

**Desarrollo de una herramienta de pronóstico de demanda para la optimización de inventarios**

Harold Andres Velasco Martínez

Asesor

Andres Felipe Solís Pino

Universidad Nacional Abierta y a Distancia UNAD  
Escuela de Ciencias Básicas, Tecnología e Ingeniería ECBTI  
Ciencia de Datos y Analítica  
2025

**Nota de Aceptacion**

Rafael Roberto Ruiz Escorcía

Nombre Director de Trabajo de Grado

Andrés Felipe Hernández Giraldo

Jurado

---

Jurado

## Resumen

Este proyecto se centra en el desarrollo de una herramienta de pronóstico de demanda para la optimización de inventarios en la empresa Medicox LTDA. La propuesta metodológica incluye el análisis y procesamiento de datos, la selección y construcción de modelos de Machine Learning, y la evaluación de los modelos para su implementación. La herramienta permitirá tomar decisiones basadas en datos para reducir el inventario crítico y evitar pérdidas por vencimiento de productos, específicamente para el producto Westcohex 270 x 30 ML.

**Palabras clave:** Pronóstico de demanda, Optimización de inventarios, Machine Learning, Reducción de inventario crítico, Gestión de inventarios

### **Abstract**

This project focuses on developing a demand forecasting tool for inventory optimization at Medicox LTDA. The methodological proposal includes data analysis and processing, selection and construction of machine learning models, and model evaluation for implementation. The tool will enable data-driven decision-making to reduce critical inventory and prevent product expiration losses, specifically for the product Westcohex 270 x 30 ML.

**Keywords:** Demand Forecasting, Inventory Optimization, Machine Learning, Critical Inventory Reduction, Inventory Management

## Contenido

Introducción .....	8
Justificación .....	9
Objetivos.....	10
Objetivo General.....	10
Objetivos Específicos.....	10
Marco de Referencia .....	11
Propuesta Metodológica.....	12
Establecer Problemática y Objetivos .....	12
Selección, Procesamiento y Análisis de la Información (ETL – EDA).....	12
Selección y Construcción de los Modelos .....	12
Evaluación de Modelo .....	12
Desarrollo e Implementación Herramienta de Consulta.....	12
Consolidación y Transformación de Información .....	13
Recopilación de Datos .....	13
Integración de Datos .....	13
Transformación de Información .....	13
Extracción, Transformación y Carga (ETL).....	13
Extracción.....	14
Transformación.....	15
Descripción y Selección de Valores .....	19
Análisis Exploratorio de Datos (EDA) .....	21
Definición del Modelo de Proyección .....	24

Selección de Modelos de Machine Learning .....	24
Revisión de Literatura.....	24
Elección de Métrica de Evaluación .....	26
Entrenamiento y Validación del Modelo .....	27
Preparación de Datos .....	27
Entrenamiento del Modelo.....	27
Validación del Modelo.....	28
Implementación del Modelo Final .....	28
Implementación de la Herramienta.....	30
Desarrollo de la Interfaz de Usuario .....	30
Integración con Sistemas Existentes .....	31
Flujograma de creación e implementación .....	32
Capacitación y Soporte a Usuarios .....	33
Capacitación de Usuarios.....	33
Soporte Continuo .....	34
Monitoreo y Evaluación.....	34
Monitoreo del Rendimiento del Modelo.....	34
Conclusiones.....	35
Recomendaciones .....	37
Referencias Bibliográficas .....	39

## Lista de Figuras

<b>Figura 1</b> <i>Serie de Tiempo – Cantidad Vendida</i> .....	15
<b>Figura 2</b> <i>Serie de Tiempo – Normalización</i> .....	20
<b>Figura 3</b> <i>Serie de Tiempo – EDA</i> .....	22
<b>Figura 4</b> <i>Métricas MAE, MSE y R2 para cada Normalización del Modelo</i> .....	26
<b>Figura 5</b> <i>Predicciones del Modelo</i> .....	28
<b>Figura 6</b> <i>Pronostico - Power Query</i> .....	29
<b>Figura 7</b> <i>Reporte - Power BI</i> .....	30
<b>Figura 8</b> <i>Flujo Power Automate</i> .....	32
<b>Figura 9</b> <i>Flujo de Implementacion</i> .....	33

## Introducción

En la actualidad, la gestión eficiente de inventarios es crucial para la supervivencia y el crecimiento de las empresas, especialmente en sectores altamente sensibles como el de dispositivos médicos y medicamentos. Medicox LTDA, una empresa con 18 años de experiencia en la comercialización de dispositivos medico quirúrgicos y medicamentos, enfrenta desafíos en la optimización de su inventario debido a la alta sensibilidad de sus productos y las restricciones de fechas de vencimiento.

La toma de decisiones informadas basadas en datos es esencial para reducir costos y aumentar la eficiencia operativa. Sin embargo, la falta de herramientas adecuadas para respaldar estas decisiones ha llevado a una baja rotación del inventario y un aumento de pérdidas por vencimiento de productos. Actualmente, el 25% del inventario de Medicox LTDA se clasifica como de rotación baja y crítica (más de 120 días).

Este proyecto propone el desarrollo de una herramienta de pronóstico de demanda utilizando técnicas de Machine Learning para la optimización de inventarios. La implementación de esta herramienta permitirá a Medicox LTDA tomar decisiones más acertadas sobre la cantidad de producto a solicitar, reduciendo el inventario crítico e incentivando la alta rotación de productos. El modelo se centrará en el producto Westcohex 270 x 30 ML y buscará replicar los resultados en otras referencias con el objetivo de mejorar la gestión de inventarios y prevenir pérdidas por vencimiento.

## **Justificación**

Toda gran conclusión es generada bajo la recolección de información, hoy en día, gracias a los avances tecnológicos y su gran dependencia de las compañías por esta, se captura mucha más información. Gracias a una de las herramientas recientemente creadas, el Machine Learning, es posible obtener un mayor beneficio de los datos, gracias a su versatilidad y procesamiento de información es posible identificar variables que a simple vista no son percibibles y la relación entre ellas.

Una de sus muchas aplicaciones en la proyección de demanda, la cual busca mediante la información ya existente y el relacionamiento de variables tales como información de ventas, compras e inventario, prevenir o anticiparse a escenarios para tomar una decisión, en esta ocasión para la empresa Medicox LTDA, se busca entender el relacionamiento de estas variables para tomar decisiones a la compra de producto, evitando tomar decisiones que no estén basadas en información y en busca de prevenir el aumento de inventario crítico que incentiven a pérdidas por vencimiento de inventario.

Es así como en el presente trabajo con la construcción de una herramienta busca aportar al inicio de la cadena de abastecimiento, al proceso de compras en su toma de decisiones en la cantidad de producto a solicitar según las variables de inventario histórico de ventas e histórico de compras. Se seleccionará un producto según su rotación y costo y se buscará entender la relación de las variables para futuros procesos de compra, para finalmente replicarlo en las demás referencias que puedan ejercer un impacto significativo en la rotación de producto crítico y pérdidas por vencimiento.

## **Objetivos**

### **Objetivo General**

Desarrollar una herramienta de pronóstico de demanda para la prevención en el aumento de los días de inventario para el producto Westcohex 270 x 30 ML, mediante técnicas de Machine Learning.

### **Objetivos Específicos**

Consolidar las fuentes de información necesarias para la creación de modelo de proyección de demanda para el producto Westcohex 270 x 30 ML.

Definir el modelo de proyección de demanda para el producto Westcohex 270 x 30 ML, a través de las métricas aceptadas en el corpus científico del dominio.

Implementación de la herramienta de pronóstico de demanda para el producto Westcohex 270 x 30 ML.

## Marco de Referencia

En el crecimiento de las empresas el tomar buenas decisiones alrededor de los gastos y costos son importantes para mantener e incentivar el crecimiento de la misma, y la gestión de inventarios, desde la creación de las empresas son el corazón para subsistir en el mercado. Para esto se ha utilizado varias herramientas y metodologías que permiten su buena administración (sistemas ERP – Metodologías ABC, MRP, ECQ etc.) (Hernández, 2024). Y sumando estas técnicas con el prever las ventas, brindan un plus a la hora de tomar decisiones de alto valor a las compañías.

El Machine Learning llegó para instalarse en nuestras vidas y en las compañías, y sumar esta herramienta a las anteriormente respalda las decisiones tomadas. Algunos de los beneficios que se han evidenciado tras su implementación es la mejor planificación de aprovisionamiento dado comportamientos de los clientes y sus ventas, como la reducción de errores en la generación de órdenes de compra a los proveedores. (Dinámica, 2022).

El desarrollo de modelos de Machine Learning existen artículos que sugieren el modelo a implementar para empresas pequeñas y medianas. Por un lado, Natthamonkan Kheawpeam - Sukree Sinthupinyo buscan realizar un pronóstico de demanda en periodos de 7 días y 30 días mediante los modelos de CatBoost y XGBoost, siendo las mejores opciones respecto a la comparación de modelos de regresión lineal. Por otro lado, E. Anitha - M. Navin Prasath - L. Sanjai - J. A. Andrea Shiny - P. Varsini proponen un modelo de Random Forest Regressor y ARIMA para la predicción de demanda. Puntos de partida para la elección de modelos para el proyecto en desarrollo.

## **Propuesta Metodológica**

### **Establecer Problemática y Objetivos**

Mediante la aplicación del análisis de datos concretar la problemática y delimitarla en base a los objetivos propuestos.

### **Selección, Procesamiento y Análisis de la Información (ETL – EDA)**

Seleccionar la información necesaria mediante el análisis de relación de las variables, procesar la información, reduciendo, eliminando y transformando la información (ETL) y por último obtener conclusiones de con la información resultante (EDA).

### **Selección y Construcción de los Modelos**

Selección de los modelos de machine Learning y establecer métricas de comparación entre ellos para la construcción del modelo final.

### **Evaluación de Modelo**

Evaluación del modelo propuesto mediante la curva de aprendizaje y entrenamiento y realizar ajuste del mismo para obtener mayor precisión.

### **Desarrollo e Implementación Herramienta de Consulta**

Selección del tipo de herramienta para el uso de la información del modelo, desarrollo e implantación para la medición y el impacto del proceso implementado

## **Consolidación y Transformación de Información**

La primera fase del proyecto consiste en la consolidación y transformación de la información necesaria para desarrollar el modelo de pronóstico de demanda. Esta fase es crucial, ya que una buena calidad de datos es fundamental para obtener resultados precisos y útiles en las etapas posteriores.

### **Recopilación de Datos**

**Fuentes Internas:** Se recopilan datos de ventas históricas para el modelamiento y entrenamiento del modelo, y compras e inventario para la comparación y toma de decisiones de la empresa Medicox LTDA. Estos datos proporcionan una visión integral del comportamiento de la demanda y las tendencias de inventario a lo largo del tiempo.

**Fuentes Externas:** No se obtienen información de datos externos, únicamente se basa la información en histórica de la compañía.

### **Integración de Datos**

**Herramientas Utilizadas:** Se utilizan herramientas de ciencia de datos como Python y Power BI para integrar y combinar los datos de diferentes fuentes. Python se emplea para la manipulación y procesamiento de grandes volúmenes de datos, mientras que Power BI facilita la visualización, implementación de la herramienta y el análisis interactivo de la información consolidada.

### **Transformación de Información**

#### ***Extracción, Transformación y Carga (ETL)***

La metodología de Extracción, Transformación y Carga (ETL) es fundamental para la integración y consolidación de datos en sistemas de inteligencia empresarial. Este proceso permite asegurar la calidad, coherencia y accesibilidad de los datos, facilitando su posterior

análisis y toma de decisiones. A continuación, se detallan las etapas del proceso ETL y su justificación:

### ***Extracción***

La fase de extracción implica obtener datos de diversas fuentes, que pueden incluir bases de datos relacionales, archivos planos, servicios web y otros sistemas de información.

Esta etapa es crucial porque asegura que todos los datos relevantes sean recopilados y centralizados. La heterogeneidad de las fuentes de datos requiere una estrategia de extracción robusta para garantizar la integridad y completitud de los datos (Kimball & Caserta, 2011).

Los datos recopilados se extraen de la base de datos del CRM de la compañía, y se almacenan en una estructura de DataFrame. Esta fase implica la lectura de la base de datos de ventas de la empresa. Se realiza una revisión de las variables importantes en la creación del modelo de pronóstico, se utilizan las variables de: Fecha, Referencia del Producto, Cantidad Vendida, Precio de Venta y Factor de Venta.

Estas variables fueron seleccionadas primeramente para el filtro y selección del producto evaluado y el procesamiento de datos o conversión de la unidad de empaque vendida y por otra parte buscando entender la variación cronológica del precio y el comportamiento de la venta de las unidades y así determinar su normalización o no.

```
conn = pyodbc.connect(conn_str)

cursor = conn.cursor()

query = "SELECT FACTURAS.FACT_FECHA, FACTURAS_DETALLE.ARTI_COD,
FACTURAS_DETALLE.FADE_CANT, FACTURAS_DETALLE.FADE_PRUNIT,
FACTURAS_DETALLE.FADE_FACTOR FROM FACTURAS_DETALLE LEFT JOIN
FACTURAS ON FACTURAS_DETALLE.FACT_ID = FACTURAS.FACT_ID WHERE
```

```
FACTURAS_DETALLE.ARTI_COD='PT0301004010' OR
```

```
FACTURAS_DETALLE.ARTI_COD='ELECT-121'"
```

```
cursor.execute(query)
```

```
tables_1 = cursor.fetchall()
```

```
columns = [column[0] for column in cursor.description]
```

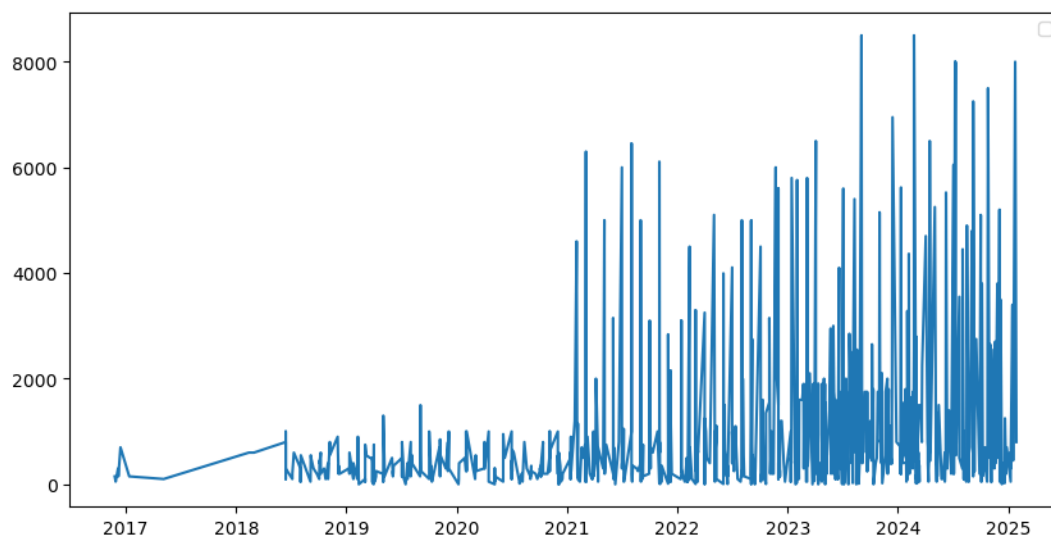
```
df_facturas_detalle =pd.DataFrame.from_records(tables_1,columns=columns)
```

```
# Cierra la conexión
```

```
conn.close()
```

## Figura 1

*Serie de Tiempo – Cantidad Vendida*



## Transformación

En la fase de transformación, los datos extraídos se limpian, enriquecen y convierten a un formato adecuado para su análisis. Esto puede incluir tareas como la eliminación de valores atípicos, la normalización de datos y la creación de variables derivadas.

La transformación de datos es esencial para mejorar la calidad de los mismos y adaptarlos a las necesidades específicas del análisis. Esta etapa permite resolver inconsistencias y asegurar que los datos sean compatibles y significativos (Inmon, 2005).

Los datos extraídos se transforman mediante técnicas de normalización, agrupación y cálculo de métricas relevantes. Se aplican técnicas de reducción de dimensionalidad para simplificar los datos y eliminar características irrelevantes o redundantes.

- **Métricas:** Se estandariza las unidades vendidas a unidad, entendiendo que este producto puede ser vendidos por cajas o paquetes.

```
# Transformación Tipo
df_facturas_detalle['FADE_CANT'] = df_facturas_detalle['FADE_CANT'].astype('int64')
df_facturas_detalle['FADE_PRUNIT'] =
df_facturas_detalle['FADE_PRUNIT'].astype('int64')
df_facturas_detalle['FADE_FACTOR'] =
df_facturas_detalle['FADE_FACTOR'].astype('int64')
# Estandarización
df_facturas_detalle['FADE_CANT'] =
df_facturas_detalle['FADE_CANT']/df_facturas_detalle['FADE_FACTOR']
df_facturas_detalle['FADE_PRUNIT'] =
df_facturas_detalle['FADE_PRUNIT']/df_facturas_detalle['FADE_FACTOR']
```

- **Agrupación:** Se realiza la agrupación de la cantidad vendida según la fecha, debido que en el contexto de la información se pueden realizar múltiples ventas del mismo producto en el mismo día, por ende, se resume (suma) la información por cantidad vendida por

fecha y se realiza la imputación de los datos para encontrar las fechas faltantes e incluirlas con valor 0.

```
# Agrupación
df=pd.DataFrame(df_facturas_detalle.groupby(['año','mes'])['FADE_CANT'].sum())
df=df.reset_index()

# Imputación
idx = pd.MultiIndex.from_product([df['año'].unique(), range(1, 13)], names=['año',
'mes'])

df = df.set_index(['año', 'mes']).reindex(idx).reset_index()

# Eliminación de vacíos
df['FADE_CANT'] = df['FADE_CANT'].fillna(0)

df['dia'] = 1

# Formato
df['Fecha'] = pd.to_datetime(df['año'].astype(str) + '-' + df['mes'].astype(str) + '-' +
df['dia'].astype(str))

df['Fecha'] = df['Fecha'].dt.strftime('%Y-%m')

df.drop(columns=['dia'], inplace=True)

df.set_index('Fecha', inplace=True)

df=df.reset_index()
```

- **Normalización:** En el presente estudio se ha decidido comparar y seleccionar diferentes métodos de suavización para normalizar la información y suavizar los picos presentados a través de la serie temporal. Los métodos considerados son: Media Móvil Simple

(SMA, por sus siglas en inglés), Media Móvil Exponencial (EWM, por sus siglas en inglés), Suavización Exponencial Simple (SES, por sus siglas en inglés) y el Modelo de Holt.

- Media Móvil Simple (SMA): Este método calcula el promedio de un conjunto de datos dentro de una ventana móvil, lo que ayuda a reducir la variabilidad y a suavizar los picos.

Se seleccionó SMA debido a su simplicidad y efectividad en situaciones donde los datos no presentan tendencia ni estacionalidad significativa (Hyndman & Athanasopoulos, 2018).

- Media Móvil Exponencial (EWM): A diferencia de SMA, EWM asigna más peso a las observaciones más recientes, lo que permite una mayor sensibilidad a los cambios recientes en la serie temporal. Esta característica lo hace adecuado para datos donde los cambios recientes son más relevantes (Holt, 2004).

- Suavización Exponencial Simple (SES): Similar a EWM, SES también asigna más peso a las observaciones recientes, pero lo hace a través de un coeficiente de suavización que se determina por ensayo y error. SES es útil en series temporales sin tendencia ni estacionalidad, proporcionando una aproximación rápida y eficiente (Brown, 1959).

- Modelo de Holt: Este método extiende el SES para incluir una componente de tendencia lineal, lo que lo hace ideal para series temporales que muestran una tendencia a lo largo del tiempo. El Modelo de Holt ha demostrado ser efectivo en situaciones donde la tendencia es una característica predominante en los datos (Holt, 2004).

```
df_full["SMA"]= df_full["FADE_CANT"].rolling(6).mean()
```

```
df_full["EWM"]= df_full["FADE_CANT"].ewm(alpha=0.2, adjust=False).mean()
```

```
model_SES = SimpleExpSmoothing(df['FADE_CANT']).fit(smoothing_level=0.2,
optimized=False)
```

```
df_full['SES'] = model_SES.fittedvalues
model_HOLT = model = ExponentialSmoothing(df['FADE_CANT'], trend='add').fit()
df_full['HOLT'] = model_HOLT.fittedvalues
```

### ***Descripción y Selección de Valores***

Media Móvil Simple (SMA):  
`df_full["SMA"] = df_full["FADE_CANT"].rolling(6).mean()`

La Media Móvil Simple calcula el promedio de los valores en una ventana móvil de tamaño 6. Esto ayuda a suavizar la serie temporal reduciendo la variabilidad.

Se eligió una ventana de tamaño 6 para capturar la tendencia a corto plazo sin perder demasiada información. Un tamaño de ventana más grande suavizaría más la serie, pero podría perder detalles importantes.

Media Móvil Exponencial (EWM):  
`df_full["FADE_CANT"].ewm(alpha=0.2, adjust=False).mean()`

La Media Móvil Exponencial asigna más peso a las observaciones recientes utilizando el parámetro alpha.

El parámetro alpha=0.2 indica que se da más peso a los valores recientes (20% del valor actual y 80% de los valores anteriores). Este valor se elige para reflejar la importancia de los datos recientes sin reaccionar de manera excesiva a fluctuaciones pequeñas.

Suavización Exponencial Simple (SES):

```
SimpleExpSmoothing(df['FADE_CANT']).fit(smoothing_level=0.2, optimized=False)
```

SES es un método de suavización exponencial que también da más peso a las observaciones recientes.

El `smoothing_level=0.2` es similar al  $\alpha$  en EWM. Se elige un valor de 0.2 para capturar tendencias recientes en los datos. No se optimizó el nivel de suavización, ya que se decidió mantener un valor constante para este estudio.

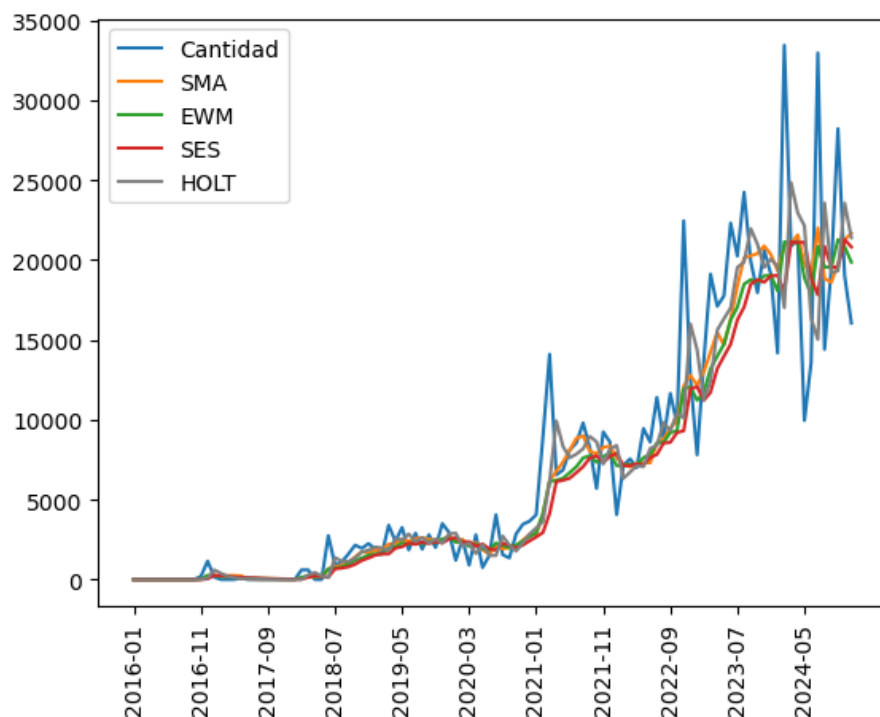
Modelo de Holt: `ExponentialSmoothing(df['FADE_CANT'], trend='add').fit()`

El Modelo de Holt extiende la suavización exponencial para incluir una componente de tendencia lineal.

Se usó una tendencia aditiva (`'trend='add'`) que es adecuada para datos con una tendencia constante. Los parámetros internos son ajustados automáticamente por el modelo (`fit()`), optimizando así la precisión del pronóstico.

## Figura 2

### *Serie de Tiempo – Normalización*

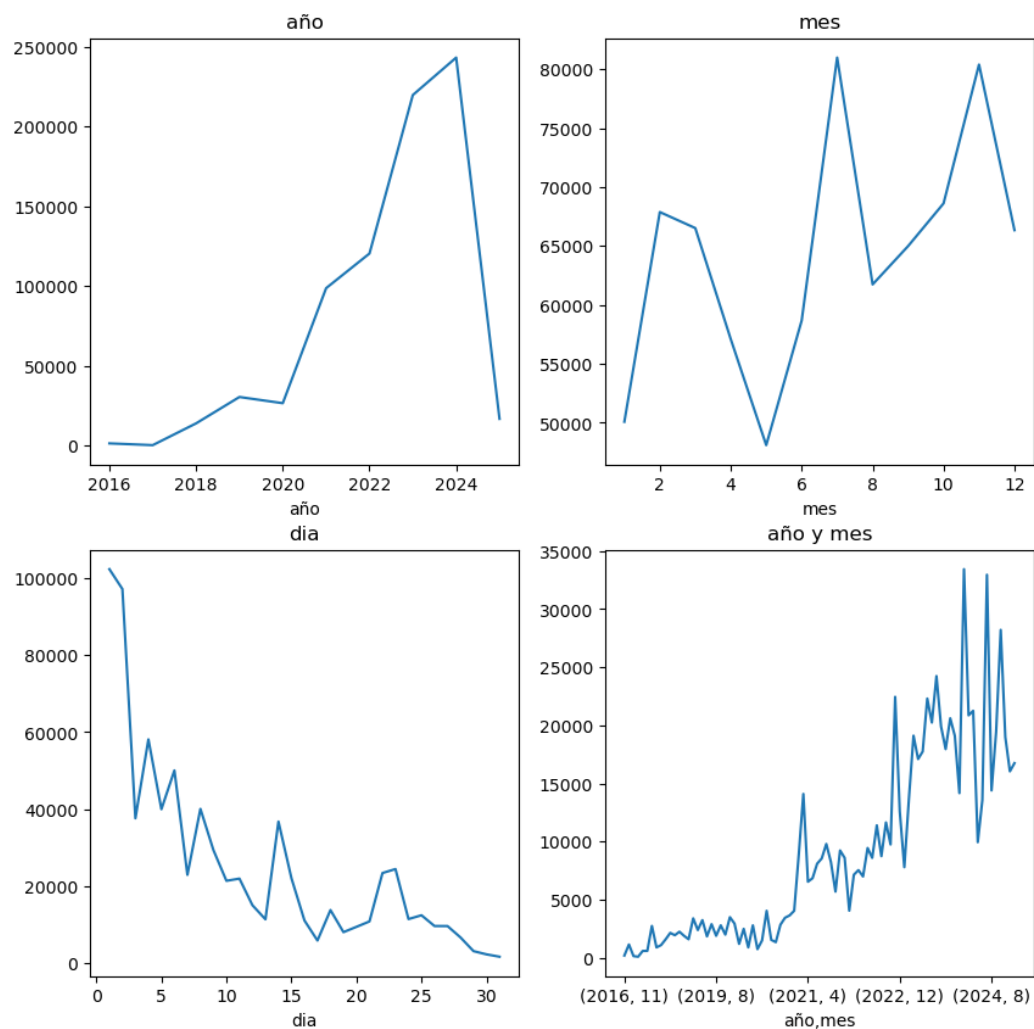


En la gráfica se observa una tendencia creciente de las ventas a partir de año 2021 la cual se ha mantenido estable desde el año 2023 con intervalos de picos mensuales, por ende, se realiza la suavización de la gráfica.

Carga: Los datos transformados se cargan en un DataFrame estructurado que facilita su acceso y análisis. Esta base de datos es el punto de partida para la fase de análisis exploratorio de datos (EDA).

### **Análisis Exploratorio de Datos (EDA)**

Visualización de Datos: Se utilizan gráficos y visualizaciones interactivas para explorar patrones y tendencias en los datos, teniendo como foco la variable tiempo, explorando el comportamiento de la serie temporal por diferentes categorías (Año, Mes, Día, Año y Mes) para determinar el patrón del uso que se le dará a la predicción del modelo (mensual, diario, anual).

**Figura 3***Serie de Tiempo – EDA*

Identificación de Variables Relevantes: A través del EDA, se identifican las variables más relevantes para el modelo de pronóstico de demanda. Esto incluye factores como estacionalidad, tendencias de ventas, y correlaciones entre diferentes variables, estableciendo como métrica base de la serie temporal la categoría AÑO – MES , para modelar según la evolución de la venta del producto en la compañía.

La elección de la categoría AÑO – MES como métrica base para la serie temporal permite modelar según la evolución de la venta del producto en la compañía. Esta estructura temporal ofrece varios beneficios: Proporciona un equilibrio entre granularidad y manejabilidad, permitiendo observar patrones mensuales sin perder de vista las tendencias anuales, y Facilita la captura de patrones estacionales y tendencias a lo largo de los años, permitiendo un ajuste más preciso del modelo para pronósticos futuros

## **Definición del Modelo de Proyección**

La segunda fase del proyecto se centra en la definición y construcción del modelo de proyección de demanda. Esta etapa es fundamental para asegurar que las proyecciones de demanda sean precisas y útiles para la toma de decisiones en el área de compras de Medicox LTDA. A continuación, se describen los pasos principales de esta fase:

### **Selección de Modelos de Machine Learning**

#### ***Revisión de Literatura***

Se realiza una revisión detallada de la literatura existente para utilizar el modelo de Machine Learning Random Forest Regressor.

La selección de los modelos se basa en estudios previos que han demostrado su efectividad en escenarios similares, considerando tanto empresas pequeñas como medianas.

La elección del modelo de Machine Learning Random Forest Regressor para este proyecto se fundamenta en una revisión detallada de la literatura existente, que demuestra su efectividad en escenarios diversos, incluyendo empresas pequeñas y medianas.

Abeysekara y Rupasinghe (2019) analizaron los factores influyentes en los sistemas de pronóstico de inventarios y concluyeron que los modelos de Machine Learning, como Random Forest, ofrecen una mayor precisión en comparación con los métodos tradicionales. De manera similar, Anitha et al. (2023) destacaron la efectividad de los algoritmos de aprendizaje automático en la predicción de la demanda de alimentos, subrayando la relevancia de estos modelos en la optimización de la cadena de suministro.

Benhamida et al. (2020) presentaron Stock&Buy, una herramienta de pronóstico de demanda que utiliza técnicas de Machine Learning para el control de inventarios, demostrando mejoras significativas en la gestión de existencias. Asimismo, Hoque et al. (2021) desarrollaron

un modelo dinámico de pronóstico de demanda utilizando Machine Learning, que resultó ser esencial para mantener la resiliencia en las cadenas de suministro.

Investigaciones recientes, como las de Kheawpeam y Sinthupinyo (2023), han aplicado modelos de Machine Learning para gestionar inventarios en tiendas minoristas multicanal, destacando la adaptabilidad y precisión de estos modelos. Además, Li et al. (2023) combinaron métodos estadísticos tradicionales con algoritmos avanzados como XGBoost, optimizando estrategias de inventario para contadores eléctricos.

Ma et al. (2021) investigaron el uso de algoritmos de Deep Learning en la predicción de demanda de repuestos automotrices, encontrando que estas técnicas superan a los enfoques convencionales en términos de precisión y eficiencia. Pawar y Tiple (2019) compararon algoritmos de Machine Learning y redes neuronales para el pronóstico de demanda de repuestos de misiles antiaéreos, concluyendo que los modelos de aprendizaje automático ofrecen una mayor exactitud en escenarios de alta variabilidad.

Pfeifer, Brand y Lohweg (2023) compararon enfoques estadísticos y de Machine Learning para la previsión de series temporales en la gestión de la demanda, encontrando que los modelos de Machine Learning proporcionan una mayor flexibilidad y precisión. Por su parte, R. et al. (2023) desarrollaron un sistema de gestión de inventarios de comercio electrónico utilizando técnicas de Machine Learning, mostrando mejoras en la predicción de la demanda y la optimización de existencias.

Finalmente, Sanjay Raja, Maurus Maria Rubenson y Sankaradass (2023) evaluaron la efectividad del aprendizaje por refuerzo en la optimización de la gestión de la cadena de suministro para el pronóstico de demanda dinámica, mientras que Puneet et al. (2021) aplicaron

técnicas de Machine Learning para el almacenamiento inteligente en almacenes, demostrando mejoras significativas en la gestión de inventarios.

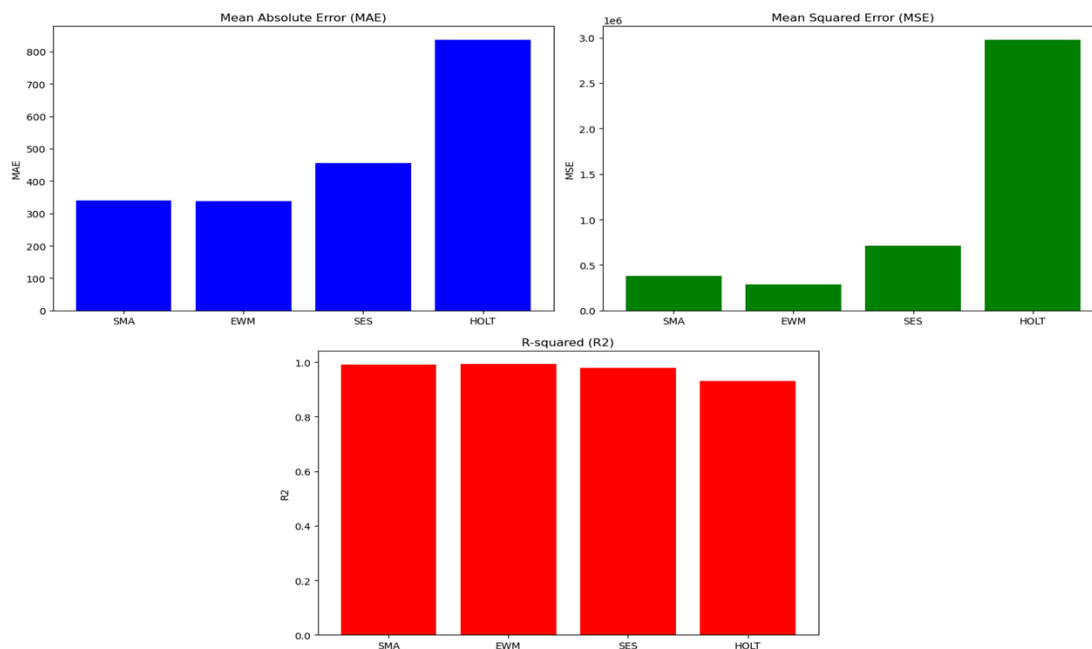
### ***Elección de Métrica de Evaluación***

Se utilizan métricas de evaluación estándar para comparar el rendimiento de los diferentes modelos de Machine Learning seleccionados con los modelos normalizados. Entre las métricas utilizadas se incluyen el error cuadrático medio (MSE), el error absoluto medio (MAE) y el coeficiente de determinación ( $R^2$ ).

Cada modelo se entrena con los datos históricos de ventas de Medicox LTDA, y se valida utilizando un conjunto de datos de prueba para evaluar su capacidad de generalización.

### **Figura 4**

*Métricas MAE, MSE y R2 para cada Normalización del Modelo*



## Entrenamiento y Validación del Modelo

### *Preparación de Datos*

Los datos consolidados y transformados en la fase anterior se dividen en conjuntos de entrenamiento y prueba. El conjunto de entrenamiento se utiliza para ajustar los parámetros del modelo, mientras que el conjunto de prueba se utiliza para evaluar su rendimiento, mediante la fórmula de `train_test_split` teniendo en cuenta que el tamaño de entrenamiento del 80%.

Se seleccionó un tamaño de entrenamiento del 80% para maximizar la cantidad de datos disponibles para el modelo mientras se deja una porción suficiente para la validación. Esta proporción es comúnmente utilizada en la práctica debido a su equilibrio entre cantidad de datos de entrenamiento y capacidad de evaluación (Hastie, Tibshirani, & Friedman, 2009).

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

### *Entrenamiento del Modelo*

La variable ['EWM'], la cual cuenta con mejor rendimiento basado en las métricas de evaluación utilizando el método de `RandomForestRegressor`. El modelo se entrena utilizando el conjunto de datos de entrenamiento, ajustando sus hiperparámetros para mejorar su precisión.

```
# Realizar predicciones utilizando el modelo RandomForestRegressor
```

```
X = df_full[['año', 'mes']]
```

```
y = df_full['EWM']
```

```
# Dividir los datos en conjuntos de entrenamiento y prueba
```

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

```
# Crear y entrenar el modelo
```

```
model = RandomForestRegressor(n_estimators=100, random_state=42)
```

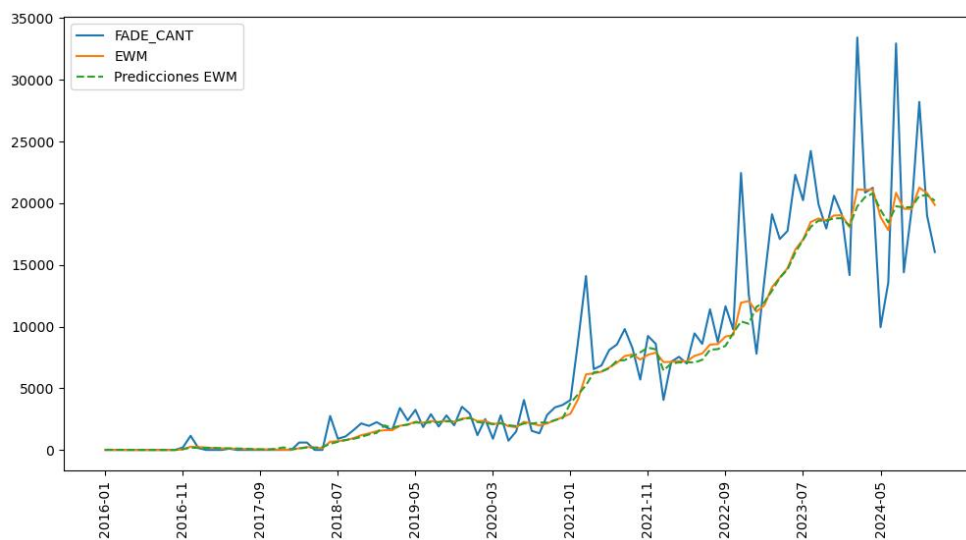
```
model.fit(X_train, y_train)
```

## Validación del Modelo

El modelo entrenado se valida utilizando el conjunto de datos de prueba. Se compara con la variable seleccionada ['EWM'] las proyecciones de demanda generadas por el modelo con los datos reales.

## Figura 5

### Predicciones del Modelo



## Implementación del Modelo Final

Una vez optimizado, el modelo final se implementa en un entorno de producción utilizando herramientas como Power BI mediante la exportación del modelo en formato pickle y su implementación en la interfaz de Power Query. Esto permite a los usuarios finales interactuar con las proyecciones de demanda y tomar decisiones informadas basadas en los datos.

```
import pandas as pd

import pickle

datos = dataset
```

with open(r"C:\Users\Analista.Comercial\OneDrive - medicox.com.co\Documentos\Practica\Proyectos Python\2\model.pkl", "rb") as archivo:

```
modelo = pickle.load(archivo)
```

```
datos["predicción"] = modelo.predict(datos[['año', 'mes']])
```

Se realizan una serie de pasos y transformaciones adicionales que permiten al integración con variables de negocio como compras, ventas e inventario para realizar la comparación.

## Figura 6

### Pronostico - Power Query

	Value.mes	Prediccion	Venta	Existencia	Compra
	<ul style="list-style-type: none"> <li>Válido 100 %</li> <li>Error 0 %</li> <li>Vacio 0 %</li> </ul>	<ul style="list-style-type: none"> <li>Válido 100 %</li> <li>Error 0 %</li> <li>Vacio 0 %</li> </ul>	<ul style="list-style-type: none"> <li>Válido 54 %</li> <li>Error 0 %</li> <li>Vacio 46 %</li> </ul>	<ul style="list-style-type: none"> <li>Válido 54 %</li> <li>Error 0 %</li> <li>Vacio 46 %</li> </ul>	<ul style="list-style-type: none"> <li>Válido 42 %</li> <li>Error 0 %</li> <li>Vacio 57 %</li> </ul>
1	1	18163	14170	24877	12500
2	2	19749	33432	23157	40000
3	3	20496	20880	22770	35000
4	4	20788	21250	49820	null
5	5	19456	9950	35070	null
6	6	18439	13575	25120	8000
7	7	19752	32950	20045	32000
8	8	19662	14400	28345	10000
9	9	19684	19440	19445	20000
10	10	20587	28210	13805	27600
11	11	20660	19000	31695	25000
12	12	20204	16042	37993	10000
13	1	18163	16750	31951	null
14	2	19749	null	null	null
15	3	20496	null	null	null
16	4	20788	null	null	null
17	5	19456	null	null	null
18	6	18439	null	null	null
19	7	19752	null	null	null
20	8	19662	null	null	null

**PROPIEDADES**

Nombre  
Pronostico

Todas las propiedades

**PASOS APLICADOS**

- Origen \*
- Tipo cambiado \*
- Ejecutar script de Python \*
- Se expandió Value \*
- Valor reemplazado \*
- Filas filtradas \*
- Tipo cambiado1 \*
- Columna combinada insertada \*
- Consultas combinadas \*
- Se expandió Resumen Fact \*
- Filas ordenadas \*
- Consultas combinadas1 \*
- Se expandió Resumen Existen... \*
- Consultas combinadas2 \*
- Se expandió Resumen Compra \*
- Columnas quitadas \*
- Tipo cambiado2 \*
- Columnas con nombre cambi...

Se desarrollan dashboards interactivos que permiten visualizar las proyecciones de demanda en tiempo real, facilitando la toma de decisiones en el área de compras y reduciendo el riesgo de excedente y pérdidas por vencimiento de productos.

## Implementación de la Herramienta

a tercera fase del proyecto se enfoca en la implementación de la herramienta de pronóstico de demanda en la empresa Medicox LTDA. Esta etapa es crucial para garantizar que las proyecciones generadas por el modelo sean accesibles y útiles para los tomadores de decisiones en el área de compras.

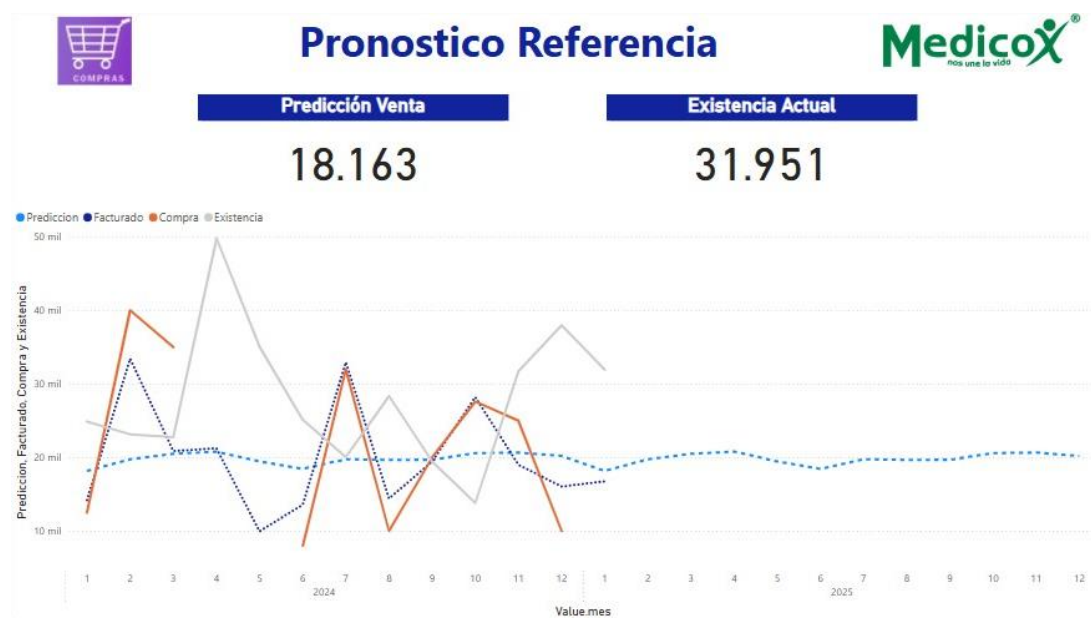
### Desarrollo de la Interfaz de Usuario

Diseño de Dashboards en Power BI: Se desarrollan dashboards interactivos en Power BI que permiten a los usuarios visualizar las proyecciones de demanda de manera intuitiva y clara. Estos dashboards incluyen un gráfico de líneas que muestra las proyecciones de demanda, así como métricas clave como el inventario actual y las tendencias de ventas y compras procesando el modelo .

El diseño de los dashboards se realiza en colaboración con los usuarios finales para asegurar que las visualizaciones sean relevantes y fáciles de interpretar.

### Figura 7

Reporte - Power BI



Este reporte fue diseñado con el objetivo de comparar las compras, el inventario existente y el pronóstico de las ventas para así tomar la decisión de las cantidades necesarias a comprar para el siguiente mes, para esto se desarrolla un gráfico de líneas que compara a través del tiempo los valores dichos de manera sencilla y precisa, además se genera de manera visual dos casillas de texto que muestran la predicción de ventas y el inventario existente en el mes presente o en curso.

Los totales visuales contiene un filtro que adhiere el dato al mes presente mientras la gráfica muestra los valores de compra inventario y pronostico por mes y año, además, la información de inventarios y compras se encuentra conectado a la base de datos del CRM de la compañía, permitiendo tener la información en tiempo real, pero la información del modelo se está reentrenando y actualizando con un intervalo de 15 días.

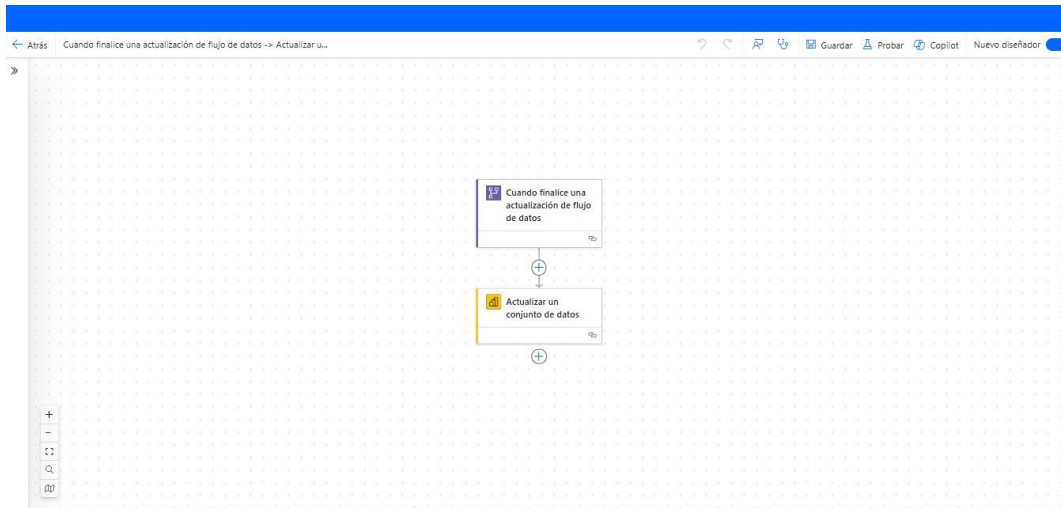
### **Integración con Sistemas Existentes**

La herramienta de pronóstico de demanda se integra con los sistemas existentes de gestión de inventarios y compras de Medicox LTDA. Esto permite una actualización automática de los datos en los dashboards de Power BI, asegurando que la información esté siempre actualizada y disponible para los usuarios, mediante la Suite de Office 365 (Power Bi y Power Automate), se establece flujo automatizado que permite la actualización del reporte en intervalos de tiempo específicos.

De primera mano se realiza la creación de las bases de datos en el servicio de power bi y se procede a realizar la publicación del reporte, gracias a la integración de sistemas de Office 365, se hace uso del aplicativo de Power Automate y se ejecutan dos pasos: se realiza una actividad que desencadena el flujo automático, esta actividad detecta la actualización de la base de información, y a partir de este realiza la actualización del reporte, entrando así en un ciclo

indefinido de actualización del reporte, el cual está definido el tiempo de actualización con el tiempo que tarda la actualización del reporte (5 minutos).

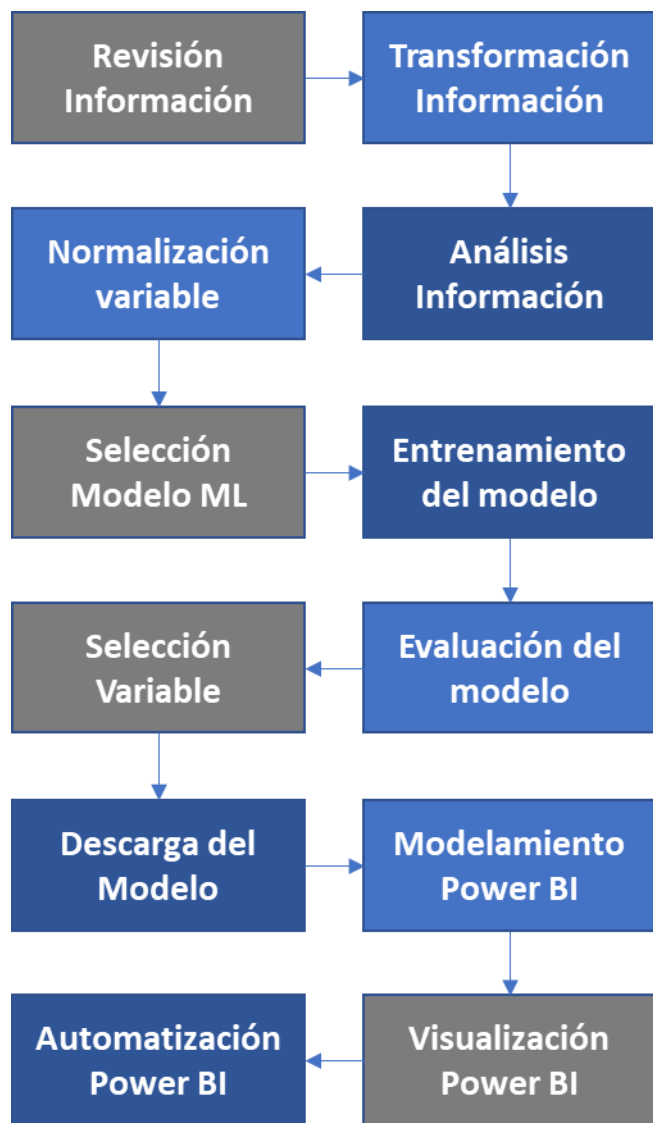
**Figura 8** *Flujo Power Automate*



Se implementan procesos automáticos de extracción, transformación y carga (ETL) para garantizar que los datos se integren de manera eficiente y precisa en la herramienta.

### ***Flujograma de Creación e Implementación***

Con el objetivo de resumir el proceso de de creación y entrenamiento del modelo y creación de la herramienta:

**Figura 9***Flujo de Implementacion***Capacitación y Soporte a Usuarios***Capacitación de Usuarios*

Se lleva a cabo un programa de capacitación para los usuarios finales, que incluye sesiones de formación sobre el uso de los dashboards de Power BI y la interpretación de las

proyecciones de demanda. La capacitación se enfoca en cómo utilizar la herramienta para tomar decisiones informadas y optimizar la gestión de inventarios.

Se proporcionan manuales y guías de uso detalladas que explican las funcionalidades de la herramienta y cómo acceder a las diferentes visualizaciones.

### ***Soporte Continuo***

Se establece un equipo de soporte para asistir a los usuarios en caso de dudas o problemas técnicos relacionados con la herramienta de pronóstico de demanda. Este equipo está disponible para resolver cualquier inconveniente y garantizar el correcto funcionamiento de la herramienta.

Se implementan canales de comunicación (Microsoft Teams), para que los usuarios puedan reportar problemas y recibir asistencia de manera eficiente.

### **Monitoreo y Evaluación**

#### ***Monitoreo del Rendimiento del Modelo***

Se establecen métricas de rendimiento y procesos de monitoreo continuo para evaluar la precisión y efectividad del modelo de pronóstico de demanda. Esto incluye el seguimiento de las proyecciones en comparación con los datos reales de ventas y la identificación de posibles desviaciones, realizando el reentrenamiento del modelo con una periodicidad de cada 15 días.

## Conclusiones

A lo largo de este proyecto, se desarrolló una herramienta de pronóstico de demanda en el contexto de distribución y comercio, demostrando su viabilidad y efectividad. Esta herramienta fue diseñada para establecer proyecciones de ventas precisas y confiables, aspecto crucial para la optimización de inventarios y la reducción de pérdidas por vencimiento de productos. La metodología ETL fue esencial para consolidar y transformar los datos necesarios, utilizando herramientas robustas como Python y Power BI. Esta integración permitió la creación de soluciones de ciencia de datos eficientes, facilitando análisis profundos y visualizaciones claras.

Los resultados obtenidos indican que la herramienta desarrollada mejora significativamente la precisión del pronóstico de demanda. Esto se traduce en una reducción notable del inventario y una disminución de las pérdidas por productos vencidos. La capacidad de realizar proyecciones más precisas apoya a los analistas de compras y negocios en la toma de decisiones fundamentadas en datos, minimizando el riesgo de exceso o escasez de inventario. La implementación de estas herramientas demostró ser un activo valioso para el proceso de toma de decisiones, proporcionándoles un panorama claro y preciso para la planificación de compras.

No obstante, es importante reconocer ciertas limitaciones del proyecto. La precisión del modelo depende en gran medida de la calidad y cantidad de los datos disponibles. Además, las fluctuaciones inesperadas en la demanda debido a factores externos no considerados en el modelo pueden afectar su precisión. Por último, la implementación y mantenimiento de la herramienta requiere un conocimiento técnico especializado y recursos continuos para asegurar su eficacia a lo largo del tiempo.

En conclusión, este proyecto ha demostrado con éxito la capacidad de desarrollar una herramienta de pronóstico de demanda efectiva en el ámbito de distribución y comercio,

ofreciendo beneficios tangibles en términos de precisión, reducción de inventario y disminución de pérdidas. Sin embargo, la atención a las limitaciones y la continua adaptación del modelo serán esenciales para mantener su relevancia y eficacia en el futuro.

## Recomendaciones

Se recomienda el desarrollo de un modelo de pronóstico de demanda específico para cada una de las referencias vendidas por Medicox LTDA. Este enfoque permitirá ampliar la cobertura y el uso de la herramienta para la toma de decisiones en el área de compras. Para lograrlo, es fundamental realizar un análisis detallado de cada referencia, determinando el método de suavización adecuado, así como el entrenamiento y la implementación de los modelos en una herramienta de Power BI, que ha sido recientemente implementada en la empresa. Esta integración asegurará una mayor precisión en las proyecciones de demanda y una mejor gestión de inventarios, reduciendo así el riesgo de excedente y pérdidas por vencimiento de productos.

Por otro lado, se recomienda la exploración de insights en diversas áreas de la empresa para identificar oportunidades adicionales para la aplicación de modelos de machine learning. Por ejemplo, se podría desarrollar un modelo de pronóstico para la realización de órdenes de alistamiento por hora. Esta iniciativa permitiría asegurar una adecuada planificación del personal logístico en horas críticas, optimizando así la asignación de recursos y reduciendo los costos operativos asociados. Además, la implementación de estos modelos contribuirá a una mayor eficiencia en la operación diaria de la empresa, asegurando que el personal esté disponible en los momentos de mayor demanda.

En un ámbito más general, se sugiere continuar explorando nuevas aplicaciones de machine learning en distintas áreas funcionales de la empresa. Esto incluye el análisis de patrones de comportamiento del cliente, la optimización de rutas de distribución y la mejora de estrategias de marketing. La utilización de herramientas avanzadas de análisis de datos no solo potenciará la toma de decisiones informada, sino que también permitirá a Medicox LTDA mantenerse competitiva en un mercado en constante evolución. La integración de estos modelos

en el entorno empresarial contribuirá a la creación de valor y a la mejora continua de los procesos internos.

### Referencias Bibliográficas

- T. K. Abeysekara and S. Rupasinghe, *Analysis of Influential Factors for Inventory Forecasting Systems*, 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), Bombay, India, 2019, pp. 1-4, doi: 10.1109/I2CT45611.2019.9033725.
- E. Anitha, M. Navin Prasath, L. Sanjai, J. A. Andrea Shiny and P. Varsini, *Effective Food Demand Forecasting Using Machine Learning Algorithms*, 2023 IEEE Engineering Informatics, Melbourne, Australia, 2023, pp. 1-7, doi: 10.1109/IEEECONF58110.2023.10520567.
- F. Z. Benhamida, O. Kaddouri, T. Ouhrouche, M. Benaichouche, D. Casado-Mansilla and D. López-de-Ipiña, *Stock&Buy: A New Demand Forecasting Tool For Inventory Control*, 2020 5th International Conference on Smart and Sustainable Technologies (SpliTech), Split, Croatia, 2020, pp. 1-6, doi: 10.23919/SpliTech49282.2020.9243824.
- M. E. Hoque, A. Thavaneswaran, S. S. Appadoo, R. K. Thulasiram and B. Banitalebi, *A Novel Dynamic Demand Forecasting Model for Resilient Supply Chains using Machine Learning*, 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC), Madrid, Spain, 2021, pp. 218-227, doi: 10.1109/COMPSAC51774.2021.00040.
- N. Kheawpeam and S. Sinthupinyo, *Demand Forecasting Using Machine Learning to Manage Product Inventory for Multi-channel Retailing Store*, 2023 IEEE International Conference on Omni-layer Intelligent Systems (COINS), Berlin, Germany, 2023, pp. 1-6, doi: 10.1109/COINS57856.2023.10189241.

- J. Li et al., *Optimal inventory strategy of electricity meters based on the combination of Holt-Winters and XGBoost*, 2023 IEEE 6th International Electrical and Energy Conference (CIEEC), Hefei, China, 2023, pp. 3987-3992, doi: 10.1109/CIEEC58067.2023.10167274.
- Z. Ma, C. Wang and Z. Zhang, *Deep Learning Algorithms for Automotive Spare Parts Demand Forecasting*, 2021 International Conference on Computer Information Science and Artificial Intelligence (CISAI), Kunming, China, 2021, pp. 358-361, doi: 10.1109/CISAI54367.2021.00075.
- N. Pawar and B. Tiple, *Analysis on Machine Learning Algorithms and Neural Networks for Demand Forecasting of Anti-Aircraft Missile Spare Parts*, 2019 International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2019, pp. 854-859, doi: 10.1109/ICCES45898.2019.9002411.
- Pfeifer, H. Brand and V. Lohweg, *A Comparison of Statistical and Machine Learning Approaches for Time Series Forecasting in a Demand Management Scenario*, 2023 IEEE 21st International Conference on Industrial Informatics (INDIN), Lemgo, Germany, 2023, pp. 1-6, doi: 10.1109/INDIN51400.2023.10218206.
- P. R, S. Kumar, S. Bhardwaj, N. Agrahari, S. Pandey and S. S. Harakannanavar, *E-Commerce Inventory Management System Using Machine Learning Approach*, 2023 International Conference on Data Science and Network Security (ICDSNS), Tiptur, India, 2023, pp. 1-7, doi: 10.1109/ICDSNS58469.2023.10245500.
- S. Sanjay Raja, A. Maurus Maria Rubenson and V. Sankaradass, *Evaluating the Effectiveness of Reinforcement Learning in Optimizing Supply Chain Management for Dynamic Demand Forecasting*, 2023 International Conference on Data Science, Agents & Artificial

Intelligence (ICDSAAI), Chennai, India, 2023, pp. 1-6, doi:  
10.1109/ICDSAAI59313.2023.10452542.

Puneet, S. Sharma, D. Deepika and G. Singh, *Intelligent Warehouse Stocking Using Machine Learning*, 2021 IEEE International Conference on Mobile Networks and Wireless Communications (ICMNWC), Tumkur, Karnataka, India, 2021, pp. 1-6, doi:  
10.1109/ICMNWC52512.2021.9688530.

Brown, R. G. (1959). *Smoothing, forecasting and prediction of discrete time series*. Prentice-Hall.

Holt, C. C. (2004). *Forecasting seasonals and trends by exponentially weighted moving averages*. International Journal of Forecasting, 20(1), 5-10.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.

Chaudhuri, S., & Dayal, U. (1997). *An overview of data warehousing and OLAP technology*. SIGMOD Record, 26(1), 65-74.

Inmon, W. H. (2005). *Building the Data Warehouse*. Wiley.