

Pronóstico de demanda de inventario de los 5 productos más vendidos de la empresa Equipos y Controles Industriales.



**Manuel Fabian Sarmiento
Marquez**

Trabajo de grado I



**FACULTAD DE INGENIERÍA Y CIENCIAS APLICADAS
MAESTRÍA EN CIENCIA DE DATOS
SANTIAGO DE CALI
2024**

ÍNDICE GENERAL

1.	Integrantes y directores del trabajo de grado.	4
2.	Introducción.	4
3.	Contexto y Antecedentes.	4
4.	Planteamiento del Problema y Justificación	5
5.	Objetivos del proyecto	5
5.1.	Objetivo general	5
5.2.	Objetivos específicos	6
6.	Metodología	6
7.	Marco Teórico	6
7.1.	Modelos	7
7.1.1.	Modelo de Promedio Móvil	7
7.1.2.	Modelo de Suavizamiento Exponencial	7
7.1.3.	Modelo de Regresión Lineal	7
7.1.4.	Modelo autorregresivo integrado de media móvil	7
7.2.	Técnicas de optimización y de evaluación	8
7.2.1.	Optimización Bayesiana	8
7.2.2.	Ventanas Móviles	8
7.2.3.	RMSE	8
8.	Estado del Arte	8
9.	Propuesta	9
9.1.	Entendimiento del negocio y acceso a los datos	10
9.2.	Carga y Análisis exploratorio de datos	10
9.3.	Tendencias e identificación de atípicos	12
9.4.	Análisis para el Producto 1	14
9.5.	Análisis para el Producto 2	16
9.6.	Análisis para el Producto 3	18
9.7.	Análisis para el Producto 4	20
9.8.	Análisis para el Producto 5	22
10.	Resultados	24
10.1.	Pronósticos Producto 1	24
10.2.	Pronósticos Producto 2	25
10.3.	Pronósticos Producto 3	26
10.4.	Pronósticos Producto 4	27
10.5.	Pronósticos Producto 5	28
11.	Validación	29

12.	Conclusiones	30
13.	Referencias	31

ILUSTRACIONES

Ilustración 1	Top 5 Marcas más vendidas de la Compañía	11
Ilustración 2	Top 5 Productos más vendidos de la Marca más vendida.	12
Ilustración 3	Top 5 productos más vendidos de la Marca más vendida en el año 2024.	12
Ilustración 4	Comportamiento mensual de los 5 productos seleccionados.	13
Ilustración 5	Distribución de ventas de cada uno de los productos seleccionados.	13
Ilustración 6	Descomposición de la serie del producto 1 en sus componentes principales	14
Ilustración 7	Polinomios que se pueden ajustar a la serie del producto 1	15
Ilustración 8	Descomposición de la serie del producto 2 en sus componentes principales	16
Ilustración 9	Polinomios que se pueden ajustar a la serie del producto 2	17
Ilustración 10	Descomposición de la serie del producto 3 en sus componentes principales	18
Ilustración 11	Polinomios que se pueden ajustar a la serie del producto 3	19
Ilustración 12	Descomposición de la serie del producto 4 en sus componentes principales	20
Ilustración 13	Polinomios que se pueden ajustar a la serie del producto 4	21
Ilustración 14	Descomposición de la serie del producto 5 en sus componentes principales	22
Ilustración 15	Polinomios que se pueden ajustar a la serie del producto 5	23
Ilustración 16	Pronósticos con intervalos de confianza para producto 1	25
Ilustración 17	Pronósticos con intervalos de confianza para producto 2	26
Ilustración 18	Pronósticos con intervalos de confianza para producto 3	27
Ilustración 19	Pronósticos con intervalos de confianza para producto 4	28
Ilustración 20	Pronósticos con intervalos de confianza para producto 5	29
Ilustración 21	Enlace a código del proyecto en Github	30

TABLAS

Tabla 1	Productos seleccionados para pronosticar	12
Tabla 2	Modelos y resultados para el producto 1	16
Tabla 3	Modelos y resultados para el producto 2	18
Tabla 4	Modelos y resultados para el producto 3	20
Tabla 5	Modelos y resultados para el producto 4	21
Tabla 6	Modelos y resultados para el producto 5	24
Tabla 7	Comparación de modelos y selección de mejores resultados	24
Tabla 8	Pronóstico de ventas para 3 meses del producto 1	25
Tabla 9	Pronóstico de ventas para 3 meses del producto 2	26
Tabla 10	Pronóstico de ventas para 3 meses del producto 3	27
Tabla 11	Pronóstico de ventas para 3 meses del producto 4	28
Tabla 12	Pronóstico de ventas para 3 meses del producto 5	29

1. Integrantes y directores del trabajo de grado.

Integrante(s):

Manuel Fabian Sarmiento M.

Director: Diego Agudelo Burbano

2. Introducción.

El presente documento expone el trabajo realizado durante los últimos tres meses, cuyo objetivo fue pronosticar la demanda de inventario de los cinco productos más vendidos en lo corrido del año para la marca más comercializada por la compañía Equipos y Controles Industriales. Este pronóstico es crucial para optimizar la gestión de inventarios, reducir costos y mejorar la satisfacción del cliente, asegurando la disponibilidad de los productos más demandados.

Para llevar a cabo este proyecto, se empleó la metodología ASUM-DM (Método unificado de soluciones analíticas para minería de datos). Esta metodología se enfoca en entender a fondo el negocio y las necesidades específicas de la organización, permitiendo diseñar una estrategia de pronóstico personalizada y eficiente. El proceso inició con una fase de análisis profundo del contexto y los requisitos del negocio, seguido de la selección y utilización de los modelos de pronóstico adecuados, y culminó con la implementación y monitoreo continuo de estos modelos para asegurar su efectividad a lo largo del tiempo.

En el análisis de los cinco productos seleccionados, se evaluaron cuatro tipos de modelos de pronóstico diferentes. Cada modelo fue sometido a un riguroso proceso de evaluación utilizando la métrica de error cuadrático medio (RMSE), con el fin de determinar cuál ofrecía el mejor rendimiento en términos de precisión y fiabilidad. Esta métrica fue esencial para identificar el modelo más adecuado para cada producto, asegurando así que las predicciones fueran lo más precisas posible.

Las siguientes secciones de este documento detallan el desarrollo y los resultados obtenidos durante este estudio, proporcionando un análisis de cada modelo evaluado y sus rendimientos en el pronóstico de demanda de inventario.

3. Contexto y Antecedentes.

Equipos y Controles Industriales es un proveedor integral de tecnología industrial que se destaca por ofrecer un amplio portafolio de productos y servicios de alta calidad. La empresa se especializa en la provisión de válvulas, instrumentos de medición, analizadores de procesos, equipos de calibración y sistemas avanzados de control. Además, proporciona servicios de ingeniería de clase mundial, instalación, configuración y puesta en marcha de sistemas, así como mantenimiento para asegurar la preservación del valor de la inversión a lo largo del ciclo de vida de las plantas industriales.

Tradicionalmente, la empresa lleva a cabo los pronósticos de demanda de forma empírica, basándose en las ventas registradas durante el mismo periodo de años anteriores y en las tendencias observadas en los últimos meses. Sin embargo, este enfoque, aunque útil, presenta limitaciones en términos de precisión y adaptabilidad a cambios rápidos en el mercado. La necesidad de optimizar la gestión de inventarios y mejorar la capacidad de respuesta a la demanda ha impulsado la búsqueda de métodos más sofisticados y precisos para el pronóstico de demanda, llevando a la implementación de técnicas avanzadas de análisis de datos y modelos predictivos.

4. Planteamiento del Problema y Justificación

Desde la pandemia del COVID-19, la empresa ha enfrentado numerosos desafíos en la gestión de inventarios debido a los retrasos y demoras causados por problemas en la logística internacional. La interrupción de las cadenas de suministro globales ha provocado variaciones significativas en los tiempos de entrega y en la disponibilidad de materiales y productos, afectando negativamente la capacidad de la empresa para satisfacer la demanda de sus clientes de manera eficiente.

Los métodos tradicionales de pronóstico de demanda, basados en ventas históricas y tendencias recientes, han demostrado ser insuficientes para adaptarse a estas nuevas condiciones volátiles del mercado. Las fluctuaciones impredecibles en la demanda, combinadas con los prolongados tiempos de reposición, han resultado en desequilibrios frecuentes en los niveles de inventario. Estos desequilibrios se manifiestan en forma de excesos, que incrementan los costos de almacenamiento y el riesgo de obsolescencia, y en desabastecimientos, que derivan en pérdidas de ventas y disminución de la satisfacción del cliente.

La empresa se encuentra en una encrucijada donde la necesidad de un método de pronóstico más robusto y adaptable es crítica. Un sistema de pronóstico basado en modelos avanzados y análisis de datos podría proporcionar las herramientas necesarias para anticipar mejor las fluctuaciones de la demanda y optimizar los niveles de inventario.

La gestión eficiente del inventario es crucial para la empresa, ya que impacta directamente en la disponibilidad de productos, la satisfacción del cliente y los costos operativos. Los pronósticos de demanda empíricos actuales no proporcionan el nivel de precisión necesario para anticipar fluctuaciones en la demanda, lo que puede resultar en exceso de inventario o en desabastecimiento. Un exceso de inventario incrementa los costos de almacenamiento y el riesgo de obsolescencia, mientras que un desabastecimiento puede llevar a la pérdida de ventas y a la insatisfacción del cliente. Por tanto, se hace imprescindible desarrollar un método de pronóstico más preciso y adaptable.

La implementación exitosa de un sistema de pronóstico de demanda más avanzado permitirá a la empresa mejorar significativamente su eficiencia operativa y su competitividad en el mercado.

5. Objetivos del proyecto

5.1. Objetivo general

Desarrollar un modelo con herramientas de estadística y “Machine Learning” en Python que sirva para pronosticar la demanda de inventario de los 5 productos más vendidos para los siguientes 3 meses.

5.2. Objetivos específicos

- Realizar un análisis exploratorio de datos para identificar si los datos con los que se cuenta satisfacen los requerimientos para el desarrollo del modelo.
- Desarrollar, entrenar y evaluar diferentes modelos en Python.
- Generar los pronósticos para los siguientes 3 meses de los 5 productos y socializar los resultados.
- Desarrollar y realizar el despliegue de la aplicación para automatizar la generación de pronósticos y facilitar su uso en software de BI.

6. Metodología

El presente estudio se llevará a cabo utilizando la metodología ASUM-DM mencionada anteriormente, que proporciona un enfoque estructurado para la implementación de proyectos de minería de datos.

ASUM-DM (Analytics Solutions Unified Method for Data Mining) es una metodología desarrollada por IBM para guiar proyectos de minería de datos y analítica avanzada de manera estructurada. Se basa en la metodología CRISP-DM y proporciona un enfoque más detallado y adaptable para proyectos complejos. Las fases de ASUM-DM incluyen:

- Entendimiento del negocio: Definir los objetivos y requisitos del proyecto.
- Enfoque analítico: Traducir los objetivos en metas de analítica.
- Requerimientos de datos: Plantear las necesidades de información específica.
- Recolección de datos
- Entendimiento de los datos: Recopilar y explorar los datos relevantes.
- Preparación de los datos: Limpiar y transformar los datos para su análisis.
- Construcción del modelo: Realizar el modelamiento que permita cumplir las metas de analítica propuestas.
- Evaluación: Validar y evaluar el modelo en base a los objetivos del negocio.
- Despliegue: Implementar y monitorizar el modelo en el entorno productivo.
- Retroalimentación.

En primer lugar, se realizarán reuniones con el equipo de compras, el equipo de bodega y el equipo de IT para entender cómo funciona actualmente el proceso para mantener el inventario de las referencias de mayor rotación de la compañía.

Posterior a esto se establecerá un plan de trabajo alineando los objetivos del proyecto con las tareas relativas a la programación que son necesarias para desarrollarlo.

Luego de esto, se realizará la conexión con la base de datos de la empresa y se analizará la tabla de pedidos existente, la cual contiene información detallada sobre las transacciones realizadas. La selección de los productos a pronosticar se realizará tomando el top 5 de las marcas más vendidas del año en curso, asegurando así que el análisis se centre en los

productos de mayor relevancia comercial.

Para la modelización del pronóstico de demanda, se emplearán dos de los cuatro tipos de modelos estadísticos: suavización exponencial Holt-Winters y ARIMA, ya que en el anterior trabajo de grado se encontró que estos dos modelos tuvieron los mejores desempeños. Cada uno de estos modelos será evaluado utilizando ventanas móviles, lo que permitirá ajustar los modelos a los cambios recientes en la demanda y mejorar su precisión predictiva. Además, se aplicará optimización bayesiana para encontrar los mejores parámetros para cada modelo, maximizando así su rendimiento.

La métrica principal para la evaluación de los modelos será el error cuadrático medio (RMSE), que permitirá identificar el modelo que ofrece la mayor precisión en las predicciones. El modelo con el menor RMSE será seleccionado para generar los pronósticos finales, proporcionando así una base sólida y precisa para la toma de decisiones en la gestión de inventarios.

Por último, el despliegue del modelo buscará hacerse automático diseñando el código para leer y escribir en una base de datos y ejecutarse de forma periódica cada mes.

7. Marco Teórico

El pronóstico de demanda de inventario es un proceso fundamental en la gestión de la cadena de suministro y la planificación de recursos empresariales. Consiste en la predicción de la cantidad de productos que se necesitarán en el futuro, con el fin de mantener niveles óptimos de inventario, evitar sobrecostos por exceso de stock y minimizar los riesgos de desabastecimiento. Este proceso es crucial para mejorar la eficiencia operativa, reducir costos y aumentar la satisfacción del cliente.

El pronóstico de demanda tiene un impacto significativo en varias áreas de la empresa:

- **Gestión de Inventarios:** Ayuda a determinar las cantidades óptimas de productos que deben mantenerse en stock.
- **Planificación de la Producción:** Facilita la programación eficiente de la producción, alineando la oferta con la demanda esperada.
- **Control de Costos:** Reduce los costos asociados al almacenamiento y manejo de inventarios excesivos, así como los costos de oportunidad por ventas perdidas debido a la falta de inventario.
- **Satisfacción del Cliente:** Mejora la capacidad de la empresa para cumplir con los pedidos de los clientes de manera oportuna, aumentando así la satisfacción y fidelidad del cliente.

Existen varios modelos para realizar pronósticos de demanda, cada uno con sus propias ventajas y limitaciones. A continuación, se describen algunos de los más utilizados en la industria:

7.1. Modelos

7.1.1. Modelo de Suavizamiento Exponencial

El suavizamiento exponencial es una técnica de pronóstico que se utiliza para analizar y predecir series temporales. A diferencia del promedio móvil, que otorga el mismo peso a todas las observaciones en la ventana, el suavizamiento exponencial asigna más peso a las observaciones más recientes, lo que permite una mejor adaptación a los cambios en los datos. Este modelo es útil para series temporales con datos que cambian lentamente y no tienen tendencias ni estacionalidades significativas.

7.1.2. Modelo autorregresivo integrado de media móvil

El modelo ARIMA (AutoRegressive Integrated Moving Average) es una técnica de análisis y pronóstico de series temporales ampliamente utilizada. Combina componentes autorregresivos (AR), de promediado móvil (MA), y una parte integrada (I) para modelar y prever datos que muestran patrones a lo largo del tiempo. ARIMA es especialmente útil para series temporales no estacionarias, donde los niveles medios y las varianzas cambian con el tiempo.

7.2. Técnicas de optimización y de evaluación

7.2.1. Optimización Bayesiana

La optimización bayesiana es un enfoque para optimizar funciones que son caras de evaluar. Se utiliza principalmente en problemas donde la función objetivo es compleja, costosa de evaluar o no tiene una forma cerrada conocida, como en la optimización de hiperparámetros en modelos de aprendizaje automático.

7.2.2. Ventanas Móviles

Las ventanas móviles (o ventanas deslizantes) son una técnica utilizada en el análisis de series temporales, procesamiento de señales y algoritmos de datos en tiempo real. Esta técnica implica la utilización de una ventana de tamaño fijo que se desplaza a lo largo de una secuencia de datos para realizar cálculos en subconjuntos consecutivos de esa secuencia.

7.2.3. RMSE

El RMSE es una métrica utilizada para evaluar la precisión de un modelo de predicción, particularmente en problemas de regresión. Mide la diferencia entre los valores predichos por el modelo y los valores reales observados, proporcionando una medida de la magnitud promedio del error.

8. Estado del Arte

La precisión en los pronósticos de demanda es crucial para la planificación de inventarios, ya que influye directamente en la capacidad de una empresa para satisfacer la demanda del cliente sin incurrir en costos excesivos de mantenimiento de inventarios. A lo largo de los años, se han desarrollado y perfeccionado diversos métodos para la previsión de la demanda, incluyendo modelos tradicionales como ARIMA y métodos más modernos basados en aprendizaje automático. A continuación, se presenta un resumen del estado del arte en

pronósticos de demanda de inventario, basado en la literatura encontrada.

Modelos Tradicionales

Modelos ARIMA

Los modelos ARIMA han sido extensamente estudiados y aplicados en diversas áreas. La base teórica de los modelos ARIMA fue establecida por Box y Jenkins (1970) y ampliada por Box et al. (1994). Estos modelos son conocidos por su capacidad para modelar series temporales estacionarias a través de la combinación de componentes autoregresivos y de media móvil. Sin embargo, a pesar de su robustez teórica, los modelos ARIMA no han ganado popularidad en las soluciones de software para la gestión de inventarios, debido en parte a la preferencia por métodos de suavizamiento exponencial que son más fáciles de implementar y ajustar.

Suavizamiento Exponencial

El suavizamiento exponencial sigue siendo una de las técnicas más populares en la planificación de la cadena de suministro. Este método es especialmente útil para la demanda intermitente, caracterizada por la llegada ocasional de demanda intercalada con intervalos sin demanda. Los métodos de suavizamiento exponencial, como el método de Holt-Winters, han demostrado ser eficaces en numerosas aplicaciones empíricas, proporcionando una precisión razonable en la mayoría de los casos.

Métodos Basados en el Aprendizaje Automático

Redes Neuronales y Modelos Probabilísticos

Los enfoques modernos han comenzado a integrar técnicas de aprendizaje automático para mejorar la precisión de los pronósticos. Por ejemplo, el uso de redes neuronales temporales convolucionales (TCN) y redes neuronales recurrentes (RNN) ha demostrado ser eficaz en la previsión probabilística, especialmente en contextos donde la demanda presenta patrones complejos y no lineales.

Modelos de Aprendizaje Profundo

Un estudio reciente de Salinas et al. (2020) introdujo el modelo DeepAR, que utiliza redes neuronales recurrentes autorregresivas para generar pronósticos probabilísticos precisos. Este enfoque es particularmente útil para conjuntos de datos grandes y complejos, donde los métodos tradicionales pueden no capturar adecuadamente las variaciones y patrones subyacentes.

Optimización de Parámetros de Modelos de Pronóstico

La optimización de los parámetros de los modelos de pronóstico también ha sido un área de enfoque significativo. La parametrización adecuada de los modelos de pronóstico puede mejorar sustancialmente el rendimiento del inventario. Kourentzes et al. (2020) proponen una forma de combinar múltiples objetivos de inventario, como satisfacer la demanda y eliminar el exceso de stock, utilizando una función de costo resultante para identificar los parámetros óptimos del modelo de pronóstico.

El campo de los pronósticos de demanda de inventario ha avanzado considerablemente, integrando métodos tradicionales con técnicas modernas de aprendizaje automático para mejorar la precisión y eficiencia de los pronósticos. Mientras que los modelos tradicionales como ARIMA y el suavizamiento exponencial siguen siendo relevantes, las innovaciones en aprendizaje automático ofrecen nuevas oportunidades para abordar los desafíos de la

demanda compleja e intermitente. La optimización continua y la validación empírica de estos modelos son esenciales para su aplicación exitosa en la gestión de inventarios.

En la revisión de la literatura, se encontró que los modelos ARIMA y de suavizamiento exponencial son ampliamente utilizados para pronósticos en diversos contextos. En nuestro trabajo, tomaremos estos casos de uso como base para el desarrollo de nuestro trabajo. Además, en la lectura de los documentos, se encontró que en la mayoría de los estudios revisados que incluyen modelos avanzados como redes neuronales, concluyen que la implementación de modelos más avanzados no aporta mejoras significativas en la precisión de los pronósticos.

9. Propuesta

Con el fin de cumplir los objetivos planteados en el trabajo, se propusieron los siguientes pasos:

9.1. Entendimiento del negocio y acceso a los datos

Para poder entender las necesidades del negocio y las necesidades de información se realizaron una serie de reuniones con el departamento de compras y almacén y con el departamento de IT, luego de las cuales el acceso a los datos fue suministrado con aprobación de la administración de la Empresa, se presentó una propuesta de trabajo con el alcance definido en este documento y se realizó la descarga de la información de la tabla en un computador de la compañía para su posterior análisis.

9.2. Carga y Análisis exploratorio de datos

El set de datos contiene 43 columnas y 212,419 registros correspondientes a las ventas registradas en el sistema desde el año 2004. Para los fines de este análisis, se utilizarán únicamente cinco columnas específicas: *Fechadocumento*, *Descripcionarticulo*, *Numeroarticulo*, *Marca* y *Cantidad*, estas se describen a continuación:

- *Fechadocumento*: Indica la fecha en la que se realizó el pedido
- *Descripcionarticulo*: Indica el nombre del producto
- *Numeroarticulo*: Indica el código interno que tiene asignado el producto para inventario
- *Marca*: Indica la Marca del producto
- *Cantidad*: Cantidad del pedido

Con estas columnas, es suficiente para realizar los pronósticos de demanda de inventario, para limpiar los datos y obtener mejores resultados en los modelos, se verificó la cantidad de nulos que tenía el Dataset.

El set de datos presenta la siguiente cantidad de datos nulos en las columnas seleccionadas:

- *Fechadocumento*: 0
- *Descripcionarticulo*: 384
- *Numeroarticulo*: 27

- Marca: 13,466
- Cantidad: 0

Y sus respectivos porcentajes son:

- Fechadocumento: 0.000%
- Descripcionarticulo: 0.181%
- Numeroarticulo: 0.013%
- Marca: 6.339%
- Cantidad: 0.000%

Se decide eliminar los registros que contienen valores nulos en las columnas *Marca* y *Descripcionarticulo* debido a las siguientes razones:

Marca: Esta columna es esencial para el análisis de ventas por marca. La falta de esta información impide identificar el rendimiento de diferentes marcas; La eliminación del 6.339% de los registros garantiza que los análisis y pronósticos sean precisos y representativos de la realidad del mercado.

Descripcionarticulo: Aunque el porcentaje de datos nulos es relativamente bajo (0.181%), la descripción del artículo es crucial para la categorización y análisis detallado de los productos. Mantener registros sin esta información podría llevar a una clasificación errónea o incompleta de los productos, afectando negativamente la calidad del análisis.

Se realiza la conversión del tipo de datos de la columna *Fechadocumento* a formato datetime para facilitar el análisis temporal de las ventas. Esta conversión permite realizar operaciones y análisis más precisos, como la identificación de tendencias de ventas a lo largo del tiempo.

Análisis y Visualización de Datos

Marcas más Vendidas:

La Ilustración 1 muestra las marcas más vendidas para identificar la marca con mayor volumen de ventas. Este análisis es crucial para comprender las preferencias de los clientes y la participación de mercado de diferentes marcas.

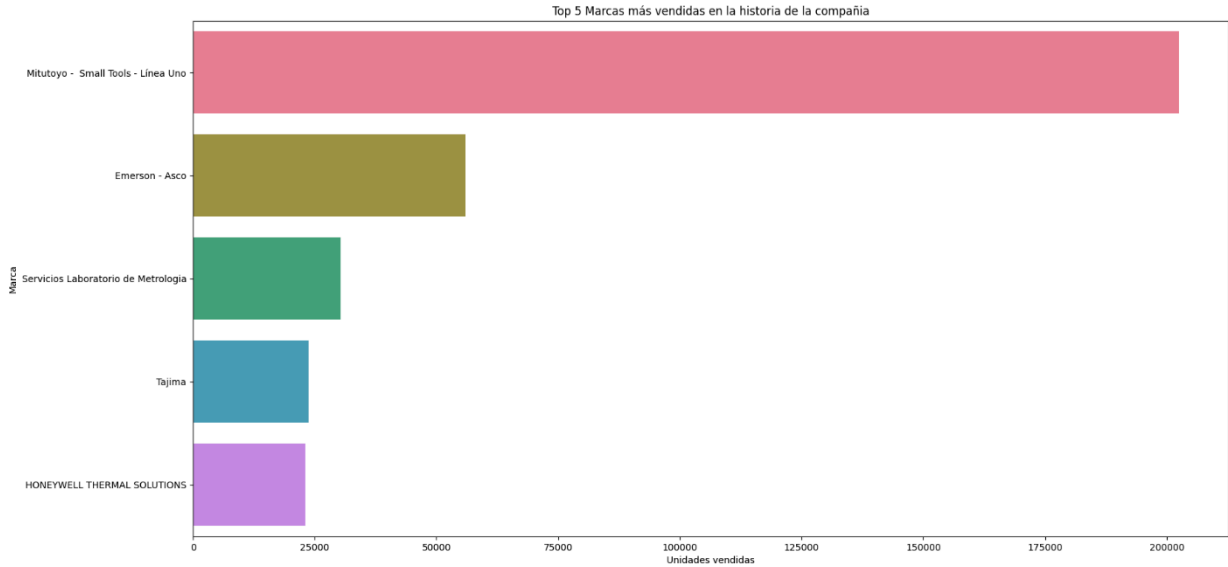


Ilustración 1 Top 5 Marcas más vendidas de la Compañía

Productos más Vendidos de la Marca Líder:

Una vez identificada la marca más vendida, se realiza una gráfica de los productos más vendidos de dicha marca, Ilustración 2. Este análisis proporciona información detallada sobre cuáles productos específicos dentro de la marca están impulsando las ventas.

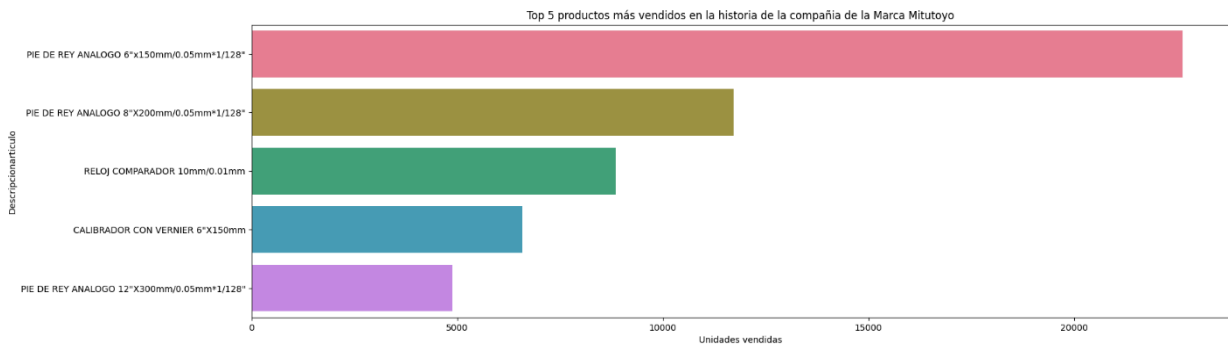


Ilustración 2 Top 5 Productos más vendidos de la Marca más vendida.

Productos más Vendidos de la Marca Líder en el Año Actual:

Por último, se realiza una gráfica de los 10 productos más vendidos de la marca líder, enfocándose en las ventas del año actual, Ilustración 3. Este análisis ayuda a identificar tendencias recientes y productos populares en el periodo más reciente, proporcionando una visión actualizada del rendimiento de la marca.

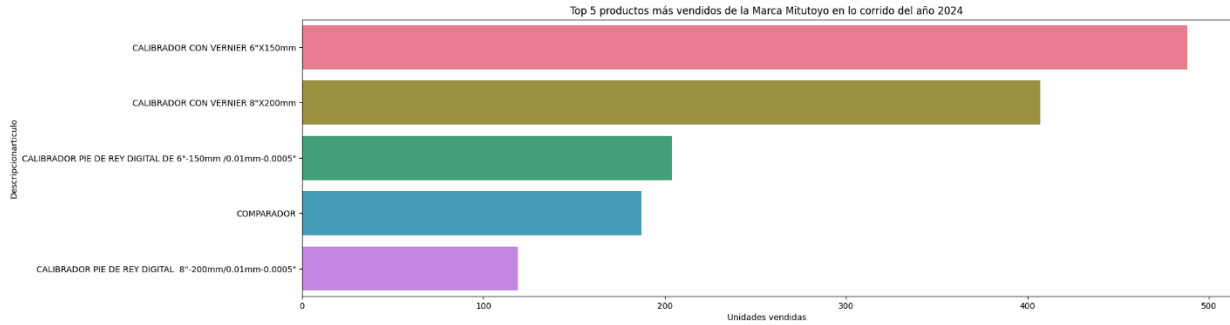


Ilustración 3 Top 5 productos más vendidos de la Marca más vendida en el año 2024.

Se identifican entonces los siguientes productos para efectuar los análisis:

PRODUCTO	CÓDIGO
CALIBRADOR CON VERNIER 6"X150mm	530-104
CALIBRADOR CON VERNIER 8"X200mm	530-114
CALIBRADOR PIE DE REY DIGITAL DE 6"-150mm /0.0...	500-196-30
COMPARADOR	2046A
CALIBRADOR PIE DE REY DIGITAL 8"-200mm/0.01mm...	500-197-30

Tabla 1 Productos seleccionados para pronosticar

9.3. Tendencias e identificación de atípicos

Determinados los productos a pronosticar, se procedió a extraer la información de fecha y cantidad de cada uno de ellos en una tabla independiente y a agrupar la cantidad de ventas por mes, se graficaron sus tendencias como se observa en la Ilustración 4. Y se graficó su distribución para detectar valores atípicos en cada uno de ellos, Ilustración 5.

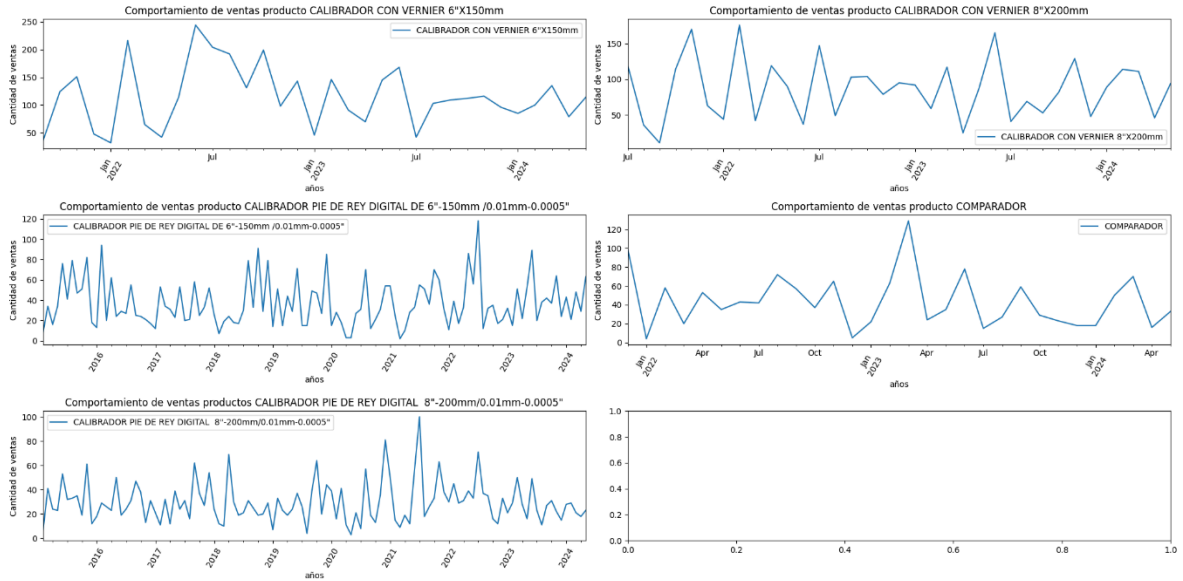


Ilustración 4 Comportamiento mensual de los 5 productos seleccionados.

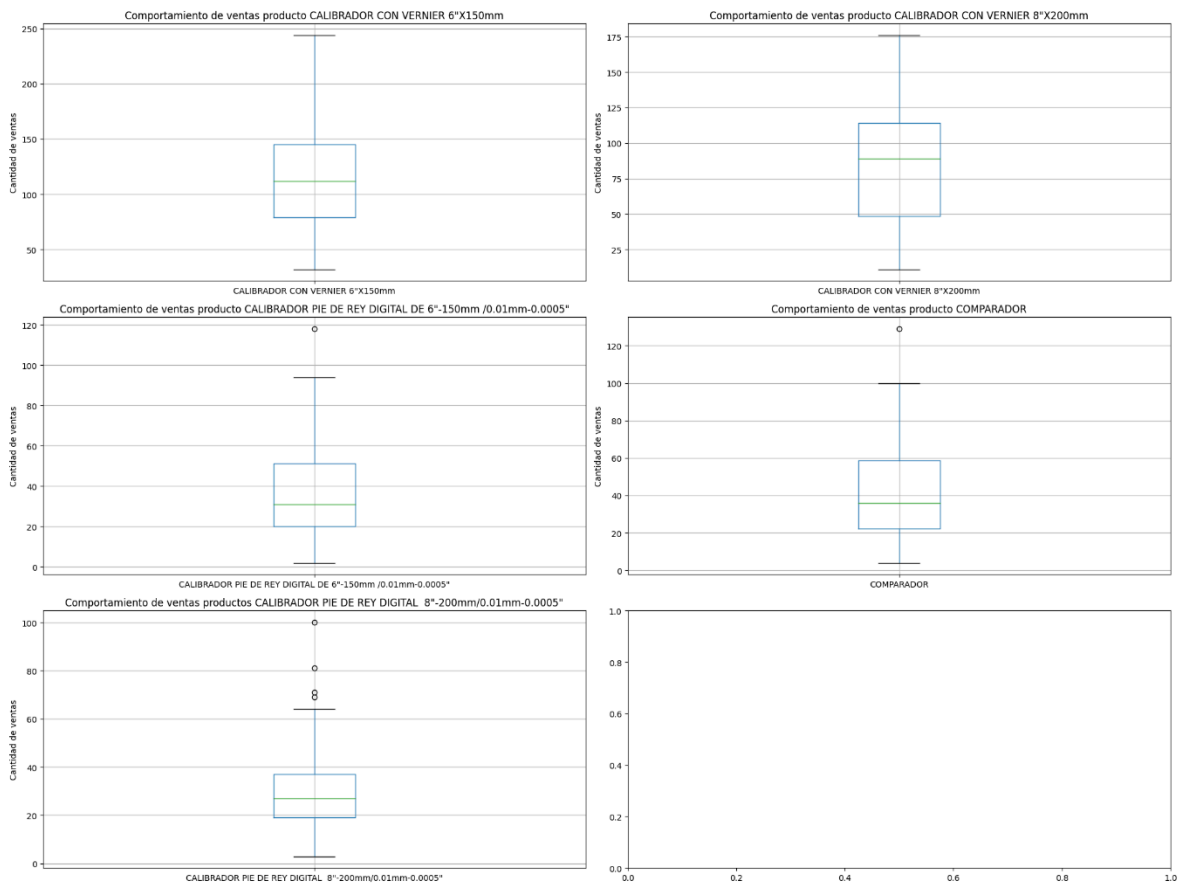


Ilustración 5 Distribución de ventas de cada uno de los productos seleccionados.

Cómo se puede observar en la anterior gráfica se encuentran valores atípicos en los productos 3,4 y 5, sin embargo, al consultar con las personas del área encargada dentro de la compañía ratificaron que estos valores están dentro del margen de la operación, por lo tanto,

se tomo la decisión de conservarlos dentro del análisis.

A partir de aquí se comenzaron a realizar los análisis y modelos individuales para cada producto, debido a que el área interesada necesita saber los pronósticos de los próximos 3 meses por la naturaleza de la logística de compras y tiempos de entrega, se tomará como periodo de validación o de test el mismo lapso de tiempo, también se utilizarán 3 ventanas de entrenamiento que serán de 24, 36 y 48 meses dependiendo de la serie.

9.4. Análisis para el Producto 1

Para el producto 1 se estableció una ventana de entrenamiento de 24 meses para todos los modelos, debido a que se tiene información de 33 meses y a que algunos modelos como el de Holt-Winters, requieren al menos dos estacionalidades para generar los pronósticos correctamente.

En la ilustración 6 se observa la descomposición de la serie de tiempo, notando una tendencia con una disminución constante desde finales del 2022.

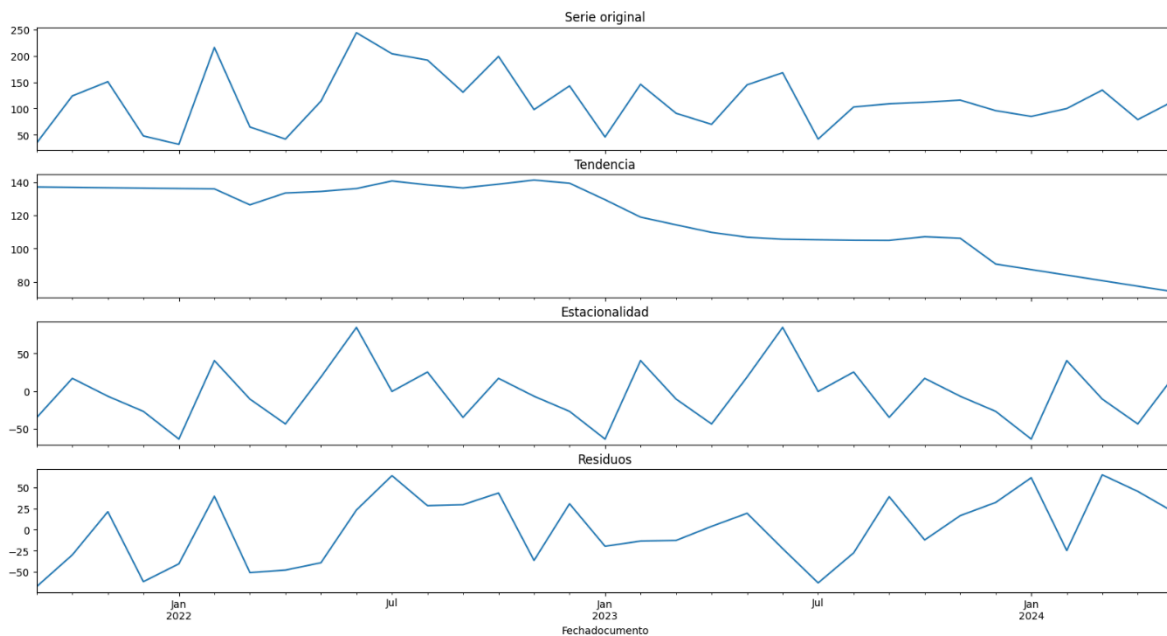


Ilustración 6 Descomposición de la serie del producto 1 en sus componentes principales

El primer modelo en evaluarse fue el de media móvil que arrojó un valor de RMSE de 12.99, este valor va a ser nuestra referencia para los siguientes modelos.

Posteriormente se utilizó un modelo de suavización exponencial Holt Winters, con optimización bayesiana para encontrar los mejores hiperparámetros del modelo, el cuál arrojó un RMSE DE 11.34 con la siguiente configuración:

alfa = 0.25, beta = 0.33, gamma = 0.86, error = add, trend = add, seasonal = None

Se realizó el ajuste de polinomios de diferentes grados para identificar cuál puede adaptarse mejor a los datos, sin embargo, el mejor polinomio con menor RMSE fue el de grado 10 como se observa en la Ilustración 7 y el valor de RMSE de 44.88, considerablemente más alto que el valor de referencia.

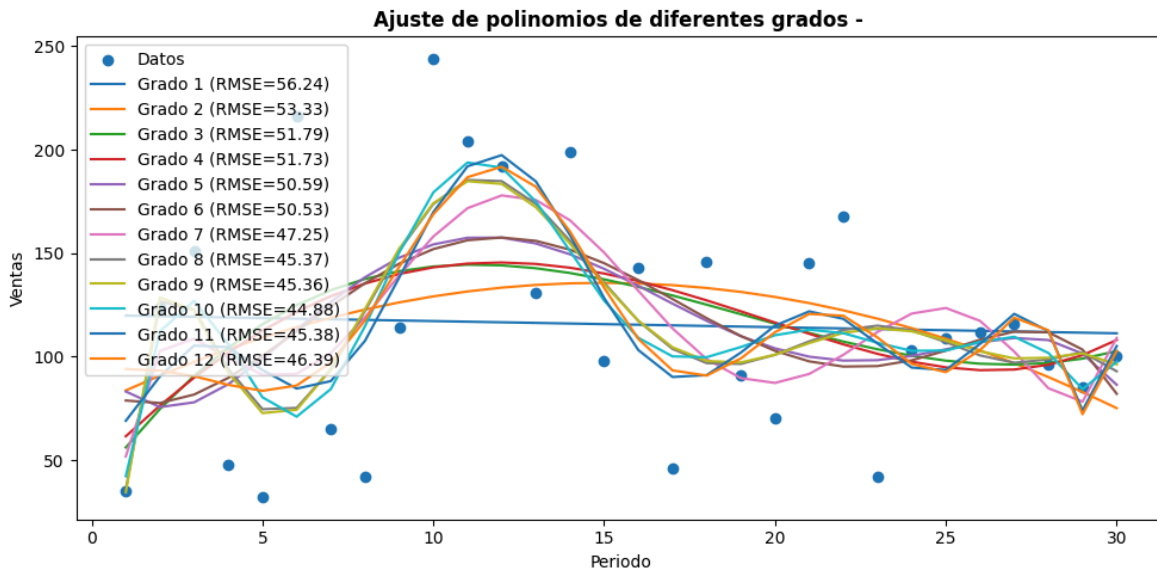


Ilustración 7 Polinomios que se pueden ajustar a la serie del producto 1

Se realizó la regresión lineal para los datos de entrenamiento pronosticando los 3 meses de prueba, y al realizar la validación con los datos de test, se obtuvo un RMSE de 645.64.

Para el análisis de regresión lineal, se realizaron las diferentes pruebas estadísticas para poder aplicar el modelo, no obstante, como ninguna de las regresiones lineales obtuvo buenos resultados de RMSE, no se entrará en el detalle de exponerlas.

El último modelo utilizado para este producto fue un ARIMA, con el fin de validar que la serie cumpliera con el requisito de ser estacionaria para usar el modelo, se realizaron las pruebas de ADF y de KPSS que arrojaron los siguientes resultados:

ADF Statistic: -5.919973338424765, p-value: 2.5199582570925285e-07

La prueba ADF tiene un valor p muy bajo, lo que sugiere que se rechaza la hipótesis nula de no estacionariedad. Esto indica que la serie de tiempo es estacionaria.

KPSS Statistic: 0.29568130183766145, p-value: 0.1

La prueba KPSS tiene un valor p relativamente alto, lo que sugiere que no se rechaza la hipótesis nula de estacionariedad. Esto también indica que la serie de tiempo es estacionaria.

Luego de obtener estos resultados, se realizó la optimización bayesiana para encontrar los mejores valores para el modelo que fueron $P = 2$, $D = 2$, $Q = 1$ con un RMSE de 8.68, y al ser este el mejor RMSE se utilizará un ARIMA (2,2,1) para realizar el pronóstico del producto.

Los resultados de los modelos de este producto se comparan en la tabla 2 que se muestra a continuación.

Modelo	Ventana de evaluación (meses)	RMSE
Media Móvil	24	12.99
Suavización Holt Winters	24	11.34
Modelo de Regresión	24	645.64
ARIMA	24	8.68

Tabla 2 Modelos y resultados para el producto 1

9.5. Análisis para el Producto 2

Para el producto 2 al igual que con el 1 se estableció una ventana de entrenamiento de 24 meses para todos los modelos, debido a que se tiene información de 35 meses y como se explicó antes algunos modelos como el de Holt-Winters, requieren al menos dos estacionalidades para generar los pronósticos correctamente.

En la ilustración 8 se observa la descomposición de la serie de tiempo, notando una tendencia con una disminución suave pero constante desde diciembre del 2023.

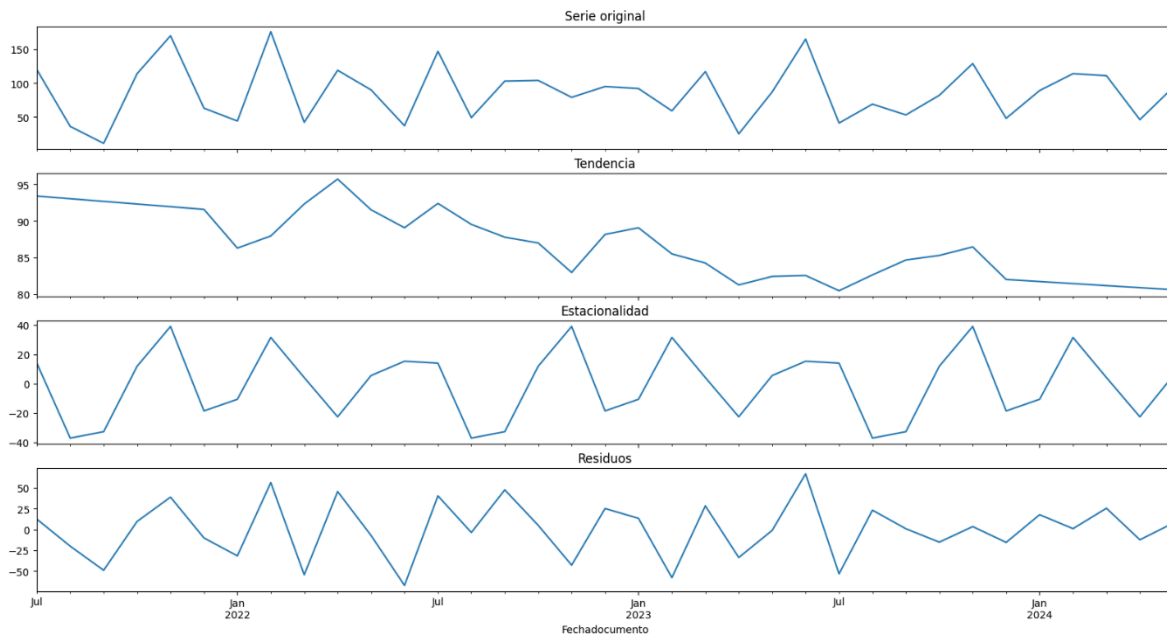


Ilustración 8 Descomposición de la serie del producto 2 en sus componentes principales

El primer modelo en evaluarse fue el de media móvil que arrojó un valor de RMSE de 25.98, este valor va a ser nuestra referencia de comparación con los siguientes modelos.

El segundo modelo que se utilizó fue de suavización exponencial Holt Winters, con optimización bayesiana para encontrar los mejores hiperparámetros del modelo, el cuál arrojó un RMSE DE 14.84 con la siguiente configuración:

alfa = 0.93, beta = 0.99, gamma = 0.64, error = mul, trend = mul, seasonal = add

Se realizó el ajuste de polinomios al igual que en el paso anterior, sin embargo, el mejor polinomio con menor RMSE fue el de grado 9 como se observa en la Ilustración 9 y el valor de RMSE de 39,32 más alto que el valor de referencia.

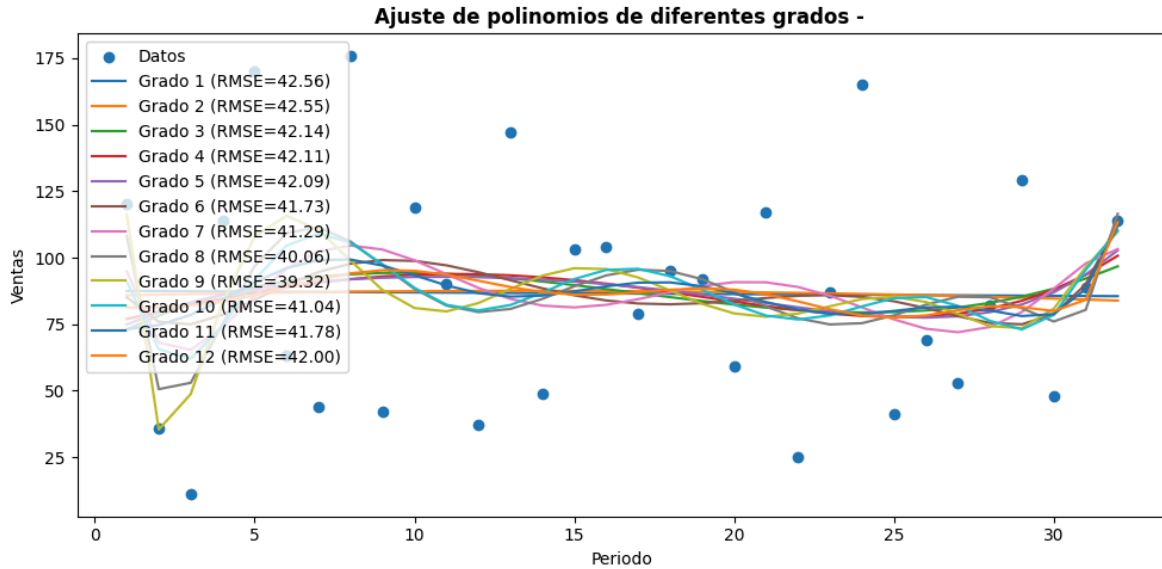


Ilustración 9 Polinomios que se pueden ajustar a la serie del producto 2

Por último, el modelo ARIMA, la verificación de estacionariedad arrojó los siguientes resultados:

ADF Statistic: -6.043476278421921, p-value: 1.2768207194010653e-07

La prueba ADF tiene un valor p muy bajo, indicando que se rechaza la hipótesis nula de no estacionariedad. La serie de tiempo es estacionaria.

KPSS Statistic: 0.2554636299265489, p-value: 0.1

La prueba KPSS tiene un valor p alto, lo que sugiere que no se rechaza la hipótesis nula de estacionariedad, confirmando que la serie de tiempo es estacionaria.

Luego de obtener estos resultados, se realizó la optimización bayesiana para encontrar los mejores valores para el modelo que fueron $P = 2$, $D = 2$, $Q = 3$ con un RMSE de 8.56, y al ser este el mejor RMSE se utilizará un ARIMA (2,2,3) para realizar el pronóstico del producto.

Los resultados de los modelos de este producto se comparan en la tabla 2 que se muestra a continuación.

Modelo	Ventana de evaluación (meses)	RMSE
Media Móvil	24	25.98
Suavización Holt Winters	24	14.84
Modelo de Regresión	24	417.72
ARIMA	24	8.56

Tabla 3 Modelos y resultados para el producto 2

9.6. Análisis para el Producto 3

Para el producto 3 se establecieron unas ventanas de entrenamiento de 24, 36 y 48 meses para todos los modelos, debido a que se cuenta con información de 112 meses

En la ilustración 10 se observa la descomposición de la serie de tiempo, notando una tendencia al alza desde el año 2023.

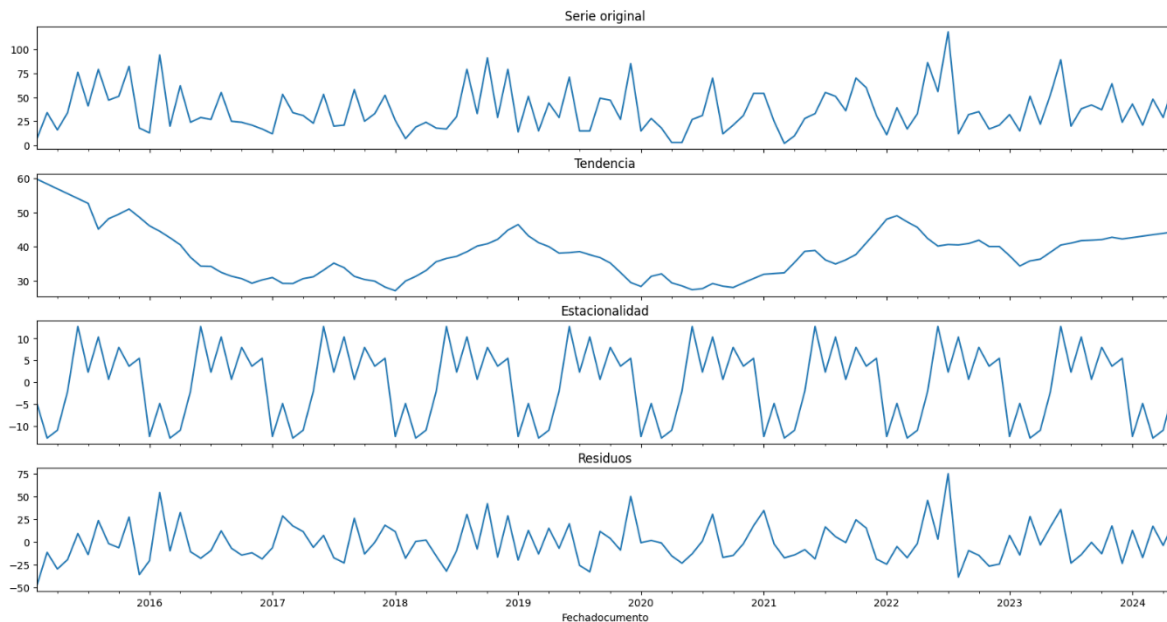


Ilustración 10 Descomposición de la serie del producto 3 en sus componentes principales

El primer modelo en evaluarse fue el de media móvil que arrojó un valor de RMSE de 15.36, este valor va a ser nuestra referencia de comparación.

El segundo modelo que se utilizó fue de suavización exponencial Holt Winters, con optimización bayesiana para encontrar los mejores hiperparámetros del modelo incluyendo la ventana de tiempo, el cuál arrojó los siguientes resultados donde w es la cantidad de meses usados para evaluar y desplazar la ventana a lo largo de la serie:

$w = 24$, $\alpha = 0.04$, $\beta = 0.75$, $\gamma = 0.76$, $\text{error} = \text{add}$, $\text{trend} = \text{None}$, $\text{seasonal} = \text{mul}$,

RMSE = 6.63

w = 48, alfa = 0.01, beta = 0.53, gamma = 0.99, error = add, trend = None, seasonal = mul, RMSE = 6.14

Se realizó el ajuste de polinomios, el mejor polinomio con menor RMSE fue el de grado 7 como se observa en la Ilustración 11 y el valor de RMSE de 22.19, ligeramente más alto que el valor de referencia.

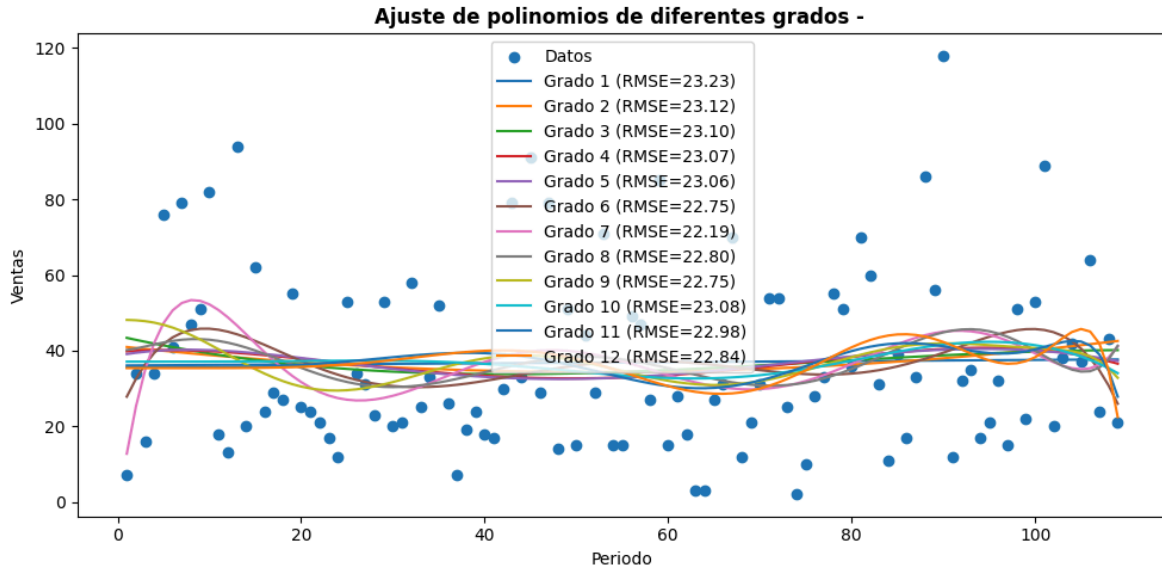


Ilustración 11 Polinomios que se pueden ajustar a la serie del producto 3

La regresión con polinomio de orden 7 arrojó los siguientes resultados:

Tamaño de ventana = 24, periodos evaluados = 3, RMSE = 96.95

Tamaño de ventana = 36, periodos evaluados = 3, RMSE = 69.64

Tamaño de ventana = 48, periodos evaluados = 3, RMSE = 22.67

Para el modelo ARIMA, las pruebas de estacionariedad mostraron los siguientes resultados:

ADF Statistic: -5.471013344760365, p-value: 2.319393159260857e-06

La prueba ADF muestra un valor p muy bajo, indicando que se rechaza la hipótesis nula de no estacionariedad. La serie de tiempo es estacionaria.

KPSS Statistic: 0.2879656024198446, p-value: 0.1

La prueba KPSS tiene un valor p alto, sugiriendo que no se rechaza la hipótesis nula de estacionariedad, confirmando la estacionariedad de la serie de tiempo.

Luego de obtener estos resultados, se realizó la optimización del modelo para encontrar los mejores valores que arrojó la siguiente información:

w = 24, fore = 3, P = 2, D = 1, Q = 2, RMSE = 10.29

w = 36, fore = 3, P = 0, D = 0, Q = 2, RMSE = 13.15

w = 48, fore = 3, P = 0, D = 0, Q = 3, RMSE = 6.6

Comparando las métricas de todos los modelos en la tabla 4, se observa que el mejor modelo para realizar el pronóstico es el modelo de Suavización Holt-Winters con un RMSE = 6.14 una

ventana de 48 meses, error = add, trend = None y seasonal = mul.

Modelo	Ventana de evaluación (meses)	RMSE
Media Móvil	24	15.36
Suavización Holt Winters	24	6.63
	48	6.14
Modelo de Regresión	24	96.95
	36	69.64
	48	22.67
ARIMA	24	10.29
	36	13.15
	48	6.6

Tabla 4 Modelos y resultados para el producto 3

9.7. Análisis para el Producto 4

Para el producto 4 al igual que con los productos 1 y 2 se estableció una ventana de evaluación de 24 meses para todos los modelos, debido a que se tiene información de 30 meses.

En la ilustración 12 se observa la descomposición de la serie de tiempo, notando una tendencia con una disminución constante desde enero del 2023.

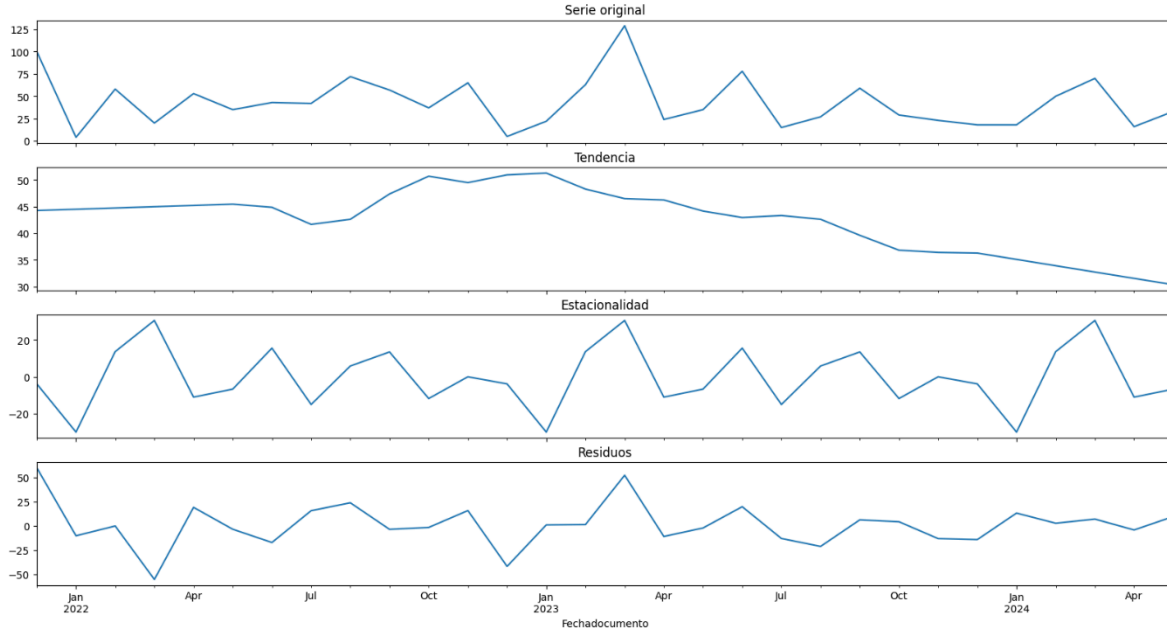


Ilustración 12 Descomposición de la serie del producto 4 en sus componentes principales

El modelo de media móvil arrojó un valor de RMSE de 27.85, este valor se tomará de referencia de comparación para los siguientes modelos.

Posteriormente se utilizó un modelo de suavización exponencial Holt Winters, con optimización bayesiana para encontrar los mejores hiperparámetros del modelo, el cuál arrojó

un RMSE DE 7.39 con la siguiente configuración:

alfa = 0.18, beta = 0.34, gamma = 0.4, error = add, trend = None, seasonal = add

Se realizó el ajuste de polinomios de diferentes grados para identificar cuál puede adaptarse mejor a los datos, sin embargo, el mejor polinomio con menor RMSE fue el de grado 10 como se observa en la Ilustración 13 y el valor de RMSE de 24.28.

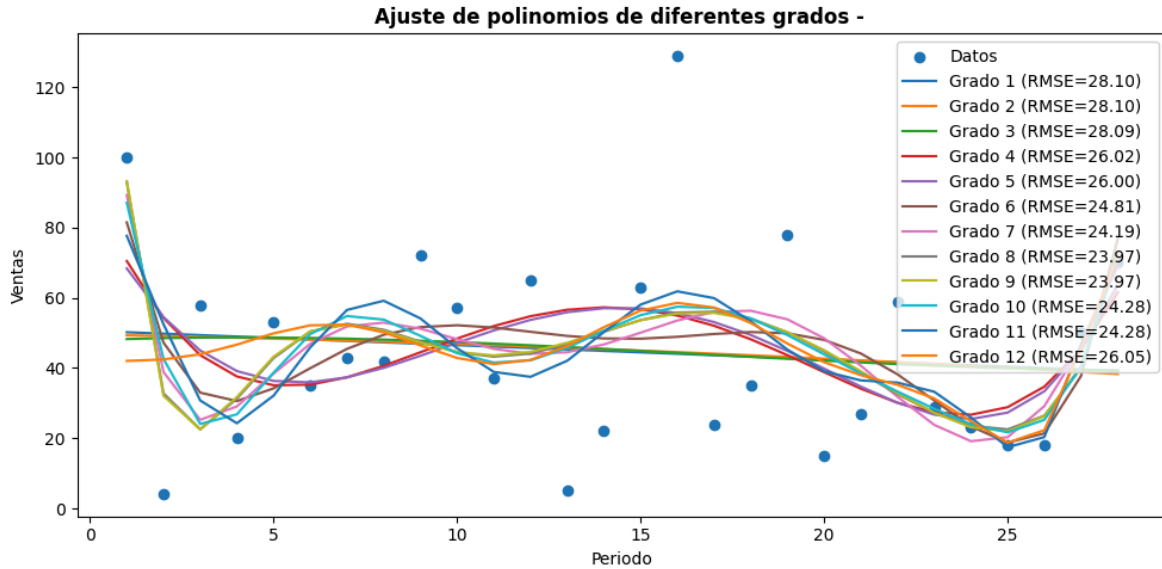


Ilustración 13 Polinomios que se pueden ajustar a la serie del producto 4

La regresión con polinomio de orden 10 para los datos de entrenamiento, arrojó unos pronósticos que al compararse con los 3 meses de validación mostraron un RMSE de 430.6.

Las pruebas de estacionariedad del modelo ARIMA dieron como resultado:

ADF Statistic: -4.958762091234618, p-value: 4.564587328732541e-05

La prueba ADF tiene un valor p bajo, lo que sugiere que se rechaza la hipótesis nula de no estacionariedad. La serie de tiempo es estacionaria.

KPSS Statistic: 0.3113458724939402, p-value: 0.1

La prueba KPSS tiene un valor p alto, indicando que no se rechaza la hipótesis nula de estacionariedad, confirmando que la serie de tiempo es estacionaria.

Luego de esto se realizó la optimización bayesiana para obtener el mejor modelo ARIMA que fue $P = 0, D = 2, Q = 0$ con un RMSE de 19.31

En la Tabla 5 se puede observar claramente que el mejor modelo para este producto fue el de Holt Winters con error = add, trend = None, seasonal = add

Modelo	Ventana de evaluación (meses)	RMSE
Media Móvil	24	27.85
Suavización Holt Winters	24	7.39

Modelo de Regresión	24	430.6
ARIMA	24	19.31

Tabla 5 Modelos y resultados para el producto 4

9.8. Análisis para el Producto 5

Para el producto 5 al igual que con el producto 3 se establecieron unas ventanas de entrenamiento de 24, 36 y 48 meses para todos los modelos, debido a que también se cuenta con información de 112 meses.

En la ilustración 14 se observa la descomposición de la serie de tiempo, notando una tendencia a la baja desde el año 2022.

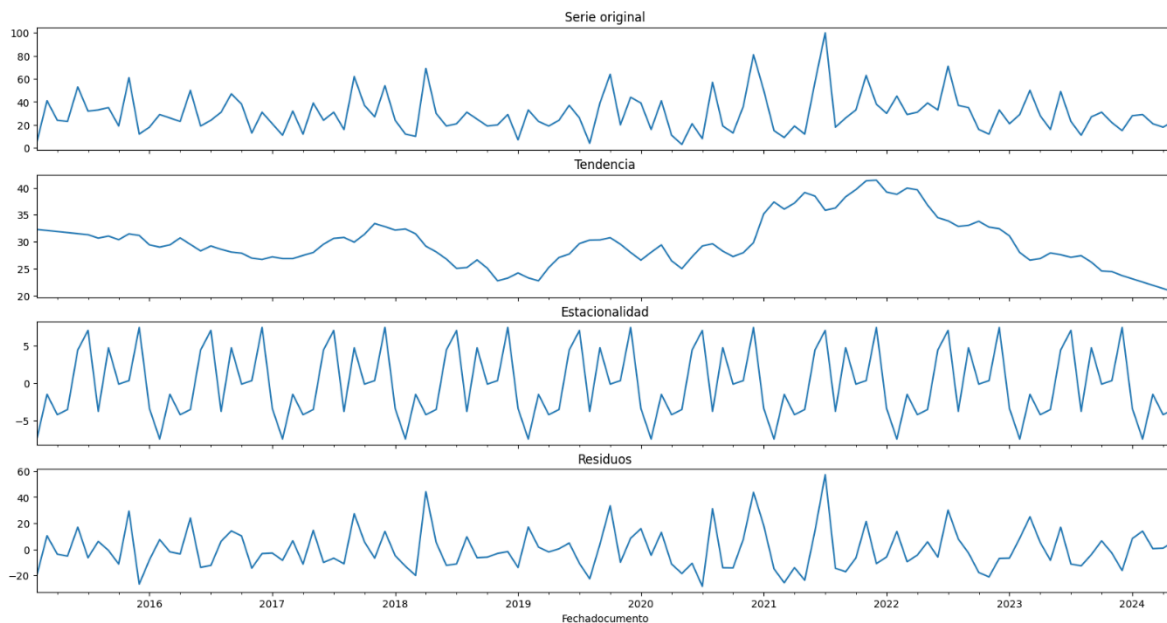


Ilustración 14 Descomposición de la serie del producto 5 en sus componentes principales

El modelo de media móvil arrojó un valor de RMSE de 5.86, este valor va a ser nuestra referencia de comparación.

El modelo de suavización exponencial Holt Winters, con optimización bayesiana incluyendo la ventana de tiempo, arrojó los siguientes resultados donde w es la cantidad de meses usados para evaluar y desplazar la ventana a lo largo de la serie:

$w = 24$, $\alpha = 0.01$, $\beta = 0.01$, $\gamma = 0.01$, error = mul, trend = mul, seasonal = None, RMSE = 7.64

$w = 36$, $\alpha = 0.01$, $\beta = 0.01$, $\gamma = 0.01$, error = add, trend = add, seasonal = None, RMSE = 6.62

$w = 48$, $\alpha = 0.39$, $\beta = 0.02$, $\gamma = 0.22$, error = mul, trend = mul, seasonal = None, RMSE = 7.77

Se realizó el ajuste de polinomios, el mejor polinomio con menor RMSE fue el de grado 7

como se observa en la Ilustración 15 y el valor de RMSE de 16.41.

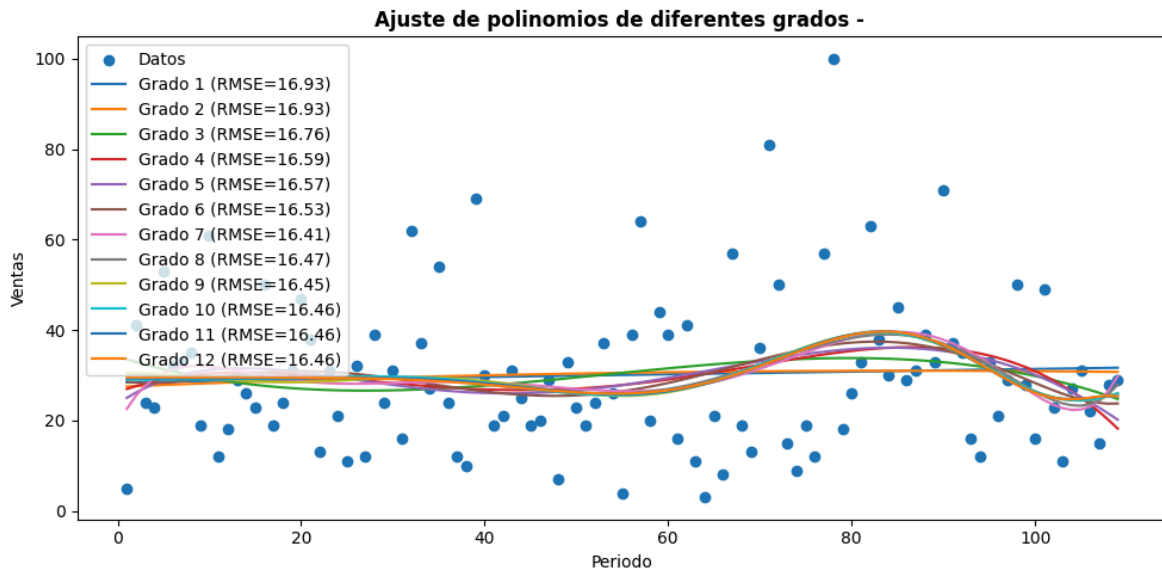


Ilustración 15 Polinomios que se pueden ajustar a la serie del producto 5

La regresión con polinomio de orden 7 arrojó los siguientes resultados:

Tamaño de ventana = 24, periodos evaluados = 3, RMSE = 416.32
Tamaño de ventana = 36, periodos evaluados = 3, RMSE = 79.34
Tamaño de ventana = 48, periodos evaluados = 3, RMSE = 11.97

Para el modelo ARIMA, las pruebas de estacionariedad arrojaron los siguientes resultados:

ADF Statistic: -5.312917632987165, p-value: 3.812345923540269e-06

La prueba ADF muestra un valor p muy bajo, indicando que se rechaza la hipótesis nula de no estacionariedad. La serie de tiempo es estacionaria.

KPSS Statistic: 0.2784510329483617, p-value: 0.1

La prueba KPSS tiene un valor p alto, lo que sugiere que no se rechaza la hipótesis nula de estacionariedad, confirmando la estacionariedad de la serie de tiempo.

Luego de obtener estos resultados, se realizó la optimización del modelo para encontrar los mejores valores que arrojó la siguiente información:

w = 24, fore = 3, P = 0, D = 1, Q = 3, RMSE = 5.96
w = 36, fore = 3, P = 0, D = 3, Q = 3, RMSE = 4.35
w = 48, fore = 3, P = 3, D = 1, Q = 0, RMSE = 4.91

Comparando las métricas de todos los modelos en la tabla 6, se observa que el mejor modelo para realizar el pronóstico es el modelo ARIMA (0,3,3)

Modelo	Ventana de evaluación (meses)	RMSE
Media Móvil	24	5.86
Suavización Holt Winters	24	7.64
	36	6.62
	48	7.77
Modelo de Regresión	24	416.32
	36	79.34
	48	11.97
ARIMA	24	5.96
	36	4.35
	48	4.91

Tabla 6 Modelos y resultados para el producto 5

10. Resultados

Para cada uno de los productos analizados, tanto la prueba ADF como la prueba KPSS indicaron que las series de tiempo son estacionarias. Los valores p bajos en las pruebas ADF y los valores p altos en las pruebas KPSS sugieren que las series no tienen una raíz unitaria (no estacionarias), confirmando que las series son adecuadas para el modelado con métodos que requieren estacionariedad, como ARIMA.

La tabla 7 muestra los resultados agrupados de los mejores modelos y sus métricas para cada uno de los productos.

Producto	Periodos	Ventanas de evaluación (meses)	Modelos evaluados			
			Media Móvil	Suavización Holt Winters	Modelo de Regresión	ARIMA
1	33	24	12.99	11.34	645.64	8.68
2	35	24	25.98	14.84	417.72	8.56
3	112	24, 36, 48	15.36	6.14	22.67	6.6
4	30	24	27.85	7.39	430.6	19.31
5	112	24, 36, 48	5.86	6.62	11.97	4.35

Tabla 7 Comparación de modelos y selección de mejores resultados

10.1. Pronósticos Producto 1

La Ilustración 16 muestra el pronóstico de ventas para el producto "CALIBRADOR CON VERNIER 6"X150mm" en los próximos meses, basado en datos históricos de ventas. La línea azul representa las ventas históricas desde septiembre de 2021 hasta mayo de 2024, mostrando fluctuaciones significativas en la demanda a lo largo del tiempo.

El pronóstico de ventas se representa con una línea naranja que comienza en junio de 2024 y se extiende hasta agosto de 2024.

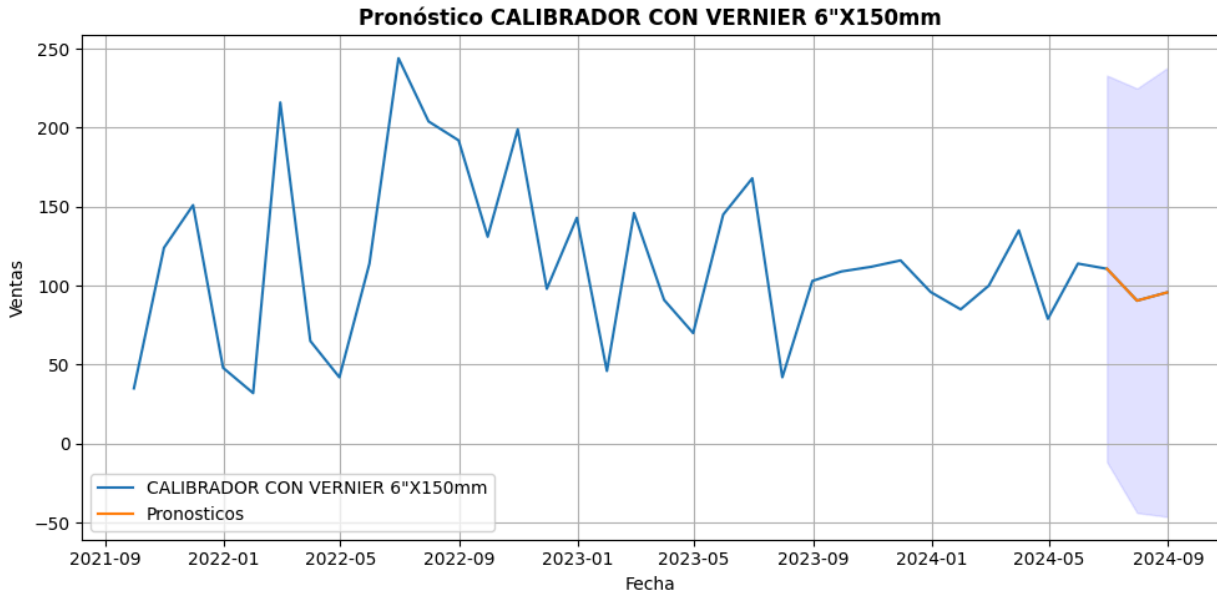


Ilustración 16 Pronósticos con intervalos de confianza para producto 1

Los puntos específicos del pronóstico son los siguientes:

Periodo	Pronóstico
2024-06-30	110.677781
2024-07-31	90.531827
2024-08-31	95.744197

Tabla 8 Pronóstico de ventas para 3 meses del producto 1

Además, la zona sombreada en color azul alrededor de la línea de pronóstico en la Ilustración 16 y en todas las ilustraciones que siguen a continuación, indica el intervalo de confianza para las predicciones, proporcionando una idea del rango dentro del cual se espera que caigan las ventas futuras.

10.2. Pronósticos Producto 2

La Ilustración 17 muestra el pronóstico de ventas para el producto "CALIBRADOR CON VERNIER 8"X200mm" en los próximos meses, basado en datos históricos de ventas. La línea azul representa las ventas históricas desde septiembre de 2021 hasta mayo de 2024,

mostrando fluctuaciones significativas en la demanda a lo largo del tiempo.

El pronóstico de ventas se representa con una línea naranja que comienza en junio de 2024 y se extiende hasta agosto de 2024.

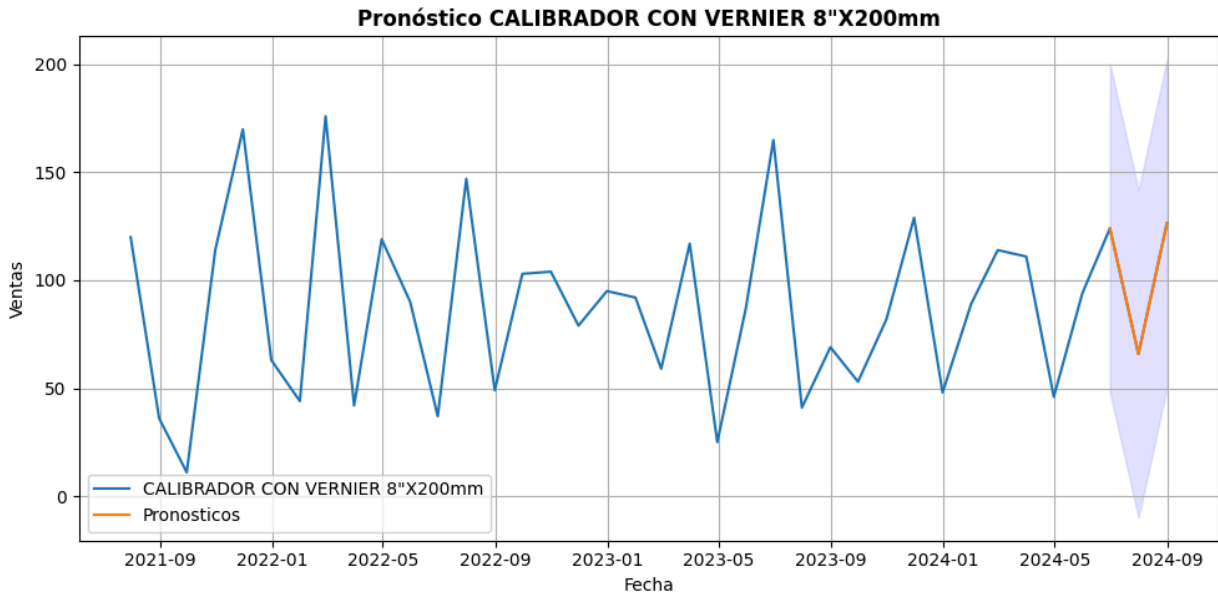


Ilustración 17 Pronósticos con intervalos de confianza para producto 2

Los puntos específicos del pronóstico son los siguientes:

Periodo	Pronóstico
2024-06-30	124.056294
2024-07-31	65.865753
2024-08-31	126.450409

Tabla 9 Pronóstico de ventas para 3 meses del producto 2

10.3. Pronósticos Producto 3

La Ilustración 18 muestra el pronóstico de ventas para el producto "CALIBRADOR PIE DE REY DIGITAL DE 6"-150mm /0.01mm-0.0005"" en los próximos meses, basado en datos históricos de ventas. La línea azul representa las ventas históricas desde el año 2015 hasta mayo de 2024, mostrando fluctuaciones significativas en la demanda a lo largo del tiempo.

El pronóstico de ventas se representa con una línea naranja que comienza en junio de 2024 y se extiende hasta agosto de 2024.

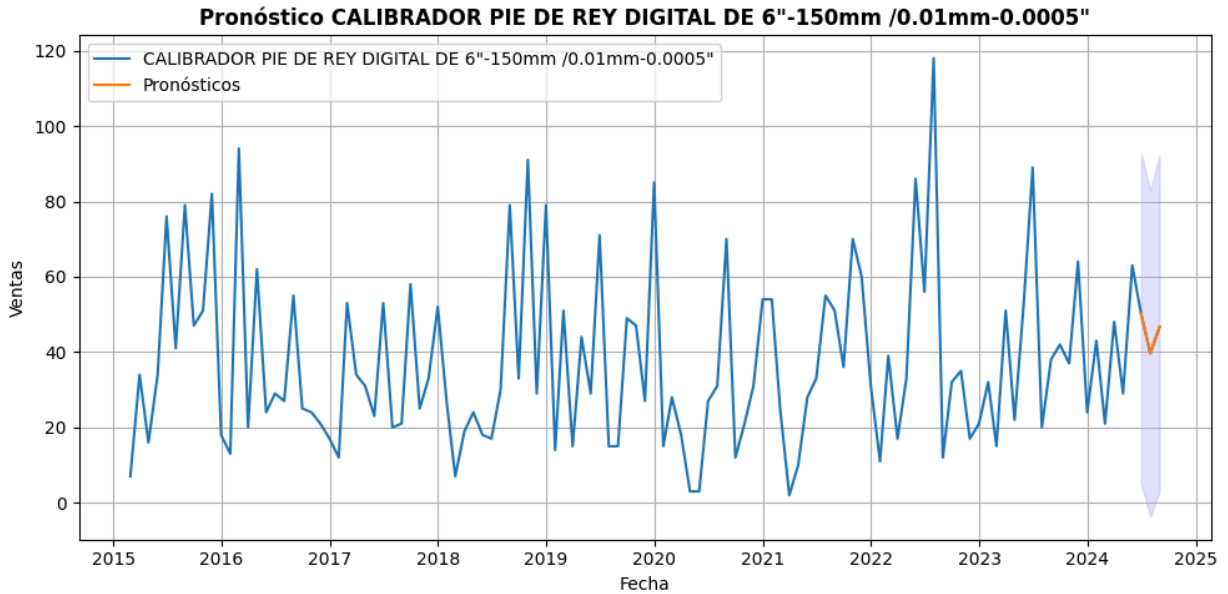


Ilustración 18 Pronósticos con intervalos de confianza para producto 3

Los puntos específicos del pronóstico son los siguientes:

Periodo	Pronóstico
2024-06-30	50.116449
2024-07-31	39.666330
2024-08-31	46.669091

Tabla 10 Pronóstico de ventas para 3 meses del producto 3

10.4. Pronósticos Producto 4

La Ilustración 19 muestra el pronóstico de ventas para el producto "COMPARADOR" en los próximos meses, basado en datos históricos de ventas. La línea azul representa las ventas históricas desde enero de 2022 hasta mayo de 2024, mostrando fluctuaciones significativas en la demanda a lo largo del tiempo.

El pronóstico de ventas se representa con una línea naranja que comienza en junio de 2024 y se extiende hasta agosto de 2024.

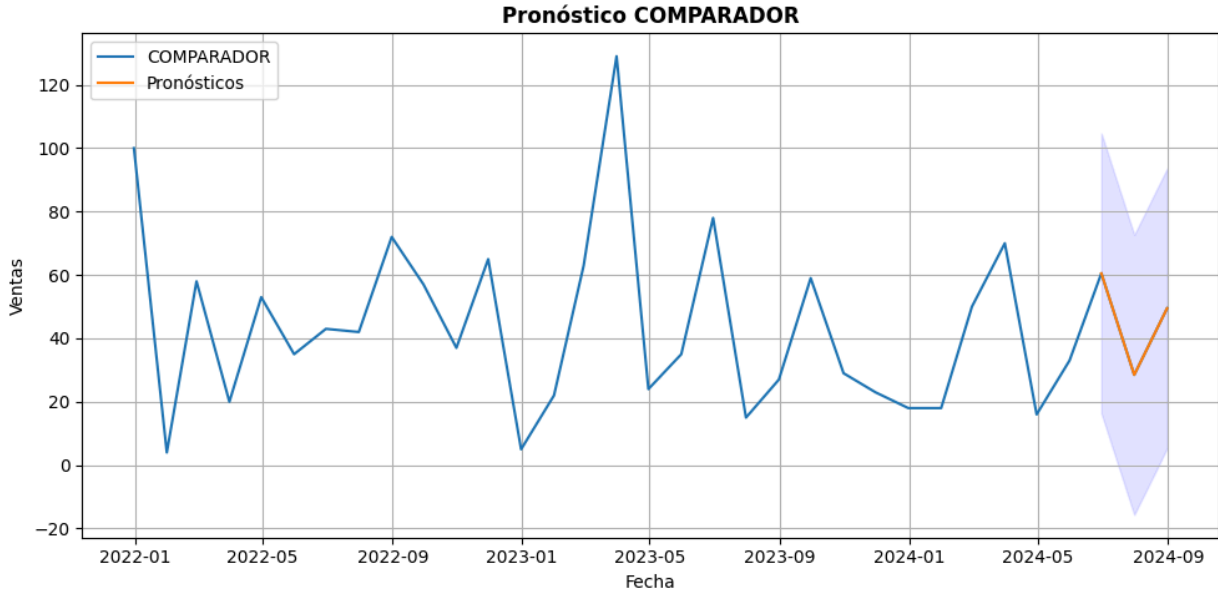


Ilustración 19 Pronósticos con intervalos de confianza para producto 4

Los puntos específicos del pronóstico son los siguientes:

Periodo	Pronóstico
2024-06-30	60.505008
2024-07-31	28.496755
2024-08-31	49.491823

Tabla 11 Pronóstico de ventas para 3 meses del producto 4

10.5. Pronósticos Producto 5

La Ilustración 20 muestra el pronóstico de ventas para el producto "CALIBRADOR PIE DE REY DIGITAL 8"-200mm/0.01mm-0.0005"" en los próximos meses, basado en datos históricos de ventas. La línea azul representa las ventas históricas desde el año 2015 hasta mayo de 2024, mostrando fluctuaciones significativas en la demanda a lo largo del tiempo.

El pronóstico de ventas se representa con una línea naranja que comienza en junio de 2024 y se extiende hasta agosto de 2024.

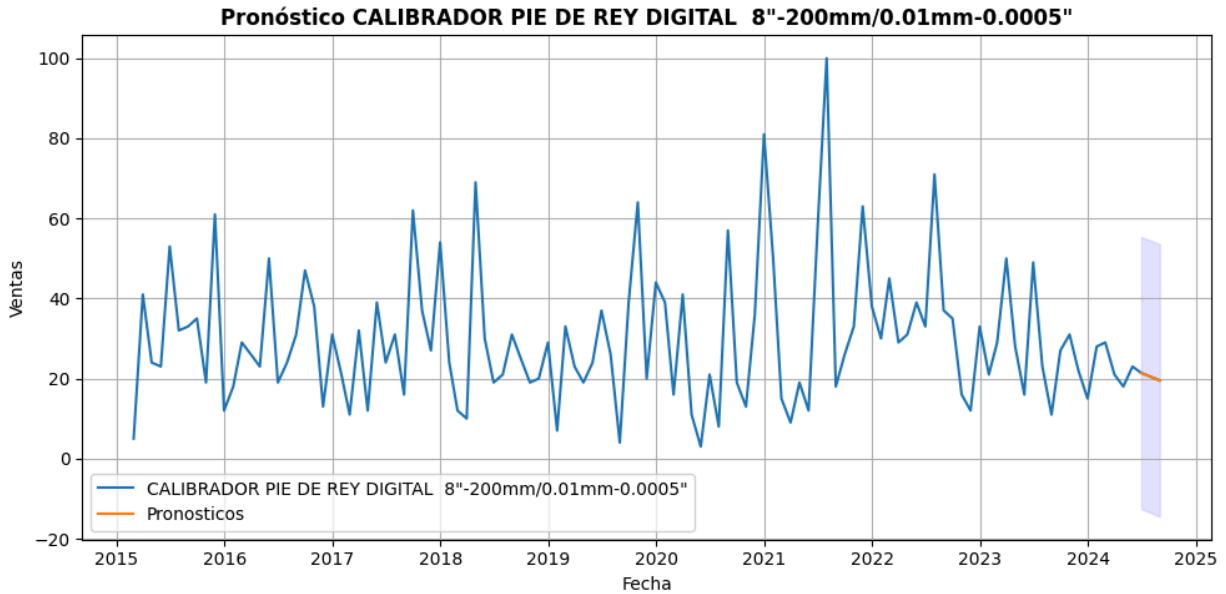


Ilustración 20 Pronósticos con intervalos de confianza para producto 5

Los puntos específicos del pronóstico son los siguientes:

Periodo	Pronóstico
2024-06-30	21.379868
2024-07-31	20.471072
2024-08-31	19.523743

Tabla 12 Pronóstico de ventas para 3 meses del producto 5

11. Desarrollo de la aplicación

Después de generar los pronósticos iniciales, el siguiente objetivo clave fue desarrollar el código necesario para ampliar el análisis y automatizar el proceso de generación de pronósticos procurando incluir más productos.

La ilustración 21 muestra el flujo de información planteado para automatizar el proceso de lectura y facilitar el uso de los pronósticos en un ambiente de producción.

Dentro del código se desarrolló la conexión a la base de datos utilizando la librería pyodbc y la librería pandas, que permiten cargar en un dataframe una consulta SQL, posteriormente se realizó el tratamiento de valores nulos y el descarte de elementos que no son productos, luego se generaron los pronósticos para cada uno de los productos.

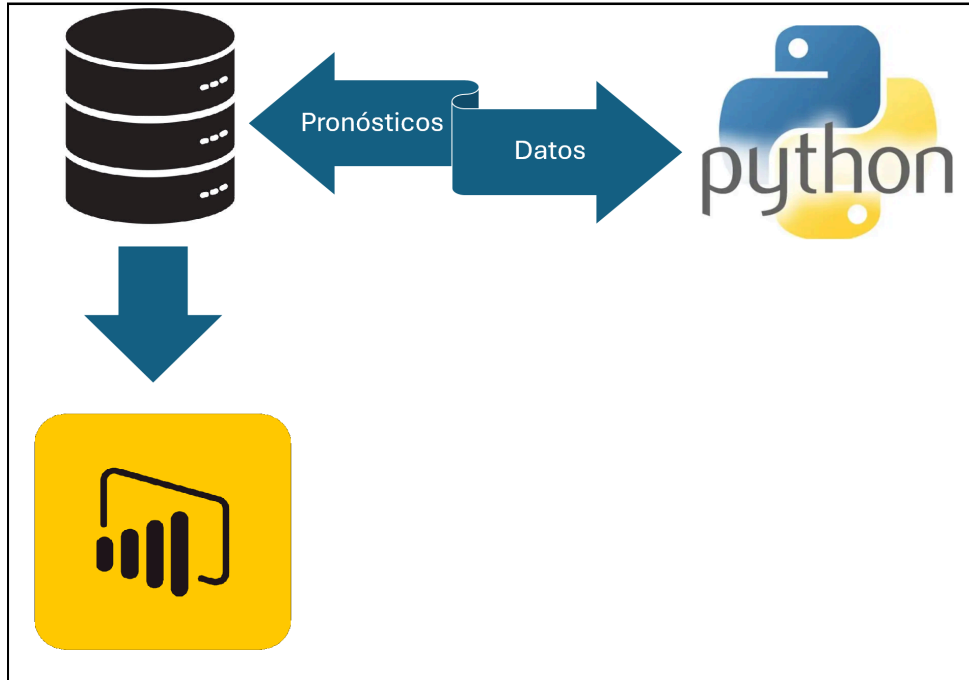


Ilustración 21 Flujo de información en desarrollo

El modelo ARIMA fue seleccionado como base para la generación de los pronósticos por encima de los otros modelos debido a su precisión y desempeño. Este modelo fue implementado para cada uno de los productos, utilizando técnicas de optimización bayesiana para configurar automáticamente los parámetros que generan los mejores pronósticos.

Luego de generar los pronósticos se escriben en una tabla en la base de datos junto con la información de su RMSE para considerar la fiabilidad del pronóstico.

El código utilizó los datos reales provenientes de la base de datos de la compañía y se cargó la información de los pronósticos en una base de datos local, se espera poner en producción más adelante una vez que el equipo de IT y el de compras den su aprobación.

12. Conclusiones

- El análisis exploratorio reveló que la información contenida en la base de datos estaba bastante ordenada y estructurada, lo que minimizó la necesidad de realizar tratamientos adicionales. La calidad y organización de los datos facilitó un flujo de trabajo más eficiente y permitió centrarnos en el desarrollo y evaluación de los modelos de pronóstico sin enfrentar los desafíos habituales asociados a la limpieza y preparación de datos.
- El modelo ARIMA ha demostrado consistentemente ser el más preciso entre los modelos evaluados para todos los productos. Esto se puede observar en los valores más bajos obtenidos por ARIMA que se observan en la Tabla 7. en comparación con los otros modelos, lo que sugiere que este modelo es particularmente efectivo para capturar las características de las series temporales en cuestión y proporcionar pronósticos más precisos.

- La Suavización Holt-Winters también ha mostrado un buen desempeño, siendo el segundo modelo más preciso en la mayoría de los casos. Especialmente en el producto 3, el modelo de Holt-Winters presentó un error muy bajo (6.14) en comparación con otros modelos. Esto indica que la Suavización Holt-Winters es una opción viable y efectiva para series temporales con patrones estacionales bien definidos.
- Los modelos de Media Móvil y el Modelo de Regresión han mostrado un desempeño inferior en comparación con ARIMA y Holt-Winters. Especialmente, el Modelo de Regresión ha generado errores significativamente altos en varios productos (por ejemplo, 645.64 en el producto 1 y 430.6 en el producto 4), lo que sugiere que estos modelos pueden no ser adecuados para las características de las series temporales evaluadas. Esto resalta la necesidad de seleccionar modelos de pronóstico que se adapten adecuadamente a la naturaleza de los datos específicos para mejorar la precisión de los pronósticos.
- Se desarrolló el código para generar pronósticos de forma automatizada y se probó en un ambiente local verificando que en el momento en que se decida integrar en producción se va a poder hacer de una forma fácil y práctica.
- Automatizar un proceso de generación de pronósticos puede no ser tan complejo en la práctica pero sí debe estar acompañado de un proceso de capacitación sobre lo que significa un pronóstico y como puede ser una herramienta que complementa el proceso de estimación y no una verdad absoluta.

13. Referencias

Alonso, J. C. (2020). Introducción a los pronósticos con modelos estadístico de series de tiempo para científico de datos (en R). Universidad Icesi.

Syntetos, A. A., Boylan, J. E., & Disney, S. M. (2009). Forecasting for inventory planning: a 50-year review. *Journal of the Operational Research Society*, 60(sup1), S149—S160. <https://doi.org/10.1057/jors.2008.173>

Molodoria. "How To Apply Machine Learning To Demand & Sales Forecasting in Retail". *MobiDev*. Accedido el 14 de abril de 2024. [En línea]. Disponible: <https://mobidev.biz/blog/retail-demand-forecasting-with-machine-learning>

W. D. Ray, "ARIMA Forecasting Models in Inventory Control", *J. Oper. Res. Soc.*, vol. 33, n.º 6, p. 567, junio de 1982. Accedido el 14 de abril de 2024. [En línea]. Disponible: <https://doi.org/10.2307/2581040>

A. P. Knopov y V. A. Pepelyaev, "Some Continuous Models of Inventory Control", *Cybern. Syst. Anal.*, vol. 41, n.º 3, pp. 465–467, mayo de 2005. Accedido el 14 de abril de 2024. [En línea]. Disponible: <https://doi.org/10.1007/s10559-005-0080-1>

R. Carbonneau, R. Vahidov y K. Laframboise, "Forecasting Supply Chain Demand Using Machine Learning Algorithms", en *Machine Learning*. IGI Glob., 2012, pp. 1652–1686.

Accedido el 14 de abril de 2024. [En línea]. Disponible: <https://doi.org/10.4018/978-1-60960-818-7.ch609>

Y. Chen, Y. Kang, Y. Chen y Z. Wang, “Probabilistic forecasting with temporal convolutional neural network”, *Neurocomputing*, vol. 399, pp. 491–501, julio de 2020. Accedido el 14 de abril de 2024. [En línea]. Disponible: <https://doi.org/10.1016/j.neucom.2020.03.011>

F. Elkarmi y N. Abu Shikhah, “Electricity Demand Forecasting”, *Int. J. Productiv. Manage. Assessment Technol.*, vol. 2, n.º 1, pp. 1–19, enero de 2014. Accedido el 14 de abril de 2024. [En línea]. Disponible: <https://doi.org/10.4018/ijpmat.2014010101>

Zeng, M. Chen, L. Zhang y Q. Xu, “Are Transformers Effective for Time Series Forecasting?”, *Proc. AAAI Conf. Artif. Intell.*, vol. 37, n.º 9, pp. 11121–11128, junio de 2023. Accedido el 15 de abril de 2024. [En línea]. Disponible: <https://doi.org/10.1609/aaai.v37i9.26317>

A. M. De Livera, R. J. Hyndman y R. D. Snyder, “Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing”, *J. Amer. Statistical Assoc.*, vol. 106, n.º 496, pp. 1513–1527, diciembre de 2011. Accedido el 15 de abril de 2024. [En línea]. Disponible: <https://doi.org/10.1198/jasa.2011.tm09771>

Y. Kang, W. Cao, F. Petropoulos y F. Li, “Forecast with forecasts: Diversity matters”, *Eur. J. Oper. Res.*, vol. 301, n.º 1, pp. 180–190, agosto de 2022. Accedido el 15 de abril de 2024. [En línea]. Disponible: <https://doi.org/10.1016/j.ejor.2021.10.024>

G. Li y H. Song, “New Forecasting Models”, *J. Travel & Tourism Marketing*, vol. 21, n.º 4, pp. 3–13, agosto de 2007. Accedido el 15 de abril de 2024. [En línea]. Disponible: https://doi.org/10.1300/j073v21n04_02

F. Ali, “Chinese GDP Forecast Using ARIMA Model”, *SSRN Electron. J.*, 2023. Accedido el 15 de abril de 2024. [En línea]. Disponible: <https://doi.org/10.2139/ssrn.4419429>

洁. 任, “Forecasting of the Scale of Employee Medical Insurance for Flexible Employers Based on GM (1,1) Model”, *Operations Res. Fuzziol.*, vol. 12, n.º 04, pp. 1400–1406, 2022. Accedido el 15 de abril de 2024. [En línea]. Disponible: <https://doi.org/10.12677/orf.2022.124148>

N. Schneidewind, “Software forecasting models”, *J. Aerosp. Comput., Inf., Communication*, vol. 6, n.º 9, pp. 540–552, septiembre de 2009. Accedido el 15 de abril de 2024. [En línea]. Disponible: <https://doi.org/10.2514/1.38053>

S. Mishra. “Resample function of Pandas”. *Medium*. Accedido el 17 de abril de 2024. [En línea]. Disponible: <https://towardsdatascience.com/resample-function-of-pandas-79b17ec82a78>