

UNIVERSIDAD ADOLFO IBÁÑEZ

TESIS DE MAGÍSTER

Caracterización de la Lealtad de Clientes hacia Marcas Propias y su Efecto sobre la Fidelidad al Retail

Autor:

Antonia Pozo Cataneo

Profesor guía:

Verónica Díaz Gómez

*Tesis realizada acorde a los requerimientos para el grado de
Magíster en Ciencias de la Ingeniería y para el título de Ingeniero Civil
Industrial*

de la

Facultad de Ingeniería y Ciencias

Agosto, 2025



Índice

1. Introducción	3
2. Estado del arte	5
3. Hipótesis y Objetivos	9
3.1. Objetivos Específicos	9
3.2. Hipótesis	9
3.3. Contexto del Problema	9
3.4. Relevancia del Estudio	10
3.5. Preguntas de Investigación	10
3.6. Alcances y Limitaciones	10
4. Metodología	11
4.1. Análisis Descriptivo de los Datos	11
4.2. Ingeniería de los atributos: ¿Cómo medir la lealtad?	12
4.2.1. Variables de Lealtad a Marca Propia	12
4.2.2. Variables Lealtad al Supermercado	15
4.3. Aplicación de Algoritmos	16
4.3.1. Análisis Factorial Confirmatorio (CFA)	16
4.3.2. Regresión Lineal Múltiple	17
4.3.3. Segmentación con K-means	17
4.4. Validación y Refinamiento	17
4.5. Análisis de Resultados y Evaluación de Impacto	18
5. Resultados	20
5.1. Modelo general sin segmentación: CFA y Regresión Lineal	20
5.1.1. Explicación del modelo	20
5.1.2. Modelo de Regresión Lineal	21
5.1.3. Scores latentes obtenidos	22
5.1.4. Resultados Regresión Lineal	23
5.2. Limitaciones del modelo general y justificación del uso de clústeres	24
5.3. Segmentación con categorías con presencia de marca propia pertenecientes a la Canasta Básica	25
5.3.1. Evaluación de Resultados Categorías Canasta Básica Año 3	27
5.4. Regresión Lineal con segmentación por clústeres y su interacción con lealtad a marca propia	28
5.4.1. Resultados Regresión Lineal con clústeres	29
5.4.2. Verificación de distintas regresiones dependiendo de los clústeres	30
5.5. Análisis específico por categoría	31
5.5.1. Análisis de Categoría	32
5.5.2. Análisis de Marca Propia en las Categorías	33
5.5.3. Resumen	37
5.6. Segmentación Categoría Harina	38
5.6.1. Criterios de segmentación y evaluación del modelo	40
5.7. Segmentaciones Complementarias para Validación	45
5.8. Modelo CFA y Regresión Lineal en Categoría Harina	47

5.9. Comparación distintos modelos	50
5.10. Recomendaciones para Clústeres	51
5.10.1. Segmentación Canasta Básica	51
5.10.2. Segmentación Categoría Harina	52
6. Conclusión	54
7. Anexo	57
7.1. Métricas por Categoría	57
7.2. Resultados Segmentación Validación 1	58
7.3. Resultados Segmentación Validación 2	59

1. Introducción

En un entorno empresarial altamente competitivo y dinámico, los principales objetivos no se limitan únicamente al incremento de los ingresos, sino también a la mejora continua en la experiencia del cliente, el fortalecimiento de su lealtad y la contribución activa a una gestión más sostenible de los recursos. En respuesta a estos desafíos, las empresas buscan de manera constante mejorar la efectividad de sus estrategias de marketing para alcanzar estos objetivos de manera más eficiente y efectiva.

Un estudio realizado por Germann et al. (2013) arrojó conclusiones reveladoras en este sentido. Según sus hallazgos, se demostró que al aumentar en una unidad (según una escala de 1 a 7) la implementación del análisis de marketing, una empresa mediana puede incrementar en un 8 % su rentabilidad sobre activos (ROA). Este aumento de ROA se eleva al 21 % cuando la empresa opera en una industria altamente competitiva y con clientes que cambian frecuentemente sus deseos y necesidades. En este contexto, una dimensión crítica del análisis de marketing es comprender las dinámicas de fidelización, especialmente cuando involucran relaciones entre distintas formas de lealtad, como la que se manifiesta hacia las marcas propias y hacia el supermercado como institución. Estas dimensiones, aunque interrelacionadas, pueden obedecer a lógicas distintas.

Para que el área de marketing pueda cumplir con los objetivos empresariales, es esencial que actúe en beneficio de sus clientes, lo que implica mantener y fortalecer relaciones con ellos a largo plazo. Esto se logra a través de estrategias de fidelización, mediante la creación y entrega continua de valor percibido. Con este fin, las empresas invierten en estudios de mercado con el propósito de identificar, conocer y comprender en profundidad a sus clientes. En base a esta comprensión, llevan a cabo un proceso de segmentación, el cual consiste en crear grupos homogéneos de individuos según sus necesidades, características y comportamientos similares (Rungruang et al., 2024). La segmentación emerge como un proceso en el ámbito de marketing, ya que permite tomar decisiones enfocadas para cada uno de estos grupos. Al realizar una segmentación precisa que refleje con fidelidad la diversidad y dinámica del comportamiento de los clientes, las empresas pueden personalizar sus estrategias orientadas a satisfacer las necesidades específicas de cada segmento, optimizando así las interacciones de marketing y los recursos de la empresa, aspecto fundamental en un contexto donde los recursos son limitados.

La complejidad inherente a la segmentación de clientes radica en la diversidad de cada individuo, que puede incluir factores demográficos, socioeconómicos, psicográficos, entre otros. Esta diversidad plantea desafíos significativos al determinar los criterios más apropiados para segmentar efectivamente el mercado. Por tanto, resulta crucial realizar un análisis exhaustivo de los datos transaccionales y de comportamiento disponibles para identificar los atributos que más impactan en la segmentación y discernir la importancia relativa de las variables. Este enfoque permite desarrollar un modelo de segmentación considerablemente más preciso y detallado. Al personalizar las estrategias de marketing basadas en las características distintivas de cada segmento, las empresas logran mejorar la eficiencia y efectividad de sus interacciones con el mercado, un aspecto crucial en contextos competitivos donde la optimización de recursos y la satisfacción del cliente son determinantes para el éxito empresarial.

Entre los diversos enfoques de segmentación, uno particularmente relevante en mercados altamente competitivos es el basado en la lealtad del cliente. En sectores donde predominan márgenes de ganancia estrechos y una creciente participación de marcas de distribuidor, fortalecer el compromiso del consumidor se ha vuelto una prioridad estratégica para los responsables de marketing (Møller Jensen, 2011). En este contexto, segmentar según el nivel de lealtad permite identificar a los clientes más valiosos y diseñar estrategias diferenciadas de retención, comunicación y asignación de recursos. Esta aproximación resulta especialmente pertinente en categorías donde las decisiones de compra se ven influenciadas por la percepción de valor y el vínculo sostenido con la marca.

Más allá de identificar segmentos de clientes, esta investigación se enfoca principalmente en comprender la dinámica entre dos formas de fidelidad clave: la lealtad hacia las marcas propias del supermercado y la lealtad hacia el supermercado en su conjunto. Ambas dimensiones, aunque vinculadas, pueden responder a motivaciones distintas, y su relación no siempre es directa ni simétrica. Este estudio plantea que, bajo ciertas condiciones, la preferencia por marcas propias podría reforzar el vínculo con la tienda, mientras que en otros casos podría representar una estrategia de ahorro independiente de la fidelidad al lugar de compra.

En ese contexto, las marcas propias desempeñan un papel clave debido a su capacidad para mejorar la imagen y rentabilidad de las empresas, además de ofrecer un potencial único para diferenciarse y fomentar la lealtad de los clientes (Baltas, 2003). Estas características hacen que las marcas propias sean un elemento estratégico en los esfuerzos de segmentación. Un análisis de segmentación enfocado exclusivamente en la marca propia resulta particularmente valioso, ya que permite a las empresas personalizar sus interacciones con los clientes según sus preferencias y comportamientos en constante cambio (Bucklin et al., 1995). Este enfoque no solo optimiza las estrategias de fidelización, sino que también maximiza el uso de los recursos disponibles, potenciando el valor de la marca propia dentro de su categoría y fortaleciendo la posición competitiva de la empresa.

La relevancia de las marcas propias en mercados como el chileno se hace aún más evidente en un contexto económico desafiante, donde los consumidores buscan alternativas más accesibles sin sacrificar calidad. En particular, las marcas propias de los supermercados han experimentado un notable incremento en sus ventas. Por ejemplo, uno de los consorcios con mayor presencia en el país reportó un aumento del 14% en las ventas de sus marcas propias durante el tercer trimestre de 2024 en comparación con el mismo período del año anterior (La Tercera, 2024). Este crecimiento refleja la creciente preferencia de los consumidores por productos de marcas propias, percibidos como alternativas de alta calidad a precios más accesibles, y subraya la importancia de desarrollar estrategias de marketing basadas en la segmentación de estos productos.

En este marco, la presente investigación se propone analizar cómo se relaciona la lealtad conductual hacia las marcas propias con la fidelidad general al supermercado, integrando múltiples dimensiones del comportamiento de compra. Para ello, se construyen indicadores compuestos de lealtad basados en datos transaccionales y se identifican perfiles de clientes que permiten comprender con mayor precisión esta dinámica. Al hacerlo, se espera aportar evidencia que oriente estrategias de fidelización más efectivas, basadas en una comprensión integral del rol que cumple la marca propia en la consolidación de la lealtad al supermercado. La segmentación de clientes, en este contexto, se emplea como una herramienta analítica para identificar perfiles diferenciados que ayuden a revelar estas dinámicas. El documento se estructura de la siguiente manera: primero se presenta el marco teórico que sustenta los conceptos clave de fidelización y marcas propias; luego, se detalla la metodología empleada para construir los indicadores y analizar los datos; a continuación, se exponen y discuten los principales resultados; y finalmente, se presentan las conclusiones y propuestas para futuras líneas de investigación.

2. Estado del arte

La segmentación de clientes ha sido una práctica fundamental en el marketing, evolucionando desde enfoques basados en simples variables demográficas hasta metodologías complejas que incorporan un análisis detallado de comportamientos, preferencias psicográficas y otros datos avanzados. Esta evolución refleja una respuesta directa a la creciente necesidad de personalización en las estrategias de marketing, impulsada por la individualidad y el cambio constante de los clientes, la reactividad de los competidores y la limitación de los recursos, como destacan Palmatier et al. (2022).

En los inicios, los modelos como RFM, que clasifican a los clientes por la recencia, frecuencia y valor monetario de sus compras, proporcionaron un marco efectivo para entender los patrones de compra. Sin embargo, los avances en la recolección de datos, especialmente a través de sistemas como las tarjetas de fidelidad, han permitido un análisis más profundo y personalizado, revelando que las necesidades de los consumidores son más complejas y diversas de lo que los enfoques tradicionales podrían capturar (Griva et al., 2022). Por ejemplo, en el estudio de Rahim et al. (2021), además de utilizar la segmentación basada en RFM, se incorpora el ingreso como un elemento de clasificación. Este enfoque más amplio y detallado tiene como objetivo capturar de manera más precisa los patrones de recompra y, en última instancia, mejorar la comprensión del comportamiento del cliente.

Con el tiempo, las técnicas de segmentación han incorporado métodos estadísticos y computacionales más avanzados, como K-means, clustering jerárquico, fuzzy clustering y redes neuronales. Estos métodos han sido validados por índices como el silhouette, que ayuda a medir la cohesión y separación entre los segmentos identificados, ofreciendo así una mejor evaluación de su efectividad. Por ejemplo, el estudio de Ballestar et al. (2018) en el ámbito del comercio electrónico demostró que la segmentación basada en el rol de los clientes en redes sociales, su antigüedad en la página y la rentabilidad obtenida puede conducir al desarrollo de estrategias de retención personalizadas. Estas estrategias se lograron gracias a una comprensión profunda del comportamiento del cliente dentro de su red social, validada mediante el análisis de conglomerados de dos pasos y el índice de silhouette.

El estudio de John et al. (2023) reveló que, al someter a comparación diversos algoritmos de clustering, el modelo de mezcla gaussiana (GMM) sobresalió con un índice de silhouette significativamente superior. Este resultado subraya la imperiosa necesidad de una meticulosa exploración y selección de los algoritmos de agrupamiento más apropiados, adecuados a las particularidades de cada caso específico.

Más recientemente, la necesidad de adaptar y afinar aún más estas técnicas ha llevado a la exploración de algoritmos basados en "beneficios buscados" por los clientes en sectores específicos como la hotelería, donde Arunachalam and Kumar (2018) sugirieron que la selección adecuada de una medida de distancia y un algoritmo de agrupamiento son cruciales para obtener resultados precisos y significativos. Asimismo, la integración de métodos de inteligencia artificial ha permitido desarrollar modelos de segmentación innovadores que combinan redes neuronales profundas con técnicas de clustering probabilístico auto-supervisado, proporcionando una flexibilidad y adaptabilidad mayores para manejar la diversidad de datos de clientes (Nguyen, 2021).

Por otro lado, el empleo de técnicas de aprendizaje automático ha demostrado ser un recurso valioso para mejorar la segmentación de clientes al integrar variables como demografía, RFM y comportamiento del cliente (Kasem et al., 2024). Destaca el enfoque de segmentación dinámica, que permite ajustar los segmentos en tiempo real según los cambios en los comportamientos de los clientes y las condiciones del mercado. Este método, apoyado por inteligencia artificial para identificar patrones complejos, facilita una segmentación más precisa y adaptativa, fortaleciendo el compromiso del cliente con resultados sostenibles.

La propuesta actual de segmentación de clientes en un retail se alinea con estas investigaciones avanzadas, buscando integrar y aplicar lo último en técnicas de modelado de datos, como machine learning,

para abordar las complejidades de la segmentación de clientes. Al combinar enfoques tradicionales con innovaciones recientes, como el análisis de concepto formal (FCA) junto con modelos como RFM, como proponen Rungruang et al. (2024), se espera mejorar la representación del conocimiento del cliente y optimizar las estrategias de marketing. Este enfoque integral no solo abarca la identificación de segmentos basados en patrones de gasto y comportamiento de compra, sino que también contempla el desarrollo de plataformas más sofisticadas que permitan a las empresas responder dinámicamente a las necesidades cambiantes de sus clientes, asegurando así su satisfacción y lealtad a largo plazo.

Entre la diversidad de enfoques de segmentación disponibles, resulta fundamental seleccionar criterios que no solo permitan clasificar clientes en grupos distintos, sino que además aporten información estratégica para la gestión del retail. En este sentido, la lealtad hacia el retail o hacia una marca constituye un análisis relevante, al permitir estudiar cómo los comportamientos de compra influyen en la relación del cliente con la tienda. Por ejemplo, Martos-Partal and González-Benito (2009) destacan una relación positiva entre la familiaridad y el reconocimiento de un producto con la lealtad a la tienda, sugiriendo extender este estudio a otros productos. En particular, la relación entre la lealtad a las marcas propias y a la tienda es de especial interés, ya que las marcas propias no solo representan una fuente directa de diferenciación para los minoristas, sino que también actúan como un catalizador para fortalecer la fidelización del cliente hacia la tienda en su conjunto.

En línea con esta perspectiva, Coelho do Vale et al. (2016) emplearon un modelo estructural con análisis factorial confirmatorio (CFA) para examinar cómo las actitudes hacia las marcas propias influyen en la lealtad al supermercado. El CFA permitió validar constructos latentes como confianza, valor percibido y actitud hacia la marca propia, evaluando su validez mediante CR (Composite Reliability). Además, el estudio consideró supermercados con distintos posicionamientos. En particular, en el caso del supermercado low-cost, la lealtad a marca propia no se tradujo en una mayor lealtad al supermercado. De hecho, la fidelidad al establecimiento estuvo explicada principalmente por factores como el precio y la conveniencia, y no por la actitud hacia la marca propia. Dado que nuestro caso de estudio se sitúa precisamente en un supermercado percibido como low-cost, esta evidencia refuerza la pertinencia de considerar ambas lealtades como dimensiones distintas pero interrelacionadas, en función del contexto estratégico del retail.

La relevancia del vínculo entre la lealtad a marca propia y la lealtad al supermercado se refuerza al considerar la evolución reciente en la percepción de las marcas propias. Inicialmente asociadas a productos de baja calidad debido a sus precios reducidos, esta percepción ha cambiado drásticamente. Según Ndlovu (2024), la diferencia de precios entre marcas propias y nacionales ha disminuido de un rango del 50 % - 70 % a uno promedio de 15 % - 50 %, reflejando una mayor competitividad en el mercado. Este cambio no solo posiciona a las marcas propias como una alternativa atractiva para los consumidores, sino que las convierte en un activo estratégico para los minoristas, al ofrecerles una herramienta para diferenciarse en mercados saturados y fidelizar a sus clientes de manera sostenible.

Este cambio en la percepción y el valor estratégico de las marcas propias se ha reflejado en su creciente participación en el mercado global, consolidando su rol como un elemento clave en las estrategias de fidelización. En términos de participación, Gielens et al. (2023) destaca que las marcas propias representaron el 21.6 % del mercado global en 2019, con cuotas superiores al 30 % en Europa Occidental y alrededor del 20 % en Estados Unidos. En mercados emergentes, su crecimiento presenta tasas de dos dígitos, lo que anticipa una convergencia hacia niveles similares a los de economías desarrolladas. Este avance posiciona a las marcas propias como una herramienta estratégica para fidelizar clientes y diferenciar a los minoristas en un entorno competitivo, donde factores como la calidad percibida y la innovación resultan claves para consolidar su éxito y preferencia entre los consumidores. Este crecimiento resalta el papel estratégico de las marcas propias para fidelizar clientes y diferenciar a los minoristas en un entorno altamente competitivo.

En este contexto, la segmentación de clientes es fundamental para fortalecer la lealtad a la marca, ya sea propia o no, ya que permite analizar las diferencias en los comportamientos y preferencias de los clientes para desarrollar estrategias adaptadas a cada grupo. Según Arunachalam and Kumar (2018) y Bucklin et al. (1995), segmentar clientes facilita el diseño de estrategias de marketing personalizadas, mejorando la interacción con los consumidores y promoviendo su fidelización. No obstante, cuantificar la lealtad representa un desafío debido a su complejidad. El mismo estudio de Bucklin et al. (1995) analiza la segmentación basada en la respuesta a las ofertas de precio y la probabilidad de cambiar de marca, realizando análisis de sensibilidad de precios. En este estudio, la lealtad se mide simplemente mediante la frecuencia de compra. Sin embargo, se puede desarrollar un enfoque más complejo y sofisticado, combinando la respuesta a ofertas con otras métricas de lealtad, como la duración de la relación con la marca y la probabilidad de recomendación, optimizando así las estrategias de marketing.

A medida que la investigación en lealtad de marca ha avanzado, se han identificado más de 200 definiciones del concepto, reflejando una falta de consenso en la literatura (Knox and Walker, 2001). La medición de la lealtad del consumidor ha evolucionado hacia enfoques que combinan tanto métricas conductuales como actitudinales, destacando sus respectivas fortalezas y limitaciones. Las métricas conductuales, como la frecuencia de compra y la participación en la categoría, se han consolidado como indicadores clave debido a su facilidad de cuantificación y análisis. Según Romaniuk and Nenycz-Thiel (2013), estas métricas permiten identificar patrones observables en las decisiones de compra, aunque no necesariamente capturan las motivaciones subyacentes del consumidor. Por otro lado, la dimensión actitudinal evalúa la conexión emocional del cliente con la marca, proporcionando un entendimiento más profundo de su fidelidad (Møller Jensen, 2011). Sin embargo, esta dimensión es más volátil y puede no reflejarse en comportamientos consistentes, lo que limita su utilidad como métrica exclusiva a largo plazo. La integración de ambas perspectivas en investigaciones previas ha demostrado ofrecer una visión más completa de la lealtad del cliente. Este enfoque combinado no solo permite identificar patrones de consumo repetido, sino también comprender cómo las percepciones y emociones del cliente influyen en su comportamiento.

En mercados de alta competencia, las decisiones de compra están influenciadas por múltiples factores. Por ejemplo, la disponibilidad de stock y la percepción de ofertas y promociones son elementos críticos para mantener la lealtad, ya que incluso clientes con alta conexión emocional hacia una marca pueden optar por otras alternativas si el producto no está disponible o si otras marcas presentan promociones atractivas (Knox and Walker, 2001). Asimismo, el índice de apoyo a la marca, propuesto por Knox and Walker (2001), mide el grado en que las compras están concentradas en un conjunto limitado de marcas dentro de una categoría, ofreciendo una perspectiva sobre la fidelidad del consumidor en mercados diversificados. Complementariamente, el concepto de multilealtad, basado en el marco conceptual de Jacoby and Olson (1970) y desarrollado por Knox and Walker (2001), refleja la tendencia de los consumidores a ser leales a varias marcas en lugar de una sola, una realidad especialmente relevante en el sector de bienes de consumo rápido.

Por último, el Market Share en Gasto ha demostrado ser una métrica clave para analizar la proporción del gasto del cliente dedicado a una marca específica en relación con el total de la categoría. Según Romaniuk and Nenycz-Thiel (2013), esta métrica permite segmentar a los consumidores en diferentes niveles de lealtad y comprender cómo estas participaciones influyen en las asociaciones de marca y el comportamiento de compra. Estas variables ofrecen una base sólida para explorar la lealtad del cliente en el contexto de las marcas propias, proporcionando un marco integral para optimizar estrategias de marketing y mejorar la fidelización.

Finalmente, es crucial que las estrategias de marketing prioricen una segmentación adaptable, capaz de responder a la evolución constante del comportamiento del consumidor. La inclusión de marcas propias en modelos de segmentación y análisis de lealtad desempeña un rol estratégico en la fidelización del cliente

y la diferenciación de los minoristas en un mercado competitivo. Para lograr resultados óptimos, se recomienda realizar comparativas entre métodos de segmentación y seleccionar aquel con el mejor índice de validación. Sin embargo, el análisis de la lealtad no debe limitarse a la segmentación. La validación de constructos latentes mediante análisis factorial confirmatorio (CFA) permite cuantificar de manera robusta dimensiones como la lealtad hacia la marca propia y hacia el supermercado, integrando múltiples indicadores de comportamiento. Posteriormente, la regresión lineal proporciona un marco analítico para explorar cómo estas formas de lealtad se relacionan entre sí, identificando posibles efectos de refuerzo o independencia. Este enfoque combinado permite no solo identificar perfiles diferenciados de clientes, sino también comprender con mayor precisión los mecanismos que explican su comportamiento de fidelidad, aportando evidencia valiosa para el diseño de estrategias comerciales más efectivas.

3. Hipótesis y Objetivos

Esta investigación se centra en comprender cómo se relaciona la lealtad conductual del cliente hacia la marca propia con su fidelidad general al supermercado, considerando que ambas pueden responder a motivaciones distintas o reforzarse mutuamente. Para abordar este fenómeno, se propone construir indicadores compuestos de lealtad a partir de datos transaccionales, entendiendo la lealtad no como una variable única, sino como un concepto que se manifiesta a través de múltiples dimensiones del comportamiento del consumidor.

El objetivo principal de esta investigación es identificar y caracterizar los factores que configuran la lealtad hacia la marca propia y hacia el supermercado dentro de una cadena de retail, construyendo indicadores para ambas dimensiones y segmentando a los clientes en función de sus comportamientos de compra, con el fin de analizar explícitamente cómo se relacionan ambas formas de lealtad entre sí. Es decir, se busca determinar si la lealtad a la marca propia refuerza la fidelidad al supermercado, o si opera como un comportamiento autónomo.

3.1. Objetivos Específicos

- Caracterizar variables transaccionales que permitan representar y medir conductualmente la lealtad del cliente hacia la marca propia y el supermercado.
- Construir indicadores compuestos de lealtad conductual mediante análisis factorial confirmatorio (CFA), con el fin de sintetizar múltiples variables en dimensiones latentes representativas.
- Analizar la relación entre lealtad a marca propia y lealtad al supermercado mediante técnicas de regresión lineal, considerando posibles efectos diferenciales según segmentos de clientes.
- Evaluar perfiles de clientes mediante segmentación basada en comportamiento, identificando grupos con distintos patrones de fidelización y explorando si existen casos de lealtad convergente entre ambas dimensiones.

3.2. Hipótesis

Existe un segmento de clientes cuya lealtad hacia la marca propia actúa como un factor potenciador de su fidelidad general al supermercado, lo que sugiere la presencia de un patrón de lealtad convergente entre ambas dimensiones de comportamiento.

3.3. Contexto del Problema

En un entorno altamente competitivo, las empresas de retail enfrentan el desafío de comprender y fortalecer la lealtad de los clientes hacia sus productos, especialmente en el caso de las marcas propias. Estas marcas emergen como un elemento diferenciador clave al ofrecer productos de calidad a precios accesibles, una característica que adquiere particular importancia en el mercado chileno, donde las condiciones económicas impulsan a los consumidores a buscar alternativas que equilibren calidad y precio.

El presente estudio analiza una cadena de supermercados e hipermercados chilenos perteneciente a un holding reconocido, con operaciones en Chile y Perú. En el territorio nacional, la cadena tiene presencia en seis regiones y ofrece una amplia gama de productos, incluyendo marcas propias que representan un componente clave de su oferta comercial. Esta cadena está orientada principalmente a consumidores de estrato socioeconómico medio y medio-bajo, lo que se refleja en su enfoque en precios accesibles y estrategias de valor percibido, especialmente a través del desarrollo de marcas propias. El análisis se

enfoca en estudiar la lealtad de los clientes hacia las marcas propias de esta cadena de supermercados. Utilizando herramientas de análisis de datos y técnicas de análisis cuantitativo, se busca identificar patrones de comportamiento relevantes que permitan comprender mejor la relación entre los clientes y la marca propia, contribuyendo al conocimiento sobre las dinámicas de fidelización.

Cabe destacar que el supermercado opera con diversas marcas propias, entre las cuales se incluye una que lleva el nombre del supermercado y otras que utilizan nombres distintos. Esta diferenciación permite evaluar el posicionamiento de la marca propia en distintos contextos comerciales y estrategias de branding.

3.4. Relevancia del Estudio

La investigación aborda un aspecto clave para el sector retail: comprender la lealtad de los clientes hacia las marcas propias, un elemento que contribuye tanto a la sostenibilidad operativa como al fortalecimiento del vínculo con los consumidores. Aunque ampliamente valorada por las empresas, la lealtad ha sido históricamente un concepto difuso, sin una definición consensuada ni una aplicación teórica consistente, lo que ha dificultado su medición precisa (Knox and Walker, 2001). Esta investigación responde a esa limitación mediante la definición estructurada del concepto de lealtad y su medición cuantificable, aplicadas empíricamente para analizar su comportamiento en el contexto del retail.

Ante este escenario, uno de los principales aportes de este estudio es justamente avanzar en la definición y medición estructurada de la lealtad del cliente. A través del análisis de datos y la aplicación de técnicas cuantitativas avanzadas, se busca caracterizar mejor la dinámica entre los clientes y la marca propia, abordando la lealtad desde una perspectiva conductual.

Además, los resultados obtenidos contribuyen al conocimiento teórico sobre las interacciones entre la fidelidad hacia una marca específica y la lealtad global hacia un supermercado, proporcionando un marco de análisis que puede ser aplicado en contextos similares. Esta aproximación representa un aporte significativo, no solo para la toma de decisiones comerciales, sino también para la evolución conceptual del estudio de la lealtad en marketing.

3.5. Preguntas de Investigación

1. ¿Qué variables representan de manera más efectiva la lealtad de los clientes hacia la marca propia?
2. ¿Cómo se relaciona la lealtad hacia la marca propia con la lealtad hacia el supermercado?
3. ¿Qué caracteriza a los segmentos de clientes identificados y cómo varían en sus patrones de compra?

3.6. Alcances y Limitaciones

El estudio se basa en datos transaccionales históricos obtenidos entre el 1 de julio de 2013 y el 26 de agosto de 2016. Aunque estos datos proporcionan un marco sólido para el análisis, su antigüedad puede limitar su aplicabilidad en un entorno de consumo actual. Además, el análisis está restringido al contexto de un solo país, Chile, lo que podría dificultar la generalización de los hallazgos a otros mercados con dinámicas diferentes.

Otro aspecto a considerar es que la cadena de supermercados estudiada no lidera en volumen de compras a nivel nacional. Esto podría influir en las tendencias observadas, focalizándolas en un segmento específico del mercado chileno, sin necesariamente reflejar patrones generales de consumo en el país.

4. Metodología

Este estudio se estructura según la metodología CRISP-DM (Cross-Industry Standard Process for Data Mining), un enfoque ampliamente reconocido en proyectos de análisis de datos por su carácter flexible, iterativo y orientado a resultados (Rawat, 2023).

El objetivo metodológico es desarrollar un modelo que permita identificar qué clientes, a partir de su comportamiento con productos de marca propia, tienen mayor potencial de lealtad al supermercado. Para ello se aplican técnicas como análisis exploratorio, modelado factorial confirmatorio (CFA), regresión lineal múltiple y segmentación mediante K-means, incluyendo interacciones por perfil de cliente.

4.1. Análisis Descriptivo de los Datos

El análisis se basa en una base de datos transaccional proveniente de un supermercado chileno, la cual abarca un período de tres años, desde el 1 de julio de 2013 hasta el 26 de agosto de 2016. La información se encuentra almacenada en AWS Amazon, lo que permite un acceso estructurado y eficiente para su procesamiento. Los datos incluyen detalles a nivel de transacción y producto, lo que posibilita el estudio del comportamiento individual de compra y el seguimiento longitudinal de los clientes. Cabe señalar que la base de datos se encuentra anonimizada, por lo que no es posible identificar a los individuos que formaron parte de esta muestra.

Entre las variables clave disponibles se encuentran:

- **Identificador del Cliente (ID):** Permite rastrear las compras de un cliente específico a lo largo del tiempo, lo que facilita el análisis del comportamiento individual. Este identificador fue anonimizado utilizando una función hash criptográfica (SHA-256), asegurando que no sea posible identificar directamente a los individuos.
- **Número de Boleta:** Cada transacción cuenta con un número único, lo que ayuda a diferenciar las compras realizadas en momentos diferentes, así como a agrupar productos comprados simultáneamente.
- **Fecha de Venta:** Se tiene registro de la fecha exacta de cada transacción, lo que permite estudiar patrones de compra a lo largo del tiempo, como la estacionalidad de ciertos productos o la recurrencia de las compras.
- **Cantidad de Unidades Vendidas:** Esta variable permite cuantificar cuántas unidades de un producto específico se adquirieron en cada transacción.
- **Monto en UF:** El valor de cada compra, expresado en unidades de fomento (UF). Lo que sirve para controlar por inflación.
- **Marca Propia:** Se registra si el producto pertenece a la marca propia del supermercado, lo que resulta crucial para evaluar la preferencia de los consumidores por la marca propia frente a otras marcas.
- **Descuento:** Se registra si el producto estaba en descuento al momento de realizar la compra. Es útil para observar la influencia de las promociones en las decisiones de compra.
- **Estructura Jerárquica del Producto:** Los productos están clasificados de manera jerárquica, desde secciones generales hasta productos específicos. Esta estructura incluye varios niveles, como se muestra a continuación:



Figura 1: Estructura de los Niveles de los Productos Aplicada a un Ejemplo Particular.

Esta estructura jerárquica permite analizar el comportamiento de compra en distintos niveles de agregación y desagregación, facilitando la evaluación del desempeño de la marca propia por secciones, categorías o subcategorías específicas.

4.2. Ingeniería de los atributos: ¿Cómo medir la lealtad?

Uno de los aportes clave de este estudio es la construcción de un índice propio de lealtad del cliente, tanto hacia la marca propia como hacia el supermercado. Este índice se diseña a partir de un conjunto de variables transaccionales que capturan patrones de comportamiento reales, cuidadosamente seleccionadas por su sustento teórico y validación empírica en investigaciones previas.

El análisis se centró en las compras realizadas entre el 1 de julio de 2015 y el 26 de agosto de 2016, periodo en el cual se calcularon métricas que permiten describir distintas dimensiones de la lealtad, tales como la frecuencia de compra, la participación de marca propia, la recompra, el comportamiento frente a promociones, entre otras. Estas variables constituyen un insumo fundamental para los análisis de segmentación y modelamiento posteriores, al representar la lealtad como un fenómeno complejo y multidimensional.

4.2.1. Variables de Lealtad a Marca Propia

A partir de la información transaccional disponible, se construyeron variables específicas que permiten caracterizar la lealtad hacia la marca propia del supermercado. Estas métricas fueron diseñadas bajo un enfoque de ingeniería de atributos, que permite transformar datos brutos en indicadores con significado analítico. Cada una de estas variables refleja una dimensión distinta del comportamiento del cliente, lo que permite capturar la lealtad desde una perspectiva más amplia y precisa.

A continuación, se describen las variables desarrolladas para este propósito, junto con su interpretación y relevancia dentro del estudio.

■ Frecuencia:

Esta variable mide la cantidad total de compras que un cliente realizó de productos de marca propia durante el periodo analizado. Para su cálculo, se identificaron todas las transacciones asociadas a cada cliente (mediante su ID) y se contabilizaron aquellas en las que adquirió al menos un producto perteneciente a una marca propia.

El resultado representa una medida directa del nivel de interacción del cliente con la marca propia. La hipótesis subyacente establece que una mayor frecuencia de compra estaría asociada a una mayor lealtad, ya que refleja un comportamiento repetitivo y sostenido de preferencia hacia la marca en cuestión.

La información obtenida de este análisis se fundamenta en estudios previos, como el desarrollado por Møller Jensen (2011) en su marco teórico sobre la lealtad del consumidor, donde se reconoce la importancia de la frecuencia de compra como un indicador clave de lealtad conductual. Además, este análisis permite identificar clientes con comportamientos consistentes hacia una marca específica y establecer una primera aproximación a la segmentación basada en su nivel de lealtad.

■ **Market Share en Gasto:**

Esta variable mide la proporción del gasto total de un cliente que fue destinado a productos de marca propia dentro de las categorías estudiadas, en relación con su gasto total en dichas categorías. El objetivo es capturar el nivel de preferencia económica del cliente hacia la marca propia, expresado en términos monetarios.

El cálculo se realizó sumando el gasto total de cada cliente en productos de marca propia y dividiéndolo por el gasto total del cliente en la categoría correspondiente. Solo se consideraron transacciones dentro del periodo definido.

Este enfoque se basa en la perspectiva de Romaniuk and Nenycz-Thiel (2013), quien plantea que la participación de mercado puede ser una manifestación observable de la lealtad, ya que un mayor gasto relativo en una marca sugiere una asociación más fuerte con ella.

Un valor alto en esta métrica indica que el cliente concentra una parte significativa de su presupuesto en productos de marca propia, lo que puede interpretarse como un fuerte vínculo de lealtad. Por el contrario, un valor bajo sugiere una mayor dispersión del gasto entre diferentes marcas, lo cual refleja una menor fidelidad.

Aunque esta variable no capta directamente percepciones o actitudes, ofrece una visión agregada del comportamiento de compra que permite identificar patrones de preferencia hacia las marcas propias dentro de un entorno competitivo.

■ **Re compra o compra repetida:**

La variable Recompra o compra repetida tiene como objetivo evaluar la consistencia del cliente al elegir productos de marcas propias dentro de la categoría seleccionada. El cálculo se estructuró en varias etapas. En primer lugar, se ordenaron todas las compras de cada cliente (RUT) desde la más reciente hasta la más antigua. Luego, se seleccionaron las últimas 10 compras de cada cliente dentro de cada categoría, con el fin de centrar el análisis en el comportamiento de compra reciente.

A continuación, se identificó cuántas de estas compras correspondían a productos de marca propia. El porcentaje de recompra se calculó como la proporción entre el número de compras de cada marca propia y el total de compras seleccionadas (10 compras), multiplicando este valor por 100. De esta manera, se obtiene una métrica que refleja la recurrencia de los clientes al elegir productos de cada una de las marcas propias.

Este análisis parte de la hipótesis de que un cliente con un porcentaje elevado en esta variable evidencia una mayor consistencia en su comportamiento de compra, lo que se traduce en una preferencia sostenida por una misma marca y una menor propensión a alternar entre opciones.

La construcción de esta variable se respalda en los fundamentos teóricos propuestos por Møller Jensen (2011), quienes destacan que la frecuencia de recompra y la repetición de compras son indicadores clave de la lealtad conductual. Medir la recompra permite observar la estabilidad de las elecciones del cliente y proporciona una base sólida para distinguir entre comportamientos ocasionales y preferencias consistentes hacia las marcas propias.

Este análisis descriptivo es esencial para identificar patrones específicos de comportamiento dentro de marcas propias del supermercado, facilitando una comprensión más profunda de la lealtad del cliente y su relación con la estabilidad de sus decisiones de compra.

■ **Fidelidad Activa:**

La variable Fidelidad Activa mide, para cada cliente, el porcentaje de ocasiones en que, estando disponible un producto de marca propia en el punto de venta, este optó efectivamente por adquirirlo. Su cálculo se basa en el análisis de todas las transacciones en las que existía disponibilidad

documentada de productos de marca propia, y se registra si el cliente eligió dicha marca por sobre otras alternativas presentes.

Esta métrica parte de la hipótesis conductual planteada por Møller Jensen (2011), que sugiere que los consumidores con alta lealtad a una marca preferida tienden a evitar reemplazarla cuando no está disponible, incluso postergando la compra. Sin embargo, la aplicación directa de esta lógica a datos transaccionales presentó limitaciones metodológicas importantes, como la baja representatividad de casos y la dificultad para observar de manera precisa la ausencia de stock a nivel individual.

Dado lo anterior, se adoptó un enfoque alternativo que pone el foco en el comportamiento del cliente bajo condiciones de disponibilidad. Así, Fidelidad Activa se interpreta como la proporción de veces que el cliente, teniendo la opción de elegir, selecciona la marca propia. Un valor alto en esta variable indica una preferencia consistente y deliberada por la marca del distribuidor, mientras que un valor bajo sugiere una tendencia a elegir otras marcas, aun cuando la marca propia está disponible.

Esta métrica ofrece una aproximación operativa y robusta para evaluar la lealtad conductual, permitiendo distinguir entre quienes escogen activamente la marca propia y quienes lo hacen de forma esporádica o circunstancial.

■ **Proporción de Productos:**

La variable Proporción de Productos de Marca Propia representa el porcentaje de artículos adquiridos por un cliente que corresponden a marcas propias del supermercado, en relación con el total de productos comprados dentro del período de análisis. A diferencia de otras métricas que consideran el gasto o la frecuencia, esta variable se enfoca en la cantidad de ítems, ofreciendo una mirada complementaria centrada en el volumen físico de compra.

Para su cálculo, se contabilizó la cantidad total de productos comprados por cada cliente, y dentro de este conjunto, se identificó cuántos de esos productos pertenecían a alguna de las marcas propias del supermercado. La división de estos dos valores entrega la proporción deseada.

Esta métrica permite capturar de forma clara el grado de incorporación de la marca propia en el comportamiento de compra cotidiano del cliente. Un valor elevado en esta proporción sugiere una preferencia consistente por productos del distribuidor, lo que puede interpretarse como una señal de lealtad estructural. En cambio, valores bajos podrían reflejar una preferencia por marcas externas o una exploración más diversificada dentro de la categoría.

■ **Oferta:**

La variable Oferta fue diseñada para capturar la proporción de compras en las que el cliente eligió productos en descuento que no pertenecen a la marca propia del supermercado, en relación con el total de sus transacciones dentro de las categorías seleccionadas. Esta métrica permite evaluar la sensibilidad al precio del consumidor y su propensión a desviarse de la marca propia ante estímulos promocionales externos.

La lógica subyacente plantea que un cliente altamente leal tendería a mantener su preferencia por la marca propia, incluso en presencia de ofertas atractivas en otras marcas. Por el contrario, una mayor proporción de compras de productos en descuento ajenos a la marca propia puede interpretarse como una menor fidelidad conductual y una mayor receptividad a estímulos de precio.

Para su cálculo, se analizaron todas las transacciones de cada cliente. A cada una se le asignó un valor binario: 1 si el producto adquirido estaba en oferta y no correspondía a una marca propia, y

0 en caso contrario. La variable se definió como la proporción de transacciones con valor 1 respecto al total de compras realizadas por el cliente en la categoría considerada (por ejemplo, harina).

El fundamento teórico de esta variable se apoya en Møller Jensen (2011), quien argumenta que el comportamiento frente a ofertas es un componente clave para comprender la lealtad efectiva del consumidor, ya que pone en evidencia hasta qué punto las decisiones están guiadas por el vínculo con la marca frente a incentivos externos como descuentos.

■ **Promoción:**

La variable Promoción busca capturar el nivel de sensibilidad al precio de los clientes a partir de su comportamiento frente a productos en descuento, independientemente de si pertenecen o no a la marca propia del supermercado. A diferencia de la variable Oferta, que considera únicamente las compras con descuento en marcas externas, esta métrica incorpora la totalidad de productos adquiridos en promoción, incluyendo los de marca propia.

El supuesto teórico subyacente es que una mayor proporción de compras en promoción refleja una mayor sensibilidad al precio, lo que puede influir en la estabilidad de las decisiones de compra del cliente. En este sentido, clientes altamente influenciados por descuentos podrían presentar una lealtad más volátil, priorizando el precio por sobre la preferencia sostenida por una marca específica.

Esta métrica permite identificar a aquellos consumidores cuyo comportamiento está fuertemente orientado por incentivos promocionales, lo cual resulta relevante para estrategias comerciales basadas en descuentos, así como para el análisis de la fidelidad real frente a estímulos de precio.

4.2.2. Variables Lealtad al Supermercado

En esta sección se presentan las variables diseñadas para capturar el nivel de lealtad general del cliente hacia el supermercado, más allá de su afinidad con las marcas propias. A diferencia de las variables de lealtad a marca propia, que se construyen sobre categorías específicas, estas métricas se basan en la totalidad de las transacciones realizadas por cada cliente durante el período de análisis, lo que permite evaluar su compromiso global con el establecimiento comercial.

El objetivo de estas variables es ofrecer una mirada integral sobre el comportamiento de compra. Estas dimensiones permiten aproximarse a la noción de lealtad desde una perspectiva más amplia y estratégica, al considerar el vínculo del cliente con el supermercado en su conjunto, y no solo con ciertos productos o marcas.

■ **RFM**

El modelo RFM (Recency–Frequency–Monetary) es uno de los indicadores más reconocidos y ampliamente utilizados en la segmentación de clientes. En este estudio, fue empleado como una aproximación para medir la lealtad hacia el supermercado. En este contexto, la recencia se define como la cantidad de días transcurridos desde la última compra del cliente; la frecuencia, como la cantidad de compras en el supermercado; y el valor monetario, como el total gastado por el cliente durante el período de análisis.

Cabe destacar que, dado que la cadena está orientada principalmente a consumidores de estratos socioeconómicos medio y medio-bajo, un mayor nivel de gasto no necesariamente refleja una mayor lealtad. En este contexto, es posible que clientes altamente comprometidos con el supermercado no presenten los mayores desembolsos, debido a restricciones presupuestarias. Por tanto, aunque el componente monetario sigue siendo una dimensión relevante para el análisis, su interpretación debe considerarse con cautela y siempre en conjunto con otras métricas conductuales.

■ Market Share en Gasto

La variable Market Share en gasto fue adaptada para medir el nivel de preferencia de cada cliente por el supermercado en estudio. Para su cálculo, se utilizó información del gasto total realizado por cada cliente en todos los supermercados durante el período disponible. A partir de ello, se determinó la proporción del gasto correspondiente al supermercado focal, dividiendo el monto desembolsado en esta cadena por el gasto total del cliente entre todos los supermercados. El resultado representa el porcentaje del gasto que cada cliente destina al supermercado de interés, lo que permite aproximarse a su grado de afinidad o preferencia por esta cadena en comparación con otras alternativas del mercado. Esta variable complementa el análisis de lealtad desde una perspectiva competitiva, al incorporar información externa al comportamiento exclusivo dentro del supermercado estudiado.

4.3. Aplicación de Algoritmos

Una vez construidas las variables que representan distintas dimensiones de la lealtad del cliente, se procede a la aplicación de técnicas analíticas avanzadas que permiten modelar y segmentar el comportamiento observado.

Para ello, se aplican tres enfoques metodológicos:

- **Análisis Factorial Confirmatorio (CFA):** para estimar dos constructos latentes que representen la lealtad del cliente, uno para marca propia y otro para el supermercado.
- **Regresión Lineal Múltiple:** para evaluar el impacto de diversas variables explicativas sobre la lealtad al supermercado.
- **Segmentación con K-means:** para identificar grupos homogéneos de clientes a partir de sus características de lealtad.

Estas técnicas, combinadas, permiten no solo validar constructos teóricos, sino también explorar relaciones causales y patrones subyacentes dentro del comportamiento del consumidor en el entorno transaccional del supermercado.

4.3.1. Análisis Factorial Confirmatorio (CFA)

Con el objetivo de estimar rigurosamente dos constructos latentes fundamentales, *lealtad hacia la marca propia* y *lealtad hacia el supermercado*, se aplicó el Análisis Factorial Confirmatorio (CFA), una técnica derivada de los modelos de ecuaciones estructurales (SEM) ampliamente utilizada para validar modelos de medición en contextos de consumo.

Esta técnica permite evaluar si un conjunto de variables observadas refleja adecuadamente una estructura teórica definida a priori. En línea con Akoglu and Özbek (2021), quienes aplicaron CFA para validar los constructos de experiencia de marca, calidad percibida, confianza y lealtad, se adopta un enfoque similar para evaluar la validez de los índices construidos en este estudio.

El modelo especificado contempla dos factores correlacionados:

- **Lealtad hacia la marca propia**, compuesto por las variables : frecuencia, market share en gasto, fidelidad activa, recompra y proporción productos marca propia.
- **Lealtad hacia el supermercado**, compuesto por las variables : frecuencia, monto total, recencia y market share del supermercado.

Una vez validados, estos constructos se utilizarán como variables dependientes en los modelos de regresión y como insumos para la segmentación posterior.

4.3.2. Regresión Lineal Múltiple

Con el objetivo de analizar los factores que inciden en la lealtad hacia la marca propia y hacia el supermercado, se utilizó un modelo de regresión lineal múltiple. Esta técnica permite estimar el efecto de múltiples variables independientes sobre una variable dependiente continua, bajo el supuesto de relaciones lineales entre ellas.

La metodología sigue el enfoque de Sondoh et al. (2007), quienes aplicaron regresión lineal para estudiar cómo variables como imagen de marca, calidad percibida y satisfacción impactan la intención de lealtad en consumidores de cosméticos. En el presente estudio, la variable dependiente es el score estandarizado de lealtad al supermercado, mientras que entre las variables independientes se incluyen: el score de marca propia, sensibilidad al precio, variables socio demográficas, entre otras.

Este modelo permite interpretar los coeficientes estimados (β) como el cambio esperado en la lealtad al supermercado por unidad de cambio en cada predictor, controlando el efecto de las demás variables.

4.3.3. Segmentación con K-means

Para identificar perfiles de clientes con patrones de comportamiento similares, se aplicó el algoritmo K-means. Esta técnica permite clasificar a los consumidores en grupos homogéneos, minimizando la variabilidad interna de cada clúster y maximizando la diferencia entre grupos.

De acuerdo con John et al. (2023), K-means es una herramienta eficaz para estudios de comportamiento del consumidor, particularmente en análisis de segmentación de lealtad. La elección del número óptimo de clústeres se realizó utilizando el método del codo y el gráfico de silueta. Posteriormente, se empleó una inicialización aleatoria de centroides, seguida por iteraciones de asignación y actualización hasta alcanzar la convergencia del algoritmo.

El resultado es una clasificación robusta de los clientes, que luego se incorpora en los modelos de regresión mediante variables indicadoras e interacciones, enriqueciendo la interpretación de los resultados y permitiendo un análisis más granular de la lealtad.

4.4. Validación y Refinamiento

La validación de los segmentos obtenidos es un paso fundamental para asegurar la efectividad y utilidad de los modelos de segmentación. Este proceso no solo verifica la cohesión y separación de los clústeres, sino que también permite ajustar y optimizar los parámetros de los modelos, mejorando la precisión de la segmentación y asegurando que los grupos identificados representen adecuadamente el comportamiento de los clientes.

- **Evaluación de la calidad de CFA:** El modelo fue evaluado siguiendo las recomendaciones de Vale et al. (2016), considerando los siguientes criterios:

1. **Cargas factoriales estandarizadas (λ):** Se esperó que todas las cargas fueran superiores a 0.60, idealmente ≥ 0.7 , lo que indica que los ítems explican una proporción significativa de la varianza del constructo.
2. **Fiabilidad compuesta (CR):** Se calculó para cada factor como medida de consistencia interna. Valores de $CR \geq 0.70$ fueron considerados aceptables.
3. **Harman's One-Factor Test:** Para evaluar la posibilidad de sesgo por varianza común del método, se aplicó la prueba de un solo factor de Harman. Un porcentaje inferior al 50 % indica que no hay evidencia clara de un sesgo significativo por método común, lo cual refuerza la validez discriminante del modelo.

- **Evaluación de la calidad de Regresión lineal:**

La calidad del modelo fue evaluada mediante los siguientes indicadores:

1. **Coefficiente de determinación (R^2):** mide la proporción de varianza explicada por el modelo.
2. **Significancia de los coeficientes:** se utilizó el valor- p para contrastar la hipótesis nula $H_0 : \beta_i = 0$ para cada predictor. Un valor- p menor a 0.05 indica significancia estadística.

- **Evaluación de la calidad de los clústeres:** La validación de la segmentación obtenida mediante K-means se realizó a través de métricas de cohesión y separación que permiten evaluar la calidad interna de los clústeres generados. En particular, se consideraron los siguientes índices de validación interna, recomendados por John et al. (2023):

1. **Índice Silhouette:** Mide qué tan similar es una observación a su propio clúster en comparación con otros clústeres. Este índice varía entre -1 y 1, donde valores cercanos a 1 indican una correcta asignación, valores cercanos a 0 reflejan fronteras difusas entre clústeres, y valores negativos sugieren una asignación inadecuada.
2. **Índice de Davies-Bouldin (DBI):** Evalúa la compacidad intra-clúster y la separación inter-clúster. Valores bajos del DBI indican clústeres más compactos y mejor separados, lo cual se asocia a una segmentación de alta calidad.
3. **Índice de Calinski-Harabasz (CHI):** Estima la razón entre la dispersión entre clústeres y la dispersión dentro de los clústeres. Valores altos del CHI reflejan una estructura clara de segmentación y una diferenciación efectiva entre los grupos identificados.

Estas métricas permiten validar cuantitativamente la calidad de la solución de clústeres y constituyen un insumo clave para determinar si el número de grupos seleccionados es el más adecuado o si se requiere ajustar el modelo.

- **Calibración de parámetros:** Basado en los resultados de validación, se ajustan los parámetros de cada modelo para maximizar su desempeño. En K-means, esto puede implicar reevaluar el número óptimo de clústeres; en CFA, ajustar el modelo de medición; y en regresión, revisar la inclusión de variables o interacción de efectos.
- **Interpretación y refinamiento de resultados:** Finalmente, se analiza la interpretación sustantiva de los modelos: se revisan las características de cada clúster, se interpretan los coeficientes de regresión y se valida que los constructos latentes estimados tengan coherencia empírica. Este proceso permite asegurar que los hallazgos no solo sean estadísticamente sólidos, sino también significativos para el estudio.

Este enfoque riguroso de validación y refinamiento garantiza la solidez metodológica del estudio, asegurando que los resultados obtenidos sean confiables.

4.5. Análisis de Resultados y Evaluación de Impacto

Una vez validados los modelos y definidos los clústeres, se procede a un análisis riguroso de los resultados con el propósito de evaluar el desempeño de los modelos aplicados y comprender en profundidad los patrones de comportamiento de los distintos perfiles de clientes. Esta etapa tiene como objetivo interpretar los hallazgos desde una perspectiva analítica, evaluando su consistencia, utilidad práctica y relevancia.

En primer lugar, se analiza la calidad del modelo de Análisis Factorial Confirmatorio (CFA), el cual permite validar estadísticamente que los constructos latentes de lealtad, hacia la marca propia y hacia el supermercado, se encuentran correctamente definidos a partir de variables observables.

Posteriormente, se estiman modelos de regresión lineal múltiple con el fin de identificar los determinantes de la lealtad hacia el supermercado. A través del análisis de los coeficientes estimados, se evalúa la magnitud del efecto de cada variable independiente, tales como la lealtad a marca propia, el perfil socio demográfico o el comportamiento frente a descuentos. Este análisis permite validar empíricamente las hipótesis planteadas respecto al vínculo entre distintas dimensiones del comportamiento de compra y la fidelidad del cliente.

El análisis se complementa con la evaluación de los clústeres obtenidos, examinando los scores de lealtad y otras variables clave en cada grupo. Esto permite identificar segmentos bien diferenciados, caracterizados por patrones de fidelidad específicos, y evaluar la calidad de la segmentación desde una perspectiva tanto estadística como práctica.

Este enfoque integral, que combina validación de constructos, modelamiento explicativo y análisis segmentado, permite no solo confirmar la solidez técnica de los resultados obtenidos, sino también evaluar su aplicabilidad para la toma de decisiones estratégicas.

5. Resultados

El presente estudio se basa en un historial de transacciones correspondiente a un periodo total de tres años. No obstante, para el desarrollo de los modelos y análisis posteriores, se optó por utilizar exclusivamente los datos del último año disponible, comprendido entre el 1 de julio de 2015 y el 26 de agosto de 2016. Esta decisión metodológica se fundamentó en el análisis exploratorio inicial, donde se observó una mayor consistencia y claridad en los patrones de comportamiento de los clientes durante dicho periodo, en comparación con los años anteriores. Esta acotación temporal también favoreció la eficiencia computacional del modelado, sin comprometer la robustez analítica ni la validez de los resultados obtenidos.

Adicionalmente, se aplicó un filtro mínimo de actividad, incluyendo en el análisis únicamente a aquellos clientes que realizaron al menos cinco compras dentro del periodo definido, considerando transacciones en cualquiera de las categorías seleccionadas. Esta condición permitió asegurar un volumen suficiente de información por cliente, manteniendo a la vez un tamaño muestral representativo.

En cuanto al conjunto de categorías analizadas, se seleccionaron 56 categorías definidas bajo dos criterios: en primer lugar, se incluyeron todas aquellas pertenecientes a la Canasta Básica de Alimentos 2016 establecida por el Gobierno de Chile (Ministerio de Desarrollo Social y Familia, 2016); en segundo lugar, se incorporaron algunas categorías adicionales de productos de higiene personal, con el objetivo de simular una canasta básica integral. Como requisito adicional, solo se consideraron categorías que contaran con al menos un producto de marca propia, dado que el análisis de lealtad hacia esta línea de productos constituye uno de los ejes centrales del estudio.

Todos los cálculos, métricas e inferencias presentadas a continuación se basan exclusivamente en este subconjunto filtrado y estructurado de datos. El procesamiento, análisis estadístico y modelamiento se llevaron a cabo utilizando el lenguaje de programación `Python`, dada su capacidad para manejar grandes volúmenes de datos, implementar modelos estadísticos avanzados y facilitar la replicabilidad del análisis mediante código estructurado y reproducible.

5.1. Modelo general sin segmentación: CFA y Regresión Lineal

5.1.1. Explicación del modelo

Con el objetivo de representar de forma sintética y estructurada los conceptos de lealtad a marca propia y lealtad al supermercado, se aplicó un modelo de Análisis Factorial Confirmatorio (CFA) con dos factores latentes correlacionados. Esta técnica, derivada de los Modelos de Ecuaciones Estructurales (SEM), permite estimar variables no observables a partir de un conjunto de indicadores empíricos, siempre que exista una base teórica sólida que respalde su agrupación.

El modelo considera dos constructos latentes:

- **score_mp**: representa el nivel de lealtad del cliente hacia la marca propia del supermercado.
- **score_tienda**: representa el nivel de lealtad general del cliente hacia el supermercado como canal de compra.

Adicionalmente, se especifica una relación de dependencia entre ambos factores, bajo la hipótesis de que una mayor lealtad hacia la marca propia podría asociarse a una mayor fidelidad hacia la tienda en su conjunto.

Estructura del modelo

$\text{score_mp} \sim \text{fidelidad_activa} + \text{market_share_mp} + \text{proporción_mp} + \text{recompra_mp} + \text{frecuencia_mp}$
 $\text{score_tienda} \sim \text{gasto_total} + \text{frecuencia_tienda} + \text{market_share_tienda} + \text{recencia_tienda}$
 $\text{score_tienda} \sim \text{score_mp}$

Variables observadas utilizadas

■ Para `score_mp`:

- *fidelidad_activa*: proporción de ocasiones en que el cliente eligió marca propia cuando esta estuvo disponible.
- *market_share_mp*: medida relativa de la preferencia por marca propia frente al universo de marcas disponibles en la categoría, medida en gasto.
- *proporción_mp*: proporción de productos de marca propia respecto al total de productos comprados por el cliente.
- *recompra_mp*: índice de repetición de compra en marca propia.
- *frecuencia_mp*: número total de compras en productos de marca propia.

■ Para `score_tienda`:

- *gasto_total*: monto total gastado por el cliente en el supermercado.
- *frecuencia_tienda*: número de compras realizadas durante el período.
- *market_share_tienda*: participación del supermercado en el gasto total del cliente.
- *recencia_días*: número de días transcurridos desde la última compra registrada.

El modelo fue estimado mediante la librería `semopy` de Python, utilizando la clase `ModelMeans`, la cual permite incorporar directamente medias e interceptos en los modelos CFA. Los factores latentes obtenidos fueron luego estandarizados (`StandardScaler`) para facilitar su interpretación como variables continuas en modelos posteriores.

5.1.2. Modelo de Regresión Lineal

Posteriormente, se aplicó un modelo de regresión lineal, utilizando la librería `statsmodels`. La variable dependiente fue el `score_tienda` estandarizado, y como variable explicativa principal se consideró el `score_mp` estandarizado, que representa la lealtad hacia la marca propia.

Además, se incorporó un conjunto de variables adicionales de carácter sociodemográfico y conductual, seleccionadas por su potencial capacidad explicativa de la lealtad al supermercado:

- **promoción**: % de veces que el cliente compro un producto en proporción (incluyendo marca propia y otras marcas). Es un índice de sensibilidad al precio.
- **decil_comuna**: índice que representa el nivel de desarrollo económico, social y de infraestructura de la comuna de residencia del cliente. Toma valores entre 1 y 10, donde 1 corresponde a las comunas con menor nivel de desarrollo y 10 a las de mayor desarrollo. Si bien se relaciona con el nivel socioeconómico, su alcance es más amplio que una simple medición de ingreso. Los valores fueron obtenidos de Ministerio de Desarrollo Social (2013).

- **monto_liquido_total**: gasto total acumulado durante el período de análisis, incluye todos los gastos la tarjeta incluso aquellos que no son al supermercado. Esto quiere simular un mayor ingreso económico.
- **edad**: edad estimada del cliente.
- **numero_hijos**: número de hijos estimado.
- **soltero**: dummy que indica si el cliente es soltero (1), o casado (0).
- **genero_mujer**: dummy que indica si el cliente es mujer (1), u hombre (0).

Formalmente, el modelo puede expresarse como:

$$\begin{aligned} \text{score_tienda_std} = & \beta_0 + \beta_1 \cdot \text{score_mp_std} + \beta_2 \cdot \text{promocion} + \beta_3 \cdot \text{decil_comuna} \\ & + \beta_4 \cdot \text{monto_liquido_total} + \beta_5 \cdot \text{edad} + \beta_6 \cdot \text{numero_hijos} \\ & + \beta_7 \cdot \text{soltero} + \beta_8 \cdot \text{genero_mujer} + \varepsilon \end{aligned} \quad (1)$$

donde ε representa el término de error aleatorio, y cada coeficiente β estima el efecto parcial de su respectiva variable independiente sobre el score de lealtad al supermercado. Las variables continuas fueron previamente estandarizadas y las categóricas se incluyeron sin transformación adicional. Se eliminaron observaciones con valores faltantes o inválidos para asegurar la consistencia del modelo.

Este enfoque permite evaluar cuantitativamente el efecto de la lealtad hacia la marca propia y otras variables en la fidelidad global al supermercado, así como identificar factores significativos que expliquen los niveles de compromiso del cliente con el supermercado.

5.1.3. Scores latentes obtenidos

Tras la aplicación del modelo de análisis factorial confirmatorio (CFA) para la construcción de los índices latentes de lealtad a marca propia (**score_mp**) y lealtad al supermercado (**score_tienda**), se procedió a evaluar la calidad de dichos constructos mediante la estimación de indicadores de fiabilidad y ajuste.

En primer lugar, las **cargas factoriales** obtenidas fueron las siguientes, organizadas según la estructura especificada para cada score:

Variable	Valor	Std. Err	p-value
fidelidad_activa	1.0000	-	-
market_share_marca	0.8230	0.0007	0.0
proporción_mp	0.8359	0.0007	0.0
recompra	0.8382	0.0007	0.0
frecuencia_mp	-0.0037	0.0009	0.0
gasto_total	1.0000	-	-
frecuencia_tienda	0.8398	0.007	0.0
ms_tienda	0.2325	0.002	0.0
recencia_días	-0.2889	0.003	0.0

Cuadro 1: Estadísticas de scores de lealtad (lealtad a marca propia y al supermercado).

Las principales cargas factoriales presentan valores elevados y positivos, cercanos o superiores a 0.8, lo que indica una fuerte relación entre los indicadores más relevantes y sus respectivos constructos latentes. No obstante, la carga de frecuencia_mp resulta negativa y cercana a cero, lo que sugiere una contribución marginal o incluso perjudicial al factor de marca propia.

Índice	Valor	Interpretación
Composite Reliability (score_mp)	0.805	Aceptable
Composite Reliability (score_tienda)	0.514	Aceptable en contexto exploratorio
Harman's One-Factor	35.88 %	No hay sesgo significativo por método común

Cuadro 2: Índices de calidad del modelo.

En el Cuadro 2 se muestran los índices de calidad del modelo. El índice de fiabilidad compuesta (Composite Reliability) alcanzó un valor de 0.805 para `score_mp` y 0.514 para `score_tienda`, lo cual es considerado aceptable dentro de estudios de carácter exploratorio. Adicionalmente, el porcentaje de varianza explicada por el primer factor fue de 35.88 %, correspondiente al resultado del test de Harman's One-Factor. Aunque este valor no es particularmente elevado, se sitúa en un rango razonable para análisis de tipo exploratorio, especialmente en contextos donde los ítems presentan cierta heterogeneidad. Al no superar el umbral del 50 %, este resultado también indica ausencia de un sesgo significativo por método común.

En conjunto, estos indicadores sugieren que, al incorporar la segmentación por clústeres previa al ajuste, el modelo CFA logra capturar de forma más precisa la estructura latente de la lealtad. Por tanto, se utilizaron estos scores ajustados como insumo para la estimación de un nuevo modelo de regresión lineal múltiple.

5.1.4. Resultados Regresión Lineal

Se estimó un modelo de regresión lineal múltiple utilizando como variable dependiente el score estandarizado de lealtad al supermercado (`score_tienda_std`), obtenido a partir del análisis factorial confirmatorio (CFA). Como principal variable explicativa se incorporó el score estandarizado de lealtad a marca propia (`score_mp_std`), acompañado de variables socio demográficas y de comportamiento seleccionadas por su posible influencia en la fidelización.

El resumen de resultados del modelo OLS se presenta a continuación:

Variable	Coef.	Std. Err	p-valor
Constante	0.0588	0.005	0.000
score_mp_std	-0.0918	0.003	0.000
promoción	0.0055	0.003	0.051
decil_comuna	0.0074	0.003	0.010
monto_liquido_total	0.0110	0.003	0.000
soltero	-0.0588	0.004	0.000
edad	0.0243	0.004	0.000
numero_hijos	0.0700	0.007	0.000
genero_mujer	-0.0523	0.006	0.000
R-squared	0.0160	—	—
Adj. R-squared	0.0160	—	—
F-statistic	226.5	—	0.000
Durbin-Watson	1.964	—	—
Jarque-Bera	80,965.829	—	—
N (total base)	127,147		

Cuadro 3: Regresión lineal múltiple con `score_tienda_std` como variable dependiente.

El modelo fue estimado mediante mínimos cuadrados ordinarios (OLS), y sus resultados se presentan en el Cuadro 3. El coeficiente de determinación ajustado (*Adjusted R-squared*) fue de 0.016, lo que

indica que aproximadamente el 1.6 % de la variabilidad en la lealtad al supermercado es explicada por el conjunto de variables incluidas en el modelo. Este nivel de explicación refleja la complejidad inherente al comportamiento del consumidor, donde múltiples factores no observables pueden influir en las decisiones de compra.

La variable `score_mp_std` presentó un coeficiente negativo y altamente significativo (-0.0918 , $p < 0.000$), lo que sugiere que, en este conjunto de datos, una mayor lealtad a la marca propia se asocia con una menor lealtad al supermercado. Este resultado va en contra de la hipótesis inicial y plantea la posibilidad de que ciertos clientes valoren la marca propia como atributo independiente, sin necesidad de concentrar sus compras en una tienda específica. Este comportamiento ha sido evidenciado por Coelho do Vale et al. (2016), al mostrar que en supermercados de posicionamiento low-cost la lealtad a marca propia no necesariamente se traduce en una mayor fidelidad al establecimiento.

Además, se observaron relaciones significativas con otras variables del modelo:

- **promoción** (0.0055 , $p < 0.051$): es significativo al 90 %.
- **decil_comuna** (0.0074 , $p < 0.000$): los clientes de comunas con mayor nivel socioeconómico exhiben mayor lealtad.
- **monto_liquido_total** (0.0110 , $p < 0.000$): un mayor gasto total se asocia positivamente con la lealtad.
- **soltero** (-0.0588 , $p < 0.000$): Significativo y negativo, lo que indica que ser soltero disminuye la lealtad del supermercado.
- **edad** (0.0243 , $p < 0.000$): A mayor edad aumenta la lealtad al supermercado.
- **numero_hijos** (0.0700 , $p < 0.000$): Mayor cantidad de hijos indica una relación directa con la lealtad.
- **genero_mujer** (-0.0523 , $p < 0.000$): ser mujer está negativamente asociado con el score de lealtad al supermercado en este modelo.

Respecto a los supuestos del modelo, el estadístico de Durbin-Watson fue de 1.964, lo que indica ausencia de auto correlación grave entre los residuos. La prueba de Jarque-Bera ($80,965.829$, $p < 0.001$) sugiere que los residuos no siguen una distribución perfectamente normal, aunque esto es común en muestras de gran tamaño y no invalida la validez de los coeficientes estimados. Sin embargo, se recomienda interpretar con precaución los intervalos de confianza y los p-valores.

En conjunto, estos resultados permiten concluir que la lealtad a marca propia, las promociones, el gasto y ciertas características demográficas tienen un impacto estadísticamente significativo sobre la lealtad al supermercado, aunque no necesariamente en la dirección inicialmente esperada.

5.2. Limitaciones del modelo general y justificación del uso de clústeres

Aunque el modelo de regresión lineal múltiple arrojó resultados estadísticamente significativos y permitió identificar variables relevantes asociadas a la lealtad al supermercado, su poder explicativo fue limitado. El valor del coeficiente de determinación ajustado (*Adjusted R-squared*) fue de 0.016, lo que indica que existe una proporción considerable de la varianza que no está siendo capturada por el modelo.

Además, el coeficiente negativo del score de lealtad a marca propia sugiere una relación contraria a la esperada, lo que podría deberse a una heterogeneidad no modelada en la base de clientes. En otras palabras, distintos tipos de clientes pueden tener relaciones diferentes entre sus patrones de lealtad a marca propia y al supermercado, y el modelo lineal único no permite capturar esas diferencias.

En base a estas observaciones, se plantea la hipótesis de que los clientes podrían agruparse en segmentos relativamente homogéneos, cuyos comportamientos sean más consistentes dentro de cada grupo. Esto permitiría:

- Capturar patrones específicos de lealtad dentro de cada grupo.
- Mejorar la calidad predictiva del modelo al incorporar esta estructura latente.
- Comprender mejor los perfiles de cliente.

Por ello, se justifica la aplicación de una técnica de segmentación no supervisada, en particular K-means clustering, como paso intermedio antes de repetir el análisis factorial y la regresión dentro de grupos definidos.

5.3. Segmentación con categorías con presencia de marca propia pertenecientes a la Canasta Básica

En esta segmentación se trabajó exclusivamente con las 56 categorías que presentan presencia de marca propia y que forman parte de una canasta básica integral, compuesta por alimentos esenciales y productos de higiene personal. Es importante destacar que, dentro de cada categoría, no es posible distinguir entre distintas marcas propias del distribuidor. Por tanto, todos los productos de marca propia fueron tratados de forma agregada, limitándose a identificar su condición de marca del supermercado.

El análisis consideró el mismo periodo de tiempo definido para el estudio general (1 de julio de 2015 al 26 de agosto de 2016), manteniéndose también el criterio de inclusión de clientes con al menos cinco compras registradas durante ese lapso. Esto dio como resultado una base de **127,149** clientes únicos (ID).

Variables utilizadas para la segmentación

La segmentación se realizó exclusivamente en base a cinco variables cuantitativas de comportamiento:

- **frecuencia_mp**: número total de compras de productos de marca propia.
- **fidelidad_activa**: proporción de veces en que el cliente eligió marca propia cuando esta estuvo disponible.
- **frecuencia_tienda**: número total de compras realizadas en el supermercado durante el período.
- **proporción_mp**: proporción de productos de marca propia sobre el total de productos comprados.
- **canasta**: nueva variable calculada que representa la cantidad promedio de productos distintos que un cliente compra por transacción. Esta métrica se incorporó con el objetivo de capturar el tamaño relativo del carro de compra, permitiendo una caracterización más completa del comportamiento de los clientes.

Cabe destacar que el conjunto de variables seleccionadas captura dimensiones tanto de *lealtad a marca propia* como de *lealtad al supermercado*. Este enfoque mixto permite identificar perfiles integrales de clientes. Estas cinco variables fueron seleccionadas tras evaluar múltiples combinaciones posibles, priorizando aquellas que ofrecieran los mejores resultados en términos de calidad del agrupamiento. El resto de las variables disponibles no fueron utilizadas en el proceso de segmentación, pero sí se consideraron posteriormente para caracterizar los grupos obtenidos.

Selección del número de clústeres e índices de calidad

Para determinar el número óptimo de clústeres, se construyeron dos herramientas visuales complementarias: el gráfico del codo (Figura 2) y el gráfico de silueta (Figura 3). Ambos gráficos sugieren un punto de inflexión relevante en $k = 4$, respaldando la elección de este número como solución final.

La validez estructural del agrupamiento obtenido fue evaluada a través de tres métricas estándar:

- **Índice de silueta promedio:** 0.3123. Indica una separación razonable entre los clústeres, siendo valores por sobre 0.25 generalmente aceptables para segmentaciones iniciales.
- **Índice de Davies-Bouldin:** 1.0950. Refleja una baja superposición entre los clústeres, donde valores más cercanos a 0 indican mejor definición.
- **Índice de Calinski-Harabasz:** 46,104.35. Este valor elevado sugiere buena compactación interna dentro de cada grupo y una separación clara entre ellos.

En conjunto, estos indicadores respaldan la calidad del modelo de segmentación elegido y justifican la interpretación de los clústeres como representaciones válidas de distintos perfiles de cliente en relación con su comportamiento hacia la marca propia y sus patrones generales de compra en el supermercado.

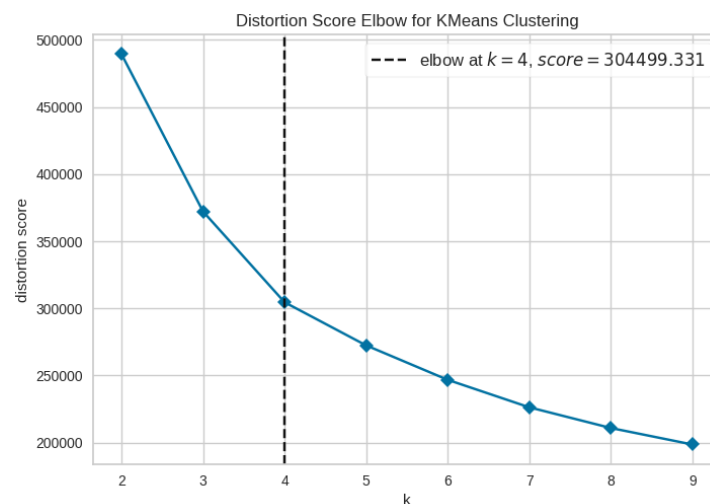


Figura 2: Gráfico del codo para determinar k en la canasta básica.

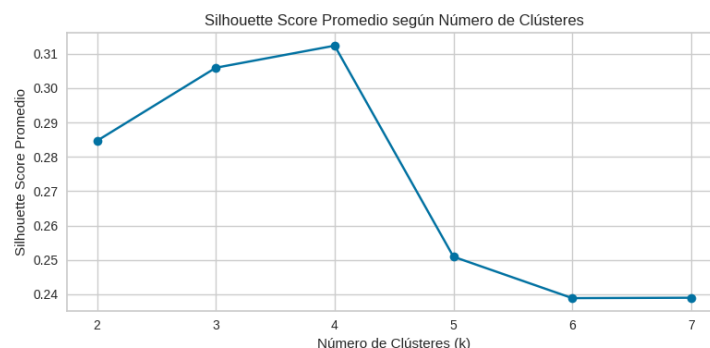


Figura 3: Gráfico de silueta para determinar k en la canasta básica.

A continuación, se presentan en los siguientes cuadros los valores promedio de las principales variables asociadas a la lealtad, calculadas para cada clúster.

Variable	Clúster 1	Clúster 2	Clúster 3	Clúster 4
Frecuencia MP (compras/año)	0.47 (0.20)	1.28 (0.49)	1.07 (0.51)	0.52 (0.29)
Fidelidad Activa (%)	10.74 (6.59)	14.97 (9.85)	14.48 (9.26)	35.48 (13.10)
Recompra (%)	13.27 (7.29)	15.91 (8.52)	16.18 (8.46)	29.66 (10.87)
Market Share MP (%)	12.61 (9.47)	14.99 (8.64)	15.16 (9.42)	32.55 (14.80)
Proporción MP (%)	10.38 (6.41)	13.48 (8.15)	13.30 (7.81)	31.99 (11.86)
Oferta (%)	34.24 (20.64)	29.84 (16.80)	34.01 (17.50)	26.66 (16.41)
Promoción (%)	37.35 (23.70)	32.10 (19.10)	37.51 (20.52)	41.55 (24.64)
Gasto Total (UF)	7.24 (3.74)	14.97 (6.13)	14.39 (6.50)	7.55 (4.28)
Frecuencia Tienda (compras/año)	8.14 (2.38)	8.08 (2.65)	17.48 (4.37)	8.63 (3.01)
Market Share Tienda (%)	39.60 (13.27)	44.26 (9.17)	42.45 (10.28)	41.14 (12.47)
Recencia (días)	84.94 (73.47)	82.95 (70.99)	58.64 (54.99)	83.64 (73.75)
Canasta Promedio	9.49 (3.98)	24.82 (7.96)	9.62 (3.88)	9.31 (4.83)
Tamaño (clientes)	65,246	18,949	17,973	24,981

Nota: Desviación Estándar en paréntesis

Cuadro 4: Promedios por variable según clúster

5.3.1. Evaluación de Resultados Categorías Canasta Básica Año 3

- **Clúster 1:** Es el grupo más numeroso, con 65,246 clientes, y se caracteriza por una baja lealtad tanto hacia la marca propia como hacia el supermercado. Presenta la menor proporción de compras en marca propia (10.38 %), una baja frecuencia de adquisición de estos productos (0.47) y un nivel reducido de fidelidad activa (10.74 %). Asimismo, exhibe valores bajos en recompra (13.27 %) y market share de marca propia (12.61 %). Se distingue, además, por su alta sensibilidad al precio, siendo el grupo con mayor exposición a ofertas (34.24 %) y promociones (37.34 %). Desde la perspectiva del canal, muestra un gasto total reducido (7.24 UF), una frecuencia de compra moderada (cada 8.14 días), y el market share en tienda más bajo entre todos los clústeres (39.60 %). Su recencia es la más elevada (84.94 días), lo que sugiere una baja actividad reciente.
- **Clúster 2:** Compuesto por 18,949 clientes, este grupo muestra una lealtad moderada hacia la marca propia y una relación sólida con el supermercado. Tiene la frecuencia de compra de marca propia más alta (1.28), junto con niveles relevantes de fidelidad activa (14.97 %) y market share hacia la marca del distribuidor (14.99 %). En términos de comportamiento general, exhibe un gasto total elevado (14.97 UF) y uno de los mayores niveles de market share en tienda (44.26 %). A pesar de realizar compras más espaciadas (frecuencia: 8.08 días; recencia: 82.95 días), estas son de gran

volumen, como lo demuestra su canasta promedio de 24.82 productos. Este clúster representa a clientes que, aunque no visitan frecuentemente la tienda, realizan compras de alto valor y mantienen una relación estable con el canal.

- **Clúster 3:** Integrado por 17,973 clientes, este grupo presenta niveles intermedios de lealtad a la marca propia (proporción de compras: 13.30 %; fidelidad activa: 14.48 %; market share MP: 15.16 %), pero destaca claramente por su fuerte vínculo con el supermercado. Es el grupo con mayor frecuencia de compra (cada 17.48 días), mayor gasto total (14.39 UF), y uno de los market share en tienda más altos (42.45 %). Además, presenta la recencia más baja del conjunto (58.64 días), lo que indica una actividad continua y reciente. A diferencia del Clúster 2, su canasta es mucho más pequeña (9.62 productos), lo que sugiere un patrón de compra frecuente y de bajo volumen. Su fidelización se expresa principalmente en la recurrencia, más que en la preferencia por marca propia.
- **Clúster 4:** Conformado por 24,981 clientes, este clúster representa el perfil más leal a la marca propia. Presenta los valores más altos en proporción de compras MP (31.99 %), fidelidad activa (35.48 %), recompra (29.66 %) y market share MP (32.55 %). En contraste, su comportamiento respecto al supermercado es más intermedio: gasto total bajo (7.55 UF), frecuencia de compra moderada (cada 8.63 días), y un market share en tienda relativamente alto (41.14 %). Su recencia (83.64 días) indica una participación reciente, aunque no intensiva. Este grupo refleja una fuerte preferencia por los productos de marca propia, aunque su conexión con la tienda no alcanza los niveles del Clúster 3.

Observación final: Los resultados de la segmentación permiten identificar dos grupos estratégicos complementarios. Por un lado, el Clúster 3 representa a clientes altamente leales al supermercado, aunque su vinculación con la marca propia es moderada. Por otro lado, el Clúster 4 reúne a los clientes más fieles a la marca propia, que mantienen una relación estable, aunque menos intensa, con la tienda. Esta dualidad sugiere la existencia de una relación bidireccional pero no lineal entre la lealtad a marca propia y la lealtad al supermercado.

5.4. Regresión Lineal con segmentación por clústeres y su interacción con lealtad a marca propia

Con el objetivo de capturar la heterogeneidad observada entre los clientes y mejorar la capacidad explicativa del modelo, se incorporó una etapa previa de segmentación utilizando el algoritmo *K-means clustering*. Esta segmentación agrupó a los clientes en cuatro clústeres diferenciados en función de variables relacionadas tanto con la lealtad a marca propia como al supermercado, lo que permitió identificar perfiles de comportamiento relativamente homogéneos.

La estimación de los scores latentes de lealtad a marca propia (`score_mp`) y lealtad al supermercado (`score_tienda`) se realizó mediante el mismo modelo de Análisis Factorial Confirmatorio (CFA) descrito en la sección anterior, utilizando la librería `semopy` en Python y estandarizando ambos scores con `StandardScaler` para su uso en modelos posteriores.

Posteriormente, se construyó un modelo de regresión lineal múltiple en el que se introdujeron:

- Variables dummy para los clústeres 2, 3 y 4 (dejando al Clúster 1 como categoría base).
- El `score_mp_std`, es decir, la lealtad a marca propia estandarizada.
- Variables de interacción entre cada dummy de clúster y el `score_mp_std`, con el objetivo de capturar diferencias en el efecto del score de marca propia según el tipo de cliente.

- Variables socio demográficas y conductuales, explicadas anteriormente.

Esta especificación permite observar:

- Cambios en el intercepto del modelo según el clúster al que pertenece el cliente, reflejando diferencias de base en su lealtad al supermercado.
- Diferencias en la pendiente del efecto del `score_mp_std`, lo cual indica si el impacto de la lealtad a marca propia sobre la lealtad al supermercado es más fuerte, más débil o incluso inverso dependiendo del segmento.

Este enfoque permite una evaluación más precisa de la relación entre los constructos de lealtad, reconociendo que dicha relación no es homogénea entre todos los clientes, y que distintas estrategias de fidelización podrían ser necesarias para distintos perfiles.

5.4.1. Resultados Regresión Lineal con clústeres

Con el objetivo de mejorar la estimación, se re estimó el modelo de regresión lineal múltiple empleando los nuevos nuevos componentes. Quedando el siguiente modelo:

$$\begin{aligned} \text{score_tienda_std} = & \beta_0 + \beta_1 \cdot \text{score_mp_std} + \beta_2 \cdot \text{promocion} + \beta_3 \cdot \text{decil_comuna} + \beta_4 \cdot \text{monto_liquido_total} \\ & + \beta_5 \cdot \text{edad} + \beta_6 \cdot \text{numero_hijos} + \beta_7 \cdot \text{soltero} + \beta_8 \cdot \text{genero_mujer} \\ & + \beta_9 \cdot \text{Cluster2} + \beta_{10} \cdot \text{Cluster3} + \beta_{11} \cdot \text{Cluster4} \\ & + \beta_{12} \cdot (\text{Cluster2} \cdot \text{score_mp_std}) + \beta_{13} \cdot (\text{Cluster3} \cdot \text{score_mp_std}) \\ & + \beta_{14} \cdot (\text{Cluster4} \cdot \text{score_mp_std}) + \varepsilon \end{aligned} \quad (2)$$

Mediante este modelo, se obtuvieron los resultados actualizados:

Variable	Coef.	Std. Err	p-valor
Constante	-0.3750	0.005	0.000
score_mp_std	-0.1383	0.005	0.000
promocion	0.0326	0.002	0.000
decil_comuna	0.0016	0.002	0.430
monto_liquido_total	0.0399	0.002	0.000
edad	-0.0658	0.003	0.000
numero_hijos	0.0598	0.003	0.000
soltero	-0.0603	0.005	0.000
genero_mujer	-0.0605	0.004	0.000
Clúster 2	1.1817	0.007	0.000
Clúster 3	1.5326	0.007	0.000
Clúster 4	0.1427	0.009	0.000
Clúster2 × score_mp	0.0679	0.013	0.000
Clúster3 × score_mp	0.1660	0.014	0.000
Clúster4 × score_mp	0.0653	0.007	0.000
R-squared	0.401	—	—
Adj. R-squared	0.401	—	—
F-statistic	5933.0	—	0.000
Durbin-Watson	1.999	—	—
Jarque-Bera	114,678.264	—	—
N (total base)	127,147		

Cuadro 5: Resultados del modelo con interacción entre `score_mp_std` y clúster.

El modelo de regresión lineal múltiple con segmentación por clústeres e interacciones presentó una mejora sustancial en su capacidad explicativa en comparación con el modelo general. El coeficiente de determinación ajustado (*Adjusted R-squared*) aumentó de 0.016 a **0.401**, lo cual indica que el nuevo modelo explica aproximadamente un 40 % de la variabilidad observada en la lealtad al supermercado (`score_tienda_std`). Este incremento sugiere que la incorporación de los clústeres de clientes y sus respectivas interacciones con el score de lealtad a marca propia permite capturar patrones heterogéneos relevantes que el modelo anterior no lograba identificar.

En este modelo, el coeficiente del score de marca propia (`score_mp_std`) fue de -0.1383 y resultó altamente significativo ($p < 0.000$), indicando nuevamente que un mayor nivel de lealtad a marca propia se asocia con una menor lealtad al supermercado, replicando el patrón del modelo sin segmentación.

No obstante, las interacciones permiten observar cómo este efecto varía en los otros grupos:

- En el **Clúster 2**, la interacción fue de 1.1817 y , efecto positivo. Lo que significa que pertenecer al clúster 2 es ser más leal al supermercado. Esto se observa en el análisis de clústeres. Además su interacción con el `score_mp` es positivo, lo que significa que a medida que aumenta la lealtad a marca propia en el clúster 2, también aumenta su lealtad al supermercado. todo esto respecto al clúster 1.
- En el **Clúster 3**, la interacción fue de 1.5326, efecto positivo y significativo ($p < 0.000$). Lo cual indica que respecto al clúster 1, el clúster 3 tiene una mayor lealtad al supermercado. Además, su interacción con `score_mp`, también es positiva y significativa. Siendo la más grande, lo que significa que son que más leales a la tienda se vuelven ante un aumento de lealtad a marca propia.
- En el **Clúster 4**, la interacción fue de 0.1427, nuevamente positiva y significativa. Recordemos este es el clúster más leal a marca propia, pero tiene una lealtad moderada al supermercado. Respecto a su interacción con `score_mp`, también es positiva y significativa, pero menor al clúster 3. Esto significa que si siguen aumentando su lealtad a marca propia, el aumento de lealtad al supermercado no se espera que sea grande.

Estos resultados muestran que la relación entre la lealtad a marca propia y la lealtad al supermercado es negativa, pero tiene distintos comportamientos en cada clúster.

Además, las variables socio demográficas y de comportamiento mantuvieron patrones similares a los observados en el modelo anterior. Variables como `promoción`, `monto_liquido_total` y `numero_hijos` siguen mostrando una relación positiva y significativa con la lealtad, mientras que `edad`, `soltero` y `genero_mujer` presentan coeficientes negativos. La única variable no significativa es `decil_comuna`.

En conjunto, esta especificación más flexible permite una comprensión más profunda de los factores que inciden en la lealtad al supermercado y demuestra empíricamente la utilidad de segmentar a los clientes para obtener modelos predictivos más precisos y accionables.

5.4.2. Verificación de distintas regresiones dependiendo de los clústeres

Con el objetivo de validar la coherencia de los resultados obtenidos en el modelo con interacción por clústeres, se estimaron nuevamente los modelos de regresión lineal múltiple, variando la categoría base de referencia del clúster. Se repitió la estimación considerando como base, respectivamente, los clústeres 2, 3 y 4. En todos los casos, la variable `score_mp_std` mantuvo un coeficiente negativo y estadísticamente

significativo, lo que refuerza la solidez de la relación inversa entre la lealtad a marca propia y la lealtad al supermercado en el conjunto general.

Este procedimiento permitió evaluar la consistencia de los coeficientes asociados tanto a los clústeres como a sus interacciones con la variable `score_mp_std`. Los resultados fueron coherentes con el comportamiento esperado: el Clúster 1, caracterizado como el grupo menos leal al supermercado, presentó sistemáticamente coeficientes negativos en todas las comparaciones y en las interacción con `score_mp_std`. En contraste, el Clúster 3, identificado como el más leal al supermercado, mostró coeficientes positivos en todas las especificaciones, lo cual respalda la interpretación previa sobre la orientación de estos segmentos.

Adicionalmente, se ejecutaron regresiones lineales múltiples separadas para cada clúster de forma individual. Esta estrategia buscó analizar en detalle el efecto de `score_mp_std` y otras variables explicativas dentro de cada segmento homogéneo de clientes. Los resultados fueron los siguientes:

- **Clúster 1:** `score_mp_std` presentó un coeficiente negativo y significativo, en línea con su perfil de baja lealtad al supermercado. A diferencia de otros clústeres, la variable `decil_comuna` tuvo un efecto negativo sobre la lealtad.
- **Clúster 2:** Se observó un coeficiente positivo para `score_mp_std`, aunque no significativo. Este grupo muestra una lealtad moderada al supermercado.
- **Clúster 3:** El coeficiente de `score_mp_std` fue positivo y significativo, reflejando la alta lealtad de este segmento hacia el supermercado.
- **Clúster 4:** También se obtuvo un coeficiente positivo y significativo para `score_mp_std`, confirmando su fuerte compromiso con la marca propia y una relación estable con el supermercado. Sin embargo, `decil_comuna` mostró nuevamente un efecto negativo.

En todos los clústeres, las variables `promoción`, `monto_liquido_total` y `numero_hijos` mantuvieron una relación positiva y significativa con la lealtad al supermercado, mientras que `edad`, `soltero` y `genero_mujer` mostraron asociaciones negativas y también significativas.

Es importante señalar que todos los modelos por clúster individual presentaron valores bajos de *Adjusted R-squared*, lo cual era esperado debido al menor tamaño muestral utilizado en cada segmentación. No obstante, este análisis adicional confirma la coherencia interna de los datos y respalda la validez de los resultados obtenidos mediante la segmentación por clústeres.

5.5. Análisis específico por categoría

La decisión de replicar el análisis de segmentación y regresión lineal para una categoría específica de productos responde a la necesidad de capturar con mayor precisión los factores que influyen en la lealtad hacia la marca propia dentro de un entorno competitivo más homogéneo. Los clientes no necesariamente mantienen un comportamiento uniforme de lealtad hacia la marca propia en todas las categorías: pueden ser fieles, por ejemplo, a las harinas de marca propia por percibir una buena relación calidad/precio, pero preferir productos de marca nacional en categorías como detergentes o lácteos. Esto se debe, en parte, a que la marca propia puede proyectar una imagen distinta según la categoría, siendo percibida como una alternativa de bajo costo en algunos casos y como una opción competitiva en otros. Por ello, al mezclar múltiples categorías con atributos muy diversos, se pierde capacidad de atribución clara sobre qué impulsa realmente la lealtad. Focalizar el análisis en una sola categoría permite reducir esta heterogeneidad y generar conclusiones más específicas y accionables para la gestión comercial.

Para el análisis que se presenta a continuación, se definieron cuatro grandes grupos dentro del nivel “Tipo de Producto”: PGC Comestible, PGC No Comestible, FLC, y Panadería y Pastelería. A partir

de estos grupos, se seleccionaron 20 categorías específicas que contienen productos de marca propia, priorizando aquellas con mayor relevancia comercial y con potencial analítico en términos de segmentación y visualización de resultados. Las categorías restantes se incluyen en el Anexo, lo que permite complementar el estudio con una visión más amplia del comportamiento de la marca propia en el resto del portafolio.

En esta etapa, se realiza una evaluación preliminar de las categorías en las que la marca propia tiene presencia significativa, con el objetivo de identificar patrones y áreas clave que contribuyan a un análisis más detallado de la lealtad. Para ello, se utilizan los datos completos disponibles durante todo el período de observación, procesados mediante consultas optimizadas en Amazon Athena (SQL), lo que garantiza eficiencia, escalabilidad y trazabilidad en el análisis. Posteriormente, se generarán visualizaciones para facilitar la interpretación de los resultados.

Finalmente, es importante destacar que la categoría “Harina”, aunque originalmente clasificada dentro del grupo “Fécula”, ha sido analizada por separado debido a su alta penetración, frecuencia de compra y relevancia estratégica. Esta decisión permite evaluar con mayor precisión los determinantes de la lealtad en una categoría clave y aislar los efectos que podrían diluirse en un análisis más agregado.

5.5.1. Análisis de Categoría

En primer lugar, se realizará un análisis preliminar para identificar cuáles son las categorías más interesantes de estudiar. Este análisis inicial se enfoca en la frecuencia y la preferencia promedio por SKU, permitiendo evaluar la relevancia de cada categoría en su totalidad, sin distinguir aún entre productos de marca propia y otras marcas. La finalidad es obtener una visión general sobre el comportamiento de compra en cada categoría, lo que ayudará a seleccionar aquellas que presentan características más atractivas para un análisis posterior centrado en la marca propia.

Frecuencia

El gráfico de *frecuencia de compra* muestra la cantidad promedio de compras anuales por cliente en cada categoría. Para calcular esta métrica, se contabilizó cuántas veces cada cliente compró en una categoría específica durante cada año del período analizado.

Esta información permite identificar la rotación de productos: categorías con mayor frecuencia reflejan un consumo más recurrente, mientras que aquellas con menor frecuencia corresponden a compras más esporádicas o de baja rotación.

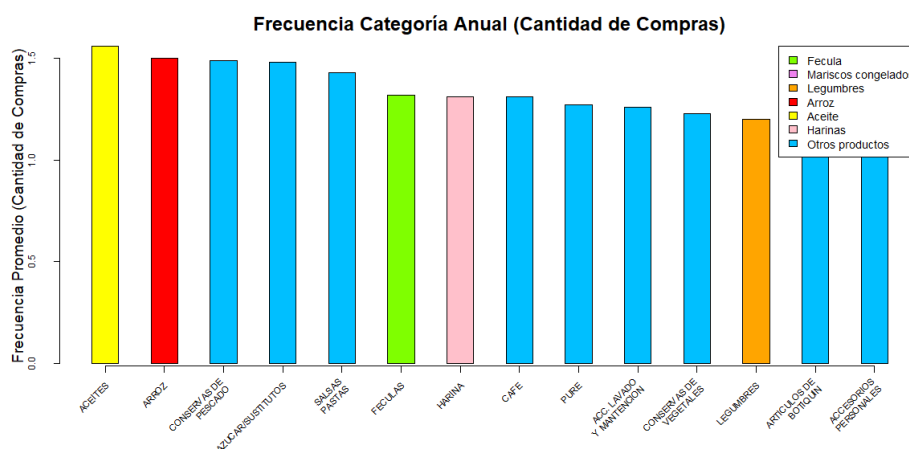


Figura 4: Frecuencia en Días en cada Categoría.

Los valores observados se encuentran mayoritariamente cercanos a 1, lo que refleja que, en promedio, los clientes realizan una única compra anual en muchas de estas categorías.

Pese a esto, se observan diferencias relativas entre categorías. **Aceites**, **arroz** y *conservas de pescado* son las que muestran las mayores frecuencias, superando levemente las 1.5 compras anuales, lo que sugiere que, aunque el promedio es bajo, estas categorías cuentan con una base más estable de consumidores recurrentes. Otras categorías como **féculas**, **harina** y *café* se ubican en rangos intermedios, mientras que productos como **legumbres**, *accesorios personales* y *artículos de botiquín* muestran las frecuencias más bajas, lo que indica un consumo más esporádico o puntual.

Preferencia Promedio de SKU

La Figura 5 muestra la *preferencia promedio de SKU*, que indica cuántas veces, en promedio, un cliente repite la compra del mismo producto (SKU) en sus últimas 10 compras dentro de una categoría. Este indicador permite identificar si existe una fuerte inclinación hacia un producto específico, sin importar la marca. Valores altos sugieren una preferencia marcada por un mismo SKU, mientras que valores bajos reflejan mayor variedad en las elecciones del consumidor.

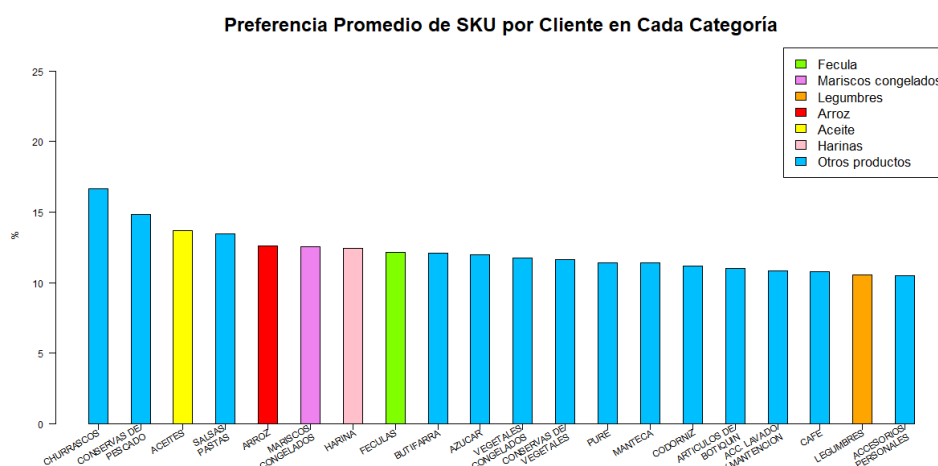


Figura 5: Preferencia Promedio de SKU por Cliente en Cada Categoría.

En la categoría de **aceites**, la preferencia promedio de SKU alcanza un 13.66 %, lo que refleja cierta repetición en la compra de un mismo producto, aunque los clientes aún exploran otras opciones.

Mariscos congelados, **arroz** y **harina** presentan valores similares (12.55 %, 12.58 % y 12.41 %), lo que sugiere una leve tendencia hacia la repetición, combinada con diversidad en las elecciones.

En **féculas**, la preferencia promedio es de 12.15 %, indicando mayor variabilidad y ausencia de un SKU dominante, posiblemente por la escasa diferenciación entre productos.

5.5.2. Análisis de Marca Propia en las Categorías

Para evaluar el desempeño de la marca propia en el contexto del supermercado, se han definido y aplicado cuatro métricas principales. Estas métricas permiten evaluar el rendimiento de la marca propia desde diferentes perspectivas. El análisis considera cuatro métricas principales: *Market Share en Gasto*, *Proporción de Compra Mensual*, *Cobertura SKU*, y *Relación de Precio*. Este enfoque integral proporciona una visión detallada del desempeño de la marca propia en el supermercado, permitiendo identificar no solo su participación en ventas, la diversidad de la oferta disponible y cómo su precio se compara con el de otras marcas.

Market Share en Gasto

El *market share en gasto* es clave para evaluar la rentabilidad e impacto financiero de la marca propia, al indicar qué proporción del dinero total gastado en una categoría corresponde a sus productos. Estudios como Bucklin et al. (1995) lo han utilizado como indicador central de la relevancia de una marca. Aplicado a categorías, permite estimar su peso económico dentro del mercado.

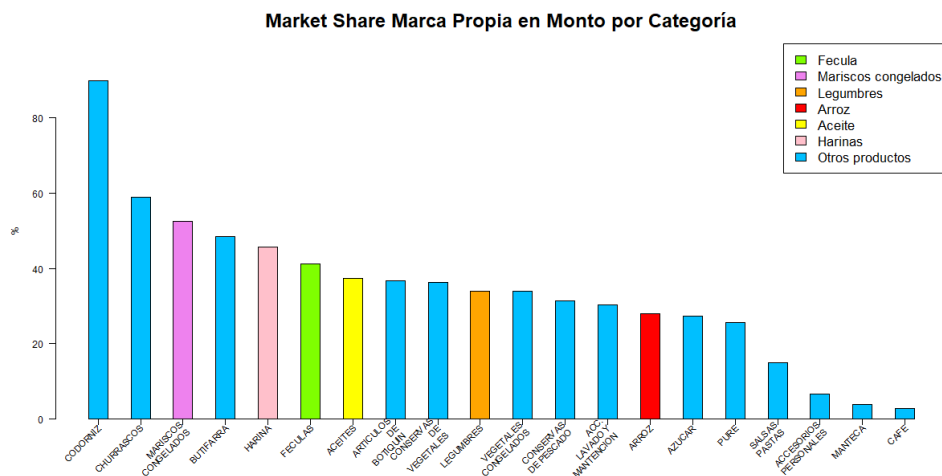


Figura 6: Market Share de Marca Propia en Gasto por Categoría.

Mariscos congelados lidera esta métrica con un *market share* en gasto del 52.57 %, lo que refleja una fuerte preferencia por los productos de marca propia en términos de valor económico.

Le siguen **harina** (45.69 %) y **fécula** (41.28 %), ambas con una presencia destacada de marca propia, lo que indica una alta penetración en estas categorías.

Aceite alcanza un 37.30 %, mostrando también una buena participación, aunque inferior en comparación con otras categorías. Esto podría deberse a una mayor competencia o a un menor precio promedio de los productos de marca propia.

Finalmente, las categorías de **legumbres** y **arroz** registran participaciones de 34.00 % y 27.90 %, respectivamente, lo que sugiere una presencia significativa pero más moderada de la marca propia.

Proporción de Compra Mensual

La *proporción de compra mensual* mide qué parte de las compras en una categoría corresponde a productos de marca propia. Se calcula mediante una variable binaria (1 si el cliente compró marca propia en un mes, 0 si no), promediada primero a nivel mensual por cliente y luego a nivel general. Esta métrica refleja la preferencia sostenida por marca propia a lo largo del tiempo.

Un valor alto, por ejemplo, 0.8 indica que el 80 % de las compras mensuales en la categoría son de marca propia, lo que señala una alta fidelidad en la elección.

Se eligió esta métrica por su capacidad de estandarizar comparaciones entre categorías con formatos o unidades de venta muy distintas, donde contar unidades compradas podría ser engañoso. Así, permite un análisis coherente y transversal de la lealtad hacia la marca propia.

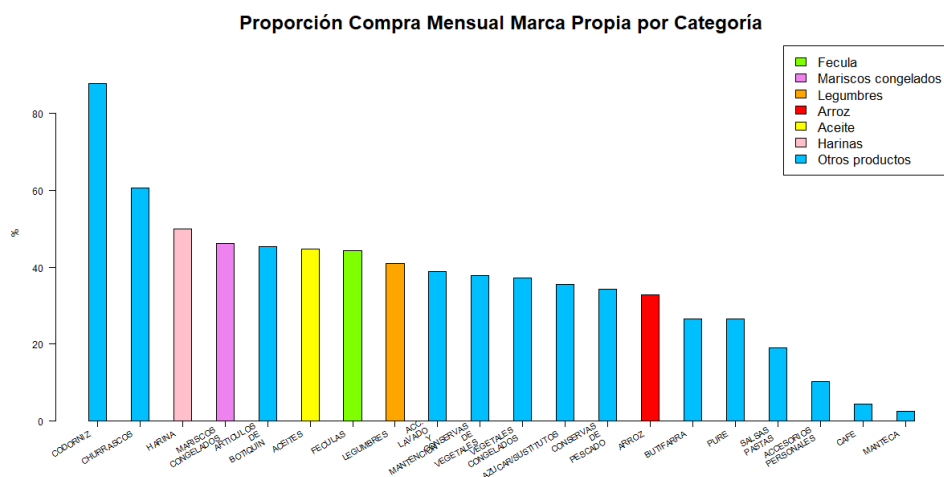


Figura 7: Proporción Mensual Marca Propia en un Mes por Categoría.

En la Figura 7 se observa que las categorías de **harina** (49.98 %) y **mariscos congelados** (46.25 %) presentan las proporciones más altas de compra mensual de marca propia, lo que refleja una fuerte preferencia por estos productos, aunque con algo de competencia de otras marcas.

Otras categorías como **aceites** (44.72 %), **féculas** (44.34 %) y **legumbres** (41.00 %) muestran una participación equilibrada, donde la marca propia mantiene una presencia relevante pero no dominante.

En el caso del **arroz** (32.81 %), la participación es menor, lo que indica una mayor preferencia por marcas competidoras.

En conjunto, estos resultados sugieren que en categorías esenciales como harina o aceites, la marca propia alcanza altos niveles de fidelización, mientras que en categorías más competitivas enfrenta mayor fragmentación del consumo.

Cobertura SKU

La *cobertura SKU* representa la proporción de SKUs de marca propia respecto al total de SKUs en una categoría, lo que permite evaluar la diversidad de productos ofrecidos por la marca propia. Un valor alto refleja una amplia presencia en términos de variedad.

A diferencia del *market share* en gasto, que mide el desempeño en ventas, la cobertura SKU analiza la amplitud de la oferta. Una alta cobertura puede respaldar un buen market share, al ofrecer múltiples opciones al consumidor. Por el contrario, una baja cobertura puede indicar una estrategia enfocada en pocos productos con alto rendimiento.

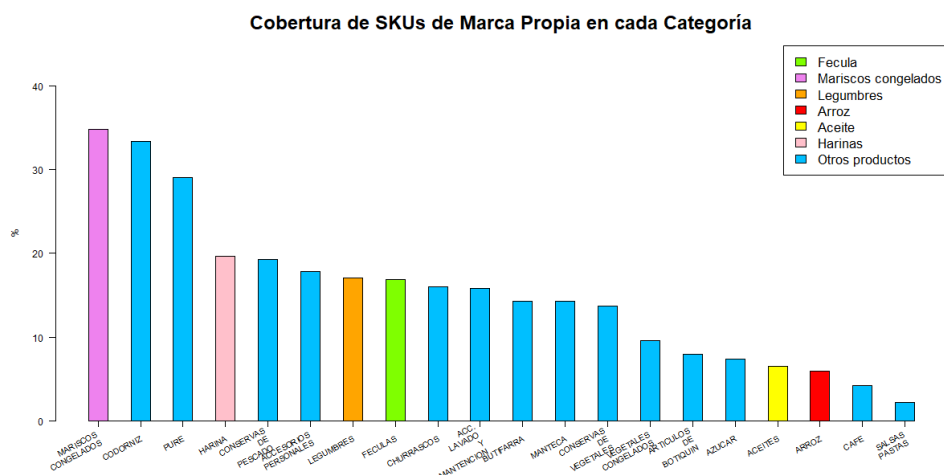


Figura 8: Cobertura de SKUs de Marca Propia en cada Categoría.

En la Figura 8 se presenta la cobertura de SKUs de marca propia en cada categoría. La categoría de **mariscos congelados** destaca con una cobertura del 34.83 %, lo que indica que más de un tercio de los productos disponibles son de marca propia. Esta amplia oferta está en línea con su alto market share en gasto y su elevada proporción de compra mensual, lo que sugiere una estrategia sólida basada en variedad.

En cambio, categorías como **aceites** (6.49 %), **arroz** (5.92 %) y **fécula** (19.64 %) muestran una cobertura significativamente menor. A pesar de ello, estas categorías exhiben buenos resultados en market share y frecuencia de compra, lo que indica que, aunque la variedad de productos es limitada, los SKUs disponibles logran captar una proporción relevante de las ventas. Por ejemplo, **aceite**, con solo un 6.49 % de cobertura, alcanza un market share en gasto del 37.30 % y una proporción de compra mensual del 44.72 %, demostrando un alto rendimiento con una oferta acotada.

Relación Precio

La métrica de *relación precio* compara el precio promedio de los productos de marca propia con el precio promedio general de cada categoría. Permite identificar si los productos de marca propia son más baratos o más caros que el resto. Un valor positivo indica que tienen un precio superior al promedio de la categoría; uno negativo, que son más asequibles.

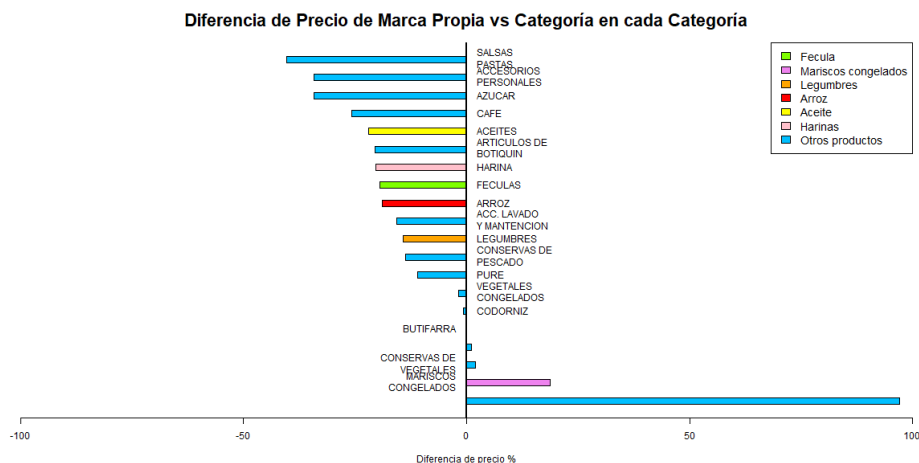


Figura 9: Diferencia de Precios de Marca Propia en Cada Categoría.

La Figura 9 muestra que en categorías como **mariscos congelados**, los productos de marca propia son un 18.71 % más caros que el promedio de la categoría. A pesar de ello, mantienen una alta cobertura SKU (34.83 %) y un destacado market share en gasto (52.57 %), lo que sugiere que los consumidores perciben un valor diferencial que justifica el mayor precio.

En contraste, los **aceites** presentan una diferencia de -21.99 %, con precios de marca propia significativamente más bajos que el promedio. Esta estrategia de precios más accesibles, junto a una alta proporción de compra mensual (44.72 %), parece compensar su baja cobertura SKU (6.49 %), logrando una fuerte presencia en la categoría.

Las categorías de **fécula** (-19.44 %) y **harina** (-20.29 %) también muestran precios por debajo del promedio. En el caso de **harina**, este leve descuento, combinado con una cobertura SKU del 19.64 % y un market share en gasto del 45.69 %, contribuye a un buen rendimiento en ventas.

En conjunto, estos resultados evidencian que una relación de precios favorable no es el único factor determinante: la variedad de oferta (cobertura SKU) y la frecuencia de compra también influyen significativamente en la competitividad de la marca propia.

5.5.3. Resumen

El Cuadro 1 presenta un análisis comparativo de diversas métricas clave para seis categorías de productos de marca propia: **aceite**, **arroz**, **fécula**, **harina**, **legumbres** y **mariscos congelados**. Estas métricas permiten evaluar el comportamiento de la marca propia en términos de participación de mercado, proporción de compra, oferta de productos y nivel de precios.

Categoría	Market Share en Gasto	Proporción de Compra Mensual	Cobertura SKU	Relación Precio
Aceite	37.30 %	44.72 %	6.49 %	-21.99 %
Arroz	27.90 %	32.81 %	5.92 %	-18.95 %
Fécula	41.28 %	44.34 %	16.85 %	-19.44 %
Harina	45.69 %	49.98 %	19.64 %	-20.29 %
Legumbres	34.00 %	41.00 %	17.11 %	-14.19 %
Mariscos Congelados	52.57 %	46.25 %	34.83 %	18.71 %
Valor Máximo del Parámetro	89.72 %	87.65 %	50 %	131.67 %
Valor Mínimo del Parámetro	0.00974 %	0.0137 %	0.11 %	-73.76 %

Cuadro 6: Resumen de las categorías destacadas y los diferentes parámetros analizados, junto a la referencia del máximo y mínimo de cada parámetro.

A partir del análisis comparativo presentado en el Cuadro 6, se concluye que la categoría **harina** es la mejor opción para un estudio más detallado de la marca propia debido a varias razones clave. En primer lugar, el *market share en gasto* de harina es del 45.69 %, lo que indica una participación significativa de los productos de marca propia en el valor total de ventas de esta categoría. Este alto porcentaje sugiere que los consumidores no solo compran productos de marca propia, sino que lo hacen de manera considerable en términos económicos.

Además, la *proporción de compra mensual* en harina es del 49.98 %, lo que refleja que casi la mitad de las compras mensuales en esta categoría corresponden a productos de marca propia. Este dato demuestra una fuerte preferencia y recurrencia en la compra de estos productos. Por otro lado, aunque la *cobertura SKU* es del 19.64 %, lo que es moderado en comparación con otras categorías, sigue siendo un nivel importante que asegura una buena variedad de productos de marca propia disponibles, lo que favorece el desempeño en ventas.

Finalmente, la *relación de precio* en harina es de -20.29 %, lo que significa que los productos de marca propia son ligeramente más económicos que el promedio de la categoría. Aunque esta diferencia es pequeña, puede influir positivamente en la decisión de los consumidores y contribuir al alto market share y a la preferencia mensual observada.

En conjunto, la categoría de harina destaca por su fuerte participación en el mercado, su preferencia por parte de los consumidores y su ventaja en términos de precios, lo que la convierte en una opción ideal para un análisis más profundo de la marca propia en el supermercado.

5.6. Segmentación Categoría Harina

El presente análisis se enfoca exclusivamente en la categoría *harina*, con el objetivo de estudiar de forma más detallada el comportamiento de los clientes frente a las marcas propias en un contexto específico. En esta categoría se identificaron 13 marcas distintas, entre ellas una clasificada como “desconocida”, debido a que no se pudo identificar el nombre de la marca. Dos de las marcas detectadas corresponden a marcas propias del supermercado.

Para mantener la generalidad del análisis y evitar sesgos asociados al reconocimiento comercial, las marcas propias fueron renombradas como **M1** y **M2**. La marca **M1** corresponde a la que no contiene

el nombre del supermercado en Chile, por lo que no siempre es percibida como marca propia por los consumidores. En cambio, **M2** sí está directamente asociada al nombre del supermercado, lo que refuerza su percepción como marca propia. Esta distinción es clave para evaluar la relación entre percepción de marca y lealtad de cliente.

En comparación con otras categorías, en la categoría *harina* se dispone de información adicional que permite un análisis más profundo. En particular, se cuenta con el detalle de las marcas asociadas a cada producto y sus respectivos precios y kilos, lo cual habilita la construcción de métricas más precisas sobre comportamiento de compra, como la sensibilidad al precio y la multilealtad. Esta información granular permite capturar con mayor fidelidad las preferencias individuales de los consumidores frente a las distintas opciones disponibles dentro de la categoría.

Variables de Lealtad a Marca Propia

Dado que en esta categoría se cuenta con información más detallada que en el análisis general, se incorporaron nuevas variables y se adaptaron otras previamente utilizadas para profundizar el estudio de lealtad a marca propia. A continuación se describen las variables incluidas en el modelo de segmentación:

- **Frecuencia:** Mide la cantidad total de compras realizadas por cada cliente (RUT) en una marca determinada (M1, M2 o marcas no propias). Se calculó como la suma de compras por cliente durante el período de análisis. Esta variable permite observar el nivel de recurrencia y es un indicador de lealtad conductual. Su importancia está respaldada por el marco teórico de Møller Jensen (2011).
- **Market Share en Gasto:** Se calculó por separado para M1, M2 y para ambas marcas propias en conjunto, expresando la proporción del gasto total del cliente en harina que se destinó a cada marca. Esto permite evaluar el nivel de apoyo del cliente hacia cada marca por separado y hacia la oferta global de marca propia.
- **Recompra:** Mide el porcentaje de clientes que realizaron compras repetidas en M1 o M2. Se calcula de forma individual para cada marca y también de forma conjunta, considerando ambas marcas como una única entidad de marca propia. Esto permite evaluar fidelidad diferenciada y también la percepción consolidada de la marca del supermercado.
- **Multilealtad:** Representa el posicionamiento relativo de las marcas propias dentro de la cartera de marcas de cada cliente. Se basa en el ranking de market share en gasto de cada marca dentro del conjunto de marcas compradas por el cliente. Se evaluó por separado para M1 y M2, y también de forma agregada. Esta variable está fundamentada en el estudio de Knox and Walker (2001) y permite identificar si las marcas propias ocupan posiciones destacadas dentro de las preferencias del cliente.
- **Oferta:** Proporción de compras en las que el cliente eligió productos en descuento de marcas no propias, sobre el total de sus transacciones en las categorías analizadas. Indica sensibilidad al precio y disposición a cambiar de la marca propia ante promociones externas.
- **Sensibilidad al Precio:** Mide la proporción de compras en las que el cliente eligió el producto de harina con el precio por kilo más bajo disponible. Para su cálculo, se determinó primero el precio por kilo de cada producto y se agruparon según su formato (1 kg, 3 kg, etc.). Luego, por cada día, local y formato, se identificó el producto más barato dentro de cada grupo. En cada transacción, se verificó si el cliente eligió esa opción más económica; si fue así, se asignó un valor de 1, y en caso contrario, un 0. Finalmente, se calculó el porcentaje de compras en las que el cliente optó por la opción más barata sobre el total de sus transacciones en la categoría. Un valor alto indica

una alta sensibilidad al precio, mientras que uno bajo sugiere menor preocupación por el costo y posiblemente mayor lealtad hacia ciertas marcas.

- **Índice de Apoyo a Marca Propia:** A diferencia de las variables anteriores, esta se calculó considerando todas las categorías del supermercado donde está presente la marca propia. Representa el porcentaje promedio de gasto en productos de marca propia respecto al gasto total del cliente en dichas categorías. Esta métrica ofrece una visión más amplia del compromiso financiero del cliente con la marca propia del supermercado y se basa en el enfoque de Knox and Walker (2001). Para este cálculo, las marcas M1 y M2 se trataron de forma conjunta.
- **Gasto Total:** Corresponde a la suma total del monto líquido desembolsado por cada cliente (RUT) en el supermercado durante el período analizado. Aunque no es un indicador directo de lealtad, se utiliza en el modelo de segmentación para contextualizar el comportamiento de compra según la magnitud del gasto.

Consideraciones Metodológicas

En esta segmentación solo se incluyeron en el modelo variables asociadas a la lealtad hacia la marca propia. Las variables que miden la lealtad hacia el supermercado, como frecuencia de visita o recencia en otras categorías, fueron reservadas para la etapa posterior de caracterización de los clústeres resultantes. Esto permite analizar si los segmentos identificados por su comportamiento hacia la marca propia también presentan patrones consistentes de fidelidad general hacia la cadena.

Las variables utilizadas en segmentaciones anteriores como *Proporción de productos marca propia* y *Canasta promedio* fueron descartadas en esta etapa, ya que las nuevas variables incorporadas permiten capturar mejor los patrones de comportamiento relevantes.

5.6.1. Criterios de segmentación y evaluación del modelo

La segmentación se aplicó únicamente a aquellos clientes (ID) que realizaron al menos una compra en la categoría de harinas durante el período comprendido entre el 1 de julio de 2015 y el 26 de agosto de 2016, lo que resultó en un total de **77,874** clientes únicos.

Para determinar el número óptimo de clústeres, se utilizó el gráfico del codo, como se muestra en la Figura 10. Aunque visualmente $k = 3$ parecía una opción evidente, este modelo no generó agrupaciones lo suficientemente diferenciadas ni relevantes desde el punto de vista del comportamiento de lealtad. Finalmente, se optó por presentar los resultados con $k = 6$, ya que este valor ofrecía una segmentación más rica y analíticamente significativa, con perfiles conductuales claramente distintos y pertinentes. Esta decisión se ve respaldada por los indicadores de calidad del agrupamiento obtenidos: el índice de silueta promedio fue de 0.4700, lo que refleja una buena separación entre grupos; el índice de Davies-Bouldin alcanzó un valor de 1.1633, indicando baja superposición entre los clústeres; y el índice de Calinski-Harabasz fue de 65,875.42, lo que sugiere una alta compacidad interna y buena dispersión externa.

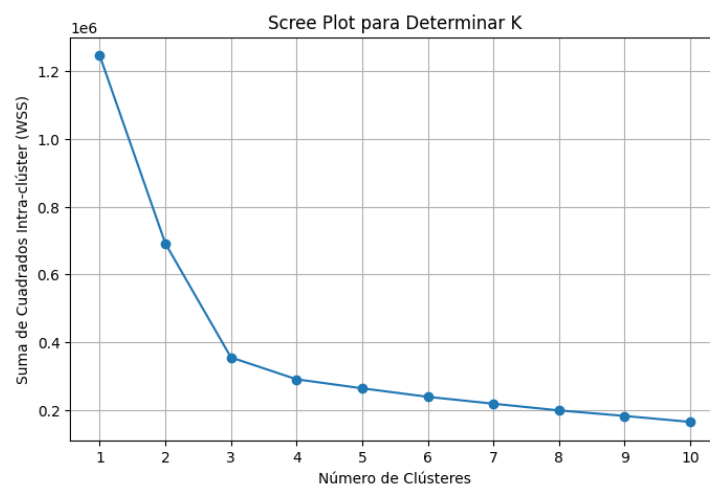


Figura 10: Gráfico del codo para determinar k en la categoría harinas

A continuación, se presentan en los siguientes cuadros los valores promedio de las principales variables asociadas a la lealtad hacia la marca propia, calculadas para cada clúster.

Variable	Clúster 1	Clúster 2	Clúster 3	Clúster 4	Clúster 5	Clúster 6
Gasto Total (UF)	7.31 (5.72)	7.71 (5.99)	7.39 (5.78)	7.77 (6.07)	11.62 (7.15)	7.77 (6.12)
Frecuencia NO MP (compras/año)	0.014 (0.12)	0.15 (0.41)	0.075 (0.29)	1.20 (0.62)	0.78 (0.84)	1.19 (0.54)
Market Share MP (%)	99.6 (2.92)	93.1 (18.50)	96.4 (13.83)	0.32 (2.76)	69.8 (28.87)	0.45 (3.59)
Recompra MP (%)	99.6 (2.66)	94.2 (15.46)	97.2 (10.72)	0.47 (3.99)	72.9 (26.36)	0.48 (3.78)
Frecuencia M1 (compras/año)	0 (0.01)	1.25 (0.70)	1.19 (0.61)	0.016 (0.13)	0.74 (0.87)	0.01 (0.12)
Frecuencia M2 (compras/año)	1.08 (0.35)	0.007 (0.08)	0.009 (0.09)	0.002 (0.03)	1.19 (0.05)	0.002 (0.5)
Multilealtad MP (percentil)	100 (0.00)	99.2 (3.19)	99.5 (4.47)	1.44 (11.18)	96.6 (7.67)	1.66 (12.05)
Fidelidad Activa (%)	99.7 (2.48)	94.8 (14.65)	97.5 (10.13)	0.52 (4.44)	75.3 (25.89)	0.5 (4.07)
Tamaño (clientes)	11,612	15,672	7,383	30,205	4,849	8,153

Nota: Desviación Estándar en paréntesis

Cuadro 7: Promedio de variables de lealtad a marca propia por clúster – Parte 1.

Variable	Clúster 1	Clúster 2	Clúster 3	Clúster 4	Clúster 5	Clúster 6
Índice de Apoyo MP (%)	21.0 (13.48)	21.1 (12.98)	22.9 (14.24)	12.8 (10.71)	20.2 (10.98)	12.9 (10.91)
Market Share M1 (%)	0.001 (0.26)	92.9 (18.53)	96.2 (13.95)	0.31 (2.73)	23.1 (24.49)	0.41 (3.45)
Market Share M2 (%)	99.6 (2.93)	0.129 (1.68)	0.17 (2.02)	0.01 (0.39)	46.64 (17.82)	0.046 (0.99)
Multilealtad M1 (%)	0.007 (1.20)	99.3 (3.34)	97.1 (15.98)	1.37 (10.99)	51.6 (46.82)	1.45 (11.35)
Multilealtad M2 (%)	100 (0.00)	0.60 (7.43)	0.78 (8.41)	0.07 (2.47)	94.4 (8.27)	0.21 (4.33)
Recompra M1 (%)	0 (0.09)	94.1 (15.51)	97.1 (10.86)	0.45 (3.91)	25.3 (26.43)	0.45 (3.68)
Recompra M2 (%)	99.6 (2.67)	0.11 (1.68)	0.13 (1.90)	0.019 (0.80)	47.5 (17.09)	0.03 (0.83)
Oferta (%)	0.04 (1.00)	0.95 (6.36)	0.78 (5.77)	0.77 (4.91)	5.85 (15.13)	94.03 (15.07)
Sensibilidad Precio (%)	39.3 (47.98)	3.4 (11.34)	96.9 (9.72)	14.8 (34.40)	33.7 (36.56)	25.8 (42.44)

Nota: Desviación Estándar en paréntesis

Cuadro 8: Promedio de variables de lealtad a marca propia por clúster – Parte 2.

Finalmente, los clústeres obtenidos en esta segmentación serán caracterizados también en función de las variables asociadas a la lealtad hacia el supermercado. Esta información se presenta en el siguiente cuadro:

Variable	Clúster 1	Clúster 2	Clúster 3	Clúster 4	Clúster 5	Clúster 6
Frecuencia (compras/año)	6.29 (4.38)	6.87 (4.61)	6.48 (4.47)	6.31 (4.37)	8.70 (4.93)	6.25 (4.28)
Market Share (%)	41.55 (12.69)	42.91 (11.52)	43.16 (11.47)	41.40 (13.01)	44.20 (9.80)	41.93 (12.59)
Monetary (UF)	7.31 (5.72)	7.71 (5.99)	7.39 (5.78)	7.77 (6.07)	11.62 (7.15)	7.77 (6.12)
Recencia (días)	125.6 (105.60)	126.4 (104.51)	112.9 (110.09)	126.3 (108.78)	95.6 (87.32)	140.4 (119.17)

Nota: Desviación Estándar en paréntesis

Cuadro 9: Promedio de variables de lealtad al supermercado por clúster.

La segmentación realizada sobre los clientes con al menos una compra durante el último año del período analizado (año 3), considerando todas las variables relacionadas con la lealtad hacia la marca propia, permitió identificar seis clústeres con perfiles de comportamiento bien diferenciados. A continuación, se describen los principales hallazgos de cada segmento:

- **Clúster 1:** Este grupo se caracteriza por una alta fidelidad conductual hacia la marca propia, con valores extremos tanto en *market share MP* (99.6%) como en *recompra MP* (99.6%). Las compras están casi totalmente concentradas en la marca M2, con un *market share* de 99.6% y una *multilealtad* del 100%, mientras que el *market share* de la marca M1 es prácticamente nula (0.014%). Esta preferencia absoluta sugiere una identificación clara y exclusiva con una de las líneas de marca propia. Además, el *índice de apoyo hacia la marca propia* en el total de categorías es elevado (21.0%), lo que indica un compromiso consistente con este tipo de productos más allá de la categoría analizada.

Los niveles de *recompra* también reflejan esta tendencia: 0.01% para M1 y 99% para M2. Por su parte, la *fidelidad activa* alcanza el 99.7%, evidenciando una preferencia constante cuando la marca

está disponible. La preferencia a *ofertas* de otras marcas es prácticamente inexistente (0.04 %), lo que sugiere una lealtad fuerte que no se ve afectada por descuentos en otras marcas. No obstante, se observa una *sensibilidad al precio* relativamente alta (39.3 %), lo cual podría interpretarse como una conciencia de valor por parte de los clientes, sin que esto implique cambiar su elección cuando existen alternativas más económicas.

En relación con la lealtad al supermercado, este segmento presenta indicadores bajos: una *frecuencia de compra anual* de solo 6.29 visitas, un *market share* en tienda de 41.55 % y una *recencia* promedio de 125.6 días. Estos valores sugieren que los clientes de este clúster realizan compras ocasionales y altamente dirigidas, con un vínculo débil hacia el supermercado como canal general de compra, pero con una fuerte afinidad hacia una marca propia específica.

- **Clúster 2:** Este segmento está compuesto por clientes con una clara inclinación hacia la marca propia, presentando un *market share* del 93.1 % y una *recompra* del 94.2 %, lo que refleja una fuerte fidelidad conductual. A diferencia del Clúster 1, este grupo presenta una *sensibilidad al precio* significativamente menor (3.4 %) y una baja preferencia por *ofertas* de otras marcas (0.95 %), lo que sugiere un patrón de compra más estable y menos influenciado por incentivos comerciales. Este comportamiento podría explicarse, en parte, por el hecho de que M1 solo representa la opción más económica en el formato de 1 kg. La baja sensibilidad al precio observada concuerda con la elección recurrente de formatos mpas grandes (3 kg, 5 kg, 10 kg o 20 kg), lo que indica una disposición a pagar más por volumen o conveniencia, y posiblemente una mayor fidelidad hacia la marca.

Su comportamiento se concentra exclusivamente en la marca M1, que representa el 92.9 % de su *market share*, mientras que la marca M2 prácticamente no es considerada (0.01 %). Esto se traduce también en la mayor *frecuencia de compra anual* hacia M1 (1.25 veces), la más alta entre todos los clústeres.

En cuanto al *índice de apoyo a la marca propia*, se sitúa en 21.1 %, un valor alto que sugiere que estos clientes también priorizan la marca propia en otras categorías más allá de harinas. Sin embargo, su relación con el supermercado no alcanza a ser definida como leal. Aunque su *frecuencia anual* de compra (6.87 veces) y su *gasto total* (7.71 UF) son algo superiores a los del Clúster 1, estos indicadores aún no permiten clasificarlo como un grupo comprometido con el supermercado, sino más bien como un segmento con fuerte afinidad hacia una marca específica del supermercado, pero con un vínculo débil con la tienda en su conjunto.

- **Clúster 3:** Este grupo muestra una elevada afinidad hacia la marca propia, con un *market share* del 96.4 % y una tasa de *recompra* del 97.2 %, lo que evidencia una fuerte consistencia en sus decisiones de compra. Su comportamiento está profundamente influenciado por el precio, presentando la *sensibilidad* más alta entre todos los clústeres (96.9 %). Esta característica se refleja en su preferencia por productos de formato pequeño, que suelen tener precios unitarios más accesibles, lo que lo convierte en un segmento altamente receptivo a variaciones de precio o promociones.

Pese a su enfoque económico, este clúster mantiene una *fidelidad activa* muy elevada (97.5 %), lo que indica que, siempre que la marca propia esté disponible, el cliente la elige de manera sistemática. Además, registra el *índice de apoyo a la marca propia* más alto (22.9 %) entre todos los segmentos, reflejando un fuerte compromiso con la marca propia en otras categorías más allá de harinas.

Al igual que el Clúster 2, su consumo se concentra casi exclusivamente en la marca M1, con un *market share* de 96.2 %, mientras que su participación en M2 es prácticamente nula (0.01 %). Sin embargo, a diferencia de aquel, este grupo responde mucho más intensamente al factor precio. A pesar de ello, su exposición a ofertas de otras marcas es baja (0.78 %), lo que refuerza su preferencia estructural por la marca propia, incluso en ausencia de promociones externas.

Desde la perspectiva del supermercado, este clúster presenta una fidelidad intermedia. Con una *frecuencia de compra* de 6.48 veces al año y un *gasto total* de 7.39 UF, se posiciona como el segundo grupo más comprometido con la tienda, sólo superado por el Clúster 5. No obstante, estos valores aún no son suficientes para definirlo como verdaderamente leal al supermercado, aunque sí muestran un potencial para mejorar la relación.

- **Clúster 4:** Este es el segmento más numeroso, compuesto por más de 30,000 clientes, y se caracteriza por una desvinculación casi total tanto con la marca propia como con el supermercado. Presenta un *market share* de apenas 0.32% y una tasa de *recompra* del 0.47% hacia la marca propia, lo que evidencia una mínima afinidad con los productos de marca del supermercado. Su *índice de apoyo* es también extremadamente bajo (12.8%), lo que confirma que estos clientes no muestran preferencia por la marca propia en otras categorías.

En cuanto a su comportamiento promocional, este grupo presenta una *sensibilidad al precio* baja (14.8%) y una preferencia a *ofertas* de otras marcas igualmente mínima (0.77%). Esto sugiere un perfil de cliente poco receptivo a estímulos comerciales.

Desde la perspectiva del supermercado, este clúster también refleja una baja vinculación con el supermercado: su *frecuencia de compra* anual es de solo 6.31 veces y su *gasto total* de 7.77 UF.

- **Clúster 5:** Este clúster, el más pequeño en términos de número de clientes, representa el perfil más interesante. Muestra una alta lealtad tanto hacia la marca propia como al supermercado. Su *market share* en marca propia alcanza un 69.8%, acompañado de una tasa de *recompra* del 72.9%, lo que indica una adopción frecuente y sostenida de productos de marca propia. A diferencia de otros segmentos, este grupo combina adquisiciones de las marcas M1 y M2, con un *market share* de 23.2% y 46.6% respectivamente, lo que sugiere una diversificación positiva dentro del portafolio de marca propia.

Su *índice de apoyo a la marca propia* también es elevado (20.2%), evidenciando que su compromiso con esta línea se extiende a otras categorías. Además, su comportamiento frente a estímulos comerciales es equilibrado: muestran una baja preferencia por *ofertas* (5.85%) de otras marcas y una *sensibilidad al precio* moderada (33.7%), lo que indica que estos clientes no se guían exclusivamente por el precio, sino que valoran otros atributos.

En términos de lealtad al supermercado, este grupo destaca claramente sobre el resto. Presentan la mayor *frecuencia de compra* anual (8.7 veces), el *market share en tienda* más alto (44.2%), el mayor *gasto total* (11.62 UF) y la menor *recencia* (95.6 días), lo que refleja un vínculo sólido y constante con el supermercado. En conjunto, todos estos indicadores posicionan al Clúster 5 como el grupo más valioso.

- **Clúster 6:** Este grupo está conformado por clientes con escasa o nula lealtad hacia la marca propia, reflejada en niveles muy bajos de *market share* (0.45%) y de *recompra* (0.48%). Sin embargo, a diferencia del Clúster 4, este segmento muestra una preferencia por *ofertas* de otras marcas extremadamente alta (94.0%) y una *sensibilidad al precio* intermedia (25.1%). Esto sugiere que, aunque no presentan afinidad con la marca propia, sí responden a estímulos promocionales, o bien que la marca que habitualmente consumen se encuentra con frecuencia en oferta, sin necesariamente ser la más barata.

Su comportamiento evidencia una preferencia sostenida por productos externos al portafolio de marca propia del supermercado, lo cual se refleja en una *multilealtad* de apenas 1.66%. Además, su *índice de apoyo a la marca propia* es bajo (12.9%), y su *fidelidad activa* mínima (0.5%), lo que confirma que estos clientes rara vez eligen la marca propia, incluso cuando está disponible.

Desde la perspectiva del supermercado, este clúster presenta bajos niveles en todos los indicadores de lealtad al supermercado: una *frecuencia de compra* reducida (6.25 veces al año), un *market share en tienda* bajo (41.93 %), una *recencia* elevada (140 días) y uno de los menores niveles de *gasto total* (7.77 UF). En consecuencia, se posiciona como un segmento de bajo valor estratégico.

Observaciones Finales

Los resultados de la segmentación permiten concluir que la lealtad hacia la marca propia y la lealtad hacia el supermercado no necesariamente coexisten, y que estas dimensiones pueden desarrollarse de manera independiente en función del perfil del cliente. Ejemplo de ello es el *Clúster 1*, donde se observa una fidelidad extremadamente alta a la marca *M2*, sin un vínculo sostenido con la tienda en su conjunto. Esta disociación entre producto y supermercado evidencia que una fuerte conexión con la marca no garantiza recurrencia ni mayor gasto en el supermercado.

Por el contrario, el *Clúster 5* sobresale como el segmento más completo y valioso en relación con la lealtad. Se trata del grupo con mayor compromiso combinado tanto con la marca propia como con el supermercado, registrando los niveles más altos en lealtad al supermercado. A pesar de ser el clúster más pequeño, concentra el mayor valor económico y de lealtad sostenida, lo que demuestra que la fidelidad hacia una marca puede ir acompañada de una conexión sólida con el supermercado.

En el extremo opuesto, los *Clústeres 4 y 6* agrupan una proporción significativa de clientes desvinculados, caracterizados por una baja o nula afinidad hacia la marca propia y el supermercado. El *Clúster 4*, en particular, muestra un desinterés generalizado incluso ante estímulos como descuentos, mientras que el *Clúster 6*, aunque poco leal, presenta una alta preferencia a marcas en promoción.

Por otra parte, los *Clústeres 2 y 3* representan perfiles intermedios: clientes con fuerte preferencia por la marca propia, especialmente la marca *M1*, pero con un vínculo aún débil con el supermercado.

En síntesis, esta segmentación no solo permite identificar a los clientes más valiosos, sino también comprender sus motivaciones y preferencias. Al caracterizar a cada segmento en ambas dimensiones, marca y supermercado, se facilita la comprensión de los distintos perfiles de comportamiento, permitiendo distinguir con mayor claridad a aquellos que pueden considerarse realmente leales.

5.7. Segmentaciones Complementarias para Validación

Con el objetivo de reforzar y validar la segmentación original presentada en la sección 5.6.1, se desarrollaron dos segmentaciones complementarias. Estas permiten contrastar los resultados anteriores desde nuevas perspectivas metodológicas, evaluando la consistencia de los patrones observados y el comportamiento de los clientes bajo condiciones ligeramente distintas. Ambas segmentaciones de validación se realizaron utilizando nuevamente el algoritmo *K-means*, determinando seis clústeres en cada caso. Esta elección no fue arbitraria: se utilizó el gráfico del codo como herramienta de apoyo para evaluar el número óptimo de clústeres, observándose nuevamente un punto de inflexión compatible con $k = 6$. Además, mantener el mismo número de clústeres que en la segmentación original permite facilitar la comparación directa entre agrupaciones, fortalecer la consistencia del análisis y observar si ciertos perfiles de comportamiento se repiten o se modifican bajo diferentes enfoques metodológicos.

Los resultados detallados de estas segmentaciones, incluyendo las tablas de promedio por clúster y otras métricas relevantes, pueden revisarse en el Anexo de este documento.

Segmentación sin distinción entre marcas propias.

La primera segmentación de validación replica la estructura y metodología de la segmentación original, pero con una modificación importante en el tratamiento de las variables: en este caso, solo se incluyen

aquellas que consideran a las dos marcas propias de manera conjunta, sin diferenciar entre ellas.

Esta decisión surge a partir de la observación de que, en la segmentación anterior, existían clústeres con similar nivel de lealtad hacia la marca propia, pero con comportamientos exclusivos hacia una de las marcas (por ejemplo, M1 o M2). Al eliminar esta distinción, se busca analizar cómo se comportan los clientes en relación con la marca propia sin el sesgo introducido por la preferencia o exclusividad hacia una marca específica dentro del conjunto.

Esta nueva segmentación permite evaluar si las agrupaciones detectadas anteriormente persisten cuando se abstrae el componente de marca, y si las dinámicas de fidelización observadas se mantienen en un contexto más generalizado. En esta configuración, se obtuvieron los siguientes índices de calidad del agrupamiento: un coeficiente de silueta promedio de 0.3575, un índice de Davies-Bouldin de 1.2290 y un índice de Calinski-Harabasz de 43,183.54. Estos valores permiten evaluar la consistencia interna y la separación entre los clústeres obtenidos. Si bien el coeficiente de silueta sugiere una estructura de segmentación moderadamente clara, los valores de Davies-Bouldin y Calinski-Harabasz indican una segmentación aceptable, con baja superposición entre clústeres y buena compacidad interna, lo cual valida la elección del número de grupos en este análisis más amplio.

Esta segmentación junto con la segmentación principal coinciden en mostrar que la lealtad hacia la marca propia y la lealtad hacia el supermercado pueden desarrollarse de manera independiente. Se identifican grupos de clientes altamente fieles a la marca propia pero con escasa vinculación al supermercado, así como otros que presentan un comportamiento regular en la tienda sin optar por sus productos de marca.

Sin embargo, uno de los hallazgos más relevantes y consistentes en ambas segmentaciones es la existencia de un grupo que combina alta lealtad tanto a la marca propia como al supermercado. En la segmentación principal, este perfil está representado por el *Clúster 5*, mientras que en la segmentación de validación 1 corresponde al *Clúster 4*. En ambos casos, se trata de clientes que presentan altos niveles de market share, recompra y fidelidad activa hacia la marca propia, junto con indicadores destacados de frecuencia de compra, gasto total y recencia en el supermercado. Esta coincidencia permite concluir que existe un segmento claramente identificado como leal tanto a la marca propia como al supermercado.

En síntesis, aunque cada segmentación captura distintos matices de la lealtad, ambas destacan consistentemente a un grupo estratégico que combina compromiso con el producto y con el canal. Esta convergencia valida el enfoque utilizado y respalda la solidez de los resultados obtenidos.

Segmentación basada en lealtad al supermercado.

La segunda segmentación de validación invierte el enfoque de la segmentación original: en lugar de agrupar a los clientes según su lealtad hacia la marca propia y luego analizar su comportamiento hacia el supermercado, esta vez se segmenta exclusivamente en función de las variables de lealtad al supermercado, para posteriormente caracterizar los clústeres según su afinidad con la marca propia.

El objetivo de esta aproximación es evaluar si la lealtad hacia la marca propia se explica por una fidelidad general al supermercado, o si se trata de un fenómeno independiente que eventualmente refuerza el vínculo del cliente con la cadena. Esta estrategia permite explorar la direccionalidad de la relación entre marca propia y supermercado, y distinguir entre patrones de lealtad (hacia el supermercado) o derivada (motivada por la marca propia).

Una vez definida la segmentación, se procedió a caracterizar cada clúster en función del comportamiento hacia la marca propia, con el fin de identificar si los clientes más fieles al supermercado también presentan alta afinidad por esta, o si existen perfiles donde la lealtad se concentra solo en uno de los dos niveles. Para esta segmentación, se obtuvieron los siguientes índices de calidad: un coeficiente de silueta promedio de 0.3059, un índice de Davies-Bouldin de 1.1241 y un índice de Calinski-Harabasz de

35,879.40. Estos resultados permiten evaluar la estructura del agrupamiento y su coherencia interna. En términos generales, el valor del coeficiente de silueta sugiere una segmentación moderadamente diferenciada, mientras que el índice de Davies-Bouldin indica una separación aceptable entre los clústeres. Por su parte, el índice de Calinski-Harabasz, aunque algo inferior al de la segmentación anterior, sigue reflejando una compacidad interna razonable. En conjunto, estos indicadores respaldan la validez del modelo, aunque evidencian una estructura menos marcada que la obtenida en la segmentación centrada en la marca propia.

Al comparar la segmentación principal con la Segmentación de Validación 2, se observa una consistencia importante en la identificación de un grupo de clientes que destaca por su alta lealtad tanto hacia la marca propia como hacia el supermercado. En la segmentación original, este perfil está representado por el *Clúster 5*, mientras que en la validación, el rol equivalente lo cumple el *Clúster 2*.

No obstante, en la Segmentación de Validación 2, esta lealtad dual no es tan marcada ni cuantitativamente tan destacada como en la segmentación principal. El *Clúster 5* original sobresale con métricas muy claras, en cambio, el *Clúster 2* en la validación refleja un patrón similar, pero con menor intensidad, lo que sugiere un grupo leal, aunque menos extremo en su comportamiento.

Además, el *Clúster 4* de la validación también presenta características cercanas a este perfil: si bien su relación con el supermercado no es tan fuerte como la del *Clúster 2*, mantiene niveles elevados de lealtad a la marca propia y un comportamiento de compra relativamente constante, lo que lo posiciona como un grupo con potencial estratégico.

Estas diferencias no disminuyen el valor de la validación, ya que permiten demostrar un hallazgo relevante: no solo se confirma que los clientes leales a la marca propia tienden a ser también leales al supermercado, como se observa en la segmentación principal, sino que también se identifica que quienes presentan un fuerte vínculo con el supermercado suelen concentrar su gasto en marcas propias, como ocurre en la Segmentación de Validación 2. Esto refuerza la existencia de una relación bidireccional entre ambas formas de lealtad, sugiriendo que el compromiso con la marca del supermercado y con el canal de compra pueden reforzarse mutuamente.

En conclusión, ambas segmentaciones permiten caracterizar distintos niveles de lealtad, pero destacan especialmente por identificar de forma consistente un grupo de clientes estratégicos que combinan compromiso con el producto y con el supermercado. Esta alineación entre los resultados evidencia la consistencia del enfoque aplicado y valida su capacidad para describir comportamientos de lealtad.

5.8. Modelo CFA y Regresión Lineal en Categoría Harina

Dado que la categoría de harinas presentó una alta participación de marca propia y una notable variabilidad en el comportamiento de compra, se decidió aplicar un modelo de *análisis factorial confirmatorio* (CFA) específico para esta categoría. El objetivo fue recalcular los scores latentes de lealtad a la marca propia y al supermercado utilizando exclusivamente las variables y datos asociados a compras de harina, logrando así una estimación más precisa y contextualizada.

El modelo CFA planteado fue el siguiente:

$$\begin{aligned} \text{score_mp} &\sim \text{indice_apoyo_mp} + \text{recompra_mp} + \text{frecuencia_mp} + \text{multilealtad_mp} + \text{market_share_mp} \\ \text{score_tienda} &\sim \text{gasto_total} + \text{market_share_tienda} + \text{frecuencia_tienda} + \text{recencia_tienda} \\ \text{score_tienda} &\sim \text{score_mp} \end{aligned}$$

En este modelo, los factores latentes representan constructos no observables pero inferidos a partir de variables empíricas. El `score_mp` representa la lealtad del cliente hacia los productos de marca propia

dentro de la categoría harina. Este score condensa comportamientos clave como la frecuencia de compra de productos de marca propia, su participación en el gasto y la reincidencia en su elección. Por su parte, el `score_tienda` representa la lealtad general del cliente hacia el supermercado.

Variable	Valor	Std. Err	p-value
indice_apoyo_mp	1.0000	-	-
recompra	0.5704	0.001	0.0
frecuencia_mp	0.4923	0.001	0.0
multilealtad_mp	0.5669	0.001	0.0
market_share_marca	0.5662	0.001	0.0
gasto_total	1.0000	-	-
ms_tienda	0.1500	0.004	0.0
frecuencia_tienda	0.8636	0.008	0.0
recencia_días	-0.4873	0.005	0.0

Cuadro 10: Estadísticas de scores de lealtad (lealtad a marca propia y al supermercado).

Las principales cargas factoriales presentan una alta contribución al factor, lo que indica una fuerte relación entre los indicadores más relevantes y sus respectivos constructos latentes. En el Cuadro 11 se muestran los valores de los índices de calidad:

Índice	Valor	Interpretación
Composite Reliability (score_mp)	0.690	Aceptable
Composite Reliability (score_tienda)	0.463	Aceptable en contexto exploratorio
Harman's One-Factor	42.02 %	No hay sesgo significativo por método común

Cuadro 11: Índices de calidad del modelo.

La calidad del modelo fue evaluada mediante indicadores estándar. La varianza explicada por el primer factor alcanzó un 42.02 %, y no se detectó sesgo por método común, lo que respalda la validez factorial. La consistencia interna de los factores, evaluada mediante composite reliability (CR), fue de 0.690 para el score de marca propia y de 0.463 para el score de tienda. Si bien este último valor se encuentra por debajo del umbral tradicionalmente aceptado, sigue siendo adecuado en el contexto exploratorio del análisis, especialmente dada la especificidad de los datos utilizados.

Con los scores latentes estandarizados obtenidos, se estimó un modelo de regresión lineal múltiple con interacción por clúster, con el fin de explorar cómo varía la relación entre la lealtad a la marca propia y la lealtad al supermercado según los distintos grupos de clientes. Esta regresión se aplicó exclusivamente sobre datos de la categoría harina.

Cabe destacar que en esta especificación del modelo, la variable originalmente utilizada para capturar el efecto de las promociones fue reemplazada por una medida más robusta de comportamiento: la sensibilidad al precio. El modelo es el siguiente:

$$\begin{aligned}
score_tienda_std = & \beta_0 + \beta_1 \cdot score_mp_std + \beta_2 \cdot sensible + \beta_3 \cdot decil_comuna \\
& + \beta_4 \cdot monto_liquido_total + \beta_5 \cdot edad + \beta_6 \cdot numero_hijos + \beta_7 \cdot soltero \\
& + \beta_8 \cdot genero_mujer + \beta_9 \cdot cluster_1 + \beta_{10} \cdot cluster_2 + \beta_{11} \cdot cluster_3 \\
& + \beta_{12} \cdot cluster_4 + \beta_{13} \cdot cluster_5 + \beta_{14} \cdot cluster_1_int + \beta_{15} \cdot cluster_2_int \\
& + \beta_{16} \cdot cluster_3_int + \beta_{17} \cdot cluster_4_int + \beta_{18} \cdot cluster_5_int + \varepsilon
\end{aligned} \tag{3}$$

Los resultados obtenidos se pueden observar en el siguiente cuadro:

Variable	Coef.	Std. Err	p-valor
const	0.4862	0.023	0.000
score_mp_std	0.4025	0.021	0.000
sensibilidad	-0.0020	0.005	0.658
decil_comuna	-0.0095	0.004	0.009
monto_liquido_total	0.0005	0.004	0.888
edad	-0.0680	0.004	0.000
numero_hijos	0.0328	0.005	0.000
soltero	-0.0235	0.009	0.006
genero_mujer	-0.0441	0.007	0.000
cluster_2	-0.8307	0.041	0.000
cluster_3	-0.8121	0.038	0.000
cluster_4	-0.8701	0.040	0.000
cluster_5	0.1320	0.015	0.000
cluster_6	-0.1413	0.035	0.000
cluster_2_int	0.1964	0.029	0.000
cluster_3_int	0.2371	0.022	0.000
cluster_4_int	0.2374	0.031	0.000
cluster_5_int	0.4419	0.025	0.000
cluster_6_int	-0.0349	0.030	0.243
R-squared	0.072	—	—
Adj. R-squared	0.072	—	—
F-statistic	307.1	—	0.000
Durbin-Watson	1.977	—	—
Jarque-Bera	18,254.615	—	—
N (total base)	77,874		

Cuadro 12: Resultados simplificados del modelo de regresión lineal con interacción score_mp_std $\times$ clúster. Para modelo de Categoría Harina.

El modelo de regresión lineal múltiple con segmentación por clústeres e interacciones aportó una mejora analítica al permitir explorar cómo varía la relación entre la lealtad a marca propia y la lealtad al supermercado según el perfil de cliente. Aunque el coeficiente de determinación ajustado (*Adjusted R-squared*) se mantuvo en un nivel bajo (0.072), esta especificación entregó una visión más granular y contextualizada, revelando patrones heterogéneos que no fueron capturados por el modelo general.

En esta configuración, el coeficiente del score de marca propia (**score_mp_std**) fue de 0.4025 y resultó altamente significativo ($p < 0.000$), lo que indica que, una mayor lealtad a marca propia se asocia con una mayor lealtad al supermercado. No obstante, las interacciones permiten observar cómo esta relación se modifica en los demás grupos de clientes.

- En el **Clúster 2**, el efecto principal fue negativo y significativo ($\beta = -0.8307$, $p < 0.000$), lo que indica una menor lealtad al supermercado en comparación con el clúster base. Sin embargo, la interacción con **score_mp_std** fue positiva y significativa ($\beta = 0.1964$, $p < 0.000$), lo que implica que, dentro de este grupo, el aumento en la lealtad a marca propia genera un incremento más pronunciado en la lealtad al supermercado.
- En el **Clúster 3**, se observó un patrón similar: menor lealtad al supermercado ($\beta = -0.8121$, $p < 0.000$), pero una interacción significativa y positiva con **score_mp_std** ($\beta = 0.2371$, $p < 0.000$), indicando que el vínculo entre ambas dimensiones se fortalece a medida que crece el compromiso con la marca propia.

- El **Clúster 4** presentó el efecto más negativo sobre la lealtad al supermercado ($\beta = -0.8701$, $p < 0.000$), lo cual es coherente por ser el clúster menos leal. Su interacción con `score_mp_std` también fue positiva y significativa ($\beta = 0.2374$, $p < 0.000$), aunque de menor magnitud, lo que sugiere que el traspaso de la lealtad desde la marca hacia la tienda es más limitado en este segmento.
- En el **Clúster 5**, se observó el único coeficiente positivo respecto al clúster base ($\beta = 0.1320$, $p < 0.000$), lo que indica una mayor lealtad estructural al supermercado. Además, su interacción con `score_mp_std` fue la más alta entre todos los grupos ($\beta = 0.4419$, $p < 0.000$), lo que sugiere que este segmento es especialmente receptivo al vínculo entre marca propia y tienda. Es decir, a medida que aumenta su lealtad a la marca propia, su compromiso con el supermercado también se fortalece significativamente.
- En el **Clúster 6**, el coeficiente principal fue negativo y significativo ($\beta = -0.1413$, $p < 0.000$), mientras que la interacción con `score_mp_std` no fue estadísticamente significativa ($\beta = -0.0349$, $p = 0.243$), lo que sugiere una relación más débil o nula entre ambas formas de lealtad en este grupo.

Además, las variables sociodemográficas y de comportamiento mantuvieron patrones coherentes con modelos anteriores. Las variables `edad`, `decil_comuna`, `soltero` y `genero_mujer` presentaron asociaciones negativas con la lealtad al supermercado, mientras que `numero_hijos` mostró un efecto positivo y significativo. En contraste, `monto_liquido_total` no fue significativo ($\beta = 0.0005$, $p = 0.888$).

La variable `sensibilidad` al precio, tampoco fue significativa ($\beta = -0.0020$, $p = 0.658$). Esto sugiere que, en la categoría de harinas, los clientes no se comportan guiados por estrategias de bajo precio, sino que eligen productos de marca propia por razones que podrían estar más ligadas a la percepción de valor o confianza, más allá del costo.

Adicionalmente, se obtuvo $\text{Durbin-Watson} = 1.977$ se encuentra muy próximo al valor ideal de 2, lo que indica que no existe autocorrelación significativa entre los errores del modelo. Esto valida el supuesto de independencia de los residuos. Por otra parte, $\text{Jarque-Bera (JB)} = 18,254.615$ indica que los residuos no siguen una distribución normal. No obstante, debido al gran tamaño muestral (77,874 observaciones), este desvío es mitigado por el teorema del límite central, por lo que las estimaciones pueden considerarse consistentes y confiables.

En conjunto, esta especificación más flexible permite una comprensión más profunda de los factores que inciden en la lealtad al supermercado. A través de la segmentación por clústeres y la inclusión de términos de interacción con el score de marca propia, se evidencian diferencias significativas entre perfiles de clientes, lo que refuerza la utilidad de este enfoque para generar modelos más interpretables y accionables.

5.9. Comparación distintos modelos

El análisis comparativo entre el modelo general de canasta básica y el modelo específico para la categoría de harinas revela diferencias significativas en la naturaleza del vínculo entre la lealtad a marca propia y la lealtad al supermercado.

En el caso de la canasta básica, los resultados indicaron que la lealtad a marca propia y la lealtad al supermercado no necesariamente coexisten en los mismos perfiles. La regresión mostró una relación negativa entre ambas dimensiones, lo que sugiere que muchos consumidores optan por la marca propia de forma selectiva y estratégica, sin establecer necesariamente un compromiso más amplio con la tienda. Este hallazgo se vio reforzado por la segmentación, donde emergieron clústeres bien diferenciados: algunos con

alta afinidad por la marca propia, pero sin necesariamente estableciendo una lealtad al supermercado, y otros con comportamiento opuesto.

En contraste, el modelo aplicado a la categoría de harinas mostró una dinámica distinta. Si bien también existe heterogeneidad entre los clientes, se identificó un clúster particularmente relevante que demuestra simultáneamente alta lealtad a la marca propia y al supermercado. La regresión lineal en este caso arrojó un coeficiente positivo y significativo para el score de marca propia, lo que indica que, en esta categoría, una mayor afinidad con la marca propia se traduce en un vínculo más fuerte con la tienda. Esto sugiere que, en productos específicos como la harina, donde la oferta de marca propia es sólida, constante y reconocida, la elección se vuelve más estable, fortaleciendo el lazo con el supermercado como proveedor confiable.

En resumen, mientras que el modelo de canasta básica revela que la lealtad hacia la marca propia y hacia el supermercado no siempre se manifiestan de forma simultánea o recíproca, el análisis específico de la categoría de harinas demuestra que, bajo ciertas condiciones, es posible construir una lealtad convergente. En este caso, la preferencia por productos de marca propia no solo coexiste con una fuerte lealtad a la tienda, sino que ambas dimensiones se potencian mutuamente. Esta diferencia tiene implicancias estratégicas relevantes para el desarrollo de marcas propias, al evidenciar que ciertos productos pueden actuar como verdaderos anclajes de fidelización integral. Asimismo, subraya la necesidad de analizar categorías de forma desagregada para identificar con mayor precisión los puntos de contacto entre el comportamiento del consumidor y las estrategias comerciales de la cadena.

5.10. Recomendaciones para Clústeres

En la siguiente sección se presentan los principales insights identificados para cada segmento, junto con una propuesta de política de marketing estratégica y personalizada, diseñada para abordar sus características específicas y maximizar su potencial de fidelización y valor para la compañía.

5.10.1. Segmentación Canasta Básica

- **Clúster 1:** Este segmento, se caracteriza por bajos niveles de lealtad hacia ambos ejes. La interacción negativa con el score de marca propia sugiere que incluso un aumento en la preferencia por estos productos no genera una mayor fidelización al supermercado. Además, presentan alta sensibilidad al precio y baja frecuencia de compra. En este contexto, se recomienda priorizar estrategias de reactivación, como campañas con incentivos de recompra, promociones selectivas y beneficios por retorno. No obstante, dada su limitada rentabilidad potencial, también se sugiere una evaluación costo-beneficio para determinar la conveniencia de invertir en su recuperación activa.
- **Clúster 2:** Este grupo realiza compras poco frecuentes pero de alto volumen, con un gasto total elevado y una fuerte lealtad al supermercado. La interacción positiva, aunque moderada, con la marca propia sugiere una oportunidad de crecimiento en esta dimensión. Se recomienda implementar estrategias de introducción a marcas propias premium, aprovechando su alto gasto promedio y su relación estable con el supermercado. Campañas de cross-selling o recomendaciones personalizadas pueden ser efectivas si destacan atributos diferenciadores como calidad, sostenibilidad o exclusividad. El objetivo debe ser reforzar la propuesta de valor de la marca propia sin comprometer la fidelización general al supermercado.
- **Clúster 3:** Este segmento representa el perfil más valioso, al combinar alta lealtad al supermercado y un fuerte vínculo con la marca propia. La interacción con marca propia es positiva, significativa y de gran magnitud, lo que indica que fortalecer la lealtad hacia la marca propia refuerza directamente la fidelidad al supermercado. Se recomienda consolidar esta relación mediante estrategias de

fidelización intensiva orientadas a la marca propia, como programas de beneficios exclusivos, membresías, reconocimiento al cliente frecuente y experiencias personalizadas. Mantener su satisfacción y reforzar el vínculo emocional con la tienda es clave para evitar su migración a la competencia.

- **Clúster 4:** Este grupo presenta la mayor afinidad con la marca propia, pero una conexión más débil con el supermercado. Sin embargo, tanto el coeficiente como la interacción con el score de marca propia son positivos, lo cual indica un efecto de arrastre: al aumentar la lealtad a la marca propia, también puede mejorar la fidelidad al supermercado. Se recomienda reforzar el posicionamiento de la marca propia como atributo exclusivo del canal, mediante mensajes como “solo aquí encuentras nuestra marca” y beneficios vinculados a la recurrencia de compra. También pueden implementarse campañas de cross-selling que integren productos de marca propia con servicios del supermercado (por ejemplo, delivery, suscripciones o promociones exclusivas en tienda), apuntando a consolidar ambas dimensiones de lealtad.

5.10.2. Segmentación Categoría Harina

- **Clúster 1:** Este grupo presenta alta lealtad a la marca M2 y un fuerte índice de apoyo a la marca propia, lo que sugiere que ya consumen marca propia en otras categorías. Se recomienda fomentar la preferencia por otras marcas propias resaltando atributos de calidad, sostenibilidad y buen precio. Además, pueden utilizarse descuentos cruzados, beneficios por recurrencia, y promociones temporales para aumentar la frecuencia de visitas. También es recomendable introducir nuevas líneas de marca propia con foco en valor y conveniencia para ampliar su consumo.
- **Clúster 2:** Se caracterizan por una preferencia exclusiva hacia la marca M1 y una baja sensibilidad al precio, lo que se refleja en la compra de productos de mayor formato o con características premium. Aunque no presentan lealtad significativa al supermercado, muestran una interacción positiva con la marca propia, lo que sugiere un potencial para fortalecer su vínculo a través de ella. Se recomienda promover la marca propia en otras categorías mediante estrategias de cross-selling y campañas que refuercen su percepción de calidad, exclusividad y conveniencia, incentivando así una mayor identificación con la tienda.
- **Clúster 3:** Este grupo también compra exclusivamente M1, pero, a diferencia del clúster anterior, es altamente sensible al precio, lo que indica que son buscadores de precios bajos. Su índice de apoyo a marca propia es elevado, reflejando una incorporación activa de estos productos en otras categorías. Aunque muestran un coeficiente negativo con la lealtad al supermercado, la interacción positiva con la marca propia sugiere una oportunidad estratégica: fortalecer la oferta de marca propia en distintas categorías podría generar un vínculo más estrecho con el supermercado e incentivar una compra más activa y recurrente. Para lograrlo, se recomienda ampliar la visibilidad de marca propia en puntos de venta, reforzar promociones específicas en categorías aún no exploradas por este grupo, e implementar acciones de fidelización que premien la compra recurrente de marca propia, como cupones personalizados y descuentos por volumen.
- **Clúster 4:** Este segmento evita productos en oferta y no muestra lealtad hacia la marca propia, prefiriendo marcas alternativas sin descuento, posiblemente asociadas a productos de mayor precio. Poseen el coeficiente más negativo en relación a la lealtad al supermercado. No obstante, su interacción positiva con marca propia sugiere una oportunidad de reconversión mediante una estrategia de marca propia premium, enfocada en productos con atributos percibidos como de alta calidad. Esto permitiría captar su atención sin recurrir a descuentos, manteniendo un precio elevado pero competitivo en valor percibido.

- **Clúster 5:** Son consumidores que alternan entre marcas (incluyendo M1 y M2), con una fuerte lealtad tanto a la marca propia como al supermercado. Presentan una de las interacciones más positivas y significativas con la marca propia. Dado su alto valor estratégico, se recomienda priorizar recursos para su retención, mediante programas de fidelización, beneficios exclusivos y una ampliación constante de la oferta de marca propia, lo que permitiría reforzar aún más su vínculo con el supermercado.
- **Clúster 6:** Aunque no presentan lealtad hacia la marca propia, se diferencian del clúster 4 por su alta orientación hacia productos en oferta. Sin embargo, su frecuencia de compra es baja (alta recencia) y la interacción con marca propia no es estadísticamente significativa. Debido a su bajo compromiso y escaso retorno potencial, no se recomienda destinar recursos relevantes a este segmento en el corto plazo.

6. Conclusión

El estudio tuvo como propósito identificar los distintos factores que permiten caracterizar de manera precisa la lealtad conductual de los clientes hacia la marca propia y hacia el supermercado, con el objetivo de evaluar si existe una relación entre ambas dimensiones y, en particular, determinar si la lealtad a la marca propia actúa como un factor que refuerza la fidelización al supermercado. Para ello, se aplicó una segmentación basada en el comportamiento de compra, lo que permitió clasificar a los clientes en distintos clústeres y analizar sus perfiles diferenciados. Sobre la base de dicha segmentación, se construyeron dos indicadores latentes de lealtad mediante un modelo de análisis factorial confirmatorio (CFA): uno orientado a capturar la lealtad hacia la marca propia y otro hacia el supermercado. Estos scores fueron posteriormente utilizados en un modelo de regresión lineal, donde el score de lealtad al supermercado fue tratado como variable dependiente, y el score de marca propia, junto con variables demográficas, los clústeres de segmentación y sus interacciones con el score de la lealtad a marca propia, fueron consideradas como variables explicativas. Este enfoque permitió estimar el efecto específico de la lealtad a marca propia sobre la fidelización al supermercado en distintos segmentos de clientes. El análisis fue aplicado, en primer lugar, a un conjunto de categorías pertenecientes a la canasta básica, y posteriormente replicado para la categoría específica de harinas, con el fin de explorar cómo varían los resultados cuando se analiza un contexto de productos más homogéneos.

Para llevar a cabo estos análisis, fue necesario primero definir qué variables representan de forma más precisa la lealtad de los clientes hacia la marca propia. Para ello, se recurrió a evidencia bibliográfica y criterios de comportamiento observables. Se identificó que dimensiones como la *frecuencia de compra*, el *market share* y la *recompra* constituyen indicadores clave ampliamente utilizados en la literatura para medir lealtad conductual. Adicionalmente, se incorporaron variables complementarias derivadas del análisis teórico y empírico. Por ejemplo, estudios previos señalan que los clientes verdaderamente leales tienden a evitar las promociones de marcas alternativas, lo que motivó la inclusión de la variable *oferta* como un indicador de consistencia en las elecciones de marca. Asimismo, se observó que consumidores con alta lealtad son propensos a esperar el reabastecimiento del producto en caso de quiebre de stock, lo cual dio origen a la variable *fidelidad activa*. Para capturar el grado de incorporación de la marca propia en el comportamiento de compra general, se incluyó la variable *proporción de productos*, que representa el peso relativo de estos productos dentro del total adquirido. Finalmente, considerando la relevancia del precio como factor de decisión, se desarrolló una medida de *promoción*, a partir del análisis del comportamiento frente a productos en descuento. En conjunto, estas variables permitieron construir una definición precisa y multidimensional de lealtad hacia la marca propia, al integrar diversos comportamientos de compra.

A partir de esta formulación, los análisis realizados tanto a nivel general de canasta básica como en la categoría específica de harinas evidencian que la lealtad a marca propia y la lealtad al supermercado no son necesariamente fenómenos alineados. En la segmentación de canasta básica, se identificó un clúster con una fuerte preferencia por productos de marca propia, pero que no presentaba altos niveles de fidelidad al supermercado. Del mismo modo, se encontraron segmentos altamente leales al supermercado que no destacaban por su afinidad hacia la marca propia. Esto sugiere que, si bien existe una leve relación entre ambas formas de lealtad, esta no es lo suficientemente fuerte como para considerarse sistemática en la población total. Esta observación fue respaldada por el modelo de regresión lineal con interacción entre el score de lealtad a marca propia y los clústeres de la segmentación. En dicho modelo, el coeficiente general del score de marca propia fue negativo ($\beta = -0.1383$), indicando que, un mayor nivel de lealtad a marca propia se asocia con una menor lealtad al supermercado. Este efecto fue estadísticamente significativo y refleja una relación negativa entre ambas dimensiones. Sin embargo, al considerar los efectos de interacción de marca propia con los clústeres, se observó que varios grupos presentan coeficientes

positivos y significativos, lo que implica que existen segmentos específicos donde, a medida que aumenta la lealtad a marca propia, también lo hace la lealtad al supermercado. Un caso destacable es el del clúster 3 (interacción Clúster3 $\times$ scoremp, $\beta = 0.1660$) que, aunque no es particularmente leal a marca propia, sí exhibe una fuerte fidelidad al supermercado. Este hallazgo es relevante, ya que muestra que, si bien actualmente la lealtad a marca propia no incrementa directamente la lealtad al supermercado en términos generales, existen segmentos con alto potencial en los que esta relación sí se manifiesta. En otras palabras, la marca propia podría actuar como palanca para fortalecer la relación con el supermercado en ciertos perfiles de cliente.

El análisis aplicado a la categoría de harinas refuerza las conclusiones obtenidas previamente, pero también aporta matices relevantes. A diferencia del análisis general de canasta básica, la segmentación obtenida en harinas evidenció una mayor claridad y diferenciación entre los clústeres. Fue posible identificar con nitidez tres perfiles: clientes leales tanto a la marca propia como al supermercado, clientes leales exclusivamente a la marca propia, y un grupo de clientes no leales. Esta delimitación más precisa de los segmentos permitió caracterizar de forma más específica los patrones de comportamiento, en contraste con los clústeres más difusos observados en el caso general. Además, mientras que en la canasta básica el coeficiente general del score de marca propia fue negativo, en la categoría de harinas fue positivo y significativo ($\beta = 0.4025$), lo que sugiere una diferencia fundamental: si bien a nivel agregado la lealtad a marca propia no se traduce necesariamente en una mayor fidelidad al supermercado, en contextos más específicos como una categoría individual sí puede operar como un motor de fidelización. Esto pone de relieve la importancia de realizar análisis segmentados por categoría, no solo por su capacidad de capturar diferentes relaciones latentes, sino también por la mayor precisión que ofrecen al caracterizar a los distintos perfiles de cliente.

A pesar de estas diferencias, ambos modelos coinciden en un hallazgo clave: la existencia de segmentos con coeficientes de interacción positivos y significativos entre el score de marca propia y los clústeres. Esto implica que, tanto en el análisis general como en el específico, existen grupos de clientes cuya lealtad hacia la marca propia puede actuar como un factor potenciador a su fidelidad al supermercado. Este patrón de lealtad convergente constituye una oportunidad estratégica para diseñar acciones comerciales dirigidas, orientadas a fortalecer simultáneamente ambas dimensiones del comportamiento de compra.

Una observación clave que surgió del análisis de las segmentaciones fueron que los compradores que presentan los mayores índices de lealtad a marca propia, medidos mediante indicadores como market share, fidelidad activa y recompra, no son necesariamente los clientes que realizan más compras. Por el contrario, estos segmentos tienden a realizar un número reducido de compras, pero en una proporción muy alta dirigidas exclusivamente a productos de marca propia, lo que eleva artificialmente sus indicadores de lealtad. Este fenómeno sugiere que la intensidad de la lealtad a marca propia, en ciertos casos, puede estar más asociada al perfil específico del cliente que a un comportamiento de compra intensivo o recurrente. En cambio, a medida que aumenta la cantidad de compras realizadas, se observa una convergencia en los patrones de consumo, lo que implica que los clientes se vuelven más homogéneos entre sí. Este fenómeno refuerza la idea de que, a mayor frecuencia de compra, los clientes tienden a comportarse de forma más homogénea, mientras que los compradores menos activos presentan perfiles más extremos y diferenciados.

En términos generales, los resultados indican que la lealtad hacia la marca propia no presenta una relación directa con la lealtad al supermercado. Esta desvinculación puede atribuirse a la multiplicidad de factores que influyen en la fidelización del cliente. Sin embargo, es especialmente relevante destacar que existen ciertos segmentos de clientes en los que esta relación sí se manifiesta de forma positiva, en estos clústeres, una mayor lealtad a la marca propia se asocia consistentemente con un incremento en la fidelidad al supermercado. Este hallazgo resulta coherente con lo planteado por Coelho do Vale et al. (2016), quienes sostienen que, en supermercados con un enfoque low-cost, como el analizado en este estudio, la lealtad hacia la marca propia no necesariamente se traduce en una mayor fidelidad

al retail. Los resultados obtenidos en esta investigación confirman dicha observación, si bien existen segmentos con alta afinidad hacia los productos de marca propia, ello no implica de forma automática una mayor conexión con el supermercado. Sin embargo, es precisamente esta variabilidad lo que da sustento a la hipótesis planteada: aunque no se observa una relación generalizada entre lealtad a marca propia y lealtad al supermercado, sí se identifican clústeres específicos en los que este vínculo es claro y estadísticamente significativo. Como se ha destacado anteriormente, estos segmentos representan una oportunidad estratégica para posicionar la marca propia como una herramienta de fidelización hacia la tienda, especialmente en aquellos perfiles donde este efecto potenciador puede ser aprovechado de forma efectiva.

El enfoque metodológico desarrollado en distintos niveles y con distintas metodologías confirmó la existencia de grupos diferenciados en cuanto a su vínculo con la marca propia y con el supermercado, lo que permite concluir que la lealtad no es un fenómeno unidimensional, sino que se expresa de manera heterogénea según el contexto de análisis, la categoría de productos y el perfil del consumidor. Comprender esta diversidad es clave para diseñar estrategias comerciales más efectivas, que no sólo promuevan el consumo de marca propia, sino que también contribuyan a fortalecer la relación del cliente con el supermercado de forma sostenida.

Investigación Futura

Se recomienda ampliar esta línea de investigación utilizando datos de supermercados cuyo posicionamiento esté enfocado en segmentos de ingresos altos, con estrategias de precios premium y una fuerte presencia de marcas propias relevantes para ese público objetivo. El estudio de este tipo de formatos permitiría validar si los patrones observados se replican en contextos socioeconómicos distintos y con otro tipo de percepción hacia la marca del distribuidor. Asimismo, se sugiere trabajar con bases de datos actualizadas que reflejen los cambios recientes en el comportamiento del consumidor, especialmente en un escenario post-pandemia y con mayor digitalización de las compras.

7. Anexo

7.1. Métricas por Categoría

Nombre	Market Share Gasto (%)	Proporción Compra Mensual (%)	Cobertura SKU (%)	Relación Precio (%)
ACC. LAVADO Y MAN-TENCIÓN	30.22	38.79	15.78	-15.60
ACCESORIO BAÑOS	11.92	17.55	15.38	-34.72
ACCESORIOS PARA BE-BÉ	16.17	19.90	15.09	-27.57
ACCESORIOS PERSONA-LES	6.62	10.28	17.84	-34.18
ACEITES	37.30	44.73	6.49	-22.00
ACONDICIONADOR	0.39	0.84	1.16	-49.18
AGUA MINERAL	5.60	6.05	3.32	-35.17
AGUA PURIFICADA	3.14	2.78	4.62	6.36
ALIMENTO PARA PE-RROS ADULTOS	10.63	14.44	5.44	24.05
ALIMENTO PARA PE-RROS CACHORRO	8.34	8.38	3.45	21.73
ALIMENTOS OTRAS MASCOTAS GATO	6.69	9.90	2.41	24.97
ARROZ	27.91	32.82	5.92	-18.96
ARTICULOS ACCESO-RIOS	41.58	45.46	21.95	-16.04
ARTICULOS DE AFEITAR	1.79	2.82	1.54	-28.71
ARTICULOS DE BOTI-QUIN	36.74	45.31	7.92	-20.55
ARTICULOS IMPLEMEN-TOS SANITARIOS	29.28	29.18	27.42	-10.72
AZÚCAR	27.31	35.47	7.39	-34.17
BASES LICORES	1.97	2.09	8.33	16.85
BICARBONATO DE SO-DIO	13.55	21.11	20.00	-40.52
BOLSAS Y MAT. DE EN-VOLVER	13.03	17.77	12.88	-21.45
BUTIFARRA	48.46	26.60	14.29	0.03
CAFE	2.87	4.42	4.20	-25.74
CERAS	7.45	7.78	6.25	30.06
CEREALES	9.09	14.70	8.36	-27.78
CHECK OUT CONFITES	0.47	0.84	0.71	-37.32
CHOCOLATES	0.17	0.24	0.11	6.62
CHORIZO	6.91	7.55	2.73	-0.58
CHURRASCOS	58.91	60.51	16.00	1.19
COLORO	21.38	27.52	13.79	-20.40
COCINA INTERNACIO-NAL	7.20	8.73	7.27	-1.54
CODORNIZ	89.72	87.65	33.33	-0.68
COMIDAS PREPARADAS	28.87	33.97	12.17	-8.48
CONDIMENTOS	10.24	19.87	10.19	-35.73
CONSERVAS DE FRUTAS	20.84	22.73	8.26	8.46
CONSERVAS DE MARIS-COS	27.11	28.96	20.00	-4.63
CONSERVAS DE PESCA-DO	31.43	34.29	19.27	-13.69
CONSERVAS DE VEGETA-LES	36.29	37.74	13.75	2.04
COSMÉTICOS	8.81	8.56	3.53	11.91
DECORACIÓN/RELLENO	6.05	12.68	9.47	-42.29
DEPILACION	2.31	2.66	1.10	-14.14
DETERGENTES	2.73	7.89	4.48	-52.71
ENVASADOS	47.92	41.68	10.56	19.71
ESPECIAS	13.47	19.78	7.23	-15.73
FECULAS	41.29	44.35	16.85	-19.44
FIDEOS	11.01	17.18	8.50	-35.43
FILETE	9.28	16.83	18.18	-57.99
FRESCA	1.88	1.13	0.91	-34.42
FRUTAS CONGELADAS	6.53	7.88	6.38	-24.59
FRUTOS SECOS	18.30	21.83	9.43	-15.19
GALLETAS	2.75	3.15	2.42	131.68

Nombre	Market Share Gasto (%)	Proporción Compra Men-sual (%)	Cobertura SKU (%)	Relación Precio (%)
HUEVO GALLINA	44.32	42.86	8.54	5.17
GASEOSAS	1.86	4.31	8.37	-58.65
GRASA	20.78	23.93	50.00	-11.44
HAMBURGUESA	9.41	13.21	9.40	-12.76
HARINA	45.69	49.98	19.64	-20.29
HELADOS	5.09	5.75	3.94	12.58
HIELO	61.49	62.53	18.18	-7.08
HIGIENE BUCAL	0.62	1.64	0.90	-54.36
JABONES DE TOCADOR Y BAÑO	1.67	3.09	3.47	-36.68
JAMON	4.30	4.56	1.17	27.58
JUGOS EN POLVO	3.61	6.48	4.28	-21.72
JUGOS LIQUIDOS	1.73	2.76	3.33	-35.04
LAVALOZAS	9.21	17.26	11.71	-38.22
LECHE LIQUIDA	14.68	16.43	2.79	-19.59
LECHES/CREMAS	2.82	4.74	11.11	-27.31
LEGUMBRES	34.01	41.01	17.11	-14.20
LICORES	5.78	8.31	2.26	-38.34
LIMPIADORES	8.96	15.68	10.95	-25.45
LOMITO	26.25	27.35	33.33	-11.57
LONGANICILLA	3.13	3.41	6.67	-38.80
LONGANIZA	9.77	11.41	4.82	-13.29
MANJAR	10.82	12.71	5.26	-15.97
MANTECA	3.88	2.48	14.29	96.99
MANTEQUILLA	9.40	11.24	4.88	-21.51
MARGARINA	1.31	1.96	3.45	-21.85
MARISCOS CONGELADOS	52.58	46.25	34.83	18.71
MERMELADAS	6.94	11.22	7.51	-36.26
MIELES Y DULCES	15.28	17.31	11.49	-6.65
NECTARES	2.71	3.75	1.60	-19.63
NUGGET	0.81	1.02	2.50	87.64
PAN ENVASADO	12.13	17.33	6.85	-26.89
PAPELES HIGIENICOS	2.82	5.98	4.26	-42.06
PATE	4.64	7.25	5.26	-46.25
PECHUGA	15.55	18.06	4.69	2.22
PERFUMERIA BEBE	9.88	16.41	2.11	-26.88
PESCADOS CONGELADOS	19.99	21.82	15.73	-8.60
PISCOS	0.27	0.55	1.30	-43.26
POSTRES LISTOS	10.57	15.13	9.55	-23.79
PRODUCTOS NATURALES Y DIETÉTICOS	4.93	9.37	0.67	-46.00
PURÉ	25.55	26.59	29.03	-10.91
QUESOS ESPECIALES	0.01	0.01	0.50	-73.76
QUESOS REGULARES	14.31	16.93	6.27	-16.42
SABORIZANTES LECHE	8.84	15.24	11.94	-48.10
SALAME	13.28	12.90	16.67	-58.81
SALCHICHÓN	1.67	2.51	2.50	3.91
SALSAS PASTAS	14.89	19.03	2.21	-40.35
SERVILLETAS	6.60	14.56	4.11	-66.76
SHAMPOO	0.23	0.62	0.66	-53.79
SNACK	6.61	10.96	7.34	-10.90
SOPAS/CREMAS/CALDO	7.67	13.03	9.50	-30.39
TÉ/HIERBAS	0.95	1.45	2.29	-18.26
TOALLAS DE PAPEL	11.16	15.05	4.65	-17.32
TRATAMIENTOS ROPA	11.03	16.47	10.39	-19.92
VEGETALES CONGELADOS	34.00	37.10	9.59	-1.76
VIENESAS	4.43	4.81	4.90	7.32
WHISKY	0.84	1.52	0.66	-52.00
YOGHURT	2.90	5.41	0.75	-71.41

7.2. Resultados Segmentación Validación 1

El cuadro a continuación presenta los promedios de las principales variables asociadas a la lealtad hacia la marca propia para cada uno de los seis clústeres identificados.

Variable	Clúster 1	Clúster 2	Clúster 3	Clúster 4	Clúster 5	Clúster 6
Market Share MP (%)	96.22 (14.27)	0.14 (1.70)	93.28 (18.61)	85.77 (24.34)	0.59 (4.18)	1.22 (5.39)
Recompra MP (%)	97.05 (11.166)	0.22 (2.66)	94.11 (16.27)	87.86 (20.31)	0.59 (4.16)	1.46 (6.24)
Multilealtad MP (%)	99.51 (4.72)	0.70 (7.83)	99.21 (3.64)	98.83 (3.59)	1.95 (13.09)	5.00 (20.20)
Fidelidad Activa (%)	97.36 (10.51)	0.24 (2.80)	94.68 (15.41)	89.01 (19.21)	0.65 (4.58)	1.59 (6.78)
Frecuencia MP (compras/año)	1.12 (0.35)	0.01 (0.09)	1.08 (0.28)	2.19 (1.30)	0.02 (0.14)	0.06 (0.23)
Índice Apoyo MP (%)	22.67 (14.28)	12.82 (11.26)	20.96 (13.25)	20.24 (10.13)	13.04 (11.05)	12.96 (8.34)
Oferta MP (%)	0.73 (5.47)	0.59 (4.70)	0.94 (6.45)	2.82 (10.01)	95.93 (12.30)	4.83 (14.92)
Sensibilidad Precio (%)	96.77 (10.34)	14.41 (34.19)	1.80 (8.24)	28.96 (34.85)	26.36 (42.83)	16.40 (34.76)
Gasto Total (UF)	6.52 (4.25)	5.14 (2.85)	5.95 (3.66)	17.87 (6.67)	7.22 (5.26)	17.18 (5.62)
Tamaño (clientes)	12.130	23.507	21.124	6.143	7.765	7.205

Nota: Desviación Estándar en paréntesis

Cuadro 13: Promedio de variables de lealtad a marca propia por clúster – Segmentación de Validación 1

A continuación, se presentan las métricas promedio de lealtad al supermercado asociadas a cada uno de los clústeres obtenidos en la segmentación de validación 1. Estas variables no fueron utilizadas como criterio de segmentación.

Variable	Clúster 1	Clúster 2	Clúster 3	Clúster 4	Clúster 5	Clúster 6
Frecuencia Supermercado (compras/año)	5.34 (4.25)	4.33 (3.42)	5.22 (4.08)	11.39 (5.82)	5.37 (4.44)	10.12 (5.58)
Market Share Supermercado (%)	42.54 (11.97)	40.66 (13.78)	42.04 (12.40)	45.43 (7.69)	41.81 (12.72)	44.04 (9.37)
Monetary (UF)	6.52 (4.25)	5.14 (2.85)	5.95 (3.66)	17.87 (6.67)	7.22 (5.26)	17.18 (5.62)
Recencia (días)	122.24 (107.09)	139.18 (113.44)	132.44 (108.24)	73.26 (66.42)	143.04 (120.22)	81.41 (75.09)

Nota: Desviación Estándar en paréntesis

Cuadro 14: Promedio de variables de lealtad al supermercado por clúster – Segmentación de Validación 1

7.3. Resultados Segmentación Validación 2

El siguiente cuadro presenta los promedios de las principales variables asociadas a la lealtad al supermercado para cada uno de los seis clústeres obtenidos en la segmentación de validación 2. Los resultados se muestran en los siguientes cuadros:

Variable	Clúster 1	Clúster 2	Clúster 3	Clúster 4	Clúster 5	Clúster 6
Frecuencia Supermercado (compras/año)	5.33 (3.38)	16.10 (5.30)	2.43 (1.74)	8.77 (3.00)	2.67 (1.94)	3.60 (1.91)
Market Share Supermercado (%)	19.35 (8.69)	44.45 (7.98)	18.17 (9.38)	46.30 (5.42)	48.45 (4.32)	47.59 (4.63)
Monetary (UF)	5.69 (3.70)	20.54 (6.72)	3.53 (2.79)	11.71 (3.93)	4.68 (3.33)	4.87 (2.48)
Recencia (días)	70.83 (44.57)	58.37 (53.49)	273.36 (75.13)	68.73 (54.36)	295.16 (65.47)	80.95 (48.35)
Tamaño (clientes)	9.413	6.809	4.269	18.978	14.262	24.143

Nota: Desviación Estándar en paréntesis

Cuadro 15: Promedio de variables de lealtad al supermercado por clúster – Segmentación de Validación 2

Variable	Clúster 1	Clúster 2	Clúster 3	Clúster 4	Clúster 5	Clúster 6
Gasto Total (UF)	5.69 (3.70)	20.54 (6.72)	3.04 (2.79)	11.71 (3.93)	4.68 (3.33)	4.87 (2.48)
Frecuencia No MP (compras/año)	0.64 (0.61)	0.95 (1.14)	0.61 (0.54)	0.77 (0.87)	0.62 (0.58)	0.60 (0.61)
Market Share MP (%)	44.54 (48.57)	48.76 (48.25)	43.05 (48.75)	48.90 (47.41)	44.89 (48.61)	48.79 (48.68)
Recompra MP (%)	44.96 (48.57)	49.84 (46.30)	43.37 (48.77)	49.67 (47.35)	45.22 (48.61)	49.30 (48.67)
Frecuencia M1 (compras/año)	0.33 (0.55)	0.61 (1.4)	0.28 (0.46)	0.53 (0.85)	0.34 (0.54)	0.39 (0.58)
Frecuencia M2 (compras/año)	0.22 (0.44)	0.37 (0.68)	0.20 (0.41)	0.28 (0.55)	0.19 (0.41)	0.21 (0.43)
Multilealtad MP (%)	47.10 (49.65)	56.82 (48.68)	44.72 (49.54)	54.54 (49.20)	47.29 (49.67)	51.65 (49.66)
Fidelidad Activa (%)	45.17 (48.66)	50.59 (46.55)	43.23 (48.81)	50.16 (47.55)	45.44 (48.72)	49.49 (48.74)

Nota: Desviación Estándar en paréntesis

Cuadro 16: Promedio de variables de lealtad a marca propia por clúster – Segmentación de Validación 2

Variable	Clúster 1	Clúster 2	Clúster 3	Clúster 4	Clúster 5	Clúster 6
Índice de Apoyo MP (%)	16.77 (13.45)	16.61 (9.44)	17.33 (15.68)	16.74 (10.32)	17.76 (14.43)	17.49 (13.35)
Market Share M1 (%)	26.26 (42.86)	29.38 (41.71)	25.23 (42.80)	30.78 (43.29)	28.30 (43.94)	31.10 (44.93)
Market Share M2 (%)	18.28 (37.44)	19.38 (35.21)	17.83 (37.56)	18.11 (35.48)	16.59 (36.10)	17.69 (36.80)
Multilealtad M1 (%)	28.52 (44.88)	37.26 (47.47)	26.34 (43.92)	36.72 (47.55)	30.27 (45.68)	33.72 (46.95)
Multilealtad M2 (%)	20.31 (40.04)	27.09 (43.55)	18.99 (39.09)	23.27 (41.63)	18.27 (38.44)	19.86 (39.64)
Recompra M1 (%)	26.58 (42.99)	30.34 (42.17)	25.43 (42.89)	31.56 (43.57)	28.62 (44.06)	31.53 (45.08)
Recompra M2 (%)	18.37 (37.50)	19.49 (35.38)	17.94 (37.64)	18.11 (35.46)	16.60 (36.10)	17.76 (36.84)
Oferta (%)	10.56 (29.68)	10.49 (27.08)	11.46 (31.40)	10.05 (27.64)	13.57 (33.33)	9.76 (28.65)
Sensibilidad Precio (%)	23.89 (41.33)	25.61 (39.37)	25.38 (42.65)	27.13 (41.33)	27.27 (43.31)	26.51 (42.66)

Nota: Desviación Estándar en paréntesis

Cuadro 17: Promedio de variables de lealtad a marca propia por clúster – Segmentación de Validación 2 (Parte 2)

Referencias

- Akoglu, H. E. and Özbek, O. (2021). The effect of brand experiences on brand loyalty through perceived quality and brand trust: a study on sports consumers. *Asia Pacific Journal of Marketing and Logistics*, 34(10):2130–2148.
- Arunachalam, D. and Kumar, N. (2018). Benefit-based consumer segmentation and performance evaluation of clustering approaches: An evidence of data-driven decision-making. *Expert Systems with Applications*, 111:11–34.
- Ballestar, M. T., Grau-Carles, P., and Sainz, J. (2018). Customer segmentation in e-commerce: Applications to the cashback business model. *Journal of Business Research*, 88:407–414.
- Baltas, G. (2003). A combined segmentation and demand model for store brands. *European Journal of Marketing*, 37(10):1499–1513.
- Bucklin, R. E., Gupta, S., and Han, S. (1995). A brand’s eye view of response segmentation in consumer brand choice behavior. *Journal of marketing research*, 32(1):66–74.
- Coelho do Vale, R., Verga Matos, P., and Caiado, J. (2016). The impact of private labels on consumer store loyalty: An integrative perspective. *Journal of Retailing and Consumer Services*, 28:179–188.
- do Vale, R. C., Matos, P. V., and Caiado, J. (2016). The impact of private labels on consumer store loyalty: An integrative perspective. *Journal of retailing and consumer services*, 28:179–188.
- Germann, F., Lilien, G. L., and Rangaswamy, A. (2013). Performance implications of deploying marketing analytics. *International Journal of Research in Marketing*, 30(2):114–128.
- Gielens, K., Dekimpe, M. G., Mukherjee, A., and Tuli, K. (2023). The future of private-label markets: A global convergence approach. *International Journal of Research in Marketing*, 40(1):248–267.
- Griva, A., Bardaki, C., Pramataris, K., and Doukidis, G. (2022). Factors affecting customer analytics: Evidence from three retail cases. *Information Systems Frontiers*, pages 1–24.
- Jacoby, J. and Olson, J. C. (1970). *Consumer Behavior: An Attitudinal Approach*. Holt, Rinehart and Winston, New York.
- John, J. M., Shobayo, O., and Ogunleye, B. (2023). An exploration of clustering algorithms for customer segmentation in the uk retail market. *Analytics*, 2(4):809–823.
- Kasem, M. S., Hamada, M., and Taj-Eddin, I. (2024). Customer profiling, segmentation, and sales prediction using ai in direct marketing. *Neural Computing and Applications*, 36(9):4995–5005.
- Knox, S. and Walker, D. (2001). Measuring and managing brand loyalty. *Journal of strategic marketing*, 9(2):111–128.
- La Tercera (2024). Cómo ha subido la venta de productos de marcas propias de los supermercados en medio de una débil economía. Consultado el 31 de octubre de 2024.
- Martos-Partal, M. and González-Benito, O. (2009). The effects of store brand loyalty on store loyalty: evidence from the spanish market. *The International Review of Retail, Distribution and Consumer Research*, 19(3):273–288.

- Ministerio de Desarrollo Social (2013). Comunas de Chile según nivel socio-económico, de salud y desarrollo humano. revisión 2013. https://www.desarrollosocialyfamilia.gob.cl/storage/docs/Comunas_de_Chile_según_nivel_socioeconomico_de_salud_y_desarrollo_humano.pdf. Ministerio de Desarrollo Social, Gobierno de Chile.
- Ministerio de Desarrollo Social y Familia (2016). Nueva serie canasta básica de alimentos 2016. <https://observatorio.ministeriodesarrollosocial.gob.cl/nueva-serie-cba-2016>. Accedido el 17 de julio de 2025.
- Møller Jensen, J. (2011). Consumer loyalty on the grocery product market: an empirical application of dick and basu's framework. *Journal of Consumer Marketing*, 28(5):333–343.
- Ndlovu, S. G. (2024). Private label brands vs national brands: new battle fronts and future competition. *Cogent Business & Management*, 11(1):2321877.
- Nguyen, S. P. (2021). Deep customer segmentation with applications to a vietnamese supermarkets' data. *Soft Computing*, 25(12):7785–7793.
- Palmatier, R. W., Petersen, J. A., and Germann, F. (2022). *Marketing Analytics: Based on First Principles*. Bloomsbury Publishing.
- Rahim, M. A., Mushafiq, M., Khan, S., and Arain, Z. A. (2021). Rfm-based repurchase behavior for customer classification and segmentation. *Journal of Retailing and Consumer Services*, 61:102566.
- Rawat, K. (2023). Applying crisp-dm methodology in developing machine learning model for credit risk prediction. In *Science and Information Conference*, pages 522–538. Springer.
- Romaniuk, J. and Nenycz-Thiel, M. (2013). Behavioral brand loyalty and consumer brand associations. *Journal of Business Research*, 66(1):67–72.
- Rungruang, C., Riyapan, P., Intarasit, A., Chuarkham, K., and Muangprathub, J. (2024). Rfm model customer segmentation based on hierarchical approach using fca. *Expert Systems with Applications*, 237:121449.
- Sondoh, S. L., Wahid, N. A., Ismail, I., Harun, A., and Kassim, A. W. M. (2007). The effect of brand image on overall satisfaction and loyalty intention in the context of color cosmetic. *Asia Pacific Journal of Marketing and Logistics*, 19(4):240–255.