



UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
Departamento de Ingeniería Comercial

**ANÁLISIS DEMOGRÁFICO Y UBICACIÓN ESTRATÉGICA
DE LOCALES COMERCIALES: OPTIMIZACIÓN PARA LA
EXPANSIÓN DE FRANQUICIAS**

Nataly Paz Fernández González
INGENIERÍA COMERCIAL
Diciembre 2024



UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
Departamento de Ingeniería Comercial

**ANÁLISIS DEMOGRÁFICO Y UBICACIÓN ESTRATEGICA
DE LOCALES COMERCIALES: OPTIMIZACIÓN PARA LA
EXPANSIÓN DE FRANQUICIAS**

Tesis presentada por
Nataly Paz Fernández González
como requisito para optar al título de
INGENIERO COMERCIAL

Profesor Correferente: **Macarena Gatica S.**

Director de Tesis: **Francesco Pontarelli C.**

Diciembre, 2024

AGRADECIMIENTOS

Al finalizar el presente trabajo, y con ello cerrar un importante ciclo en mi vida universitaria, quiero expresar mis más sinceros agradecimientos a todas aquellas personas que hicieron posible este logro.

A mis padres, por ser mi pilar fundamental en este proceso, quienes con su amor me enseñaron el valor del esfuerzo y la perseverancia. Su confianza han sido mi mayor fortaleza. Este logro también es suyo.

A mis hermanos, por ser fuente constante de apoyo, inspiración y alegría. Su presencia en mi vida me recuerda cada día que el esfuerzo vale la pena cuando se tiene con quién compartir los logros.

Agradezco también a mis profesores y compañeros, quienes me brindaron conocimientos, guía y colaboración a lo largo de estos años.

Finalmente, agradezco a la vida por darme la oportunidad de aprender, crecer y superar los momentos difíciles. Esta etapa llega a su fin, pero me llevo conmigo aprendizajes que me acompañaran siempre.

“Este logro no es el fin, sino el comienzo de nuevos desafíos y sueños por alcanzar”.

RESUMEN EJECUTIVO

Este estudio se centra en el análisis demográfico y la optimización de la ubicación estratégica de locales comerciales, específicamente en la comuna de Viña del Mar, Región de Valparaíso, Chile. A través de modelos predictivos y técnicas de análisis, la investigación busca identificar las mejores ubicaciones para la expansión de franquicias, considerando factores clave como las características demográficas, económicas y de consumo en Viña del Mar.

La metodología empleada combina datos provenientes del censo 2002 y 2017, registros municipales, base de datos operacionales y logísticas proporcionadas por plataformas como PedidosYa. Estos datos se integraron mediante herramientas avanzadas de análisis, como modelos de regresión lineal múltiple, permitiendo la creación de modelos predictivos de alto impacto.

Entre los principales hallazgos el análisis reveló que las variables relacionadas a la preferencia en tipos de comida, en locales comerciales y en tasa de conversión de cliente, tienen un impacto significativo en la selección de ubicaciones de franquicias. Los resultados confirman la relevancia del análisis demográfico y la integración de datos de mercado para optimizar decisiones estratégicas de expansión.

Finalmente, este trabajo proporciona conclusiones y recomendaciones específicas. La investigación no solo resalta la importancia de un enfoque basado en datos, sino que también establece una base para futuras investigaciones en localización estratégica.

ABSTRACT

This study focuses on the demographic analysis and optimization of the strategic location of commercial establishments, specifically in the city of Viña del Mar, Valparaíso Region, Chile. Through predictive models and analytical techniques, the research aims to identify the best locations for franchise expansion, considering key factors such as demographic, economic, and consumer characteristics in the city of Viña del Mar.

The methodology used combines data from the 2002 and 2017 census, municipal records, and operational and logistics databases provided by platforms like PedidosYa. These data were integrated using advanced analysis tools, such as multiple linear regression models, allowing for the creation of high-impact predictive models.

Among the main findings, the analysis revealed that variables related to food preferences, commercial establishments, and customer conversion rates have a significant impact on franchise location selection. The results confirm the relevance of demographic analysis and the integration of market data to optimize strategic expansion decisions.

Finally, this work provides specific conclusions and recommendations. The research not only highlights the importance of a data-driven approach but also lays the foundation for future studies on strategic location planning.

ÍNDICE

	Páginas
1. INTRODUCCIÓN	11
2. ORIGEN Y PROPÓSITO DEL ESTUDIO	12
2.1 OBJETIVOS	13
2.1.1 Objetivo General.....	13
2.1.2 Objetivo Específicos.....	13
2.2 Alcance del estudio	13
2.2.1 Limitaciones.....	14
3. Marco Teórico	15
3.1 Franquicias	15
3.1.1 Revisión bibliográfica del modelo de franquicias.....	15
3.1.2 Factores clave en la localización de franquicias.....	17
3.2 Franquicias en Chile	17
3.2.1 Beneficios y desafíos	18
3.3 Modelo de delivery en franquicias	19
3.3.1 Factores clave del éxito en el modelo de delivery.....	19
3.3.2 Desafíos del modelo de delivery	20
3.3.3 Impacto del delivery en la localización de franquicias	21
3.4 Competencia en el sector de las franquicias	21
3.4.1 Factores diferenciadores en la competencia.....	23
3.4.2 Desafíos competitivos y regulación	24
3.5 Métodos de pronóstico	24
3.5.1 Geomarketing y modelos de localización en la expansión de franquicias.....	24
3.5.2 Métodos de localización de nuevos comercios.....	24
3.5.3 Geomarketing y los sistemas de información geográfica (SIG)	25
3.5.4 Modelos de localización y sus antecedentes	26
3.5.5 Modelos de ubicación	27
3.5.6 Modelos de regresión lineal múltiple	29
3.6 Análisis de variables	30
3.6.1 Indicadores demográficos	30
3.6.2 R-Cuadrado.....	30
3.6.3 R-Cuadrado Ajustado	31
3.6.4 Valor P.....	31
3.6.5 F-Statistic.....	32
3.6.6 Prueba Durbin- Watson	32
3.6.7 Coeficientes de las variables independientes	32
3.7 Monetización de datos	33
4. Propuesta de la metodología de trabajo	34
4.1 Fases de la propuesta metodológica	35

4.2 Diagrama de propuesta metodológica	36
5. Aplicación metodológica	37
5.1 Recopilación de los datos	37
5.2 Tamaño muestra	38
5.3 Análisis de las variables.....	39
5.3.1 Descripción de las variables	39
5.3.2 Análisis demográfico	41
5.4 Aplicación de Modelo regresión lineal múltiple	48
5.4.1 Modelo 1	49
5.4.2 Modelo 2	51
5.4.4 Modelo 3	52
5.4.5 Modelo 4	54
6. Resultados	56
6.1 Comparación Modelos.....	56
6.2 Aplicación de datos en Modelo 1	56
7. Conclusiones y recomendaciones	58
ANEXOS	60
Cálculos Adicionales: Modelo 1	60
Cálculos Adicionales: Modelo 2	60
Cálculos Adicionales: Modelo 3	60
Cálculos Adicionales: Modelo 4	61
Código python.....	61
BIBLIOGRAFÍA	65
WEBGRAFÍA	67

ÍNDICE TABLAS

	Páginas
<i>Tabla 1: Descripción de variables.</i>	41
<i>Tabla 2: Cálculos análisis descriptivo por variables.</i>	48
<i>Tabla 3: Variables que componen los modelos.</i>	49

ÍNDICE DE ILUSTRACIONES

	Páginas
<i>Ilustración 1: Origen de las franquicias a través del tiempo.....</i>	16
<i>Ilustración 2: Diagrama metodología propuesta.....</i>	36
<i>Ilustración 3: Variación Porcentual Población Total censo 2002 7 2017.</i>	43
<i>Ilustración 4: Resultados Regresión Lineal Múltiple Modelo 1.....</i>	51
<i>Ilustración 5: Resultados Regresión Lineal Múltiple Modelo 2.....</i>	52
<i>Ilustración 6: Resultados Regresión Lineal Múltiple Modelo 3.....</i>	54
<i>Ilustración 7: Resultados Regresión Lineal Múltiple Modelo 4.....</i>	56

ÍNDICE DE GRAFICOS

	Páginas
<i>Gráfico 1: Población Total por Preferencia_1</i>	44
<i>Gráfico 2: Distribución por Género Preferencia_1</i>	44
<i>Gráfico 3: Población Total por Preferencia_2</i>	45
<i>Gráfico 4: Distribución por Género Preferencia_2</i>	45
<i>Gráfico 5: Población Total por Preferencia_3</i>	46
<i>Gráfico 6: Distribución por Género Preferencia_3</i>	46

1. INTRODUCCIÓN

La presente investigación aborda el análisis demográfico y la identificación de ubicaciones estratégicas para la instalación de locales comerciales, con un enfoque en la optimización de modelos predictivos aplicados al crecimiento de franquicias. Este estudio se centra en la comuna de Viña del Mar, en la Región de Valparaíso, Chile, y busca proporcionar información relevante a propietarios de franquicias interesados en expandir sus operaciones en esta zona.

Para cumplir con este objetivo, se utilizó una muestra de conveniencia compuesta por 10 observaciones, seleccionada en función de la disponibilidad de datos específicos sobre preferencias de consumo, características demográficas y patrones de compra en diferentes zonas de la ciudad. Los datos fueron obtenidos de fuentes oficiales como el censo 2002 y 2017, complementados con base de datos operacionales y logísticos provistas por plataformas como PedidosYa, y procesados mediante herramientas de análisis estadístico avanzadas.

Los resultados preliminares permiten comprender mejor los factores clave que inciden en la elección de ubicaciones óptimas para la expansión de franquicias, combinando variables como la densidad poblacional, el gasto promedio en delivery y la tasa de conversión de clientes. A partir de estos hallazgos se ofrecen conclusiones que orientan tanto a los emprendedores como a las empresas en la toma de decisiones estratégicas de localización.

Esta tesis invita al lector a profundizar en las metodologías empleadas, los análisis realizados y las recomendaciones derivadas, con el propósito de comprender cómo los datos pueden convertirse en un recurso valioso para la expansión exitosa de franquicias en mercados locales.

2. ORIGEN Y PROPÓSITO DEL ESTUDIO

Durante mi práctica profesional, pude identificar el enorme potencial que existe en los datos de logística y operaciones para optimizar las rutas de entrega e identificar patrones de consumo. Con este enfoque, en el análisis de datos se convierte en una herramienta clave para reducir costos operativos, mejorar la experiencia del cliente, y en última instancia, aumentar la rentabilidad de la empresa.

El proceso comenzó con la recopilación y análisis de grandes volúmenes de datos relacionados con el comportamiento de los cliente, tanto las preferencias en cuanto al tipo de comida y comercios. Complementándose con datos demográficos extraídos del Censo de los años 2002 y 2017. A partir de esta información, se podrían generar modelos predictivos capaces de identificar qué zonas es más óptima para la instalación de franquicias.

La monetización de estos datos es un componen clave de este enfoque, ya que se podría generar valor no solo dentro de la empresa, sino también al ofrecer información valiosa a los socios comerciales, como restaurantes, comercios y dueños de franquicias.

El estudio busca abrir nuevas oportunidades de negocio mediante el aprovechamiento estratégico de los datos generados en el día a día de las operaciones.

2.1 OBJETIVOS

2.1.1 Objetivo General

Determinar la ubicación estratégica de la instalación de locales comerciales en la comuna de Viña del Mar en la Región de Valparaíso, Chile, mediante modelos predictivos, con el fin de proporcionar información valiosa a los propietarios de franquicias para optimizar sus decisiones de expansión.

2.1.2 Objetivo Específicos

- 1) Realizar un análisis demográfico detallado de Comuna de Viña del Mar en la Región de Valparaíso, Chile.
- 2) Identificar los métodos mayormente utilizados en la actualidad relacionado a la ubicación estratégica, con el fin de construir un modelo que permita la predicción de instalación de franquicias.
- 3) Obtener resultados finales sobre los modelos resultantes de acuerdo con el método aplicado y aportar las conclusiones y recomendaciones pertinentes para el negocio.

2.2 Alcance del estudio

El presente estudio se centrará en la comuna de Viña del Mar, ubicada en la Región de Valparaíso, Chile. Este análisis se realiza mediante el uso de modelos predictivos basados en datos demográficos y económicos de la zona, incluyendo variables como densidad poblacional, distribución por género, ingresos promedio por hogar y tasas de crecimiento poblacional.

La investigación utiliza una muestra seleccionada por conveniencia, limitada por la disponibilidad de datos, lo que permite realizar una evaluación preliminar de los patrones y tendencias relevantes para la localización estratégica. Sin embargo, se reconoce que esta muestra reducida puede limitar la generación de los resultados. Por ello, se recomienda considerar bases de datos más amplias en futuras investigaciones para validar los hallazgos obtenidos.

Los resultados de este estudio están dirigidos principalmente a propietarios de franquicias interesados en expandir sus negocios en Viña del Mar. A través de este trabajo, se busca proporcionarles herramientas prácticas para optimizar sus decisiones estratégicas de ubicación comercial, adaptadas a las características demográficas y socioeconómicas específicas de la comuna.

El alcance temporal del estudio abarca datos actuales y proyecciones demográficas hasta el año 2024, permitiendo evaluar tanto el contexto presente como el potencial de crecimiento futuro de las zonas estudiadas.

2.2.1 Limitaciones

Es importante destacar que el análisis se basa en una muestra reducida, seleccionada por conveniencia debido a la disponibilidad de datos. Esta decisión, aunque útil para evaluar tendencias y desarrollar modelos iniciales, puede limitar la representatividad de los resultados para toda la comuna. Por ellos, se sugiere que futuros estudios consideren la evaluación de los modelos propuestos utilizando una base de datos más amplia y representativa. Esto permitirá no solo validar las conclusiones obtenidas, sino también ampliar su aplicabilidad a otras regiones o contextos similares.

3. Marco Teórico

3.1 Franquicias

3.1.1 Revisión bibliográfica del modelo de franquicias

La franquicia es un modelo de negocios en el que un empresario, denominado franquiciante, otorga a un tercero, el franquiciado, el derecho a utilizar su marca, know-how y sistemas comerciales a cambio de ciertos pagos. Este modelo se basa en una relación de beneficio mutuo: el franquiciante recibe una tarifa inicial y pagos continuos, mientras que el franquiciado se beneficia del uso de una marca reconocida, el acceso a conocimientos técnicos y apoyo en aspectos como contabilidad, marketing y manejo de personal. Este tipo de negocio permite que emprendedores sin experiencia previa puedan iniciar su propio negocio bajo la tutela de un empresario experimentado.

El origen de las franquicias modernas se remonta a mediados del siglo XIX destacando el caso de Singer Sewing Machine en 1850, que implementó una red de concesionarios para reducir los costos de venta directa. Aunque este modelo duró solo una década, sentó las bases para el sistema actual de franquicias, centrado en la concesión de productos y marcas. Posteriormente, empresas como General Motors y Coca-Cola adoptaron modelos similares, consolidando la franquicia como una estrategia viable para expandir mercados sin asumir grandes riesgos financieros.

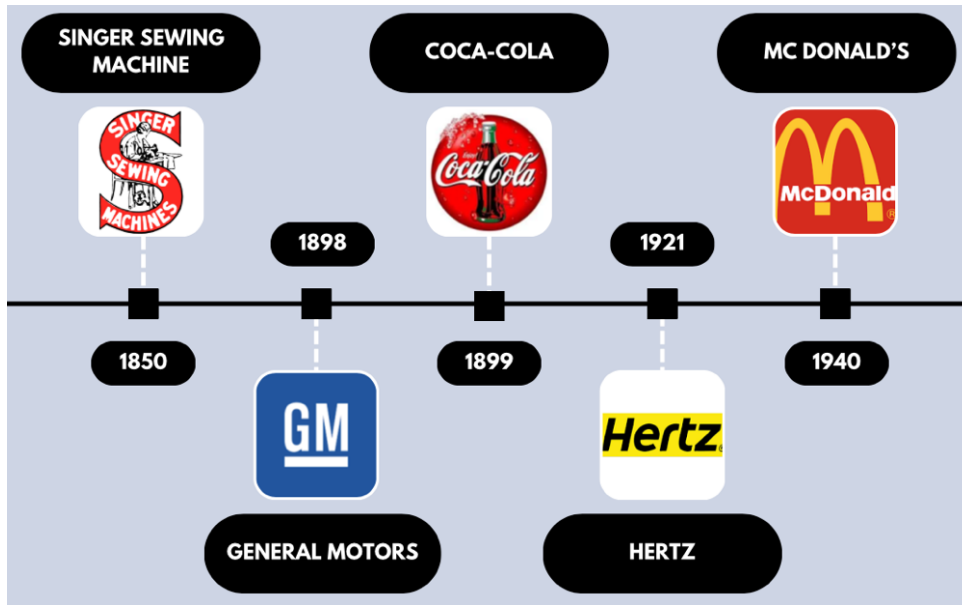


Ilustración 1: Origen de las franquicias a través del tiempo.

Fuente: Cruz soto, n. v. (2016). Análisis descriptivo de las franquicias en Chile.

A lo largo del tiempo, las franquicias han evolucionado, originando distintos tipos:

- Franquicia de producto o marca registrada: Radica principalmente en la concesión de derechos, permitiendo al franquiciado usar una marca o nombre comercial, y el franquiciante actúa como proveedor exclusivo.
- Franquicia de formato de negocios: El franquiciado adopta el método de trabajo y el know-how del franquiciante, garantizando uniformidad en las condiciones de calidad y el servicio.
- Franquicia por conversión: Empresas establecidas se adhieren a franquicias grandes, beneficiándose del reconocimiento de marca.
- Franquicia social: Similar a las franquicias comerciales, pero sin fines de lucro.

3.1.2 Factores clave en la localización de franquicias

La localización es uno de los aspectos más importantes en el éxito de una franquicia. Las franquicias suelen establecerse en centros comerciales debido a la alta afluencia de público y la concentración de tiendas del sector retail. La elección adecuada del lugar es esencial para asegurar el flujo de clientes y maximizar las oportunidades de crecimiento. Además, en la actualidad, la integración de servicios de delivery ha ampliado las posibilidades para las franquicias, permitiendo atender a una mayor cantidad de consumidores a través de plataformas digitales.

3.2 Franquicias en Chile

El desarrollo de las franquicias en Chile es relativamente reciente, con la llegada de McDonald's en 1990 como uno de los pioneros en este modelo de negocio. Desde entonces, el número de franquicias ha crecido considerablemente, estimándose alrededor de 208 marcas que operan bajo este formato. Sin embargo, Chile carece de una entidad reguladora formal para este tipo de negocios, lo que ha dificultado la obtención de datos precisos sobre el sector. En 2014, la cámara de comercio de Santiago organizó la primera feria de franquicias, lo que marcó un hito en la consolidación de este modelo en el país.

En cuanto a la situación actual, el mercado de franquicias en Chile ha experimentado un crecimiento sostenido en los últimos años, consolidándose como una opción atractiva tanto para emprendedores como para grandes marcas internacionales. En 2023, se ha destacado un fortalecimiento del sector debido a la creciente adopción de tecnología y la recuperación económica tras la pandemia.

Por el lado del crecimiento y evolución del sector, según la cámara de comercio de Santiago (CCS), durante el periodo de la pandemia, las franquicias en Chile crecieron un 7%, con un aumento significativo en locales propios y franquiciados. La mayor parte de este crecimiento se dio en el sector de comercio, aunque las

ventas mostraron un comportamiento dispar dependiendo del rubro. Sin embargo, el 83% de los franquiciantes proyecta un crecimiento en ventas para los próximos años, lo que evidencia el optimismo en el sector (CCS, 2023).

En cuanto a los sectores más rentables, destacan los siguientes:

- **Gastronomía:** La franquicia de alimentos sigue siendo un sector dominante. marcas como NIU Sushi y Doggi's han expandido significativamente su presencia en todo el país.
- **Belleza y Bienestar:** Franquicias como Eva Spa han demostrado ser altamente rentables, aprovechando el creciente interés por el cuidado personal.
- **Tecnología y Servicios Digitales:** Con el auge de los servicios online, franquicias relacionadas con tecnología, como Gamer's Up Online, también están ganando popularidad.

Por otro lado, la innovación tecnológica es uno de los principales factores que impulsa el crecimiento del sector. Las franquicias que adoptan tecnologías avanzadas y automatización en sus procesos son más competitivas y resilientes ante cambios económicos (Paideia, 2023).

3.2.1 Beneficios y desafíos

El modelo de franquicia en Chile ofrece varios beneficios, como la reducción de riesgos financieros, el acceso a redes de apoyo comercial, y la capacitación constante para los franquiciados. Estas ventajas han hecho del sistema de franquicias una opción atractiva para emprendedores que buscan minimizar riesgos y operar con marcas consolidadas (Latam networks, 2024).

A pesar de este crecimiento, algunos desafíos persisten, como la necesidad de adaptar los modelos de negocio a las nuevas tecnologías y los cambios en el comportamiento del consumidor. Las franquicias que no se adapten a estos cambios

tecnológicos podrían enfrentar dificultades en los próximos años (100franquicias, 2023).

3.3 Modelo de delivery en franquicias

El modelo de delivery ha ganado relevancia en las últimas décadas, especialmente en el sector de las franquicias, al ofrecer una solución eficiente y adaptada a las nuevas demandas del consumidor moderno. Este modelo permite la entrega de productos directamente al cliente, ampliando el alcance de las franquicias sin la necesidad de expandir su infraestructura física.

El delivery ha experimentado una evolución significativa, particularmente en el ámbito de las franquicias de comida rápida. Según estudios realizados por Franco y Moreno (2013), en ciudades como Medellín, las dinámicas del delivery han fortalecido el modelo de franquicias, haciendo necesario un plan de negocios que responda a las necesidades del consumidor actual. Caviedes (2013) resalta que la clave del éxito empresarial radica en contar con un modelo de negocio sólido que soporte la demanda de entrega a domicilio.

La pandemia de Covid-19 aceleró la adopción del modelo de delivery, convirtiéndose en una necesidad más que en una opción. En este contexto, las franquicias que implementaron servicios de entrega a domicilio, como Uber Eats, Rappi, y PedidosYa, experimentaron un crecimiento significativo. Estas plataformas no solo facilitaron la entrega de alimentos, sino que también diversificaron sus servicios para incluir productos de supermercados, farmacias y otros establecimientos, logrando una mayor penetración en el mercado.

3.3.1 Factores clave del éxito en el modelo de delivery

El éxito del modelo de delivery en franquicias depende de varios factores, entre los que destacan:

- **Innovación y adaptación:** La capacidad de las franquicias para adaptarse a los cambios en el comportamiento del consumidor es crucial. Según Batista, Bolívar y Medina (2016), la innovación en los modelos de negocio, como el delivery, es fundamental para mejorar la relación con los clientes, más allá de centrarse solo en el producto o servicio ofrecido. Las franquicias que se enfocan en mejorar la experiencia del cliente mediante entregas rápidas y eficientes han logrado un crecimiento sostenido.
- **Calidad del servicio:** La calidad en la atención al cliente es un elemento esencial en el éxito del delivery. Como señala Castillo et al. (2012), la satisfacción del cliente con el servicio de entrega está directamente relacionada con la rapidez y efectividad del servicio. El uso de plataformas digitales, así como la atención directa a través de canales como teléfono o WhatsApp, ha sido clave para mejorar la percepción del cliente sobre el servicio de entrega.
- **Impacto de la tecnología:** Las plataformas de delivery no solo han facilitado la conexión entre las franquicias y los consumidores, sino que también han creado nuevas oportunidades para las pequeñas y medianas empresas. Según Núñez (2020), el uso de estas plataformas permite a las franquicias ajustarse rápidamente a las demandas del mercado, lo que ha sido particularmente útil durante la pandemia.

3.3.2 Desafíos del modelo de delivery

A pesar de sus beneficios, el modelo de delivery también enfrenta ciertos desafíos. Uno de los principales problemas es la precariedad laboral que enfrentan los repartidores, quienes frecuentemente trabajan en condiciones laborales difíciles, sin acceso a seguridad social ni derechos laborales básicos. Hurtubia et al. (2021) advierten sobre la necesidad de regular este sector para garantizar condiciones laborales dignas, así como para mitigar los impactos urbanos negativos que el

crecimiento de estas plataformas ha causado, como el aumento del tráfico y la contaminación.

Otro desafío es el costo elevado que las plataformas de delivery imponen a los establecimientos, con comisiones que pueden llegar al 30% del valor de cada pedido. Esto ha llevado a que algunos restaurantes y franquicias reconsideren la opción de implementar sistemas de entrega propios para reducir los costos operativos y mantener precios competitivos.

3.3.3 Impacto del delivery en la localización de franquicias

La adopción del modelo de delivery ha tenido un impacto directo en la localización de las franquicias. Si bien la ubicación física sigue siendo importante para captar a los consumidores que prefieren realizar sus compras en persona, el delivery ha permitido a las franquicias expandir su área de cobertura sin la necesidad de abrir nuevos puntos de venta. Esto ha sido especialmente beneficioso en mercados densamente poblados o en áreas donde el acceso a ciertos productos es limitado.

La flexibilidad que ofrece el delivery permite que las franquicias puedan operar en localidades con menor densidad comercial o en zonas alejadas de los grandes centros urbanos, donde antes no sería rentable establecer una franquicia. En estos casos, la clave del éxito radica en la eficiencia del servicio de entrega y la capacidad de las franquicias para adaptarse a las nuevas dinámicas del mercado.

3.4 Competencia en el sector de las franquicias

El panorama competitivo en el sector de las franquicias es dinámico y está en constante evolución, particularmente en industrias como la restauración y el comercio digital. A medida que las franquicias se expanden geográficamente, enfrentan una competencia creciente, tanto a nivel local como internacional. Esta

competencia no solo proviene de otras franquicias, sino también de negocios independientes que utilizan tecnologías y estrategias similares.

Uno de los actores más destacados en la industria del delivery y franquicias en América Latina es PedidosYa, que fue fundada en 2009 en Uruguay. Esta plataforma de delivery ha evolucionado rápidamente para convertirse en una de las líderes del mercado, operando en 15 países y colaborando con más de 110,000 comercios asociados y 60,000 repartidores. La clave de su éxito radica en la innovación continua, introduciendo servicios como PedidosYa Market (mercados 100% digitales), PedidosYa envíos (servicio de última milla), y PedidosYa Pay (billetera virtual), lo que les ha permitido diversificar su oferta y mantenerse competitivos en un mercado saturado.

En Chile, PedidosYa ha jugado un rol significativo desde su entrada en 2010, integrando más de 6,000 tiendas y generando empleo para 2,000 distribuidores. Su capacidad para adaptarse a las demandas locales y su enfoque en la expansión han sido cruciales para su crecimiento.

Otro competidor clave en este sector es Uber Eats, lanzado en 2014 como una extensión de la empresa Uber, que previamente se dedicaba exclusivamente al transporte de pasajeros. Uber Eats se ha expandido rápidamente, llegando a 24 ciudades en Chile, con más de 13,000 restaurantes y tiendas asociadas en la plataforma. La ventaja competitiva de Uber Eats radica en su infraestructura tecnológica avanzada y su capacidad para ofrecer tiempos de entrega más rápidos que muchos competidores locales.

Rappi, fundada en 2015 en Colombia, es otra plataforma importante en la competencia de delivery en América Latina. En Chile, Rappi ha logrado una cobertura significativa, con presencia en 15 ciudades y más de 10,000 comercios asociados. Su estrategia se basa en ofrecer una experiencia de "SuperApp",

permitiendo a los usuarios acceder a una amplia gama de productos y servicios, desde la entrega de alimentos hasta servicios financieros.

Además de las plataformas digitales de delivery, las franquicias tradicionales también enfrentan una fuerte competencia. Por ejemplo, las cadenas de comida rápida y restaurantes establecidos en centros comerciales compiten por la atención del mismo segmento de consumidores. En este contexto, la capacidad de adaptación y diferenciación se vuelve fundamental para el éxito. La localización sigue siendo un factor crucial, ya que estar en lugares de alta afluencia, como los centros comerciales, permite captar un mayor flujo de clientes.

3.4.1 Factores diferenciadores en la competencia

Uno de los factores diferenciadores en el mercado de las franquicias es la innovación en el servicio. Las plataformas de delivery como PedidosYa y Rappi han implementado nuevas soluciones, como la entrega en última milla y mercados 100% digitales, que les permiten destacarse frente a otros competidores. Adicionalmente, la experiencia del usuario juega un papel esencial. Las plataformas que ofrecen una experiencia de compra más rápida y conveniente, y que minimizan el contacto físico a través de aplicaciones móviles, logran captar un mayor porcentaje del mercado.

Otro factor clave es la diversificación de servicios. Las plataformas que ofrecen una gama más amplia de productos y servicios (no solo comida, sino también productos farmacéuticos, de supermercado, etc.) Tienen una ventaja competitiva significativa. Esta diversificación les permite acceder a un segmento de mercado más amplio y aprovechar oportunidades que otras franquicias más especializadas no pueden alcanzar.

3.4.2 Desafíos competitivos y regulación

Si bien las plataformas de delivery han aportado un nivel elevado de competencia en el sector de las franquicias, también han enfrentado críticas relacionadas con la precariedad laboral y la falta de regulación. Los repartidores en Chile, al igual que en otros países, frecuentemente trabajan en condiciones desfavorables, sin acceso a beneficios sociales o seguridad laboral. Este factor puede convertirse en un obstáculo competitivo, ya que algunos consumidores y comercios pueden preferir volver a métodos tradicionales de entrega que no dependan de estas plataformas.

Además, el costo elevado de utilizar estas plataformas, con comisiones que oscilan entre el 15% y el 30%, representa un desafío tanto para los restaurantes como para los consumidores, quienes ven incrementado el precio final de los productos. Esto ha llevado a algunos establecimientos a reconsiderar sus estrategias de delivery, optando por soluciones más económicas y controladas internamente.

3.5 Métodos de pronóstico

3.5.1 Geomarketing y modelos de localización en la expansión de franquicias

Los datos geográficos permiten segmentar regiones en función de características demográficas como edad, escolaridad y religión, lo que a su vez define tipologías psicosociológicas, tales como la clase social y los estilos de vida, los cuales influyen en el comportamiento de consumo (Alcaide et al., 2012). Estas segmentaciones geográficas son cruciales para el diseño de estrategias de mercadeo, así como para la selección de áreas óptimas donde abrir nuevas sucursales (Doyle, 2001). El geomarketing, como disciplina, aborda precisamente la localización de nuevos sitios comerciales, basándose en el análisis espacial para maximizar el éxito de las franquicias y reducir el riesgo asociado con la expansión.

3.5.2 Métodos de localización de nuevos comercios

Históricamente, en los años setenta, los ejercicios de localización se basaban más en la intuición que en modelos formales de localización-asignación. Los tomadores de decisiones utilizaban varios métodos, muchos de los cuales se siguen empleando hoy en día:

1. Listas de Comprobación: Este método consiste en identificar y enlistar los factores más importantes para seleccionar un sitio adecuado para la apertura de una sucursal. Factores como la accesibilidad, la potencial expansión, la situación económica de la población y la presencia de competidores son considerados y calificados de forma ordinal para los diferentes sitios alternativos. La opción que obtenga la mayor suma de calificaciones será la más apropiada para la instalación del negocio.
2. Análisis de Similitud: En este enfoque, se elige una sucursal existente que tenga el mejor rendimiento (en ventas o ganancias) y se buscan ubicaciones con características similares. Este método no se enfoca en entender por qué una tienda es más exitosa que otra, sino en identificar el entorno en el que se encuentra para replicar el éxito en nuevas zonas con contextos urbanos similares (Douard, 2006).
3. Áreas de Influencia: Basado en la proximidad, este método considera que a menor distancia entre el cliente y la tienda, mayor será el número de clientes atraídos. Para hacerlo más realista, se utilizan tiempos de recorrido en lugar de la distancia física, integrando tanto los tiempos a pie como en automóvil. Esta metodología también se apoya en datos sobre las características de los clientes, como el monto y frecuencia de las compras, que pueden obtenerse mediante encuestas o información interna de la empresa (Alcaide et al., 2012).

3.5.3 Geomarketing y los sistemas de información geográfica (SIG)

El uso de sistemas de información geográfica (SIG) ha revolucionado el geomarketing al permitir la combinación de datos internos generados por la empresa

con datos externos, provenientes de fuentes como el gobierno o instituciones académicas. Estos datos permiten analizar de manera detallada el entorno del negocio, incluyendo la localización de clientes, competidores, redes de transporte y vialidades (Chasco, 2003).

En particular, los datos internos, como las ventas totales, la frecuencia de compra, el nivel de satisfacción y las direcciones postales de los clientes, se integran con datos externos para realizar análisis de segmentación de clientes y determinar zonas óptimas para la ubicación de nuevas sucursales (Alcaide et al., 2012; Bocalandro y Krauthamer, 2007). Esta información es esencial para la toma de decisiones estratégicas en la expansión de franquicias.

Los SIG permiten analizar simultáneamente la ubicación espacial de elementos como oficinas, clientes, competidores y rutas de acceso, junto con los atributos de estos elementos, como el horario de operación o el tamaño del local. Estas capacidades permiten realizar análisis espaciales complejos, como la identificación de zonas de influencia o la visualización de competidores cercanos (Doyle, 2001).

3.5.4 Modelos de localización y sus antecedentes

Los modelos de localización han sido empleados en una variedad de aplicaciones, desde la ubicación de centrales de abasto hasta el establecimiento de sitios óptimos para el manejo de residuos tóxicos. El propósito general de un análisis de localización es responder a la pregunta: ¿Dónde colocar un nuevo establecimiento? (Plastria, 2002). Este tipo de modelos permiten a las empresas optimizar sus decisiones, tomando en cuenta tanto la localización disponible como los criterios más relevantes para el negocio.

La literatura sobre modelos de localización ofrece una amplia variedad de teorías. Garrocho et al. (2003) identifican cuatro enfoques conceptuales clave para explicar el comportamiento entre oferentes y consumidores:

- a) Enfoque microeconómico, que asume que los consumidores buscan maximizar su utilidad, mientras que los empresarios buscan maximizar sus beneficios.
- b) Teoría del lugar central, desarrollada por W. Christaller en 1933, que explica la distribución espacial y el tamaño de los asentamientos en función de la localización de actividades terciarias.
- c) Teoría de la interacción espacial, basada en el principio de gravitación comercial de Reilly (1931), que postula que el flujo de consumidores está relacionado positivamente con la población de una localidad y negativamente con la distancia entre localidades.
- d) Teoría de la subasta del suelo urbano, que se basa en el concepto de la renta del suelo para determinar la localización de actividades económicas.

3.5.5 Modelos de ubicación

La selección de ubicaciones estratégicas para negocios es un proceso complejo que involucra diversas variables y modelos analíticos. Entre los principales enfoques utilizados en la investigación, se destacan los siguientes modelos y variables clave.

3.5.5.1 Variables utilizadas en modelos de ubicación estratégica

La demografía y el tráfico son factores clave para determinar la viabilidad de una ubicación comercial. Aspectos como la densidad de población, el perfil demográfico, y el tráfico peatonal o vehicular son fundamentales para evaluar el potencial de una zona. Un alto tráfico peatonal, por ejemplo puede incrementar las oportunidades de atraer clientes, siempre y cuando el perfil de estos coincida con el público objetivo de la empresa (Xmap, 2023). Además, la proximidad a la competencia del local es otra variable crítica en la selección de una ubicación. En algunos casos, estar cerca de competidores o negocios complementarios puede ser beneficioso, ya que puede atraer a más clientes, especialmente en zonas comerciales o de entretenimiento.

Sin embargo, en otros casos, puede ser más estratégico alejarse de áreas saturadas para evitar la competencia directa y crear una oferta diferenciada (Xmap, 2023).

La accesibilidad también juega un papel fundamental en la elección de una ubicación comercial. La disponibilidad de transporte público y las buenas vías de acceso son esenciales, especialmente en áreas urbanas densamente pobladas. Igualmente, la proximidad a infraestructuras clave, como centro de distribución, estaciones de transportes o aeropuertos, es importante según el tipo de negocio, ya que puede facilitar la logística y mejorar la conectividad con clientes y proveedores (Perez-Benitez et al., 2021).

3.5.5.2 Modelos utilizados para la determinación de ubicaciones estratégicas

El Análisis multicriterio es un enfoque que permite a los gerentes evaluar múltiples factores simultáneamente para priorizar las mejores ubicaciones comerciales. Factores como el acceso a tecnología, el clima, la seguridad y la estabilidad política son esenciales en este tipo de análisis. Uno de los métodos más utilizados es el Analytic Hierarchy Process (AHP), que facilita la ponderación de diversas variables en función de su relevancia para los objetivos del negocio. Este modelo se complementa con otros enfoques como el Promethee y el Geometric Analysis For Interactive (GAIA), los cuales permiten realizar un análisis más visual y dinámico, facilitando la toma de decisiones (Perez-Benitez et al., 2021).

Por otro lado, los modelos gravitacionales están basados en la teoría de la gravitación comercial, que postula que la atracción de un negocio está directamente relacionada con la población en su entorno y la distancia a otras localidades o competidores. Este tipo de modelos ha sido particularmente útil en la planificación de grandes cadenas minoristas y en la distribución de instalaciones en áreas de alta

densidad comercial, ayudando a proveer la demanda de consumidores en función de su proximidad en ciertos puntos (Biscaia & Mota, 2013).

Finalmente, los sistemas de información geográfica (SIG) son herramientas poderosas para mapear y analizar grandes cantidades de datos geoespaciales. Este enfoque ayuda a visualizar patrones de comportamiento, identificar zonas de alta demanda potencial, y planificar rutas de acceso o distribución de manera eficiente (Xmap, 2023).

3.5.6 Modelos de regresión lineal múltiple

Nos permite establecer la relación que se produce entre una variable dependiente (Y), y un conjunto de variables independientes (X_1, X_2, \dots, X_k). El análisis de regresión lineal múltiple, a diferencia del simple, se aproxima más a situaciones de análisis real puesto que los fenómenos, hechos y procesos sociales, por definición, son complejos y, en consecuencia, deben ser explicados en la medida de lo posible por la serie de variables que, directa e indirectamente, participan en su concreción.

Al aplicar el análisis de regresión múltiple lo más frecuente es que tanto la variable dependiente como las independientes sean variables continuas medidas en escala de intervalo o razón. No obstante, caben otras posibilidades: (1) también podremos aplicar este análisis cuando relaciones una variable dependiente continua con un conjunto de variables categóricas. (2) o bien, también aplicaremos el análisis de regresión lineal múltiple en el caso de que relacionemos una variable dependiente nominal con un conjunto de variables continuas.

La anotación matemática del modelo o ecuación de regresión lineal múltiple es la siguiente:

$$Y = a + B1_{x1} + B2_{x2} + \dots + Bn_{xn} + e$$

Donde:

Y : Es la variable a predecir.

a : Es el intercepto (constante).

$B1_{x1}, B2_{x2} \dots Bn_{xn}$: Son parámetros desconocidos a estimar.

e : Es el error que cometemos en la predicción de los parámetros.

Este modelo asume que la relación entre la variable dependiente y las variables independientes es lineal, lo que implica que cualquier cambio en una de las variables independientes tendrá un efecto proporcional y constante.

3.6 Análisis de variables

3.6.1 Indicadores demográficos

Los indicadores demográficos son fundamentales para entender la estructura poblacional de una región y facilitar la toma de decisiones en diversos sectores, como la planificación territorial y el desarrollo de negocios. En este contexto, se presentan datos basados en el censo de población y vivienda 2002 elaboradas por el instituto nacional de estadísticas (INE) los cuales se pronosticaron con la variación porcentual obtenida del censo 2017.. Estos datos incluyen información sobre la población total, desagregada por sexo, zonas (urbana y rural) y grupos etarios.

Estos indicadores permiten evaluar la estructura demográfica y su evolución, proporcionando una base sólida para la planificación de políticas públicas y estrategias empresariales, como la localización de nuevos comercios o franquicias, al identificar segmentos poblacionales específicos que pueden ser atendidos según su distribución geográfica y características sociodemográficas.

3.6.2 R-Cuadrado

El R-cuadrado es una medida estadística utilizada para evaluar la calidad de un modelo de regresión. Indica qué proporción de la variabilidad de la variable dependiente es explicada por las variables independientes del modelo.

El valor del R-cuadrado oscila entre 0 y 1, y se calcula utilizando la relación entre dos componentes de la variabilidad: la variabilidad explicada por el modelo y la variabilidad no explicada correspondientes a los residuos. Cuando el R-cuadrado es cercano a 1, significa que el modelo está explicando una gran proporción de la variabilidad en los datos, lo cual indica un buen ajuste. Por ejemplo, un R-cuadrado de 0,90 implica que el modelo explica el 90% de la variabilidad de la variable dependiente. En cambio, el valor es cercano a 0, significa que el modelo no está explicando casi nada de la variabilidad de los datos.

Se debe tener en cuenta que un R-cuadrado alto, no necesariamente implica que las variables independientes causen los cambios en la variable dependiente. Además, un R-cuadrado alto no siempre significa que el modelo sea bueno, ya que puede haber sobreajuste, especialmente si se incluyen muchas variables irrelevantes.

3.6.3 R-Cuadrado Ajustado

El R-cuadrado ajustado es una medida de la proporción de variabilidad en la variable dependiente que es explicada por las variables independientes del modelo. Este valor ajusta el R-cuadrado tradicional, penalizando el uso de variables no aportan significativamente al modelo. Un R-cuadrado ajustado alto indica que las variables seleccionadas explican bien la variabilidad de la variable dependientes. Mientras que un valor bajo sugiere que el modelo podría no estar capturando adecuadamente la relación. Sin embargo, es importante no solo fijarse en este valor, sino también en la validez de los supuestos del modelo.

3.6.4 Valor P

El valor p asociado a cada variable independiente nos dice si el coeficiente es significativamente diferente de cero. Si el valor p es menor que un umbral comúnmente utilizado de 0,05, podemos rechazar la hipótesis nula de que el coeficiente es igual a cero y concluir que la variable tiene un impacto significativo en

la variable dependiente. Si el valor p es mayor que 0,05, esto indica que la variable no tiene una relación estadísticamente significativa con la variable dependiente en este modelo.

3.6.5 F-Statistics

El estadístico F se utiliza para evaluar la hipótesis nula de que todos los coeficientes de las variables independientes son iguales a cero, es decir, que las variables independientes no tienen efecto sobre la variable dependiente. Si el valor p asociado al estadístico F es bajo, por ejemplo, menor a 0,05, esto significa que al menos una de las variables independientes tiene una relación significativa con la variable dependientes. Si el valor p, es alto, esto podría indicar que el modelo no es significativo en su conjunto y que las variables seleccionadas no aportan mucha información para predecir la variable dependiente.

3.6.6 Prueba Durbin- Watson

El estadístico Durbin-Watson mide la autocorrelación de los residuos del modelo. Este estadístico toma valores entre 0 y 4, donde un valor cercano a 2 indica que no hay autocorrelación, es decir, que los residuos son independientes entre sí. Un valor cercano a 0 equivale a una autocorrelación positiva, mientras que un valor cercano a 4 significa una autocorrelación negativa. En general, se busca que el estadístico Durbin-Watson sea cercano a 2, lo que indica que los residuos no están correlacionados y que el modelo no presenta problemas de autocorrelación.

3.6.7 Coeficientes de las variables independientes

Los coeficientes indican la relación entre cada variable independiente y la variable dependiente. Si un coeficiente es positivo, significa que un aumento en esa variable independiente está asociado con un aumento en la variable dependiente, y viceversa. Además, los errores estándar asociados a los coeficientes proporcionan una medida de la precisión de las estimaciones de los coeficientes. Un coeficiente con un error estándar pequeño representa que el coeficiente está estimado con

mayor precisión, mientras que un error estándar grande indica que la estimación es menor confiable.

3.7 Monetización de datos

La monetización de datos se refiere al proceso de convertir los datos en un activo valioso que genera ingresos o mejoras en la eficiencia para una empresa. En la era digital actual, donde los datos son un recurso crucial, las organizaciones pueden utilizarlos para mejorar la propuesta de valor hacia sus clientes y fortalecer su competitividad (quijada, 2022).

La monetización de datos implica transformar este activo en ingresos o utilidades, ya sea mediante la venta directa de los datos o, de forma más efectiva, mediante su análisis para mejorar productos, procesos y experiencias del cliente (Schmarzo, 2020). Los datos, al ser tratados como un recurso económico que puede producir beneficios, cumplen con la definición de activos propuesta por organizaciones como la fundación de estándares internacionales de reportes financieros (IFRS) (2018) y la organización para la cooperación y el desarrollo económicos (OCDE).

Existen varias estrategias para monetizar los datos. Según el Economic Statistics Centre Of Excellence (ESCOE), las principales formas incluyen:

1. Venta de los datos: Una estrategia en la cual los datos se comercializan directamente a terceros.
2. Desarrollo de nuevos productos: Creación de productos completamente nuevos relacionados con los datos.
3. Mejora de productos existentes: Utilizar los datos para optimizar y mejorar la oferta actual.
4. Mejora de la eficiencia operativa: Uso de los datos para optimizar procesos internos, reducir costos o aumentar ingresos (quijada, 2022).

Schmarzo (2020) sostiene que la mejor forma de monetizar los datos es a través del análisis que permita generar insights valiosos sobre clientes, productos y procesos operativos. Esto puede conducir a la optimización de la eficiencia, la mejora de la experiencia del cliente, la reducción de riesgos y la identificación de nuevas oportunidades de negocio.

El principal beneficio de monetizar los datos es que permite a las organizaciones mejorar su competitividad y eficiencia. Entre los beneficios clave se incluyen:

- Optimización operativa: Al analizar los datos, las empresas pueden identificar áreas de mejora en sus procesos, lo que lleva a una mayor eficiencia y menores costos operativos (Schmarzo, 2020).
- Incremento en ventas y utilidades: A través de la segmentación y el análisis de datos, las empresas pueden ofrecer productos y servicios más alineados con las necesidades del cliente, lo que puede traducirse en mayores ventas.
- Mejora de la experiencia del cliente: Los datos permiten a las empresas comprender mejor a sus clientes, lo que facilita la personalización de sus ofertas y la mejora de la satisfacción del cliente (Liozu & Ulaga, 2018).
- Creación de nuevas oportunidades: La monetización de datos puede ayudar a identificar nuevas oportunidades de negocio, abriendo la puerta a productos o servicios innovadores basados en los insights obtenidos (Schmarzo, 2020).

Por otra parte, Schmarzo (2020) argumenta que los datos son un activo que, a diferencia de otros, no se deprecia con el tiempo ni con el uso, sino que, al contrario, aumenta su valor cuanto más se utilicen, volviéndose más completos y precisos.

4. Propuesta de la metodología de trabajo

La propuesta de la metodología de trabajo está expuesta en fases, las cuales están relacionadas a los objetivos específicos, con el fin de responder al objetivo principal del estudio. Esta metodología de trabajo responde a las bases del marco de trabajo explicativo que contempla este estudio. Se da énfasis en el análisis de regresión múltiple, los estadísticos, pruebas y análisis que se aplicaran para determinar la relación y grado de asociación entre una variable dependiente y sus variables explicativas, así como la estimación de los parámetros de la ecuación.

4.1 Fases de la propuesta metodológica

En la primera fase se realizara la búsqueda de estudios anteriores con el objetivo de conocer y analizar modelos de localización aplicados en investigaciones y el aporte de este en los estudios. Se identificarán los tipos de análisis que se realizaran en cuanto a las variables utilizadas, entre otros aportes esenciales que se obtengan. Con el aporte entregado por las investigaciones, se establecerá un modelo predictivo teórico, permitiendo ser el lineamiento base bajo los fundamentos, aportes y metodologías de trabajo de los estudios relacionados.

En la segunda fase, una vez establecido el modelo predictivo, se diseña una propuesta de trabajo metodológica específica, con las etapas respectivas de ejecución. La etapa número uno, es en la cual se desarrollarla una base de datos con la información recopilada a través de datos demográficos obtenidos del censo y complementando la base de datos con la información recopilada de la empresa, de esta forma los datos serán la entrada para ser aplicada en los métodos de pronósticos. La etapa número dos, se realizaran los análisis descriptivos de las variables obtenidas en la base de datos. esto permitirá identificar las características fundamentales de los datos, hacer una primera exploración para entender el comportamiento, y tomar decisiones sobre qué aspecto pueden ser más relevantes para análisis posteriores. En cuanto a la tercera etapa es clave, el ordenar información en datos de pruebas para ser aplicados y generar modelos. Cabe señalar que los datos considerados como pruebas, son los datos más actuales

obtenidos. La cuarta etapa, se aplicarán el modelo establecidos para realizar los modelos de predicción de localización. En la quinta etapa, se analizarán los resultados obtenidos validando el o los modelos ejecutados.

En la tercera fase, se analizarán y discuten los resultados obtenidos con el modelo aplicado y se realizara una comparación entre los diferentes modelos obtenidos, para finalizar se entregaran conclusiones y recomendaciones de la memoria.

4.2 Diagrama de propuesta metodológica

La elaboración de esta propuesta es realizada en base a una estrategia teórica y práctica, estructurada en fases y que responde principalmente a los objetivos específicos. La ilustración 1 explica las fases descritas en el punto 4.1.

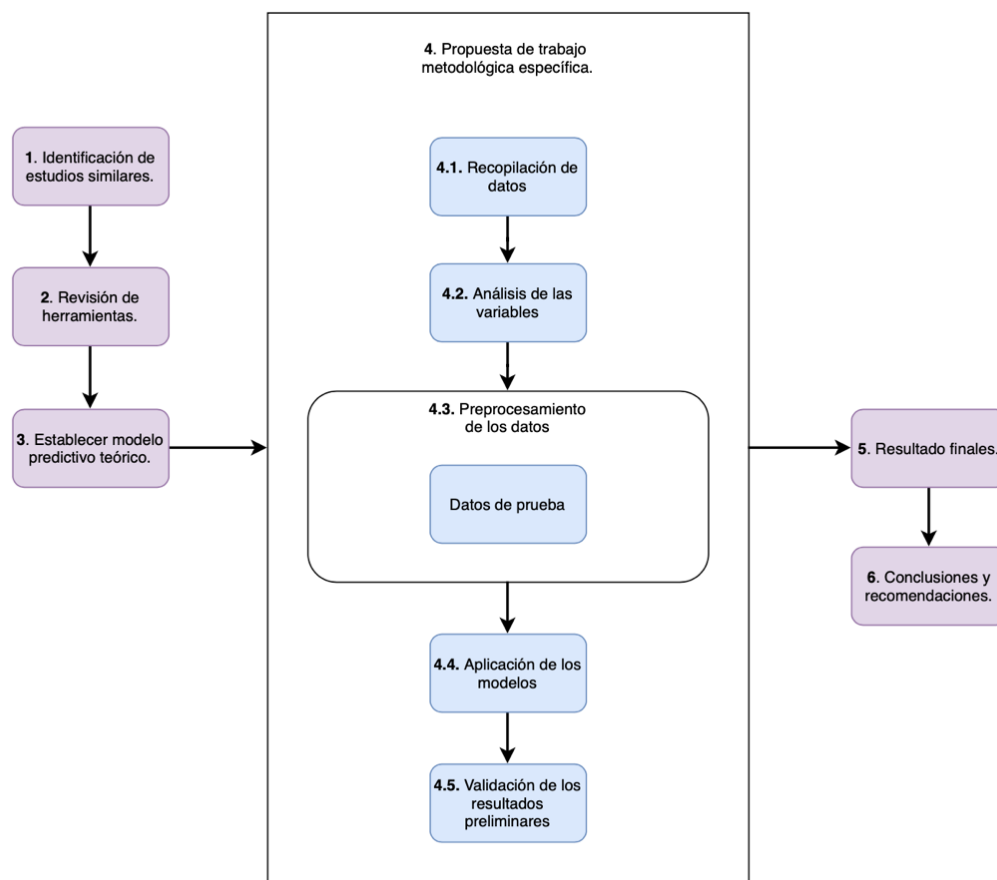


Ilustración 2: Diagrama metodología propuesta.

Fuente: Elaboración propia.

5. Aplicación metodológica

5.1 Recopilación de los datos

Los datos demográficos utilizados en este análisis provienen de diversas fuentes oficiales que permiten obtener información precisa y actualizada sobre las características de la población, en primer lugar, los datos se han obtenido a partir del censo oficial realizado en 2002 en las cuales se aplicaron las variaciones porcentuales obtenidos del censo 2017, un instrumento fundamental para conocer la estructura y distribución de la población en el país. Los resultados específicos para la región de Valparaíso y, en particular, para la comuna de Viña del Mar, están disponibles en la plataforma del Instituto Nacional de Estadísticas (INE), a través del sitio web oficial de los resultados del censo del año 2002 y 2017. Este recurso proporciona acceso a los datos detallados sobre la población, la vivienda y otras aspectos relevantes de la región, permitiendo un análisis exhaustivo de su demografía.

Adicionalmente, para obtener información más específica sobre los datos demográficos de Viña del Mar, se consultó el portal de estadísticas de Municipio de Viña del Mar. En este sitio, se pueden acceder a las “bases de datos, territorial, poblacional y de vivienda”, las cuales se ofrecen información detallada sobre la distribución de la población, características socioeconómica, y datos relacionados con la vivienda en la comuna. Esta fuente permite complementar los datos obtenidos del censo, proporcionando una visión más precisa de las condiciones demográficas y habitacionales de Viña del Mar.

En cuanto a los datos relacionados al negocio fueron obtenidos a partir de las bases de datos de la empresa PedidosYa, las cuales contienen información valiosa sobre

las operaciones y el comportamiento de los clientes. Para acceder y procesar esta información, se utilizó la herramienta BigQuery, una plataforma de análisis de datos en la nube que permite realizar consultas complejas de manera eficiente.

A través del sistema BigQuery, fue posible realizar consultas personalizadas sobre distintos aspectos del negocio, como el comportamiento de los usuarios, las tendencias de consumo y la distribución geográfica de los pedidos, entre otros.

5.2 Tamaño muestra

El tamaño de la muestra utilizado en este modelo, compuesto por 10 observaciones es pequeño y se describe como una muestra de conveniencia. Esto significa que no se seleccionó utilizando un método estadístico riguroso, como el muestro aleatorio o estratificado, sino que fue elegido por razones prácticas, debido a la disponibilidad de datos.

La elección del tamaño de muestra reducido se debe por varias razones. En primer lugar, la disponibilidad de datos especialmente cuando se trabaja con registros de zonas geográficas o base de datos de clientes. Este tipo de muestras pequeñas es común en análisis exploratorios o preliminares, cuyo objetivo principal es identificar tendencias o probar la viabilidad del modelo antes de expandir el análisis.

A pesar de sus ventajas prácticas, el uso de una muestra pequeña tiene varias implicaciones. Una de las principales es la menor representatividad, ya que al no ser aleatoria, la muestra podría no reflejar con precisión las características de la población objetivo, lo que limita la capacidad de generalizar los resultados. Asimismo, el menor poder estadístico asociado a un tamaño reducido afecta la capacidad de detectar relaciones significativas entre las variables, incluso si estas relaciones existen en la población. Por último, una muestra pequeña es más sensible a anomalías, como valores atípicos o errores en los datos, lo que puede sesgar las estimaciones y afectar la fiabilidad del modelo.

Para mitigar los problemas asociados al tamaño reducido de la muestra, existen varias alternativas. Una opción es ampliar la muestra, incluso si sigue siendo de conveniencia, ya que obtener más datos mejorará la estabilidad de los resultados y reducirá el margen de error. Otra estrategia es aplicar técnicas de remuestreo, como el método Bootstrap, que puede proporcionar estimaciones más confiables de la varianza en muestras pequeñas. Finalmente, es fundamental reconocer las limitaciones del modelo, dejando claro en el análisis que las conclusiones deben interpretarse como indicativas y no definitivas, y que es necesario realizar más investigación para validar hallazgos.

Aunque el tamaño de muestra es limitado y fue seleccionado por conveniencia, este modelo puede ofrecer una visión preliminar valiosa del fenómeno estudiado. No obstante, cualquier inferencia más amplia debe considerar esta restricción, y se recomienda complementar el análisis con datos adicionales o técnicas estadísticas que reduzcan los problemas asociados al tamaño reducido de la muestra.

5.3 Análisis de las variables

5.3.1 Descripción de las variables

Las variables que se aplicaran en los modelos de regresión múltiple se definen en la siguiente tabla.

Zona	Sectores de la zona a estudiar. En la investigación zona corresponde a la comuna de Viña del Mar.
Preferencia_1	Preferencia en tipo de comida top 1 en las diferentes zonas.
Preferencia_2	Preferencia en tipo de comida top 2 en las diferentes zonas.

Preferencia_3	Preferencia en tipo de comida top 3 en las diferentes zonas.
Comercio_1	Preferencia en comercios top 1 en las diferentes zonas.
Comercio_2	Preferencia en comercios top 2 en las diferentes zonas.
Comercio_3	Preferencia en comercios top 3 en las diferentes zonas.
N_Persona_Zona	Número de personas en cada zona. (Obtenidos del censo 2002 y proyectados a partir de la variación porcentual obtenida en el censo 2017, la cual tuvo una variación de 16,49 en Viña del Mar.)
Genero_Hombre	Número de personas de genero hombre en cada zona. Obtenidos del censo 2002 y proyectados a partir de la variación porcentual obtenida en el censo 2017, la cual tuvo una variación de 16,49 en Viña del Mar.
Genero_Mujer	Número de personas de genero mujer en cada zona. Obtenidos del censo 2002 y proyectados a partir de la variación porcentual obtenida en el censo 2017, la cual tuvo una variación de 16,49 en Viña del Mar.
Tasa_Crecimiento_Poblacional	Porcentaje proyectado de la población.
Ingreso_Promedio_Hogar	Ingreso promedio hogares en las diferentes zonas.
Tasa_Saturación_Comercio	Porcentaje de nivel de saturación de las diferentes zonas en términos de la cantidad de consumidores o clientes potenciales que están siendo atendidos por una oferta comercial específica.
Cliente_Edad	Promedio edad de clientes en las diferentes zonas.

Tasa_Conversión_Cliente	Porcentaje de interacciones a través de aplicación móvil de delivery que se convierten en ventas.
Gasto_Promedio_Delivery	Promedio de gasto de delivery en las diferentes zonas.
N_Pedidos_Zona	Cantidad de pedidos en las diferentes zonas.
Ticket_Promedio_Pedido_Zona	Gasto promedio por pedidos a través de aplicación móvil en las diferentes zonas.
Satisfacción_Cliente_Zona	NPS promedio en las diferentes zonas con escala de 1 al 10.

Tabla 1: Descripción de variables.

Fuente: Elaboración propia.

5.3.2 Análisis demográfico

El análisis demográfico es un componente esencial para determinar las zonas estratégicas en la comuna de Viña del Mar, Región de Valparaíso, Chile, donde se puede considerar la instalación de nuevos locales comerciales, especialmente de franquicias. Para ellos, se ha utilizado una serie de variables demográficas obtenidas de fuentes oficiales, que incluyen el censo nacional realizado en 2002 y 2017, datos específicos del municipio y del comportamiento comercial registrado a través de plataformas como PedidosYa.

En primer lugar, es importante resaltar la población total de Viña del Mar, que asciende a 302.634 habitantes. Esta cifra proporciona un panorama general sobre el tamaño de la población en la comuna, lo cual es clave para evaluar el potencial de consumo y viabilidad de nuevos negocios. Dentro de la población total, el desglose por género muestra que 143.628 personas son hombres y 159.006 son mujeres. Esta información es crucial para entender las características demográficas

de la población y cómo pueden influir en las decisiones comercial, ya que ciertos productos y servicios pueden estar dirigidos a uno otro género en particular.

El análisis por genero permite, por ejemplo, adaptar la oferta de franquicias a las necesidades de cada grupo. Las zonas con una alta concentración de hombres o mujeres pueden tener una mayor demanda de productos o servicios específicos dirigidos a esos géneros.

Un aspecto clave en este estudio es la tasa de crecimiento poblacional proyectada, En el caso de Viña del Mar, la variación porcentual en cuanto al aumento de la población correspondiente a 16,49%, este valor obtenido de la variación del censo 2002 y 2017 (Ven Ilustración 3). Este dato es fundamental porque permite identificar las zonas con mayor potencial de expansión demográfica. Las áreas con un alto crecimiento poblacional son particularmente atractivas para la instalación de franquicias, ya que pueden experimentar un aumento en la demanda de productos y servicios a medida que la población crece. De este modo, las franquicias que se ubiquen en estas áreas podrían beneficiarse de una clientela en expansión en el futuro cercano.

Además, el ingreso promedio por hogar es otra variable determinante para el análisis, el cual se estima en \$650.047 CLP. Este dato es crucial para evaluar el poder adquisitivo de los habitantes de cada zona, y por lo tanto, para determinar el tipo de productos o servicios que podrían tener mayor demanda.

El perfil de los consumidores, determinado por el promedio de edad en cada zona, también, es crucial. La edad promedio de los clientes corresponde a 32,7, lo cual indica las preferencia de los diferentes grupos etarios, lo que es clave para seleccionar el tipo de franquicias que más pueda interesas en cada área. Por ejemplo, una zona con una población predominantemente joven podría ser un buen lugar para franquicias que ofrezcan productos de servicios rápidos, mientras que

una zona con una mayor proporción de personas mayores podría ser más adecuada para franquicias orientadas a la salud o el bienestar.

Las variables analizadas anteriormente son fundamentales para identificar las zonas más prometedoras para la instalación de franquicias y proporcionaran a los franquicitas las herramientas necesarias para tomar decisiones informadas sobre la expansión de sus negocios.

Territorio	Censo 2002	Censo 2017	Variación (%)
Comuna de Viña del Mar	286.931	334.248	16,49
Región de Valparaíso	1.539.852	1.815.902	17,93
País	15.116.435	17.574.003	16,26

Ilustración 3: Variación Porcentual Población Total censo 2002 y 2017.

Fuente: Censos de Población y Vivienda 2002 y 2017, INE.

En el análisis demográfico se analizaron cada preferencia, estas se obtuvieron por la mayor cantidad de pedidos en cada sector de Viña del Mar. La Preferencia 1, muestra que las hamburguesas son la preferencia con mayor población total (Ver gráfico 1), seguidas por completos y sushi, que cuentan con un tamaño poblacional más reducido pero significativo. En cuanto a la distribución por género (Ver gráfico 2), las mujeres representan una mayor proporción que los hombres en todas las preferencias evaluadas, siendo esta diferencia especialmente notoria en el caso de las hamburguesas. Esto sugiere que las hamburguesas son una opción ampliamente preferida en términos generales.

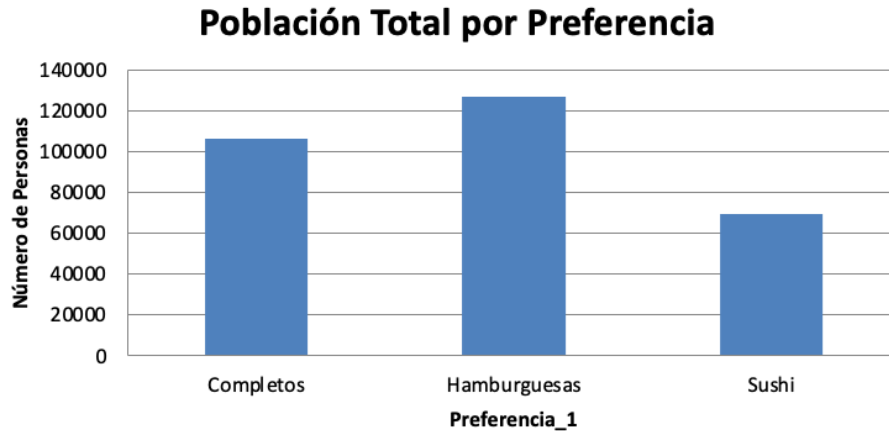


Gráfico 1: Población Total por Preferencia_1.

Fuente: Elaboración Propia

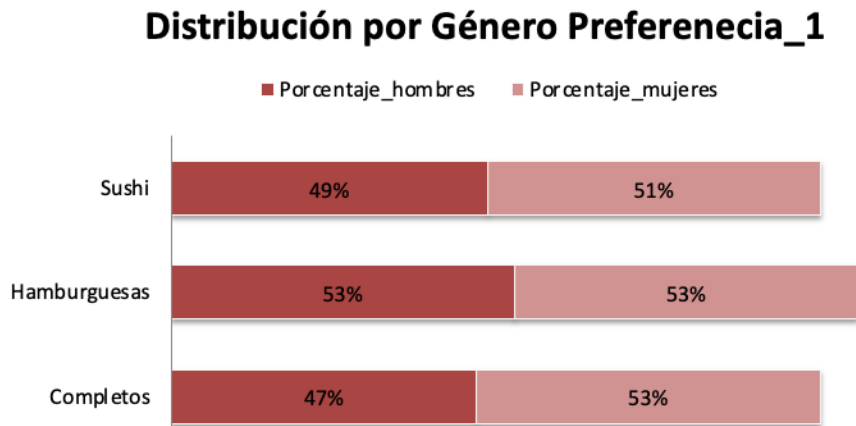


Gráfico 2: Distribución por Género Preferencia_1.

Fuente: Elaboración Propia

Al analizar la Preferencia 2 (Ver gráfico 3), se observa que el sushi destaca como la opción más popular, seguido por los completos, y las pizzas, con proporciones similares de hombres y mujeres en cada categoría (Ver gráfico 4), aunque las mujeres siguen siendo ligeramente predominantes.

Población Total por Preferencia_2

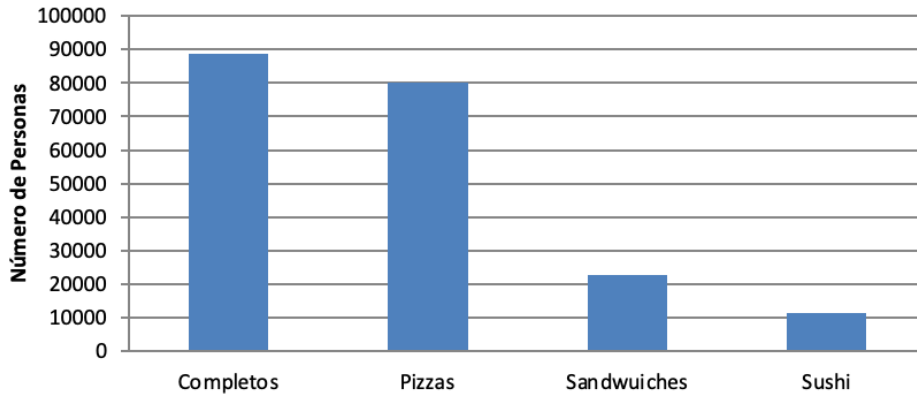


Gráfico 3: Población Total por Preferencia_2.

Fuente: Elaboración Propia.

Distribución de Género por Preferencia_2

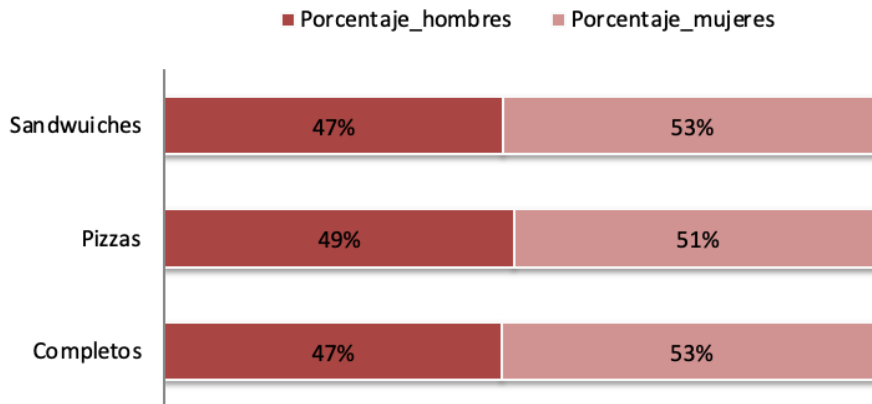


Gráfico 4: Distribución por Género Preferencia_2.

Fuente: Elaboración Propia.

En cuanto a la Preferencia 3 (Ver gráfico 5), los sándwiches lideran en población, seguidos de las hamburguesas y pizzas. Estos resultados, muestran que aunque las hamburguesas y los completos son dominante en términos generales (Ver gráfico 6), las opciones como sushi, sándwiches y pizzas también tienen una aceptación considerable en distintas posiciones de preferencia.

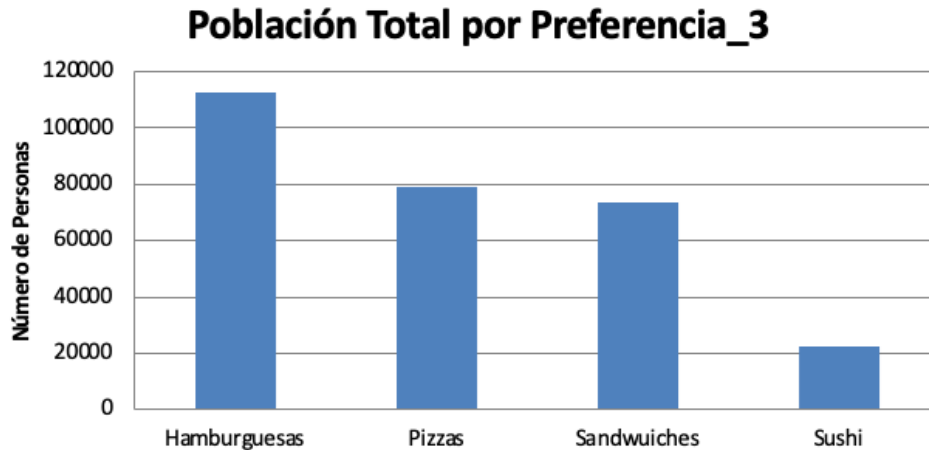


Gráfico 5: Población Total por Preferencia_3.

Fuente: Elaboración Propia.

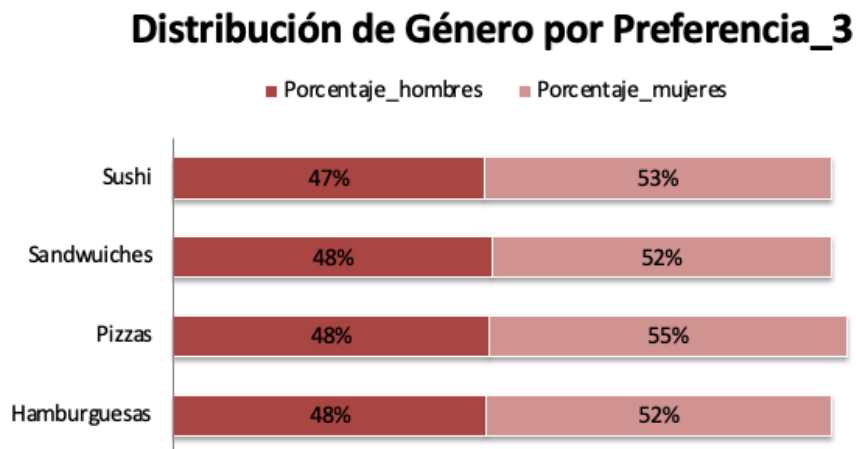


Gráfico 6: Distribución por Género Preferencia_3.

Fuente: Elaboración Propia.

5.3.3 Análisis descriptivo variables

El análisis descriptivo de las variables (ver tabla 2) del conjunto de datos proporciona información clave sobre su comportamiento. Para comenzar, la variable Zona tiene un promedio de 5,5 con una desviación estándar de 3,03, lo que significa que tiene una distribución equilibrada entre los valores mínimo (1) y máximo (10).

Esto indica que las observaciones están distribuidas a lo largo de un rango amplio, con una variabilidad moderada.

En cuanto a las preferencias, Preferencia_1 presenta un promedio de 1,7 con una desviación estándar de 0,82, mientras que Preferencia_2 y Preferencia_3 tienen medias de 3,3, aunque Preferencia_3 muestra una mayor dispersión (desviación estándar de 1,70 frente a 0,95 de Preferencia_2). Esto implica que Preferencia_3 tiene una mayor variabilidad en las respuestas de los individuos.

Las variables relacionadas con comercio también reflejan diferentes patrones. Por ejemplo, Comercio_1 tiene un promedio de 1,6 y una desviación estándar de 0,84, lo cual significa poca variabilidad entre sus observaciones, con valores comprendidos entre 1 y 3. Esto contrasta con otras variables de comercio que podrían presentar mayor diversidad en los datos.

El análisis de las restantes variables revela rangos y distribuciones específicas que ayudan a comprender mejor el comportamiento del fenómeno analizado. Las diferencias en las medias y desviaciones estándar sugieren dónde podrían existir relaciones interesantes o áreas para un análisis más detallado. Este conjunto de estadísticas iniciales proporciona una base sólida para explorar patrones y posibles correlaciones entre las variables del modelo.

variables	Observaciones	Mean	Desviación Estándar	Mínimo	Máximo
Zona	10	5,5	3,02765	1,0	10,0
Preferencia_1	10	1,7	0,82327	1,0	3,0
Preferencia_2	10	3,3	0,94868	2,0	5,0
Preferencia_3	10	3,3	1,70293	1,0	5,0
Comercio_1	10	1,6	0,84327	1,0	3,0
Comercio_2	10	2,7	0,82327	1,0	4,0
Comercio_3	10	2,0	0,94280	1,0	4,0

N_Persona_Zona	10	30976,64	9981,17	15705,24	48164,36
Genero_Hombre	10	14576,56	4614,64	7557,4	22016,8
Genero_Mujer	10	16679,64	5401,40	8147,84	26147,56
Tasa_Crecimiento_Poblacional	10	1,51	0,0	1,51	1,51
Ingreso_Promedio_Hogar	10	650047,0	0,0	650047,0	650047,0
Tasa_Saturación_Comercio	10	42,868	13,7570	26,98	68,95
Cliente_Edad	10	32,7	1,82878	30,0	36,0
Tasa_Conversión_Cliente	10	52,06700	13,0846	32,56	72,97
Gasto_Promedio_Delivery	10	14668,4	6257,31	8177,0	28456,0
N_Pedidos_Zona	10	152285,0	73180,2	88856,0	295184,0
Ticket_Promedio_Pedido_Zona	10	31367,2	10766,6	18432,0	51476,0
Satisfacción_Cliente_Zona	10	8,3941	1,04966	6,62	10,0

Tabla 2: Cálculos análisis descriptivo por variables.
Fuente: Elaboración propia.

5.4 Aplicación de Modelo regresión lineal múltiple

Los modelos propuestos a continuación se componen de las siguientes variables (ver tabla 3):

variables	Modelo 1	Modelo 2	Modelo 3	Modelo 4
Zona	x	x	x	x
Preferencia_1	x	x	x	x
Preferencia_2			x	x
Preferencia_3				
Comercio_1	x		x	x
Comercio_2				
Comercio_3				
N_Persona_Zona		x		
Genero_Hombre				

Genero_Mujer				
Tasa_Crecimiento_Poblacional				
Ingreso_Promedio_Hogar		x		
Tasa_Saturación_Comercio				
Cliente_Edad				x
Tasa_Conversión_Cliente	x		x	x
Gasto_Promedio_Delivery		x	x	x
N_Pedidos_Zona		x	x	x
Ticket_Promedio_Pedido_Zona			x	
Satisfacción_Cliente_Zona		x	x	x

Tabla 3: Variables que componen los modelos.
Fuente: Elaboración propia.

5.4.1 Modelo 1

El Modelo 1 (ver ilustración 4) tiene como variable dependiente la Zona y representa un R-cuadrado de 0,921, lo que indica que este modelo explica el 92,1% de la variación en la variable dependiente. Esto representa un ajuste significativamente mejor en comparación con los modelos anteriores. Además, el R-cuadrado ajustado es de 0,881, un valor muy alto que refleja que las variables independientes seleccionadas tienen una fuerte relación con la variable dependiente. El estadístico F del modelo es de 23,30, con un valor p asociado de 0,00106, lo que confirma que el modelo es globalmente significativo al nivel de confianza del 95%. Sin embargo, el tamaño de la muestra es reducido, con solo 10 observaciones, lo que podría afectar la robustez de las estimaciones. Por otro lado, el estadístico Durbin-Watson es de 2,153, cercano a 2, lo que significa que no hay evidencia de autocorrelación en los residuos.

En cuanto a los coeficientes, el intercepto tiene un valor de 3,6580, que representa el valor esperado de Zona cuando todas las variables independientes son cero.

Aunque su valor p de 0,054 está cerca del umbral de significancia, no alcanzara a ser significativo al nivel del 95%. La variable Preferencia_1 tiene un coeficiente de 0,7490, lo que indica que por cada unidad adicional en esta variable, zona incrementara en promedio en 0,7490. Sin embargo, su valor p de 0,212 sugiere que este efecto no es estadísticamente significativo. Por otro lado, Comercio_1 presenta un coeficiente de 3,3550, lo que representa un impacto positivo fuerte y estadísticamente significativo ($p = 0,001$), destacándose como una de la variables más influyentes en la predicción de Zona. Finalmente, Tasa_Conversión_Cliente tiene un coeficiente de -0,0922, lo que refleja una relación negativa significativa ($p = 0,023$), indicando que un aumento en esta variable está asociado con una disminución en Zona.

Las métricas adicionales revelan errores estándar relativamente pequeños, lo que indica precisión en las estimaciones de los coeficientes. El número de condición del modelo es 251, un valor aceptable que no sigue problemas graves de multicolinealidad. Sin embargo, el test de Omnibus muestra un valor de p de 0,004, lo que significa que los residuos podrían no seguir una distribución normal, aspecto que podría afectar la inferencia estadística.

Dep. Variable:	Zona	R-squared:	0.921			
Model:	OLS	Adj. R-squared:	0.881			
Method:	Least Squares	F-statistic:	23.20			
Date:	Tue, 03 Dec 2024	Prob (F-statistic):	0.00106			
Time:	00:15:21	Log-Likelihood:	-12.072			
No. Observations:	10	AIC:	32.14			
Df Residuals:	6	BIC:	33.35			
Df Model:	3					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	3.6580	1.527	2.396	0.054	-0.077	7.393
Preferencia_1	0.7490	0.537	1.396	0.212	-0.564	2.062
Comercio_1	3.3550	0.582	5.767	0.001	1.931	4.779
Tasa_conversion_cliente	-0.0922	0.030	-3.032	0.023	-0.167	-0.018
Omnibus:	11.072	Durbin-Watson:	2.153			
Prob(Omnibus):	0.004	Jarque-Bera (JB):	5.015			
Skew:	1.553	Prob(JB):	0.0815			
Kurtosis:	4.544	Cond. No.	251.			

Ilustración 4: Resultados Regresión Lineal Múltiple Modelo 1.

Fuente: Elaboración Propia.

5.4.2 Modelo 2

El análisis de modelo 2 (ver Ilustración 5), que tiene como objetivo identificar los factores más relevante para determinar la zona óptima para ubicar un comercio en específico franquicias, se implementa una regresión múltiple con la variable dependiente Zona y variables independientes en las cuales se incluyen Preferencia_1, Ingreso_promedio_hogar, N_Persona_Zona, N_pedidos_zona, Gasto_Promedio_Delivery y Satisfaccion_cliente_zona.

Las métricas globales del modelo, el R-cuadrado es 0,883, lo que indica que el modelo explica aproximadamente el 88,3% de la variabilidad de la variable dependiente Zona, siendo un buen ajuste. Respecto al R-cuadrado ajustado, que toma en cuenta el número de variables en el modelo, es 0,736, lo que sigue demostrando un ajuste razonablemente bueno. El valor estadístico F es de 6.024, con un valor p de 0,0532. Lo que significa que, en conjunto, las variables independientes tienen un efecto significativo sobre la variable dependiente. En cuanto al valor de Durbin-Watson es 1,178, lo que representa que podría haber cierta autocorrelación positiva en los residuos, ya que este valor es menor a 2.

Los coeficientes, que representan el impacto sobre la variable dependiente, se observa que Preferencia_1 tiene un coeficiente positivo de 4.1724. Esto significa que un aumento en la preferencia de los consumidores hacia un determinado tipo de comida tiene una relación positiva con la zona, es decir, áreas con mayor preferencia hacia ciertos tipos de comidas tienden a estar asociadas con una zona óptima para la instalación de una franquicia. El coeficiente de Gasto_Promedio_Delivery es de 0,0005, lo que se traduce en una relación positiva con la variable dependiente. Un aumento es el gasto promedio de los consumidores en delivery tiene un efecto positivo sobre la Zona optima.

En cuanto a las variables Ingreso_promedio_hogar, N_Persona_Zona, N_pedidos_zona y Satisfaccion_cliente_zona, no son estadísticamente significativas en el modelo, lo que implica que en Zona no tiene un impacto relevante en la determinación de la Zona optima.

Dep. Variable:	Zona	R-squared:	0.883			
Model:	OLS	Adj. R-squared:	0.736			
Method:	Least Squares	F-statistic:	6.024			
Date:	Tue, 03 Dec 2024	Prob (F-statistic):	0.0532			
Time:	12:53:05	Log-Likelihood:	-14.023			
No. Observations:	10	AIC:	40.05			
Df Residuals:	4	BIC:	41.86			
Df Model:	5					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Preferencia_1	4.1724	1.785	2.337	0.080	-0.784	9.128
Ingreso_promedio_hogar	1.433e-05	1.2e-05	1.197	0.297	-1.89e-05	4.76e-05
N_Persona_Zona	-0.0001	0.000	-1.026	0.363	-0.000	0.000
N_pedidos_zona	-8.698e-06	1.19e-05	-0.730	0.506	-4.18e-05	2.44e-05
Gasto_promedio_Delivery	0.0005	0.000	3.554	0.024	0.000	0.001
Satisfaccion_cliente_zona	-1.5465	1.140	-1.357	0.246	-4.711	1.618
Omnibus:	0.074	Durbin-Watson:	1.178			
Prob(Omnibus):	0.964	Jarque-Bera (JB):	0.292			
Skew:	0.097	Prob(JB):	0.864			
Kurtosis:	2.186	Cond. No.	2.80e+06			

Ilustración 5: Resultados Regresión Lineal Múltiple Modelo 2.

Fuente: Elaboración Propia.

5.4.4 Modelo 3

El modelo 3 (ver Ilustración 6) fue diseñado para incluir una selección amplia de variables que podrían influir en la determinación de la variable dependiente zona. Este modelo se caracteriza por un ajuste perfecto, con un R-cuadrado de 1,000 y un R-cuadrado ajustado de 0,999, lo que indica que las variables seleccionadas explican la totalidad de la variación observada en Zona.

El estadístico F del modelo tiene un valor muy alto, 1518,07, con un valor p de 0,0198, lo que confirma que el modelo es globalmente significativo. Esto sugiere que las variables incluidas en el modelo tienen un impacto conjunto estadísticamente significativo sobre la variable dependiente.

En cuanto a los coeficientes de las variables, algunas de ellas destacan por su significancia estadística. Preferencia_1 tiene un coeficiente positivo de 3,8684, lo que indica que un aumento en esta variable se asocia con un incremento significativo en la predicción de Zona. El valor p asociado (0,036) refuerza su importancia estadística. Gasto_Promedio_Delivery, con un coeficiente positivo de 0,0004 ($p = 0,029$), también demuestra ser un factor relevante, sugiriendo que a mayores gastos promedio están vinculados a mejores resultados en Zona.

Por otro lado, variables como Tasa_Conversión_Cliente y Satisfacción_Cliente_Zona tienen coeficientes negativos significativos, -0,1334 ($p = 0,025$) y -1,6100 ($p = 0,045$) respectivamente, lo que implica que aumentos en estas variables podrían estar asociados con un menor desempeño en Zona.

Sin embargo, no todas las variables del modelo son estadísticamente significativas. Comercio_1 y Preferencia_2 no alcanzan el nivel de significancia requerido, lo que significa que su impacto en la predicción de Zona es limitado o no significativo en este conjunto de datos.

Un desafío notable en este modelo es la presencia de multicolinealidad, como lo indica el número de condición elevado ($6,16e+06$). Este problema podría afectar la estabilidad de las estimaciones y la interpretación precisa de los coeficientes. Además, el estadístico Durbin-Watson tiene un valor de 1,742, lo que indica una ligera autocorrelación positiva en los residuos, aunque no lo suficiente como para invalidar el modelo.

Dep. Variable:		Zona	R-squared:	1.000		
Model:		OLS	Adj. R-squared:	0.999		
Method:		Least Squares	F-statistic:	1518.		
Date:		Tue, 03 Dec 2024	Prob (F-statistic):	0.0198		
Time:		00:15:21	Log-Likelihood:	22.283		
No. Observations:		10	AIC:	-26.57		
Df Residuals:		1	BIC:	-23.84		
Df Model:		8				
Covariance Type:		nonrobust				
	coef	std err	t	P> t	[0.025	0.975]
const	14.4715	0.908	15.936	0.040	2.933	26.010
Comercio_1	0.7645	0.133	5.767	0.109	-0.920	2.449
Preferencia_1	3.8684	0.219	17.695	0.036	1.091	6.646
Preferencia_2	-0.2103	0.089	-2.352	0.256	-1.346	0.926
Gasto_promedio_Delivery	0.0004	1.65e-05	21.726	0.029	0.000	0.001
Tasa_conversion_cliente	-0.1337	0.005	-25.866	0.025	-0.199	-0.068
N_pedidos_zona	2.036e-06	1.89e-06	1.074	0.477	-2.2e-05	2.61e-05
Ticket_promedio_pedido_zona	-3.742e-05	1.09e-05	-3.441	0.180	-0.000	0.000
Satisfaccion_cliente_zona	-1.6100	0.115	-14.015	0.045	-3.070	-0.150
Omnibus:	2.396	Durbin-Watson:	1.742			
Prob(Omnibus):	0.302	Jarque-Bera (JB):	0.948			
Skew:	-0.234	Prob(JB):	0.622			
Kurtosis:	1.566	Cond. No.	6.16e+06			

Ilustración 6: Resultados Regresión Lineal Múltiple Modelo 3.

Fuente: Elaboración Propia.

5.4.5 Modelo 4

El modelo 4 (Ver ilustración 5) ofrece un ajuste notablemente preciso para predecir la variable dependiente zona en función de las variables seleccionadas. Con un R-cuadrado de 1,000, el modelo explica el 100% de la variación en la variable dependiente, lo que indica un ajuste perfecto a los datos. Incluso al considerar el número de variables independientes incluidas, el R-cuadrado ajustado sigue siendo excepcionalmente alto, con un valor de 0,999. La F-Statistics es significativa, con un valor de 1861,13 y un valor p de 0,0179, lo que respalda que el modelo es globalmente significativo y que las variables independientes contribuyen de manera significativa a explicar la variación en zona.

Entre las variables incluidas, se destacan algunas como significativas. Preferencia_1 tiene un impacto positivo sustancial en la predicción de Zona, con un

coeficiente de 3,3945, mientras que Gasto_Promedio_Delivery muestra una relación positiva moderada, con un coeficiente de 0,0004, ambos con significancia estadística. Por otro lado, Satisfacción_Cliente_Zona y Tasa_Conversión_Cliente son estadísticamente significativas, pero con efectos negativos sobre zona, lo que indica que un aumento en estas variables está asociado con una disminución en la variable dependiente. Otras variables, como Comercio_1, Preferencia_2, Cliente_Edad y N_Pedidos_Zona, no alcanzaron niveles de significancia estadística, aunque su inclusión mejora el ajuste global del modelo.

Un aspecto importante a considerar es la presencia de multicolinealidad, como lo sugiere el número de condición del modelo, que alcanza un valor de 6,64e+06. esto indica fuertes correlaciones entre algunas variables independientes, lo que podría estar afectando la precisión de las estimaciones individuales de los coeficientes. Además, el estadístico Durbin-Watson, con un valor de 1,713, indica una leve autocorrelación positiva en los residuos, aunque no lo suficiente como para ser un problema crítico.

Dep. Variable:	Zona	R-squared:	1.000			
Model:	OLS	Adj. R-squared:	0.999			
Method:	Least Squares	F-statistic:	1861.			
Date:	Tue, 03 Dec 2024	Prob (F-statistic):	0.0179			
Time:	00:15:21	Log-Likelihood:	23.302			
No. Observations:	10	AIC:	-28.60			
Df Residuals:	1	BIC:	-25.88			
Df Model:	8					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	8.8475	0.927	9.540	0.066	-2.937	20.632
Comercio_1	0.9817	0.103	9.500	0.067	-0.331	2.295
Preferencia_1	3.3945	0.145	23.440	0.027	1.554	5.235
Preferencia_2	0.1057	0.066	1.609	0.354	-0.729	0.940
Gasto_promedio_Delivery	0.0004	1.41e-05	27.020	0.024	0.000	0.001
Cliente_edad	0.0774	0.020	3.839	0.162	-0.179	0.334
Tasa_conversion_cliente	-0.1122	0.006	-20.040	0.032	-0.183	-0.041
N_pedidos_zona	-4.662e-06	6.26e-07	-7.448	0.085	-1.26e-05	3.29e-06
Satisfaccion_cliente_zona	-1.4997	0.095	-15.845	0.040	-2.702	-0.297
Omnibus:	1.599	Durbin-Watson:	1.713			
Prob(Omnibus):	0.450	Jarque-Bera (JB):	0.922			
Skew:	0.406	Prob(JB):	0.631			
Kurtosis:	1.753	Cond. No.	6.64e+06			

Ilustración 7: Resultados Regresión Lineal Múltiple Modelo 4.

Fuente: Elaboración Propia.

6. Resultados

6.1 Comparación Modelos

La comparación de los modelos 1, 2, 3 y 4 muestra diferencias importantes en términos de ajuste, significancia global, multicolinealidad y la capacidad de interpretar los coeficientes. Cada modelo presenta características particulares que lo hacen más o menos adecuado para determinar la zona óptima para ubicar un comercio, en este caso una franquicia.

El modelo 1 es el mejor entre los modelos comparados, destacando por su eficiencia y simplicidad. Aunque tiene una muestra pequeña, su R-cuadrado ajustado de 0,881, la significancia de las variables clave como Comercio_1 y Tasa_Conversión_Cliente, y la ausencia de problemas graves de autocorrelación lo hacen más confiable y útil para predicción de la Zona. El Modelo 3 y el Modelo 4 tienen un ajuste perfecto, pero la multicolinealidad y la dificultad de interpretar los coeficientes reducen su aplicabilidad práctica. El modelo 2, aunque también tiene un ajuste razonable, tiene un R-cuadrado ajustado más bajo y problemas de significancia en muchas de sus variables, lo que limita su efectividad.

6.2 Aplicación de datos en Modelo 1

Debido a que el Modelo 1 es el mejor candidato de modelo para trabajar, se ejecutarán los datos en el modelo para determinar en qué sector de la comuna de Viña del Mar es óptima para establecer un local comercial correspondiente a una franquicia. La ecuación del modelo 1 corresponde a:

$$Zona = B0 + B1 * Preferencia_1 + B2 * N_{PersonaZona} + B3 * Tasa_{convercion_{cliente}}$$

Aplicación con los coeficientes obtenidos:

$$Zona = 3,6580 + 0,7490 * Preferencia_1 + 3,3550 * N_{PersonaZona} + (-0,0922) * Tasa_{convercion_{cliente}}$$

En este tipo de modelo, la Zona óptima es aquel que tiene el valor más alto de Zona, ya que representa una ubicación con las mejores características según las variables del modelo.

En este caso el valor más alto de Zona es 10,659. Por lo tanto, basándonos en el modelo y los valores obtenidos, la ubicación óptima para el comercio corresponde al sector de Recreo, con una preferencia por el tipo de comida Completos.

La segunda ubicación estratégica identificada es Viña del Mar Oriente con un valor de 8,478, donde la preferencia predominante es el tipo de comida Hamburguesas. La tercera Zona es el Plan de Viña del Mar con una puntuación 7,462 también destacando la preferencia por Hamburguesas.

En cuarto lugar, se encuentra el sector Reñaca, con una puntuación de 7,175, donde se recomienda la instalación de un local de Sushi. A continuación, se encuentra Miraflores, con una preferencia en Hamburguesas y una puntuación de 5,504, seguida por la Santa Inés con 4,538 puntos y una preferencia por Completos.

Otras zonas destacadas son Gómez Carreño, con una puntuación de 3,915 y una preferencia por Sushi, y Reñaca Alto donde también predomina la preferencia por Sushi, con una puntuación de 2,206. En Achupallas donde la preferencia es el tipo de comida Hamburguesas con una puntuación de 3,241. Finalmente, se encuentra la zona de Nueva Aurora con puntuación 1,729 y una preferencia por Sándwiches.

7. Conclusiones y recomendaciones

El análisis realizado a lo largo de este estudio se centra en la aplicación de modelos predictivos para la optimización de la zona óptima para ubicar un comercio. En el cual es crucial evaluar los modelos propuestos con una base de datos más extensas, en las cuales se consideren datos demográficos actuales.

Como se ha mencionado el tamaño de muestra es limitado y fue seleccionado por conveniencia, este modelo seleccionado puede ofrecer una visión preliminar valiosa del fenómeno estudiado. No obstante, cualquier inferencia más amplia debe considerar esta restricción, y se recomienda complementar el análisis con datos adicionales o técnicas estadísticas que reduzcan los problemas asociados al tamaño reducido de la muestra.

A través de los diferentes modelos generados, principalmente el Modelo 1, se ha logrado tener cómo las variables Preferencia_1, Comercio_1 y Tasa_Conversión_Cliente, impactan la ubicación óptima de los comercios en diversas zonas, lo que, a su vez, facilita la toma de decisiones más informadas sobre la ubicación de nuevas franquicias o comercios.

Se concluye que el uso de regresión múltiple para la predicción de Zona ha permitido obtener un R-cuadrado de 0,921 en el Modelo 1, lo que significa que el modelo explica un alto porcentaje de la variabilidad de la variable dependiente, siendo este uno de los mejores ajustes entre los modelos considerados. Esto refuerza la importancia de usar un conjunto robusto de las variables como Comercio_1 y Tasa_Conversión_Cliente, que fueron significativas en la predicción de la zona óptima, a pesar de la limitación de tamaño de la muestra (solo 10 observaciones).

La analítica predictiva desarrollada en este estudio ha demostrado ser una herramienta poderosa para identificar las zonas más óptimas para la ubicación de

comercios en especial las franquicias, basándose en un análisis riguroso de las variables que afectan directamente la dinámica de consumo en distintas ubicaciones. Este tipo de análisis avanzado de datos, aunque basado en un modelo relativamente sencillo, tiene la capacidad de mejorar la toma de decisiones estratégicas para la expansión de los negocios.

El término de recomendaciones, es crucial que la empresa fomente una cultura organizacional basada en datos, donde aproveche el creciente volumen de datos generados constantemente. Sin embargo, muchas veces esta información no se considera de manera estructurada y no se aprovecha para su análisis. Integrar analítica avanzada permitirá a la empresa explorar nuevas estrategias de negocio, basadas en datos concretos, lo cual la posicionara con una ventaja competitiva importante en el mercado. Para esto, la empresa debe garantizar la correcta recolección, transformación y análisis de los datos, además de mantener un repositorio robusto que permita generar modelos predictivos cada vez más precisos.

Es clave desarrollar y validar continuamente modelos predictivos que identifiquen patrones en el desempeño de ubicaciones, perfiles de clientes y tendencias del mercado, integrando herramientas para identificar áreas de alto potencial.

Esto es el inicio para el análisis basado en datos, potenciando este tipo de decisiones.

ANEXOS

Cálculos Adicionales: Modelo 1

```
Métricas adicionales para el Modelo 1:  
Adjusted R-squared: 0.8809472934295814  
F-statistic: 23.198923119194525  
Prob(F-statistic): 0.0010605985249874889  
Standard Errors: const 1.526585  
Preferencia_1 0.536674  
Comercio_1 0.581800  
Tasa_conversion_cliente 0.030400  
dtype: float64  
Durbin-Watson: 2.153492714878675
```

Cálculos Adicionales: Modelo 2

```
Métricas adicionales para el Modelo 1:  
Adjusted R-squared: 0.736214910931408  
F-statistic: 6.023736725816013  
Prob(F-statistic): 0.05324070286216532  
Standard Errors: Preferencia_1 1.785001  
Ingreso_promedio_hogar 0.000012  
N_Persona_Zona 0.000128  
N_pedidos_zona 0.000012  
Gasto_promedio_Delivery 0.000130  
Satisfaccion_cliente_zona 1.139779  
dtype: float64  
Durbin-Watson: 1.1777332211101001
```

Cálculos Adicionales: Modelo 3

```
Métricas adicionales para el Modelo :  
Adjusted R-squared: 0.9992589869234415  
F-statistic: 1518.0668316815425  
Prob(F-statistic): 0.019847416601275036  
Standard Errors: const 0.908079  
Comercio_1 0.132555  
Preferencia_1 0.218614  
Preferencia_2 0.089388  
Gasto_promedio_Delivery 0.000017  
Tasa_conversion_cliente 0.005169  
N_pedidos_zona 0.000002  
Ticket_promedio_pedido_zona 0.000011  
Satisfaccion_cliente_zona 0.114874  
dtype: float64  
Durbin-Watson: 1.7415440449057533
```

Cálculos Adicionales: Modelo 4

```
Métricas adicionales para el Modelo :
Adjusted R-squared: 0.9993955686980727
F-statistic: 1861.1287047860103
Prob(F-statistic): 0.01792549321466967
Standard Errors: const          9.274453e-01
Comercio_1                      1.033365e-01
Preferencia_1                   1.448190e-01
Preferencia_2                   6.567976e-02
Gasto_promedio_Delivery        1.405261e-05
Cliente_edad                   2.016985e-02
Tasa_conversion_cliente        5.598231e-03
N_pedidos_zona                 6.259168e-07
Satisfaccion_cliente_zona     9.464470e-02
dtype: float64
Durbin-Watson: 1.712771162100436
```

Código python

```
import pandas as pd
import statsmodels.api as sm
from sklearn.preprocessing import LabelEncoder
from sklearn.decomposition import PCA

# Cargar los datos
df = pd.read_excel('Datos.xlsx')

y = df['Zona']

# Modelo 1: Variables originales
X1 = df[['Preferencia_1', 'Comercio_1', 'Tasa_conversion_cliente']]

# Asegurar que todas las variables numéricas sean de tipo adecuado
X1 = X1.apply(pd.to_numeric, errors='coerce')

# Añadir una constante a las variables independientes
X1 = sm.add_constant(X1)

# Crear el modelo de regresión para el Modelo 1
model1 = sm.OLS(y, X1).fit()

# Mostrar el resumen del Modelo 1
print("Modelo 1 - Resumen")
print(model1.summary())

# Obtener métricas adicionales del Modelo 1
adjusted_r_squared_1 = model1.rsquared_adj
f_statistic_1 = model1.fvalue
f_p_value_1 = model1.f_pvalue
standard_errors_1 = model1.bse
```

```

durbin_watson_1 = sm.stats.stattools.durbin_watson(model1.resid)

# Mostrar métricas adicionales del Modelo 1
print("\nMétricas adicionales para el Modelo 1:")
print(f"Adjusted R-squared: {adjusted_r_squared_1}")
print(f"F-statistic: {f_statistic_1}")
print(f"Prob(F-statistic): {f_p_value_1}")
print("Standard Errors:", standard_errors_1)
print(f"Durbin-Watson: {durbin_watson_1}")

# Modelo 2: Variables originales
X2 = df[['Preferencia_1', 'Ingreso_promedio_hogar', 'N_Persona_Zona',
        'N_pedidos_zona', 'Gasto_promedio_Delivery',
        'Satisfaccion_cliente_zona'
        ]]

# Asegurar que todas las variables numéricas sean de tipo adecuado
X2 = X2.apply(pd.to_numeric, errors='coerce')

# Añadir una constante a las variables independientes
X2 = sm.add_constant(X2)

# Crear el modelo de regresión para el Modelo 2
model2 = sm.OLS(y, X2).fit()

# Mostrar el resumen del Modelo 2
print("Modelo 2 - Resumen")
print(model2.summary())

# Obtener métricas adicionales del Modelo 2
adjusted_r_squared_2 = model2.rsquared_adj
f_statistic_2 = model2.fvalue
f_p_value_2 = model2.f_pvalue
standard_errors_2 = model2.bse
durbin_watson_2 = sm.stats.stattools.durbin_watson(model2.resid)

# Mostrar métricas adicionales del Modelo 2
print("\nMétricas adicionales para el Modelo 2:")
print(f"Adjusted R-squared: {adjusted_r_squared_2}")
print(f"F-statistic: {f_statistic_2}")
print(f"Prob(F-statistic): {f_p_value_2}")
print("Standard Errors:", standard_errors_2)
print(f"Durbin-Watson: {durbin_watson_2}")

# Modelo 3: Variables originales
X3 = df[['Comercio_1', 'Preferencia_1', 'Preferencia_2', 'Gasto_promedio_Delivery',
        'Tasa_conversion_cliente', 'N_pedidos_zona',
        'Ticket_promedio_pedido_zona', 'Satisfaccion_cliente_zona'
        ]]

```

```

]]

# Asegurar que todas las variables numéricas sean de tipo adecuado
X3 = X3.apply(pd.to_numeric, errors='coerce')

# Añadir una constante a las variables independientes
X3 = sm.add_constant(X3)

# Crear el modelo de regresión para el Modelo 3
Model3 = sm.OLS(y, X3).fit()

# Mostrar el resumen del Modelo 3
print("Modelo 3- Resumen")
print(model3.summary())

# Obtener métricas adicionales del Modelo 3
adjusted_r_squared_3 = model3.rsquared_adj
f_statistic_3 = model3.fvalue
f_p_value_3 = model3.f_pvalue
standard_errors_3 = model3.bse
durbin_watson_3 = sm.stats.stattools.durbin_watson(model3.resid)

# Mostrar métricas adicionales del Modelo 3
print("\nMétricas adicionales para el Modelo 3:")
print(f"Adjusted R-squared: {adjusted_r_squared_3}")
print(f"F-statistic: {f_statistic_3}")
print(f"Prob(F-statistic): {f_p_value_3}")
print("Standard Errors:", standard_errors_3)
print(f"Durbin-Watson: {durbin_watson_3}")

# Modelo 4: Variables originales
X4 = df[['Comercio_1', 'Preferencia_1', 'Preferencia_2', 'Gasto_promedio_Delivery',
        'Cliente_edad', 'Tasa_conversion_cliente', 'N_pedidos_zona',
        'Satisfaccion_cliente_zona']]

]]

# Asegurar que todas las variables numéricas sean de tipo adecuado
X4 = X4.apply(pd.to_numeric, errors='coerce')

# Añadir una constante a las variables independientes
X4 = sm.add_constant(X4)

# Crear el modelo de regresión para el Modelo 4
Model4 = sm.OLS(y, X4).fit()

# Mostrar el resumen del Modelo 4

```

```

print("Modelo 4- Resumen")
print(model4.summary())

# Obtener métricas adicionales del Modelo 4
adjusted_r_squared_4 = model4.rsquared_adj
f_statistic_4 = model4.fvalue
f_p_value_4= model4.f_pvalue
standard_errors_4 = model4.bse
durbin_watson_4 = sm.stats.stattools.durbin_watson(model4.resid)

# Mostrar métricas adicionales del Modelo 4
print("\nMétricas adicionales para el Modelo 4:")
print(f"Adjusted R-squared: {adjusted_r_squared_4}")
print(f"F-statistic: {f_statistic_4}")
print(f"Prob(F-statistic): {f_p_value_4}")
print("Standard Errors:", standard_errors_4)
print(f"Durbin-Watson: {durbin_watson_4}")

```

BIBLIOGRAFÍA

- Cruz soto, n. v. (2016). *Análisis descriptivo de las franquicias en Chile* (informe de práctica profesional). Universidad de Valparaíso, facultad de ciencias económicas y administrativas.
- Cámara de comercio de Santiago (ccs). (2023). *Estudio revela que las franquicias aumentaron un 7% en Chile durante la pandemia*.
- Paideia. (2023). *10 franquicias en Chile rentables este 2023*. Revista Paideia.
- Latam networks. (2024). *franquicias en Chile para invertir en 2024*.
- 100franquicias. (2023). *2023: Año de fortalecimiento de la franquicia en Chile*.
- Rodríguez J., María J, Mora C. & Rafaiel (2001). *Análisis de regresión múltiple*.
- xmap. (2023). *business location analysis: the key to strategic decision making*.
- perez-benitez, v., gemar, g., & hernández, m. (2021). *multi-criteria analysis for business location decisions*. mathematics, 9(20), 2615.
- biscaia, r., & mota, i. (2013). *models of spatial competition: a critical review*. papers in regional science.
- montejaño escamilla, j. a., & cruz bello, g. m. (2018). *modelos de localización para geomarketing*. espacialidades, revista de temas contemporáneos sobre lugares, política y cultura, 8(1), 95-120.
- Suárez Varela, a. (2023). *sistema de recomendación geoespacial para la localización de nuevos hoteles en Madrid*. [trabajo de fin de grado, universidad complutense de Madrid].
- IFRS Foundation. (2018). *marco conceptual para la información financiera*.
- Liozu, m., & Ulaga, w. (2018). *monetizing data: a practical roadmap for framing, pricing & selling your b2b digital offers*. value Innovation advisor publishing.
- quijada, c. a. (2022). *monetizando los datos*. realidad empresarial, (13), 52-55.

- Schmarzo, b. (2020). *the economics of data, analytics, and digital transformation: the theorems, laws, and empowerments to guide your organization's digital transformation*. Packt publishing.

WEBGRAFÍA

Web 1. www.ccs.cl.

Web 2. www.latamnetworks.es.

Web 3. www.100franquicias.cl.

Web 4. <http://hdl.handle.net/10045/12077>

Web 5. www.xmap.ai.

Web 6. https://www.researchgate.net/publication/326426180_modelos_de_localizacion_para_geomarketing

Web 7. <https://www.ctcp.gov.co/proyectos/contabilidad-e-informacion-financiera/documentos-organismos-internacionales/enmiendas-iasb-emitidas-en-2018/5-c-marco-conceptual-para-la>

Web 8. <http://resultados.censo2017.cl/Region?R=R0>

Web 9. <http://www.censo2017.cl>

Web 10. <https://doi.org/10.51378/reuca.v1i13.7049>

Web 11. <https://geoarchivos.ine.cl/File/pub/población-y-vivienda-valparaíso.pdf>

Web 12. http://www.censo2017.cl/wp-content/uploads/2017/12/Presentacion_Resultados_Definitivos_Censo2017.pdf

Web 13. <https://regiones.ine.cl/valparaiso/prensa/el-ingreso-laboral-promedio-en-la-región-de-valparaíso-fue-de-601.402-pesos-mensual-en-2021#:~:text=El%20ingreso%20laboral%20promedio%20en,601.402%20pesos%20mensual%20en%202021>.

Web 14. <https://www.bcn.cl/siit/nuestropais/region5/indica.htm>

Web 15. <https://www.bcn.cl/siit/nuestropais/region5/indica.htm>

Web 16.

https://www.researchgate.net/publication/358879328_El_servicio_delivery_como_tendencia_microempresarial_ante_la_emergencia_sanitaria_por_el_COVID-19_en_la_ciudad_de_Cuenca

Web 17.

https://www.researchgate.net/publication/374809270_EFECTO_DEL_DELIVERY_

COMO_CANAL_DE_DISTRIBUCION_EN_EL_CRECIMIENTO_DE_NEGOCIOS_
CAPITULO_VIII

Web 18.

https://www.researchgate.net/publication/366086106_Fairwork_Paraguay_2022_Estandares_Laborales_en_la_Economia_de_las_Plataformas

Web 19. https://www.ine.gob.cl/docs/default-source/proyecciones-de-poblacion/publicaciones-y-anuarios/base-2017/estimaciones-y-proyecciones-2002-2035_base-2017_reg_area_infografía.pdf?sfvrsn=5d8b3bcc_5

Web 20. <https://www.ine.gob.cl/estadisticas/sociales/demografia-y-vitales/demografia-y-migracion>

Web 21. http://www.fundacionfuturo.cl/wp-content/themes/fund_futuro_theme/img/pdf/censo-2018.pdf

Web 22. <https://www.memoriachilena.gob.cl/602/w3-article-31530.html#>