



ESCUELA DE NEGOCIOS

MAESTRÍA EN INTELIGENCIA DE NEGOCIOS Y CIENCIA DE DATOS

**OPTIMIZACIÓN DE LA GESTIÓN DE ESTADOS FINANCIEROS EN UNA EMPRESA
DE ALIMENTOS Y BEBIDAS, A TRAVÉS DE ANÁLISIS EXPLORATORIO,
MODELOS DE REGRESIÓN Y SERIES TEMPORALES CON VISUALIZACIÓN EN
POWER BI**

**Profesor
Manuel Morocho Cayamcela**

**Autores
César Ayerve
Lizeth Peña**

2024

RESUMEN

Este proyecto de grado explora la aplicación de técnicas de inteligencia de negocios y ciencia de datos para optimizar la gestión de estados financieros en una de las principales empresas de alimentos y bebidas. En el desarrollo del proyecto se implementan métodos de análisis de datos, incluyendo análisis exploratorio, modelos predictivos como la regresión RIDGE y ARIMA, y herramientas de visualización como Power BI. El objetivo principal es mejorar la toma de decisiones estratégicas y operativas de una destacada empresa del sector, enfrentando el desafío de procesar y gestionar grandes volúmenes de datos financieros con precisión y eficiencia.

Los resultados demuestran que mediante el uso de estas técnicas avanzadas se puede lograr una mejor comprensión del comportamiento financiero, anticipar tendencias y adaptar estrategias en respuesta a las dinámicas del mercado. Los dashboards interactivos y los modelos de predicción permiten una gestión más informada y proactiva, evidenciando una mejora significativa en la eficiencia y precisión de las operaciones financieras de la compañía.

En conclusión, el proyecto valida la efectividad de integrar tecnologías de análisis y visualización de datos en la gestión financiera, sugiriendo su potencial para ser replicado en otras áreas de la empresa o incluso en otras industrias, fortaleciendo así la toma de decisiones en el área financiera y contable basada en datos y la competitividad empresarial.

ABSTRACT

This degree project explores the application of business intelligence and data science techniques to optimize financial statement management in one of the leading food and beverage companies. Throughout the project, data analysis methods are implemented, including exploratory analysis, predictive models such as RIDGE regression and ARIMA, and visualization tools like Power BI. The main objective is to enhance strategic and operational decision-making in a prominent company within the sector, facing the challenge of processing and managing large volumes of financial data with precision and efficiency.

The results demonstrate that through the use of these advanced techniques, a better understanding of financial behavior can be achieved, allowing for the anticipation of trends and the adaptation of strategies in response to market dynamics. Interactive dashboards and predictive models enable more informed and proactive management, showing significant improvements in the efficiency and accuracy of the company's financial operations.

In conclusion, the project validates the effectiveness of integrating data analysis and visualization technologies into financial management, suggesting its potential to be replicated in other areas of the company or even in other industries, thereby strengthening data-based decision-making in the financial and accounting area and enhancing business competitiveness.

ÍNDICE DEL CONTENIDO

INTRODUCCIÓN	1
REVISIÓN DE LITERATURA.....	2
IDENTIFICACIÓN DEL OBJETO DE ESTUDIO	9
PLANTEAMIENTO DEL PROBLEMA.....	11
OBJETIVO GENERAL.....	13
OBJETIVOS ESPECÍFICOS	13
JUSTIFICACIÓN Y APLICACIÓN DE LA METODOLOGÍA	15
RESULTADOS	25
Análisis Exploratorio	25
Análisis Predictivo.....	36
DISCUSIÓN DE LOS RESULTADOS Y PROPUESTA DE SOLUCIÓN.....	43
Implicaciones para la organización	43
CONCLUSIONES Y RECOMENDACIONES	46
Conclusiones.....	46
Recomendaciones	47
REFERENCIAS	48
ANEXOS	50

ÍNDICE DE TABLAS

Tabla 1:Matriz de descripción de variables.....	17
---	----

ÍNDICE DE FIGURAS

Figura 1: Variables Cuantitativas	18
Figura 2: Variables Cuantitativas Predictoras.....	18
Figura 3: Variables Cualitativas	19
Figura 4: Pruebas de predicción	22
Figura 5: Visualización del dataset en Python	25
Figura 6: Correlación general de las variables	26
Figura 7: Correlación general de las variables	26
Figura 8: Correlación entre Gastos de Marketing y Ventas por año	27
Figura 9: Correlación entre Descuentos por desempeño y Gastos_de_Marketing	28
Figura 10: Tendencia de Ventas Netas por Unidad de Negocio.....	29
Figura 11: Comparación de Ventas Netas entre Unidades de Negocio.....	29
Figura 12: Distribución de Ventas Netas	30
Figura 13: Resumen de rentabilidad	30
Figura 14: Resumen de ratios	30
Figura 15: Predicción gasto en medios digitales (RIDGE)	37
Figura 16: Predicción Cartera Dudosa (Ridge).....	38
Figura 17: Predicción Ventas Brutas (Arima)	39
Figura 18: Predicción Costo Materia Prima (Arima).....	41
Figura 19: Dashboard Interactivo de Estados Financieros.....	42

INTRODUCCIÓN

En el entorno altamente competitivo de la industria de alimentos y bebidas, la gestión financiera eficiente se ha convertido en un factor crítico para el éxito y la sostenibilidad empresarial. Este proyecto se centra en la aplicación de técnicas avanzadas de inteligencia de negocios y ciencia de datos con el objetivo de optimizar la gestión financiera en una de las empresas más importantes de este sector. Mediante el uso de análisis exploratorio de datos, se busca identificar patrones y tendencias que permitan una mejor comprensión del comportamiento financiero de las empresas.

Además, el proyecto incorpora modelos predictivos, como la regresión RIDGE, para predecir variables financieras clave y apoyar la toma de decisiones informadas. También se exploran series temporales utilizando modelos ARIMA para analizar y prever cambios en las tendencias financieras a lo largo del tiempo. Por último, el uso de herramientas de visualización de datos, como Power BI, facilita la presentación de los resultados de manera clara y concisa, permitiendo a los gestores tomar decisiones basadas en datos de manera eficiente. Con esto se busca mejorar la capacidad de la empresa para gestionar sus recursos financieros y así, proporcionar una ventaja competitiva al anticipar cambios en el mercado y adaptarse a ellos de manera proactiva.

REVISIÓN DE LITERATURA

En el contexto actual de los negocios, la capacidad de analizar y aprovechar los datos financieros generados y obtenidos interna y externamente, es crucial para las compañías del sector de alimentos y bebidas . Numerosos estudios han explorado el uso de herramientas de análisis de datos e inteligencia artificial en la gestión y proyección de cifras financieras, demostrando cómo estas tecnologías pueden mejorar la toma de decisiones estratégicas y operativas (Delfino Barilla & Lastarria, 2020).

Esta fase del proyecto se centra en revisar estudios previos que implementan métodos de análisis exploratorio, predictivo, correlacional y visualización de datos para obtener información relevante sobre los componentes de estados financieros. Al entender las soluciones y modelos desarrollados en estos estudios, se pretende identificar las mejores prácticas y enfoques más efectivos que pueden ser aplicados en nuestro proyecto para optimizar la gestión financiera y fortalecer la competitividad en el mercado de la compañía estudiada.

Dado que el uso de plataformas de software para la gestión de Big Data y análisis de datos en las áreas de contabilidad y finanzas aún tiene mucho camino por desarrollar en las empresas ecuatorianas y no existen investigaciones previas públicas y disponibles en el país, hemos recopilado información de fuentes primarias y secundarias extrajas en donde el uso de herramientas como el análisis exploratorio y predictivo, Python y la visualización de datos en dashboards interactivos ya se encuentra más desarrollado.

Cinzia Delfino Barilla y Lucia Lastarria Reynoso realizaron un estudio que presenta una metodología que combina herramientas de análisis de Big Data con análisis predictivo, de riesgo y estratégico para optimizar el desempeño financiero de las empresas, medido a través de Indicadores Clave de Desempeño (KPI). Los resultados destacan que esta metodología facilita la innovación y la flexibilidad en la empresa, mejorando su

rendimiento financiero siempre que se mantenga un buen ajuste entre las actividades de Big Data, la estructura organizacional y el compromiso de la alta gerencia con el desarrollo de capacidades dinámicas orientadas a procesos (Delfino Barilla & Lastarria, 2020).

En las investigaciones previas, se ha identificado que las herramientas principales que han sido implementadas para mejorar la toma de decisiones sobre estados financieros incluyen Python, con bibliotecas como Pandas para la manipulación de datos (Hilpisch, 2019), Scikit-learn para el aprendizaje automático (Tatsat et al., 2020) y Matplotlib para la visualización de datos; y R, con paquetes como dplyr para la manipulación de datos y ggplot2 para la visualización (Fernández, 2020). También se utilizan herramientas de Business Intelligence como Tableau (Tableau Software, n.d.) y Power BI (Pinda Guanolema, 2023).

Las plataformas de visualización de datos, como Tableau, Power BI y QlikView, se destacan por su capacidad de presentar información compleja de manera gráfica e interactiva, permitiendo a los gerentes y usuarios en general interpretar datos rápidamente y tomar decisiones informadas. Investigaciones concluyen que el uso de estas herramientas no solo optimiza los procesos empresariales, sino que también incrementa la competitividad organizacional al permitir una gestión más efectiva de la información. Además, se resalta la importancia de la visualización de datos como un componente crucial en la inteligencia de negocios, proporcionando una ventaja significativa en la era de la información (Aucancela Guamán, Viteri, & Castro, 2019).

Las técnicas más comunes incluyen también análisis exploratorio de datos (EDA) para describir las características principales de los datos, análisis predictivo con modelos como regresión lineal y árboles de decisión, y análisis correlacional mediante coeficiente de correlación de Pearson y análisis de componentes principales (PCA) (Tatsat et al., 2020).

El análisis financiero consiste en evaluar los datos financieros para tomar decisiones informadas sobre inversiones, adquisiciones y otras transacciones económicas. Implica

examinar los estados financieros como el estado de resultados, el balance general y el estado de flujo de efectivo para entender la situación financiera de una empresa u organización. Este tipo de análisis es crucial para tomar decisiones basadas en datos reales, ya que proporciona una visión detallada de la salud y el desempeño financiero de la empresa, siendo de ayuda para inversores, gerentes y otros interesados o entidades de control (Riera Riera, 2020).

Adicionalmente, (Riera Riera, 2020) también establece que la previsión financiera o forecasting es el proceso de hacer predicciones sobre tendencias y eventos financieros futuros basándose en datos financieros pasados y presentes. Esto implica el uso de modelos estadísticos y matemáticos para prever el rendimiento financiero futuro y tomar decisiones informadas sobre inversiones, presupuestos y planificación financiera. Además, la previsión financiera ayuda a las empresas a gestionar el flujo de efectivo y la liquidez al predecir entradas y salidas de efectivo futuras. Al revisar los datos financieros y realizar proyecciones, las compañías pueden identificar áreas que requieren mejoras en su desempeño financiero, observar oportunidades y riesgos potenciales y tomar decisiones estratégicas para alcanzar sus objetivos.

Python & GmbH una importante página web especializada en Python, señala que esta herramienta se puede utilizar para realizar diversos cálculos financieros básicos, como calcular rendimientos, tasas de descuento y precios de bonos. Además, indica que la biblioteca numpy es muy útil para calcular el valor presente neto (VPN) de una serie de flujos de efectivo, la tasa interna de retorno (TIR) y la tasa interna de retorno modificada (TIRM). Finalmente, indica que Python también es capaz de calcular ratios financieros comunes como el ratio de deuda a capital, el retorno sobre la inversión (ROI) y las ganancias por acción (EPS)(Python & GmbH, 2017).

Sin lugar a duda, Python es una de las herramientas que será utilizada en este proyecto para realizar análisis exploratorio, descriptivo y predictivo para poder entender y obtener insights de los estados financieros de la compañía seleccionada para el análisis.

Parth Kulkarni en su investigación sobre la gestión financiera basada en análisis avanzado indica que la integración de tecnologías de Inteligencia Artificial y Aprendizaje Automático en la Planificación y Análisis Financiero promete revolucionar la gestión de los estados financieros de las organizaciones. Sin embargo, su implementación enfrenta varios desafíos significativos como la calidad y la infraestructura de los datos, ya que los análisis avanzados dependen de datos precisos, completos y accesibles y no siempre todas las empresas tienen acceso a esto. Otro obstáculo que identificó es que las organizaciones deben invertir en desarrollar habilidades especializadas en análisis de datos, modelado estadístico y programación, lo que resalta la brecha de talento en el mercado laboral (Kulkarni, 2023).

Adicionalmente, la implementación de inteligencia artificial y machine learning debe cumplir con estrictas normativas regulatorias y éticas, asegurando la transparencia, la responsabilidad y la equidad en el uso de estas tecnologías. Además, la adopción de estas tecnologías puede enfrentar resistencia al cambio dentro de las organizaciones, lo que requiere estrategias de gestión del cambio efectivas y una comunicación clara para fomentar una cultura impulsada por los datos (Kulkarni, 2023).

En una investigación realizada por Axel Rodríguez en 2023, cuyo objetivo fue mejorar la toma de decisiones empresariales mediante la implementación de un sistema de inteligencia de negocios, se optó por utilizar Microsoft Power BI para automatizar el proceso de consolidación y transformación de datos financieros, permitiendo el establecimiento de indicadores clave, como los ratios financieros, esto contribuyó a proporcionar información precisa y en tiempo real a la alta gerencia para facilitar decisiones estratégicas. (González Rodríguez, 2023)

El procedimiento descrito en la investigación incluye varias fases, comenzando con la revisión y análisis de la información actual y el proceso de consolidación de datos. Posteriormente, el autor realizó el modelado de datos utilizando SQL Server y finalmente conectó Power BI con la base de datos para diseñar el dashboard final. Los resultados incluyeron una reducción significativa en el tiempo y los errores asociados con la

consolidación manual de datos financieros, así como la mejora en la precisión y la eficiencia de las decisiones gerenciales (González Rodríguez, 2023).

En el desarrollo de este proyecto, se utilizará también esta herramienta de Power BI para contribuir a que la organización tenga una ventaja competitiva al optimizar el proceso de toma de decisiones basados en datos financieros precisos y actualizados por medio de un dashboard interactivo fácil de entender y que arroje datos que se conviertan en información y posteriormente en insights que generen valor a la compañía.

En un estudio realizado por Omer Berat, Mehmet Ugur y Ahmet Murat en el que analizaron el uso de técnicas de deep learning en la predicción de series temporales financieras, se determinó que aunque existen numerosos estudios sobre la predicción financiera con Machine Learning, hay una carencia de revisiones exhaustivas específicamente sobre DL en este ámbito (Berat Sezer, 2020).

Por otro lado, el estudio se centró en la predicción de series temporales financieras utilizando deep learning y los autores identificaron varias técnicas, como las redes neuronales convolucionales, las redes neuronales recurrentes y las memorias de corto y largo plazo, que han demostrado ser más eficaces para manejar las complejidades inherentes a los datos financieros, mejorando la precisión y la capacidad de predicción en comparación con los enfoques tradicionales (Berat Sezer, 2020).

Las técnicas de machine learning como la regresión lineal, las máquinas de soporte vectorial y las redes neuronales recurrentes son de bastante utilidad para predecir los precios de las acciones. Así lo demostraron Patel, Shah y Thakkar en su investigación realizada en 2020, en donde utilizaron datos históricos de precios de apertura, cierre, máximo, mínimo y volúmenes de transacciones para obtener una mayor precisión en la predicción de precios de acciones en comparación con la regresión lineal y las Máquinas de Soporte Vectorial, siendo especialmente eficaces para capturar relaciones a largo plazo en series temporales de datos (Patel, Shah, & Thakkar, 2020).

En otro estudio similar, en donde el autor Thomas Fischer realizó una investigación que incluyó el uso de datos históricos financieros para entrenar modelos de memoria a corto y largo plazo y los comparó con métodos tradicionales de predicción financiera, demostró que estos modelos de Deep Learning superaron a los enfoques convencionales en términos de precisión y capacidad para manejar la naturaleza no lineal y compleja de los datos financieros debido a su capacidad para capturar dependencias a largo plazo en series temporales financieras (Fischer, 2020).

Se pudo observar que el autor recopiló datos históricos de precios de acciones, que incluyeron precios de apertura, cierre, máximo, mínimo y volumen de transacciones. Estos datos se dividieron en conjuntos de entrenamiento y prueba, se desarrollaron y entrenaron varios modelos, incluidos LSTM, regresión lineal, SVM (máquinas de vectores de soporte) y modelos ARIMA (Fischer, 2020).

Este estudio respalda la adopción de técnicas de deep learning en el campo de la predicción financiera, sugiriendo que pueden ofrecer ventajas significativas sobre los métodos tradicionales, además de tener la capacidad de predecir con mayor precisión los movimientos del mercado financiero puede llevar a la implementación de estrategias de inversión más efectivas.

Se encontró una investigación adicional que tuvo como objetivo realizar un análisis exploratorio de datos en los precios históricos de las acciones de Facebook, mediante el uso de Python. Aquí, se analizaron datos del 2018 al 2019 para identificar patrones estacionales, analizar la distribución de precios y detectar outliers. Se utilizaron gráficos de líneas, histogramas y diagramas de dispersión para visualizar las relaciones entre diferentes variables financieras (Khare, 2019).

Los resultados revelaron patrones estacionales claros y distribuciones de precios que no seguían una distribución normal, con presencia de outliers significativos. Además, se identificaron correlaciones entre los precios de diferentes acciones, lo que permitió una mejor comprensión del comportamiento de las mismas. Estas observaciones son

cruciales para optimizar estrategias de inversión y mejorar la toma de decisiones basada en datos detallados. Este enfoque subraya la importancia del análisis exploratorio de datos en la gestión financiera y la inversión (Khare, 2019).

Es posible tomar varias técnicas aplicadas en la investigación de Khare en este proyecto, por ejemplo, los Histogramas y Box Plots permitirán analizar la distribución de ingresos, costos y márgenes de beneficio y así identificar patrones y outliers en las cuentas del estado de resultados. Los Gráficos de Líneas permitirán observar tendencias temporales en los estados financieros, como variaciones mensuales o trimestrales en ingresos y gastos. Además, el cálculo de ratios en Python proporcionará una visión clave como el margen de beneficio, el ratio de liquidez y el ratio de endeudamiento que permitirá conocer a los usuarios e interesados evaluar la salud financiera de la compañía.

Para concluir, la literatura revisada subraya la necesidad de integrar tecnologías avanzadas y enfoques analíticos robustos de Business Intelligence para potenciar la gestión financiera y la toma de decisiones estratégicas en las empresas a través de soluciones de análisis exploratorio de datos, machine learning y visualización de datos.

IDENTIFICACIÓN DEL OBJETO DE ESTUDIO

A lo largo de los años, la gestión eficiente de los estados financieros ha sido esencial para la supervivencia y el crecimiento de cualquier empresa. Esto es especialmente cierto en sectores competitivos como el de alimentos y bebidas, donde la toma de decisiones estratégicas y operativas depende en gran medida de la precisión y la accesibilidad de los datos financieros. (Elliott & Elliott, 2019).

El objeto de estudio de este proyecto es mejorar la gestión y el análisis de los estados financieros de una de las empresas más importantes de alimentos y bebidas. Se busca desarrollar un modelo de análisis de datos que incluya análisis exploratorio, y modelos predictivos como regresión RIDGE y series temporales con ARIMA, además del uso de SQL para almacenamiento y gestión de datos que no solo ayude a tomar decisiones estratégicas más informadas, sino que también haga que las operaciones diarias sean más eficientes.

Una gestión eficiente de los estados financieros es crucial para la sostenibilidad y el crecimiento de la empresa. A pesar de ser una empresa líder en su sector, enfrenta ciertos desafíos significativos en su gestión financiera debido a la falta de herramientas adecuadas que soporten el volumen de información que maneja y la complejidad de los datos financieros, por ello, este proyecto pretende proporcionar herramientas avanzadas que mejoren significativamente la precisión y la eficiencia en la toma de decisiones financieras.

La principal problemática de la compañía es la falta de herramientas robustas para el procesamiento y gestión eficiente de grandes volúmenes de datos financieros, lo que resulta en demoras y decisiones basadas en datos incompletos o incorrectos. Además, la empresa está interesada en crear una herramienta de reporting que proporcione información financiera clara, entendible y accesible para todos los usuarios. La información financiera que se genera actualmente no es fácilmente interpretable, lo que

dificulta la comprensión y el uso de estos datos por parte de los diferentes niveles de la organización, auditores y organismos de control.

Adicionalmente, las investigaciones previas respaldan el uso de un enfoque analítico avanzado para abordar estos problemas. El uso de técnicas de análisis de datos y visualización permitirá procesar y analizar grandes volúmenes de datos de manera eficiente, proporcionando insights precisos y oportunos para la toma de decisiones.

Estas herramientas permiten manejar y procesar grandes volúmenes de datos de manera eficiente, automatizando y optimizando los procesos de gestión financiera. Con Python, se pueden desarrollar modelos de análisis y predicción más precisos, considerando múltiples variables y patrones históricos, lo que mejorará significativamente la exactitud de las proyecciones financieras (Khare, 2019).

Power BI facilitará la creación de dashboards y reportes interactivos que proporcionen una visión clara y accesible de los datos financieros, permitiendo a todos los niveles de la organización comprender y utilizar la información de manera efectiva (González Rodríguez, 2023). De este modo, no solo mejorará la eficiencia operativa, sino que también permitirá a la compañía tomar decisiones más informadas y estratégicas, optimizando la asignación de recursos y mejorando su capacidad para adaptarse a las demandas del mercado.

PLANTEAMIENTO DEL PROBLEMA

La compañía objeto del análisis enfrenta desafíos significativos en la gestión de sus estados financieros debido a la falta de herramientas adecuadas para el procesamiento y análisis de grandes volúmenes de datos financieros. Las herramientas tradicionales, como Excel, que se utilizan actualmente, no son capaces de manejar eficientemente el volumen de datos financieros de la empresa, lo que resulta en demoras y procesos ineficientes.

Además, la empresa carece de una herramienta de reporting dinámica que proporcione información financiera clara y accesible para todos los usuarios, lo que dificulta la comprensión y el uso de los datos financieros en la toma de decisiones estratégicas y operativas. Finalmente, las proyecciones financieras se basan en cálculos históricos y experiencia del usuario, lo que no siempre refleja con precisión la realidad financiera y afecta negativamente la precisión del dynamic forecast y la presentación de presupuestos a la casa matriz.

Se propone que el proyecto se llevará a cabo en varias fases:

Recolección de Datos:

- Recopilación detallada de datos financieros de las cuentas de ingresos y gastos de 2021 a 2023, incluyendo ventas, costos de producción, gastos de marketing y otros gastos generales asociados a cada unidad de negocio.
- Extraer datos de ventas, costos de producción, gastos de marketing, y otros gastos generales asociados a cada producto.

Análisis Exploratorio de Datos:

- Utilizando Python para limpiar y explorar los datos, identificando patrones y anomalías.

- Asegurando que los datos son completos y están en el formato adecuado.

Desarrollo de Modelos Predictivos:

- Desarrollo e implementación de modelos predictivos utilizando técnicas de machine learning, como modelos predictivos como regresión RIDGE y series temporales con ARIMA, además del uso de SQL para almacenamiento y gestión de datos. Estos modelos serán validados utilizando técnicas de validación cruzada y métricas de rendimiento como RMSE y MAE.
- Utilizar técnicas como regresión lineal, árboles de decisión, o métodos ensemble para desarrollar un modelo que prediga la rentabilidad futura.

Visualización y Reporting:

- Implementación de Power BI para crear dashboards y reportes interactivos que faciliten la toma de decisiones.
- Incorporar visualizaciones de datos en tiempo real y controles interactivos para explorar diferentes escenarios y simulaciones basadas en las predicciones del modelo.

Validación y Ajustes:

- Validación de los modelos y ajustes necesarios para asegurar la precisión y utilidad de las proyecciones y reportes.
- Realizar pruebas del sistema completo (modelo y dashboard) para asegurar que los datos se actualizan correctamente y que las predicciones son visualizadas de manera efectiva.
- Recoger feedback de usuarios potenciales para realizar ajustes necesarios.

Documentación y Presentación Final:

- Documentar todo el proceso de desarrollo, incluyendo el código, las decisiones de diseño, y los resultados de las pruebas.
- Preparar una presentación final que resuma el proyecto, los métodos utilizados, y demuestre cómo el dashboard puede ser utilizado para mejorar la toma de decisiones.

OBJETIVO GENERAL

Desarrollar e implementar un modelo avanzado de análisis exploratorio y predictivo de datos para optimizar la gestión de los estados financieros de una de las principales compañías de alimentos y bebidas en Ecuador, mejorando la toma de decisiones estratégicas y operativas mediante la visualización clara y precisa de información financiera utilizando herramientas de análisis exploratorio, predictivo y Power BI.

OBJETIVOS ESPECÍFICOS

1. Implementar una solución de análisis exploratorio de datos eficiente

Desarrollar un sistema eficiente de análisis exploratorio de datos que optimice la precisión y velocidad de los reportes financieros.

2. Desarrollar modelos predictivos avanzados

Utilizar técnicas de inteligencia artificial para crear modelos predictivos que mejoren la exactitud de las proyecciones financieras, considerando múltiples

variables y patrones históricos, con el fin de mejorar el dynamic forecast y la presentación de presupuestos a la casa matriz.

3. Crear dashboards interactivos y reportes financieros claros

Utilizar Power BI para diseñar y desarrollar dashboards y reportes interactivos que presenten la información financiera de manera clara y accesible, facilitando la comprensión y el uso de los datos por parte de todos los niveles de la organización, auditores y organismos de control.

Se espera que el proyecto resulte en una mejora significativa en la gestión de los estados financieros de la compañía, proporcionando herramientas avanzadas para el análisis, predicción y visualización de datos financieros. Esto permitirá a la empresa tomar decisiones más informadas, optimizar sus recursos y mantener su posición de liderazgo en el mercado.

JUSTIFICACIÓN Y APLICACIÓN DE LA METODOLOGÍA

Para llevar a cabo el análisis de optimización de la gestión de estados financieros en una empresa de alimentos y bebidas en Ecuador, se ha utilizado una base de datos integral extraída del sistema ERP SAP. Esta base contiene registros detallados de las cuentas contables, incluyendo el Estado de Pérdidas y Ganancias y el Estado de Situación Financiera compuesto por 115.187 filas y 124 columnas. Además, la base de datos proporciona información cualitativa como Unidad de Negocio, año, mes, y producto.

La elección de esta base de datos se debe a su relevancia directa y exhaustividad, asegurando que el análisis refleje precisamente la realidad financiera de la empresa. Además, no solo provee una visión integral de la situación financiera de la empresa, sino que también permite un análisis preciso y relevante gracias a la calidad y el nivel de detalle de los datos (Elliott & Elliott, 2019). Esta elección es muy importante para el desarrollo de un sistema de análisis eficiente que mejore la exactitud de las proyecciones financieras y optimice la gestión de los estados financieros.

En el proyecto, el proceso de limpieza, pre-procesamiento y transformación de datos fue fundamental para asegurar la calidad y la precisión del análisis posterior, e incluye varias acciones clave:

- Estandarización y traducción de variables: Estandarización de nombres y traducción al español.
- Manejo de valores faltantes: Identificación y trato de los valores faltantes, ya sea imputándolos con métodos estadísticos como la media o la mediana de la columna, o eliminando registros cuando la falta de información sea crítica y no pueda estimarse de manera confiable.
- Corrección de errores de formato: Normalización de formatos de fechas y cifras monetarias para asegurar consistencia en todo el conjunto de datos.

- Validación de datos: Verificación de la exactitud de los registros comparándolos con fuentes de datos estándar o históricos de la empresa para identificar posibles errores de ingreso o discrepancias.
- Creación de variables derivadas: Generación de nuevas variables que puedan tener más significado para el análisis, como ratios financieros (por ejemplo, razón corriente, margen operativo), que son indicadores clave de rendimiento y salud financiera (Aucancela Guamán, Viteri, & Castro, 2019).
- Descomposición de series temporales: Separar las series temporales en componentes estacionales, tendenciales y residuales, para mejorar el análisis exploratorio, modelos ARIMA y otros análisis de series temporales.
- Agrupación y segmentación: Agrupación de los datos en categorías lógicas basadas en criterios financieros, como tipo de gasto o ingreso, para análisis específicos y detallados por segmentos.

Las variables descritas a continuación fueron obtenidas del sistema ERP SAP de la empresa en análisis:

Tabla 1. Matriz de descripción de variables
Fuente: Peña & Ayerve

Variable	Naturaleza	Descripción
Unidad_de_Negocio	Categórica	Identifica el grupo o unidad de negocio responsable de las operaciones registradas.
Anio	Temporal	Corresponde al año fiscal en que se registraron los datos.
Mes	Temporal	Indica el mes del período fiscal al que corresponden los datos.
Ventas_Brutas (Dependiente)	Cuantitativa	Ingresos totales generados por ventas antes de cualquier deducción.
Descuentos_distribuidores_desempeño	Cuantitativa	Subsidios otorgados a distribuidores como parte de incentivos de rendimiento comercial.
Ventas_Netas	Cuantitativa	Ingresos por ventas después de deducir descuentos y devoluciones.
Costo_de_Ventas	Cuantitativa	Costos asociados directamente a la producción de bienes vendidos.
Gastos_de_Marketing	Cuantitativa	Total de gastos incurridos en actividades de marketing y promoción.
Gastos_Fijos_de_Fabrica	Cuantitativa	Gastos fijos incurridos por las operaciones de la fábrica, independientes del nivel de producción.

Gastos_Generales_Fijos	Cuantitativa	Gastos fijos generales no directamente atribuibles a la producción o venta de bienes.
Resultado_Operacional	Cuantitativa	Utilidad operativa consolidada, calculada como la diferencia entre ingresos operacionales y gastos operacionales.
Medios Digitales	Cuantitativa	Inversión en marketing digital, incluyendo publicidad en plataformas online y redes sociales.
Deudas Malas	Cuantitativa	Valor de las cuentas por cobrar consideradas irre recuperables o con baja probabilidad de pago.
Materia Prima	Cuantitativa	Costo de las materias primas utilizadas en la producción de los bienes vendidos.

La elección de las variables para este análisis se ha centrado en capturar una imagen integral y precisa del desempeño financiero y operativo de la empresa. Las variables seleccionadas, que incluyen indicadores de ingresos, costos, gastos y rendimiento operacional, son fundamentales para evaluar la salud financiera y la eficiencia de la gestión (García, 2020).

Las variables de ventas brutas y netas, junto con los costos de ventas y materia prima, ofrecen una perspectiva directa sobre la rentabilidad y la capacidad de la empresa para generar ingresos en relación con sus costos de producción. Los gastos de marketing y medios digitales, por su parte, reflejan la inversión en crecimiento y adquisición de clientes, aspectos cruciales para la expansión y la competitividad en el mercado (Burguete, 2016).

Los gastos fijos, tanto de fábrica como generales, y el resultado operacional son indicativos de cómo la empresa maneja su estructura de costos y su eficiencia operativa (Pinda Guanolema, 2023). Estas variables son esenciales para identificar áreas potenciales de mejora y para garantizar que los recursos se utilicen de manera óptima.

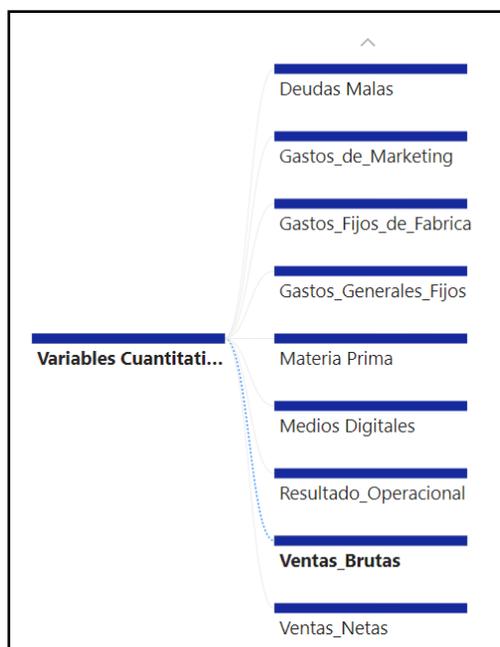


Figura 1: Variables Cuantitativas
Fuente: Peña, Ayerve, 2024

Variables como la cartera dudosa son vitales para evaluar la eficacia de las políticas de crédito y cobranza de la empresa, proporcionando insights clave para la gestión del riesgo y la optimización de la liquidez (García, 2020).

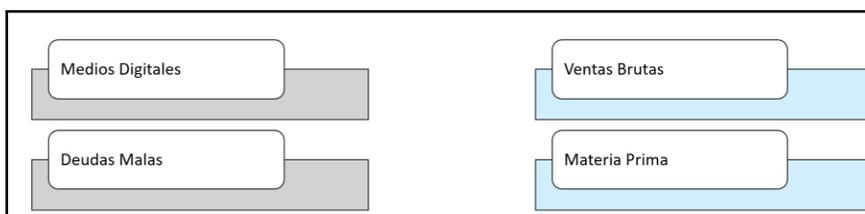


Figura 2: Variables Cuantitativas Predictoras
Fuente: Peña, Ayerve, 2024

La inclusión de variables temporales y descriptivas como la unidad de negocio, año y mes permite realizar análisis detallados y contextualizados que son esenciales para la planificación estratégica y la toma de decisiones basada en datos. Estas variables facilitan el seguimiento de la evolución del negocio a lo largo del tiempo y la comparación entre diferentes segmentos o unidades operativas.

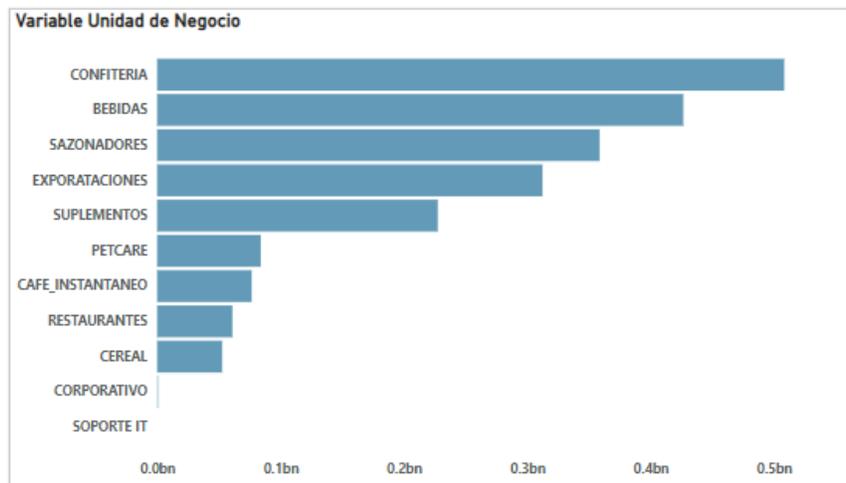


Figura 3: Variables Cualitativas
Fuente: Peña, Ayerve, 2024

A continuación, se presenta una explicación breve de los pasos ejecutados en el análisis exploratorio:

- Importación de las librerías numpy, pandas, matplotlib.pyplot, y seaborn.
- Configuración del entorno para ignorar advertencias generadas durante la ejecución.
- Carga del dataset con datos financieros utilizando pandas.
- Se renombraron columnas para clarificar su significado, como cambiar nombres de indicadores financieros a términos más comprensibles y en español.
- Selección de datos específicos de ventas, marketing y descuentos para calcular su correlación.
- Visualización de la correlación mediante un diagrama de dispersión y un mapa de calor para facilitar la interpretación.
- Exploración de las correlaciones entre ventas netas y descuentos por desempeño entre 2020 y 2023, mostrando los resultados en un gráfico de barras.
- El método describe() fue aplicado para obtener un resumen estadístico de ciertas variables financieras.
- Los datos fueron organizados por unidad de negocio y año para sumar las ventas netas, las cuales se visualizaron en un gráfico de tendencia.

- A partir de las ventas netas y el costo de ventas, se calcularon el margen bruto y el margen neto, así como el resultado operacional.
- Se calcularon y visualizaron medias de rentabilidad y ratios específicos por unidad de negocio, como los gastos de marketing y generales sobre las ventas netas.
- Para observar la distribución de las ventas netas se empleó un histograma, y un boxplot permitió comparar estas ventas entre distintas unidades de negocio.
- Una matriz de correlación entre varias variables financieras fue creada y visualizada para explorar más relaciones entre ellas.

En la fase de análisis predictivo de este proyecto, se implementaron técnicas de modelado estadístico para pronosticar ciertas variables de diferentes unidades de negocio. Se utilizó la regresión RIDGE, un método de regresión regularizada que ayuda a manejar la colinealidad entre las variables predictoras y a mejorar la estabilidad del modelo (Zou & Hastie, 2020).

Paralelamente, se aplicaron modelos de series temporales ARIMA para capturar las tendencias y patrones estacionales inherentes a los datos de ventas a lo largo del tiempo. Este enfoque dual permitió ajustar modelos robustos que reflejan las complejidades de los datos financieros y también proporcionar predicciones precisas y confiables que son esenciales para la planificación estratégica y la toma de decisiones en la empresa (Box, Jenkins, & Reinsel, 2021).

La elección de las variables Medios Pagados Digitales Puros, Cartera Dudosa, Ventas Brutas, y Costo de Material Prima como foco de la predicción para 2024 es estratégicamente significativa. Estas variables reflejan aspectos importantes del rendimiento financiero y operativo de la empresa y también contribuyen con insights valiosos sobre la eficiencia de la inversión en marketing, la salud del crédito, y la dinámica de la rentabilidad de los productos.

Al predecir estas variables mediante el uso de regresión RIDGE y ARIMA, se aborda directamente la necesidad de anticipar tendencias en áreas que pueden impactar directamente en la rentabilidad y eficiencia de costos. Esta aproximación no solo mejora

la precisión en la planificación y el manejo de recursos, sino que también fortalece la capacidad de la empresa para ajustar sus estrategias en respuesta a cambios anticipados en el mercado y en su estructura de costos (Box, Jenkins, & Reinsel, 2021).

Para los modelos de regresión RIDGE, las variables objetivo son 'Medios Pagados "Digitales Puros" y "Cartera Dudosa", mientras que para los modelos ARIMA, se focalizan en "Ventas Brutas" y "Costo Material Prima". Esta diferenciación asegura que cada modelo se ajuste y optimice para predecir los aspectos más críticos de la actividad empresarial según sus patrones específicos (Box, Jenkins, & Reinsel, 2021).

Sin embargo, es importante resaltar que para este proyecto se escogieron las variables mencionadas anteriormente a manera de piloto o modelo, pero en la aplicación empresarial estas variables pueden ser reemplazadas con cualquier otra variable que sea necesario proyectar. Se recomienda que la variable predictora o dependiente sea Ventas Brutas.

La regresión Ridge añade un término de penalización a la regresión lineal para evitar el sobreajuste (Pedregosa, 2020), a continuación, la ecuación matemática:

$$\text{Minimizar } \frac{1}{n} \sum_{i=1}^n (y_i - x_i w)^2 + \lambda |w|^2$$

Donde:

(y_i) es el valor objetivo.

(x_i) es el vector de características.

(w) es el vector de coeficientes.

(λ) es el parámetro de regularización (Box, Jenkins, & Reinsel, 2021).

El proceso comenzó definiendo los parámetros utilizando `GridSearchCV` con varios solucionadores y un rango para el parámetro de regularización `Alpha` con la finalidad de optimizar el rendimiento del modelo a través de un proceso de búsqueda aleatoria con validación cruzada, especificada por `KFold`, que ayuda a garantizar que el modelo no se ajuste excesivamente a un subconjunto particular de los datos.

Posteriormente, con la función *entrenar_modelo*, se configuró la búsqueda aleatoria para encontrar el mejor modelo basado en el error cuadrático medio negativo. La elección de características se realizó mediante *SelectKBest* con la función *f_regression*, que elige las características más significativas para la regresión. Luego, los datos se dividieron en conjuntos de entrenamiento y prueba y se escalaron para normalizarlos, muy importante para el rendimiento de estos algoritmos de aprendizaje automático (Pedregosa, 2020).

Para cada variable objetivo especificada, se entrenó el modelo Ridge, se evaluó su rendimiento utilizando el *MSE* y el coeficiente de determinación *R2*. Finalmente, el código se preparó para realizar predicciones para el año 2024, utilizando los datos del último año fiscal disponible. Así, se logró reflejar el nuevo año fiscal en semestres, trimestres y meses.

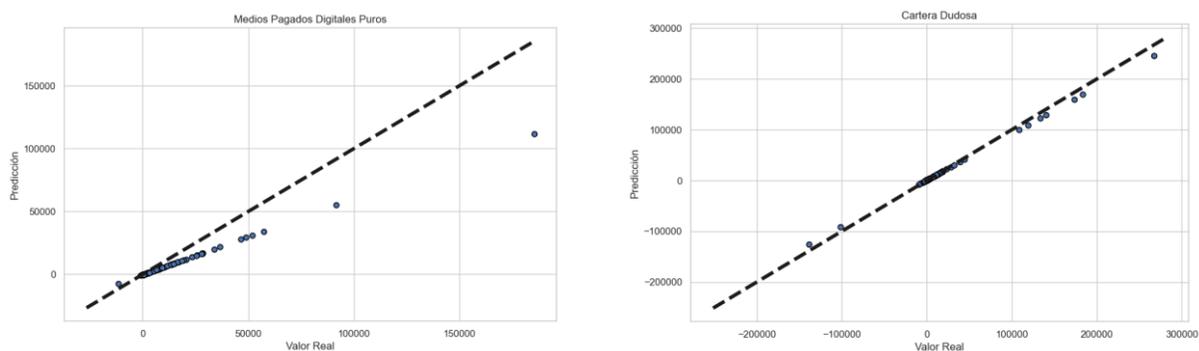


Figura 4: Pruebas de predicción
Fuente: Peña, Ayerve, 2024

En lo que respecta a predicción de series temporales, se enfocó en las variables "Ventas Brutas" y "Costo Material Prima Menos Margen" utilizando modelos ARIMA. Esta parte del análisis comenzó con la agrupación de los datos originales por unidad de negocio, año fiscal y periodo, sumando los valores para cada variable. A partir de estos datos agrupados, se busca generar predicciones utilizando el modelo ARIMA, configurado específicamente con un orden de (5,1,0) para cada serie de tiempo.

En este proceso, se identificaron combinaciones únicas de unidad de negocio para garantizar que las predicciones fueran específicas y relevantes para cada segmento.

Para cada combinación, se extrajeron las series temporales correspondientes y se ajustó el modelo. En caso de enfrentar problemas técnicos, como errores de descomposición, el modelo tiene la capacidad de capturar estos incidentes y continuar con el siguiente grupo de datos sin interrumpir el flujo de análisis.

Tras ajustar los modelos, se realizaron pronósticos para un periodo futuro de 12 meses. Además, se evaluaron las métricas de rendimiento de cada modelo, como el error cuadrático medio (ECM) y el error absoluto medio (MAE), para medir la precisión de las predicciones. Estas métricas se recopilaban para cada unidad de negocio, proporcionando una evaluación cuantitativa del rendimiento del modelo.

Finalmente, los resultados de las predicciones se organizaron en un formato estructurado que incluye detalles como el año fiscal, el periodo, el semestre, el trimestre, el mes y las predicciones de ventas y costos. Este conjunto de datos de predicciones se preparó para su análisis posterior o para informar decisiones estratégicas, destacando la capacidad de los modelos ARIMA para informar sobre tendencias futuras y facilitar una planificación empresarial efectiva basada en datos históricos y proyectados (Guerra Rodríguez, 2022).

El modelo ARIMA combina estos tres componentes:

1. Componente Autoregresiva (AR):

$$AR(p): y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \epsilon_t$$

Donde : $\phi_1, \phi_2, \dots, \phi_p$ son los coeficientes autoregresivos y ϵ_t es el error blanco.

2. Componente de Diferenciación (I):

$$I(d): \Delta^d y_t = (1 - \beta)^d y_t$$

Donde: (B) es el operador de retraso y (d) es el orden de diferenciación

3. Componente de Media Móvil (MA):

$$MA(q): y_t = \epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \dots + \theta_q\epsilon_{t-q}$$

Donde : $\phi_1, \phi_2, \dots, \phi_q$ son los coeficientes de la media móvil y ϵ_t es el error blanco.

Modelo ARIMA Combinado

El modelo ARIMA combina estos tres componentes:

$$\begin{aligned} ARIMA(p, d, q): \Delta^d y_t \\ = \phi_1 \Delta^d y_{t-1} + \phi_2 \Delta^d y_{t-2} + \dots + \phi_p \Delta^d y_{t-p} + \epsilon_t \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots \\ + \theta_q \epsilon_{t-q} \end{aligned}$$

Donde:

(p) es el orden de la parte autorregresiva (AR).

(d) es el grado de diferenciación.

(q) es el orden de la media móvil (MA). (Guerra Rodríguez, 2022)

En la siguiente etapa se realizó una evaluación del rendimiento de los modelos para asegurar la precisión y la aplicabilidad de los mismos, además para una visualización efectiva y una interpretación intuitiva de los resultados, se implementaron scripts de graficación en Python, diseñados para mostrar las predicciones de cada variable a lo largo de los meses del año 2024.

Adicionalmente, el modelo permite exportar un archivo en formato Excel para realizar análisis adicionales o importarlo en herramientas de visualización de datos como Power Bi.

RESULTADOS

Análisis Exploratorio

El proceso de análisis exploratorio comenzó con la carga de la base de datos extraída del sistema SAP en un entorno de programación Python. Este paso permitió asegurar el acceso y manipulación eficiente de los datos, utilizando librerías como pandas para manejar los datasets y numpy para operaciones numéricas. En la figura 4 se presentan las variables de interés utilizadas en el análisis exploratorio de datos.

	Unidad	Año	Mes	Ventas_Brutas	Descuentos_distribuidores_desempeño	Ventas_Netas	Costo_de_Ventas	Gastos_de_Marketing	Gastos_Fijos_de_Fabrica	Gastos_Generales_Fijos	Resultado_Operacional	
0	CAFÉ_INSTANTANEO	2020	1	1337136.73		236330.93	1100805.80	371093.14	187157.98	291157.71	185475.08	236354.02
1	CAFÉ_INSTANTANEO	2020	2	1438026.67		249616.26	1188410.41	376470.96	209100.89	345457.74	259193.82	189103.15
2	CAFÉ_INSTANTANEO	2020	3	1306355.82		188973.35	1117382.47	328941.94	195495.30	305629.41	232938.08	234434.87
3	CAFÉ_INSTANTANEO	2020	4	1086097.58		178651.17	907446.41	292876.22	38507.79	119566.91	248728.21	231850.79
4	CAFÉ_INSTANTANEO	2020	5	1407225.31		194450.30	1212775.01	331810.17	72691.75	223728.53	221298.96	414033.53
...
511	PETCARE	2023	8	1965654.40		472721.19	1492933.21	598374.32	95808.12	219394.95	208932.56	446713.04
512	PETCARE	2023	9	1975140.59		455634.26	1519506.33	558194.49	173711.62	318134.68	215962.44	408119.24
513	PETCARE	2023	10	1722321.09		341472.44	1380848.65	533213.59	98421.77	254539.69	191113.95	385070.63
514	PETCARE	2023	11	1931470.03		389975.49	1541494.54	606598.04	62496.04	202479.87	229845.67	483707.24
515	PETCARE	2023	12	2311256.09		421578.58	1889677.51	683370.70	24008.99	90908.19	222706.79	870242.29

Figura 5: Visualización del dataset en Python
Fuente: Peña, Ayerve, 2024

Inicialmente, se realizó un análisis exploratorio de los datos para identificar patrones, tendencias y posibles irregularidades en las finanzas de la empresa. Este análisis se llevó a cabo empleando técnicas estadísticas avanzadas, facilitando una comprensión profunda de los datos (Gorakala & Dangeti, 2019).

Se realizó un análisis de correlación de Pearson que permite medir la relación entre dos variables continuas (Paz, 2022). Inicialmente, se realizó un análisis exploratorio de los datos para identificar patrones, tendencias y posibles irregularidades en las finanzas de la empresa. Este análisis se llevó a cabo empleando técnicas estadísticas, facilitando una comprensión profunda de los datos subyacentes.

El análisis estadístico de los datos financieros de la empresa muestra una significativa variabilidad en las Ventas Brutas, Costo de Ventas y Gastos Generales Fijos. Las Ventas Brutas promedian \$396k, con una desviación estándar de \$368k, y un rango que va desde \$0 hasta un máximo de \$1,613k, indicando períodos de alta y baja actividad de ventas.

	Ventas_Brutas	Costo_de_Ventas	Gastos_Generales_Fijos
count	516.00	516.00	516.00
mean	3961141.46	1784469.56	426022.22
std	3686040.01	1984648.73	463459.38
min	0.00	-4343.04	-375506.42
25%	1132347.57	345694.40	148065.96
50%	1828377.41	660509.91	216299.57
75%	7047768.96	3054097.51	627109.94
max	16313758.27	7786482.21	3143521.39

Figura 6: Correlación general de las variables
Fuente: Peña, Ayerve, 2024

En primera instancia, se ejecutó el script de correlación entre todas las variables, sin embargo, al ser un análisis financiero, se procedió a escoger ciertas variables de interés, para conocer la relación entre ellas.

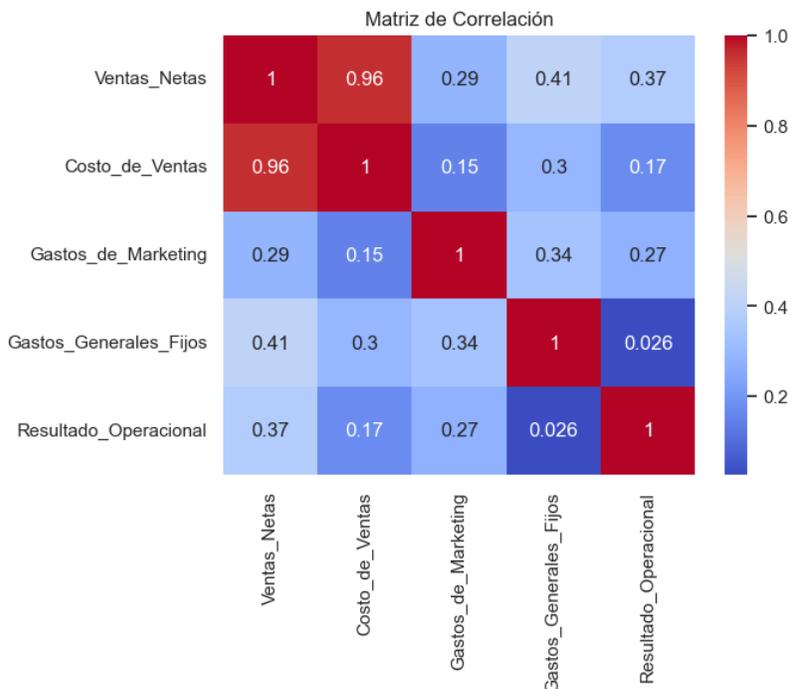


Figura 7: Correlación general de las variables
Fuente: Peña, Ayerve, 2024

El análisis presentado en la figura 8 muestra que la correlación positiva indica que a medida que los gastos de marketing aumentan, también lo hacen las ventas, pero el grado de esta relación es débil. Esto puede significar que, aunque los gastos de marketing influyen en las ventas, otros factores también están jugando un papel importante en las ventas totales.

Por otro lado, la correlación de 0.72 presentado en la figura 9, indica que aumentar los gastos en marketing podría estar relacionado con mayores descuentos por desempeño, sugiriendo que las inversiones en marketing pueden estar impulsando el desempeño de los distribuidores.



Figura 8: Correlación entre Gastos de Marketing y Ventas por año
Fuente: Peña, Ayerve, 2024

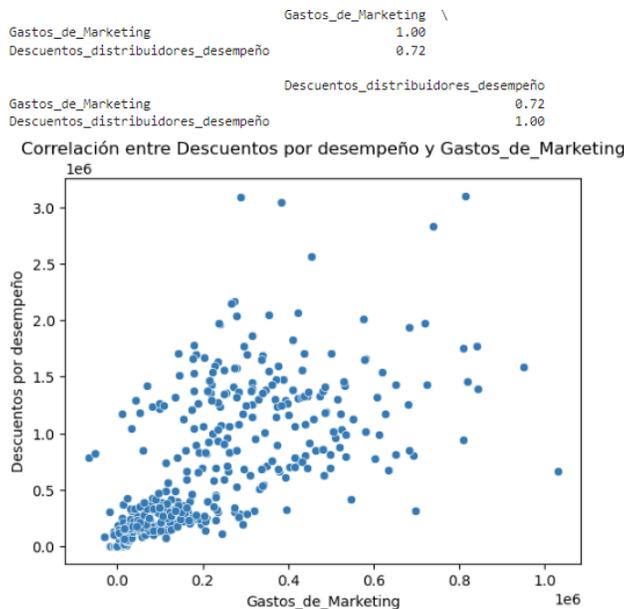


Figura 9: Correlación entre Descuentos por desempeño y Gastos_de_Marketing
Fuente: Peña, Ayerve, 2024

Posteriormente, se realizaron análisis de tendencias y comparación, la gráfica (figura 10 y 11) muestra una variabilidad notable en las ventas netas entre distintas unidades de negocio desde principios de 2020 hasta finales de 2023. Algunas unidades, como “Confitería” y “Exportaciones”, destacan por un crecimiento sostenido y significativo en ventas netas, evidenciando picos pronunciados y un ascenso constante que podría reflejar estrategias comerciales efectivas o una mayor demanda de mercado.

Por otro lado, unidades como “Petcare” y “Restaurantes” exhiben un desempeño menos destacado, con ventas continuamente bajas durante el periodo analizado. Esto indica la necesidad de explorar posibles mejoras o de reconsiderar las estrategias actuales para estas áreas.

Además, la gráfica revela fluctuaciones que pueden estar relacionadas con factores estacionales o eventos de mercado específicos, indicados por los picos que ocurren alrededor de la misma época cada año en varias unidades. Este patrón estacional, junto con los cambios abruptos en las tendencias de ventas, podría estar influido por eventos externos como lanzamientos de productos o campañas de marketing importantes.

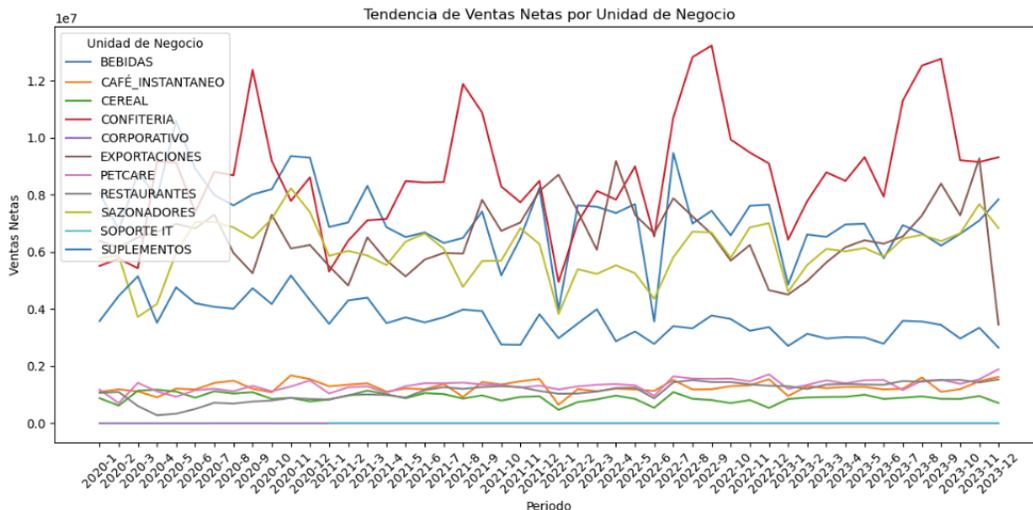


Figura 10: Tendencia de Ventas Netas por Unidad de Negocio
Fuente: Peña, Ayerve, 2024

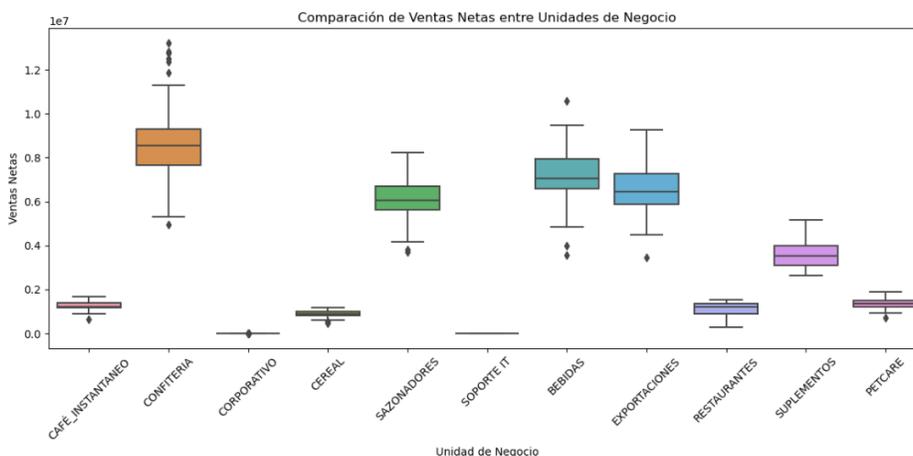


Figura 11: Comparación de Ventas Netas entre Unidades de Negocio
Fuente: Peña, Ayerve, 2024

La distribución de las ventas netas muestra un sesgo positivo, lo que implica que la mayoría de las observaciones se acumulan en los rangos más bajos, especialmente evidente en el pico del primer bin (0 a 2 millones), aunque hay ventas significativamente más altas en algunos casos. La mayoría de los datos se sitúan entre 0 y 2 millones, con menos frecuencia de observaciones a medida que las ventas netas aumentan.

Existen picos menores alrededor de los 6 millones, sugiriendo que ciertas unidades de negocio o periodos específicos experimentan ventas más elevadas en comparación con el resto, y la forma de la curva confirma que los datos están inclinados hacia la derecha.

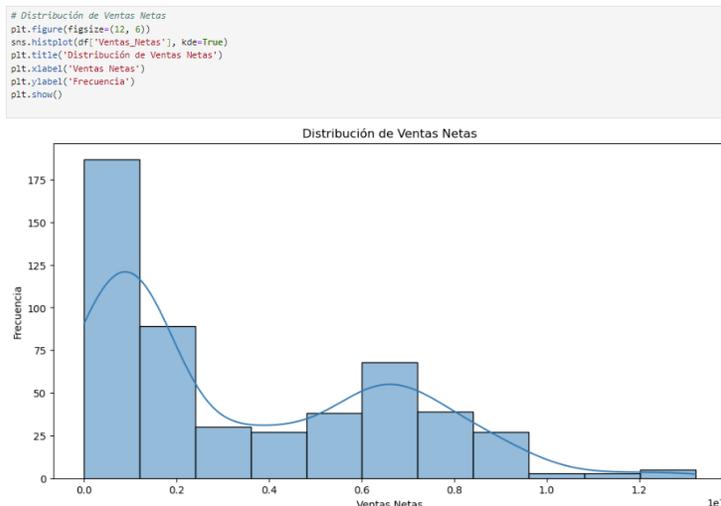


Figura 12: Distribución de Ventas Netas
Fuente: Peña, Ayerve, 2024

Finalmente, a través de scripts de Python se calcularon ratios financieros como el margen bruto y margen neto de las variables seleccionadas sobre las ventas para determinar la rentabilidad por Unidad de Negocio y conocer cuanto representa cada cuenta comparándolo con el retorno.

```
# Calcular Margen Bruto y Margen Neto
df['Margen_Bruto'] = df['Ventas_Netas'] - df['Costo_de_Ventas']
df['Margen_Neto'] = df['Resultado_Operacional'] / df['Ventas_Netas']

# Resumen de rentabilidad por Unidad de Negocio
rentabilidad = df.groupby('Unidad')[['Ventas_Netas', 'Resultado_Ope
print(rentabilidad)
```

Unidad	Ventas_Netas	Resultado_Operacional	Margen_Neto
BEBIDAS	7216800.06	567822.86	0.07
CAFÉ_INSTANTANEO	1266137.66	306715.69	0.24
CEREAL	892891.32	203370.74	0.21
CONFITERIA	8708842.31	1980849.83	0.22
CORPORATIVO	16.62	-302.07	NaN
EXPORTACIONES	6511017.73	0.00	0.00
PETCARE	1332281.56	431767.49	0.32
RESTAURANTES	1120096.86	224800.95	0.17
SAZONADORES	6054509.29	2183790.27	0.36
SOPORTE IT	0.00	-404015.26	-inf
SUPLEMENTOS	3608445.93	1131907.80	0.31

Figura 13: Resumen de rentabilidad
Fuente: Peña, Ayerve, 2024

```
# Calcular Ratios
df['Ratio_Gastos_Marketing'] = df['Gastos_de_Marketing'] / df['Ventas_Netas']
df['Ratio_Gastos_Generales'] = df['Gastos_Generales_Fijos'] / df['Ventas_Netas']

# Resumen de ratios por Unidad de Negocio
ratios = df.groupby('Unidad')[['Ratio_Gastos_Marketing', 'Ratio_Gastos_Generales']]
print(ratios)
```

Unidad	Ratio_Gastos_Marketing	Ratio_Gastos_Generales
BEBIDAS	0.04	0.14
CAFÉ_INSTANTANEO	0.11	0.17
CEREAL	0.12	0.18
CONFITERIA	0.04	0.18
CORPORATIVO	NaN	NaN
EXPORTACIONES	0.00	0.02
PETCARE	0.08	0.15
RESTAURANTES	0.03	0.19
SAZONADORES	0.06	0.12
SOPORTE IT	NaN	inf
SUPLEMENTOS	0.11	0.11

Figura 14: Resumen de ratios
Fuente: Peña, Ayerve, 2024

Figura 15: Predicción gasto en medios digitales (RIDGE)
Fuente: Peña, Ayerve, 2024

Los gastos en medios digitales se muestran relativamente estables para la mayoría de las unidades, con ligeras fluctuaciones. Por ejemplo, "Café Instantáneo" y "Bebidas" muestran incrementos graduales hasta julio. En Septiembre, Cereal se destaca claramente debido al pico mencionado anteriormente, mientras que otras unidades mantienen o incrementan ligeramente sus gastos habituales.

Posterior al pico de Septiembre, se proyecta que los gastos de "Cereal" disminuyan drásticamente, regresando a niveles similares a los de otras unidades. Las demás unidades muestran una tendencia decreciente o estabilización hacia el final del año, con "Confitería" y "Restaurantes" manteniendo gastos moderados.

Este análisis ayuda a entender dónde y cuándo se esperan inversión de recursos en marketing digital, lo que es crucial para evaluar también la efectividad de estas inversiones y ajustar las estrategias de gasto en función de los resultados deseados y el retorno sobre la inversión. Además, los picos inusuales pueden ser investigados para asegurar que los gastos están alineados con los objetivos estratégicos y las campañas planificadas.

El modelo Ridge para la variable "Medios Digitales" muestra un rendimiento excepcional. El ECM de 11.95 indica que las predicciones están muy cerca de los valores reales. El MAE de 1.34 sugiere que el error promedio absoluto es muy bajo. La puntuación R^2 de casi 1 (0.9999999698763895) indica que el modelo explica prácticamente el 100% de la variabilidad en los datos. Los mejores hiperparámetros encontrados fueron 'alpha': 0.01 y 'solver': 'svd'.

```
Fitting 10 folds for each of 42 candidates, totalling 420 fits
Rendimiento de Ridge para Medios_Pagados_Digitales_Puros:
Error Cuadrático Medio: 11.95098759029237
Error Absoluto Medio: 1.3423841608585023
Puntuación R^2: 0.9999999698763895
Mejores Hiperparámetros: {'alpha': 0.01, 'solver': 'svd'}
```

Respecto a la proyección de la cartera dudosa, las predicciones de Enero a Mayo para la mayoría de las unidades son relativamente bajas, con la mayoría de las unidades manteniendo predicciones debajo de los \$50 KUSD. Esto indica una gestión de crédito efectiva y pocos riesgos de crédito anticipados durante estos meses. Por otro lado, en Junio se muestra un pico extremadamente alto para la unidad "Bebidas", con una predicción que supera las \$200 KUSD. Este podría ser un indicativo de posibles problemas de cobranza o un incremento en las ventas a crédito no aseguradas durante este período.

Después del pico en Junio, las predicciones vuelven a niveles más manejables, aunque "Bebidas" muestra un pico significativo en Noviembre. Los valores negativos observados sugieren posibles recuperaciones de deudas previamente clasificadas como dudosas.

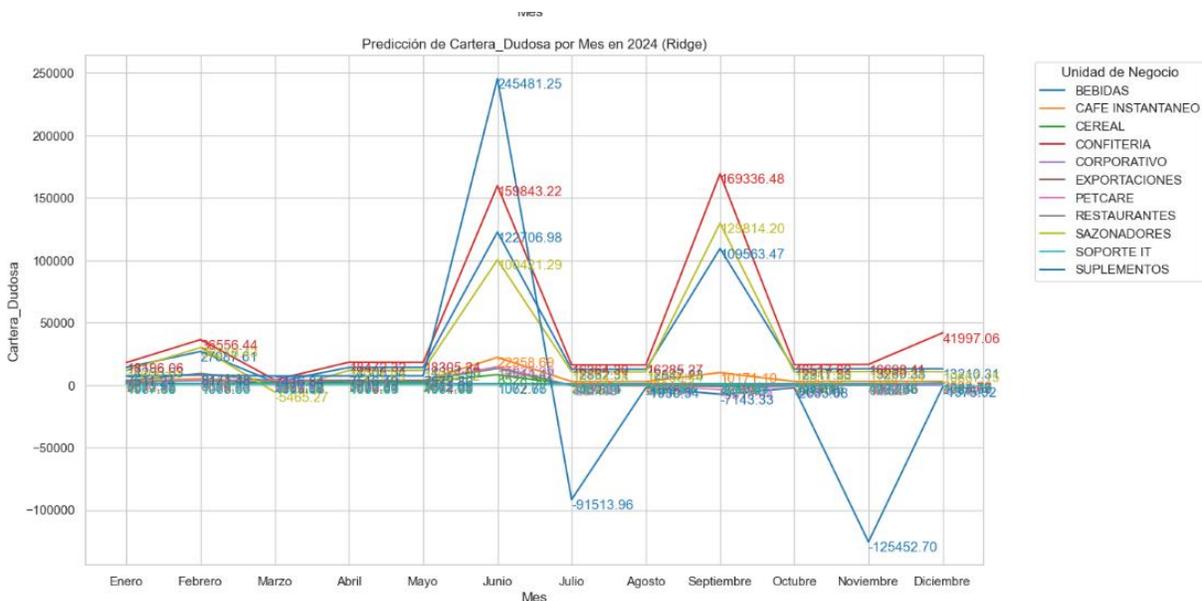


Figura 16: Predicción Cartera Dudosa (Ridge)
Fuente: Peña, Ayerve, 2024

El modelo Ridge para la variable "Cartera Dudosa" también muestra un rendimiento excepcional. El ECM de 349.50 indica que las predicciones están muy cerca de los valores reales. El MAE de 7.75 sugiere que el error promedio absoluto es muy bajo. La puntuación R² de casi 1 (0.9999996436189778) indica que el modelo explica

prácticamente el 100% de la variabilidad en los datos. Los mejores hiperparámetros encontrados fueron 'alpha': 0.01 y 'solver': 'svd'.

```
Fitting 10 folds for each of 42 candidates, totalling 420 fits
Rendimiento de Ridge para Cartera_Dudosa:
Error Cuadrático Medio: 349.50018586953826
Error Absoluto Medio: 7.7549375198471795
Puntuación R^2: 0.999996436189778
Mejores Hiperparámetros: {'alpha': 0.01, 'solver': 'svd'}
```

En cuanto a las predicciones ARIMA, la mayoría de las unidades de negocio muestran una tendencia lineal en las predicciones de ventas brutas a lo largo del año, lo que puede significar una planificación controlada de puesta de productos a la venta.

La mayoría de las unidades muestran una tendencia más estable a lo largo del año. Sin embargo, la unidades "Confitería", tiene una tendencia a la alta en ciertos meses, lo cual podría indicar una estacionalidad en las ventas o efectos de promociones y estrategias de mercado.

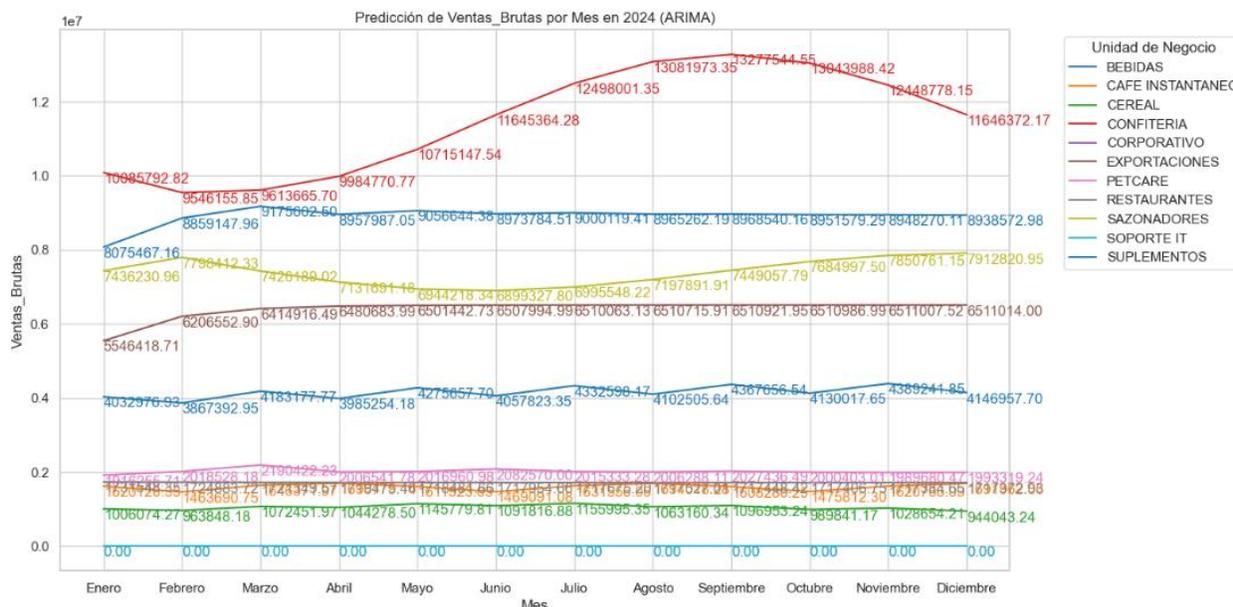


Figura 17: Predicción Ventas Brutas (Arima)
Fuente: Peña, Ayerve, 2024

El modelo ARIMA ajustado para "Ventas Brutas" en la unidad de negocio Café Instáneo muestra una puntuación R^2 de 0.1356, indicando que el modelo explica aproximadamente el 13.56% de la variabilidad. El modelo para "Costo de Material Prima Menos Margen" tiene una puntuación R^2 de 0.0147, indicando que explica solo el 1.47% de la variabilidad.

<p>Unidad de Negocio: CAFE INSTANTANEO, Código Empresa: EC12 Mejores Hiperparámetros para Ventas Brutas: (5, 0, 2) Error Cuadrático Medio (ECM) para Ventas Brutas: 53411143973.8847 Error Absoluto Medio (MAE) para Ventas Brutas: 182938.1909818419 Puntuación R^2 para Ventas Brutas: 0.13562855717656752</p> <p>Mejores Hiperparámetros para Costo de Material Prima Menos Margen: (5, 1, 3) Error Cuadrático Medio (ECM) para Costo de Material Prima Menos Margen: 3582164390.357714 Error Absoluto Medio (MAE) para Costo de Material Prima Menos Margen: 39801.172484931165 Puntuación R^2 para Costo de Material Prima Menos Margen: 0.014730842838253544</p>

Para "CONFITERIA", el modelo ARIMA para "Ventas Brutas" tiene una puntuación R^2 de 0.6473, indicando un ajuste decente donde el modelo explica el 64.73% de la variabilidad. Para "Costo de Material Prima Menos Margen", la puntuación R^2 es 0.4291, indicando que el modelo explica el 42.91% de la variabilidad.

<p>Unidad de Negocio: CONFITERIA, Código Empresa: EC12 Mejores Hiperparámetros para Ventas Brutas: (4, 1, 5) Error Cuadrático Medio (ECM) para Ventas Brutas: 2109795297716.9055 Error Absoluto Medio (MAE) para Ventas Brutas: 1273885.296290192 Puntuación R^2 para Ventas Brutas: 0.6472834818816916</p> <p>Mejores Hiperparámetros para Costo de Material Prima Menos Margen: (4, 1, 5) Error Cuadrático Medio (ECM) para Costo de Material Prima Menos Margen: 87683069118.63557 Error Absoluto Medio (MAE) para Costo de Material Prima Menos Margen: 218325.04677958562 Puntuación R^2 para Costo de Material Prima Menos Margen: 0.42913408741196246</p>

En la gráfica de predicción de costo de materia prima se observan fluctuaciones en los costos a lo largo del año, lo que podría reflejar cambios en los precios de las materias primas, la implementación de nuevas estrategias de producción o cambios en la demanda.

Algunas unidades de negocio como "Exportaciones" muestra picos en ciertos meses, lo que puede sugerir una necesidad de planificación especial para esos periodos,

posiblemente por costos más elevados debido a la demanda del mercado o a cambios en los precios de los insumos.

Tener visibilidad anticipada de estos datos es muy importante para los gerentes de operaciones y financieros, ya que proporciona insights sobre los períodos de mayor costo y permite planificar con anticipación para optimizar los márgenes. Por ejemplo, podrían considerar estrategias para mitigar los efectos de los picos de costos, como la compra anticipada de materiales o la renegociación de contratos con proveedores antes de los meses más caros.

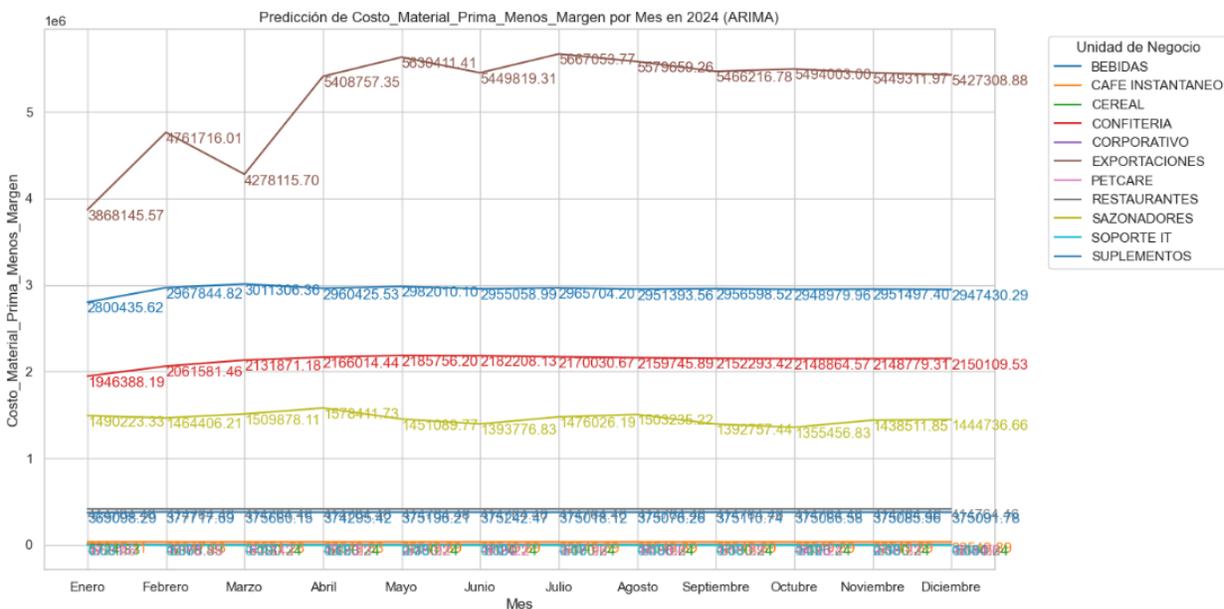


Figura 18: Predicción Costo Materia Prima (Arima)
Fuente: Peña, Ayerve, 2024

Esta presentación gráfica permite a los stakeholders evaluar visualmente el rendimiento proyectado de diferentes aspectos del negocio, haciendo posible la identificación rápida de patrones, la verificación de la coherencia de las predicciones con las expectativas del negocio y la planificación adecuada en respuesta a las tendencias anticipadas. Este enfoque valida la eficacia de los modelos predictivos utilizados y asegura que los resultados sean accesibles y útiles para la toma de decisiones (Patel, Shah, & Thakkar, 2020).

Finalmente, se presenta un modelo de Dashboard interactivo en Power Bi que permitirá a la organización tener una visión holística e integral de las principales cuentas y grupos de cuentas de los Estados Financieros e identificar rápidamente los resultados de indicadores financieros.

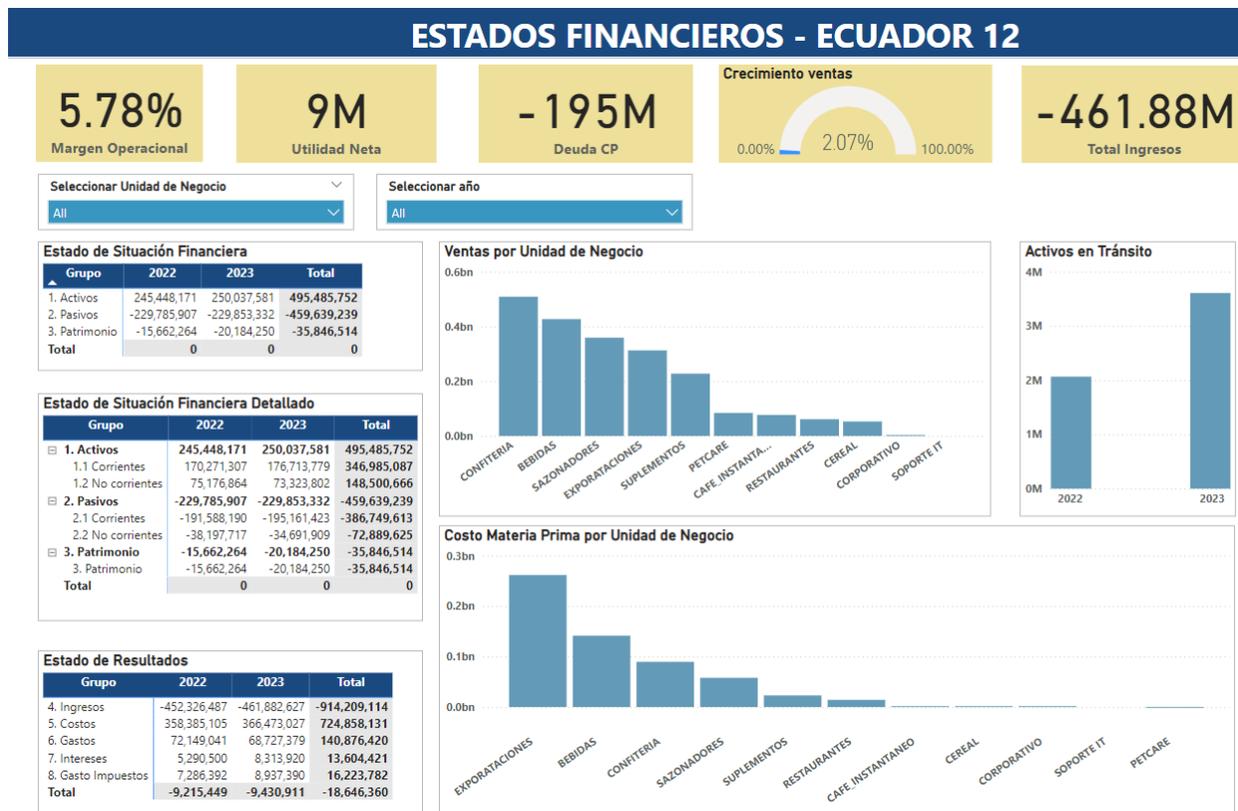


Figura 19: Dashboard Interactivo de Estados Financieros
Fuente: Peña, Ayerve, 2024

DISCUSIÓN DE LOS RESULTADOS Y PROPUESTA DE SOLUCIÓN

Implicaciones para la organización

Dado que la compañía hace varios años tiene un compromiso continuo con la excelencia operativa y la innovación, la organización está excepcionalmente equipada para implementar el modelo de análisis descriptivo y predictivo propuesto. La compañía cuenta con un equipo altamente capacitado de profesionales en análisis de datos, recursos tecnológicos de última generación y una sólida base financiera que garantizan la ejecución efectiva de este proyecto.

Además, el apoyo decidido de la alta gerencia no solo refuerza la viabilidad de la implementación, sino que también asegura una alineación con los objetivos estratégicos de la empresa.

A continuación, se presentan algunos de los impactos al implementar el modelo:

- **Eficiencia en el Procesamiento de Datos:** Utilizando Python para el análisis exploratorio de datos (EDA), la organización ha mejorado significativamente la eficiencia en el procesamiento y análisis de datos financieros. Esta herramienta permite un análisis descriptivo rápido y efectivo, facilitando una comprensión más profunda y ágil de los datos financieros.
- **Benchmarking de Unidades de Negocio:** El análisis proporciona un conocimiento detallado de las unidades de negocio con mejor y peor desempeño. Este benchmarking es crucial para identificar áreas de éxito y aquellas que requieren intervención, permitiendo una gestión basada en el rendimiento.

- **Visibilidad de Tendencias y Montos Significativos:** Se ha logrado una mayor visibilidad de los picos en cuentas significativas, tanto en términos de temporalidad como de distribución de montos por unidad de negocio. Esto permite a la organización identificar patrones y tendencias clave, optimizando la toma de decisiones y la asignación de recursos.
- **Modelos de Predicción para Cuentas Contables:** La implementación de modelos de predicción avanzados ha capacitado a la organización para anticipar el comportamiento futuro de cuentas contables importantes, mejorando la planificación financiera y la respuesta a cambios potenciales del mercado.
- **Prevención y Corrección en la Gestión:** Los modelos predictivos fomentan una capacidad de reacción tanto preventiva como correctiva, permitiendo a la organización anticiparse a problemas potenciales y ajustar planes y estrategias en consecuencia antes de que las desviaciones se conviertan en problemas mayores.
- **Dashboards Dinámicos:** El uso de dashboards para la exploración y análisis de datos financieros facilita la interpretación dinámica y continua de las finanzas de la organización. Estas herramientas visuales permiten a los gestores monitorizar y responder rápidamente a los cambios, asegurando que las decisiones se basen en los datos más actuales.
- **Visualización de Predicciones:** La visualización de las predicciones financieras proporciona una representación clara y gráfica de las tendencias futuras y el comportamiento de las cuentas. Esto es esencial para la planificación estratégica y operativa, ya que ofrece una visión anticipada y detallada del posible escenario financiero de la empresa.
- **Planificación de Presupuesto:** La organización puede planificar mejor sus presupuestos anuales sabiendo qué unidades requerirán más inversión en medios

digitales y en qué momentos. Además, las unidades con menores gastos pueden ser evaluadas para determinar si existe una subinversión en medios digitales que podría estar limitando su crecimiento o visibilidad en el mercado.

Uno de los principales impactos en la organización al implementar este modelo es la mejora en la asignación de presupuestos. Con el modelo predictivo, la asignación de presupuesto desde la casa matriz es mucho más eficiente y asertiva. Estos modelos permiten una asignación de recursos más informada y estratégicamente viable, basada en proyecciones confiables.

Se realizó una comparación entre los valores proyectados por la organización con métodos tradicionales versus este modelo y se determinó que el modelo de predicción con RIDGE y ARIMA muestra una presión mayor en un 15%.

Para terminar, el uso de herramientas de Business Intelligence, como Python y Power Bi para el análisis exploratorio, ha permitido reducir tiempos de procesamiento de Estados Financieros de 8 horas por mes a 2 horas mensuales por persona. Apalancando que este tiempo optimizado pueda ser utilizado no en procesos operativos sino estratégicos.

CONCLUSIONES Y RECOMENDACIONES

Conclusiones

El análisis de correlación revela una asociación entre las inversiones en marketing y el desempeño de los distribuidores. Aunque esta relación existe, se encontró que no es particularmente fuerte, sugiriendo que mientras los gastos de marketing incrementan, las ventas tienden a aumentar solo ligeramente.

Se observó que los incentivos ofrecidos a los distribuidores tienen un impacto directo y positivo en el crecimiento de las ventas. Esta relación directa subraya la efectividad de las estrategias de incentivos en el fomento de un mayor volumen de ventas.

En términos de ventas netas, la Unidad de Negocio de Confitería se destacó, mostrando los valores más altos tanto en mediana como en rango intercuartil, seguida de cerca por las unidades de "Bebidas" y "Exportaciones". Esto indica que estas unidades son particularmente robustas en términos de rendimiento de ventas dentro de la empresa.

El modelo ARIMA proporcionó predicciones valiosas sobre el comportamiento de las ventas basándose en datos históricos desde 2020 hasta 2023, proyectando una tendencia al alza para el año 2024. Esta herramienta de modelado demostró ser efectiva para entender las dinámicas temporales de las ventas.

Por otro lado, el modelo de predicción Ridge destacó una tendencia ascendente en la asignación de presupuesto para gastos de Marketing Digital, reflejando posiblemente una estrategia corporativa enfocada en fortalecer la presencia digital en respuesta a las tendencias del mercado.

Recomendaciones

Se sugiere coordinar con el área comercial un análisis detallado de las campañas de marketing para identificar cuáles han sido más efectivas con la finalidad de refinar las estrategias de marketing y concentrar esfuerzos en las tácticas que generan mejores resultados en ventas. Obtener y aplicar los insights de las unidades de negocio con mayores ventas para replicar este éxito en otras unidades y compartir las mejores prácticas y lecciones aprendidas puede potenciar el rendimiento en toda la organización.

Considerar ajustes en la estrategia de marketing para mejorar la eficiencia y el impacto en las ventas, así como incluir la reasignación de recursos hacia las campañas más rentables o la experimentación con nuevos enfoques de marketing.

Integrar los modelos predictivos en el sistema de gestión empresarial para monitorizar y prever métricas clave de manera continua para una gestión más proactiva y basada en datos.

Ser consciente de que los modelos dependen significativamente de la calidad y cantidad de los datos históricos proporcionados. Es vital asegurar que los datos utilizados sean precisos y estén completos para mantener la fiabilidad de las predicciones.

Estar preparados para ajustar los modelos ante cambios abruptos en el mercado que no estén reflejados en los datos históricos. Esto puede requerir una revisión y actualización periódica de los modelos para incorporar nueva información y tendencias emergentes.

Ajustar la provisión para deudas incobrables según las predicciones para asegurar la salud financiera de la empresa sin comprometer la liquidez.

Replicar y expandir el uso de dashboards interactivos en las áreas de Estudio de Mercado para permitir un acceso más fácil y dinámico a las predicciones y métricas relevantes para mejorará la capacidad de respuesta y la toma de decisiones basada en datos en tiempo real.

REFERENCIAS

- Aucancela Guamán, M. A., Viteri, J. O., & Castro, C. V. (2019). Plataformas de visualización de datos e indicadores financieros y de gestión. *Ciencia Digital*.
- Berat Sezer, O. (2020). Financial time series forecasting with deep learning. *Applied Soft Computing*.
- Box, G. E., Jenkins, G., & Reinsel, G. (2021). *Time Series Analysis: Forecasting and Control*.
- Burguete, A. (2016). *Análisis financiero*. Editorial digital UNID.
- Delfino Barilla, C., & Lastarria, L. R. (2020). ¿Cómo puede contribuir el Big Data a mejorar el rendimiento financiero de las empresas? *Revista Mexicana de Economía y Finanzas*.
- Elliott, B., & Elliott, J. (2019). *Financial Accounting and Reporting*. Pearson.
- Fischer, T. (2020). Deep learning with long short-term memory networks for financial. *European Journal of Operational Research*.
- García, V. (2020). *El análisis financiero como herramienta para la toma de decisiones*.
- González Rodríguez, A. A. (2023). *Diseño de investigación de las razones financieras básicas por medio de la herramienta Microsoft Power Business Intelligence para la optimización de la toma de decisiones para una organización desarrolladora de proyectos, ubicada en la ciudad de Guatemala*. Guatemala.
- Gorakala, S., & Dangeti, P. (2019). *Hands-On Exploratory Data Analysis with Python*.
- Guerra Rodríguez, J. (2022). *Fundamentos y variantes de los modelos ARIMA para el análisis de series temporales: Aplicación a la estadística universitaria*.
- Khare, J. (2019). Exploratory Data Analysis of Financial Data Using Python. *Towards Data Science*.
- Kulkarni, P. (2023). Advanced Analytics Driven Financial Management: An Innovative Approach to Financial Planning & Analysis. *Academia.edu*.
- Patel, J., Shah, S., & Thakkar, P. (2020). *Análisis Predictivo de los Precios de las Acciones usando Machine Learning*.
- Paz, E. D. (2022). *Integración de herramientas estadísticas con Python*.
- Pedregosa, F. (2020). Scikit-learn: Machine learning in Python. *Machine Learning Research*.

- Pinda Guanolema, B. R. (Octubre de 2023). Business intelligence en el análisis financiero.
- PythonBaba. (Febreri de 2023). *PythonBaba*. Obtenido de Using Python for Financial Analysis and Forecasting: A Comprehensive Guide: <https://pythonbaba.com/using-python-for-financial-analysis-and-forecasting/>
- Riera Riera, B. (2020). Análisis Contable y Financiero: Una Herramienta Clave Para La Eficiente Gestión Financiera Y Contable. *KnE Engineering*.
- Zou, H., & Hastie, T. (2020). Regularization and variable selection via the elastic net. *Royal Statistical Society*, 301-320.
- Fernández, A. (2020). El papel del Big Data en el reporting y la toma de decisiones. *Revista de Contabilidad y Dirección*, 31.
- Hilpisch, Y. (2019). Python for Finance: Mastering Data-Driven Finance. In *O'Reilly Media, Inc.*
- Python, T., & Gmbh, Q. (2017). Python For Finance: Algorithmic Trading. *DataCamp Community*.
- Tatsat, H., Puri, S., & Lookabaugh, B. (2020). Machine Learning and Data Science Blueprints for Finance. In *O'Reilly Media*.

ANEXOS

Anexo 1 Matriz de investigaciones similares

Metodología	Tipos de datos utilizados	Resultados	Implicaciones gerenciales	Referencia
Análisis descriptivo y exploratorio.	Datos estructurados de estados financieros como balances generales, cuentas de resultados y estados de flujo de efectivo.	Evaluación del rendimiento económico y financiero real de una empresa, detectando dificultades y aplicando medidas correctivas.	Análisis de estados financieros para determinar la liquidez, solvencia, rentabilidad, crecimiento del negocio y pronósticos.	(Riera Riera, 2020)
Análisis de Big Data (BDA) con análisis prospectivo, de riesgo y estratégico	Datos financieros estructurados, datos de transacciones en línea, datos de redes sociales, datos XML y datos en la nube.	Identificación de variables críticas que generan valor, diseño de estrategias contingentes y evaluación de impactos en KPI financieros.	Innovación y la flexibilidad organizacional, mejora del rendimiento financiero mediante un ajuste entre las actividades de Big Data, la estructura organizacional y el compromiso de la alta gerencia.	(Delfino Barilla & Lastarria, 2020)
Inteligencia Artificial y Machine Learning	Estructurados, datos históricos y predictivos sobre ventas y comportamiento del cliente	Experiencias de compra mejoradas, entrega eficiente y distribución optimizada por región	Mejora en la toma de decisiones basada en datos, aumento en la satisfacción del cliente.	(Kulkarni, 2023)
Técnicas de visualización de datos	Estructurados, datos financieros y contables.	Identificación de las principales plataformas de visualización de datos para facilitar la generación de indicadores financieros y de gestión.	Incremento de la competitividad organizacional a través de la eficiencia y rapidez en la gestión de información.	(Aucancela Guamán, Viteri, & Castro, 2019)

Metodología	Tipos de datos utilizados	Resultados	Implicaciones gerenciales	Referencia
Series Temporales y Deep Learning	Estructurados, datos financieros y contables.	Predicción de acciones, índices, tendencias, productos básicos, volatilidad, divisas y criptomonedas.	Modelos con una mejor comprensión de cómo implementar modelos de Deep Learning, lo que conlleva a decisiones más informadas y mejores prácticas en la predicción financiera.	(Berat Sezer, 2020)
Series Temporales, Python, Visualización de datos	Estructurados, datos financieros y contables.	Técnicas de análisis y modelado avanzado usando Python que mejoran la eficiencia en la gestión de datos y recursos financieros.	Adoptar un enfoque más data-driven, lo cual mejora la precisión de las predicciones y la capacidad de adaptación a cambios del mercado. Además, fomenta la innovación y el desarrollo de soluciones personalizadas en finanzas corporativas.	(Hilpisch, 2019)
Deep Learning	Estructurados, datos financieros y contables.	Mejores predicciones y decisiones más informadas para inversores y gestores financieros y precisión de las previsiones y en la gestión de riesgos financieros.	Desarrollo de modelos que superaron a los enfoques convencionales en términos de precisión y capacidad para manejar la naturaleza no lineal y compleja de los datos financieros.	(Fischer, 2020)

Metodología	Tipos de datos utilizados	Resultados	Implicaciones gerenciales	Referencia
Análisis Exploratorio de Datos (EDA)	Estructurados, datos históricos de precios de acciones.	Identificación de patrones estacionales, análisis de la distribución de precios y detección de outliers. Visualización de correlaciones entre precios de diferentes acciones.	Mejor comprensión del comportamiento de las acciones, optimización de estrategias de inversión y mejora en la toma de decisiones basada en datos financieros detallados	(Khare, 2019)
Análisis cuantitativo y cualitativo. Análisis de datos financieros históricos.	Estructurados, datos financieros históricos, datos transaccionales, y entrevistas con expertos.	Identificación de patrones en los datos financieros que permiten mejorar la toma de decisiones. Desarrollo de modelos predictivos para anticipar tendencias del mercado.	Mejor capacidad para tomar decisiones estratégicas informadas, optimización de recursos, y aumento en la precisión de predicciones financieras. Recomendaciones para la implementación de tecnologías de Business Intelligence en la gestión financiera.	(Pinda Guanolema, 2023)
Bibliotecas de Python (Pandas, NumPy, Matplotlib, Scikit-learn y Statsmodels).	Estructurados, datos históricos financieros, precios de acciones, tasas de interés, indicadores económicos.	Mejora en la precisión de las predicciones financieras y la capacidad de identificar tendencias y patrones en los datos financieros.	Permite a los gerentes tomar decisiones informadas, gestionar riesgos y planificar estratégicamente con mayor eficacia.	(PythonBaba, 2023)

Metodología	Tipos de datos utilizados	Resultados	Implicaciones gerenciales	Referencia
Machine Learning, bibliotecas pandas, scikit-learn, TensorFlow y Keras	Estructurados, datos financieros históricos, datos de mercado, transacciones en tiempo real, datos económicos y de noticias.	Desarrollo de modelos predictivos, algoritmos de trading automatizado, gestión de riesgos y optimización de carteras.	Implementar técnicas avanzadas de machine learning para mejorar la toma de decisiones, optimizar estrategias de inversión.	(Tatsat et al., 2020)
Algoritmos de machine learning, regresión lineal, máquinas de soporte vectorial, y redes neuronales	Estructurados, datos históricos de precios de acciones, incluyendo precios de apertura, cierre, máximo, mínimo y volumen de transacciones.	Los modelos algoritmos de Machine Learning mostraron una mayor precisión en la predicción de precios de acciones en comparación con la regresión lineal y las máquinas de soporte vectorial.	Permite a los inversores y gestores financieros optimizar sus estrategias de inversión y adaptarse mejor a las condiciones del mercado.	(Patel, Shah, & Thakkar, 2020)

Anexo 2 Comparación predicción modelo vs Datos reales de la compañía

Link a Github para Notebook: [GitHub - patoayerve/PROYECTO_CAPSTONE_EC12: PROYECTO CAPSTONE EC12](#)

