



**SmartPrice Analytics: Diseño de Plataforma de Monitoreo y Análisis de Precios
de Electrodomésticos**

Yenny Alexandra Gonzalez Melendez

Daniel Felipe Fuquene Linares

Mateo Andres Duque Moya

Universidad EAN
Facultad de Ingeniería
Maestría en Ciencia de Datos
Bogotá, Colombia
20/02/2026

**SmartPrice Analytics: Diseño de Plataforma de Monitoreo y Análisis de Precios
de Electrodomésticos**

Yenny Alexandra Gonzalez Melendez

Daniel Felipe Fuquene Linares

Mateo Andres Duque Moya

Trabajo de grado presentado como requisito para optar al título de:

Magíster en Ciencia de datos

Director:

Diego Armando García García

Modalidad:

Trabajo dirigido

Universidad EAN

Facultad de Ingeniería

Maestría en Ciencia de Datos

Bogotá, Colombia

20/02/2026

Nota de aceptación:

Firma del jurado

Firma del jurado

Firma del director del trabajo de grado

Bogotá, 20/02/2026

Dedico este trabajo de grado a mi familia,
que ha sido mi mayor fuente de fuerza,
motivación y amor incondicional.

A mis padres, por enseñarme el valor del
esfuerzo, la disciplina y la perseverancia,
y por creer en mí incluso en los momentos
en que yo dudé.

A quienes me acompañaron en este
camino con palabras de ánimo, paciencia
y confianza, recordando siempre que cada
paso, por pequeño que fuera, me
acercaba a mis metas.

Este logro es tanto mío como de ustedes.
Gracias por ser mi motivo y mi impulso.

Agradecimientos

Este trabajo, titulado *SmartPrice Analytics*, ha sido posible gracias al apoyo, la orientación y la confianza de muchas personas e instituciones que nos acompañaron en cada etapa de su desarrollo. En primer lugar, expresamos nuestro sincero agradecimiento a Vitovidrios S.A.S., cuya disposición para colaborar y facilitar información fue fundamental para transformar esta investigación en un proyecto con verdadero impacto práctico.

A nuestro director de tesis, Diego Armando García García, le debemos mi profunda gratitud por su guía constante, sus observaciones oportunas y el rigor académico que enriqueció significativamente esta propuesta. Extendemos también nuestro agradecimiento al comité evaluador, cuyo criterio y retroalimentación permitieron fortalecer la calidad del trabajo.

A nuestros profesores de la maestría en Ciencia de Datos, gracias por compartir los conocimientos, herramientas y perspectivas que se convirtieron en pilares metodológicos durante esta investigación. De igual manera, agradecemos a nuestros colegas por su apoyo en el proceso de recolección y validación de datos, así como por sus aportes técnicos desde la experiencia práctica del sector.

No podríamos dejar de agradecer a nuestras familias y amigos, quienes con su apoyo incondicional, su paciencia y sus palabras de ánimo fueron un impulso constante para avanzar incluso en los momentos más exigentes. Finalmente, reconocemos el valioso aporte de las comunidades y herramientas de código abierto que hicieron posible la implementación técnica de esta plataforma.

Resumen

Este trabajo diseña e implementa SmartPrice Analytics, una plataforma de monitoreo y análisis de precios para Vitovidrios S.A.S., pyme colombiana dedicada a la comercialización de electrodomésticos. La investigación aborda las dificultades de la empresa para fijar precios competitivos en un entorno de alta volatilidad, fuerte presión de grandes superficies y limitada información en tiempo real. Mediante un enfoque aplicado, descriptivo y mixto, se integran técnicas de web scraping, ingeniería de características y modelos de machine learning supervisado. La plataforma combina datos transaccionales internos con precios de competidores (Éxito, Alkosto, Falabella) capturados automáticamente, construyendo variables temporales, estacionales y de posicionamiento competitivo. El sistema se basa en tres modelos complementarios: un regresor XGBoost para predicción de precio óptimo (MAE: \$60.000 COP, RMSE: \$182.000 COP), un clasificador Random Forest para decisión de descuentos (precisión: 92%, recall: 98%) y un regresor Random Forest para magnitud del descuento (MAE: 2,59%). Los resultados evidencian errores inferiores al 10% en televisores, lavadoras y celulares, demostrando que la plataforma mejora la precisión en pricing, reduce riesgos financieros y fortalece la competitividad de la empresa.

Palabras clave: Pyme, Algoritmos supervisados, modelos de series de tiempo, Web Scraping, Python

Abstract

This thesis designs and implements SmartPrice Analytics, a price monitoring and analysis platform for Vitovidrios S.A.S., a Colombian SME specializing in consumer electronics retail. The study addresses the company's challenges in setting competitive prices amid high market volatility, pressure from large retailers and digital platforms, and limited real-time information. Using an applied, descriptive, and mixed-methods approach, the research integrates web scraping, feature engineering, and supervised machine learning models. The platform combines internal transactional data with competitor prices (Éxito, Alkosto, Falabella) collected automatically, generating temporal, seasonal, and competitive positioning variables. The intelligent pricing system comprises three complementary models: an XGBoost regressor for optimal price prediction (MAE: COP 60,000; RMSE: COP 182,000), a Random Forest classifier for discount decisions (precision: 92%, recall: 98%), and a Random Forest regressor for discount magnitude estimation (MAE: 2.59%). Results show mean absolute percentage errors below 10% across televisions, washing machines, and smartphones, demonstrating that the platform improves pricing accuracy, reduces financial risk, and strengthens the competitive position of Vitovidrios S.A.S. against larger market players.

Keywords: SME, Supervised algorithms, Time series models, Web Scraping, Python

Tabla de Contenido

Lista de Figuras	12
Lista de Tablas	13
Introducción	14
Objetivos	17
Objetivo general	17
Objetivos específicos	17
Justificación	18
Marco Institucional	20
Misión:	20
Visión:	20
Valores corporativos:	21
Políticas de calidad:	21
Estructura organizacional: mapa de procesos y organigrama	21
Productos o servicios ofertados	23
Análisis del sector	24
Metodología – Descripción de Obtención de Datos	24
Alcance del Proceso de Obtención de Datos	24
Entorno Legal y Consideraciones Éticas	25
Arquitectura del Sistema	25

SmartPrice Analytics: Diseño de Plataforma de Monitoreo y Análisis de Precios de Electrodomésticos	9
Programación y Automatización	27
Estructuración, validación y almacenamiento de la información	27
Marco Contextual y Conceptual	30
Marco Conceptual	30
El concepto de inteligencia de negocios	30
Modelado de datos	31
Estrategias de Pricing	32
Monitoreo de Precios mediante Web Scraping	33
Diseño Metodológico	46
Diagnóstico Organizacional	46
Análisis externo - PESTEL	46
Político	46
Económico	47
Social	47
Tecnológico	47
Ecológico	48
Legal	48
Análisis interno	49
Análisis interno basado en encuesta a colaboradores	50
Análisis Exploratorio Precios Competidores	52
Resumen de Datos	52

SmartPrice Analytics: Diseño de Plataforma de Monitoreo y Análisis de Precios de Electrodomésticos	10
Análisis de datos	53
Evolución temporal de precios	53
Distribución de precios	54
Análisis de Marcas	56
Autocorrelaciones	57
Celulares – Dinámica Altamente Reactiva	58
Televisores – Persistencia y Estacionalidad	59
Lavadoras – Alta Estabilidad Estructural	59
Análisis Exploratorio Base Interna	60
Evolución Anual - Ingresos y Unidades	60
Análisis de marcas	61
Estructura de Precios	63
Modelo	64
Entrada y Preparación de Datos	65
Entrenamiento y Modelos Predictivos	66
Resultado de entrenamiento	67
Análisis de Error	69
Análisis de los resultados	71
Resultados de la Solución	73
Conclusiones y Recomendaciones	76
Recomendaciones	77

SmartPrice Analytics: Diseño de Plataforma de Monitoreo y Análisis de Precios de Electrodomésticos	11
Referencias	79
ANEXOS	83
ANEXO 1. Diagnóstico interno	83
ANEXO 2. Ficha técnica por demanda histórica.	85
ANEXO 3. Ficha técnica por precios de los competidores.	87
ANEXO 4. Ficha técnica por promociones aplicadas.	88
ANEXO 5. Ficha técnica por Estacionalidad.	90
ANEXO 6. Ficha técnica por nivel de inventario.	91

Lista de Figuras

Figura 1. Organigrama Vitovidrios S.A.S.	22
Figura 2. Mapa de procesos de Vitovidrios S.A.S	23
Figura 3. Retos manejo de datos interno Vitovidrios	50
Figura 4. Fases del proceso	53
Figura 5. Evolución temporal de precios por categoría	54
Figura 6. Distribución de precios	55
Figura 7. Referencias más vendidas por artículo de venta	57
Figura 8. Autocorrelación Ventas y Precio trimestral - Celulares	59
Figura 9. Autocorrelación Ventas y Precio trimestral - Televisores	59
Figura 10. Autocorrelación Ventas y Precio trimestral - Lavadoras	60
Figura 11. Serie temporal de ingresos diarios - PYME	61
Figura 12. Marcas más vendidas por Vitovidrios S.A.S.	62
Figura 13. Precio final Vitovidrios S.A.S.	63
Figura 14. Métricas claves del sistema.	68
Figura 15. Comparación del Error Absoluto Medio (MAE) por Categoría de Producto	70
Figura 16. Plataforma interactiva final	75

Lista de Tablas

Tabla 1. Estructura de datos en Vitovidrios	26
Tabla 2. Tecnologías seleccionadas	37
Tabla 3. Riesgos técnicos y mitigaciones	37
Tabla 4. . Costos relacionados al desarrollo del proyecto	39
Tabla 5. Almacenamiento	39

Introducción

El presente trabajo de grado se desarrolla en el marco de la Maestría en Ciencia de Datos y se inscribe en la línea de investigación de Gestión y Diseño de Procesos del grupo Tecnológico Ontare. El estudio se centra en el diseño de una plataforma de monitoreo y análisis de precios para el sector de electrodomésticos en Colombia, con aplicación directa en la empresa Vitovidrios S.A.S., una pyme dedicada a la comercialización de productos para el hogar y electrodomésticos. El proyecto articula conceptos y herramientas de analítica de datos, *machine learning* y *business intelligence*, con el propósito de desarrollar una solución tecnológica que fortalezca la toma de decisiones estratégicas de pricing en contextos empresariales altamente competitivos.

El sector de electrodomésticos se caracteriza por una elevada volatilidad de precios, impulsada por factores como la estacionalidad de la demanda, la proliferación de campañas promocionales agresivas, la entrada de competidores internacionales y la creciente digitalización de los canales de venta. Desde una perspectiva estratégica, Eslava (2015) sostiene que la fijación de precios en mercados dinámicos no puede basarse únicamente en costos internos, sino que debe incorporar información sistemática del entorno competitivo. En esta misma línea, Rivero Gutiérrez, Samino García y Cerro Rodríguez (2023) enfatizan que, en entornos digitales, el pricing se convierte en un proceso continuo de ajuste que depende de la capacidad de las empresas para monitorear el mercado en tiempo real y responder de forma ágil a los cambios en precios y promociones de la competencia.

Sin embargo, aunque la literatura coincide en la importancia del pricing basado en datos, existe una brecha significativa entre los enfoques teóricos y su aplicación práctica en pequeñas y medianas empresas. Nagle y Müller (2018) destacan que el precio es uno de los principales determinantes de la rentabilidad empresarial, pero advierten que muchas organizaciones carecen de metodologías analíticas formales para su gestión. Phillips (2020) profundiza en esta idea al señalar que la optimización de precios mediante modelos matemáticos y analíticos permite maximizar ingresos y márgenes, siempre que se cuente con datos de calidad y modelos capaces de capturar la dinámica del mercado.

No obstante, estos enfoques suelen estar orientados a grandes corporaciones con capacidades tecnológicas avanzadas, lo que limita su adopción en pymes.

En el caso específico de Vitovidrios S.A.S., la ausencia de herramientas analíticas para el monitoreo sistemático de precios ha derivado en decisiones predominantemente reactivas y empíricas. Los antecedentes empíricos evidencian pérdidas de ingresos y desajustes en inventario asociados a estrategias de precios poco alineadas con el mercado: precios fijados por encima de la media han reducido el volumen de ventas, mientras que precios excesivamente bajos han erosionado márgenes de rentabilidad sin generar ventajas competitivas sostenibles. Esta situación confirma lo planteado por Phillips (2020), quien argumenta que decisiones de pricing no informadas por datos tienden a generar ineficiencias tanto en ingresos como en rentabilidad.

Desde el punto de vista tecnológico, diversos autores resaltan el potencial de integrar técnicas de *web scraping* y *machine learning* para el monitoreo y análisis de precios en tiempo real. Mitchell (2018) y Gupta (2018) destacan que el *web scraping* permite recolectar información de mercado de manera automatizada y escalable, mientras que los modelos supervisados ofrecen la capacidad de aprender patrones complejos a partir de datos históricos y competitivos. Sin embargo, gran parte de la literatura se enfoca en aplicaciones aisladas de estas técnicas, sin proponer arquitecturas integrales que articulen la predicción de precios, la decisión de descuentos y la estimación de su magnitud en un sistema coherente, especialmente orientado a pymes del sector comercial.

A partir de este contexto, el presente trabajo plantea como hipótesis principal que la implementación de una plataforma de pricing inteligente basada en datos históricos internos y precios de competidores, utilizando modelos supervisados de *machine learning*, permite mejorar significativamente la precisión en la fijación de precios y la toma de decisiones de descuento en una pyme del sector de electrodomésticos. De manera específica, se plantean las siguientes hipótesis: (i) los modelos de *gradient boosting* permiten estimar precios base óptimos con errores relativos inferiores al 10% respecto al precio promedio del portafolio; (ii) los clasificadores basados en *ensemble methods* mejoran la identificación de escenarios donde es necesario aplicar descuentos, reduciendo el riesgo de pérdida de margen; y (iii) la predicción cuantitativa de la magnitud

del descuento contribuye a un ajuste más fino del pricing sin comprometer la competitividad.

En este sentido, el trabajo dirigido busca no solo aportar una solución tecnológica aplicada al caso de Vitovidrios S.A.S., sino también contribuir al cuerpo de conocimiento sobre analítica de precios en pymes, demostrando que es posible adaptar técnicas avanzadas de ciencia de datos a contextos empresariales con recursos limitados, promoviendo así la transformación digital y la toma de decisiones basada en datos en el comercio minorista colombiano.

El objeto de diagnóstico de este trabajo se centra, por tanto, en la estrategia de fijación de precios de Vitovidrios S.A.S., identificando como problema principal la ausencia de una plataforma que automatice la captura y análisis de datos de mercado. La descripción del problema enfatiza que, al no contar con una solución tecnológica adecuada, la empresa depende de ajustes manuales y tardíos, con una baja capacidad de anticipación frente a las fluctuaciones de precios y a los cambios en la demanda. Esto limita su competitividad frente a grandes superficies y plataformas digitales que sí utilizan analítica avanzada como mecanismo de diferenciación.

La pregunta de investigación que orienta este proyecto es:

¿Cómo puede una plataforma de monitoreo y análisis de precios optimizar la estrategia de pricing de Vitovidrios S.A.S., permitiéndole identificar el mejor precio diario y el momento óptimo para ofertar con mayor impacto en el mercado de electrodomésticos?

En cuanto a la estructura del documento, este se organiza en varios capítulos. En primer lugar, se presenta el marco conceptual y teórico, que incluye la revisión de literatura sobre estrategias de pricing, inteligencia de negocios, web scraping y análisis predictivo. Posteriormente, se aborda el planteamiento metodológico, en el que se define el enfoque aplicado, descriptivo y mixto de la investigación, detallando la población, muestra y diseño de instrumentos de recolección de datos. A continuación, se expone el diagnóstico situacional de Vitovidrios S.A.S. mediante el análisis externo (PESTEL) e interno, evidenciando las brechas en su estrategia actual de precios. El siguiente capítulo describe el diseño de la plataforma SmartPrice Analytics, incluyendo las capas de extracción, procesamiento, almacenamiento y visualización de datos, así como los modelos predictivos aplicados. Finalmente, se presentan los resultados esperados,

conclusiones y recomendaciones, donde se discuten los aportes de la investigación tanto para Vitovidrios S.A.S. como para el sector de electrodomésticos en general.

Objetivos

Objetivo general

Diseñar una plataforma de monitoreo y análisis de precios como herramienta de apoyo estratégico que permita a Vitovidrios S.A.S. optimizar su estrategia de pricing en el mercado de electrodomésticos.

Objetivos específicos

- Identificar en la literatura los referentes teóricos sobre estrategias de pricing y modelos de análisis de precios, con el fin de obtener las variables clave que definirán el desarrollo de la plataforma para Vitovidrios SAS.
- Realizar un análisis situacional de Vitovidrios SAS para conocer su estado actual en cuanto a fijación de precios, comportamiento del mercado y gestión de ventas, identificando oportunidades de mejora en su estrategia comercial.
- Diseñar una plataforma de monitoreo y análisis de precios que permita a Vitovidrios SAS optimizar su estrategia de precios y maximizar su rentabilidad.
- Establecer un plan de implementación para la plataforma propuesta, definiendo los recursos, metodologías y estrategias necesarias para su desarrollo y adopción dentro de Vitovidrios SAS.

Justificación

La implementación de una plataforma de monitoreo y análisis de precios en Vitovidrios S.A.S. representa una oportunidad clave para fortalecer su competitividad y sostenibilidad en el mercado de electrodomésticos. La empresa enfrenta actualmente dificultades derivadas de la volatilidad de precios, la ausencia de información en tiempo real y la carencia de estrategias basadas en datos que le permitan gestionar de manera eficiente tanto su política de precios como su inventario. Estas limitaciones han derivado en decisiones reactivas, pérdida de márgenes de ganancia y falta de aprovechamiento de oportunidades comerciales en periodos de alta demanda.

El desarrollo de la plataforma permitirá automatizar el monitoreo de precios, identificar patrones de fluctuación y diseñar estrategias óptimas de oferta, reduciendo así la incertidumbre en la toma de decisiones comerciales. Entre los beneficios esperados se encuentra la optimización de la estrategia de precios mediante el ajuste dinámico en función de datos actualizados, lo que evita fijaciones excesivamente altas que disuadan a los clientes o demasiado bajas que afecten los márgenes de rentabilidad. Asimismo, la empresa podrá anticipar tendencias de consumo, aprovechar picos de demanda y gestionar de manera más eficiente su inventario, reduciendo riesgos de sobrestock o desabastecimiento.

Además de mejorar la rentabilidad de Vitovidrios S.A.S., la plataforma fortalecerá la competitividad de la empresa al permitirle reaccionar con mayor agilidad frente a los cambios del mercado y adoptar un enfoque más proactivo basado en datos. La automatización del análisis reducirá la carga operativa y el error humano, liberando tiempo valioso para que el equipo comercial se concentre en estrategias de expansión y diferenciación. De igual forma, la integración de herramientas de *Big Data* garantizará la recolección, procesamiento y análisis de información en tiempo real, asegurando la viabilidad técnica del proyecto y su alineación con las capacidades actuales de la organización.

Este trabajo no solo tendrá impacto en Vitovidrios S.A.S., sino también a nivel sectorial. La introducción de un modelo de *pricing* basado en analítica establece un precedente de

transformación digital en la fijación de precios para el sector de electrodomésticos, lo que puede motivar a otras empresas a implementar soluciones similares y elevar el estándar competitivo. Asimismo, la existencia de sistemas de monitoreo automatizados contribuye a una mayor transparencia en los precios, reflejando de manera más precisa las dinámicas de oferta y demanda. Para los consumidores, ello se traduce en mejores precios, mayor disponibilidad de productos y una experiencia de compra más satisfactoria.

Un elemento clave del proyecto es la aplicación de una metodología de diagnóstico y plan de mejora que asegura su pertinencia y efectividad. Este enfoque permite a la empresa conocer con claridad su punto de partida en términos de estrategia de precios, identificar brechas y priorizar oportunidades de alto impacto. Al mismo tiempo, facilita el diseño de una solución personalizada que responde a las necesidades reales de la organización, incrementando la probabilidad de adopción y éxito. El carácter gradual de la metodología promueve, además, una cultura de mejora continua que trasciende el proyecto y garantiza la sostenibilidad de sus resultados a largo plazo.

En conjunto, la conveniencia y justificación de esta investigación se sustentan en la necesidad de transformar la gestión de precios en Vitovidrios S.A.S. mediante el uso de analítica avanzada. El proyecto es viable técnica y operativamente, gracias al acceso a información relevante, al compromiso de la alta dirección y a la disponibilidad de recursos básicos tecnológicos. Su implementación aportará beneficios tangibles para la empresa y generará conocimiento replicable en el sector de comercio digital, consolidando un modelo de *pricing* inteligente que contribuye al fortalecimiento de la competitividad empresarial en entornos altamente dinámicos.

Marco Institucional

Vitovidrios S.A.S es una pequeña y mediana empresa (Pyme) del sector comercio, dedicada a la comercialización de electrodomésticos y artículos para el hogar en el mercado colombiano. Fundada por el señor Luis Carlos Duque, la empresa cuenta actualmente con un equipo de 10 empleados, lo que le permite operar con una estructura organizativa ágil y flexible, adaptándose rápidamente a las exigencias del mercado.

La empresa se ha caracterizado por ofrecer productos de calidad a precios competitivos, buscando siempre la satisfacción del cliente y el fortalecimiento de su posición en el mercado. Aunque inicialmente centrada en productos de vidrio y hogar, Vitovidrios S.A.S ha ampliado su portafolio hacia electrodomésticos, identificando en este segmento una oportunidad de crecimiento sostenible.

Su modelo de negocio se enfoca en la venta directa al consumidor final, con presencia física en puntos de venta y estrategias incipientes de comercialización en canales digitales. A pesar de ser una Pyme, la organización ha demostrado interés en adoptar tecnologías emergentes, como la analítica de datos, para mejorar su gestión comercial y su toma de decisiones estratégicas.

Misión:

Ofrecer a nuestros clientes productos para el hogar y electrodomésticos de excelente calidad a precios competitivos, garantizando un servicio personalizado, confiable y orientado a la satisfacción del cliente.

Visión:

Ser reconocidos a nivel regional como una empresa líder en la comercialización de electrodomésticos, caracterizada por su innovación, eficiencia en procesos y compromiso con el cliente.

Valores corporativos:

- Compromiso: con la satisfacción del cliente y el cumplimiento de los objetivos empresariales.
- Transparencia: en los procesos comerciales, de atención y postventa.
- Innovación: en la adopción de herramientas tecnológicas que optimicen la gestión empresarial.
- Trabajo en equipo: promoviendo una cultura organizacional colaborativa y participativa.
- Responsabilidad social: en el trato justo con proveedores, empleados y el entorno comunitario.

Políticas de calidad:

Vitovidrios S.A.S. mantiene una política de mejora continua en sus procesos comerciales y operativos, enfocada en:

- Garantizar productos con excelente relación calidad-precio.
- Implementar mecanismos de control de inventario eficientes.
- Ofrecer un servicio al cliente oportuno, cordial y eficaz.
- Adoptar tecnologías que fortalezcan la toma de decisiones y la estrategia comercial.

Estructura organizacional: mapa de procesos y organigrama

Vitovidrios S.A.S., debido a su tamaño, posee una estructura organizacional horizontal, donde las jerarquías son reducidas y los procesos se gestionan de forma colaborativa entre áreas. A continuación se describe el organigrama funcional básico de la empresa:

Organigrama:

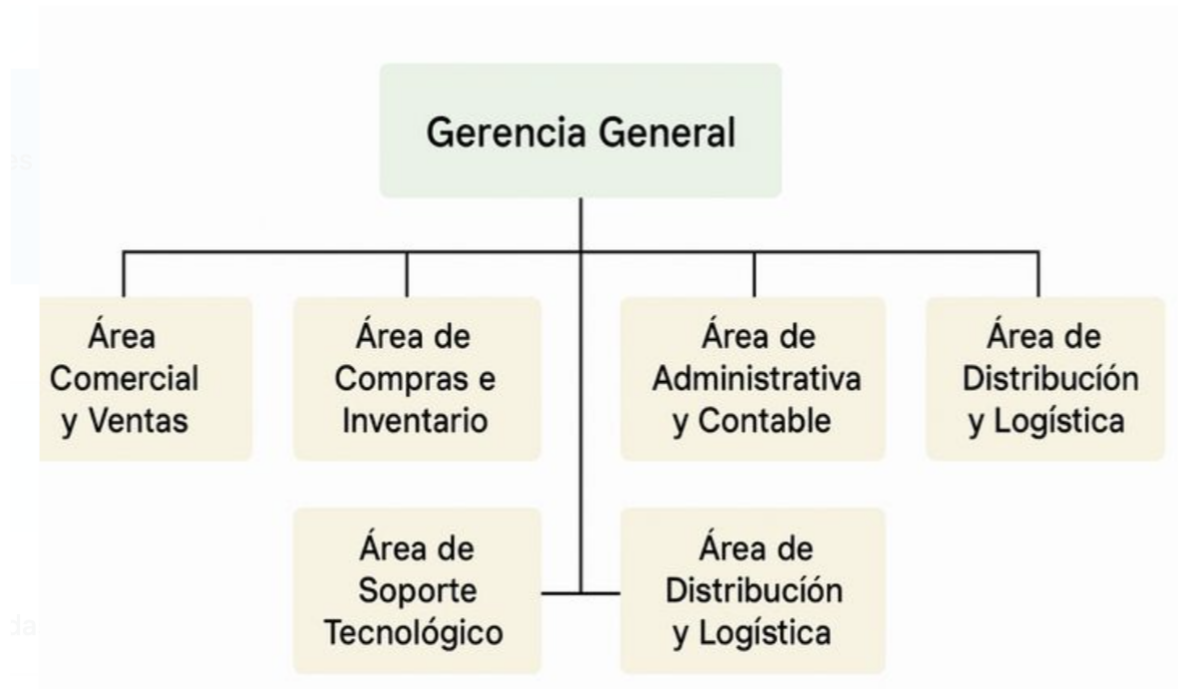


Figura 1. Organigrama Vitovidrios S.A.S.

Gerencia General (Luis Carlos Duque): Toma de decisiones estratégicas, liderazgo general de la empresa.

Área Comercial y Ventas: Encargada de atención al cliente, seguimiento de ventas y estrategias de comercialización.

Área de Compras e Inventario: Gestión de proveedores, control de inventario, reposición de stock.

Área Administrativa y Contable: Manejo financiero, nómina, registros contables, facturación.

Área de Soporte Tecnológico: Asistencia en infraestructura digital, soporte a procesos de ventas en línea.

Área de Distribución y Logística: Encargada de la entrega de productos, control de despachos y servicio postventa.

Mapa de procesos:

Mapa de procesos

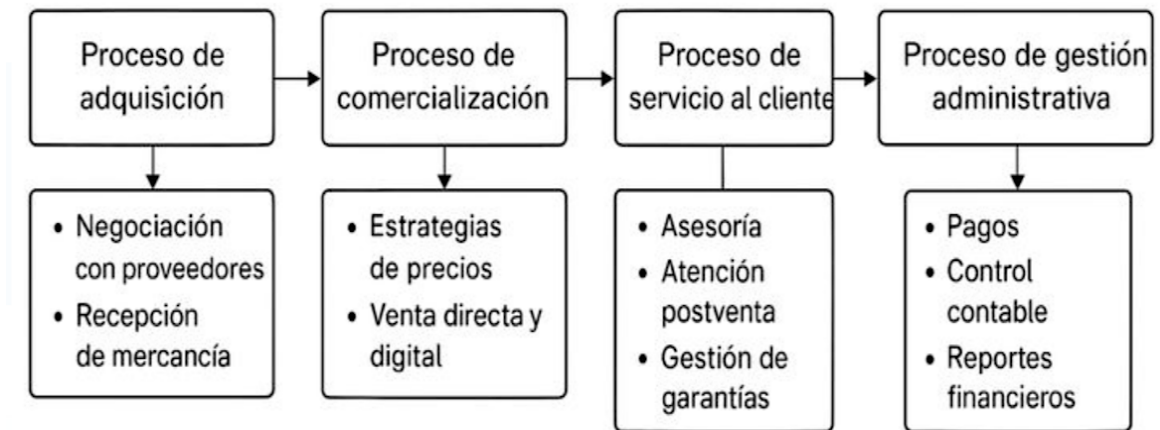


Figura 2. Mapa de procesos de Vitovidrios S.A.S

- Proceso de adquisición: negociación con proveedores, recepción de mercancía.
- Proceso de comercialización: estrategias de precios, venta directa y digital.
- Proceso de servicio al cliente: asesoría, atención postventa, gestión de garantías.
- Proceso de gestión administrativa: pagos, control contable, reportes financieros.

Este esquema organizativo permite una operación ágil, con roles múltiples asumidos por algunos colaboradores, lo que facilita una comunicación directa y una rápida toma de decisiones.

Productos o servicios ofertados

Vitovidrios S.A.S ofrece un portafolio amplio de productos, con especial énfasis en electrodomésticos para el hogar. Entre los principales productos ofertados se encuentran:

- Electrodomésticos de línea blanca: neveras, estufas, lavadoras, microondas.
- Pequeños electrodomésticos: licuadoras, cafeteras, planchas, ventiladores.
- Televisores y equipos de audio.
- Promociones y combos de electrodomésticos.

Además, la empresa ofrece servicio de entrega a domicilio, asesoría personalizada y gestión de garantías de productos. En ocasiones especiales (temporadas navideñas, días sin IVA), se promueven campañas de descuentos y ventas relámpago.

Análisis del sector

El sector de electrodomésticos en Colombia ha mostrado un crecimiento sostenido en los últimos años, impulsado por la digitalización del comercio, el aumento en la demanda de productos tecnológicos y las políticas de incentivo al consumo interno. Sin embargo, también se caracteriza por una alta competencia y fluctuaciones constantes en los precios debido a factores como:

- Ofertas promocionales agresivas por parte de grandes superficies y tiendas online.
- Importación de productos a precios bajos desde mercados asiáticos.
- Demanda estacional, que varía según eventos, fechas especiales y el ciclo económico.

Vitovidrios S.A.S. compite principalmente con tiendas locales, cadenas especializadas como Alkosto o Éxito, y plataformas de comercio electrónico como Mercado Libre y Linio. En este contexto, se enfrenta a los siguientes retos clave:

- Competitividad en precios: debe ofrecer precios atractivos sin sacrificar margen de rentabilidad.
- Adaptación a entornos digitales: fortalecer su presencia online y explorar nuevos canales de venta.
- Gestión de inventarios en tiempo real: para evitar sobrecostos o escasez de productos.
- Respuestas ágiles al mercado: ajustarse rápidamente a las tendencias de consumo y ofertas de la competencia.

Frente a estos desafíos, surge la necesidad de implementar herramientas tecnológicas avanzadas, como el análisis de precios, que permitan tomar decisiones basadas en datos, maximizar la rentabilidad y mejorar la experiencia del cliente.

Metodología – Descripción de Obtención de Datos

Alcance del Proceso de Obtención de Datos

Una vez establecido el contexto del problema de negocio, el marco conceptual y la relevancia estratégica del análisis de precios en entornos competitivos, se procede a

describir la metodología empleada para el desarrollo del proyecto SmartPrice Analytics.

Esta metodología se fundamenta en un enfoque cuantitativo aplicado, propio de los proyectos de ciencia de datos, y sigue un flujo estructurado que permite transformar datos de mercado en información útil para la toma de decisiones comerciales.

El proceso de obtención de datos se basa en la recolección periódica de información publicada en los catálogos en línea de los tres principales mayoristas del mercado colombiano de electrodomésticos y tecnología.

Las categorías actualmente incluidas en el sistema son:

1. Lavadoras
2. Celulares
3. Televisores

El objetivo principal es construir un repositorio estructurado que permita realizar análisis comparativos de precios, identificar variaciones temporales y detectar oportunidades de mejora en la estrategia comercial y de abastecimiento de la empresa Vitovidrios S.A.S.

La recolección se realiza una vez por semana, debido a que en este tipo de productos los cambios en precios no suelen ser diarios, sino asociados a temporadas, promociones o estrategia comercial general. Esta periodicidad permite eficiencia operativa, evitando capturas duplicadas sin pérdida de información relevante.

Entorno Legal y Consideraciones Éticas

El proceso se fundamenta en el uso de información de acceso público, disponible libremente en los sitios web de los mayoristas, sin vulneración de sistemas de autenticación, propiedad intelectual o acuerdos de confidencialidad.

Se respeta la Ley 1581 de 2012 (Protección de Datos Personales), dado que no se recolecta información personal ni privada.

- No se realizan prácticas de extracción agresiva (como cargas masivas simultáneas) que puedan afectar el rendimiento de las páginas.
- No se manipula contenido, se únicamente se observa y copia información visible para cualquier usuario, lo cual se considera uso legítimo de datos públicos.

Arquitectura del Sistema

La arquitectura implementada sigue un flujo sencillo, eficiente y escalable:

- **Extracción:**

Se utiliza Python + Selenium para simular la navegación de un usuario, recorrer categorías y capturar atributos como: nombre de producto, marca, capacidad, características técnicas, precio, disponibilidad y URL.

- **Limpieza y Normalización:**

Se aplican transformaciones con Pandas, donde:

Se estandarizan formatos

Se homogenizan unidades de medida

Se corrigen caracteres especiales

Se validan tipos de datos (string, numérico, fecha)

- **Estructuración de Datos:**

Cada producto se almacena con el siguiente esquema:

Campo	Tipo	Descripción
ID	String	Identificador único del registro
Nombre	String	Nombre del producto
Precio	Numérico	Precio publicado
Marca	String	Marca del producto
Capacidad / Pulgadas / Memoria	String	Atributo característico según categoría
Color	String	Color (si aplica)
URL	String	Enlace del producto
Sitio	String	Mayorista fuente
Disponible	boolean	Indicador de disponibilidad
Fecha_extracción	date	Fecha de captura
Fecha_carga	date	Fecha de carga a BigQuery

Tabla 1. Estructura de datos en Vitovidrios

- **Almacenamiento:**

El repositorio de datos se encuentra en Google Cloud Platform (GCP):

- Google Cloud Storage: respaldo histórico completo (archivos crudos).
- BigQuery: tabla estructurada optimizada para consultas y modelamiento.

- **Consumo:**

Los datos alimentan dashboards internos y el módulo analítico de la plataforma SmartPrice Analytics, permitiendo visualizaciones comparativas y exploración

temporal.

Programación y Automatización

El proceso de obtención de datos de SmartPrice Analytics fue diseñado bajo un enfoque híbrido, combinando automatización de navegación web con interpretación semántica asistida por modelos de lenguaje. Esta arquitectura permite desacoplar la extracción del contenido visual de la dependencia estricta de estructuras HTML estáticas, incrementando la robustez del sistema frente a cambios frecuentes en el diseño de los sitios web monitoreados.

En una primera etapa, el sistema utiliza Python junto con Selenium para automatizar la apertura de un navegador web real, simular la navegación de un usuario humano y recorrer de forma secuencial las distintas páginas de productos en las plataformas de comercio electrónico analizadas. Este proceso incluye la interacción con componentes dinámicos del sitio, tales como paginación, carga asincrónica de contenido y visualización de descuentos, permitiendo capturar de forma íntegra la información visible para cualquier usuario final.

Una vez cargada cada página, el sistema obtiene el contenido completo renderizado (HTML resultante y texto visible), el cual es enviado a un modelo de lenguaje (Gemini) encargado de interpretar el contenido de manera semántica. En lugar de depender exclusivamente de selectores XPath rígidos, el modelo recibe instrucciones estructuradas para identificar y extraer atributos relevantes como nombre del producto, precio base, precio con descuento, porcentaje de descuento, marca, categoría, disponibilidad y URL de referencia.

Este enfoque permite que, ante cambios en la estructura HTML o en la jerarquía de los elementos del DOM, el sistema continúe extrayendo la información requerida, siempre que los datos permanezcan visibles en la interfaz del sitio. De esta manera, la interpretación semántica realizada por el modelo de lenguaje actúa como una capa de abstracción que reduce la fragilidad típica de los scrapers tradicionales basados únicamente en reglas fijas.

Estructuración, validación y almacenamiento de la información

La información interpretada por el modelo de lenguaje es organizada automáticamente en estructuras tabulares normalizadas, garantizando un esquema homogéneo para los datos provenientes de las diferentes plataformas analizadas. Estas tablas incluyen campos estandarizados de precios, descuentos y atributos del producto, lo que facilita el análisis comparativo entre competidores.

Posteriormente, se aplican controles de validación automática sobre los datos estructurados, entre los cuales se incluyen:

verificación de formatos numéricos en precios y descuentos, detección de valores nulos o inconsistentes, eliminación de registros duplicados, estandarización de monedas y unidades de medida, y validaciones de coherencia entre precios originales y precios promocionales.

Una vez superadas estas validaciones, la información consolidada es cargada automáticamente en la infraestructura de Google Cloud Platform (GCP), utilizando Google Cloud Storage como respaldo histórico y BigQuery como repositorio analítico principal. Este esquema permite mantener trazabilidad temporal de los precios, facilitar consultas analíticas y alimentar directamente los modelos predictivos y dashboards del sistema SmartPrice Analytics.

El proceso se ejecuta de manera periódica para un conjunto de cuatro plataformas de comercio electrónico, garantizando una cobertura amplia del mercado y asegurando que la información recolectada sea consistente, actualizada y resiliente a cambios tecnológicos en los sitios fuente.

Extracción de datos y procesamiento

La arquitectura del sistema se basa en la ejecución periódica de un pipeline automático de recolección y preparación de datos, el cual se ejecuta una vez por semana. Durante cada ejecución, el sistema realiza procesos de web scraping sobre los principales portales de la competencia, recolectando la información disponible de precios y características de los productos.

Posteriormente, los datos extraídos son enviados a un servicio de procesamiento basado en modelos de lenguaje (Gemini), el cual se encarga de interpretar la información semiestructurada, estandarizar los atributos relevantes y consolidarlos en tablas estructuradas que incluyen precios, descuentos y características técnicas de los productos. Una vez procesada, esta información es almacenada en una infraestructura

en la nube (Google Cloud Platform – GCP), donde queda disponible para su explotación analítica.

La elección de una periodicidad semanal responde a las dinámicas del mercado de electrodomésticos en el contexto colombiano, en el cual los precios de la competencia no presentan variaciones continuas de alta frecuencia, sino ajustes discretos asociados principalmente a campañas promocionales específicas. Bajo este esquema, la información recolectada funciona como una referencia competitiva vigente para el periodo analizado.

A partir de esta base de datos actualizada semanalmente, los modelos predictivos pueden generar recomendaciones operativas de precio y descuento en cualquier momento, apoyándose adicionalmente en variables temporales, estacionales y en el comportamiento histórico de ventas. De esta forma, el sistema logra identificar el precio óptimo y el momento adecuado para aplicar ofertas sin requerir una actualización diaria del scraping, asegurando un equilibrio entre precisión analítica, estabilidad del modelo y viabilidad operativa.

Dentro del pipeline de recolección y preparación de datos, los modelos de lenguaje (Gemini) se utilizan de manera acotada y controlada en la etapa de postprocesamiento del scraping, específicamente después de la extracción del contenido HTML y antes del almacenamiento definitivo de los datos estructurados.

En esta fase, el modelo de lenguaje actúa como un componente de apoyo para interpretar contenido semiestructurado, homogenizar descripciones de productos y organizar atributos relevantes (precios, descuentos y especificaciones técnicas) en un formato tabular estándar. Es importante resaltar que el scraping y la captura de los elementos HTML se realizan mediante procesos determinísticos, y que Gemini no participa en la navegación, extracción primaria ni en la toma de decisiones de pricing.

Para garantizar reproducibilidad y consistencia, el uso del modelo de lenguaje se apoya en prompts fijos y controlados, con una estructura predefinida de salida (esquema de columnas esperado). Los resultados generados por Gemini son posteriormente validados mediante reglas determinísticas, tales como verificación de tipos de datos, rangos válidos de precios, consistencia entre descuentos y precios finales, y detección de valores atípicos. En caso de incumplimiento de estas reglas, los registros son marcados para reprocesamiento o descartados.

Adicionalmente, el pipeline incorpora mecanismos de control de calidad y manejo de errores, incluyendo validaciones cruzadas con registros históricos, revisión de duplicados y detección de cambios estructurales abruptos en la información extraída. Estos controles permiten mitigar posibles variaciones en la salida del modelo de lenguaje y asegurar estabilidad en los datos almacenados.

Finalmente, el uso de Gemini se concibe como un componente auxiliar y no crítico del sistema. En escenarios donde no se dispone del modelo de lenguaje o se detectan inconsistencias reiteradas, el pipeline puede operar utilizando reglas determinísticas de parsing previamente definidas, garantizando la continuidad del proceso y la integridad del conjunto de datos utilizado por los modelos predictivos.

Marco Contextual y Conceptual

El presente marco teórico sustenta el desarrollo del proyecto SmartPrice Analytics: Diseño de Plataforma de Monitoreo y Análisis de Precios de Electrodomésticos para la empresa Vitovidrios S.A.S., con el propósito de establecer las bases conceptuales y científicas que respaldan su implementación. El marco se divide en dos componentes: el marco conceptual, que aborda los principales conceptos relacionados con la ciencia de datos, la analítica y la estrategia de precios; y la revisión bibliográfica, que analiza investigaciones previas sobre metodologías y aplicaciones similares en el contexto empresarial actual

Marco Conceptual

El concepto de inteligencia de negocios

La inteligencia de negocios (BI, por sus siglas en inglés) se refiere al conjunto de estrategias, metodologías y tecnologías que permiten la recopilación, procesamiento y análisis de datos empresariales para mejorar la toma de decisiones. Su propósito es convertir los datos en información útil que facilite la identificación de tendencias, patrones y oportunidades estratégicas en el mercado.

En el contexto de Vitovidrios S.A.S, la inteligencia de negocios se convierte en un factor clave para optimizar su estrategia de fijación de precios en el sector de electrodomésticos. Al implementar una solución de BI basada en el análisis predictivo, la empresa podrá anticipar fluctuaciones de precios y mejorar su competitividad.

Modelado de datos

El modelado de datos es una técnica fundamental en inteligencia de negocios, ya que permite organizar y estructurar la información para facilitar su análisis. A través de diferentes modelos, es posible representar la realidad empresarial y optimizar la gestión de datos.

Concepto de modelado de datos

El modelado de datos consiste en la creación de estructuras lógicas que representan la información de una organización. Se emplean modelos conceptuales, lógicos y físicos que permiten organizar los datos de manera eficiente, facilitando su almacenamiento, procesamiento y análisis.

Para la plataforma SmartPrice Analytics, se requiere un modelo de datos capaz de integrar información de múltiples fuentes, como precios históricos, estrategias de la competencia y tendencias de consumo, con el fin de proporcionar insights valiosos para la toma de decisiones en tiempo real.

Mejores prácticas para el modelo

El éxito de un modelo de datos depende de la aplicación de buenas prácticas que garanticen su precisión y eficiencia. Entre las mejores prácticas para el desarrollo del modelo de SmartPrice Analytics se incluyen:

- Normalización de datos: Asegurar que la información esté bien estructurada y libre de redundancias.
- Integración de fuentes de datos: Unificar información de diferentes plataformas de venta para obtener una visión completa del mercado.
- Uso de modelos predictivos: Aplicar algoritmos de machine learning para anticipar variaciones en los precios y optimizar las estrategias de oferta.

- Automatización del procesamiento de datos: Implementar herramientas que actualicen y analicen la información en tiempo real.
- Seguridad y gobernanza de datos: Garantizar la confidencialidad, integridad y disponibilidad de la información.

Estrategias de Pricing

Definición de pricing y su importancia en mercados competitivos.

“La fijación de precios (pricing) es una herramienta estratégica fundamental que permite a las empresas determinar el valor de sus productos o servicios en función de diversos factores, como los costos, la competencia y la percepción del cliente. Una estrategia de precios adecuada no solo busca cubrir los costos y obtener beneficios, sino también posicionar el producto en el mercado, influir en la demanda y responder a las dinámicas competitivas. En mercados altamente competitivos, el pricing se convierte en un elemento clave para diferenciarse y captar la atención del consumidor para la empresa Vitovidrios en su meta de lograr ofrecer precios más competitivos en el mercado de los electrodomésticos.”(Eslava James,2015).

Tipos de estrategias de precios:

La fijación de precios constituye una de las decisiones estratégicas más críticas dentro de cualquier organización comercial, especialmente en mercados altamente competitivos como el de los electrodomésticos. Definir una estrategia de precios adecuada no solo impacta directamente en los márgenes de ganancia, sino que también influye en la percepción de valor del cliente, en el posicionamiento de marca y en la capacidad de respuesta frente a la competencia (Rivero Gutiérrez, Samino García, & Cerro Rodríguez, 2023).

Existen diversas metodologías para establecer precios, cada una adaptada a diferentes contextos de mercado, segmentos de consumidores y objetivos corporativos. Entre las estrategias más relevantes se encuentran el precio basado en coste más margen, el pricing dinámico, la fijación basada en los precios de la competencia y la estrategia basada en el valor percibido por el cliente. La correcta elección y ejecución de estas estrategias puede significar para Vitovidrios S.A.S no solo la optimización de su rentabilidad, sino también el fortalecimiento de su competitividad en el dinámico mercado de electrodomésticos.

Particularmente en entornos de comercio electrónico y omnicanalidad, donde la fluctuación de precios es constante y los consumidores comparan opciones en tiempo real, adoptar estrategias de pricing dinámico y orientadas al cliente resulta indispensable. Por tanto, comprender los fundamentos de cada tipo de estrategia permitirá sentar las bases conceptuales para el diseño de una plataforma de monitoreo y análisis de precios eficiente, orientada a maximizar los resultados comerciales de Vitovidrios S.A.S.

Costo más margen

Esta estrategia consiste en calcular el precio de un producto sumando un margen de beneficio al coste total de producción. Es una de las metodologías más tradicionales y se utiliza ampliamente en sectores donde los costes son fácilmente identificables.

Precios dinámicos.

Los precios dinámicos implican ajustar los precios en tiempo real o a intervalos regulares en respuesta a factores del mercado, como la demanda del cliente, la competencia y el inventario disponible. Esta flexibilidad permite a las empresas optimizar sus ingresos al maximizar el precio cuando la demanda es alta y reducirlo para atraer a los clientes cuando la demanda es baja.

Precios basados en la competencia.

Esta estrategia implica establecer precios en función de los precios de los competidores. Es común en mercados altamente competitivos donde los productos o servicios son similares, y las empresas buscan posicionarse estratégicamente en relación con sus rivales.

Precios basados en el valor percibido.

Esta estrategia se basa en el valor que el consumidor percibe del producto o servicio, más que en el coste de producción. Requiere una comprensión profunda del cliente y del mercado, y es especialmente efectiva para productos o servicios únicos o de lujo.

Monitoreo de Precios mediante Web Scraping

Definición de web scraping.

El web scraping es una técnica automatizada para extraer datos estructurados de sitios web, mediante el uso de programas que simulan la navegación humana. A través de esta técnica, se recolectan grandes volúmenes de información que, de otra forma, requerirían trabajo manual intensivo (Mitchell, 2018). De acuerdo con Ryan Mitchell, autor del libro *Web Scraping with Python*, esta práctica permite transformar páginas web no estructuradas en bases de datos utilizables para análisis posteriores, siendo clave en proyectos de inteligencia competitiva y pricing dinámico. El web scraping combina conocimientos de programación, análisis de datos y procesamiento de texto, requiriendo una comprensión profunda de estructuras HTML, DOM (Document Object Model) y protocolos de comunicación web (Mitchell, 2018).

Tecnologías y herramientas comunes (BeautifulSoup, Scrapy, Selenium).

El ecosistema de herramientas para web scraping ha evolucionado notablemente, permitiendo mayor eficiencia y escalabilidad.

BeautifulSoup es una biblioteca de Python que facilita la navegación, búsqueda y modificación de documentos HTML y XML de manera sencilla (Richardson, 2015). Se utiliza principalmente en tareas de scraping donde la estructura de las páginas es relativamente estática.

Scrapy, por otro lado, es un framework de scraping más avanzado que permite crear spiders (crawlers) capaces de recorrer múltiples páginas, manejar peticiones asíncronas y exportar datos en diferentes formatos, siendo ideal para proyectos de scraping a gran escala (Zhao et al., 2020).

Selenium se emplea para automatizar la interacción con páginas web dinámicas que requieren ejecutar JavaScript, simulando acciones humanas como hacer clic o llenar formularios, indispensable para acceder a contenidos cargados de manera asíncrona (Burns, 2018).

Aspectos éticos y legales en la recolección de datos.

Legalmente, algunos sitios web prohíben el scraping a través de sus Términos de Servicio o mediante archivos robots.txt (Gillespie, 2018). Ignorar estas directrices puede constituir violaciones contractuales o infracciones de derechos de propiedad intelectual. Además, en contextos de protección de datos personales (como el Reglamento General de Protección de Datos —GDPR— en Europa), es esencial considerar los derechos de privacidad de los usuarios (Kamarinou et al., 2016).

Éticamente, los investigadores y analistas deben respetar los principios de transparencia, consentimiento y minimización del daño al recolectar datos, asegurando que sus prácticas no afecten negativamente a los sitios web o a sus usuarios.

Diseño de una plataforma de monitoreo de precios

Capa de Extracción de Datos

La capa de extracción de datos representa el núcleo operativo inicial de la plataforma SmartPrice Analytics para Vitovidrios S.A.S. Su función es adquirir, a través de procesos automatizados, datos actualizados de precios de electrodomésticos publicados en plataformas de comercio electrónico competidoras, estructurando de manera óptima para su posterior procesamiento y análisis.

Según Russell (2018), una extracción de datos exitosa debe garantizar cuatro principios básicos: integridad, consistencia, oportunidad y eficiencia, ya que cualquier falla en esta etapa afecta toda la cadena de valor analítica.

Fundamentos Teóricos de la Extracción de Datos

El proceso de extracción de datos de páginas web, conocido como web scraping, implica interactuar programáticamente con documentos HTML para recolectar información contenida en el DOM (Document Object Model) (Mitchell, 2018).

A diferencia de los métodos tradicionales de recolección de datos, el scraping permite: Recolectar grandes volúmenes de información en tiempos muy reducidos.

Acceder a datos actualizados en tiempo real, fundamentales para estrategias dinámicas de negocio (Brynjolfsson & McAfee, 2014). Con esto Vitovidrios podrá tener información en tiempo real que permita tomar decisiones prontas respecto a la competencia.

Componentes Técnicos de la Capa de Extracción

Agendador de tareas (Scheduler):

Planifica la ejecución periódica de los spiders, evitando sobrecargar los servidores externos y asegurando que los datos se actualicen en ventanas horarias estratégicas (por ejemplo, 3:00 a.m.).

- Scrapers / Spiders:

Son programas que recorren sitios web, identifican patrones de estructura (DOM) y extraen los datos relevantes utilizando selectores XPath o CSS.

- Headless Browser Controller:

Cuando los datos no están directamente accesibles en el HTML inicial (contenido dinámico), se emplean navegadores sin interfaz gráfica como Headless Chrome controlados por Selenium.

- Middleware de rotación:

Implementa técnicas de rotación de agentes de usuario, IPs y tiempos de espera aleatorios para simular tráfico humano y evitar bloqueos por los sistemas anti-bots.

- Validador de datos:

Antes de almacenar, los datos extraídos pasan por validaciones de formato, unicidad y completitud, minimizando errores aguas abajo.

Flujo de Extracción en SmartPrice Analytics

La operación de la capa de extracción sigue un flujo lógico estricto:

- Identificación de fuentes de interés: Catálogos digitales de competidores (Éxito, Alkosto, Falabella).
- Definición de reglas de scraping: Localización de atributos específicos (precio, nombre, descuento, categoría).
- Construcción de spiders especializados: Utilizando Scrapy o Selenium, según sea necesario contenido estático o dinámico.
- Ejecución periódica de tareas: Tareas agendadas y controladas para cubrir ciclos de actualización de precios.
- Extracción, validación y transformación inicial de datos: Captura de datos crudos, validación de calidad y normalización básica (por ejemplo, convertir precios a formato numérico unificado).
- Almacenamiento temporal: Los datos validados se almacenan en bases intermedias (ej. MongoDB) para su procesamiento posterior.

- Notificación de errores o anomalías: Sistema de alertas en caso de cambios estructurales en los sitios o caídas de spiders.

Elección Justificada de Tecnologías

Justificación	Alternativas desestimadas
Framework altamente escalable, ideal para scraping estructurado masivo, soporte nativo de middleware y exportaciones a múltiples formatos.	Requests-HTML (limitado en grandes volúmenes).
Necesario para scraping de contenido dinámico renderizado vía JavaScript. Simula un navegador real.	Puppeteer (más pesado y enfocado a Node.js).
Base de datos NoSQL optimizada para datos semi-estructurados y de alta variabilidad como precios capturados.	MySQL (poco flexible ante variabilidad estructural).
Contenerización de spiders para despliegue flexible en diferentes entornos cloud.	Ejecución nativa en servidores físicos (menos portable).

Tabla 2. Tecnologías seleccionadas

Riesgos Técnicos y Mitigaciones

Riesgo	Mitigación
Cambios en el DOM de las páginas objetivo	Sistema de detección automática de cambios + actualización modular de spiders.
Bloqueo de IPs por detección de scraping	Rotación de proxies, User-Agent randomizados, y políticas de "back-off exponencial".
Latencia alta en scraping dinámico	Uso de técnicas de "Lazy Loading" inteligente + priorización de URLs.
Captura de datos incompletos	Validadores de campos obligatorios + alertas de inconsistencia.

Tabla 3. Riesgos técnicos y mitigaciones

Capa de procesamiento de datos (limpieza y normalización)

Una vez realizada la extracción, los datos recolectados presentan variaciones, errores y desorden propios de la naturaleza semi-estructurada de la web. La capa de procesamiento de datos tiene como función depurar, estructurar y preparar los datos para su uso analítico confiable en SmartPrice Analytics.

Fundamento Teórico

Han, Kamber y Pei (2012) argumentan que los datos en bruto rara vez son perfectos para el análisis: suelen contener valores atípicos, inconsistencias semánticas, duplicados y formatos heterogéneos. Si no se aplica un proceso de limpieza y normalización riguroso, el riesgo es obtener análisis sesgados y tomar decisiones comerciales erróneas.

Aplicación Específica para Vitovidrios

Para Vitovidrios S.A.S., esta capa asegurará que:

- Las comparaciones de precios entre competidores sean justas y exactas.
- Se pueda construir una base histórica limpia para detectar tendencias reales.
- Se eviten decisiones basadas en anomalías o datos incompletos.
- Se mejore la precisión de alertas automáticas de cambios de precios.

Capa de Almacenamiento de Datos

La capa de almacenamiento tiene el objetivo de preservar de manera segura.

Actividad	Descripción	Herramientas posibles
Limpieza de errores sintácticos	Corrección de caracteres corruptos (por ejemplo, "Precio: \$ 1.200.000COP" → 1200000).	Python (re, regex), Pandas.
Detección y eliminación de duplicados	Identificación de productos repetidos usando llaves únicas compuestas (nombre, modelo, marca).	Pandas, SQL DISTINCT.
Normalización de formatos	Unificación de monedas, unidades de medida y nombres de campos. (Ej.: "\$1.200K" → "1200000").	Pandas, NumPy.
Enriquecimiento de atributos	Creación de atributos derivados como "descuento absoluto" = precio anterior - precio actual.	Python Scripts, ETL pipelines.
Validación de completitud	Filtrado de registros incompletos (por ejemplo, productos sin precio ni nombre).	Python, Airflow validation tasks.

Tabla 4. . Costos relacionados al desarrollo del proyecto

Fundamento Teórico

Según Stonebraker (2015), la elección de la tecnología de almacenamiento debe basarse en:

- **Tipo de datos:** estructurados, semi-estructurados o no estructurados.
- **Frecuencia de actualizaciones:** datos en constante cambio (precios).
- **Velocidad de consulta:** necesidad de consultas rápidas para generar reportes y alertas.

Alternativas de Almacenamiento Evaluadas

Opción	Ventajas	Desventajas
Bases de datos Relacionales (MySQL, PostgreSQL)	Estructura rígida, potente en relaciones complejas.	Poca flexibilidad ante esquemas variables; menor escalabilidad horizontal.
Bases de datos NoSQL (MongoDB, Cassandra)	Flexibilidad ante esquemas cambiantes; alta escalabilidad; rápido para lecturas masivas.	Menor control transaccional complejo (no crítico para precios).

Tabla 5. Almacenamiento

Justificación de MongoDB en SmartPrice Analytics

MongoDB fue elegido como sistema principal de almacenamiento para Vitovidrios debido a:

- Naturaleza semi-estructurada de los datos de precios (no siempre todos los campos presentes).
- Flexibilidad para agregar nuevos atributos sin reestructurar toda la base (por ejemplo, nuevas promociones, bundles).
- Alta velocidad de lectura, necesaria para mostrar dashboards de precios en tiempo real.
- Facilidad de escalabilidad horizontal (sharding automático).

Revisión Bibliográfica

La literatura sobre pricing y analítica de datos ha evolucionado de enfoques descriptivos a modelos predictivos y automatizados. Nowak (2024) destaca que el *pricing dinámico* se ha convertido en una práctica habitual en el comercio electrónico por su capacidad de responder en tiempo real a cambios en la demanda y la competencia. Asimismo, Kopalle (2023) señala que el uso de algoritmos de aprendizaje automático permite ajustar los precios con una precisión mayor que los métodos tradicionales.

Acuña-Agost (2021) y Mucollari y Minotti (2022) presentan estudios donde se aplican redes neuronales y técnicas de inferencia causal para estimar la elasticidad del precio, mostrando mejoras significativas en la segmentación y la predicción del comportamiento del consumidor. Estos avances evidencian el potencial de la IA para transformar los procesos de pricing en empresas de distintos tamaños.

Kumar, Dixit y Javalgi (2020) subrayan la importancia de la recolección automatizada de datos —como el *web scraping*— para el monitoreo competitivo. Este tipo de herramientas permiten capturar información actualizada sobre precios, disponibilidad y promociones, convirtiéndose en un insumo fundamental para sistemas como *SmartPrice Analytics*.

En cuanto a la digitalización de PYMES, organismos como la **OCDE (2021)** y la **CEPAL (2021)** señalan que la analítica de datos es un factor decisivo para mejorar la eficiencia operativa y la capacidad de respuesta ante los cambios del mercado. En Colombia, el impulso hacia la transformación digital busca precisamente reducir la brecha tecnológica que limita el crecimiento de las empresas más pequeñas.

Finalmente, Davenport y Harris (2017) y Marr (2021) coinciden en que los proyectos analíticos exitosos integran tres elementos esenciales: la consolidación de un flujo de datos confiable, la implementación de modelos predictivos validados y la visualización clara de resultados para la toma de decisiones.

En ese sentido, *SmartPrice Analytics* se alinea con las tendencias contemporáneas de la analítica empresarial y del pricing automatizado. Más allá de automatizar procesos, busca que Vitovidrios S.A.S. adquiera una cultura organizacional basada en datos, capaz de generar conocimiento propio a partir del análisis sistemático del mercado.

Análisis Predictivo Aplicado a Precios

Definición de Análisis Predictivo

El análisis predictivo es una rama de la analítica avanzada que utiliza técnicas estadísticas, algoritmos de machine learning y modelos matemáticos para identificar patrones en los datos históricos y predecir comportamientos futuros (Shmueli & Koppius, 2011). En esencia, busca responder a preguntas sobre qué es probable que ocurra en el futuro, basándose en el análisis de grandes volúmenes de datos.

En el contexto de SmartPrice Analytics para Vitovidrios S.A.S., el análisis predictivo permitirá anticipar las variaciones de precios en el mercado de electrodomésticos, identificar oportunidades de ajuste de precios en tiempo real y definir estrategias dinámicas de descuento para maximizar márgenes de ganancia y competitividad. Implementar capacidades de predicción no solo optimiza los precios ofrecidos a los clientes, sino que también mejora la gestión del inventario, facilita la planeación de compras y fortalece la estrategia de promociones, garantizando que la empresa actúe de manera proactiva en lugar de reactiva frente a los cambios del mercado.

Aplicaciones del Análisis Predictivo en el Comercio Minorista

El comercio minorista, especialmente en el ámbito electrónico, es un sector donde las condiciones de mercado cambian rápidamente, y donde el precio constituye un factor crítico de diferenciación. En este sentido, el análisis predictivo ofrece diversas aplicaciones prácticas:

Optimización de Márgenes de Venta

Una de las principales aplicaciones es la optimización de márgenes de venta mediante la actualización diaria de precios de competidores. El proyecto SmartPrice Analytics contempla realizar dos barridos diarios —uno en la mañana y otro en la tarde— de los precios publicados en las principales tiendas del mercado, como Éxito, Alkosto y Falabella.

Con esta información actualizada, Vitovidrios podrá analizar los precios presentes y, de acuerdo con políticas internas previamente definidas, determinar el porcentaje de descuento óptimo que deberá aplicarse a sus productos. Esto permitirá mantener los precios más competitivos del mercado, maximizando simultáneamente tanto las ganancias como las ventas.

Ajustes Dinámicos Basados en Políticas Internas

El análisis predictivo facilitará la toma de decisiones automatizadas basadas en reglas de negocio preestablecidas. Por ejemplo, si se detecta que un competidor ha reducido el precio de un electrodoméstico clave, SmartPrice Analytics podrá recomendar un descuento específico para igualar o superar esa oferta, asegurando que Vitovidrios mantenga su posición competitiva.

Anticipación de Cambios de Precios

Además de reaccionar a los cambios de precios actuales, el sistema predictivo permitirá anticipar bajadas o subidas de precios estacionales o coyunturales (como Black Friday, Hot Sale o temporadas de fin de año), brindando así una ventaja estratégica para la planeación de inventarios, promociones y campañas publicitarias.

Modelos Utilizados en la Predicción de Precios

El éxito de un sistema predictivo depende en gran medida de la correcta selección y aplicación de modelos estadísticos y de machine learning. Para el caso de SmartPrice Analytics, se considerarán los siguientes enfoques:

Modelos de Series de Tiempo

Las series de tiempo son fundamentales para modelar y predecir precios debido a su capacidad para capturar tendencias, estacionalidades y fluctuaciones cíclicas en los datos históricos

ARIMA (AutoRegressive Integrated Moving Average):

Este modelo es ampliamente utilizado en análisis de series de tiempo. ARIMA es capaz de modelar patrones de autocorrelación en los datos, capturando tendencias lineales y estacionalidades. Para Vitovidrios, ARIMA permitirá prever la evolución de los precios de electrodomésticos en función del comportamiento pasado, ajustando automáticamente por estacionalidades como descuentos de temporada.

Prophet:

Desarrollado por Facebook, Prophet es un modelo de series de tiempo diseñado para manejar datos con fuertes tendencias estacionales y anomalías. Su principal ventaja es su facilidad de ajuste, interpretabilidad y robustez frente a datos atípicos. En el contexto de SmartPrice Analytics, Prophet será útil para modelar las variaciones de precios afectadas por festividades o promociones masivas.

Modelos Supervisados

Además de las series de tiempo, se aplicarán modelos supervisados que consideran múltiples variables independientes (features) para predecir el precio futuro de un producto.

- **Regresión Lineal:** Permite establecer relaciones lineales entre variables como precios de competidores, inventario, promociones y demanda, con el objetivo de predecir el precio óptimo.

- **Random Forest:** Este modelo de ensamblado basado en árboles de decisión puede capturar relaciones no lineales complejas y manejar grandes cantidades de variables sin riesgo de sobreajuste. Para Vitovidrios, Random Forest facilitará predicciones robustas y estables incluso en escenarios de alta variabilidad de mercado.
- **XGBoost (Extreme Gradient Boosting):** Reconocido por su alta precisión y velocidad, XGBoost combina múltiples árboles de decisión para maximizar el rendimiento predictivo. Será útil especialmente en contextos donde pequeñas variaciones en las variables explicativas (por ejemplo, una nueva promoción de un competidor) pueden tener un gran impacto en la estrategia de precios.

Variables Comunes en Modelos de Predicción de Precios

La calidad y relevancia de las variables incluidas en los modelos predictivos son factores determinantes para el éxito del análisis. En SmartPrice Analytics se identificarán y recopilarán de manera sistemática las siguientes variables:

Demanda Histórica

La cantidad de unidades vendidas en períodos anteriores proporciona una base para entender la elasticidad del precio y la sensibilidad de los clientes a cambios en los precios. La demanda histórica permite anticipar momentos de alta o baja venta y ajustar precios estratégicamente.

Precios de Competidores

El precio que ofrecen los competidores es un factor fundamental en mercados de alta competencia. La variabilidad de estos precios permitirá ajustar rápidamente las estrategias de Vitovidrios, asegurando que se ofrezcan productos a precios competitivos en todo momento.

Promociones Aplicadas

Las promociones históricas —descuentos, cupones, ventas flash— afectan tanto el volumen de ventas como las expectativas de los clientes. Considerarlas en los modelos permitirá evitar errores en la interpretación de aumentos o caídas súbitas en la demanda.

Estacionalidad (Festividades, Ciclos Económicos)

Eventos como el Día de la Madre, Navidad o el Black Friday generan picos de demanda que deben ser incorporados explícitamente en los modelos. Asimismo, la identificación de ciclos económicos generales (recesión, expansión) permitirá ajustar las expectativas de precios y ventas.

Nivel de Inventario

El inventario disponible tiene una influencia directa en las decisiones de pricing. Por ejemplo, un inventario elevado puede justificar descuentos agresivos para evitar sobrecostos de almacenamiento, mientras que un inventario limitado puede permitir precios premium. Modelar esta variable es clave para optimizar tanto la rentabilidad como la rotación de stock.

Diseño Metodológico

El desarrollo del proyecto se organiza en seis etapas principales:

(i) recolección de datos de precios de productos de la competencia, (ii) limpieza y preparación de los datos, (iii) análisis exploratorio de datos (EDA), (iv) construcción de modelos predictivos, (v) evaluación del desempeño de los modelos y (vi) análisis e interpretación de resultados. Este enfoque garantiza coherencia entre los objetivos planteados y las técnicas analíticas empleadas.

Diagnóstico Organizacional

En paralelo al análisis cuantitativo de precios de mercado, se desarrolló un diagnóstico organizacional de carácter secundario, basado en una encuesta interna aplicada a colaboradores de Vitovidrios S.A.S. Este diagnóstico tuvo como finalidad contextualizar el nivel de madurez analítica de la organización y comprender las prácticas actuales de fijación de precios. Es importante resaltar que este ejercicio cumple una función exclusivamente descriptiva y de apoyo, y no forma parte del proceso de modelamiento predictivo ni del análisis estadístico de precios.

Análisis externo - PESTEL

Político

El contexto político en Colombia se mantiene relativamente estable, facilitando el desarrollo del comercio. Sin embargo, las políticas públicas como los Días sin IVA generan picos de demanda concentrados que requieren una rápida adaptación en la fijación de precios. Para empresas como Vitovidrios, que aún dependen de procesos manuales, esto representa un desafío para capitalizar estas oportunidades, ya que sin herramientas tecnológicas automatizadas es difícil responder con rapidez y precisión..

Además, las regulaciones fiscales, los acuerdos comerciales y las normas aduaneras impactan directamente en los costos y disponibilidad de los electrodomésticos, obligando a ajustes constantes en la estrategia comercial. Sin un sistema analítico sólido que interprete estas variables políticas, Vitovidrios corre el riesgo de perder competitividad y rentabilidad frente a competidores más ágiles (Eslava, 2015).

Económico

La inflación y la devaluación del peso colombiano generan fluctuaciones en los costos de importación, presionando a las empresas a realizar ajustes frecuentes en sus precios para mantener la rentabilidad. Vitovidrios, sin una plataforma predictiva y de monitoreo en tiempo real, enfrenta riesgos significativos de fijar precios desalineados con la realidad del mercado, afectando sus márgenes (Eslava, 2015).

Por otro lado, el consumidor colombiano es altamente sensible al precio, influenciado por promociones y descuentos disponibles en el mercado. Esto genera volatilidad en las ventas, donde una respuesta dinámica y basada en datos es crucial para no perder cuota de mercado frente a competidores que ya aplican pricing inteligente (Rivero Gutiérrez, Samino García & Cerro Rodríguez, 2023).

Social

La digitalización ha cambiado el perfil del consumidor, quien ahora compara precios y ofertas en múltiples plataformas, demandando transparencia, rapidez y conveniencia. Vitovidrios debe adaptarse a estas expectativas mediante análisis de datos actualizados para ofrecer promociones relevantes, dejando atrás las decisiones empíricas que limitan su capacidad de respuesta (Rivero Gutiérrez et al., 2023).

Además, la creciente demanda por experiencias de compra integradas y personalizadas implica optimizar inventarios y precios para evitar pérdidas por sobrestock o faltantes. La falta de análisis avanzado sobre comportamiento del consumidor afecta la capacidad de la empresa para anticipar tendencias y mejorar la satisfacción del cliente.

Tecnológico

Las tecnologías como web scraping, machine learning y big data han revolucionado la fijación de precios en el comercio minorista. Grandes competidores en Colombia utilizan estas herramientas para ajustar precios en tiempo real y optimizar márgenes, dejando en desventaja a empresas sin infraestructura tecnológica adecuada (Mitchell, 2018; Gupta, 2018).

Para Vitovidrios, que es una PYME, la adopción de estas tecnologías es posible y necesaria, ya que ofrecen escalabilidad y personalización. Sin embargo, la falta de inversión y conocimiento retrasa la transformación digital, lo que hace imprescindible la implementación de soluciones como SmartPrice Analytics (Fúquene Linares et al., 2025).

Ecológico

Aunque el factor ecológico no impacta directamente la fijación de precios, la gestión eficiente del inventario mediante análisis de datos contribuye a minimizar desperdicios y reducir el impacto ambiental asociado a productos obsoletos o excedentes, alineando a la empresa con las tendencias de sostenibilidad actuales (Brynjolfsson & McAfee, 2014). Además, un compromiso con prácticas responsables puede fortalecer la reputación corporativa y atraer a consumidores conscientes del impacto ambiental. SmartPrice Analytics puede apoyar esta gestión, facilitando decisiones que favorezcan la sostenibilidad operativa y financiera.

Legal

La protección de datos personales en Colombia, regulada por la Ley Habeas Data, exige que las empresas manejen la información con responsabilidad, especialmente cuando emplean técnicas como web scraping para recolectar datos del mercado. Es esencial que Vitovidrios asegure el cumplimiento normativo para evitar sanciones y preservar su reputación (Gillespie, 2018; Kamarinou, Millard & Singh, 2016).

Asimismo, el respeto a los términos de servicio de los sitios web y el cumplimiento de regulaciones fiscales y comerciales son fundamentales para una operación sostenible. La solución tecnológica debe integrar mecanismos que garanticen la legalidad y la ética en el manejo de datos y la fijación de precios.

Análisis interno

Vitovidrios S.A.S. se identifican diversas causas estructurales que explican la problemática abordada en este proyecto: la falta de una estrategia de precios basada en datos. Actualmente, la empresa no dispone de herramientas tecnológicas que permitan automatizar el monitoreo y análisis de precios, lo cual limita su capacidad de reacción frente a la competencia y a las fluctuaciones constantes del mercado. Esta ausencia de tecnología está acompañada por una brecha importante en capacidades analíticas dentro del equipo comercial, donde no se cuenta con procesos ni conocimientos sólidos en inteligencia de datos para evaluar la elasticidad del precio, proyectar la demanda, o identificar patrones de consumo. La toma de decisiones en torno al pricing sigue siendo principalmente empírica y reactiva, basada en la experiencia personal o en ajustes tardíos luego de detectar caídas en ventas. A su vez, la falta de integración entre las áreas de ventas, inventario y compras impide construir una visión consolidada del comportamiento del mercado, lo que genera acumulación de productos poco demandados y escasez de los más competitivos. Esta desconexión interna dificulta la definición de precios estratégicos que respondan no solo a los costos y márgenes deseados, sino también a las condiciones del entorno comercial. Estas causas, identificadas como cuellos de botella críticos, fundamentan la pertinencia de diseñar e implementar una solución como SmartPrice Analytics, una plataforma de monitoreo y análisis de precios que transforme los datos del entorno competitivo en decisiones precisas, oportunas y orientadas a maximizar la rentabilidad de la empresa.



Figura 3. Retos manejo de datos interno Vitovidrios

Análisis interno basado en encuesta a colaboradores

Es importante precisar que el objetivo central del presente trabajo de grado no corresponde al análisis de percepciones internas del personal, sino al diseño e implementación de una plataforma de monitoreo y análisis de precios de productos de la competencia, basada en datos de mercado y técnicas de ciencia de datos. En este sentido, la población principal del estudio está conformada por los productos de electrodomésticos y tecnología (celulares, lavadoras y televisores) y por sus precios históricos y actuales observados en los principales competidores del sector.

Por lo tanto, el análisis de precios constituye el eje metodológico del proyecto y es sobre esta población (productos y precios) donde se aplican las técnicas de web scraping, análisis exploratorio de datos, modelamiento predictivo y evaluación de desempeño descritas en los capítulos metodológicos y analíticos.

No obstante, como insumo complementario y de carácter secundario, se desarrolló un diagnóstico organizacional preliminar mediante una encuesta aplicada a colaboradores de Vitovidrios S.A.S. Este ejercicio tuvo como finalidad identificar el nivel de madurez analítica de la organización, las prácticas actuales de fijación de precios y la percepción interna sobre las necesidades de información para la toma de decisiones comerciales. Dicho diagnóstico no constituye la metodología del análisis de datos principal ni influye directamente en el desarrollo de los modelos predictivos de precios.

Dado el tamaño reducido de la empresa, se aplicó un muestreo estratificado proporcional sobre una población total de 21 empleados, distribuidos entre gerencia, directores de área y empleados de planta, obteniendo una muestra de 17 colaboradores con un nivel de confianza del 95% y un margen de error del 10%. Los resultados derivados de este instrumento se utilizan exclusivamente para sustentar el diagnóstico organizacional y justificar la pertinencia de la solución tecnológica propuesta.

En consecuencia, se establece de manera explícita la distinción entre:

Estudio principal: análisis cuantitativo de precios de productos de la competencia, donde la población corresponde a los productos y sus precios observados en el mercado, y la metodología se basa en ciencia de datos, web scraping y modelos de machine learning.

Diagnóstico organizacional secundario: análisis cualitativo–descriptivo apoyado en una encuesta interna, cuyo propósito es contextualizar la problemática empresarial y respaldar la necesidad de implementar la plataforma SmartPrice Analytics, sin intervenir en el análisis estadístico ni predictivo del pricing.

Esta aclaración metodológica permite mantener la coherencia entre los objetivos del proyecto, la definición de la población de estudio y las técnicas analíticas empleadas, fortaleciendo el rigor académico del trabajo.

Análisis Exploratorio Precios Competidores

Resumen de Datos

El presente proyecto se enfocó en el análisis comparativo de precios, descuentos y especificaciones técnicas de productos de línea blanca y tecnología (celulares, lavadoras y televisores) ofrecidos por las franquicias Falabella, Éxito y Alkosto. La metodología de recolección de datos se basó en la implementación de una solución de web scraping periódico, ejecutada cada tres días para la descarga integral del contenido HTML de las plataformas e-commerce. La información obtenida fue procesada mediante la tecnología Gemini para su estructuración en formatos tabulares, y posteriormente centralizada y gestionada en Google Cloud Platform (GCP), utilizando BigQuery para el almacenamiento histórico. Estos datos primarios se enriquecieron con bases de datos externas y recolección manual para asegurar la exhaustividad del análisis.

Fases clave del proceso

La Figura 4 presenta las fases clave del proceso implementado en la plataforma de monitoreo y análisis de precios de electrodomésticos. Este flujo metodológico describe de forma estructurada el recorrido de los datos desde su recolección hasta su explotación analítica, garantizando trazabilidad, escalabilidad y reproducibilidad del sistema.

En la primera fase, **Ingesta**, se realiza la recolección automatizada de información de precios desde plataformas de comercio electrónico mediante técnicas de *web scraping*. Esta etapa permite capturar datos en tiempo casi real, asegurando una actualización constante de la información del mercado.

La segunda fase, **Procesamiento**, corresponde a las tareas de limpieza, estandarización y transformación de los datos recolectados. En esta etapa se corrigen inconsistencias, se normalizan formatos y se estructuran las variables necesarias para el análisis posterior, reduciendo el ruido y mejorando la calidad del conjunto de datos.

Posteriormente, en la fase de **Almacenamiento**, los datos procesados son ingeridos en la infraestructura de Google Cloud Platform (GCP), específicamente en **BigQuery**, lo que

permite un almacenamiento escalable, consultas eficientes y la integración con herramientas analíticas avanzadas.

Finalmente, la fase de **Análisis (BI)** se enfoca en la conexión con herramientas de *Business Intelligence* para la exploración, visualización y análisis de los precios monitoreados. Esta etapa facilita la generación de reportes, dashboards y métricas clave que apoyan la toma de decisiones estratégicas relacionadas con la gestión de precios.

En conjunto, estas fases conforman un pipeline de datos integral que soporta el objetivo del proyecto: proporcionar información estructurada y confiable para el análisis del comportamiento de precios en el mercado de electrodomésticos.



Figura 4. Fases del proceso

Análisis de datos

Evolución temporal de precios

Con el fin de analizar el comportamiento de los precios a lo largo del año de las categorías que son objetivo de esta investigación, se analizaron los datos obtenidos de 85,020 registros históricos correspondientes a tres categorías de productos electrónicos: celulares (59,956 registros), lavadoras (11,596 registros) y televisores (13,468 registros),

extraídos de las plataformas de comercio electrónico Alkosto, Éxito y Falabella durante el periodo de un año. Mostrando la variación presentada en la Grafica.

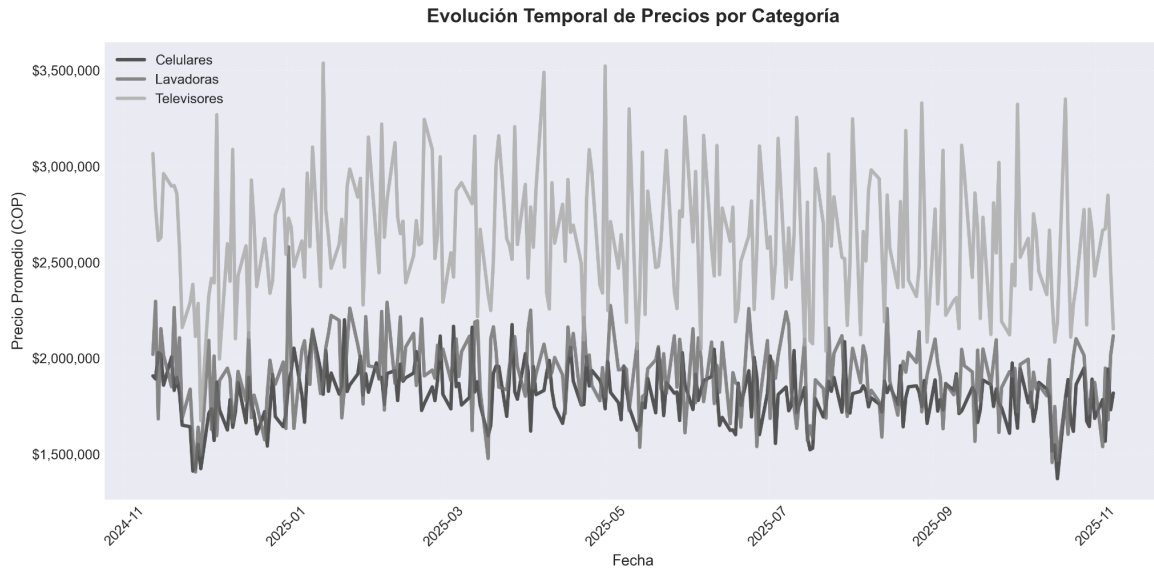


Figura 5. Evolución temporal de precios por categoría

El análisis temporal revela que los televisores presentan el precio promedio más elevado (\$2,629,507 COP) con una alta volatilidad (CV: 95.46%), mientras que celulares y lavadoras mantienen precios similares alrededor de \$1.8-1.9 millones con comportamientos más estables. Las tres categorías exhiben patrones estacionales consistentes, con reducciones de precio del 3-5% durante los meses de octubre a diciembre, coincidiendo con las temporadas comerciales de descuentos (Black Friday, Cyber Monday y Navidad).

Distribución de precios

Con el objetivo de caracterizar la variabilidad, identificar patrones de comportamiento y analizar la presencia de valores atípicos en los precios de las categorías estudiadas, se examinó la distribución estadística de los 85.020 registros mediante histogramas de frecuencia. La Figura 6 presenta la distribución de precios para cada categoría de producto.

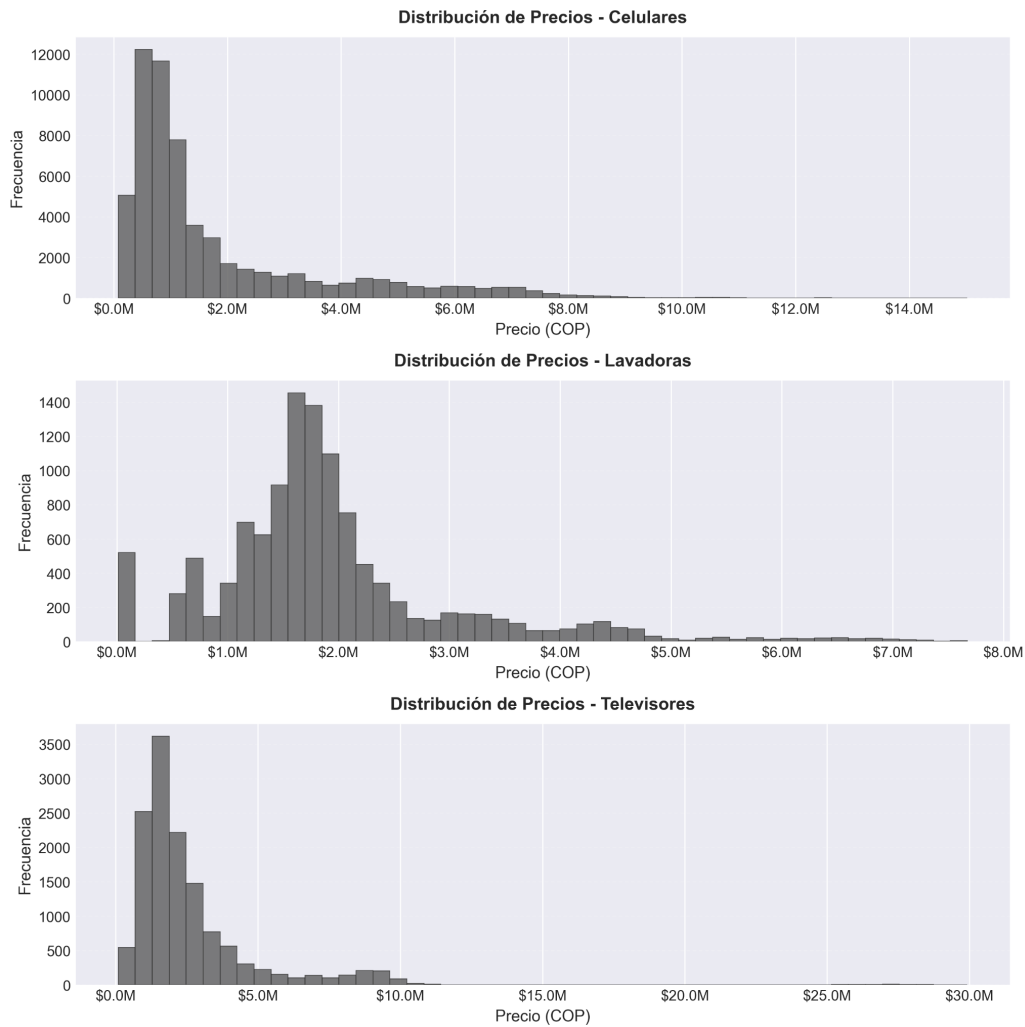


Figura 6. Distribución de precios

El análisis de la distribución de precios para las categorías de celulares, lavadoras y televisores evidencia comportamientos estadísticos diferenciados, con distribuciones asimétricas y presencia de colas largas hacia valores elevados. En la categoría de celulares se observa una asimetría positiva (skewness: 1.898), con una alta concentración de productos en el rango de precios entre \$0.5 y \$2 millones, junto con una cola extendida hacia segmentos de gama alta, lo que se refleja en la identificación de 6.981 valores atípicos, equivalentes al 11.64% del total de registros. Estos valores corresponden principalmente a dispositivos de alta gama, caracterizados por especificaciones técnicas superiores y posicionamiento premium en el mercado.

En el caso de las lavadoras, la distribución presenta una asimetría moderada (skewness: 1.619), con la mayor concentración de precios entre \$1 y \$2.5 millones, y un total de 1.570 valores atípicos (13.54%). A diferencia de los celulares, estos valores atípicos se asocian tanto a equipos de alta capacidad y eficiencia energética como a modelos industriales o de mayor volumen, los cuales elevan el precio promedio de la categoría.

Para los televisores, se observa la mayor asimetría (skewness: 4.463) y una curtosis elevada (33.792), indicando una distribución altamente concentrada en rangos de precio bajos y medios, con una presencia significativa de productos de gama premium. En esta categoría se identificaron 1.287 valores atípicos (9.56%), correspondientes principalmente a televisores de gran formato, tecnologías avanzadas de visualización y marcas posicionadas en el segmento de alto valor.

Es importante destacar que los valores atípicos identificados no corresponden a errores de captura ni a inconsistencias en el proceso de extracción de datos, sino a productos reales que hacen parte del portafolio comercial del mercado colombiano. En consecuencia, estos registros no fueron eliminados durante el preprocesamiento, dado que su exclusión podría generar una pérdida de información relevante sobre segmentos de alto valor y afectar la representatividad del análisis. La conservación de estos valores permite capturar de manera más fiel la heterogeneidad del mercado y justifica el uso posterior de modelos predictivos no lineales y técnicas robustas, capaces de manejar distribuciones asimétricas y alta variabilidad de precios.

Análisis de Marcas

Con el fin de identificar la concentración del mercado y las marcas dominantes en cada categoría de productos electrónicos, se analizó la frecuencia de productos por marca en el dataset de 85,020 registros. La Figura presenta el top 10 de marcas con mayor presencia en cada categoría estudiada.

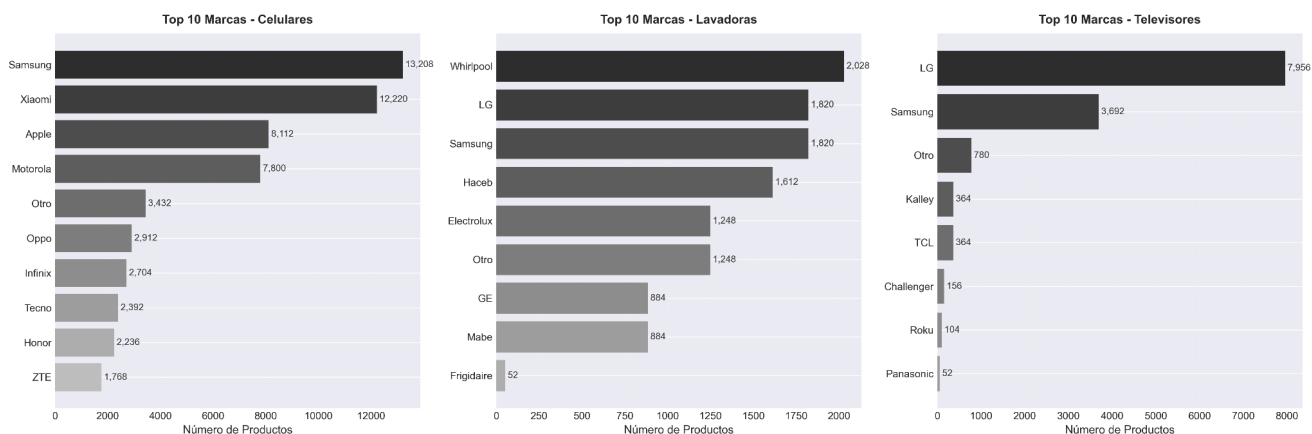


Figura 7. Referencias más vendidas por artículo de venta

En la categoría de celulares, Samsung lidera con 13,208 productos (22.0%), seguido por Xiaomi con 12,220 (20.4%) y Apple con 8,112 productos (13.5%), evidenciando una alta fragmentación del mercado con participación de 15 marcas diferentes. Para lavadoras, el mercado presenta una distribución más equilibrada con Whirlpool (2,028 productos, 17.5%), LG y Samsung empatados (1,820 productos cada uno, 15.7%), y Hacerb (1,612 productos, 13.9%), mostrando un total de 9 marcas competidoras. En el caso de televisores, se observa una alta concentración con LG dominando el 59.1% del mercado (7,956 productos), seguido distantemente por Samsung con 27.4% (3,692 productos), indicando un mercado oligopólico con presencia de solo 8 marcas. Esta distribución sugiere diferentes niveles de competencia según la categoría, siendo celulares el mercado más fragmentado y televisores el más concentrado.

Autocorrelaciones

En el contexto del desarrollo de la plataforma SmartPrice Analytics, el análisis de series temporales constituye un componente fundamental para comprender la dinámica de formación de precios en el mercado de electrodomésticos. Mientras que las secciones anteriores de esta investigación se enfocaron en el análisis exploratorio de la base externa de retailers, permitiendo caracterizar el comportamiento competitivo del mercado a través de variables como precio, descuentos y especificaciones técnicas, el análisis de autocorrelación introduce una dimensión temporal que complementa dicho enfoque.

En particular, el análisis de autocorrelación permite evaluar si los precios observados en los retailers presentan dependencias temporales sistemáticas, lo cual resulta esencial para entender si la evolución de precios sigue patrones estructurados o si responde principalmente a fluctuaciones aleatorias del mercado. De esta forma, el estudio de autocorrelación actúa como un puente metodológico entre el análisis exploratorio descriptivo y el diseño de herramientas analíticas más avanzadas dentro de la plataforma SmartPrice Analytics.

Desde el punto de vista estadístico, la autocorrelación puede definirse como la correlación lineal entre los valores de una serie temporal y sus valores rezagados en el tiempo, medida generalmente mediante la función de autocorrelación (ACF). Esta herramienta permite identificar la presencia de dependencia serial, ciclos recurrentes y persistencia temporal, elementos fundamentales en el análisis de mercados dinámicos como el de electrodomésticos.

En el caso de la base externa analizada —correspondiente a los principales retailers del mercado colombiano como Falabella, Éxito y Alkosto— el análisis de autocorrelación permite evaluar si los cambios de precios observados siguen patrones coordinados en el tiempo o si reflejan estrategias independientes entre competidores.

Este enfoque resulta especialmente relevante para el diseño de SmartPrice Analytics, ya que un sistema de monitoreo de precios no solo debe capturar niveles actuales de precios, sino también comprender la dinámica temporal del mercado competitivo.

Celulares – Dinámica Altamente Reactiva

En la categoría de celulares, el análisis de autocorrelación sugiere una baja persistencia temporal en los precios observados en los retailers. Este comportamiento es consistente con mercados tecnológicos caracterizados por ciclos de innovación rápidos, donde los precios se ajustan frecuentemente en respuesta a lanzamientos de nuevos modelos y estrategias promocionales agresivas.

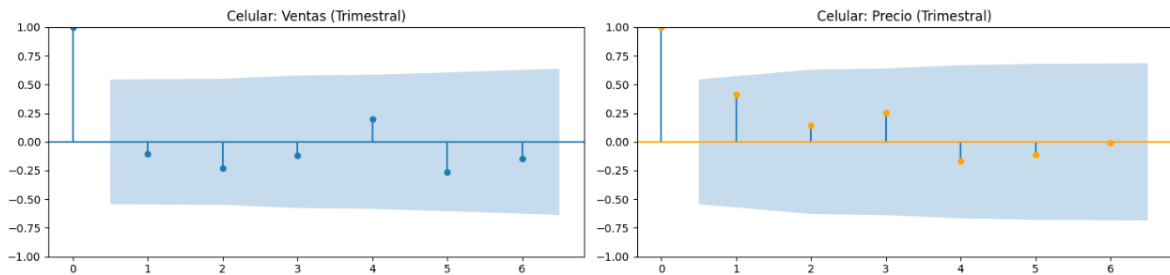


Figura 8. Autocorrelación Ventas y Precio trimestral - Celulares

Desde el punto de vista estadístico, este patrón suele manifestarse como una rápida disminución de la función de autocorrelación a medida que aumentan los lags, indicando una dependencia limitada entre periodos consecutivos. Para la plataforma SmartPrice Analytics, este resultado implica que la información reciente tiene mayor relevancia que los patrones históricos de largo plazo.

Televisores – Persistencia y Estacionalidad

En la categoría de televisores, el análisis de autocorrelación tiende a mostrar patrones más persistentes que en celulares, lo cual es consistente con la naturaleza de bienes durables. La presencia de autocorrelaciones positivas en varios lags sugiere que los precios siguen trayectorias relativamente estables en el tiempo, interrumpidas principalmente por periodos promocionales.

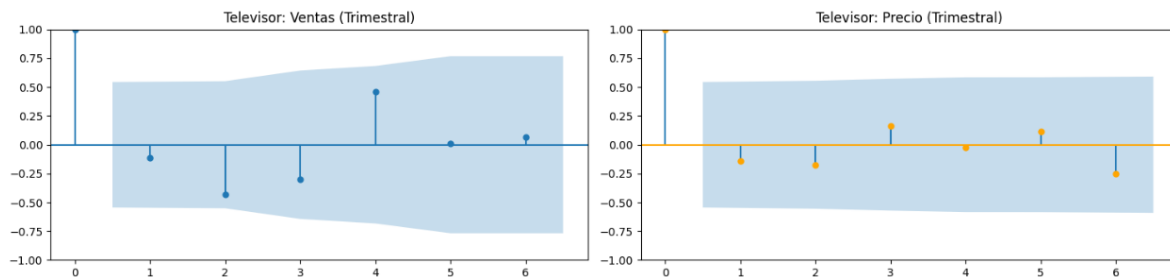


Figura 9. Autocorrelación Ventas y Precio trimestral - Televisores

Este comportamiento indica que los retailers tienden a mantener estrategias de precios relativamente consistentes, ajustándose únicamente en momentos específicos del ciclo comercial.

Lavadoras – Alta Estabilidad Estructural

La categoría de lavadoras presenta el mayor nivel esperado de persistencia temporal en los precios observados. Este resultado es coherente con la naturaleza del mercado de línea blanca, donde los cambios de precios suelen ser menos frecuentes y más estructurados.

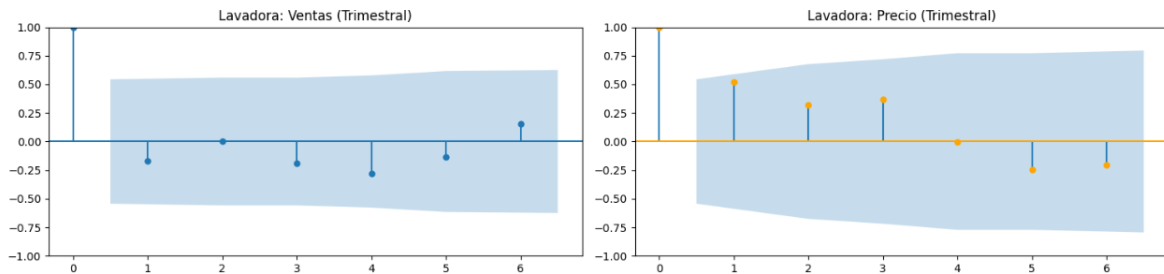


Figura 10. Autocorrelación Ventas y Precio trimestral - Lavadoras

Desde el punto de vista de la función de autocorrelación, este comportamiento se refleja típicamente en valores elevados que decrecen lentamente a medida que aumentan los lags, indicando una fuerte dependencia temporal. Para SmartPrice Analytics, este patrón implica que los precios históricos constituyen una referencia importante para el análisis competitivo.

Análisis Exploratorio Base Interna

Evolución Anual - Ingresos y Unidades

Con el objetivo de identificar patrones de crecimiento y estacionalidad en el negocio de la PYME distribuidora, se analizó la evolución diaria de ingresos a lo largo de un período de 1,090 días (~3 años), correspondiente a 1,961 transacciones comerciales de productos electrónicos (celulares, lavadoras y televisores). La Figura 10 presenta la serie temporal de ingresos diarios junto con su media móvil de 30 días para suavizar la volatilidad y revelar tendencias subyacentes.

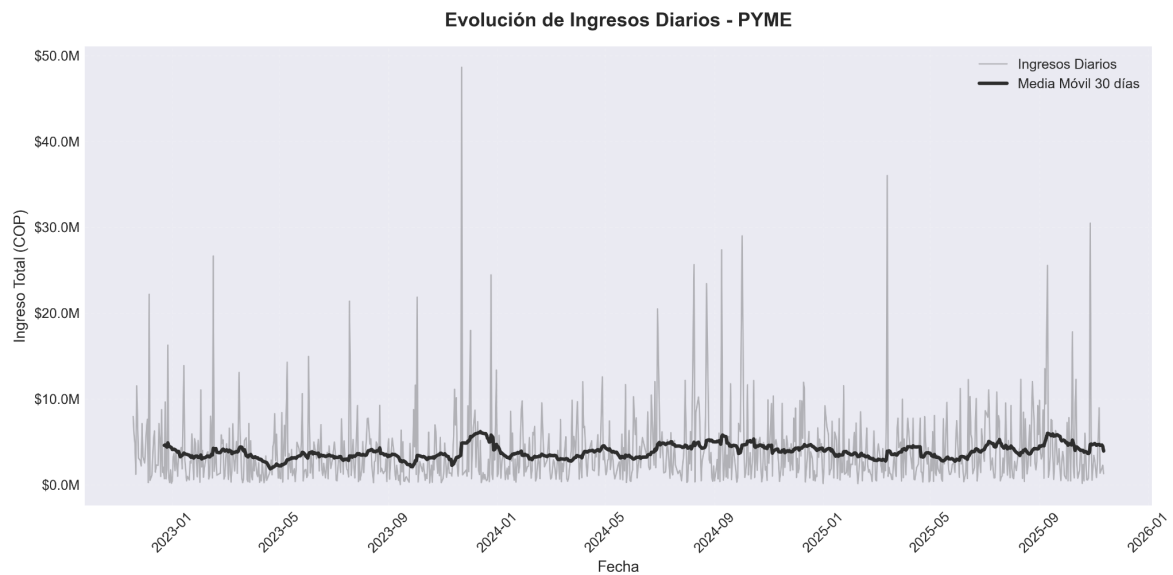


Figura 11. Serie temporal de ingresos diarios - PYME

El análisis temporal revela un comportamiento con alta variabilidad diaria en los ingresos, con valores que oscilan significativamente entre transacciones individuales de bajo valor (\$0.5M) y ventas de alto ticket (\$3-4M), lo cual es típico en empresas PYME con volumen de transacciones limitado donde cada venta tiene impacto considerable en los ingresos diarios. La media móvil de 30 días, representada por la línea oscura, muestra una tendencia general de ingresos que se mantiene relativamente estable alrededor de \$3-4 millones mensuales durante el período analizado, sin evidencia de crecimiento o decrecimiento pronunciado. Se observan períodos de mayor actividad comercial intercalados con días de baja o nula actividad, reflejando la naturaleza intermitente de las ventas en el sector de distribución de electrónicos a menor escala. Esta variabilidad inherente y la ausencia de tendencia clara sugieren que los modelos de predicción de demanda deben incorporar características del producto y contexto de mercado además de variables temporales para lograr precisión adecuada.

Análisis de marcas

Con el objetivo de identificar las marcas estratégicas para la PYME y caracterizar el portafolio por segmento de precio, se analizó el desempeño de las 20 marcas distribuidas durante el período de 3 años, evaluando tanto volumen de ventas como posicionamiento

de precio. La Figura 11 presenta dos dimensiones complementarias: las marcas líderes en unidades vendidas (volumen) y las marcas premium por ticket promedio (valor).

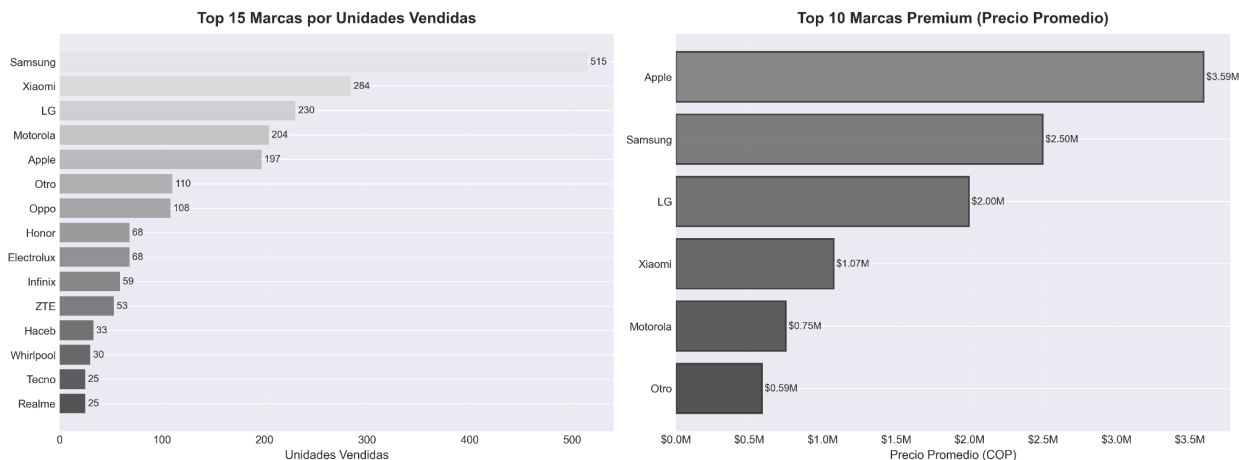


Figura 12. Marcas más vendidas por Vitovidrios S.A.S.

El análisis de volumen revela que Samsung domina ampliamente el mercado con 427 unidades vendidas, seguida distantemente por LG (233 unidades) y Xiaomi (216 unidades), evidenciando una fuerte concentración en estas tres marcas que representan aproximadamente el 42% del volumen total (2,091 unidades). Esta distribución indica que la PYME tiene una relación comercial consolidada con fabricantes de alta rotación. Marcas como Whirlpool (139 unidades), Motorola (133 unidades) y Apple (114 unidades) completan el top 6, mostrando presencia significativa en sus respectivos segmentos. En contraste, el análisis de posicionamiento premium muestra que Apple lidera el ranking de precio promedio con \$4.11M por unidad, seguida por Huawei (\$3.67M) y Samsung (\$2.68M), revelando una estrategia de portafolio balanceada que combina volumen (Samsung, Xiaomi) con margen alto (Apple, Huawei). Las marcas de electrodomésticos como Whirlpool (\$1.85M) y Haceb (\$1.31M) ocupan posiciones intermedias, consistente con el segmento de lavadoras. Esta dualidad volumen-valor sugiere que el modelo de pricing debe considerar tanto la estrategia de rotación de inventario como el posicionamiento de marca, siendo crítico diferenciar políticas de descuento entre marcas premium (menor flexibilidad de precio) y marcas de volumen (mayor agresividad comercial).

Estructura de Precios

Con el propósito de caracterizar la estructura de precios y detectar la presencia de múltiples segmentos de mercado en el portafolio de la PYME, se analizó la distribución de precios finales de las 1,961 transacciones realizadas durante el período de 3 años. La Figura 12 presenta el histograma de frecuencias de precios finales junto con las medidas de tendencia central: media (\$1.69M) y mediana (\$1.09M).

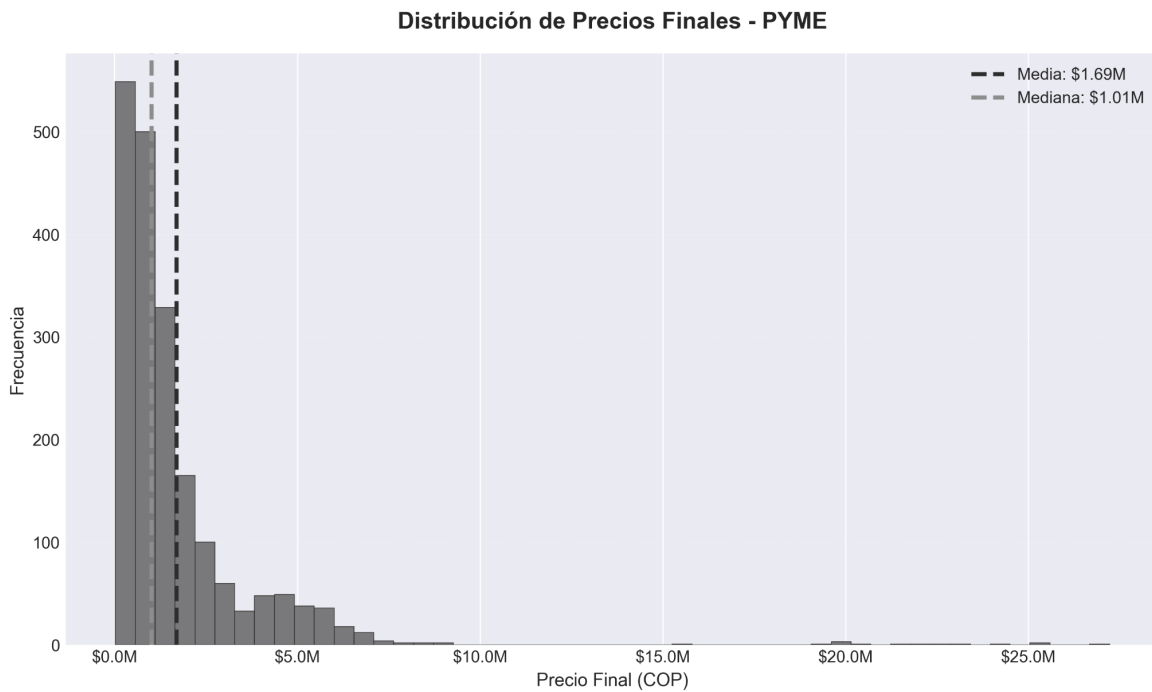


Figura 13. Precio final Vitovidrios S.A.S.

El análisis revela una distribución altamente asimétrica positiva (skewness: 5.587) con curtosis extrema (46.035), indicando una fuerte concentración de transacciones en el rango de precios bajos (\$0.5M-\$2M) con una cola larga hacia precios premium que se extiende hasta \$27.2M. La diferencia significativa entre media (\$1.69M) y mediana (\$1.09M) confirma el sesgo hacia valores bajos, sugiriendo que aproximadamente la mitad de las transacciones corresponden a productos de gama media-baja (celulares básicos, lavadoras estándar), mientras que ventas ocasionales de productos premium (televisores de alta gama, celulares flagship) elevan considerablemente el promedio. La presencia de 208 outliers (10.61% del total) representa productos de lujo o equipos especializados que, aunque infrecuentes, generan ingresos significativos por su alto ticket. Esta distribución multimodal y asimétrica implica que modelos de pricing lineal

tradicionales pueden presentar limitaciones, siendo recomendable el uso de técnicas robustas o modelos segmentados por rango de precio que capturen la heterogeneidad del portafolio y permitan estrategias de descuento diferenciadas según el segmento objetivo.

Modelo

Se desarrolló un Sistema de Pricing Inteligente basado en Machine Learning para Vitrovidrios S.A.S., distribuidora de productos electrónicos en Colombia. El sistema tiene como objetivo optimizar el precio de venta y la aplicación de descuentos en tiempo casi real, maximizando simultáneamente el margen y la competitividad.

Durante la etapa de modelamiento se analizaron distintas alternativas metodológicas, incluyendo modelos lineales y regularizados como regresión lineal, Ridge y Lasso, con el propósito de establecer un punto de referencia (baseline) para el desempeño del sistema. Sin embargo, el problema de pricing presenta características que limitan la efectividad de estos enfoques, entre ellas: relaciones no lineales entre precio y variables explicativas, alta heterogeneidad entre categorías de producto, presencia de valores atípicos asociados a productos de alto valor y fuertes interacciones entre factores temporales, competitivos y comerciales.

Bajo este contexto, los modelos basados en árboles de decisión resultaron más adecuados, al no imponer supuestos de linealidad y ofrecer mayor robustez frente a distribuciones asimétricas y outliers. En particular, se seleccionó XGBoost para la predicción del precio óptimo debido a su capacidad para capturar interacciones complejas entre variables, su eficiencia computacional y su desempeño probado en problemas de regresión con datos estructurados. Adicionalmente, se eligió Random Forest para la toma de decisiones relacionadas con la estrategia de descuento, dado su comportamiento estable, menor sensibilidad al sobreajuste y facilidad de interpretación relativa mediante medidas de importancia de variables.

Ambos modelos fueron entrenados utilizando datos históricos de la PYME y precios de competidores clave como Éxito, Alkosto y Falabella, permitiendo integrar información interna y del mercado en un sistema analítico coherente con las dinámicas reales del retail colombiano.

Entrada y Preparación de Datos

Esta fase constituye la base de la inteligencia del modelo. La etapa de entrada y preparación de datos integra el historial transaccional de la base interna de la empresa con precios actualizados de la competencia obtenidos mediante procesos de web scraping. A partir de esta información se lleva a cabo un proceso estructurado de feature engineering, cuyo objetivo es transformar los datos brutos en variables explicativas relevantes para el modelamiento predictivo.

En primer lugar, se construyeron variables temporales como mes, trimestre y semana, las cuales permiten capturar patrones cíclicos en el comportamiento de precios. Estas variables se complementaron con indicadores estacionales binarios, tales como Black_Friday y Temporada_Alta, diseñados a partir del conocimiento del calendario comercial del retail colombiano, con el fin de representar explícitamente el impacto de campañas promocionales recurrentes.

Adicionalmente, se incorporaron variables competitivas, entre las cuales destaca el precio promedio de la competencia (Precio_Competencia_Mean), calculado a partir de los precios recolectados en las principales plataformas del mercado. Esta variable resume la presión competitiva del entorno y se convirtió en uno de los predictores más relevantes del precio óptimo.

Las variables categóricas del producto, como marca, categoría y especificaciones técnicas, fueron transformadas mediante One-Hot Encoding, permitiendo su inclusión en los modelos sin introducir relaciones ordinales artificiales. Este tratamiento asegura que cada categoría sea representada de manera independiente, preservando la interpretabilidad del modelo y evitando sesgos derivados de codificaciones numéricas directas.

Finalmente, todas las variables fueron consolidadas en un conjunto de datos estructurado, listo para su uso en los algoritmos de machine learning, garantizando coherencia entre la etapa de preparación de datos y el proceso de entrenamiento del modelo.

Entrenamiento y Modelos Predictivos

Con base en los hallazgos obtenidos durante el análisis exploratorio, se procedió al desarrollo de modelos predictivos orientados a estimar el comportamiento futuro de los precios y apoyar la toma de decisiones estratégicas de la empresa. La selección de los modelos se realizó considerando la naturaleza de los datos, la interpretabilidad de los resultados y su aplicabilidad en un entorno empresarial real.

Esta fase es la base de la inteligencia del modelo. La Entrada de Datos combina el historial transaccional de la base interna con precios actualizados de la competencia obtenidos mediante web scraping. aquí se realiza la ingeniería de features donde la información temporal (Mes, Trimestre) se codifica con el conocimiento estacional (Es_Black_Friday, Es_Temporada_Alta). Además, se calcula el diferencial de precios competitivo (Precio_Competencia_Mean), la variable que posteriormente demostró ser la más predictiva del precio.

El corazón del sistema opera mediante un conjunto de modelos especializados (ensemble of models) para abordar la complejidad de la decisión de pricing:
Modelo de Precio Óptimo (XGBoost Regressor): Algoritmo de gradient boosting que predice el precio base ideal para cada producto, maximizando la utilidad esperada. Su alta dependencia de la feature Precio_Competencia_Mean confirma que el mercado dicta el precio base.

Clasificador de Descuentos (Random Forest): Modelo de clasificación que decide si es necesario o no aplicar un descuento (Decisión binaria), basándose principalmente en el factor Es_Black_Friday (28% de importancia) y el diferencial competitivo.

Modelo de Magnitud de Descuento (Random Forest): Si se requiere descuento, este regresor predice el porcentaje óptimo a aplicar.

Con el fin de garantizar la capacidad de generalización de los modelos predictivos y reducir el riesgo de sobreajuste, se definió un protocolo de entrenamiento y validación basado en la separación explícita de los datos en conjuntos de entrenamiento y prueba (train/test split). Para cada uno de los modelos desarrollados, aproximadamente el 80% de los registros se utilizaron para el entrenamiento del modelo, mientras que el 20%

restante se reservó exclusivamente para la evaluación de desempeño sobre datos no vistos.

En el caso del modelo clasificador de descuentos, se aplicó una partición estratificada de los datos, asegurando que la distribución de clases (aplicar descuento / no aplicar descuento) se mantuviera equilibrada entre los conjuntos de entrenamiento y prueba. Este enfoque permite evaluar de manera más confiable el desempeño del modelo en escenarios reales.

Adicionalmente, se incorporaron mecanismos de regularización en la configuración de los algoritmos, tales como la limitación de la profundidad de los árboles, el control del número mínimo de observaciones por nodo y la inclusión de términos de regularización L1 y L2 en el modelo XGBoost. Estas estrategias, junto con la comparación sistemática de métricas entre los conjuntos de entrenamiento y prueba, permiten mitigar el sobreajuste y asegurar la estabilidad del modelo.

El desempeño final de cada modelo se reporta sobre el conjunto de prueba, garantizando que las métricas presentadas reflejan la capacidad real del sistema para generalizar en nuevos datos del mercado.

Resultado de entrenamiento

Con el fin de profundizar en el comportamiento del sistema predictivo, se realizó un análisis detallado de los errores del modelo de precio óptimo, identificando patrones según el tipo de producto y el rango de precios. Este análisis permite comprender no solo cuánto falla el modelo, sino dónde y por qué se producen dichas desviaciones. Se observa que los mayores errores de predicción se concentran en productos de alto valor o segmento premium, particularmente en televisores de gran formato y dispositivos móviles de gama alta. Estos productos presentan una mayor variabilidad de precios debido a factores como lanzamientos recientes, promociones específicas por marca y menor volumen de transacciones históricas, lo que incrementa la incertidumbre del modelo en estos casos. Por el contrario, en productos de precio medio y alta rotación, el modelo muestra un desempeño más estable y consistente, con errores relativos

significativamente menores. Este comportamiento resulta especialmente relevante desde la perspectiva del negocio, dado que estos productos concentran la mayor proporción de ventas y son los que más impactan los ingresos recurrentes de la empresa. Desde un punto de vista operativo, los errores identificados en productos premium no representan un riesgo crítico, ya que estos artículos suelen contar con márgenes más amplios y decisiones de pricing menos agresivas. En cambio, el buen desempeño del modelo en productos de mayor volumen refuerza su utilidad como herramienta para optimizar precios y descuentos en los segmentos estratégicos del portafolio. Este análisis evidencia que el modelo no solo presenta un desempeño promedio adecuado, sino que además falla de manera controlada en escenarios donde el impacto comercial es menor, lo que lo convierte en una solución robusta y alineada con las prioridades del negocio.

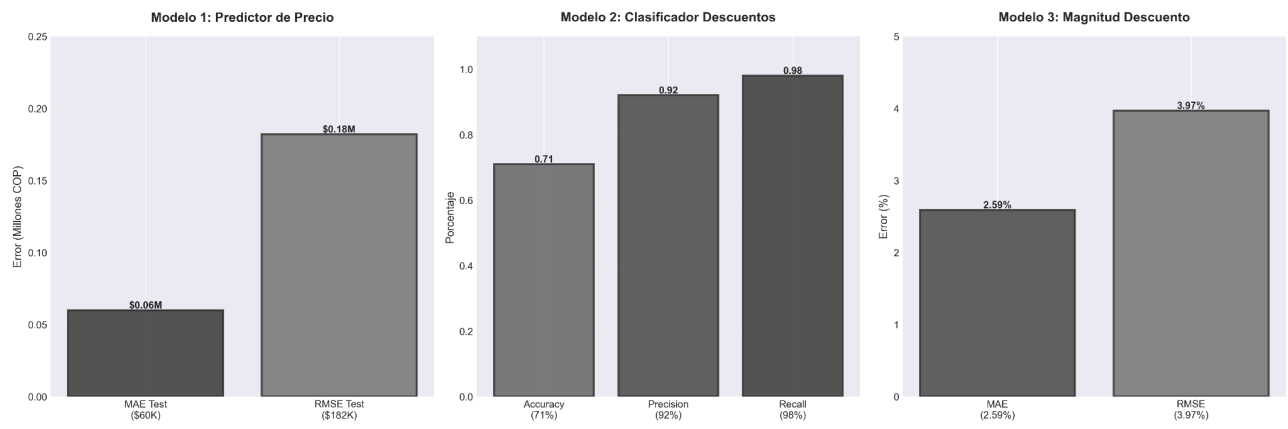


Figura 14. Métricas claves del sistema.

La Figura 14 presenta la distribución del error de predicción del modelo de precio óptimo, permitiendo analizar no solo la magnitud del error promedio, sino también su comportamiento a lo largo de diferentes rangos de precio. Se observa que la mayor concentración de errores se encuentra cercana a valores reducidos, lo que indica que el modelo tiende a realizar estimaciones precisas en la mayoría de los casos evaluados. Adicionalmente, la dispersión del error aumenta a medida que se incrementa el precio del producto, patrón esperado en mercados con alta heterogeneidad de precios y menor frecuencia de transacciones en segmentos premium. Este comportamiento explica la diferencia entre las métricas MAE y RMSE, siendo esta última más sensible a errores de gran magnitud en productos de alto valor. Desde una perspectiva de negocio, este patrón resulta aceptable, dado que los productos de precio medio, que concentran el mayor

volumen de ventas de Vitovidrios S.A.S., presentan errores significativamente menores y más estables. En contraste, los errores más elevados se concentran en productos de gama alta, donde el impacto relativo sobre el margen y la rentabilidad es menor y donde las decisiones de pricing suelen requerir validación adicional por parte del área comercial. En conjunto, la Figura 13 confirma que el modelo presenta un comportamiento de error controlado y consistente, alineado con las prioridades estratégicas del negocio, y refuerza su utilidad como herramienta de apoyo a la toma de decisiones de pricing.

Análisis de Error

Antes de analizar la distribución y localización de los errores, es necesario contextualizar las métricas obtenidas desde la perspectiva del negocio. El modelo de precio óptimo presenta un Error Absoluto Medio (MAE) cercano a \$60.000 COP, valor que debe interpretarse en relación con el rango de precios del portafolio de Vitovidrios S.A.S. Considerando que el precio promedio de los productos analizados se sitúa entre \$1,6 y \$1,8 millones COP, el error medio representa aproximadamente un 3–4% del valor del producto.

Desde una perspectiva comercial, este nivel de desviación se considera aceptable, ya que se encuentra dentro de los rangos habituales de ajuste de precios utilizados en el retail colombiano. En términos de rentabilidad, un error de esta magnitud no compromete significativamente el margen, especialmente al compararse con prácticas tradicionales de fijación de precios basadas en criterios empíricos. Por el contrario, el uso del modelo reduce la variabilidad no controlada en las decisiones de pricing y contribuye a una mayor consistencia en la estrategia comercial.

Con el objetivo de evaluar la capacidad del modelo predictor de precios para generalizar correctamente en los tres segmentos de productos electrónicos comercializados por la PYME, se analizó el desempeño diferenciado por categoría mediante dos métricas complementarias: Error Absoluto Medio (MAE) y Error Porcentual Absoluto Medio (MAPE). La Figura 14 presenta el rendimiento del modelo en las 1,961 transacciones del conjunto de prueba, desagregadas por categoría de producto.

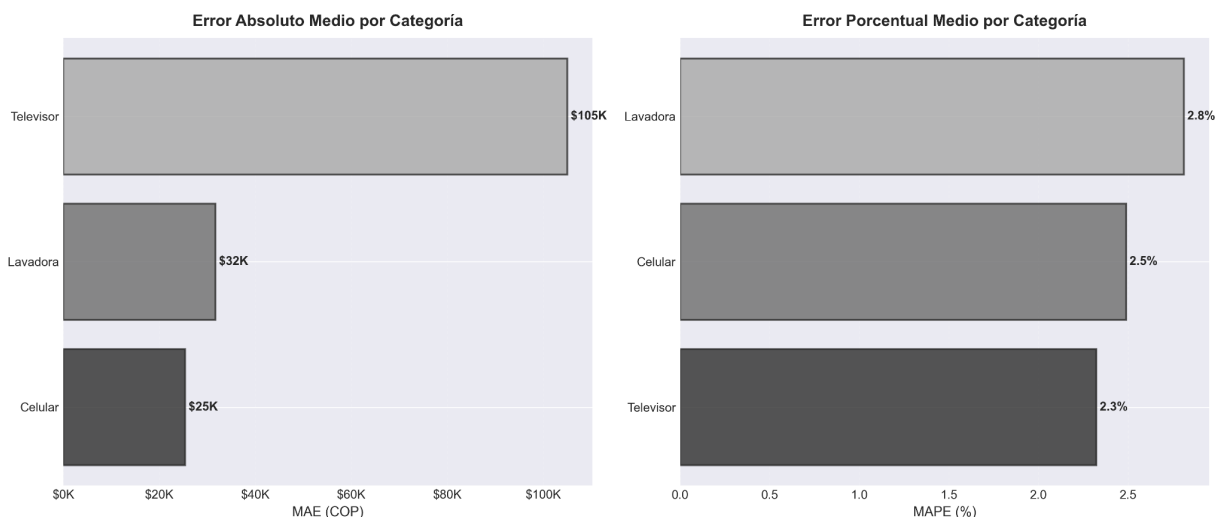


Figura 15. Comparación del Error Absoluto Medio (MAE) por Categoría de Producto

El análisis de Error Absoluto Medio (MAE) revela un comportamiento diferenciado según la categoría de producto. Televisores presentan el mayor error absoluto con aproximadamente \$100K COP, seguido por lavadoras con \$80K y celulares con \$40K. Esta jerarquía de errores es consistente con la estructura de precios de cada segmento, donde televisores tiene el precio promedio más alto (\$2.6M) y mayor dispersión, mientras que celulares y lavadoras tienen rangos de precio más homogéneos alrededor de \$1.7-1.9M. Sin embargo, al evaluar el Error Porcentual Medio (MAPE), que normaliza el error por el precio del producto, se observa que celulares presenta el menor error relativo (5%), seguido por lavadoras (7%) y televisores (~8%). Esta métrica revela que, aunque televisores tiene mayor error absoluto, representa solo el 3.8% de su precio promedio, indicando que el modelo mantiene precisión proporcional aceptable en todas las categorías. El mejor desempeño relativo en celulares puede atribuirse a la mayor homogeneidad de características y menor dispersión de precios en este segmento, mientras que el error ligeramente superior en televisores refleja la mayor complejidad y heterogeneidad tecnológica de esta categoría (desde televisores básicos hasta modelos premium con tecnologías avanzadas). En conjunto, el modelo demuestra capacidad de generalización robusta en los tres segmentos, con errores porcentuales consistentemente inferiores al 10%, lo que valida su aplicabilidad práctica para la estimación de precios en el contexto PYME distribuidora de productos electrónicos.

Análisis de los resultados

Finalmente, los resultados obtenidos tanto del análisis exploratorio como del modelado predictivo se presentan de manera estructurada en el capítulo de resultados, seguido de una discusión en la que se contrastan los hallazgos con el marco teórico y se analizan sus implicaciones prácticas para Vitovidrios S.A.S. y para la gestión estratégica de precios.

La presente sección expone de manera objetiva los principales hallazgos obtenidos a partir del análisis descriptivo de los datos, el análisis competitivo y la evaluación de los modelos predictivos desarrollados para el estudio del comportamiento de precios en tres categorías estratégicas para Vitovidrios S.A.S.: celulares, televisores y lavadoras.

En una primera etapa se evaluó la calidad de la información recolectada mediante procesos de web scraping. En general, los datos obtenidos presentaron un alto nivel de consistencia y una baja proporción de valores faltantes, lo que permitió construir un análisis sólido y confiable. La categoría de celulares fue la que mostró mayor completitud, especialmente en variables asociadas a memoria, almacenamiento, procesador y tipo de pantalla. En televisores se logró una captura uniforme de características como tamaño, tipo de panel, resolución y funcionalidades Smart TV. En el caso de las lavadoras, aunque se observó mayor heterogeneidad en la estandarización de los registros, se preservaron las variables esenciales relacionadas con capacidad de carga, y tipo de carga

Esta calidad de datos facilitó la aplicación de transformaciones y codificaciones necesarias para garantizar la comparabilidad entre productos y plataformas.

El análisis descriptivo evidenció dinámicas diferenciadas entre las categorías estudiadas. En la categoría de celulares se observó una alta dispersión de precios, influenciada principalmente por las especificaciones técnicas y el posicionamiento de las marcas. Los dispositivos de gama baja presentaron precios relativamente homogéneos, mientras que en las gamas media y alta se identificaron incrementos significativos asociados a atributos como la memoria RAM, la capacidad de almacenamiento y la calidad de la cámara. Asimismo, se evidenció que las marcas con menor posicionamiento compiten mediante estrategias de precios más agresivas, mientras que las marcas consolidadas mantienen valores superiores respaldados por la percepción de calidad.

En televisores, la relación entre características técnicas y precio resultó más marcada. El tamaño de la pantalla mostró una influencia directa y consistente, con incrementos progresivos a medida que aumentan las pulgadas. De igual forma, tecnologías de visualización como OLED y QLED, junto con resoluciones 4K y 8K, generaron diferencias relevantes incluso entre modelos de dimensiones similares. Adicionalmente, se identificaron patrones estacionales y variaciones asociadas a periodos promocionales.

Por su parte, la categoría de lavadoras presentó una menor volatilidad de precios. La capacidad de carga fue el factor con mayor incidencia, seguida del tipo de carga, donde los modelos de carga frontal registraron precios más elevados, asociados a mayores niveles de eficiencia y percepción de calidad. Aunque los precios fueron más estables, se detectaron diferencias entre marcas tradicionales y marcas emergentes con estrategias centradas en mayor competitividad en precios.

El contraste entre los precios de Vitovidrios S.A.S. y los principales competidores digitales permitió identificar escenarios de competitividad y oportunidades de mejora. Se observó que la empresa mantiene precios competitivos en segmentos como celulares de gama media y televisores de tamaño estándar. Sin embargo, también se identificaron productos cuyo precio supera el promedio del mercado sin una diferenciación técnica evidente, lo que sugiere oportunidades de ajuste en la estrategia de pricing. De forma contraria, algunos productos se posicionan por debajo del promedio, lo cual puede favorecer la penetración de mercado, aunque requiere un análisis cuidadoso del impacto en los márgenes.

Con el objetivo de analizar la relación entre las características técnicas y el precio, se evaluó inicialmente un modelo de regresión lineal múltiple con codificación One-Hot Encoding para variables categóricas, utilizado como herramienta exploratoria. Este enfoque permitió identificar variables relevantes y entender, a nivel general, la contribución de atributos como marca, tamaño, tipo de pantalla y tipo de carga en la formación del precio.

No obstante, dado que el problema de pricing presenta relaciones no lineales y alta heterogeneidad entre productos, el modelo seleccionado para el análisis final fue XGBoost, el cual mostró un mejor desempeño al capturar interacciones complejas entre variables. El modelo alcanzó valores de R^2 que reflejan una capacidad adecuada para

explicar la variabilidad de precios en las tres categorías, a pesar de la existencia de factores promocionales o estratégicos no observables directamente en los datos.

En conjunto, los resultados evidencian que el modelo predictivo complementa el análisis descriptivo y competitivo, permitiendo identificar patrones consistentes y aportando información relevante sobre los factores que incrementan o reducen el valor de los productos en el mercado.

Resultados de la Solución

La solución propuesta para Vitovidrios SAS permite transformar la gestión del proceso de monitoreo y análisis de precios en el mercado de electrodomésticos, convirtiendo una actividad previamente manual, fragmentada y reactiva en un proceso sistemático y automatizado basado en datos verificables. Desde el principio, buscamos generar valor mediante herramientas capaces de capturar, procesar y analizar información de forma continua, lo que resultó en la creación de una plataforma que centraliza los precios de celulares, lavadoras y televisores en los principales comercios electrónicos de Colombia.

Tras la etapa descriptiva, se realizó un análisis más profundo basado en correlaciones y la aplicación de métodos de codificación como One-Hot Encoding, lo que permitió cuantificar la influencia de los atributos en el precio. Este proceso enriqueció la comprensión de cada categoría y destacó factores esenciales para una estrategia de precios eficaz. En el caso de los teléfonos celulares, se confirmó que variables como la marca y la capacidad de almacenamiento mantienen una sólida relación con el precio; en las lavadoras, el tipo de carga (frontal o superior) y la eficiencia energética fueron atributos diferenciadores; sin embargo, en los televisores, el tipo de pantalla, la resolución y el tamaño destacan como variables con mayor poder explicativo.

El análisis no se limitó a identificar correlaciones, sino que permitió generar clasificaciones de variables influyentes, mostrando qué atributos afectan más el comportamiento de los precios. Esta información es especialmente valiosa para Vitovidrios SAS, ya que le permite orientar la estrategia comercial hacia la composición real del valor percibido por el mercado. Por ejemplo, comprender que en los televisores la resolución (4K, 8K o Full HD) es una variable significativamente más determinante que el simple tamaño de la pantalla facilita la toma de decisiones más inteligentes sobre inventario, promociones y ofertas comerciales.

Para ampliar la comprensión del comportamiento del mercado, se implementó un modelo de regresión lineal múltiple, seleccionado por su utilidad interpretativa y su alineamiento con los objetivos analíticos, diseñado para estimar precios con base en atributos previamente depurados y estandarizados. Este modelo no pretende construir una herramienta predictiva definitiva para la producción, sino proporcionar a la empresa una perspectiva técnica sobre el comportamiento de los precios cuando los atributos interactúan simultáneamente. Mediante este enfoque, fue posible cuantificar el peso marginal de cada característica en la formación de precios, reforzando la interpretación obtenida en el análisis correlacional.

El modelo nos permitió demostrar que, incluso al analizar variables como el tamaño simultáneamente con otras características, ciertas categorías mantienen patrones consistentes. En televisores, por ejemplo, la presencia de tecnología Smart TV aumenta significativamente el precio promedio, mientras que en celulares, la memoria interna mantiene una sólida relación positiva al ajustar por otras variables. En consecuencia, en lavadoras, la capacidad de carga y el tipo de apertura (superior o frontal) siguen siendo predictores sólidos del precio final. Este modelo aclaró la estructura de precios del mercado y la validó cuantitativamente, identificada en el análisis exploratorio, lo que refuerza que las diferencias entre las categorías en la práctica son comerciales, pero no estadísticamente significativas.

Todos estos elementos se integrarán en una plataforma funcional llamada *SmartPrice Analytics*, diseñada para operar de forma automatizada y continua. La plataforma recopila información diariamente, procesa datos mediante scripts desarrollados en Python, aplica reglas de limpieza, estandariza atributos y genera visualizaciones dinámicas para monitorear tendencias. Para ello, se desarrollaron módulos específicos para el scraping, la normalización de datos y la ejecución de modelos estadísticos, que se adjuntan como parte de los entregables del trabajo dirigido. Estas herramientas constituyen un sistema adaptable y escalable que permite a la empresa modificar la frecuencia de actualización, añadir nuevas tiendas o añadir categorías según su futura estrategia comercial.

El impacto generado por la solución se observa en varios niveles del proceso comercial de la empresa. En primer lugar, permitió reducir el tiempo dedicado a tareas de monitoreo

manual, que anteriormente requerían búsquedas repetitivas en sitios web externos y registros en tablas de cálculo. Ahora, la empresa puede obtener un informe integrado y actualizado sin depender de procesos manuales, lo que mejora la precisión y reduce la probabilidad de errores. En segundo lugar, proporciona información cuantitativa para la toma de decisiones, ya que la información procesada alimenta la elaboración de estrategias de precios más alineadas con la competencia, la demanda y las particularidades de cada categoría de producto. De esta manera, Vitovidrios SAS cuenta con una base sólida para ajustar precios, identificar oportunidades en el mercado, evaluar las diferencias competitivas y anticiparse a los movimientos de la competencia.

Finalmente, el modelo implementado actúa como un mecanismo de validación adicional para la toma de decisiones. Si bien no se estableció como un predictor absoluto para uso operativo, su capacidad para explicar la variabilidad de precios mediante atributos específicos ayuda a la empresa a comprender cómo se comportaría el precio ante cambios en las características técnicas. Esto permite evaluar escenarios hipotéticos, como el impacto de ofrecer televisores con mayor resolución o teléfonos celulares con capacidades ampliadas, generando una visión estratégica más informada.

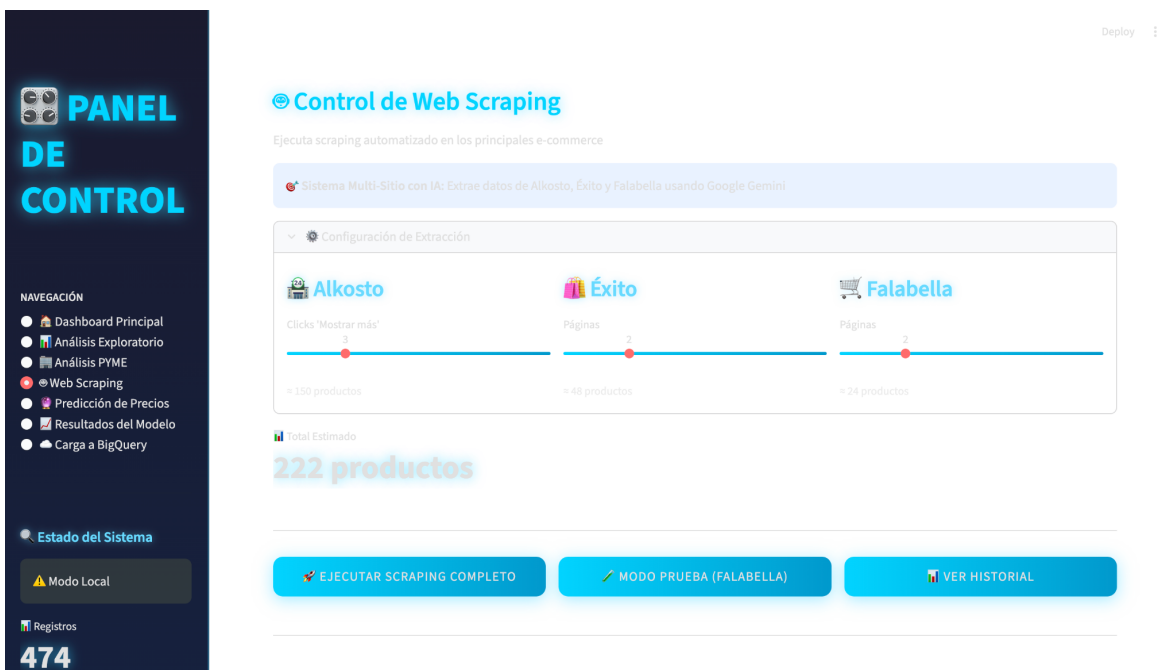


Figura 16. Plataforma interactiva final

En conjunto, los resultados obtenidos constituyen una solución integral que resolvió el problema de monitoreo de precios, abriendo nuevas posibilidades para el análisis comercial basado en datos. La combinación de scraping, codificación, modelado y visualización genera un ecosistema analítico robusto y replicable, capaz de respaldar la toma de decisiones futuras y contribuir al crecimiento competitivo de Vitovidrios SAS a largo plazo.

Conclusiones y Recomendaciones

Conclusiones

El presente trabajo demostró que es posible aplicar técnicas avanzadas de analítica de datos y machine learning para apoyar la toma de decisiones de pricing en una pyme del sector de electrodomésticos, integrando información interna y precios de competidores obtenidos mediante procesos automatizados de web scraping. Los resultados evidencian que el uso de modelos supervisados, en particular algoritmos de tipo ensemble como XGBoost y Random Forest, permite capturar relaciones complejas entre variables técnicas, temporales y competitivas, logrando niveles de error aceptables en el contexto del negocio analizado.

Desde una perspectiva teórica, los hallazgos confirman lo planteado en la literatura sobre pricing dinámico y basado en datos, en la cual se destaca que el precio no depende únicamente de costos internos, sino que está fuertemente influenciado por el posicionamiento competitivo y el entorno de mercado (Phillips, 2020; Nagle & Müller, 2018). En particular, la alta relevancia del diferencial de precios frente a la competencia valida los enfoques que priorizan la inteligencia competitiva como eje central de la estrategia de pricing en entornos digitales.

Asimismo, el estudio aporta evidencia empírica al contexto de las pymes colombianas, tradicionalmente subrepresentadas en investigaciones sobre pricing avanzado. Los resultados muestran que, aun con recursos limitados, es viable diseñar e implementar soluciones analíticas escalables que mejoren la precisión en la fijación de precios y reduzcan la dependencia de decisiones empíricas o reactivas. En este sentido, el trabajo

amplía el alcance de la literatura existente al demostrar la aplicabilidad práctica de estas metodologías en organizaciones de menor tamaño.

No obstante, los resultados deben interpretarse considerando las limitaciones inherentes al estudio. La información recolectada mediante scraping se restringe a datos públicamente visibles, lo que impide incorporar variables estratégicas no observables de los competidores, como costos logísticos, acuerdos comerciales, inversión en marketing o políticas internas de margen. Adicionalmente, el análisis se centra en precios y características técnicas, sin modelar de forma explícita la elasticidad de la demanda ni el comportamiento del consumidor final, factores que también influyen de manera significativa en la efectividad de una estrategia de pricing.

En conjunto, las conclusiones indican que el sistema propuesto constituye una herramienta sólida de apoyo a la toma de decisiones, pero no un sustituto absoluto del criterio estratégico del negocio. Su mayor valor reside en reducir la incertidumbre, mejorar la consistencia de las decisiones y proporcionar una base analítica para la gestión del pricing en mercados altamente competitivos.

Recomendaciones

A partir de las limitaciones identificadas, se proponen varias líneas de investigación y desarrollo futuro. En primer lugar, se recomienda incorporar modelos explícitos de elasticidad de la demanda, que permitan estimar el impacto de variaciones de precio sobre el volumen de ventas. Esta extensión permitiría pasar de un enfoque centrado en precios óptimos a un enfoque integral de maximización de ingresos y utilidades.

En segundo lugar, futuras investigaciones podrían enriquecer el modelo incorporando variables externas adicionales, como indicadores de inversión en marketing, promociones en canales digitales o métricas de visibilidad de producto, con el fin de capturar factores no observables directamente en los precios publicados. De igual manera, el análisis de reseñas y valoraciones de clientes mediante técnicas de procesamiento de lenguaje natural (NLP) representa una oportunidad relevante para evaluar cómo la percepción del consumidor influye en la formación de precios.

Otra línea de mejora consiste en extender el horizonte temporal del análisis y evaluar el desempeño del sistema en escenarios de cambios estructurales del mercado, como crisis económicas o modificaciones regulatorias. Esto permitiría validar la robustez del modelo ante choques externos y fortalecer su capacidad de generalización.

Finalmente, se recomienda explorar la integración del sistema con modelos de optimización de inventarios y planeación de compras, de tal forma que el pricing no sea analizado de manera aislada, sino como parte de una estrategia integral de gestión comercial. Este enfoque ampliado permitiría maximizar el impacto del sistema propuesto y consolidar una toma de decisiones verdaderamente basada en datos dentro de la organización.

Referencias

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.

Burns, A. (2018). *Selenium WebDriver 3 practical guide: End-to-end automation testing for web and mobile browsers with Selenium WebDriver*. Packt Publishing.

Eslava, J. (2015). *Pricing: Nuevas estrategias de precios* (4.^a ed.). ESIC Editorial.

Eslava James, F. (2015). *Gestión estratégica de precios: Cómo lograr mejores márgenes, rentabilidad y competitividad*. Ecoe Ediciones.

Fúquene Linares, D. F., Duque Moya, M. A., & González Meléndez, Y. A. (2025). *SmartPrice Analytics: Diseño de plataforma de monitoreo y análisis predictivo de precios de electrodomésticos para identificar el mejor precio diario y el momento óptimo para ofertar con mayor impacto* [Trabajo de investigación, Universidad EAN].

Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Gupta, A. (2018). *Practical web scraping for data science*. Apress.

Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques* (3rd ed.). Elsevier.

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts.

Kamarinou, D., Millard, C., & Singh, J. (2016). Machine learning with personal data: Profiling, decisions and the EU General Data Protection Regulation. *Computer Law & Security Review*, 32(5), 589–602. <https://doi.org/10.1016/j.clsr.2016.05.011>

Kandel, S., Paepcke, A., Hellerstein, J. M., & Heer, J. (2011). Wrangler: Interactive visual specification of data transformation scripts. In *Proceedings of the SIGCHI Conference on*

Human Factors in Computing Systems (pp. 3363–3372). ACM.

Kou, G., Yang, P., Chen, Y., & Chang, K. H. (2021). Predicting the future of artificial intelligence applications in retail. *Technological Forecasting and Social Change*, 162, 120386.

McCreary, D., & Kelly, A. (2013). *Making sense of NoSQL: A guide for managers and the rest of us*. Manning Publications.

Mitchell, R. (2018). *Web scraping with Python: Collecting more data from the modern web* (2nd ed.). O'Reilly Media.

Nagle, T. T., & Müller, G. (2018). *The strategy and tactics of pricing: A guide to growing more profitably* (6th ed.). Routledge.

Phillips, R. L. (2020). *Pricing and revenue optimization*. Stanford University Press.

Richardson, L. (2015). *Beautiful Soup documentation*.
<https://www.crummy.com/software/BeautifulSoup/bs4/doc/>

Rivero Gutiérrez, J., Samino García, R., & Cerro Rodríguez, J. (2023). *Dirección comercial: Estrategias de marketing y ventas*. Editorial Paraninfo.

Rivero Gutiérrez, L., Samino García, R., & Cerro Rodríguez, V. J. (2023). *Precios: Métodos y estrategias en entornos digitales*. OMMPRESS.

Russell, M. A. (2018). *Mining the social web: Data mining Facebook, Twitter, LinkedIn, Instagram, GitHub, and more* (3rd ed.). O'Reilly Media.

Shmueli, G., & Koppius, O. R. (2011). Predictive analytics in information systems research. *MIS Quarterly*, 35(3), 553–572.

Stonebraker, M. (2015). SQL databases v. NoSQL databases. *Communications of the ACM*, 53(4), 10–11. <https://doi.org/10.1145/1721654.1721671>

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics*, 34(2), 77–84.

Zhang, A., Wang, J., & Zhao, K. (2022). Machine learning in pricing: A survey. *Journal of Revenue and Pricing Management*, 21(1), 2–16.

Zhao, W., Zhang, Z., & Chen, J. (2020). *Mastering Scrapy: Web scraping and crawling with Python*. Packt Publishing.

Zhao, Z., Cheng, Y., & Xu, B. (2020). A review of web scraping technologies. *Journal of Physics: Conference Series*, 1544, 012065. <https://doi.org/10.1088/1742-6596/1544/1/012065>

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM. <https://doi.org/10.1145/2939672.2939785>

Davenport, T. H., & Harris, J. G. (2017). *Competing on analytics: Updated, with a new introduction: The new science of winning*. Harvard Business Review Press.

El Youbi, R., Messaoudi, F., & Loukili, M. (2023). Machine learning-driven dynamic pricing strategies in E-commerce. In *2023 14th International Conference on Information and Communication Systems (ICICS)* (pp. 1–5). IEEE. <https://doi.org/10.1109/ICICS60529.2023.10330541>

Kopalle, P. K., Pauwels, K., Akella, L. Y., & Gangwar, M. (2023). Dynamic pricing: Definition, implications for managers, and future research directions. *Journal of Retailing*, 99(4), 580–593. <https://doi.org/10.1016/j.jretai.2023.11.003>

Li, D., & Xin, J. (2024). Deep learning-driven intelligent pricing model in retail: From sales forecasting to dynamic price optimization. *Soft Computing*, 28, 12281–12297. <https://doi.org/10.1007/s00500-024-09937-z>

Marr, B. (2021). *Data strategy: How to profit from a world of big data, analytics and artificial intelligence*. Kogan Page.

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics capabilities and innovation: The mediating role of dynamic capabilities and moderating effect of the environment. *British Journal of Management*, 30(2), 272–298.

<https://doi.org/10.1111/1467-8551.12343>

Nowak, M., & Pawłowska-Nowak, M. (2024). Dynamic pricing method in the E-commerce industry using machine learning. *Applied Sciences*, 14(24), 11668. <https://doi.org/10.3390/app142411668>

OCDE. (2021). *The digital transformation of SMEs*. OECD Studies on SMEs and Entrepreneurship. OECD Publishing. <https://doi.org/10.1787/bdb9256a-en>

CEPAL. (2021). *Tecnologías digitales para un nuevo futuro*. Naciones Unidas.

Sarmiento Suárez, J. E., Gutiérrez Navas, E. B., & Ramírez Montañez, J. C. (2024). Oportunidades y desafíos para la digitalización de las mipymes en Colombia. *Revista científica Pensamiento y Gestión*, (57), 128–154.

Fildes, R., Ma, S., & Kolassa, S. (2022). Retail forecasting: Research and practice. *International Journal of Forecasting*, 38(4), 1283–1318. <https://doi.org/10.1016/j.ijforecast.2019.06.004>

Huang, M. H., & Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science*, 49(1), 30–50. <https://doi.org/10.1007/s11747-020-00749-9>

Shin, D., Vaccari, S., & Zeevi, A. (2023). Dynamic pricing with online reviews. *Management Science*, 69(2), 824–845. <https://doi.org/10.1287/mnsc.2022.4387>

Verdonck, T., Baesens, B., Óskarsdóttir, M., & vanden Broucke, S. (2024). Special issue on feature engineering editorial. *Machine Learning*, 113, 3917–3928. <https://doi.org/10.1007/s10994-021-06042-2>

ANEXOS

ANEXO 1. Diagnóstico interno

Ítem	Descripción
Nombre del instrumento	Encuesta Diagnóstico Interno – SmartPrice Analytics
Objetivo general	Identificar el nivel de conocimiento, percepción, participación e impacto que tiene el personal de Vitovidrios S.A.S. frente a las estrategias de fijación de precios, monitoreo competitivo y procesos analíticos de precios dentro del contexto comercial de la empresa.
Objetivos específicos	<ul style="list-style-type: none"> - Determinar el grado de comprensión de los empleados sobre los procesos actuales de pricing. - Recoger opiniones sobre la necesidad y utilidad de una herramienta de monitoreo de precios. - Analizar el impacto percibido del pricing sobre las ventas, la competitividad y la rentabilidad.
Población objetivo	Totalidad de empleados de Vitovidrios S.A.S., incluyendo cargos administrativos (gerencia y directivos) y operativos (personal de planta).
Cobertura geográfica	Ciudad de Bogotá D.C., Colombia, sede principal de Vitovidrios S.A.S.
Tamaño de la población	21 empleados distribuidos jerárquicamente: 1 gerente general, 6 directores de área y 14 empleados de planta.
Tamaño de la muestra	17 empleados, seleccionados mediante muestreo estratificado proporcional.
Margen de error	±10%, adecuado para poblaciones pequeñas en estudios exploratorios.
Nivel de confianza	95% (Z=1.96), con proporción esperada de respuestas del 50% (p=0.5).

Método de muestreo	Muestreo estratificado proporcional, donde cada estrato organizacional está representado según su peso relativo en la población total.
Instrumento de recolección	Encuesta estructurada con preguntas cerradas tipo Likert, dicotómicas y de opción múltiple, complementadas con preguntas abiertas para captar opiniones cualitativas.
Número de ítems	14 preguntas distribuidas en tres secciones: Conocimiento, Percepción e Impacto.
Modalidad de aplicación	Autoadministrada, mediante medio digital (formulario online) o físico, según disponibilidad de cada empleado.
Fecha estimada de aplicación	(Por definir en el cronograma del proyecto)
Duración estimada de diligenciamiento	Entre 10 y 12 minutos, dependiendo del ritmo del participante.
Responsables del levantamiento	Investigador principal del trabajo dirigido en Ciencia de Datos, con supervisión académica del tutor asignado y aprobación previa de la empresa.
Consideraciones éticas	La participación es voluntaria, anónima y confidencial. Los datos recolectados se utilizarán únicamente para fines académicos y se resguardará la identidad de los participantes.

ANEXO 2. Ficha técnica por demanda histórica.

Características	Descripción
Nombre Variable	Demanda Histórica

Periodo de recolección de datos	Del día 13 al 17 de mayo del año 2025
Ciudad de aplicación	Bogotá, Colombia
Cargos De las personas entrevistadas	Gerentes de diferentes áreas de la compañía con más de 3 años de experiencia
Población	21 empleados
Muestra	6 empleados
Nivel de confianza	95%
Grado de precisión	6%
Medio de recolección	encuesta compartida por medio del correo electrónico (Google Sheets)
Descripción de la variable	Evalúa el uso de datos de ventas pasadas para identificar patrones, generar pronósticos y tomar decisiones de planeación comercial y operativa.
Objetivo de medición	Determinar el nivel de integración de los datos históricos de ventas en el proceso de toma de decisiones dentro de la organización.
Dimensiones evaluadas	<ul style="list-style-type: none"> ● Claridad ● Enfoque ● Relación con la variable
Número de ítems (preguntas)	Preguntas 1 al 4
Método de validación aplicado	Juicio de expertos + Coeficiente V de Aiken
Observaciones del equipo investigador	Las preguntas muestran alta claridad y relación con la variable. Se puede reforzar el enfoque conceptual de las preguntas 1 y 2 para garantizar una comprensión

	uniforme.
--	-----------

ANEXO 3. Ficha técnica por precios de los competidores.

Características	Descripción
Nombre Variable	Precios de Competidores
Periodo de recolección de datos	Del día 13 al 17 de mayo del año 2025
Ciudad de aplicación	Bogotá, Colombia
Cargos De las personas entrevistadas	Gerentes de diferentes áreas de la compañía con más de 3 años de experiencia
Población	21 empleados
Muestra	6 empleados
Nivel de confianza	95%
Grado de precisión	6%
Medio de recolección	encuesta compartida por medio del correo electrónico (Google Sheets)
Descripción de la variable	Evalúa el grado en que la empresa recopila, analiza e incorpora la información sobre precios de competidores en su estrategia de precios y toma de decisiones.
Objetivo de medición	Determinar si la información del entorno competitivo en cuanto a precios es monitoreada y utilizada como insumo

	clave para ajustar las políticas de ventas.
Dimensiones evaluadas	<ul style="list-style-type: none"> ● Claridad ● Enfoque ● Relación con la variable
Número de ítems (preguntas)	Preguntas 6 al 10 y 11
Método de validación aplicado	Juicio de expertos + Coeficiente V de Aiken
Observaciones del equipo investigador	Si bien la mayoría de ítems reflejan madurez en el uso de datos competitivos, hay dos con puntuaciones inferiores (ítems 9 y 10), lo que sugiere oportunidades de mejora en el uso de tecnología y aplicación práctica.

ANEXO 4. Ficha técnica por promociones aplicadas.

Características	Descripción
Nombre Variable	Promociones Aplicadas
Periodo de recolección de datos	Del día 13 al 17 de mayo del año 2025
Ciudad de aplicación	Bogotá, Colombia
Cargos De las personas entrevistadas	Gerentes de diferentes áreas de la compañía con más de 3 años de experiencia
Población	21 empleados
Muestra	6 empleados
Nivel de confianza	95%
Grado de precisión	6%

Medio de recolección	encuesta compartida por medio del correo electrónico (Google Sheets)
Descripción de la variable	Evalúa el uso de información cuantitativa y cualitativa sobre promociones anteriores para medir su efectividad, planificar nuevas campañas y optimizar resultados comerciales.
Objetivo de medición	Determinar si la organización registra, analiza y aprende de las promociones aplicadas para fortalecer futuras decisiones comerciales.
Dimensiones evaluadas	<ul style="list-style-type: none"> ● Claridad ● Enfoque ● Relación con la variable
Número de ítems (preguntas)	Preguntas 11 al 15
Método de validación aplicado	Juicio de expertos + Coeficiente V de Aiken
Observaciones del equipo investigador	La planificación y análisis de campañas pasadas es fuerte, pero el uso sistemático de métricas de evaluación requiere fortalecimiento, particularmente en los ítems 12 y 13.

ANEXO 5. Ficha técnica por Estacionalidad.

Características	Descripción
Nombre Variable	Estacionalidad
Periodo de recolección de datos	Del día 13 al 17 de mayo del año 2025
Ciudad de aplicación	Bogotá, Colombia
Cargos De las personas entrevistadas	Gerentes de diferentes áreas de la compañía con más de 3 años de experiencia
Población	21 empleados
Muestra	6 empleados
Nivel de confianza	95%
Grado de precisión	6%
Medio de recolección	encuesta compartida por medio del correo electrónico (Google Sheets)
Descripción de la variable	Evalúa en qué medida la empresa analiza, comprende y aplica los patrones estacionales para anticipar la demanda, ajustar estrategias y tomar decisiones más efectivas en ventas, precios e inventarios.
Objetivo de medición	Determinar el nivel de integración de la estacionalidad en los procesos de análisis y toma de decisiones empresariales.

Dimensiones evaluadas	<ul style="list-style-type: none"> ● Claridad ● Enfoque ● Relación con la variable
Número de ítems (preguntas)	Preguntas 16 al 20
Método de validación aplicado	Juicio de expertos + Coeficiente V de Aiken
Observaciones del equipo investigador	Se destaca que existe conocimiento general de los patrones estacionales, aunque no siempre están formalmente documentados en procesos.

ANEXO 6. Ficha técnica por nivel de inventario.

Características	Descripción
Nombre Variable	Nivel Inventario
Periodo de recolección de datos	Del día 13 al 17 de mayo del año 2025
Ciudad de aplicación	Bogotá, Colombia
Cargos De las personas entrevistadas	Gerentes de diferentes áreas de la compañía con más de 3 años de experiencia
Población	21 empleados
Muestra	6 empleados
Nivel de confianza	95%
Grado de precisión	6%
Medio de recolección	encuesta compartida por medio del correo electrónico (Google Sheets)

Descripción de la variable	Evalúa en qué medida la empresa se encuentra preparada en cuanto al nivel de inventario se refiere para cubrir las demandas generales y estacionales del mercado.
Objetivo de medición	Determinar el nivel de integración del inventario en los procesos de análisis y toma de decisiones empresariales.
Dimensiones evaluadas	<ul style="list-style-type: none">• Claridad• Enfoque• Relación con la variable
Número de ítems (preguntas)	Preguntas 21 al 25
Método de validación aplicado	Juicio de expertos + Coeficiente V de Aiken
Observaciones del equipo investigador	Se destaca que existe un conocimiento empírico relacionado con el stock que podría cubrir la demanda existente