



Universidad de
SanAndrés

E337 - Big Data, Aprendizaje y Minería de Datos -
Semestre Otoño 2024

**“Segmentación del Mercado de Libros en
Chile: Análisis Avanzado de Datos para
Estrategias de Marketing Personalizadas
para Equis Libros”**

Lucila Brisighelli y Maria Constanza Cerundolo

Profesora: Noelia Romero ; **Tutora:** Victoria Oubiña

5 de Julio, 2024

1 Introducción

En un mundo cada vez más digitalizado, el comercio electrónico ha transformado radicalmente la forma en que las empresas interactúan con sus clientes. Equis Libros, una librería argentina, esta sumido a esta tendencia al enfocarse en gran medida en las ventas online a través de la plataforma de Mercado Libre. Recientemente, la librería ha expandido sus operaciones al mercado chileno, enfrentándose a una competencia intensa con múltiples actores establecidos y una demanda diversificada. Este escenario presenta tanto desafíos como oportunidades para Equis Libros.

La entrada de Equis Libros en el mercado chileno plantea la necesidad de entender las particularidades de este nuevo entorno competitivo. La competencia con librerías establecidas y la diversidad de preferencias de los consumidores chilenos requiere una adaptación precisa de la oferta y las estrategias de marketing. Este estudio se motiva por la oportunidad de utilizar técnicas avanzadas de big data para identificar segmentos de mercado específicos, permitiendo a Equis Libros ajustar su oferta y estrategia para satisfacer mejor las necesidades de sus clientes chilenos.

Desde una perspectiva de microeconomía, la segmentación de mercados se vuelve óptima para un vendedor cuando los consumidores tienen diferentes disposiciones a pagar y se puede identificar claramente estos segmentos. Esta práctica permite maximizar los ingresos al cobrar precios diferentes a distintos grupos de consumidores basándose en su disposición a pagar. La discriminación de precios puede ser beneficiosa porque permite capturar el excedente del consumidor, incrementando así los ingresos del vendedor. Por ejemplo, ofrecer un producto a mayor precio para un grupo de consumidores que valora más el producto, y a un precio menor para otro grupo más sensible al precio, puede aumentar las ventas totales y la rentabilidad.

¿Cómo podemos segmentar el mercado de libros en Chile para que Equis Libros pueda dirigir de manera efectiva sus productos y estrategias de marketing a diferentes segmentos de clientes? Esta pregunta de investigación busca abordar el desafío fundamental que enfrenta Equis Libros en su expansión. Llevar a cabo esta propuesta no solo beneficiará a Equis Libros en su expansión, sino que también ofrecerá valiosas contribuciones a la literatura existente sobre segmentación de mercados en el contexto del comercio electrónico. Los hallazgos podrían servir como referencia para otras empresas del sector que buscan expandirse a nuevos mercados y mejorar su estrategia de marketing mediante el uso de análisis de datos avanzados.

2 Literatura previa

La segmentación del mercado y el análisis del comportamiento del cliente han sido ampliamente estudiados en la literatura del marketing y la economía digital. De esta forma, presentaremos algunos trabajos relevantes que sirven de base para nuestra propuesta de segmentar el mercado de libros en Chile y diseñar estrategias adaptadas a las necesidades específicas de los consumidores chilenos.

Casariago (2017) en su paper *"A Common Data Representation Model for Customer Behavior Tracking"*, establece un marco para representar de manera común el comportamiento observado del cliente. Este trabajo destaca que, gracias a la transformación digital, la adopción de redes sociales, la infraestructura en la nube y la tecnología de big data, ahora es posible rastrear y almacenar el comportamiento de los clientes. El autor propone un modelo de datos que permite a las organizaciones representar cualquier modelo de negocio centrado en el cliente, abordando problemas de marketing como la segmentación, la venta cruzada y la retención de clientes. Este modelo facilita el uso de herramientas analíticas y predictivas populares, permitiendo a las empresas anticipar el comportamiento de sus clientes y ofrecer servicios más personalizados y relevantes. Esta capacidad es crucial para mejorar la experiencia del cliente y aumentar su lealtad.

Por otro lado, el segundo paper de Zerboni, Silva y Prado (2013) *"Gestión de cuentas clave desde la perspectiva del cliente"*, publicado en la *Revista de Negocios del IEEM*, argumentan que la correcta segmentación es fundamental para diseñar estrategias de venta que se adapten a las diferentes necesidades de los clientes. Los autores proponen segmentar a los clientes basándose en su comportamiento, lo cual permite una mayor comprensión y diferenciación de los mismos, fortaleciendo así la competitividad de la empresa. Este enfoque se centra en identificar criterios de clasificación de clientes que van más allá de las bases tradicionales, como el tamaño o la ubicación geográfica, y se enfoca en aspectos más complejos y específicos del comportamiento del cliente. La implementación de un sistema de gestión de cuentas clave (KAM) desde la perspectiva del cliente es presentada como una estrategia efectiva para enfrentar los desafíos del mercado actual.

A su vez, se destaca la importancia de correlacionar perfiles demográficos, psicográficos y conductuales del personal operativo con las necesidades del cliente para mejorar la segmentación y la experiencia del usuario. Este estudio enfatiza que la adopción de big data en el comercio electrónico permite a las empresas obtener una comprensión más profunda de sus clientes, lo cual es esencial para diseñar estrategias de marketing efectivas. La capacidad de identificar patrones en los datos y elaborar predicciones es crucial para optimizar la prestación de servicios y aumentar la satisfacción del cliente.

Fiedler et al.(2016), en paper *"Why Customer Analytics Matter"*, sostienen que las empresas que hacen un uso extensivo de análisis de clientes tienden a tener un impacto considerable en su rendimiento corporativo, superando a sus competidores. Para ello, se destaca la necesidad de recopilar todas las transacciones e interacciones de los clientes a través de todos los canales, y de entender y aprender de ese comportamiento para poder anticipar sus movimientos.

Por ultimo, Zerboni, Silva y Prado (2013) también destacan que la segmentación basada en el comportamiento del cliente permite a las empresas desarrollar estrategias de venta más eficaces y adaptadas. Proponen segmentar las cuentas clave de acuerdo con cómo los clientes clasifican a los proveedores y el valor que buscan en los productos o servicios. Esta segmentación facilita una diferenciación más precisa y fortalece la ventaja competitiva de la empresa, ya que permite una mejor comprensión del cliente y sus necesidades.

De esta forma, la literatura existente muestra la relevancia de la segmentación de mercados basada en el comportamiento del cliente y el uso de tecnologías avanzadas de análisis de datos para mejorar la efectividad de las estrategias de marketing. La adopción de estas técnicas permitirá a Equis Libros optimizar su oferta y estrategias de marketing, logrando una mejor penetración en el mercado chileno y aumentando la satisfacción y lealtad de sus clientes.

3 Base de datos

Para llevar a cabo la segmentación del mercado de libros en Chile y diseñar estrategias de marketing efectivas para Equis Libros, es fundamental contar con una base de datos robusta y bien estructurada. Esta base de datos debe integrar información obtenida tanto de las transacciones realizadas en Mercado Libre como de fuentes externas, como encuestas demográficas nacionales y estudios de mercado.

El primer conjunto de datos proviene de Mercado Libre, la plataforma a través de la cual Equis Libros realiza sus ventas online. Los datos de comportamiento de compra incluyen información sobre los libros comprados anteriormente, la frecuencia de compra y las categorías de interés. Estos datos permiten entender los patrones de compra de los clientes y segmentar el mercado en función de sus comportamientos. Además, es posible que se pueda acceder a datos básicos de los clientes, como edad, género, así como a sus preferencias de lectura, que incluyen géneros favoritos, autores preferidos y editoriales más seleccionadas.

Para obtener estos datos, la empresa de Equis Libros cuenta con una base con perfiles de usuarios de Mercado Libre, la cual obtiene a través de informes estadísticos todos los

meses. Este proceso generalmente implica la firma de acuerdos de confidencialidad y el cumplimiento de las políticas de privacidad y uso de datos de la plataforma. Algo a tener en cuenta, es que si bien Mercado Libre puede proporcionar una rica fuente de datos transaccionales, la disponibilidad de datos básicos de los individuos y de preferencias específicas dependerá de las políticas de compartición de información de la plataforma.

Además de los datos de Mercado Libre, es crucial complementar esta información con datos de fuentes externas. Las encuestas demográficas nacionales ofrecen una visión más amplia y detallada de la población chilena y sus hábitos de lectura. Datos como edad y género por ubicación geográfica se pueden obtener de encuestas nacionales realizadas por instituciones como el Instituto Nacional de Estadísticas de Chile. Estos datos permiten comprender mejor la distribución de la población, permitiéndonos así agregar la característica de la distribución geográfica de la población en Chile.

Es importante considerar tanto las transacciones ya realizadas como los posibles segmentos de mercado en los que Equis Libros podría enfocarse. Para las transacciones realizadas, los datos disponibles incluirán información sobre el producto comprado, características de la transacción y, si es accesible, datos del cliente. Estos datos permiten crear perfiles detallados de los clientes y analizar patrones de compra, lo cual es esencial para segmentar el mercado efectivamente.

Para identificar posibles segmentos de mercado, se utilizarán los datos básicos y de preferencias obtenidos de encuestas y estudios de Mercado Libre. La combinación de estos datos con la información de compras pasadas permitirá identificar segmentos específicos, como por ejemplo jóvenes lectores interesados en ciencia ficción o adultos mayores que prefieren el estilo biográfico, pudiendo además caracterizarlo por ubicación geográfica en Chile.

De esta forma, se integrarían ambas fuentes de datos, tanto las de Mercado libre, como las provenientes de la pagina de estadísticas nacionales. Luego se podrán almacenar y gestionar todos los datos recopilados, para posteriormente realizarán análisis descriptivos para identificar patrones y tendencias. Podríamos realizar también un análisis predictivo sobre comportamientos futuros y segmentar clientes. Herramientas de visualización de datos se podrían utilizarán para representar gráficamente los segmentos de mercado y las tendencias de compra.

De esta forma, esta base de datos permitirá a Equis Libros obtener una visión integral de sus clientes y del mercado chileno. Esto facilitará la creación de estrategias de marketing efectivas y personalizadas, logrando una mejor inserción en el mercado chileno y aumentando la satisfacción y lealtad de sus clientes.

4 Metodología

Para abordar la segmentación del mercado de libros en Chile y diseñar estrategias de marketing efectivas para Equis Libros, implementaremos una metodología que combine varias técnicas vistas en clase. Estas nos permitirán explorar y comprender mejor los diferentes segmentos del mercado, así como las características y comportamientos de los clientes.

El primer paso en nuestra metodología será realizar un análisis exploratorio de datos. Esta etapa inicial implica examinar detalladamente los datos demográficos y de comportamiento obtenidos tanto de Mercado Libre como de las encuestas demográficas. Durante esta etapa, utilizaríamos diversas técnicas de visualización de datos, como histogramas y gráficos de dispersión para identificar tendencias y patrones iniciales. Este análisis preliminar es crucial, ya que nos permite detectar cualquier valor atípico en los datos, comprender las relaciones básicas entre las variables y obtener una visión general de la distribución de los clientes, sus preferencias de lectura y comportamientos de compra. Por ejemplo, podríamos identificar si ciertos grupos de edad tienen una mayor tendencia a comprar libros de ciertos géneros o si existen patrones de compra estacionales, al igual que si hay patrones de compras en una determinada región geográfica.

En casos donde haya una gran cantidad de observaciones, como puede ser en los casos de analizar la demografía, podríamos utilizar análisis de componentes principales (PCA), para la reducción de dimensionalidad. Esto nos permitiría identificar las características principales que influyen en el comportamiento del cliente. Mediante esta técnica transformaríamos un gran conjunto de variables en un conjunto más pequeño de componentes principales que explican la mayor parte de la variabilidad en los datos. Por ejemplo, si tenemos datos sobre múltiples aspectos de comportamiento de compra y características demográficas, nos ayudaría a destacar las variables más significativas, como edad, frecuencia de compra y géneros de libros preferidos.

Sin embargo, el análisis más importante para la segmentación del mercado será el análisis de clúster, utilizando específicamente el método de k-medias. Este método agrupa a los clientes en segmentos homogéneos basados en características demográficas y de comportamiento similares. En el análisis de clúster k-medias, selecciona un número predefinido de clústeres (k) y los datos se asignan a estos clústers de manera que se minimice la variabilidad dentro de cada grupo y se maximice la variabilidad entre los grupos. Por ejemplo, podríamos agrupar a los clientes en segmentos como "lectores jóvenes de ficción", "adultos que prefieren biografías" y "compradores frecuentes de libros educativos". Al aplicar k-medias, identificaríamos patrones comunes de compra y relaciones entre diferentes pro-

ductos, permitiendo a Equis Libros adaptar sus estrategias de marketing a las necesidades específicas de cada segmento. Esta segmentación nos ayudaría a entender mejor los diferentes tipos de clientes y sus preferencias, facilitando la personalización de ofertas y la predisposición a pagar por cada segmento,

Complementaremos este análisis con técnicas de sentiment analysis para obtener más datos sobre las preferencias y la satisfacción de los clientes. Si disponemos de datos de opiniones de clientes, como reseñas de productos o comentarios en redes sociales, utilizaremos técnicas de procesamiento de lenguaje natural (NLP) para analizar estos textos y extraer información sobre el sentimiento y las opiniones de los clientes. El sentiment analysis nos permitiría identificar temas recurrentes en las opiniones, como la calidad del servicio, la satisfacción con los productos y las preferencias de los clientes. Esta información es crucial para ajustar estrategias y mejorar la experiencia del cliente, ya que proporcionaría una visión más profunda de cómo los clientes perciben a Equis Libros y qué aspectos valoran más.

A su vez, recordemos que mediante los datos complementarios obtenidos de encuestas y estudios de mercado, examinaremos la distribución geográfica de los clientes para identificar posibles diferencias regionales en las preferencias de lectura y el comportamiento de compra. Este análisis permitiría adaptar las estrategias de marketing a las particularidades de cada región, optimizando así las preferencias sobre el mercado chileno. Por ejemplo, podríamos descubrir que ciertos géneros de libros son más populares en ciertas regiones, lo que nos permitiría enfocar las campañas de marketing y promociones en esas áreas específicas.

A su vez, podríamos realizar un análisis complementario en el cual nos centraríamos en identificar las características predictivas de una compra. Para eso utilizaríamos el modelo Logit de clasificación y modelos de regularización. El modelo Logit nos permitiría predecir la probabilidad de que un cliente realice una compra en función de diversas características demográficas y de comportamiento. La regularización, como la regresión de Ridge o Lasso, ayudarían a evitar el sobreajuste y mejora la capacidad del modelo para generalizar a nuevos datos al penalizar la complejidad del modelo. Utilizando este modelo, seleccionaríamos las variables más importantes para predecir si una compra se realiza o no. Por ejemplo, podríamos determinar que la frecuencia de compra y el historial de compras anteriores son los factores más importantes que influyen en la decisión de compra de un cliente. Así, hacer enfoque en los modelos predictivos sería fundamental para diseñar campañas de marketing dirigidas y aumentar la tasa de clientes potenciales.

Finalmente, la combinación de todos estos métodos nos proporcionarían una comprensión profunda y multifacética del mercado de libros en Chile. Esta metodología permitiría

a Equis Libros optimizar sus estrategias de marketing, adaptándose a las necesidades específicas de sus clientes y mejorando su competitividad en el mercado chileno.

5 Conclusiones y limitaciones

La implementación de una metodología avanzada de análisis de datos para segmentar el mercado de libros en Chile ofrece a Equis Libros una oportunidad significativa para optimizar sus estrategias de marketing y mejorar su competitividad en un entorno de mercado altamente competitivo. A través del uso de análisis exploratorio de datos, análisis de componentes principales (PCA), análisis de clúster (k-medias), sentiment analysis y modelos logísticos y de regularización, esperaríamos obtener una comprensión profunda y detallada de los diferentes segmentos de clientes y sus comportamientos de compra.

Se espera encontrar varios resultados clave. En primer lugar, la identificación de segmentos de mercado claramente definidos basados en características demográficas y comportamientos de compra, como podrían ser grupos de jóvenes lectores interesados en ficción, adultos que prefieren biografías y compradores frecuentes de libros educativos. En segundo lugar, patrones de compra específicos que revelen las preferencias de los clientes por ciertos géneros, autores y editoriales. En tercer lugar, información valiosa sobre la satisfacción del cliente y sus opiniones a través del sentiment analysis, lo que permitirá ajustar las estrategias de marketing y mejorar la experiencia del cliente. Por último, el análisis sobre la ubicación geográfica proporcionaría una visión clara de las diferencias regionales en las preferencias de lectura y comportamientos de compra, facilitando la personalización de campañas de marketing por región.

A pesar de los beneficios potenciales, este proyecto enfrenta varias limitaciones. La principal limitación es la dependencia de la calidad y disponibilidad de los datos de Mercado Libre y de fuentes externas. Si los datos proporcionados por los usuarios de Mercado Libre son limitados o no contienen información detallada, esto podría afectar la precisión del análisis.

Otra limitación es la posibilidad de sesgos en los datos recopilados. Los datos de comportamiento de compra reflejan solo las acciones de los clientes en la plataforma de Mercado Libre, lo que podría no representar completamente sus preferencias de lectura en otros contextos. Asimismo, el sentiment analysis basado en opiniones y comentarios de clientes puede estar influenciado por una minoría, lo que no necesariamente refleja la opinión de la mayoría de los clientes.

Por último, el uso de modelos predictivos, como el modelo Logit y de regularización, depende de la calidad de los datos históricos y puede no siempre capturar todos los factores

que influyen en la decisión de compra de un cliente. Los modelos pueden necesitar ajustes y validación continua para mantener su precisión y relevancia a medida que cambian las tendencias del mercado y el comportamiento del cliente.

En conclusión, aunque la metodología propuesta ofrece un enfoque sólido para segmentar el mercado de libros en Chile y diseñar estrategias de marketing personalizadas, es crucial considerar las limitaciones y desafíos potenciales. Para llevar a cabo esta propuesta es fundamental gestionar cuidadosamente los datos, la validación de los modelos y la continua adaptación a nuevos datos y tendencias, aspectos esenciales para el éxito del propuesta.

6 Referencias

- Caicedo Consuegra, L. D., Márquez Vásquez, P. A., & Meza Pérez, A. M. (2023). Artificial intelligence algorithms based on socio-behavioral profiles for intelligent customer segmentation: Case study. *Ingeniería y competitividad*, 25(3).
- Casariego Sarasquete, N. M. (2017). A common data representation model for customer behavior tracking. *Icono 14*, 15(2), 55-91. doi: 10.7195/ri14.1078v15i2.1078
- Fiedler, Roth, M., & Vetvik, O. J. (2016). Why customer analytics matter. *McKinsey & Company, Zugriff am*, 5.
- Instituto Nacional de Estadísticas de Chile (INE). (n.d.). Estadísticas. Instituto Nacional de Estadísticas de Chile. <https://www.ine.gob.cl/estadisticas/>
- Zerboni, F., Silva, J., & Prado, M. (2013). Gestión de cuentas clave desde la perspectiva del cliente. *IEEM Revista de Negocios*, 16(3), 56-62.