63
Electricidad	0,90	0,81	0,81	2,04
Herramientas y maquinarias	0,88	0,77	0,76	0,44
Menaje	0,95	0,90	0,89	0,20
Muebles y organización	0,94	0,88	0,88	2,13

Se observa de la tabla anterior que los modelos se ajustan bien para la serie de datos estudiada, la varianza explicada en general es mayor al 80%.

Una vez estudiado el ajuste del modelo, se procede a analizar la significancia de los parámetros de los modelos. Se presentarán solo tres categorías el resto se encuentran en Anexo C: Categoría de productos por nivel.

Tabla 31: Significancia de las variables del modelo (Administración)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	T	Sig.
LAGS(A,1)	0,92	0,03	0,92	30,08	0,00
LNSN	0,09	0,03	0,08	2,65	0,01

Tabla 32: Significancia de las variables del modelo (Básicos Decohogar)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	T	Sig.
LAGS(A,1)	0,19	0,04	0,14	4,61	8,97E-06

Tabla 33: Significancia de las variables del modelo (Blanco)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	T	Sig.
LAGS(A,1)	0,86	0,04	0,86	20,46	0,00

Se observa que los coeficientes para las categorías antes expuestas son significativos al 95%. Por lo tanto, se concluye que existe un ajuste adecuado del modelo. Sin embargo, es

necesario estudiar no solo el ajuste del modelo si no que también la capacidad predictiva que tenga el modelo. En el siguiente capítulo se expone lo antes mencionado.

4.5.2.2 Predicción del Modelo

Una vez realizado el estudio de los errores de ajuste del modelo, es necesario estudiar la capacidad predictiva de éste. Para lo anterior, se ha realizado una estimación para el periodo comprendido entre el 01-04-2008 y el 10-06-2008 (10 semanas).

Como el modelo predice la utilidad de cada categoría, se calcula el market share de cada nivel a través de la Ecuación 19. Una vez obtenidos los market share se procede a multiplicar éstos por la cantidad estimada a través del modelo de predicción de demanda por niveles de puntos

Las siguientes dos tabla presenta la predicción del modelo para las 10 semanas en estudio.

Tabla 34: Predicción de canjes semanales (Empresa Principal)

Semanas	Administración	Básicos Decohogar	Blanco	Deportes	Electrohogar
1	92	4380	696	446	2118
2	106	5174	729	618	1529
3	96	4459	760	470	2253
4	100	4102	875	420	2191
5	125	4359	659	379	2631
6	102	4605	763	295	2418
7	99	4023	1016	162	2521
8	90	3968	749	825	1308
9	148	5100	1416	962	2754
10	111	5035	1088	460	1974

Tabla 35: Predicción de canjes semanales (Empresa Holding)

Semanas	Comunicación, Electrónica	Electricidad	Herramientas y Maquinarias	Menaje	Muebles y Organización
1	678	6	15	362	246
2	793	9	10	427	245
3	856	9	8	541	312
4	548	8	15	419	275
5	294	6	14	402	213
6	255	8	14	527	240
7	261	11	13	650	252
8	216	6	17	515	244
9	174	7	18	643	311
10	82	5	4	470	231

Los errores porcentuales por semana para cada categoría se exponen en las siguientes tablas.

Tabla 36: Error porcentual Semanal (Empresa Principal)

Semanas	Administración	Básicos Decohogar	Blanco	Deportes	Electrohogar
1	40,4%	8,9%	14,0%	15,8%	3,8%
2	34,4%	8,2%	4,5%	8,2%	10,9%
3	30,3%	7,8%	0,1%	9,5%	13,8%
4	36,9%	3,4%	11,6%	14,0%	2,6%
5	25,3%	11,0%	11,9%	3,6%	19,3%
6	21,5%	17,7%	13,1%	6,8%	25,1%
7	10,0%	36,1%	23,4%	23,9%	45,7%
8	33,7%	9,7%	6,6%	10,8%	13,0%
9	29,2%	11,6%	3,1%	7,4%	18,3%
10	22,6%	20,7%	6,2%	2,5%	26,3%

Tabla 37: Error porcentual Semanal (Empresa Holding)

Semanas	Comunicación, Electrónica	Electricidad	Herramientas y Maquinarias	Menaje	Muebles y Organización
1	37,9%	58,8%	31,0%	36,0%	39,9%
2	32,7%	12,6%	20,4%	29,5%	33,1%
3	27,6%	27,1%	14,8%	27,1%	29,9%
4	27,4%	19,5%	26,2%	32,4%	35,7%
5	6,9%	27,0%	10,6%	18,1%	22,5%
6	14,9%	44,7%	7,7%	17,2%	20,4%
7	5,0%	7,6%	6,2%	6,8%	8,8%
8	28,4%	22,5%	23,7%	30,5%	32,2%
9	10,1%	5,0%	15,6%	24,1%	26,9%
10	6,1%	23,9%	0,6%	17,0%	20,8%

De todo lo anterior se obtiene los indicadores de error de estimación son los siguientes.

Tabla 38: Ajuste del modelo (periodo Validación)

Categoría	R	R ²	R ² Ajustado
Administración	0,90	0,81	0,81
Básicos Decohogar	0,93	0,86	0,86
Blanco	0,89	0,80	0,80
Deportes	0,90	0,81	0,81
Electrohogar	0,93	0,86	0,86
Comunicación, Electrónica	0,89	0,80	0,80
Electricidad	0,87	0,75	0,75
Herramientas y maquinarias	0,86	0,74	0,73
Menaje	0,89	0,79	0,79
Muebles y organización	0,89	0,80	0,80

Se observa de la tabla anterior, que no existe mucha diferencia entre el periodo de validación y el periodo de calibración con respecto al indicador de ajuste del modelo, salvo en Menaje y en Muebles y organización. Por lo anterior, no se tiene un sobre ajuste del modelo en el periodo de ajuste de éste.

Tabla 39: Indicadores de error

Categoría	MAPE	MAXP	RMSE	NMSE	SSMP	SSM
ADMINISTRACION	28%	40%	47	0,135	8,4%	4
BASICOS DECOHOGAR	14%	36%	585	0,018	1,2%	57
BLANCO	9%	23%	93	0,011	4,0%	3
DEPORTES	10%	24%	56	0,012	5,7%	3
ELECTRO HOGAR	18%	46%	390	0,038	0,4%	8
COMUNICACIÓN, ELECTRONICA	20%	38%	219	0,203	9,7%	15
ELECTRICIDAD	25%	59%	4	0,190	6,8%	0
HERRAMIENTAS Y MAQUINARIAS	16%	31%	4	0,064	5,0%	0
MENAJE	24%	36%	167	0,086	3,9%	21
MUEBLES Y ORGANIZACIÓN	27%	40%	110	0,131	7,0%	9

Se observa que existe un MAPE menor al 30% en todas las categorías. Sin embargo, existe un MAPE menor a un 20% en aquellas categorías más importantes, es decir, que poseen un share mayor por semana, estas categorías son Básicos Decohogar, Electro Hogar y Blanco.

Se desprende también que el error máximo es en general alrededor de dos veces el error medio. Si se observa la tabla semanal de error existen dos semanas en que el error aumenta, estas semanas son las semanas 1 y 7. Este aumento en estas semanas, se puede deber a singularidades de la semana, como por ejemplo condiciones climáticas que

pueden producir que la gente no concurran a las tiendas. Si se deja de lado la estimación en estas semanas los indicadores de error serian los siguientes.

Tabla 40: Indicadores de Error

Categoría	MAPE	MAXP	RMSE	NMSE	SSMP	SSM
ADMINISTRACION	27%	37%	44	0,123	8,4%	4
BASICOS DECOHOGAR	11%	21%	503	0,014	1,2%	57
BLANCO	8%	14%	74	0,007	4,0%	3
DEPORTES	9%	16%	58	0,011	5,7%	3
ELECTRO HOGAR	15%	26%	316	0,025	0,4%	8
COMUNICACIÓN, ELECTRONICA	18%	33%	185	0,145	9,7%	15
ELECTRICIDAD	21%	45%	3	0,099	6,8%	0
HERRAMIENTAS Y MAQUINARIAS	14%	26%	3	0,045	5,0%	0
MENAJE	23%	32%	162	0,081	3,9%	21
MUEBLES Y ORGANIZACIÓN	26%	36%	102	0,114	7,0%	9

Se observa que el MAPE de cada categoría disminuye en todas las categorías, principalmente en aquellas donde el error máximo difería en más puntos porcentuales en relación al MAPE.

5. Resumen de resultados

En éste capítulo, se presentará un resumen de los principales resultados, a partir de éstos se analizarán los modelos para y concluir sobre las ventajas y desventajas de éstos. Recordando que la validación de los modelos se realizó para un periodo de 10 semanas y por lo tanto los indicadores que se presentan se encuentran dentro de este periodo.

En primer lugar se determinó la demanda global de canje de productos a través de series de tiempo ARIMAX y por suavizado exponencial. Los resultados para cada uno de los modelos fueron los siguientes.

Tabla 41: Errores de Estimación (ARIMAX)

RMSE	MAPE	MAXP	NMSE	SSMP	SSM	R	R ²	R ² Ajustado
1369	10,5%	15,6%	0,007	1,7%	275	0,90	0,82	0,82

Tabla 42: Errores Modelo Suavizado Exponencial

RMSE	MAPE	MAXP	NMSE	SSMP	SSM	R	R ²	R ² Ajustado
2130	15%	27,5%	0,02	5,5%	325	0,84	0,71	0,69

De las tablas anteriores se concluye que para estimar la demanda a nivel global el modelo más adecuado entre los antes mencionados es un modelo ARIMAX. A pesar de ser un modelo más sencillo y de fácil implementación, el modelo de suavizado exponencial posee un error mucho mayor que el modelo ARIMAX, es por esto que se sugiere a la empresa utilizar un modelo que posea el menor MAPE posible. Lo anterior se justifica por el hecho de que al tener un menor error en la estimación global, el error que se arrastra a los niveles y categorías de productos será por lo tanto menor.

El error promedio en el modelo escogido resulta ser de un 10,5% y el MAXP de un 15,6%. Por lo tanto, no existe mucha varianza en el error de la estimación del modelo. Otro indicador importante de observar es SSMP, el valor de éste es de 1,7% lo que indica que existiría un quiebre de stock promedio de por lo menos un 1,7%.

La siguiente etapa fue determinar la demanda de productos por categoría o tipo de productos, para lo anterior se utilizaron tres modelos distintos los cuales fueron: El modelo de participación de mercado propuesto, modelos Arima y de suavizado exponencial. En las siguientes tablas se exponen los resultados para cada uno de estos modelos.

Tabla 43: Errores Participación de mercado autorregresivo

Indicador	Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Nivel 6	Nivel 7	Nivel 8
MAPE	8%	14%	16%	17%	57%	27%	67%	33%
MAXP	23%	27%	50%	42%	204%	67%	193%	55%
RMSE	836	518	450	102	26	13	6	4
NMSE	0,008	0,025	0,030	0,077	0,403	0,154	0,385	0,349
SSMP	3%	1%	2%	14%	25%	17%	25%	22%
SSM	341	25	54	62	14	8	4	2

Tabla 44: Errores Suavizado

Indicador	Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Nivel 6	Nivel 7	Nivel 8
MAPE	22%	18%	27%	21%	33%	24%	57%	26%
MAXP	53%	43%	56%	50%	96%	49%	87%	71%
RMSE	2197	744	428	178	20	10	7	5
NMSE	0,056	0,148	0,132	0,167	0,305	0,103	0,956	0,610
SSMP	2,5%	6,1%	11,2%	7,8%	10,1%	12,5%	28,6%	17,5%
SSM	201	145	265	32	7	5	4	2

Tabla 45: Errores ARIMA

Indicador	Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5	Nivel 6	Nivel 7	Nivel 8
MAPE	20%	18%	17%	19%	30%	20%	52%	29%
MAXP	40%	29%	60%	45%	83%	41%	71%	75%
RMSE	1915	677	413	74	20	11	8	5
NMSE	0,036	0,014	0,029	0,038	0,236	0,106	0,856	0,678
SSMP	1,3%	2,1%	7,9%	10,1%	15,9%	14,2%	38,6%	27,5%
SSM	147	74	219	46	11	7	6	3

Se observa que para los niveles 1 al 4 el modelo que predice mejor es el modelo propuesto de participación de mercado auto regresivo, para los niveles siguientes (4 al 8) son los modelos de serie de tiempo que predicen mejor. El modelo ARIMA es aquel que posee un menor MAPE y menor MAXP sin embargo, la diferencia con el modelo de suavizado exponencial no es muy grande, sumado a lo anterior, el modelo ARIMA posee un peor indicador de SSM y es más complejo de implementar. El tener un SSM más bajo permitiría disminuir los quiebres de stock de los productos ya que se subestima en menor medida la demanda, como en el caso anterior o global. Por todo lo anterior, se concluye que la empresa debería utilizar el modelo de participación de mercado para los primeros 4 niveles y para los siguientes 4 se debe utilizar uno de suavizado exponencial. Con los modelos propuestos para cada nivel de puntos, se observa que el MAPE es menor al 20% en los primeros 4 niveles y que en los siguientes 4 niveles salvo el nivel 5 no supera el 35%. Este error es mayor en los niveles más altos debido a que existe un menor número de canjes totales y una mayor aleatoriedad de éstos. Si se observa el RMSE en los niveles con mayor error las unidades erradas no superan las 15, lo que hace muy difícil disminuir este error.

Finalmente, se presentan los resultados para las categorías para los primeros 4 niveles.

Tabla 46: Errores Categorías Nivel 1

Categoría	MAPE	MAXP	RMSE	NMSE	SSMP	SSM
ADMINISTRACION	28%	40%	47	0,135	8%	4
BASICOS DECOHOGAR	14%	36%	585	0,018	1%	57
BLANCO	9%	23%	93	0,011	4%	3
DEPORTES	10%	24%	56	0,012	6%	3
ELECTRO HOGAR	18%	46%	390	0,038	0%	8
COMUNICACIÓN, ELECTRONICA	20%	38%	219	0,203	10%	15
ELECTRICIDAD	25%	59%	4	0,190	7%	0
HERRAMIENTAS Y MAQUINARIAS	16%	31%	4	0,064	5%	0
MENAJE	24%	36%	167	0,086	4%	21
MUEBLES Y ORGANIZACIÓN	27%	40%	110	0,131	7%	9

Tabla 47: Errores Categorías Nivel 2

	MAPE	MAXP	RMSE	NMSE	SSMP	SSM
BASICOS DECOHOGAR	23%	38%	106	0,109	1%	3
BLANCO	28%	50%	74	0,162	0%	1
ELECTRO HOGAR	18%	48%	352	0,092	0%	0
NINOS	46%	77%	9	0,414	20%	3
ACCESORIOS AUTOMOVILES	17%	28%	24	0,101	2%	3
COMUNICACIÓN, ELECTRONICA	28%	40%	100	0,163	4%	6
ELECTRICIDAD	20%	50%	15	0,077	4%	2
LINEA BLANCA Y CLIMATIZAC	37%	65%	43	2,389	9%	8
MENAJE	20%	37%	138	0,045	6%	30
MUEBLES Y ORGANIZACIÓN	15%	29%	74	0,037	5%	23

Tabla 48: Errores Categorías Nivel 3

Categorías	MAPE	MAXP	RMSE	NMSE	SSMP	SSM
ADMINISTRACION	13%	38%	194	0,033	1%	7
BLANCO	25%	52%	99	0,155	1%	1
ELECTRO HOGAR	17%	24%	130	0,045	5%	30
ACCESORIOS AUTOMOVILES	13%	28%	9	0,110	1%	0
AGRO	16%	40%	20	0,238	7%	3
COMUNICACIÓN, ELECTRONICA	10%	19%	76	0,097	0%	0
FERRETERIA	35%	55%	30	0,455	4%	2
MENAJE	36%	49%	17	0,524	4%	2

Tabla 49: Errores Categorías Nivel 4

Categoría	MAPE	MAXP	RMSE	NMSE	SSMP	SSM
ADMINISTRACION	15%	26%	20	0,041	9,4%	8
BASICOS DECOHOGAR	28%	56%	6	0,135	22,7%	2
ELECTRO HOGAR	14%	33%	44	0,086	9,7%	11
NINOS	14%	37%	20	0,167	5,4%	3
COMUNICACIÓN, ELECTRONICA	25%	36%	8	0,155	3,4%	1
MENAJE	44%	81%	4	0,536	17,1%	1
MUEBLES Y ORGANIZACIÓN	25%	42%	2	0,168	0,3%	0
TABIQUERIA/TECHUMBRE/AISL	26%	39%	5	0,152	13,9%	1

Para estimar los canjes por categoría se utilizó el modelo de participación de demanda basado en participación de mercado autorregresivo. No se utilizaron modelos de serie de tiempo como comparación ya que estos requieren un historial mínimo de 50 observaciones y en la mayoría de las categorías esto no se cumplía ya que pasaban de estar en un catálogo a no estar en el siguiente.

Se observa de las tablas anteriores, que a medida que van aumentando los niveles de puntos, el MAPE de las categorías en general también aumenta (salvo un par de categorías), esto se debe a que existe una mayor volatilidad en los canjes a medida que aumentan los niveles y disminuyen la cantidad de canjes. Para el nivel 1 el MAPE no supera el 27%, y el SSMP no supera el 10%. En el nivel 2 el MAPE se mantiene en promedio bajo el 25% pero existen alguna categorías en donde este supera el 40% pero si se observa el RMSE éste no supera las 9 unidades (el RMSE mayor es de 352 unidades) por lo que la importancia de este error no es tan relevante. Observando el SSMP de este nivel, en general no supera el 10% salvo la categoría niños que es de 20%. Para el nivel 3 el MAPE no supera el 36% y un MAXP de un 55%, no existe mucha volatilidad con respecto a los errores de las estimaciones. El nivel 4 a pesar que el MAPE se mantiene en promedio bajo al 30% existe una mayor volatilidad con respecto al MAXP ya que es en general casi el doble que el MAPE, esto se debe a que este nivel posee menos del 2% de los canjes semanales y la aleatoriedad como se vio anteriormente es mayor en relación a los niveles anteriores.

En resumen, para estimar los canjes a nivel global, se sugiere utilizar un modelo de estimación de demanda basado en series de tiempo. Para los canjes por nivel desde el nivel 1 al 4 se sugiere utilizar un modelo basado en participación de mercado autor regresivo y para los niveles del 5 al 8 una estimación basada en suavizado exponencial. Para los canjes por categorías dada la estructura y calidad de los datos históricos, se sugiere utilizar un modelo de estimación de demanda basado en participación de mercado autorregresivo.

En conclusión el error de arrastre desde la estimación de canjes globales hasta la estimación por categorías no es desmesurada y se ha logrado encontrar modelos de estimación de demanda para los canjes por semanas para una empresa que solo posee una estimación global mensual.

6. Conclusiones

Al finalizar el trabajo y después de haber desarrollado distintos modelos para cada etapa queda la convicción de que los modelos aplicados son una gran herramienta de ayuda a la toma de decisiones, ya que pueden ayudar a mantener un stock adecuado de productos y permitir disminuir a la larga la insatisfacción de los clientes que no logran canjear los productos deseados por éstos.

Se han propuesto modelos para determinar la demanda de las categorías pertenecientes a un club de fidelización los cuales son adquiridos a través de los canjes de los puntos acumulados cuando se realizan compras con la tarjeta del club. El modelo se compone de tres etapas, la primera es estimar la demanda global de canjes para la siguiente semana a través de un modelo de series de tiempo, a continuación, se estima la demanda por cada nivel de punto a través de un modelo de participación de mercado autorregresivo para los primeros 4 niveles y un modelo de suavizado exponencial para los siguientes 4 niveles. Luego, se estima la demanda por categorías de productos para los 4 primeros niveles a través de un modelo de participación de mercado autorregresivo.

A lo largo de este trabajo se ha logrado desarrollar un modelo de estimación de demanda semanal que no existe actualmente en la empresa y por lo tanto se logra mejorar la situación actual de estimación y que puede ayudar a disminuir los costos y los quiebres de stock. Se ha demostrado que existen variables externas que pueden influir en el comportamiento de los consumidores, por ejemplo el hecho de que existan meses o semanas donde los canjes aumenten de forma drástica da a conocer que al cliente le genera más utilidad canjear en estos meses que en otros, ya sea por que poseen más puntos, reemplazar compras, o simplemente porque tienen el tiempo de ir de compras un determinado momento del mes. Otro factor que influye directamente en los canjes de los clientes, son las compras totales de éstos, que afectan directamente la cantidad de puntos acumulados. No se puede dejar de lado, es las acciones que tome la competencia, si existe un mejor control de inventario en esas tiendas en relación a la tienda de estudio, puede significar una fuga de los clientes desde una a la otra, también el poder canjear por cualquier producto de la tienda puede aumentar el atractivo para los clientes o el realizar acciones de marketing como doblar puntajes, poder acceder a los premios con los puntos y un poco de dinero, etc. Pueden tener el mismo efecto de fuga y por lo tanto una disminución de las ventas y canjes.

A pesar de lograr desarrollar un buen modelo de estimación existe un paso previo, que es mantener un control de inventario adecuado de los productos. Sin lo anterior, los modelos de estimación solo servirán a nivel global y no solucionarán los problemas de stock en las tiendas y por ende los problemas de insatisfacción de los clientes.

La modelación de la demanda en todas sus etapas mostró resultados satisfactorios ya que salvo un par de categorías y nivel. Además, mientras menores eran los productos canjeados en general los modelos entregaban errores mayores.

Observando los objetivos específicos de esta memoria, se desprenden obtienen las siguientes conclusiones:

Se desarrolló un modelo de estimación de demanda a nivel global robusto que posee un MAPE de un 10% y un MAXP de un 15% lo que permite establecer que no existe mucha aleatoriedad en el pronóstico de los canjes globales a través de una modelación con series de tiempo, incorporando variables externas como fin de mes, festividades como navidad y los puntos

entregados. Se determinó que existe en algunos meses una concentración de los canjes en la última semana de éste y que aumentan drásticamente los canjes. Además, la variable más correlacionada con los canjes de productos resultó ser los puntos entregados en la semana de estudio y en las dos precedentes. En la época de navidad aumentan drásticamente los canjes. Lamentablemente no se obtuvo la información de promociones de doble puntaje que podrán haber influido directamente en los canjes.

Se logró encontrar un modelo que permite estimar los canjes por nivel de puntos MAPE entre un 8% y un 17% en los primeros 4 niveles que representan más del 97% de los canjes semanales, por lo tanto es un modelo confiable y el cual puede ayudar a reducir los quiebres de stock y costos directos. Para los niveles superiores, el MAPE varía entre el 20% y 52%. A pesar de que el error es mayor que para los primeros niveles este error no supera las 50 unidades en promedio (suma de los 4 niveles) en comparación con las 836 unidades de error en promedio del nivel 1, por lo que se considera un error menor y difícil de disminuir. Los modelos recomendados a utilizar son un modelo de participación de mercado para los primeros 4 niveles y de suavizado exponencial para los siguientes 4 niveles.

Para estimar la demanda de las categorías de productos por nivel, se utilizó un modelo de participación de mercados. El MAPE promedio es de alrededor del 25% con mucha aleatoriedad, dependiendo del nivel y la categoría. Las variables que afectan los canjes de las categorías son la distribución de estas en el catálogo que se ve reflejada con la cantidad de SKUS de cada una de ellas y como varían de un catálogo a otro. A medida que aumenta el nivel de puntos, aumenta la aleatoriedad del error, es decir, aumenta el MAXP con respecto al MAPE incluso en el nivel 4 toma un valor máximo alrededor de dos veces el MAPE.

Muchas veces se tiende a realizar modelos de estimación de demanda complicados que si disminuyen el error de estimación pero que su implementación es compleja. Del trabajo anterior, se concluye que existe un “trade off” entre la complejidad del modelo en la implementación de éstos y el error que estos provocan. Esto quiere decir, que a veces un modelo sencillo con un poco más de error puede ser más adecuado que un modelo complicado, dado el tiempo de dedicación en implementarlo, en la capacitación de personal para mantenerlo, actualizarlo, etc. que puede a la larga aumentar los costos de operación en mayor medida de la reducción que provoca el modelo en la estimación. Para este trabajo se ha concluido que los modelos ARIMAX para estimar la demanda global son el más. Sin embargo, un modelo de participación de mercado autorregresivo es más adecuado para los niveles bajos y para los niveles altos (con bajos canjes) un modelo sencillo como el de suavizado exponencial es más adecuado, dado que arroja resultados muy similares al ARIMA y mejores en relación al modelo de participación autorregresivo.

Considerando que la estimación actual de demanda es global y por nivel de puntos y además mensual. El modelo entregado que permite estimar la demanda global, por nivel de puntos y por categorías de productos semanalmente. Se concluye que se ha contribuido enormemente en la estimación de demanda de productos. Sin lugar a dudas, hay mucho por hacer y no solo una buena estimación es suficiente si no que también un buen control de inventario.

Se ha logrado distinguir variables que afectan en las distintas etapas de estimación de demanda en los canjes. A nivel global como se ha mencionado anteriormente las variables más significantes que influyen en el comportamiento de canje de los consumidores resultaron ser las semanas fuertes de algunos meses, esto puede ser explicado por el hecho de que a fin de mes es cuando los puntos caducan y por lo tanto la gente apresura sus canjes para no perder estos puntos. Otra variable significativa a los modelos son las compras que los clientes han hecho, ya que éstas

afectan directamente a la acumulación de puntos y por lo tanto al acceso de productos más atractivos. Finalmente, existe la variable navidad que sugiere no solo que la gente canjea sus puntos en esta fecha porque es cuando más compras realiza sino también porque reemplazan regalos canjeado productos en vez de comprarlos. El número de productos que posea una categoría influye en el atractivo de algunas de éstas. Sin embargo, existen categorías que son atractivas independiente de la cantidad de productos que éstas posean, como por ejemplo, las categorías Electro hogar y Blanco que mantienen un share estable a pesar de la varianza en el número de productos que posean.

Se ha logrado determinar una cuantificación de pronóstico de los errores de los modelos y además generado pronósticos confiables dentro de un rango de error aceptable y que sin lugar a dudas aportará no solo en la futura estimación de demanda sino que en la generación de catálogos ya que se pueden generar catálogos con aquellas categorías que generen más utilidades y obviar las de menor utilidad.

Es menester explicar las principales causas que pueden originar errores en las estimaciones y por lo tanto aportar en los futuros trabajos para mejorar las estimaciones.

No incorporar la información de promociones doble puntaje a la estimación, lo que podría disminuir el error global de canje y por lo tanto el error de arrastre a los niveles y categorías sería menor.

- Cuando existen pocas unidades canjeadas en una determinada categoría o nivel, el equivocarse en una unidad produce un error mayor que el equivocarse en una unidad en aquellas categorías o nivel con canjes mayores en unidad.
- Existen ciertas categorías que no poseen un comportamiento modelable, y estas son las que poseen un mayor error porcentual medio.
- Ciertas categorías no han variado mucho en la cantidad de productos que poseen por nivel, por lo tanto la variable que tiene relación con la cantidad de productos no resultó relevante para esta categoría y si se varia de un catálogo a otro la cantidad de productos de la categoría en forma significativa, puede llegar producir grandes errores de predicción en las primeras semanas.
- Muchas veces convivieron dos catálogos simultáneamente por un par de semanas, sin embargo, esta información no se encuentra disponible por lo que en estas semanas puede haber sobreestimado la variable de cantidad de productos.

7. Estudios Futuros y propuestas de mejora

En este capítulo y final se desarrollarán una serie de propuestas de mejora a los modelos desarrollados a través del proyecto y para que estos sean puestos en práctica.

Para mejorar la estimación de los canjes a nivel global, se recomienda incorporar las variables que en esta memoria no se contó con la información. Las variables a que se hace referencia son las semanas que se han realizado promociones doble puntaje, los puntos por vencer por semana y las acciones de los competidores. Para la primera variable se sugiere tener un registro en la base de datos sobre las fechas en que se hicieron las promociones y solamente en papel como se maneja actualmente y lo que provocó que la persona encargada de entregar esta información no haya tenido tiempo de buscar en los registro estas fechas. Para la variable de

puntos a vencer, se sugiere a la empresa no pisar esta variable en la base de datos y permitir así incorporar la información histórica de los puntos a vencer en el modelo global.

Para mejorar la estimación por niveles se sugiere crear un indicador que sea el número de clientes que tienen acceso a cada nivel al comienzo de cada semana, así se puede tener una cota de la cantidad de productos que se pueden canjear. Lamentablemente la variable de puntos disponible por clientes se pisa en la base de datos por lo que crear este indicador no fue posible.

Con respecto a las categorías de productos, hay hartas formas de mejorar la predicción, una de ellas es guardar en la base de datos la fecha exacta del cambio de catálogo y también si se encuentran dos catálogos disponibles en paralelo y por cuánto tiempo estos estuvieron en vigencia. Además, para los productos exclusivos se recomienda incorporarlos con las mismas descripciones que un producto no exclusivo, ya que es una tarea ardua clasificarlos en una de las categorías existentes. Además, incorporándolas como un producto de la tienda se puede bajar en otro nivel para la estimación, por ejemplo, de electrodoméstico bajar a televisores.

Sin embargo, no solo tener una buena estimación es suficiente para suplir las demandas de los clientes, sino también, una buena planificación de inventarios. Actualmente, no se cuenta con la información de los productos que se encuentran en stock en cada tienda, y por lo tanto la información histórica de stock de productos. Actualmente, se estima⁹ que el 40% de las personas que desean un producto no pueden llevárselo porque no se encuentra disponible. Considerando que son alrededor de 500.000 clientes diferentes los que canjean al menos un producto por año, considerando que para alcanzar el primer nivel de puntos se debe gastar en compras \$600.000 anuales. Si de estos 500.000 clientes el 40% no consiguen canjear el primer producto deseado se tendrían 200.000 clientes no logran canjear el primer producto deseado considerando una fuga de un 1%¹⁰ por insatisfacción con el servicio y que por lo tanto, estos clientes cambiarían su consumo en la tarjeta de la competencia, 2.000 clientes se fugarían por año lo que equivaldría si solo son clientes que canjean en el nivel 1 (nivel más bajo) una pérdida mínima anual de MM\$1.200. Por lo anterior, es fundamental tener un buen control de inventario, para impedir la insatisfacción de los clientes y por ende su fuga.

Se sugiere además, cambiar el sistema de catálogo por productos existentes en la tienda, lo anterior permitiría reducir costos de inventario y operacionales al poder utilizar el “know how” de la gestión de la tienda en cuanto a inventario y la relación con los proveedores.

Finalmente, se debería incorporar el concepto de atracción de las categorías para la selección de categorías del producto en el diseño de los catálogos. Es decir, en aquellas categorías que poseen mayor atracción incorporarlas al catálogo y descartar aquellas que históricamente no aportan valor a los clientes.

⁹ Valor entregado por el gerente de Fidelización de la empresa

¹⁰ Según la gerencia de marketing, la fuga actual de clientes es de un 10% anual

Carta Gantt

Mes	Ago-07				Sep-07				Oct-07				Nov-07				Dic-07				Ene-08				Feb-08				Mar-08				Abr-08				May-08				Jun-08				Jul-08				Ago-08							
Semana	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4				
Estudio Bibliográfico	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Definición del Problema								■																																																
Objetivos de la memoria									■	■	■	■	■	■	■	■																																								
Generación de la Metodología													■	■	■	■													■	■	■	■																								
Generación de Hipotesis																	■	■	■	■	■	■	■	■	■	■	■	■																												
Limpieza de datos																																																								
Validación de Hipotesis																																																								
Desarrollo de modelo Global																																																								
Desarrollo de modelo Por Nivel																																																								
Desarrollo de modelo Por Categorías																																																								
Conclusiones																																																								
Fecha Objetiva	■																																																							
Objetivo logrado	■																																																							

Bibliografía

- [1] Luís G. Renart Cava, Profesor del IESE. CRM: Tres estrategias del éxito. s.l. : Gemma Tonijuan, 2004.
- [2] Taylor, Gail Ayala and Scott A. Neslin. “The Current and Future Sales Impact of a Retail Frequency Reward Program,” *Journal of Retailing*, 81 (4), 293–305. 2005.
- [3] Yuping, Liu. “The long-term Impact of Loyalty Programs on consumer Purchase Behavior and Loyalty”, *Journal of Marketing*, Vol. 71. 2007.
- [4] Lewis, Michael. “The Influence of Loyalty Programs and Short-Term Promotions on Customer Retention,”. 2005.
- [5] Kim, Byung-Do, Mengze Shi, and Kannan Srinivasan. “Reward Programs and Tacit Collusion,” *Marketing Science*, 20 (Spring), 99–120. 2001.
- [6] Box, Jenkins. *Time series Analysis: Forecasting and Control*". 1976.
- [7] Garson, G. David. “Statnotes: Topics in Multivariate Analysis”. s.a.
- [8] Unanue, Oscar. Estudio y Pronóstico de demanda a nivel de sku para la cadena regional de supermercados BRYC. s.l. : Memoria, 2006.
- [9] L. G. Cooper, M. Nakanishi. “Market Share Analysis”. *International series in quantitative marketing*. 1988.
- [10] Memorias alumno Christopher Gross, Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas. Febrero de 2006.
- [11] Besanko, David, Achin Gupta y Dipak Jain. “Logit Demand Estimation Under Competitive Pricing Behavior: An Equilibrium Framework” *Management Science*. 1998.
- [12] Salvatore, Dominick. *Estadística y Econometría*. 2004.
- [13] Reglamento del programa de puntos de la administradora de servicios computacionales de la tienda . 2008.
- [14] Godoy, José Miguel. Memoria "Medición del impacto de Alianzas estratégicas en Programas de Lealtad". Santiago : s.n., 2008 (En desarrollo).

Anexos

Anexo A: Canjes semanales por meses

Gráfico 28: Canjes Semanales (Enero)

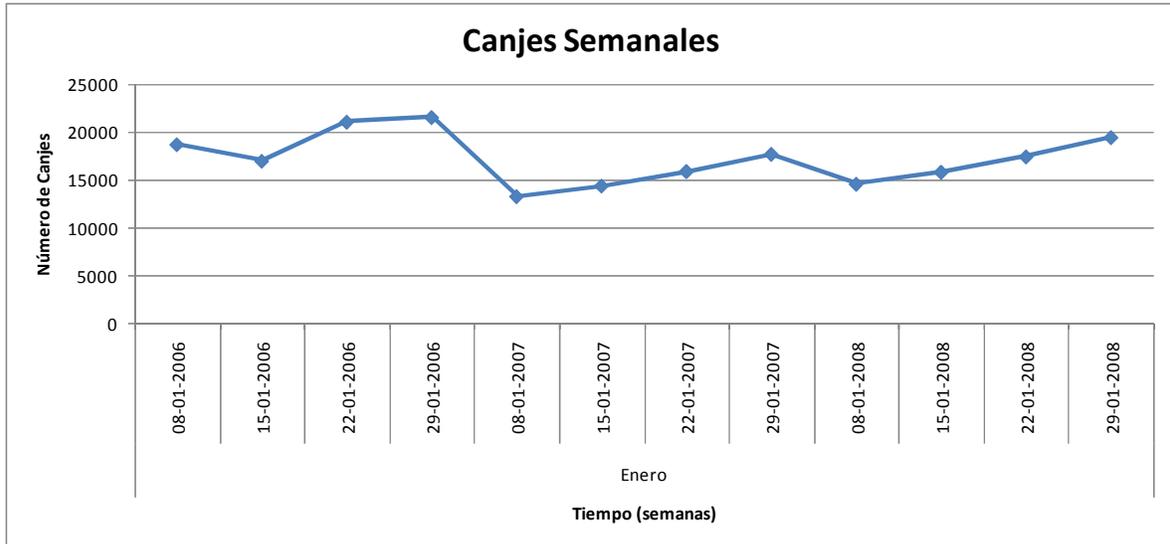


Gráfico 29: Canjes Semanales (Abril)

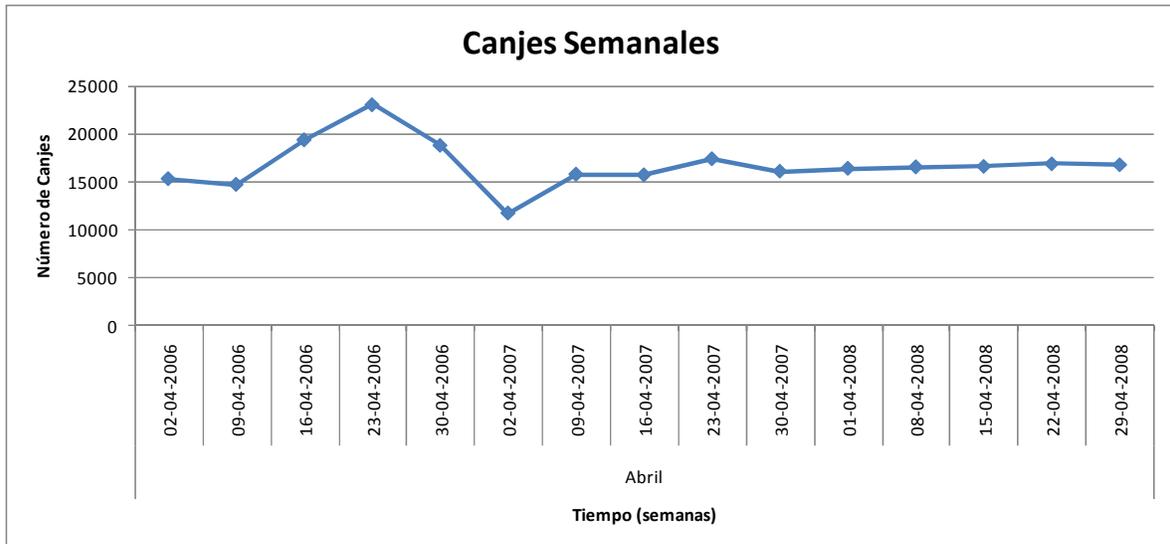


Gráfico 30: Canjes Semanales (Mayo)

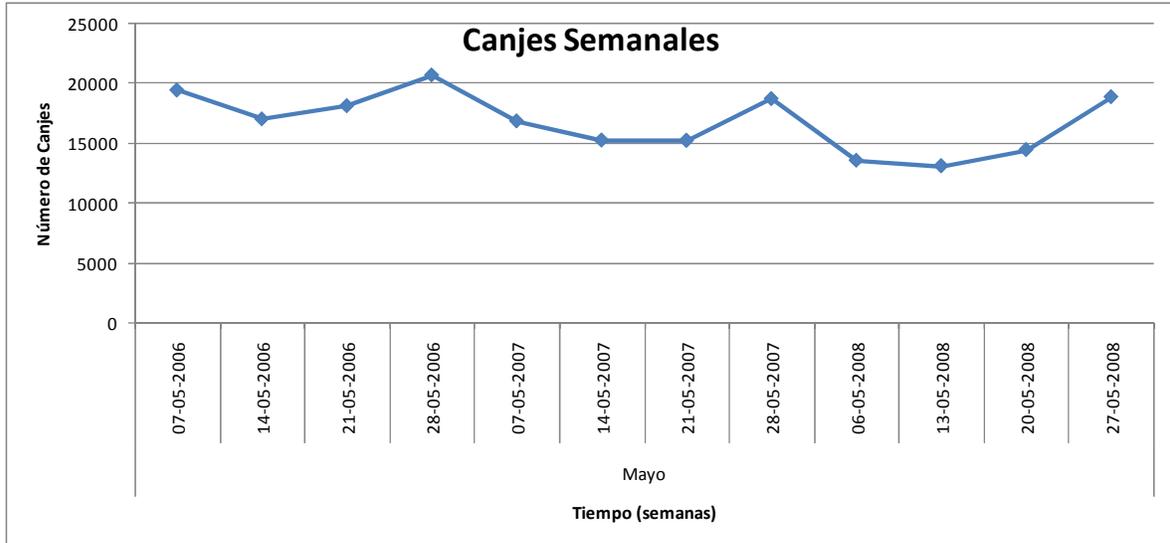


Gráfico 31: Canjes Semanales (Junio)

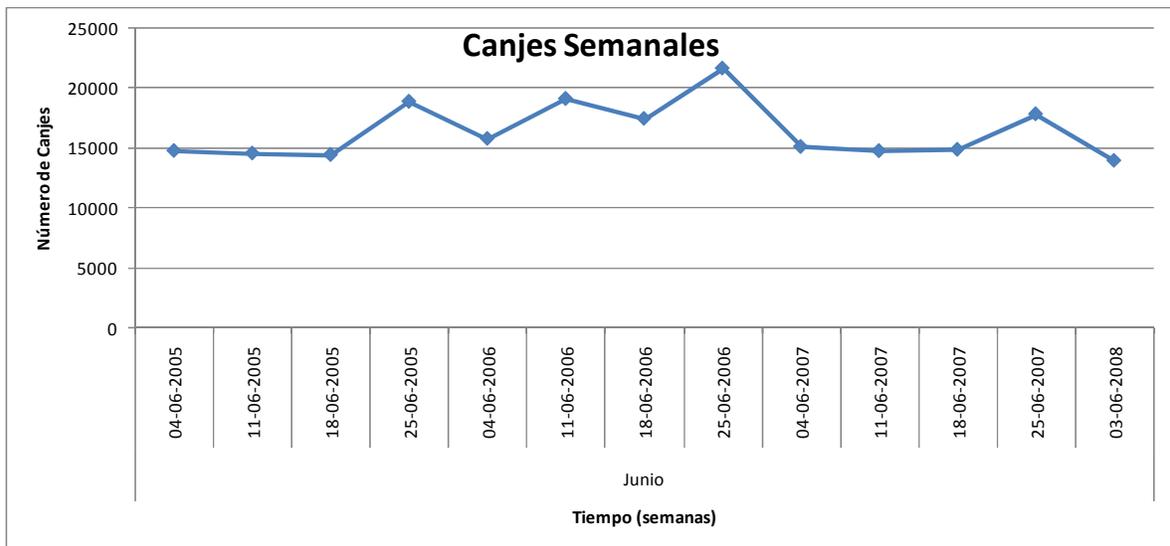


Gráfico 32: Canjes Semanales (Julio)

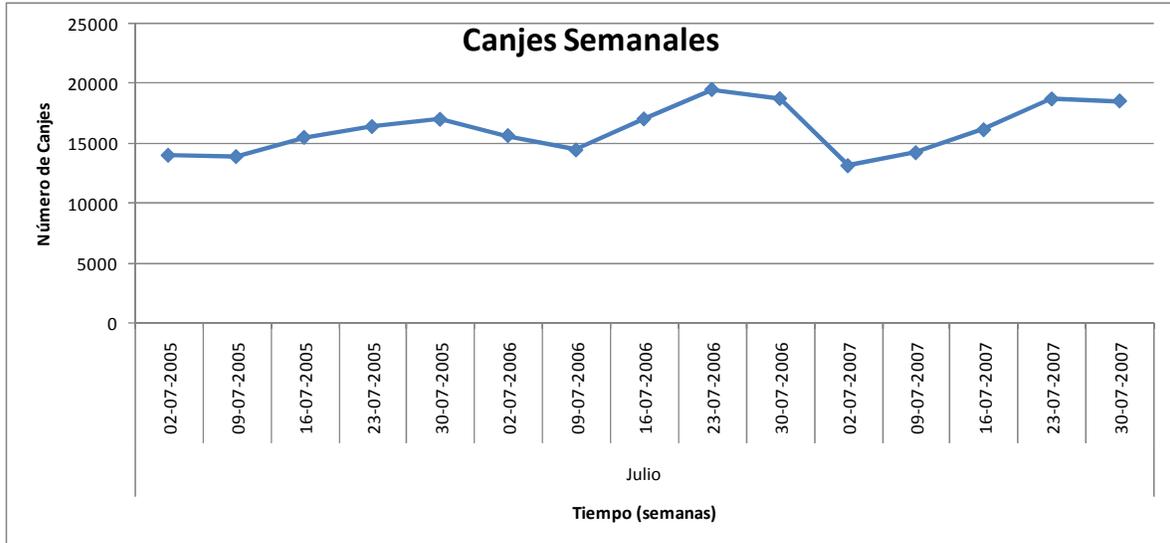


Gráfico 33: Canjes Semanales (Agosto)

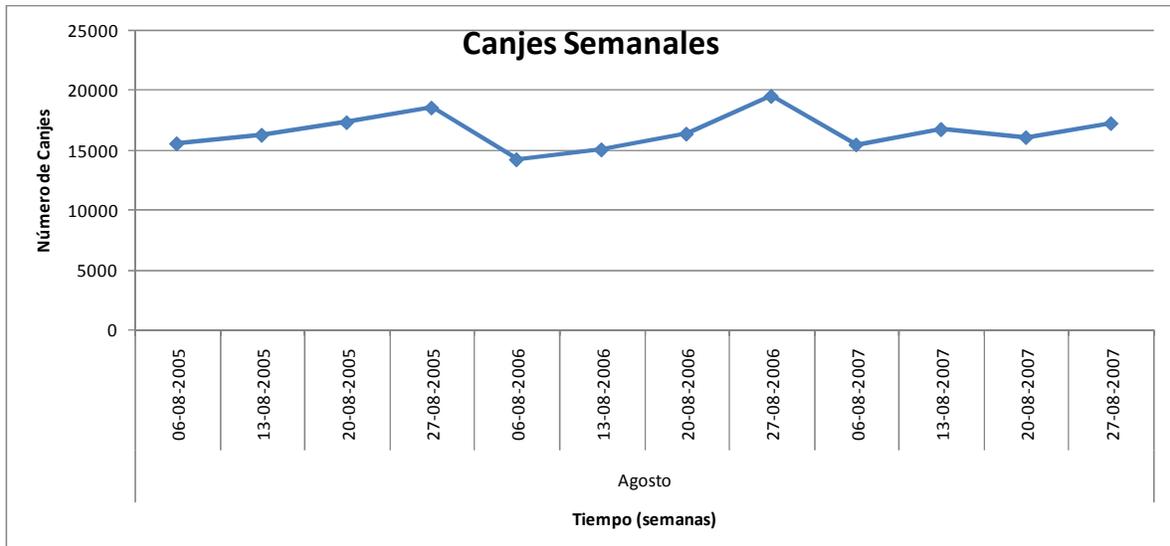
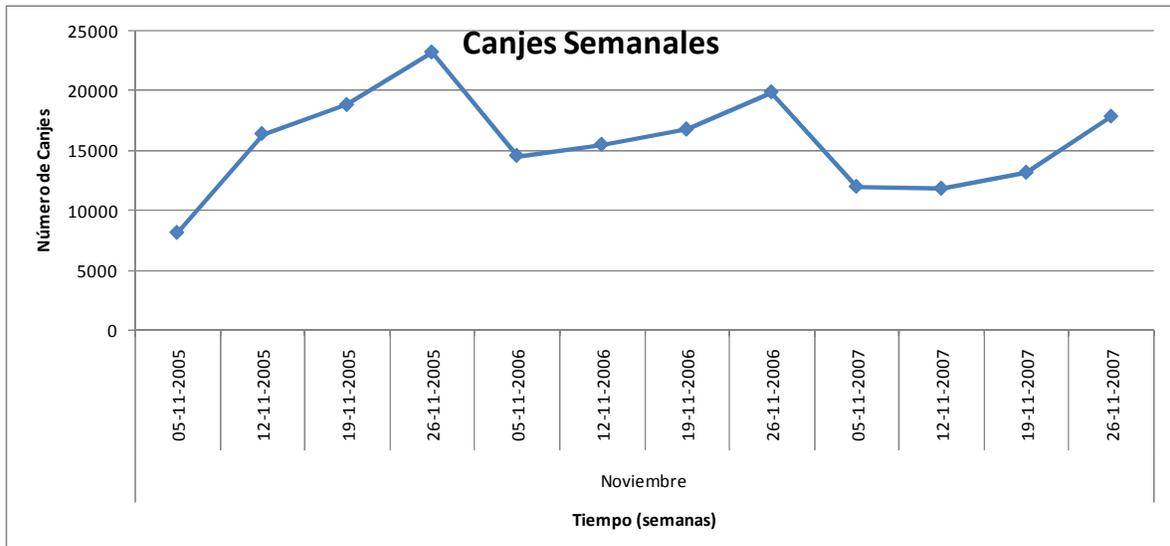


Gráfico 34: Canjes Semanales (Noviembre)



Anexo B: Ajuste del modelo por nivel

Tabla 50: Significancia de las variables del modelo (nivel 2)

Variables Explicativas	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t
LAGS(U,1)	0,83	0,06	0,83	13,73
LAGS(U,3)	0,17	0,06	0,17	2,78

Tabla 51: Significancia de las variables del modelo (nivel 3)

Variables Explicativas	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LAGS(U,1)	0,873	0,06	0,873	14,48	0,00
LAGS(U,3)	0,125	0,06	0,126	2,08	0,04

Tabla 52: Significancia de las variables del modelo (nivel 4)

Variables Explicativas	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LAGS(U,1)	1,00	0,04	0,99	280,62	0,00

Tabla 53: Significancia de las variables del modelo (nivel 5)

Variables Explicativas	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LAGS(U,1)	1,00	0,08	0,99	138,82	0,00

Tabla 54: Significancia de las variables del modelo (nivel 6)

Variables Explicativas	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LAGS(U,1)	0,99	0,006	0,97	170,95	0,00

Tabla 55: Significancia de las variables del modelo (nivel 7)

Variables Explicativas	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LAGS(U,1)	0,553	0,076	0,551	7,33	0,00
LAGS(U,2)	-0,443	0,076	-0,441	5,87	0,00

Tabla 56: Significancia de las variables del modelo (nivel 8)

Variables Explicativas	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LAGS(U,1)	0,485	0,08	0,482	6,435	0,00
LAGS(U,2)	0,511	0,08	0,509	6,791	0,00

Anexo C: Categoría de productos por nivel

Nivel 1

Tabla 57: Significancia de las variables del modelo (Deportes)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	T	Sig.
LAGS(U,1)	1,06	0,08	1,12	12,60	0,00
LAGS(U,2)	-0,15	0,08	-0,17	-1,95	0,05

Tabla 58: Significancia de las variables del modelo (Electrohogar)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	T	Sig.
LAGS(U,1)	0,88	0,03	0,88	26,12	0,00
SN	0,70	0,34	0,11	2,05	0,04

Tabla 59: Significancia de las variables del modelo (Comunicación, Electrónica)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	T	Sig.
LAGS(U,1)	0,80	0,08	0,80	9,63	0,00
LAGS(U,2)	0,18	0,08	0,18	2,21	0,03

Tabla 60: Significancia de las variables del modelo (Electricidad)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	T	Sig.
(Constante)	-0,70	0,36		1,93	0,04
LAGS(U,1)	0,37	0,07	0,37	4,90	0,00
LAGS(U,2)	0,41	0,07	0,41	5,44	0,00
LnSN	0,15	0,05	0,21	2,93	0,00
DIFF(LnSN,1)	0,63	0,12	0,14	5,21	0,00

Tabla 61: Significancia de las variables del modelo (Electricidad)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	T	Sig.
(Constante)	-0,29	0,12		-2,34	0,02
LAGS(U,1)	0,93	0,03	0,93	30,69	0,00
DIFF(LnSN,1)	0,51	0,38	0,04	1,36	0,02

Tabla 62: Significancia de las variables del modelo (Menaje)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LAGS(U,1)	0,93	0,03	0,93	36,73	0,00
LnSN	0,06	0,02	0,07	2,59	0,01

Tabla 63: Significancia de las variables del modelo (Muebles y organización)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LAGS(U,1)	0,97	0,01	0,97	56,73	0,03

Nivel 2

Tabla 64: Categorías Nivel 2

Básicos Decohogar	Blanco	Electro hogar
Accesorios Automóviles	Comunicación Electrónica	Electricidad
Menaje	Muebles y Organización	

Tabla 65: Error de Ajuste de los modelos Nivel 2 (periodo de Validación)

Categoría	R	R ²	R ² Ajustado	Error Tip. De la Estimación
Básicos Decohogar	0,94	0,89	0,89	1,35
Blanco	0,94	0,89	0,89	2,13
Electrohogar	0,87	0,75	0,75	0,14
Accesorios Automóviles	0,95	0,90	0,90	0,12
Comunicación, Electrónica	0,94	0,89	0,89	2,11
Electricidad	0,93	0,87	0,87	2,03
Menaje	0,90	0,81	0,81	0,04
Muebles y organización	0,94	0,89	0,89	0,13

Tabla 66: Error de Ajuste de los modelos Nivel 2 (periodo de Validación)

Categoría	R	R ²	R ² Ajustado	Error Tip. De la Estimación
Básicos Decohogar	0,92	0,85	0,85	1,39
Blanco	0,92	0,84	0,84	2,19
Electrohogar	0,85	0,73	0,73	0,19
Accesorios	0,93	0,86	0,86	0,23
Automóviles	0,91	0,83	0,83	2,17
Comunicación, Electrónica	0,92	0,84	0,84	2,09
Electricidad	0,88	0,77	0,77	0,10
Menaje	0,95	0,90	0,90	0,17
Muebles y organización				

Tabla 67: Significancia de las variables del modelo (Básicos Decohogar)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LAGS(U,1)	1,00	0,01	0,99	104,01	0,00

Tabla 68: Significancia de las variables del modelo (Blanco)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LAGS(U,1)	0,62	0,07	0,62	9,25	0,00
LAGS(U,3)	0,33	0,07	0,33	4,93	0,00

Tabla 69: Significancia de las variables del modelo (Electrohogar)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LAGS(U,1)	0,88	0,06	0,88	14,77	0,00
LAGS(U,3)	0,12	0,06	0,12	1,99	0,05

Tabla 70: Significancia de las variables del modelo (Accesorios Automóviles)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LnSN	0,06	0,02	0,05	2,52	0,01
LAGS(U,1)	0,95	0,02	0,95	51,93	0,00

Tabla 71: Significancia de las variables del modelo (Comunicación, Electrónica)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LAGS(U,1)	0,98	0,01	0,99	99,56	0,00
LnSN	0,16	0,10	0,02	1,68	0,01

Tabla 72: Significancia de las variables del modelo (Electricidad)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LnSN	0,09	0,04	0,06	2,17	0,03
LAGS(U,1)	0,93	0,03	0,93	32,43	0,00

Tabla 73: Significancia de las variables del modelo (Menaje)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LAGS(U,1)	0,66	0,06	0,66	10,90	0,00
LAGS(U,2)	0,17	0,04	0,27	4,51	0,00

Tabla 74: Significancia de las variables del modelo (Muebles y Organización)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LnSN	0,13	0,04	0,12	3,12	0,00
LAGS(U,1)	0,88	0,04	0,88	23,31	0,00

Tabla 75: Indicadores de error

	MAXP	MAPE	RMSE	NMSE	SSMP	SSM
BASICOS DECOHOGAR	38%	23%	106	0,109	1%	3
BLANCO	50%	28%	74	0,162	0%	1
ELECTRO HOGAR	48%	18%	352	0,092	0%	0
NINOS	77%	46%	9	0,414	20%	3
ACCESORIOS AUTOMOVILES	28%	17%	24	0,101	2%	3
COMUNICACIÓN, ELECTRONICA	40%	28%	100	0,163	4%	6
ELECTRICIDAD	50%	20%	15	0,077	4%	2
LINEA BLANCA Y CLIMATIZAC	65%	37%	43	2,389	9%	8
MENAJE	37%	20%	138	0,045	6%	30
MUEBLES Y ORGANIZACIÓN	29%	15%	74	0,037	5%	23

Nivel 3

Tabla 76: Categorías Nivel 3

Administración	Blanco	Electrohogar
Accesorios Automóviles	Agro	Comunicación, Electrónica
Ferretería	Menaje	

Tabla 77: Error de Ajuste de los modelos Nivel 3

Categoría	R	R ²	R ² Ajustado	Error Tip. De la Estimación
Administración	0,94	0,88	0,88	0,81
Blanco	0,94	0,89	0,89	0,25
Electrohogar	0,94	0,89	0,89	0,10
Accesorios Automóviles	0,95	0,90	0,90	0,32
Agro	0,94	0,89	0,89	0,11
Comunicación, Electrónica	0,94	0,89	0,89	1,23
Ferretería	0,92	0,85	0,85	0,90
Menaje	0,92	0,85	0,85	1,27

Tabla 78: Error de Ajuste de los modelos Nivel 3 (periodo de validación)

Categoría	R	R ²	R ² Ajustado	Error Tip. De la Estimación
Administración	0,92	0,84	0,84	0,86
Blanco	0,93	0,86	0,86	0,27
Electrohogar	0,91	0,83	0,83	0,40
Accesorios Automóviles	0,92	0,85	0,85	0,73
Agro	0,91	0,82	0,82	0,31
Comunicación, Electrónica	0,90	0,81	0,81	1,31
Ferretería	0,89	0,79	0,79	0,95
Menaje	0,87	0,76	0,76	1,34

Tabla 79: Significancia de las variables del modelo (Administración)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	T	Sig.
LAGS(U,1)	0,468	0,044	0,480	10,651	0,000
LnSN	0,394	0,034	0,519	11,503	0,000

Tabla 80: Significancia de las variables del modelo (Blanco)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LnSN	0,103	0,036	0,090	2,887	0,004
LAGS(U,1)	0,910	0,031	0,908	29,293	0,000

Tabla 81: Significancia de las variables del modelo (Electrohogar)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LnSN	0,113	0,034	0,126	3,314	0,001
LAGS(U,1)	0,876	0,038	0,871	22,941	0,000

Tabla 82: Significancia de las variables del modelo (Accesorios Automóviles)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LAGS(U,1)	0,996	0,017	0,997	12,321	0,000

Tabla 83: Significancia de las variables del modelo (Agro)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LAGS(U,1)	0,999	0,001	0,999	32,034	0,000

Tabla 84: Significancia de las variables del modelo (Comunicación, Electrónica)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LAGS(U,1)	0,933	0,019	0,942	48,889	0,000
LnSNC	0,053	0,017	0,059	3,047	0,003

Tabla 85: Significancia de las variables del modelo (Ferretería)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
(Constante)	-0,375	0,127		-2,955	0,004
LnSN	0,128	0,042	0,157	3,071	0,003
LAGS(U,1)	1,110	0,077	1,111	14,373	0,000
LAGS(U,2)	-0,314	0,071	-0,314	-4,448	0,000
DIFF(U,3)	0,175	0,077	0,050	2,281	0,024
DIFF(U,1)	0,214	0,043	0,108	4,952	0,000

Tabla 86: Significancia de las variables del modelo (Menaje)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
(Constante)	-1,272	0,700		-1,817	0,004
SN	0,153	0,084	0,075	1,818	0,007
LAGS(U,2)	0,912	0,042	0,906	21,722	0,000
DIFF(U,1)	0,159	0,126	0,023	1,263	0,009

Tabla 87: Indicadores de error

Categorías	MAPE	MAXP	RMSE	NMSE	SSMP	SSM
ADMINISTRACION	13%	38%	194	0,033	0,7%	7
BLANCO	25%	52%	99	0,155	1,0%	1
ELECTRO HOGAR	17%	24%	130	0,045	4,6%	30
ACCESORIOS AUTOMOVILES	13%	28%	9	0,110	1,4%	0
AGRO	16%	40%	20	0,238	6,5%	3
COMUNICACIÓN, ELECTRONICA	10%	19%	76	0,097	0,0%	0
FERRETERIA	35%	55%	30	0,455	4,0%	2
MENAJE	36%	49%	17	0,524	4,0%	2

Nivel 4

Tabla 88: Categorías Nivel 4

Administración	Básicos Decohogar	Electro Hogar
Niños	Comunicación, Electrónica	Menaje
Muebles y Organización	Tabiquería/Techumbre	

Tabla 89: Error de Ajuste de los modelos Nivel 4

Categoría	R	R ²	R ² Ajustado	Error Tip. De la Estimación
Administración	0,95	0,90	0,90	1,33
Básicos Decohogar	0,80	0,64	0,63	3,52
Electro hogar	0,95	0,90	0,80	0,12
Niños	0,95	0,90	0,90	1,43
Comunicación, Electrónica	0,94	0,89	0,89	0,11
Menaje	0,90	0,81	0,81	4,21
Muebles y Organización	0,90	0,81	0,81	3,98
Tabiquería/Techumbre	0,95	0,90	0,90	1,12

Tabla 90: Error de Ajuste de los modelos Nivel 4 (periodo de validación)

Categoría	R	R ²	R ² Ajustado	Error Tip. De la Estimación
Administración	0,92	0,85	0,85	1,42
Básicos Decohogar	0,77	0,60	0,60	3,75
Electro hogar	0,91	0,82	0,82	0,19
Niños	0,91	0,83	0,83	1,52
Comunicación, Electrónica	0,90	0,81	0,81	0,11
Menaje	0,90	0,81	0,81	4,49
Muebles y Organización	0,90	0,81	0,81	4,25
Tabiquería/Techumbre	0,89	0,80	0,80	1,19

Tabla 91: Significancia de las variables del modelo (Administración)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	T	Sig.
LAGS(U,1)	0,996	0,044	0,480	10,65	0,00

Tabla 92: Significancia de las variables del modelo (Básicos Decohogar)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
(Constante)		5,64	1,79	3,15	0,00
LnSN		2,83	0,76	0,25	3,72
LAGS(U,1)		0,68	0,07	0,68	9,41
DIFF(U,1)		-0,44	0,07	-0,30	6,15

Tabla 93: Significancia de las variables del modelo (Electrohogar)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
(Constante)	-0,07	0,03		-2,24	0,03
LnSN	0,11	0,03	0,19	3,40	0,00
LAGS(U,2)	0,66	0,06	0,66	10,93	0,00
DIFF(U,1)	0,85	0,08	0,55	10,40	0,00

Tabla 94: Significancia de las variables del modelo (Niños)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
(Constante)	-0,92	0,26		-3,54	0,00
LAGS(U,1)	0,81	0,05	0,80	16,48	0,00
SN	2,07	1,27	0,08	1,62	0,11

Tabla 95: Significancia de las variables del modelo (Comunicación Electrónica)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
(Constante)	-0,95	0,44		-2,15	0,03
LAGS(U,1)	0,93	0,03	0,93	29,91	0,00
DIFF(U,1)	0,54	0,11	0,15	4,79	0,00

Tabla 96: Significancia de las variables del modelo (Mensaje)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LnSNMENAJE	0,23	0,09	0,11	2,64	0,01
LAGS(LnMENAJE,1)	0,86	0,04	0,86	20,83	0,00

Tabla 97: Significancia de las variables del modelo (Muebles y Organización)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LnSN	0,29	0,07	0,19	4,04	0,00
LAGS(U,1)	0,79	0,05	0,79	17,27	0,00

Tabla 98: Significancia de las variables del modelo (Tabiquería/Techumbre)

Variables	Coefficientes no estandarizados	Error típ.	Coefficientes estandarizados	t	Sig.
LAGS(U,1)	0,95	0,02	0,96	54,65	0,00
LnSN	0,05	0,02	0,04	2,47	0,01

Tabla 99: Indicadores de Error

Categoría	MAXP	MAPE	RMSE	NMSE	SSMP	SSM
ADMINISTRACION	26%	15%	20	0,041	9,4%	8
BASICOS DECOHOGAR	56%	28%	6	0,135	22,7%	2
ELECTRO HOGAR	33%	14%	44	0,086	9,7%	11
NINOS	37%	14%	20	0,167	5,4%	3
COMUNICACIÓN, ELECTRONICA	36%	25%	8	0,155	3,4%	1
MENAJE	81%	44%	4	0,536	17,1%	1
MUEBLES Y ORGANIZACIÓN	42%	25%	2	0,168	0,3%	0
TABIQUERIA/TECHUMBRE/AISL	39%	26%	5	0,152	13,9%	1