

---

# IMPLEMENTACIÓN DE UN MODELO PREDICTIVO PARA LA AYUDA DE TOMA DE DECISIONES DEL EQUIPO DE VENTAS DE WALMART CONNECT

---

Sebastián Matías Parra Molina  
Proyecto para optar al título de Ingeniero Civil Industrial  
Santiago, 2023



## Índice

### Tabla de contenido

<b>Índice</b>	3
<b>Resumen ejecutivo</b>	5
<b>Abstract</b>	5
<b>Introducción</b>	6
1.Contexto	6
2.Problema	6
3.Oportunidad	9
<b>Objetivos</b>	9
1.Objetivo general	9
2.Objetivo específico	9
<b>Estado del arte</b>	10
1. Modelos predictivos	10
1.1.Series de tiempo	11
2. Sistemas de apoyo a la decisión	12
3. Business intelligence	13
<b>Solución</b>	14
Análisis de riesgo	14
<b>Metodología</b>	16
<b>Métricas de desempeño</b>	17
<b>Evaluación económica</b>	18
<b>Desarrollo</b>	19
1.Desarrollo	19
A) Entendimiento del negocio	19
B) Entendimiento de los datos	20
c) Preparación de los datos	22

<b>D)Modelado .....</b>	<b>25</b>
<b>Resultados.....</b>	<b>33</b>
Resultados de modelos en LIDERAPP .....	34
<b>Resultados de cierre de año .....</b>	<b>36</b>
<b>Conclusiones .....</b>	<b>40</b>
<b>Discusión .....</b>	<b>40</b>
<b>Referencias bibliográficas.....</b>	<b>41</b>
<b>Anexos .....</b>	<b>41</b>

## Resumen ejecutivo

En el presente informe, se aborda la problemática del incumplimiento de ventas de espacios digitales en Walmart Connect, centrándose específicamente en LIDERAPP y [LIDER.CL](#). El enfoque del proyecto se orienta hacia la toma de decisiones del equipo de ventas al iniciar el proceso de oferta de espacios publicitarios, considerando tener prioridad con los clientes que más invierten.

Con el objetivo de abordar esta situación, el equipo de Walmart Connect ha facilitado el mes de diciembre como una oportunidad para buscar mayores inversiones en LIDERAPP, dado que históricamente ha sido el mes con menor rendimiento, especialmente en la plataforma de LIDERAPP.

Para llevar a cabo este análisis, se emplearon tres modelos de predicción de demanda: ARIMA, Prophet y Holt-Winter. Además, se implementó el modelo RandomForestRegressor para obtener una predicción de los cinco mejores clientes inversores a los que se les ofrecerán espacios publicitarios en el mes de diciembre. Este enfoque integral busca mejorar la toma de decisiones del equipo de ventas y potenciar las oportunidades de venta en un periodo crítico para el cumplimiento de objetivos.

## Abstract

The present report addresses the issue of underperformance in the sales of digital spaces on Walmart Connect, specifically focusing on LIDERAPP and [LIDER.CL](#). The project is approached from the perspective of the sales team's decision-making process when initiating the offering of advertising spaces, primarily considering customer priorities and efficient space allocation.

To address this challenge, Walmart Connect's team has provided the month of December as an opportunity to seek additional investments in LIDERAPP, as it has historically been the month with the lowest performance, particularly on the LIDERAPP platform. Consequently, the trial version of the project will concentrate on securing investments specifically for LIDERAPP.

To conduct this analysis, three demand prediction models were employed: ARIMA, Prophet, and Holt-Winter. Additionally, the RandomForestRegressor model was implemented to forecast the top five clients to whom advertising spaces will be offered in December. This comprehensive approach aims to enhance the sales team's decision-making and capitalize on sales opportunities during a critical period for goal attainment.

## Introducción

### 1.Contexto

Walmart Chile se destaca como una empresa dedicada a la gestión de supermercados, implementando una estrategia multiformato que le permite ofrecer soluciones adecuadas a diversos segmentos de clientes en todo el territorio nacional. Dentro de su estructura, destaca Walmart Connect, una división que representa una solución omnicanal.

La omnicanalidad, en este contexto, se refiere a la integración de todos los canales de venta disponible, tanto online como offline. En otras palabras, Walmart Connect utiliza una estrategia que conecta a los proveedores y vendedores de la compañía con los clientes a través de múltiples canales de comercialización, permitiendo una experiencia de compra coherente y unificada.

Esta solución de publicidad omnicanal ofrece a los proveedores y Sellers la oportunidad de potenciar sus productos. Esto se logra mediante el aprovechamiento de una extensa cantidad de información recopilada a través de los diversos canales de venta en la compañía, tanto entorno online como offline de los supermercados.

Actualmente, Walmart Connect ha incursionado en el rubro de retail media, que se centra en la venta de espacios publicitarios digitales. Este enfoque específico se materializa a través de plataformas LIDERAPP y LIDER.CL. En este contexto, los proveedores y sellers tienen la posibilidad de destacar y promocionar sus productos de manera efectiva, capitalizando la presencia de Walmart en estos destacados canales digitales.

### 2.Problema

Dentro de Walmart Connect, se encuentra el equipo de ventas, cuyos miembros tienen el cargo de Key Account Manager (KAM), identificados por las siglas iniciales, Cada KAM tiene a su cargo una cartera de proveedores y se encarga de ofrecerles los espacios publicitarios digitales disponibles. Para ello, elaboran y envían planes anuales, mensuales y trimestrales a sus respectivos proveedores. Cabe destacar que cada KAM opera con una meta anual de ventas establecida por Walmart, la cual se evalúa mensualmente. La rotación de espacios publicitarios ocurre semanalmente, iniciando la semana los jueves y terminando los miércoles.

El principal desafío del área radica en el incumplimiento de las metas asignadas por la empresa. Esta problemática surge debido a una serie de situaciones, que incluyen la carencia de estrategias sólidas, la limitada capacidad para anticipar comportamientos futuros de las ventas, una asignación ineficiente de los espacios publicitarios, la falta de capacidad de monetizar oportunidades futuras y la carencia de una estrategia de fijación de precios dinámicos para cada espacio, adaptados a las tendencias actuales y particularidades de cada proveedor.

Plan	Plan Liderapp	Cumplimiento
ENE	138,112,470	84,235,267
FEB	169,968,108	106,038,027
MAR	218,215,170	201,421,897
ABR	235,811,801	139,908,852
MAY	259,851,931	200,747,943
JUN	232,857,142	201,732,074
JUL	222,911,646	153,128,975
AGO	200,786,602	186,259,910
SEPT	193,854,232	173,820,409

Tabla 1: Plan LIDERAPP VS Cumplimiento Logrado hasta el mes de septiembre. CLP

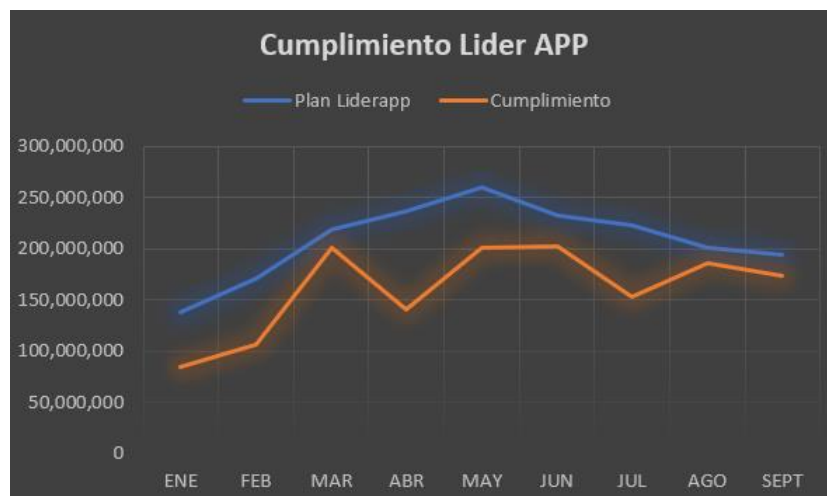


Ilustración 1: Gráfico de cumplimiento LIDERAPP. Elaboración propia.

En el siguiente gráfico y tabla se presentan los datos correspondientes a los meses de enero a septiembre del año 2023, destacando tanto el desempeño actual como el plan establecido por la empresa para LIDERAPP. El gráfico proporcionado ofrece una representación clara de que el cumplimiento está por debajo de las expectativas previstas.

Plan	Plan SOD	Cumplimiento
ENE	91,050,805	48,308,978
FEB	98,461,512	106,739,506
MAR	163,208,455	191,161,354
ABR	192,037,846	140,587,209
MAY	188,639,665	128,507,888
JUN	160,867,909	139,464,589
JUL	144,265,506	121,867,659
AGO	154,824,750	127,638,737
SEPT	164,408,993	108,419,025

Tabla 2: Plan LIDER.CL VS Cumplimiento logrado. En CLP.



Ilustración 2: Gráfico cumplimiento LIDER.CL. Elaboración propia.

De manera análoga al gráfico anterior, presentamos el rendimiento de LIDER.CL, representado como SOD, durante los meses de enero a septiembre. Este gráfico también destaca de manera evidente un incumplimiento claro de la meta establecida.

	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEPT
Cumplimiento APP	61%	62%	92%	59%	77%	87%	69%	93%	90%
Cumplimiento SOD	53%	108%	117%	73%	68%	87%	84%	82%	66%

Tabla 3: Porcentajes de cumplimiento LIDERAPP y LIDER.CL



### 3.Oportunidad

La oportunidad clave se centra en el potencial de incrementar las ventas de Walmart Connect. Todo esto se fundamenta en la capacidad del equipo para tomar decisiones informadas. Sin embargo, a pesar de contar con datos valiosos, la carencia de una estrategia predictiva sólida ha generado pérdidas como:

- **Pérdida de ingresos:** la falta de precisión en la asignación de espacios publicitarios ha resultado en una pérdida sustancial de ingresos para Walmart Connect. Esta deficiencia tiene un impacto financiero negativo, especialmente en relación con las metas de ventas establecidas por la empresa.
- **Pérdida de oportunidades:** La incapacidad de prever y satisfacer eficientemente la demanda de espacios publicitarios digitales ha ocasionado la pérdida de oportunidades. Esto se traduce en la incapacidad de ofrecer a los proveedores cuando están más interesados o tienen más posibilidad de invertir en ellos.

### Objetivos

#### 1.Objetivo general

Desarrollar una solución basada en análisis de datos históricos, que permita a Walmart Connect gestionar de mejor manera los espacios publicitarios digitales. La meta es mejorar de manera significativa la toma de decisiones del equipo de ventas, lo que resultará en un aumento sustancial de los ingresos generados por la venta de espacios publicitarios digitales. La implementación exitosa de esta solución no solo beneficiará la asignación de espacios publicitarios, sino que también establecerá un estándar más elevado para la calidad de decisiones estratégicas, posicionando a Walmart Connect como líder en el mercado publicitario omnicanal.

#### 2.Objetivo específico

La solución tiene como objetivo potenciar la toma de decisiones de equipo de Walmart Connect, proporcionando información clave para mejorar la asignación de espacios publicitario. Se busca mantener la tasa de ocupación de estos espacios en un mínimo del 90% mensual, garantizando una eficiente utilización de los recursos publicitarios disponibles. Al lograr esta mejora en la gestión de

espacios, el objetivo final es aumentar los ingresos de Walmart Connect, apuntando específicamente al cumplimiento del 100% de las metas mensuales por venir. Se busca una prueba para el mes de diciembre para una futura implementación en el 2024.

### Estado del arte

Dentro del contexto de un proyecto ingenieril, el estado del arte y búsqueda de una solución se centrarán en encontrar enfoques que conduzcan a la mejora de la toma de decisiones de un equipo o empresa. Todo esto con el propósito de desarrollar una solución que se ajuste con los objetivos del proyecto.

Basándose en la problemática, se encontraron tres posibles soluciones que contribuyen a la toma de decisiones, las cuales se mostrarán a continuación.

### 1. Modelos predictivos

Los modelos predictivos también conocidos como modelos de predicción son un conjunto de herramientas técnicas y estadísticas que sirven para pronosticar y predecir el comportamiento de un evento<sup>1</sup>.

Los modelos predictivos se dividen en dos categorías principales según el tipo de datos que analizan. Los modelos de clasificación se centran en categorizar la información, basándose principalmente en datos históricos. Estos modelos responden a preguntas binarias, como clasificar suscriptores como susceptibles o no a clicar en una campaña publicitaria, proporcionando además porcentajes de probabilidad, por otro lado, los modelos de regresión predicen o estiman valores numéricos, como el rendimiento futuro de un producto o individuo. En un ejemplo práctico, Enel implementó un modelo predictivo de demanda específico para cada una de sus redes sociales, como Twitter y Facebook, en respuesta a la necesidad de mejorar el control sobre los pronósticos futuros de demandas de los clientes en estas plataformas.<sup>2</sup>

Enel empleó diversos modelos de series de tiempo. Entre ellos, se destacan el Prophet de Facebook, el método de Holt-Winters y el modelo ARIMA. Todo esto con la finalidad de implementar un modelo para cada una de sus redes sociales.

---

<sup>1</sup>SHIRLY NOWAK (diciembre, 2022). Modelos predictivos ¿Qué son y para que se usan?

<sup>2</sup> Díaz (2020). Diseñar e implementar modelos de predicción de demanda de clientes para las redes sociales.

### 1.1. Series de tiempo

Para contextualizar los modelos investigados, profundizaremos el concepto de series de tiempo. Una serie temporal por definición es una “sucesión de observaciones de una variable realizadas a intervalos regulares de tiempo”<sup>3</sup>.

El estudio de las series temporales tiene como objetivo principal entender el comportamiento de una variable a lo largo del tiempo. Se busca obtener conocimiento sobre dicha variable para realizar predicciones futuras, asumiendo que no habrá cambios estructurales. La capacidad de prever el valor futuro de la variable es crucial en la toma de decisiones, ya que ayuda a la incertidumbre y mejora los resultados de los agentes involucrados.

#### 1.1.1. Método de Holt-Winters

Dentro del ámbito de las series temporales, se encuentra el método Holt-Winters que por definición es “una técnica de suavizado que utiliza el conjunto de estimaciones recursivas a partir de la serie histórica. Estas estimaciones utilizan una constante de nivel,  $\alpha$ , una constante de tendencia,  $\beta$ , y una constante estacional multiplicativa,  $\gamma$ ”<sup>4</sup>.

$$\hat{Y}_t = \alpha(\hat{Y}_{t-1} - T_{t-1}) + (1 - \alpha) \frac{Y_t}{F_{t-s}} \quad 0 < \alpha < 1$$

$$T_t = \beta T_{t-1} + (1 - \beta)(\hat{Y}_t - \hat{Y}_{t-1}) \quad 0 < \beta < 1$$

$$F_t = \gamma F_{t-s} + (1 - \gamma) \frac{Y_t}{\hat{Y}_t} \quad 0 < \gamma < 1$$

iii: Ecuaciones de estimaciones recursivas. Parra (enero, 2019)

Como se muestra en la ilustración iii, se utilizan parámetros específicos, como  $s = 4$  para datos trimestrales y  $s = 12$  para datos mensuales. En este contexto,  $\hat{Y}_t$  representa el nivel de suavizado de la serie temporal,  $T_t$  la tendencia suavizada y  $F_t$  el ajuste estacional suavizado. Estos elementos son esenciales para comprender y modelar el comportamiento de la serie temporal, proporcionando una base para la predicción y la interpretación de datos trimestrales y mensuales.

<sup>3</sup> Parra (enero, 2019). Estadísticas y machine learning con R.

<sup>4</sup> Parra (enero, 2019). Estadísticas y machine learning con R.

### 1.1.2. Modelo ARIMA

“También conocido como, trata de realizar previsiones acerca de los valores futuros de una variable utilizando únicamente como información la contenida en los valores pasados de la propia serie temporal.”<sup>5</sup>. El modelo ARIMA es una alternativa a la creación de modelos de ecuaciones únicas o simultáneas. En este caso, se asume que las series temporales tienen un carácter estocástico, lo que significa que sus propiedades probabilísticas deben ser analizadas para que los datos hablen por sí mismos.

### 1.1.3. Modelo Prophet FB

FB Prophet es un modelo muy extendido para la predicción de series temporales desarrollado por Facebook.

El modelo Prophet se basa en la suma o producto de distintos componentes: la tendencia ( $g(t)$ ), los componentes estacionales ( $s(t)$ ), los componentes de festivos ( $h(t)$ ), los componentes del resto de regresores ( $r(t)$ ) y finalmente el ruido blanco ( $\varepsilon_t$ ).

$$y(t) = g(t) + s(t) + h(t) + r(t) + \varepsilon_t$$

*Ecuación 1: Modelo Completamente aditivo.*

$$y(t) = g(t) \times (1 + s(t)) + h(t) + r(t) + \varepsilon_t$$

*Ecuación 2: Modelo con estacionalidad multiplicativa.*

## 2. Sistemas de apoyo a la decisión

“Un Sistema de Apoyo a las Decisiones (DSS por sus siglas en inglés Decision Support System), es un sistema informático que sustenta el proceso de toma de decisiones, lo cual implica la utilización de datos y modelos para la generación, la estimación, la evaluación y/o la comparación sistemática de alternativas, ayudando a los responsables de la toma de decisión a reunir inteligencia, generar opciones y tomar decisiones.”<sup>6</sup>

La instalación de un sistema de soporte de decisiones en una empresa conlleva diversos beneficios clave que impactan positivamente en su desempeño y competitividad. En primer lugar, el DSS contribuye a mejorar la competitividad al analizar y procesar grandes volúmenes de información

---

<sup>5</sup> Parra (enero, 2019). Estadísticas y machine learning con R.

<sup>6</sup> C.López (2014). Sistemas de apoyo a la toma de decisiones. Una aplicación en el área de Gestión Universitaria.

relevante. Esto agiliza la toma de decisiones correctas y proporciona una visión clave de la empresa en función de factores críticos, lo que permite adaptarse de manera más efectiva a un entorno empresarial.

### 3. Business intelligence

“Es el proceso de transformación de datos en información aplicando técnicas de extracción transformación y carga (ETL) proporcionando información validada para la adecuada toma de decisiones y de manera oportuna.”<sup>7</sup>

La investigación realizada se centró en la propuesta de una solución para un proyecto de Eknowledge Group S.A.C., satisfaciendo su necesidad de obtener información detallada sobre el flujo de ventas. Esto se logró mediante la generación de reportes analíticos, que incluyen aspectos como la distribución de clientes, ingresos según ubicación geográfica, seguimiento diario de ingresos, medios publicitarios para la capacitación de estudiantes, entre otros. Estos reportes proporcionaron información histórica, comparativa y consolidada en tiempo real, permitiendo un análisis detallado desde una perspectiva global hasta aspectos más específicos. El propósito final fue facilitar la toma de decisiones en la empresa. En la ilustración “4” se muestra cómo funciona la solución a grandes rasgos del proyecto investigado.

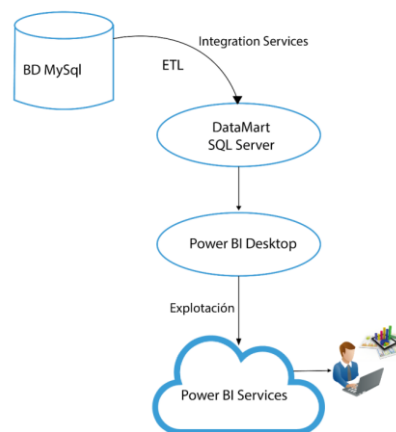


Ilustración 4: Arquitectura técnica del proyecto de Eknowledge S.A.C. Peña(abril,2018).

<sup>7</sup> Peña (abril, 2018). Implementación de una herramienta de Business Intelligence para mejorar el proceso de toma de decisiones financieras – Área de ventas – en Eknowledge Group S.A.C.

### Solución

Considerando el estado del arte y tomando en consideración los objetivos del proyecto y de la pasantía, y aplicando criterios que facilitan la selección de una solución coherente. Tenemos los siguientes criterios:

1. Objetivos ingenieriles del proyecto: Garantizar que la solución este alineada con los objetivos del proyecto y que a solución sea de acuerdo con el syllabus de la pasantía.
2. Fácil aplicación en el área: Seleccionar una solución que sea intuitiva y fácil de aplicar en el contexto de Walmart Connect.
3. Menor costo posible: Priorizar soluciones que ofrezcan eficiencia y funcionabilidad a un costo razonable. La implementación no solo debe ser efectiva en términos de mejora de la toma de decisiones, sino también económicamente viable para garantizar un retorno de inversión positivo.
4. Flexibilidad a futuros cambios: Diseñar una solución con una estructura que permita adaptarse a futuras mejoras continuas.

	Modelo Predictivo	Sistemas De apoyo	BI
Objetivos Ingenieriles del proyecto	✓	✓	-
Fácil aplicación en el área	✓	-	-
Menor costo posible	✓	-	✓
Flexibilidad a futuros cambios	✓	✓	✓

Tabla 4: Criterios de selección para la solución. Elaboración propia.

Considerando lo anterior, la opción que mejor se ajusta a la propuesta es desarrollar un modelo predictivo de demanda de espacios publicitarios digitales, utilizando herramientas como Python. Este enfoque se respalda por la experiencia del estudiante en el uso de Python, adquirida durante la formación universitaria, y por el conocimiento previo en modelos de predicción en el ámbito de la gestión de operaciones.

### Análisis de riesgo

A continuación, se presenta una matriz de riesgos para la solución escogida, junto con sus correspondientes mitigaciones.

Matriz de riesgo		Probabilidad				
		Improbable	Posible	Ocasional	Moderado	Constante
Severidad		1	2	3	4	5
Insignificante	1					
Menor	2				Adopción del usuario.	
Moderado	3	Cambios de los requisitos del proyecto.			Complejidad Del Modelo	
Crítico	4		Cambios significativos en el mercado de publicidad digital.		Escalabilidad.	Disponibilidad de datos históricos.
Severo	5				Falta de datos de calidad.	

Tabla 5: Matriz de riesgos. Elaboración propia

Color	Nivel
2 a 4	Aceptable
5 a 8	tolerable
10 a 12	Alto
15 a 25	severo

Tabla 6: Nivel de aceptabilidad de la Matriz de riesgos. Elaboración propia.

- Cambios de los requisitos del proyecto: Cambios repentinos de las necesidades del área.  
Mitigación: Reuniones constantes con el equipo de Walmart Connect.
- Adopción del Usuario: Resistencia al cambio por parte de los usuarios del área.  
Mitigación: Realizar sesiones de capacitación detalladas, ofreciendo soporte continuo y obtener retroalimentación regular para abordar preocupaciones y mejorar la adaptabilidad.
- Cambios significativos en el mercado digital: Cambios repentinos en el panorama digital, como nuevas plataformas, modificaciones en algoritmos de motores de búsqueda o tendencias de consumo.  
Mitigación: Flexibilidad del modelo predictivo que sea capaz de adaptarse rápidamente a cambios en el entorno digital.
- Complejidad del modelo: Pude traer problemas en el desarrollo e implementación a la solución escogida.  
Mitigación: Realizar una evaluación exhaustiva de la complejidad técnica antes de iniciar el desarrollo de este. También realizar prototipos y pruebas iterativas del modelo.
- Escalabilidad: Incapacidad del sistema para manejar el aumento significativo en la demanda o datos.

Mitigación: diseñar la solución con escalabilidad en mente y realizar pruebas de carga exhaustiva, Actualizar recursos según sea necesario.

- Falta de datos de calidad: la insuficiencia “p” baja calidad de los datos disponibles para el entrenamiento del modelo puede comprometer la precisión y eficiencia del modelo predictivo.

Mitigación: Auditoria de datos rigurosa para el conjunto de datos a disposición del proyecto.

- Disponibilidad de datos históricos: La falta de disponibilidad o acceso limitado a datos históricos relevantes puede afectar negativamente la capacidad del modelo para aprender patrones y realizar predicciones precisas.

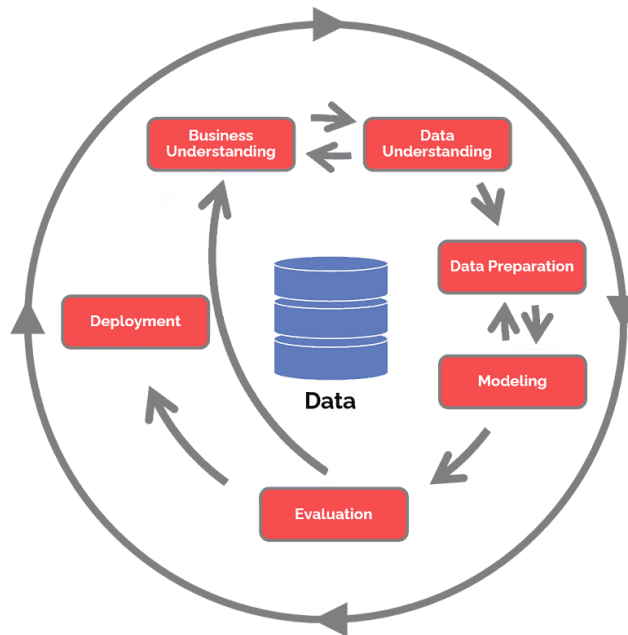
Mitigación: Identificar cualquier limitación de la obtención de datos y buscar soluciones proactivas.

## Metodología

La metodología que seguirá el proyecto se basará en el enfoque CRISP-DM (Cross-Industry Standard Process for Data Mining), utilizado en la industria de ciencia de datos. La metodología se compone de los siguientes pasos:

1. Entendimiento del negocio: Comprender en profundidad los objetivos y las necesidades del negocio de Walmart Connect, así como el contexto de la industria de la publicidad digital.
2. Entendimiento de los datos: Recopilar y explorar los datos históricos de ventas, comportamiento de los proveedores y cualquier otra fuente de datos relevantes.
3. Preparación de los datos: Limpieza, transformación y preparación de los datos para su uso en el modelado.
4. Modelado: Construcción del modelo predictivo de demanda y desarrollo de sistemas de gestión de categorías y fijación de precios dinámicos.
5. Evaluación: Evaluación de la precisión del modelo y el rendimiento de los sistemas desarrollados utilizando métricas relevantes.
6. Implementación: Integración de los sistemas en la operación de Walmart Connect.
7. Monitoreo: Seguimiento continuo del rendimiento de los sistemas en producción y ajustes según sea necesario.





V: Ilustración de la metodología CRISP-DM. Hotz (enero, 2023)

### Métricas de desempeño

Para determinar si el proyecto a desarrollar es exitoso se evaluarán en base a ciertas medidas de desempeño (KPI's), las cuales son:

1. **Incremental de ventas:** Se calculará el aumento en los ingresos por ventas de espacios publicitarios digitales en comparación con el periodo anterior a la implementación del modelo predictivo. El objetivo por lograr es alcanzar un aumento del 15% en ingresos mensuales, apuntando al 100% de las respectivas metas.
2. **Porcentaje de cumplimiento de cada Kam:** Esta métrica se calculará como el porcentaje de ventas totales de espacios publicitarios digitales generados por cada kam y se comparará con periodos anteriores.

$$\text{Cumplimiento} = \frac{\text{Espacios vendidos por cada Kam}}{\text{Total de espacios disponibles}} \times 100$$

Ecuación 3: Medida de desempeño 2. Elaboración propia.

Es importante señalar que cada KAM maneja clientes diferente, por lo que no todos venderán lo mismo. Es por esto la razón de esta métrica.

3. **Ocupación de espacios:** Se medirá la tasa de ocupación de espacios publicitarios, que representa el porcentaje de espacios vendidos sobre el total de espacios disponibles. El objetivo es alcanzar una tasa de ocupación superior al 90% y poder apuntar al 100%.

$$\text{Ocupación de espacios} = \frac{\text{Espacios totales vendidos}}{\text{Espacios totales Disponibles}} \times 100$$

*Ecuación 4: Medida de desempeño 3. Elaboración propia.*

#### Evaluación económica

	Mes 1	Mes 2	Mes 3	Mes 4	Mes 5
Actividad/Mes	Agosto	Septiembre	Octubre	Noviembre	Diciembre
Horas Mensuales Trabajadas	162.00	162.00	162.00	162.00	162.00
Costo por hora mano de obra(-)	\$ 7,385.00	\$ 7,385.00	\$ 7,385.00	\$ 7,385.00	\$ 7,385.00
Costo mano de obra mensual(-)	\$ 1,196,370.00	\$ 1,196,370.00	\$ 1,196,370.00	\$ 1,196,370.00	\$ 1,196,370.00
Costos por progrma python(-)	\$ -	\$ -	\$ -	\$ -	\$ -
Ingresos generados del proyecto	\$ -	\$ -	\$ -	\$ -	\$ 16,642,620.00
Utilidad	-\$1,196,370.00	-\$ 1,196,370.00	-\$ 1,196,370.00	- \$1,196,370.00	\$ 15,446,250.00

*Tabla 7: Viabilidad económica del proyecto. Elaboración Propia.*

TIR	53%
-----	-----

*Tabla 8: TIR del proyecto. Elaboración Propia*

Realizaremos la evaluación económica en relación con las mejoras en las ventas previstas para LIDERAPP, dado que se nos concedió un periodo de prueba durante el mes de diciembre con la plataforma. En este contexto, contamos con un registro de las horas mensuales trabajadas, que suman un total de 810 horas durante la pasantía. El costo por hora para los ingenieros civiles industriales es de 7,385 pesos<sup>8</sup>. Finalmente, calculamos el costo de obra mensual multiplicando las horas mensuales por el costo de mano de obra. En cuanto a los ingresos, durante la fase de prueba del proyecto en diciembre, se espera un aumento del 15% respecto a las ventas registradas en el

<sup>8</sup> Salario medio para Ingeniero Civil Industrial en Chile (2023)

mismo mes. Esto se traduce en una cifra estimada de  $110,950,800 * 15\% = 16,642,620$  millones de pesos, de acuerdo con la métrica de desempeño 1. Posteriormente, al calcular la Tasa Interna de Retorno (TIR) en Excel, obtenemos un resultado del 53%.

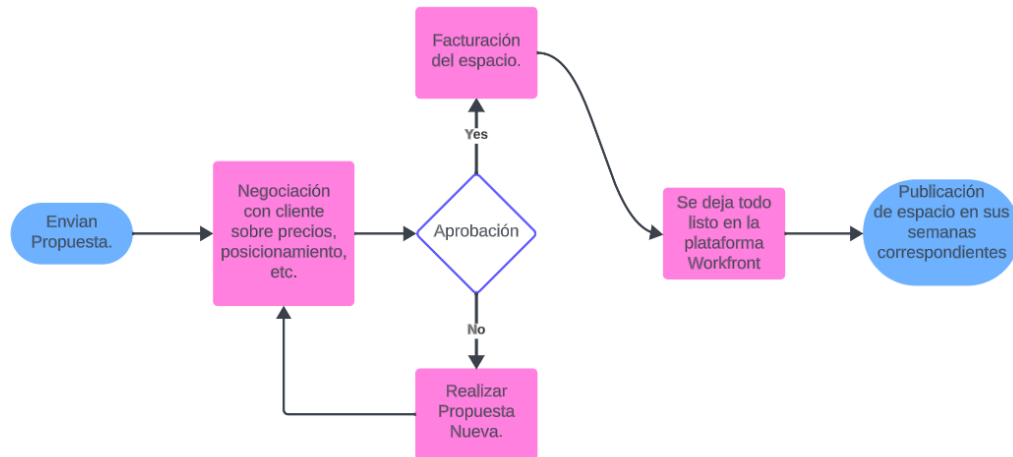
## Desarrollo

### 1.Desarrollo

#### A) Entendimiento del negocio

Primero, es crucial comprender el funcionamiento del negocio de Walmart connect. Inicialmente, el Key Account Manager (KAM) elabora una cotización ofreciendo a los proveedores espacios publicitarios en venta disponibles que van cambiando, dependiendo el a los espacios que les da el equipo de ecommerce. Estos espacios varían según estén en LIDERAPP o SOD; para LIDERAPP, se ofrecen Banners sliders, Search (Producto patrocinado), caluga, Mosaico carrusel y mini mosaico, mientras que para SOD pueden ser banners categoría, banners slider, caluga inferior, caluga superior, y search. La frecuencia de negociación de cada KAM puede ser semestral, trimestral, anual o mensual, dependiendo del proveedor.

Una vez enviado el plan, el proveedor confirma los espacios publicitarios que quiere adquirir y proporciona los requisitos de posicionamiento, negociación de precios, gráficos e información necesaria. Posteriormente, otro equipo se encarga de aprobar estos elementos, y una vez que el espacio ha sido aprobado por el equipo de operaciones, procede a la facturación por el equipo de finanzas, estando así listo para ser publicado en su semana correspondiente. Es relevante señalar que los planes pueden experimentar modificaciones según las necesidades específicas de cada cliente.



*Ilustración 6: Entendimiento del proceso de venta de un espacio publicitario. Elaboración propia.*

## B) Entendimiento de los datos

Una vez que hemos comprendido el funcionamiento de la venta de espacios en Walmart connect, procedemos a descargar los datos de la plataforma Workfront, la cual se detalla en el anexo 1. En dicha plataforma, se encuentra toda la información relacionada con los espacios publicitarios que cada KAM carga al realizar una venta.

Datos	Tipo de dato	Detalles
Name	Object	Nombre general del espacio publicitario
Estado comercial	Object	Estado del espacio publicitario (vendido o disponible)
Lista	Object	Lista a la cual pertenece el espacio (va de L1 a L12)
Mes	float64	Mes del año a cuál pertenece (1 a 12)
Semana	float64	Semana del año a cuál pertenece (1 a 52)
Desde	datetime	Fecha de inicio del espacio publicitario.
Hasta	datetime	Fecha de término del espacio publicitario.
KAM	Object	KAM que vendió el espacio.
Segmento	Object	Plataforma en donde se encuentra el espacio.
Espacio comercial	Object	Lugar dentro de la plataforma que se encuentra el espacio.
SLOT Homologado	Object	Nombre específico del espacio
Ubicación	float64	Numero de ubicación del espacio
Cliente	Object	Proveedor al cual le pertenece el espacio
Campaña	Object	Campaña del espacio publicitario
Tipo Facturación	Object	Tipo de facturación que quiere el cliente.
Inversión Final	float64	Monto la cual compra el espacio el Proveedor
Categoría	Object	Categoría del producto al que corresponde el espacio
Marca	Object	Marca de la publicidad
PN	Object	Numero identificador del producto/s promocionado

Tabla9: Tipos de datos obtenidos en Workfront. Elaboración propia

La tabla anterior evidencia la necesidad de una transformación clara de los datos para facilitar su análisis y lograr un exitoso modelado en Python.

### c) Preparación de los datos

#### 1. Transformación de los datos

Los datos totales comprenden un conjunto de 5731 registros de espacios publicitarios digitales vendidos. Los cuales se dividirán en dos conjuntos: LIDERAPP y SOD (LIDER.CL). Una vez que tenemos ambos conjuntos de datos, procedemos a su transformación.

En primera instancia, nos ocupamos de eliminar los valores nulos. Posteriormente agrupamos las columnas “Desde” y “SLOT Homologado”, contando las ocurrencias de la última columna. A continuación, renombramos las columnas y creamos un nuevo dataframe que incluye como columnas el nombre de cada espacio publicitario y el recuento y como datos el recuento semanal de cada uno de estos espacios.

Finalmente, realizamos una iteración sobre las columnas para crear un nuevo dataframe para cada espacio publicitario, facilitando así su modelado y predicción.



Ilustración 7: Demostración de como quedo cada división de los dataframes iniciales. Elaboración propia.

Datos	Tipo de dato	Detalles
Semanas	Datetime	Fecha Desde que muestra el inicio de cada semana del año
(Nombre de cada espacio correspondiente)	float64	Lleva el nombre y la información de cada espacio a la venta

Tabla 10: Tipos de datos que llevará cada variable para su manipulación. Elaboración propia.

## 2. Exploración de los datos

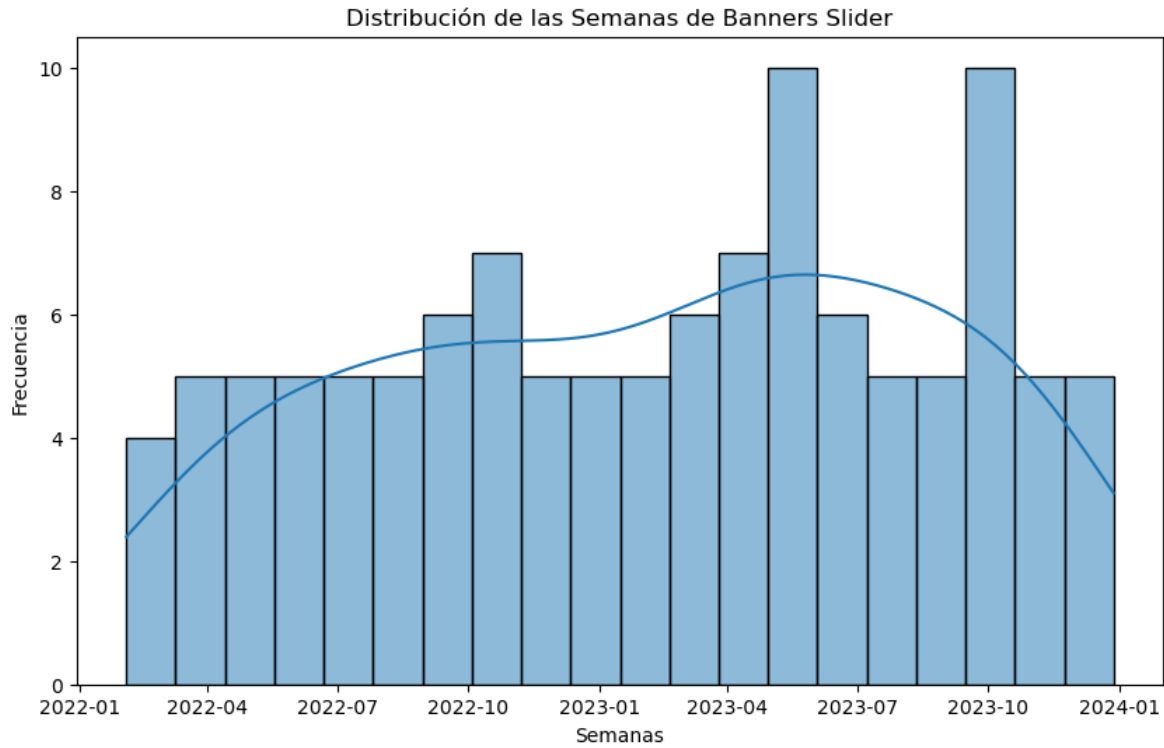
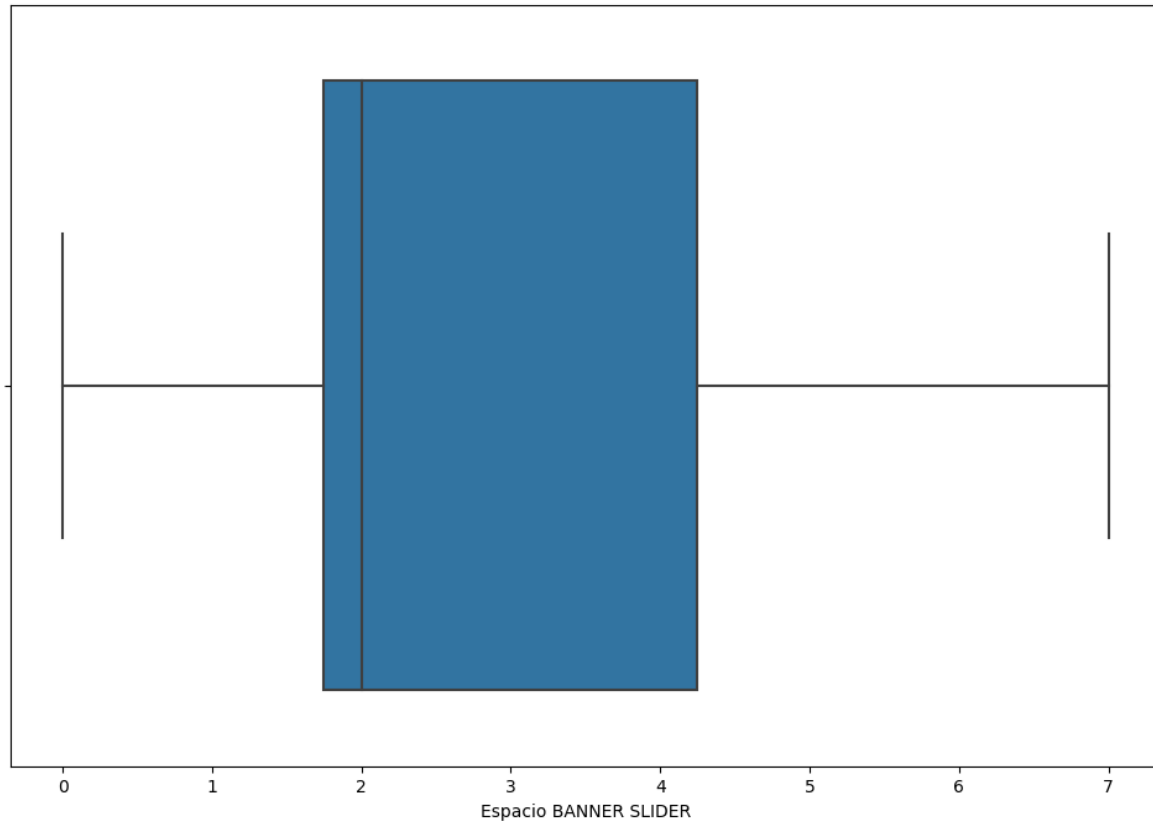


Ilustración 8: Gráfico de distribución de banners sliders semanal. Elaboración propia.



Boxplot de Espacio BANNER SLIDER para Identificar Outliers

*Ilustración 9: Boxplot para detectar outliers. Elaboración propia.*

Exploramos los datos en cada dataframe de espacio con el objetivo de identificar outliers u cualquier información relevante que nos ayude que pueda derivar en predicciones alejadas de la realidad. Una vez completado el preprocesamiento, avanzamos a la siguiente etapa, centrada en la construcción del modelo predictivo.

#### D)Modelado

Optaremos por utilizar series temporales en nuestro modelado, ya que se ajusta de manera óptima a los datos obtenidos. La elección de esta técnica se fundamenta en su capacidad para analizar y prever patrones en datos secuenciales a lo largo del tiempo, siendo especialmente adecuada para el contexto de nuestro análisis. Implementaremos tres métodos específicos en este proceso: el modelo ARIMA, el método Holt-Winters y el modelo Prophet de Facebook. Además, como solicitud específica de la empresa, se incluirá una predicción con el adicional con el modelo Random Forest Regressor, que consiste en un listado con los cinco clientes que invertirán más. Esta medida tiene

como objetivo proporcionar un respaldo adicional a la toma de decisiones, ofreciendo una perspectiva más detallada y orientada hacia los clientes de mayor impacto económico.

### 1. Modelo ARIMA

El modelo ARIMA (AutoRegressive Integrated Moving Average) se introduce de manera concisa en la revisión bibliográfica, pero para una comprensión más detallada en el contexto de Python y su aplicación en el presente proyecto, se explicarán sus componentes.

El modelo ARIMA se descompone en los siguientes componentes que son necesarios entenderlo para un buen modelado.

- **p (orden de autorregresión):** En ARIMA, "p" representa el orden de autorregresión, indicando cuántas observaciones pasadas se consideran para predecir el valor futuro. En el caso específico de 'order=(2, 1, 1)', se utilizan las dos observaciones anteriores después de aplicar una diferencia de primer orden.
- **d (grado de diferencia):** La integración se refiere al número de veces que se ha diferenciado la serie temporal para hacerla estacionaria. En este caso 'order=(2, 1, 1)' significa que se ha aplicado una diferencia de primer orden, lo que implica restar cada valor de la serie del valor anterior.
- **q (orden de media móvil):** Representa el tamaño de la ventana de la media móvil, es decir, cuantos errores pasados se deben tener en cuenta para predecir el siguiente valor. En este caso 'order=(2, 1, 1)' indica que se están considerando los dos errores pasados.

Proceso del código:

1. **Exploración y visualización:** Convierte la columna 'Semanas' a tipo datetime, establece 'Semanas' como índice y realiza un gráfico de la serie temporal.

Serie Temporal de Espacio Banner slider

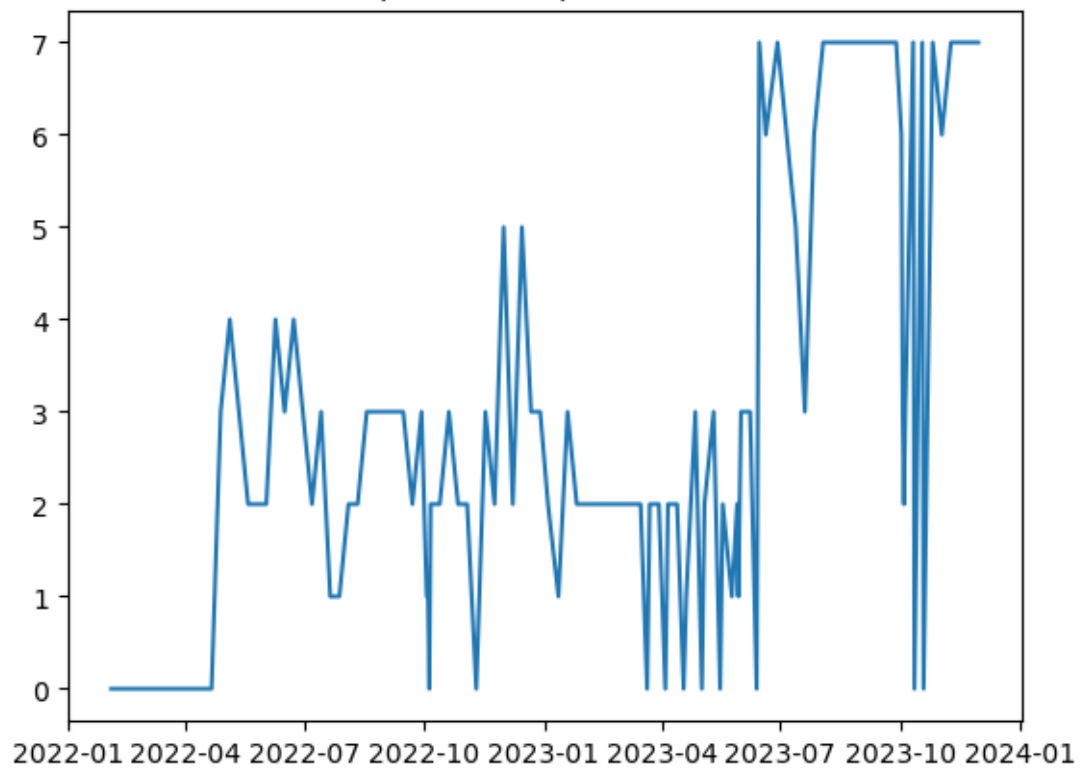


Ilustración10: Serie temporal Banner slider. Elaboración propia

- Descomposición estacional: Utiliza la función 'seasonal\_decompose' para descomponer la serie temporal en tendencia, estacionalidad y residuos.

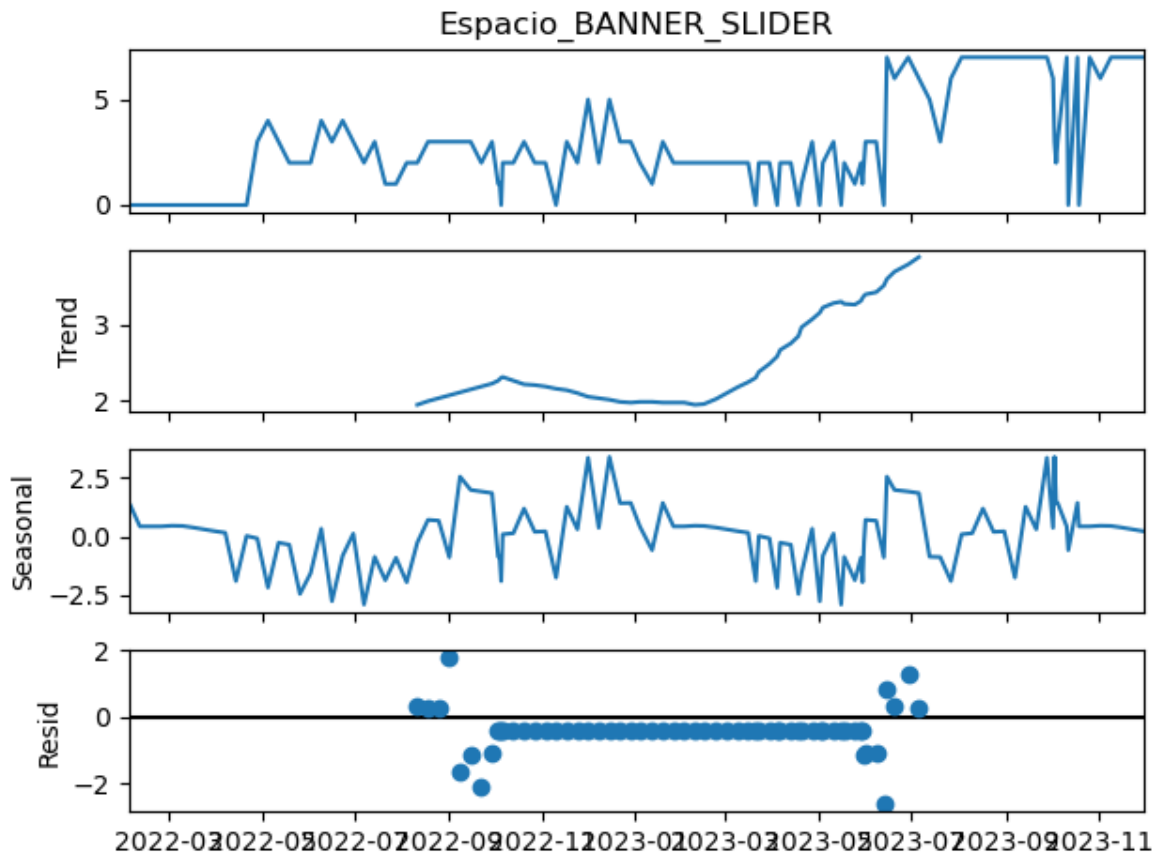


Ilustración 11: Descomposición de la serie temporal

3. Selección automática de parámetros ARIMA: Utiliza 'auto\_arima' de la biblioteca 'pmdarima' para seleccionar automáticamente los mejores parámetros ARIMA para el modelo. La selección de los parámetros óptimos fue 'order=(2, 1, 1)' mencionado anteriormente.
4. Ajuste del modelo ARIMA: Crea y ajusta un modelo ARIMA con los parámetros seleccionados.
5. Predicción para diciembre de 2023: Utiliza el modelo ajustado para predecir las próximas 4 semanas (diciembre de 2023) y muestra un gráfico con los datos históricos y las predicciones.
6. Imprime las predicciones: Muestra las predicciones para las 4 semanas de diciembre de 2023 en la consola.

### Predicción de demanda de Banners Sliders para diciembre de 2023

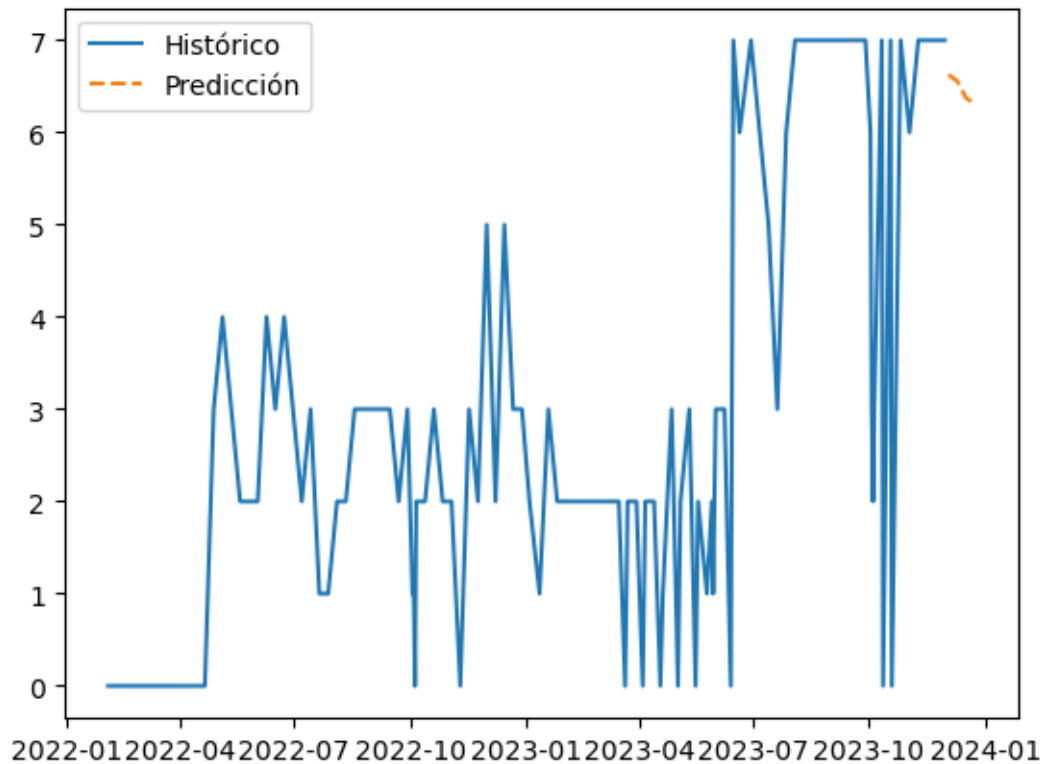


Ilustración 12: Predicción de demanda Banner slider

Semana DIC	Banner Slider
1	6.6269
2	6.556
3	6.3749
4	6.3086

Tabla 9: Predicciones Banner slider.

En resumen, el modelo ARIMA utilizado tiene un componente autorregresivo (AR) de orden 2, un componente de diferenciación (I) de orden 1 y un componente de media móvil (MA) de orden 1. Este modelo se ha ajustado a los datos históricos para realizar predicciones futuras. El procedimiento anterior se realizará para cada una de las variables.

## 2. Método Holt-Winter

En este modelo como ya vimos antes, es crucial tener en cuenta la tendencia, estacionalidad y nivel. En particular, nos enfocamos en la estacionalidad aditiva porque, cuando la magnitud de la estacionalidad no experimenta variaciones en relación con el nivel de la serie temporal. En otras

palabras, la estacionalidad influye de manera constante en la serie sin importar el nivel que tenga en un momento dado.

Proceso de modelado en código:

1. se realiza una serie temporal como ilustración 10.
2. Ajuste del modelo Holt-winter: se utiliza la clase 'ExponentialSmoothing' de 'statsmodels.tsa.holtwinters' para crear y ajustar el modelo de Holt-Winters a los datos históricos. En este caso, se especifica 'trend=add' para una tendencia aditiva, 'seasonal=add' para una estacionalidad aditiva y 'seasonal\_periods=52' para un período estacional de 52 semanas.
3. Predicción para diciembre de 2023: El modelo ajustado se utiliza para predecir las próximas cuatro semanas (diciembre de 2023).
4. Visualización de datos históricos y predicciones: Se genera un gráfico que muestra los datos históricos junto con las predicciones para visualizar la calidad del ajuste.

Predicción de demanda de Banners Sliders para diciembre de 2023 (Holt-Winters)

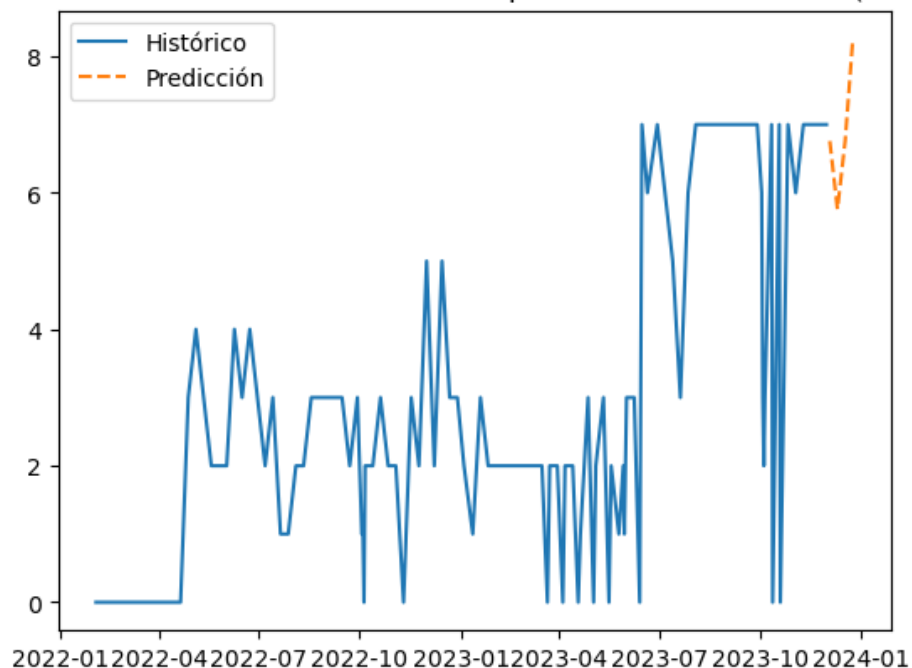


Ilustración 13: Predicción demanda Banner slider Holt-Winter

Semana DIC	Banner Slider
1	6.7588
2	5.7288
3	6.7588
4	8.2589

Tabla 10: Predicción número HLT

En resumen, el código implementa el modelo de Holt-Winters para realizar pronósticos de la serie temporal, considerando tendencia y estacionalidad. El ajuste del modelo y la generación de predicciones se llevan a cabo utilizando las funcionalidades proporcionadas por la biblioteca 'statsmodels.tsa.holtwinters'

### 3. Método Prophet Facebook

Gracias a su fácil uso en este modelo no hay la necesidad de ajustar parámetros, dado que el mismo modela sus parámetros.

El proceso que se utilizó para desarrollar el código fue:

1. Primero se convierten las columnas según el formato requerido por prophet (columnas "ds" para las fechas y "y" para los valores a predecir)
2. Se realiza la serie temporal para un mejor entendimiento
3. Se ajusta el modelo mediante la instalación de la clase 'Prophet()' y el método 'fit()'
4. Se crea un dataframe futuro con las predicciones
5. Se realiza la predicción

Predicción de demanda de Banners Sliders para diciembre de 2023 (Prophet)

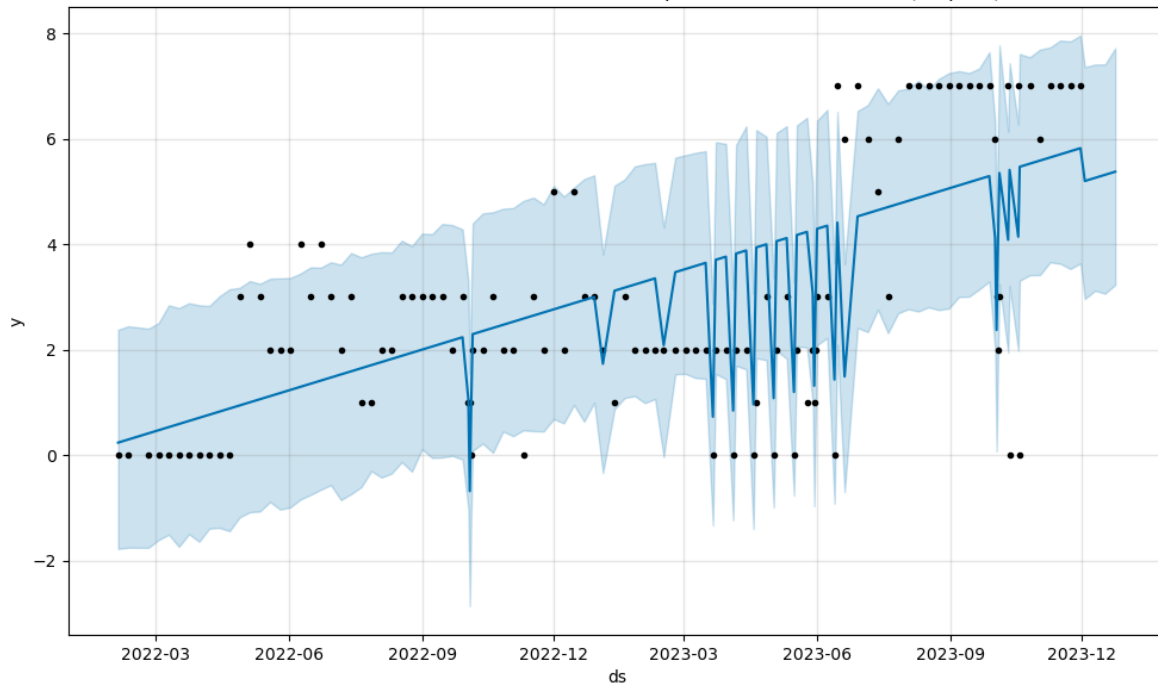


Ilustración 14: Predicción Prophet

Semana DIC	Banner Slider
1	5.197
2	5.2567
3	5.3155
4	5.3743

Tabla 11: Predicción numérica Prophet.

Este modelo es relativamente sencillo comparado en comparación con los otros dos anteriores, haciendo que Prophet sea una opción accesible para pronósticos de series temporales.

#### 4. Desarrollo de petición de clientes que tendrán mayor inversión

El código de predicción para esta petición utiliza 'RandomForestRegressor' como modelo de regresión para hacer predicciones sobre la inversión final de los clientes en el año 2024. Aquí hay algunas razones por las que podría preferirse 'RandomForestRegressor' en este contexto y luego una explicación del proceso general del código Random Forest Regressor:

1. Ensemble Learning: 'RandomForestRegressor' es un modelo de ensamblado que combina múltiples árboles de decisión para mejorar la precisión y generalización del modelo. Al



entrenar varios árboles y promediar sus predicciones, se reduce el riesgo de sobreajuste a los datos de entrenamiento.

2. Manejo automático de variables: Random Forest maneja automáticamente la importancia de las variables, lo que significa que no es necesario escalar o normalizar las características antes de entrenar el modelo.
3. Robustez: Es robusto frente a datos atípicos y no sensibles a pequeñas variaciones en los datos.
4. Manejo de características categóricas: Random Forest puede manejar naturalmente variables categóricas sin necesidad de codificación especial.

#### Proceso del Código:

1. Manipulación de fechas: Las columnas de fecha se convierten al formato adecuado (datetime) y se extraen los años y meses correspondientes.
2. Codificación de variables categóricas: Se utiliza 'LabelEncoder' para convertir la variable categórica 'Cliente' en una variable numérica ('Cliente\_encoded').
3. Selección de características y objetivo: Se eligen las características ('Cliente\_encoded' y 'Mes') y la variable objetivo ('Inversión final').
4. Filtrado de datos: Se filtran los datos para incluir solo los años 2022 y 2023.
5. Entrenamiento del modelo: Se utiliza un modelo 'RandomForestRegressor' para entrenar el modelo utilizando las características seleccionadas y la variable objetivo.
6. Generación de predicciones: Se generan predicciones para cada cliente y mes en el año 2024.

#### Resultados

Al asignarnos el mes de diciembre como periodo de prueba solo para LIDERAPP, dado que hasta el momento este mes solo registraba una venta de alrededor de 110.950.800 millones en LIDERAPP, con una meta establecida de 340.888.398. Es crucial resaltar que se nos otorgó el mes de noviembre para incrementar las ventas del mes de diciembre mediante al proyecto. Es por esta razón que, considerando las proyecciones derivadas de los tres modelos presentados, donde el estudiante asumió la responsabilidad de exhibir y explicar dichas predicciones, junto con una anticipación de los cinco principales futuros inversores para diciembre, los Key Account Managers (KAM) se unieron

para elaborar un plan de ventas. Este plan prioriza la comercialización de los espacios más costosos a los clientes que se pronostica realizarán mayores inversiones y así conseguir un aumento al actual monto de ventas.

### Resultados de modelos en LIDERAPP

#### 1. ARIMA:

Espacio/semana de diciembre	49	50	51	52
caluga	0	0	0	0
Banner slider	6.62	6.55	6.37	6.3
Mini mosaico	0.87	5.82	5.29	1.53
Mosaico carrusel	1.34	2.68	4.56	6.45
Search	47.24	8.9	26.28	23.25

Tabla 12: Predicciones de espacios LIDERAPP con ARIMA.

#### 2. Holt-Wintwer

Espacio/semana de diciembre	49	50	51	52
caluga	0	0	0	0
Banner slider	6.75	5.72	6.75	8.25
Mini mosaico	3.63	0.031	1.42	1.82
Mosaico carrusel	2.69	2.19	4.19	4.19
Search	28.95	26.87	25.76	26.13

Tabla 13: Predicciones de espacios LIDERAPP con Holt-Winter.

#### 3. Prophet:

Espacio/semana de diciembre	49	50	51	52
caluga	0	0	0	0
Banner slider	5.19	5.25	5.31	5.37
Mini mosaico	5.27	5.31	5.35	5.39
Mosaico carrusel	3.94	2.87	5.92	4.98
Search	23.3	23.57	23.84	24.10

Tabla 14: Predicciones de espacios LIDERAPP con Prophet.

## 4. RandomForestRegressor:

Lugar de mejor Inversión	Cliente
1	Softys
2	Unilever
3	Carozzi
4	P&G
5	Nestle

*Tabla 15: Predicciones los mejores inversores para el mes de diciembre*

Estos son los entregables que se anticipaban proporcionar al equipo de ventas de Walmart Connect para facilitar una toma de decisiones rápida respecto a las prioridades de ventas en el equipo. Además, se explorarán nuevas oportunidades de ventas, ya que las predicciones entregadas indican que aún hay margen para aumentar las ventas.

## Resultados de cierre de año

Cientes	Ventas finales en de diciembre CLP
Carozzi	CLP23,403,968
Bimbo	CLP9,313,689
softys	CLP26,413,388
VSPT	CLP3,600,000
ABINBEV	CLP13,789,196
AGROCOMMERCE	CLP844,694
Agrosuper	CLP3,840,000
beiersdorf	CLP5,900,000
Biotec	CLP1,000,000
Cerveceria CCU	CLP9,900,000
Cocacola andina	CLP3,240,000
Colun	CLP1,920,000
Demaria	CLP2,496,000
Diageo	CLP10,900,000
Duracel	CLP680,000
EDGWELL	CLP1,200,000
EMBOTELLADORAS CHILENAS UNIDAS (CCU)	CLP1,200,000
Iansa	CLP2,040,000
Good Food	CLP720,000
J&J	CLP799,998
Kraft	CLP4,001,694
Mars	CLP1,920,000
Mondelez	CLP17,390,480
Nestle	CLP38,711,616
Pernod Ricard	CLP1,600,000
PEPSICO	CLP9,098,455
P&G	CLP21,326,479
PF	CLP8,359,936
SADIA	CLP7,472,064
Reckitt	CLP7,420,000
SC johnson	CLP8,386,946
Unilever	CLP33,024,637
Soprole	CLP8,160,000
TMLUC	CLP2,400,000
Vivafoods	CLP936,000
Watts	CLP1,040,000
TOTAL	CLP294,449,240

Tabla 16: Resultado diciembre con el fin de año.

Cientes	Vendido Inicialmente en CLP
Carozzi	CLP10,000,000
Bimbo	CLP2,500,000
softys	CLP13,413,388
VSPT	CLP0
ABINBEV	CLP0
AGROCOMMERCE	CLP0
Agrosuper	CLP0
beiersdorf	CLP5,900,000
Biotec	CLP1,000,000
Cerveceria CCU	CLP4,000,000
Cocacola andina	CLP3,240,000
Colun	CLP1,920,000
Demaria	CLP0
Diageo	CLP10,900,000
Duracel	CLP0
EDGWELL	CLP1,200,000
EMBOTELLADORAS CHILENAS UNIDAS (CCU)	CLP1,200,000
Iansa	CLP1,200,000
Good Food	CLP720,000
J&J	CLP799,998
Kraft	CLP3,021,414
Mars	CLP0
Mondelez	CLP0
Nestle	CLP14,600,000
Pernod Ricard	CLP0
PEPSICO	CLP4,500,000
P&G	CLP4,000,000
PF	CLP0
SADIA	CLP0
Reckitt	CLP0
SC johnson	CLP0
Unilever	CLP15,500,000
Soprole	CLP8,160,000
TMLUC	CLP1,200,000
Vivafoods	CLP936,000
Watts	CLP1,040,000
TOTAL	CLP110,950,800

Tabla 17: Ventas de diciembre antes del mes de prueba.

Cientes	Incremental de ventas
Carozzi	134.0%
Bimbo	272.5%
softys	96.9%
VSPT	
ABINBEV	
AGROCOMMERCE	
Agrosuper	
beiersdorf	0.0%
Biotec	0.0%
Cerveceria CCU	147.5%
Cocacola andina	0.0%
Colun	0.0%
Demaria	
Diageo	0.0%
Duracel	
EDGWELL	0.0%
EMBOTELLADORAS CHILENAS UNIDAS (CCU)	0.0%
Iansa	70.0%
Good Food	0.0%
J&J	0.0%
Kraft	32.4%
Mars	
Mondelez	
Nestle	165.1%
Pernod Ricard	
PEPSICO	102.2%
P&G	433.2%
PF	
SADIA	
Reckitt	
SC johnson	
Unilever	113.1%
Soprole	0.0%
TMLUC	100.0%
Vivafoods	0.0%
Watts	0.0%
TOTAL	167.1%

Tabla 18: Incremental de ventas de diciembre.

Espacio/semana de diciembre	de 49	50	51	52	Tasa de Ocupación. de
caluga	0	0	0	0	
Banner slider	7	7	7	4	89%
Mini mosaico	1	5	6	1	93%
Mosaico carrusel	1	2	5	6	93%
Search	51	4	25	23	98%

*Ecuación 5: Tasa de Ocupación de espacios.*

Basándonos en las proyecciones, el equipo tomó la decisión conjunta de solicitar espacios equivalentes a la cantidad prevista por el modelo ARIMA, redondeando las cifras. Posteriormente, optaron por presentar una oferta a los proveedores de manera conjunta, ya que durante el proceso se ha llevado a cabo una significativa reducción de personal, lo que impide la medición actual de la medida de desempeño 2. Esta medición será posible a partir del próximo año, una vez que se hayan redistribuido las carteras de los Key Account Managers (KAM). Una vez determinada la cantidad de espacios, el equipo salió a ofrecerlos a los clientes durante el mes de noviembre, con la intención de incentivar inversiones para el mes de diciembre. Al concluir el mes de noviembre, se cerró el periodo de diciembre, marcando así el cierre del año con los resultados detallados en las tablas. Las métricas de rendimiento que pudieron ser evaluadas incluyen el Incremento de Ventas, que logró aumentar las ventas de diciembre en un 167.1%, situando el total del mes en 294.449.240 millones de pesos y mantener una tasa de ocupación de espacios en 90%.

Puesto	Mejores Inversores diciembre
1	Nestle
2	Unilever
3	softys
4	Carozzi
5	P&G

*Tabla 19: Mejores Inversores del mes de diciembre.*

Observamos que las predicciones de los inversionistas más destacados no mostraron una gran discrepancia en cuanto a la identificación de aquellos que realizarían inversiones más significativas.

Finalmente, tras revisar con el equipo los resultados de los modelos, y la métrica MAE se llegó al acuerdo de usar el modelo ARIMA para presentar las predicciones para SOD y LIDERAPP para el año 2024.

Errores de pronósticos (Mes de diciembre)	ARIMA	Prophet	Holt-Winter
Banner slider	0.84	2.42	-2.92
Mini mosaico	0.51	-6.099	8.32
Mosaico carrusel	1.03	-0.74	3.71
Search	2.67	4.71	-8.19

Tabla 20: Errores de pronóstico de las predicciones.

Métricas/Modelos	ARIMA	Holt-Winter	Pophet
MAE	1.2625	3.49225	5.785

Tabla 21: Error absoluto medio de las predicciones.

## Conclusiones

En este proyecto, se evidenció que la gestión de la toma de decisiones por parte de un equipo de trabajo es más eficaz cuando no existe incertidumbre sobre eventos futuros. Concretamente, se demostró que, en este caso particular, el equipo optó por no cerrar el año de manera prematura y se dedicó a explorar nuevas oportunidades de venta de espacios publicitarios digitales. Esta estrategia resultó en un notable aumento del 167% en las ventas, culminando el mes con un cumplimiento del 86%.

Inicialmente el proyecto fue pensado para su implementación en las futuras propuestas del año 2024, este proyecto reservó el mes de diciembre como período de prueba y evaluación de resultados. Su éxito, según la percepción del equipo, da paso a la entrega de predicciones mensuales para el año 2024. Estas predicciones serán proporcionadas para facilitar la toma de decisiones del equipo de ventas en relación con sus propuestas a los clientes.

## Discusión

Este proyecto tiene potencial para mejoras significativas en el futuro al incorporar otras predicciones, como identificar las categorías predominantes en un mes o determinar los momentos del año más propicios para que un cliente invierta en espacios publicitarios digitales. En última instancia, la expansión de este proyecto puede abarcar otras áreas y proyectos dentro de Walmart,




como el proyecto Instore que se está llevando a cabo este año, centrado en la venta de publicidad física en tiendas.

### Referencias bibliográficas

1. SHIRLY NOWAK (diciembre, 2022). Modelos predictivos ¿Qué son y para que se usan? [Modelos predictivos ¿Qué son y para qué se usan? - Nuclio School](#)
2. Parra (enero, 2019). Estadísticas y machine learning con R. [Estadística y Machine Learning con R \(bookdown.org\)](#)
3. C.López (2014). Sistemas de apoyo a la toma de decisiones. Una aplicación en el área de Gestión Universitaria. [Documento completo.pdf \(unlp.edu.ar\)](#)
4. Decision support system. [DSS-2\(mod-30\).ppt \(live.com\)](#)
5. Peña (Abril, 2018). Implementación de una herramienta de Business Intelligence para mejorar el proceso de toma de decisiones financieras – Área de ventas – en Eknowledge Group S.A.C. [335023364.pdf \(core.ac.uk\)](#)
6. Díaz (2020). Diseñar e implementar modelos de predicción de demanda de clientes para las redes sociales. [TD000181.pdf \(udp.cl\)](#)
7. Hotz (enero, 2023). What is CRISP DM? [What is CRISP DM? - Data Science Process Alliance \(datascience-pm.com\)](#)
8. Castellón (2023). [Serie-de-tiempo-con-Machine-Learning/Modulo\\_20\\_Comparanod Modelos arima lstm Random forest prophet.ipynb at main · Naren8520/Serie-de-tiempo-con-Machine-Learning · GitHub](#)
9. Amat (Octubre, 2020). [Random Forest python \(cienciadedatos.net\)](#)
10. Salario promedio chile. [Salario para Ingeniero Civil Industrial en Chile - Salario Medio \(talent.com\)](#)

### Anexos

#### 1. Vista de la plataforma Workfront

Solicitud Masiva  Pin current page

REPORT **Espacios RESERVADO y VENDIDO (Sales)** Show Filters Report Actions As of Dec 5, 2023 11:10 pm hora de verano de Chile

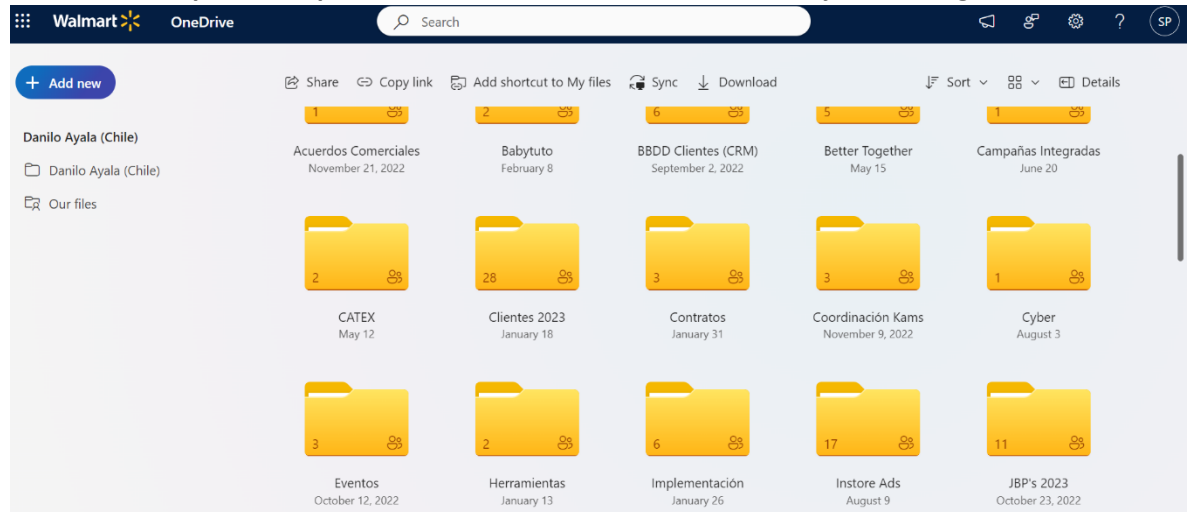
Details Summary

→

Name	Deadline	Fecha	ESTADO	Lista	Mes	Semana	Desde ↑	Hasta	KAM	Segmento	Espacio	Slot	Ubicación	RUT Cliente	Client
Semana - Connect: 46 (93)															
Semana - Connect: 49 (165)															
<input type="checkbox"/> SOD_BANNER_CATEGORIA_1_49	11/23/23	5/10/23	VENDIDO	L12	12	49	12/7/23	12/13/23	NICO	SOD	CATEGORIA	BANNER_CATEGORIA	1	SOLICITAR RUT	GOOC S A
<input type="checkbox"/> SOD_BANNER_CATEGORIA_1_49	11/23/23	10/16/23	VENDIDO	L12	12	49	12/7/23	12/13/23	NICO	SOD	CATEGORIA	BANNER_CATEGORIA	1	96720910	MONI
<input type="checkbox"/> SOD_BANNER_CATEGORIA_1_49	11/23/23	10/18/23	VENDIDO	L12	12	49	12/7/23	12/13/23	VALE	SOD	CATEGORIA	BANNER_CATEGORIA	1	96.894.370-4	SADIA S A

Showing All (445) projects 1 - 445 of 445 projects

## 2. Vista de las carpetas compartidas de Walmart connect donde se dejará el código.



## 3. Código de transformación de datos

```

1 import pandas as pd
2 import numpy as np
3 import matplotlib.pyplot as plt
4 from statsmodels.tsa.arima.model import ARIMA
5 from statsmodels.tsa.statespace.sarimax import SARIMAX
6 from statsmodels.tsa.seasonal import seasonal_decompose
7
8
9
10 # In[48]:
11
12
13 df = pd.read_excel("C:\\Users\\cmpm6\\OneDrive\\Escritorio\\LIDERAPP.xlsx")
14 df = df.dropna()
15 # Agrupa por 'Desde' y ' SLOT Homologado' y cuenta las ocurrencias
16 grouped_df = df.groupby(['Desde', ' SLOT Homologado']).size().reset_index(name='conteo')
17
18 # Pivotar para tener ' SLOT Homologado' como columnas
19 pivoted_df = grouped_df.pivot(index='Desde', columns=' SLOT Homologado', values='conteo').fillna(0)
20
21 # Renombrar las columnas para que sea más legible
22 pivoted_df.columns.name = None # Eliminar el nombre de las columnas
23 pivoted_df.reset_index(inplace=True) # Resetear el índice para obtener 'Desde' como una columna
24
25 # Ahora puedes acceder a las columnas 'Desde' y ' SLOT Homologado' según tus necesidades
26 nueva_data = pivoted_df[['Desde'] + list(pivoted_df.columns[1:])]
27
28 # Renombrar las columnas según tus especificaciones
29 nueva_data.columns = ['Semanas'] + [f'Espacio_{col}' for col in nueva_data.columns[1:]]
30
31 # Mostrar la nueva data
32 print(nueva_data)
33
34 # Lista para almacenar los DataFrames divididos
35 dataframes_divididos = []

```

```
37 # Iterar sobre las columnas de Espacio_X
38 for columna in nueva_data.columns[1:]:
39     # Crear un nuevo DataFrame con 'Semanas' y la columna actual
40     nuevo_dataframe = nueva_data[['Semanas', columna]].copy()
41
42     # Renombrar la columna 'SLOT Homologado' según tus necesidades
43     nuevo_dataframe.columns = ['Semanas', columna]
44
45     # Añadir el nuevo DataFrame a la lista
46     dataframes_divididos.append(nuevo_dataframe)
47
48 # Acceder a los DataFrames individuales
49 for df_individual in dataframes_divididos:
50     nombre_columna = df_individual.columns[1]
51     print(f"DataFrame '{nombre_columna}':\n{df_individual}\n")
52
```

#### 4. Código Modelo ARIMA

```
# Convierte la columna 'Semanas' a tipo datetime
df1['Semanas'] = pd.to_datetime(df1['Semanas'])

# Establece 'Semanas' como índice
df1.set_index('Semanas', inplace=True)

# Gráfico de la serie temporal
plt.plot(df1['Espacio_BANNER_SLIDER'])
plt.title('Serie Temporal de Espacio Mini msaico')
plt.show()

from statsmodels.tsa.seasonal import seasonal_decompose

result = seasonal_decompose(df1['Espacio_BANNER_SLIDER'], model='additive', period=52)
result.plot()
plt.show()

from pmdarima import auto_arima

# Encuentra automáticamente los parámetros ARIMA
arima_model = auto_arima(df1['Espacio_BANNER_SLIDER'], suppress_warnings=True)
print(arima_model.summary())

# Ajusta el modelo ARIMA
model = ARIMA(df1['Espacio_BANNER_SLIDER'], order=(2, 1, 1))
fitted_model = model.fit()

# Predice para diciembre de 2023
forecast = fitted_model.forecast(steps=4) # Cambiado a 4 para predecir 4 semanas de diciembre

# Gráfico de datos históricos y predicciones
plt.plot(df1.index, df1['Espacio_BANNER_SLIDER'], label='Histórico')
plt.plot(pd.date_range(start='2023-12-01', periods=4, freq='W'), forecast, label='Predicción', linestyle='dashed')
plt.legend()
plt.title('Predicción de demanda de Espacio_BANNER_SLIDER para diciembre de 2023')
plt.show()

# Imprime las predicciones para diciembre de 2023
print("Predicciones para diciembre de 2023:")
print(forecast)
```

## 5. Código Modelo Prophet

```
import pandas as pd
import matplotlib.pyplot as plt
from prophet import Prophet

# Convierte la columna 'Semanas' a tipo datetime
df1['Semanas'] = pd.to_datetime(df1['Semanas'])

# Renombra las columnas según el formato requerido por Prophet
df_prophet = df1.rename(columns={'Semanas': 'ds', 'Espacio_BANNER_SLIDER': 'y'})

# Gráfico de la serie temporal
plt.plot(df_prophet['ds'], df_prophet['y'])
plt.title('Serie Temporal de Espacio Banner slider')
plt.show()

# Ajuste del modelo Prophet
model_prophet = Prophet()
model_prophet.fit(df_prophet)

# Crear un dataframe para las fechas futuras
future = model_prophet.make_future_dataframe(periods=4, freq='W')

# Realizar la predicción
forecast_prophet = model_prophet.predict(future)

# Gráfico de datos históricos y predicciones
fig = model_prophet.plot(forecast_prophet)
plt.title('Predicción de demanda de Banners Sliders para diciembre de 2023 (Prophet)')
plt.show()

# Imprime las predicciones para diciembre de 2023
print("Predicciones para diciembre de 2023 (Prophet):")
print(forecast_prophet[['ds', 'yhat']].tail(4))
```

## 6. Código Modelo Holt-Winter

```
import pandas as pd
import matplotlib.pyplot as plt
from statsmodels.tsa.holtwinters import ExponentialSmoothing

# Convierte la columna 'Semanas' a tipo datetime
df1['Semanas'] = pd.to_datetime(df1['Semanas'])

# Establece 'Semanas' como índice
df1.set_index('Semanas', inplace=True)

# Gráfico de la serie temporal
plt.plot(df1['Espacio_BANNER_SLIDER'])
plt.title('Serie Temporal de Espacio Banner slider')
plt.show()

# Ajuste del modelo Holt-Winters
model_hw = ExponentialSmoothing(df1['Espacio_BANNER_SLIDER'], trend='add', seasonal='add', seasonal_periods=52)
fitted_model_hw = model_hw.fit()

# Predicción para diciembre de 2023
forecast_hw = fitted_model_hw.forecast(steps=4) # Cambiado a 4 para predecir 4 semanas de diciembre

# Gráfico de datos históricos y predicciones
plt.plot(df1.index, df1['Espacio_BANNER_SLIDER'], label='Histórico')
plt.plot(pd.date_range(start='2023-12-01', periods=4, freq='W'), forecast_hw, label='Predicción', linestyle='dashed')
plt.legend()
plt.title('Predicción de demanda de Banners Sliders para diciembre de 2023 (Holt-Winters)')
plt.show()

# Imprime las predicciones para diciembre de 2023
print("Predicciones para diciembre de 2023 (Holt-Winters):")
print(forecast_hw)
```

## 7. Predicciones para 2024 de los 5 mejores inversores de cada mes.

Cliente	Mes
LICORES MITJANS S.A	1
LOREAL	1
UNILEVER	1
MARS	1
SOFTYS CHILE SPA	1
SOFTYS CHILE SPA	2
COLGATE	2
COCACOLA ANDINA	2
COCACOLA	2
CERVECERA CCU CHILE LTDA (CCU)	2
COLGATE	3
COCACOLA ANDINA	3
SOFTYS CHILE SPA	3
CERVECERA CCU CHILE LTDA (CCU)	3
COCACOLA	3
SOFTYS CHILE SPA	4
SOFTYS	4

SOCIEDAD PUNTA DE LOBOS S A	4
SOFRUCO	4
BIOTEC CHILE SA	4
TMLUC	5
TRENDY	5
SOFTYS	5
CAROZZI	5
CARNES NUBLE S A	5
SOFTYS	6
Pernod ricard	6
Pepsico	6
GOOD FOOD S A	6
KRAFT	6
Pernod ricard	7
Pepsico	7
Proteina Spa	7
GILCO	7
Minuto Verde	7
Pepsico	8
Pernod ricard	8
Proteina Spa	8
PF	8
SC JOHNSON	8
ABINBEV	9
ACONCAGUA FOODS	9
CERVECERA CCU CHILE LTDA (CCU)	9
KRAFT HEINZ CHILE LIMITADA	9
SC JOHNSON	9
CAROZZI	10
BIMBO	10
CARNES NUBLE S A	10
BIOTEC CHILE SA	10
SOFTYS	10
BIMBO	11
SOCIEDAD PUNTA DE LOBOS S A	11
SC JOHNSON	11
SOC ANONIMA VINA SANTA RIT	11

COLGATE PALMOLIVE CHILE SXAX	11
COCACOLA	12
COMERCIAL Y SERVICIO CHILEEMPR	12
COMERSUR S.A	12
COSMETICA NACIONAL SA	12
COMPANIA PISQUERA DE CHILE S A (CCU)	12