



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA NEGOCIOS INTERNACIONALES**

TEMA:

Análisis y pronóstico de las ventas para la optimización de tiempos de importación en la empresa IMPORGUIDSA mediante la aplicación del modelo de series de tiempos ARIMA

AUTOR:

Núñez Ordóñez, Daniel Roberto

**Trabajo de integración curricular previo a la obtención del título de
LICENCIADO EN NEGOCIOS INTERNACIONALES**

TUTOR:

Ing. Carrera Buri, Félix Miguel, Mgs.

Guayaquil, Ecuador

04 de septiembre del 2023



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA NEGOCIOS INTERNACIONALES

CERTIFICACIÓN

Certificamos que el presente trabajo de integración curricular fue realizado en su totalidad por **Núñez Ordóñez, Daniel Roberto**, como requerimiento para la obtención del título de **Licenciado en Negocios Internacionales**.

TUTOR

f. _____
Ing. Carrera Buri, Félix Miguel Mgs.

DIRECTOR DE LA CARRERA

f. _____
Ing. Hurtado Cevallos, Gabriela Elizabeth Mgs.

Guayaquil, a los 4 del mes de septiembre del año 2023



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA NEGOCIOS INTERNACIONALES

DECLARACIÓN DE RESPONSABILIDAD

Yo, **Núñez Ordóñez, Daniel Roberto.**

DECLARO QUE:

El Trabajo de Integración Curricular, **Análisis y pronóstico de las ventas para la optimización de tiempos de importación en la empresa IMPORGUIDSA mediante la aplicación del modelo de series de tiempos ARIMA** previo a la obtención del título de **Licenciado en Negocios Internacionales**, ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de mi total autoría.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance del Trabajo de Integración Curricular referido.

Guayaquil, a los 4 del mes de septiembre del año 2023

EL AUTOR:

f. _____

Núñez Ordóñez, Daniel Roberto



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA NEGOCIOS INTERNACIONALES

AUTORIZACIÓN

Yo, Núñez Ordóñez, Daniel Roberto

Autorizo a la Universidad Católica de Santiago de Guayaquil a la **publicación** en la biblioteca de la institución del Trabajo de Integración Curricular, Análisis y pronóstico de las ventas para la optimización de tiempos de importación en la empresa IMPORGUIDSA mediante la aplicación del modelo de series de tiempos ARIMA, cuyo contenido, ideas y criterios son de mi exclusiva responsabilidad y total autoría.

Guayaquil, a los 04 del mes de septiembre del año 2023

EL AUTOR:

f. _____
Núñez Ordóñez, Daniel Roberto



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA NEGOCIOS INTERNACIONALES

REPORTE URKUND

REPORTE URKUND 100%

URKUND

Documento	Nuñez_Ordóñez - Análisis y pronóstico de ventas para la optimización de tiempos de importación en la empresa IMPORGUIDSA mediante la aplicación del modelo de series de tiempos ARIMA.docx (D173173414)
Presentado	2023-08-27 19:16 (-05:00)
Presentado por	daniel.nunez01@cu.ucsg.edu.ec
Recibido	felix.carrera01.ucsg@analysis.arkund.com
Mensaje	Tesis - Daniel Nuñez Mostrar el mensaje completo 0% de estas 38 páginas, se componen de texto presente en 0 fuentes.

PDF viewer controls: PDF icon, zoom in, zoom out, search, back, forward.

.....
Ing. Félix Miguel Carrera Buri

AGRADECIMIENTO

Quiero empezar agradeciendo a Dios, quien ha sido mi brújula y en quien he confiado firmemente para cumplir esta meta. A mi madre, quien con su paciencia y apoyo incondicional ha logrado siempre motivarme para nunca rendirme a pesar de las dificultades que se presentaron en el camino. A mi padre, quien con su esfuerzo y sacrificio me dio la oportunidad de prepararme y conseguir este gran paso en mi vida. A mi hermano, por haber compartido y celebrado siempre mis logros como si fueran de él.

A mi tutor Ing. Félix Carrera quien, con vocación me ha dado la guía y ayuda que me han permitido desarrollar este trabajo que representará un nuevo logro en mi vida profesional. A mis compañeros, que han hecho que toda esta etapa universitaria quede como una experiencia significativa en mi vida y a todos los docentes que han dejado sus enseñanzas en este proceso.

Atentamente,
Daniel Núñez.

DEDICATORIA

Quiero dedicarle este trabajo de titulación a mi familia, por ser pilares fundamentales en mi vida y siempre estar presentes junto a mí en cada paso.

Se lo dedico a mi abuelo Roberto, que desde el cielo sé que está orgulloso y feliz de verme alcanzar mi título.

A mi enamorada, por siempre brindarme su apoyo y comprensión cuando la necesité.

Por último, me lo dedico a mí, por los sacrificios y esfuerzos que hice para hacerlo posible.

Atentamente,
Daniel Núñez.



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA DE NEGOCIOS INTERNACIONALES**

TRIBUNAL DE SUSTENTACIÓN

f. _____

Ing. Hurtado Cevallos, Gabriela Elizabeth Mgs.
DIRECTORA DE CARRERA

f. _____

Ing. Arias Arana, Wendy Vanessa Mgs.
COORDINADOR DEL ÁREA

f. _____

Ing. Padilla Lozano, Carmen Paola PhD.
OPONENTE



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA DE NEGOCIOS INTERNACIONALES**

CALIFICACIÓN

Núñez Ordóñez, Daniel Roberto

ÍNDICE GENERAL

Contenido

Introducción	2
Problemática.....	4
Justificación.....	5
Alcance.....	7
Objetivos	7
Objetivos generales	7
Objetivos específicos	7
Marco Teórico.....	8
Importaciones.....	8
Coeficiente de correlación.....	8
Series Temporales	9
Componentes de una serie de tiempo.....	9
Componente tendencia	10
Componente estacional	10
Componente aleatoria.....	10
Clasificación de las series temporales.....	10
Estacionarias	10
No estacionarias	11
Características de una serie temporal.....	11
Estacionariedad	11

Linealidad.....	12
Tendencia	12
Proceso Estocástico.....	13
MODELO ARIMA (p, d, q) (P, D, Q) s.....	15
Modelos Autorregresivos AR(p).....	16
Modelo de Medias Móviles (MA)q.....	16
Modelo Autorregresivo de Medias Móviles (ARMA) p, q.....	17
Marco Conceptual	18
Proceso estocástico.....	18
Serie Temporal	18
Media.....	18
Covarianza.....	18
Correlación.....	19
Ruido blanco	19
Ergodicidad	19
Variable aleatoria	20
Estacionariedad	20
ARIMA	20
Marco Legal	20
Sección I.....	21
Sección II	22
Metodología	23
Métodos de investigación.....	25
Normalización.....	25

Metodología y de Box y Jenkins.....	26
Identificación y Selección del Modelo:.....	26
Estimación de parámetros:	27
Validación del modelo:	27
Ruido Blanco.....	27
Proceso Autoregresivo	29
Media Móvil.....	29
Estacionariedad	30
Estacionalidad	31
Raíz Unitaria	32
Forecast	33
Análisis de resultados.....	34
library(readxl)	34
library(xts).....	34
library(dplyr).....	34
library(tidyverse).....	35
library(lubridate)	35
library(tseries)	35
library(astsa).....	36
library(forecast).....	36
library(foreign).....	37
library(timsac).....	38
library(vars).....	38

library(mFilter).....	38
library(dynlm)	39
library(nlme)	39
library(prophet)	39
library(gridExtra)	40
library(ggpubr)	40
library(ggstatsplot)	41
Línea de códigos	42
Conclusiones del Análisis de resultados	59
Conclusiones	60
Recomendaciones.....	62
Anexo 1. Base de datos de ventas IMPORGUIDSA	63
Anexo 2. Script utilizado en R Studio.....	65
Bibliografía	67

ÍNDICE TABLAS

<i>Tabla 1</i> Tipos de correlación	9
<i>Tabla 2</i> Proceso Estocástico.....	14

ÍNDICE FIGURAS

<i>Ilustración 1</i> Base de datos importada	42
<i>Ilustración 2</i> Aplicación de "colnames"	44
<i>Ilustración 3</i> Base de datos con fechas futuras.....	45
<i>Ilustración 4</i> Gráfico de valores pronosticados.....	48
<i>Ilustración 5</i> Tendencias y estacionalidades	48
<i>Ilustración 6</i> Boxplot para determinar patrones estacionales.....	51
<i>Ilustración 7</i> Gráfico de ventas en dólares	53
<i>Ilustración 8</i> Medidas de rendimiento del modelo.....	55
<i>Ilustración 9</i> Gráfico de pronóstico mediante modelo ARIMA.	58

RESUMEN

En un entorno empresarial marcado por la creciente globalización y la dinámica cambiante de los mercados, la optimización de la gestión de ventas y la eficiente administración de inventario se han convertido en pilares fundamentales para el éxito y la competitividad de las organizaciones. Este trabajo de investigación se centra precisamente en abordar estos desafíos en el contexto operativo de la importadora IMPORGUIDSA a través de un enfoque analítico basado en el modelo ARIMA (AutoRegressive Integrated Moving Average). El propósito principal de esta investigación es proporcionar una metodología sólida para analizar y proyectar las ventas de IMPORGUIDSA, permitiendo una anticipación precisa de las demandas del mercado. Con la implementación del modelo ARIMA, se busca no solo prever las tendencias futuras de ventas, sino también optimizar los tiempos de importación y la gestión del inventario en consonancia con las necesidades actuales del mercado. Para llevar a cabo este análisis, se ha empleado la herramienta R Studio, que ofrece un entorno eficiente y potente para la manipulación y visualización de datos. Mediante el análisis detallado de los datos obtenidos, esta investigación busca ofrecer una visión informada sobre las pautas de compra de los clientes, los patrones estacionales y las fluctuaciones del mercado. Los resultados obtenidos proporcionarán a IMPORGUIDSA información valiosa para tomar decisiones estratégicas informadas, adaptando su enfoque de importación y gestión de inventario para satisfacer de manera efectiva las demandas cambiantes de los consumidores.

Palabras Claves: ARIMA, Importaciones, Business Intelligence, Inventario, Ventas, Series de tiempos.

ABSTRACT

In a business environment marked by increasing globalization and changing market dynamics, sales management optimization and efficient inventory management have become fundamental pillars for the success and competitiveness of organizations. This research work focuses precisely on addressing these challenges in the operational context of the importer IMPORGUIDSA through an analytical approach based on the ARIMA (AutoRegressive Integrated Moving Average) model. The main purpose of this research is to provide a solid methodology to analyze and project IMPORGUIDSA's sales, allowing an accurate anticipation of market demands. With the implementation of the ARIMA model, we seek not only to forecast future sales trends, but also to optimize import times and inventory management in line with current market needs. With the implementation of the ARIMA model, the aim is not only to forecast future sales trends, but also to optimize import times and inventory management in line with current market needs. To carry out this analysis, the R Studio tool was used, which offers an efficient and powerful environment for data manipulation and visualization. Through detailed analysis of the data obtained, this research seeks to provide an informed view of customer buying patterns, seasonal patterns and market fluctuations. The results obtained will provide IMPORGUIDSA with valuable information to make informed strategic decisions, adapting its import and inventory management approach to effectively meet changing consumer demands.

Keywords: ARIMA, Imports, Business Intelligence, Inventory, Sales, Time Series.

RÉSUMÉ

Dans un environnement commercial marqué par une mondialisation croissante et une dynamique de marché changeante, l'optimisation de la gestion des ventes et la gestion efficace des stocks sont devenues des piliers fondamentaux pour le succès et la compétitivité des organisations. Ce travail de recherche vise précisément à relever ces défis dans le contexte opérationnel de l'importateur IMPORGUIDSA par le biais d'une approche analytique basée sur le modèle ARIMA (Moyenne mobile intégrée autorégressive). L'objectif principal de cette recherche est de fournir une méthodologie solide pour analyser et projeter les ventes d'IMPORGUIDSA, permettant une anticipation précise des demandes du marché. En appliquant le modèle ARIMA, l'objectif est non seulement de prévoir les tendances futures des ventes, mais aussi d'optimiser les délais d'importation et la gestion des stocks en fonction des besoins actuels du marché. Pour réaliser cette analyse, l'outil R Studio a été utilisé, qui offre un environnement efficace et puissant pour la manipulation et la visualisation des données. Grâce à une analyse détaillée des données obtenues, cette étude vise à fournir un aperçu éclairé des habitudes d'achat des clients, des tendances saisonnières et des fluctuations du marché. Les résultats obtenus fourniront à IMPORGUIDSA des informations précieuses pour prendre des décisions stratégiques éclairées, en adaptant son approche de l'importation et de la gestion des stocks afin de répondre efficacement aux demandes changeantes des consommateurs.

Mots-clés: ARIMA, Importations, Business Intelligence, Stocks, Ventes, Séries temporelles.

Introducción

El mercado actual cada vez va cambiando y transformándose en consecuencia de la alta competitividad, más aún con el apogeo de los últimos años respecto a la globalización. Por lo tanto, la gestión de la información en las empresas a nivel mundial es utilizada mayormente en la planeación, la operación y en la toma de decisiones. (Pupiales, 2019). En la actualidad, nos encontramos inmersos cada vez más en la era de la información y el conocimiento; esto conlleva que las gerencias de las empresas implementen varios cambios buscando innovaciones y desarrollos positivos en diferentes ámbitos. Acerca de esto, Baena declara que se las conocen como herramientas de análisis y que son las que las organizaciones utilizan para afrontar la alta competitividad en los diferentes mercados. (Pupiales, 2019)

Por consiguiente, señala (Corona Armenta et al., 2016), que las principales cabezas de las empresas necesitan información sólida y calificada, de la cual se obtenga una visión clara de todo el entorno de la empresa en los distintos niveles y por lo tanto que haya eficiencia en el uso de los recursos. Estos autores definen que, las organizaciones a nivel mundial quieren ser competitivas en el mercado y así mismo darles mayor atención a las necesidades y demandas de un mundo en plena globalización, por el contrario, las empresas perderían importantes oportunidades.

Acorde a estas necesidades, es donde podemos mencionar e integrar a la inteligencia de negocios en nuestra investigación. El Business Intelligence (BI) corresponde a las diferentes metodologías, prácticas y capacidades dirigidas al análisis de información relevantes que permitirá obtener mejores decisiones y resultados en las empresas. La práctica del BI se ejecuta y desarrolla junto a sistemas de tecnologías de la información y de un conocimiento a profundidad del giro de negocio o de la principal actividad de la empresa (Silva Solano, 2017). Dicho esto, la implementación de la inteligencia de negocios resulta un cambio positivo en contextos como los análisis de problemáticas y para óptimos procesos de toma de decisiones. (Pourshahid et al., 2011). A su vez, se ha podido evidenciar a través de los años y con mayor frecuencia que la creación de plataformas de Business Intelligence vinculadas a factores determinantes como lo es el rendimiento corporativo, resulta ser un apoyo esencial al momento de hacer frente a situaciones demandadas por toda una sociedad o mercado en particular. Para ello, se necesita tener en cuenta que, la inteligencia de negocios es un conjunto de herramientas que

ayudan al proceso de toma de decisiones en las empresas y que estas mismas sean lo más eficientes posibles(Hallikainen & Tim, 2021).

Actualmente se puede decir que la Inteligencia de Negocios es un sistema estratégico con respecto a la orientación automatizada y sistemática para la transformación y comunicación de datos (Muñoz et al., 2016). En relación, al concepto actual de una sociedad globalizada, los autores (Rodríguez & Pedraja, 2016), señalan que la administración y planeación en la mayoría de las organizaciones se rigen bajo dos aspectos relevantes e indispensables: la dirección y la gestión. La dirección empieza y aporta dando un norte, mientras que la gestión realiza su camino y se direcciona bajo ese norte.

Business Intelligence en las pequeñas y medianas empresas en Ecuador tiene unshare bastante importante, cerca del 40% de las pymes utilizan al menos una herramienta de la inteligencia de negocios, tales como Big Data, Data Analytics, Data science, Data warehousing, OAP. De hecho, en Ecuador, el uso de las herramientas de BI está distribuido por las empresas públicas con un 75% y las empresas privadas con un 69.6% (Cueva et al., 2014). Uno de los más grandes motivos de la escasez de estos sistemas dentro de las organizaciones es el costo elevado de su implementación, que presenta mayor dificultad en el proceso de desarrollo correcto y capacitación. En efecto, las actividades más destacadas que aplican la inteligencia de negocios es el sector financiero, y en promedio solamente cuentan con 3 o 4 expertos que laboran en cada una de estas empresas.

De acuerdo con una encuesta realizada por Brother International Corporation en 2016 sobre Pequeñas Empresas y Tecnología, en Ecuador, a pesar de las complicaciones y limitaciones económicas, el 30% de las Pymes ecuatorianas utiliza la inteligencia de negocios como una herramienta para lograr un incremento en su productividad e ingresos (Merino, 2017). La inversión necesaria para el desarrollo de estos sistemas o proyectos se promedia entre \$40.00 y \$60.000, y acorde a los resultados de las encuestas existe una mejora del 97% en una toma de decisiones cuando se usa el BI.

Ahora bien; en otros países de América, como en Chile, 2 de cada 3 empresas chilenas eligen implementar proyectos e iniciativas de Business Intelligence en sus operaciones(Cueva et al., 2014). Esto debido a la necesidad de las empresas por

satisfacer a sus clientes; por lo que la inclusión de sistemas de información como CRM dentro de Latinoamérica cobra mayor relevancia dentro del mercado (Muñoz et al., 2016). Algo similar, e incluso en mayor escala, sucede en Estados Unidos donde las grandes compañías tales como Amazon, Walmart y Apple emplean el Business Intelligence como una herramienta cotidiana y esencial dentro de sus operaciones.

Problemática

La aplicación de estos métodos en una manera analítica se puede tornar complicado y hasta cierto punto ser riesgoso cuando no se tiene el suficiente conocimiento en la materia que envuelve Business Intelligence y que se relaciona directamente con el entendimiento y análisis de los datos. En este caso, se podría identificar como la materia principal a la estadística; una ciencia dedicada a la recopilación, organización, representación, análisis e interpretación de información derivada de investigaciones sobre eventos, personas o grupos de individuos. Su objetivo es obtener conclusiones precisas o estimaciones anticipadas del comportamiento futuro. Por lo tanto, permite obtener un análisis lo más cercano a la realidad posible. Según la Dra. Marta Álvarez, “es como tomar decisiones a ciegas. No hay manera que una empresa sea exitosa, si no cuenta con factores de medición de las distintas variables y datos” (2018, citado en Universidad EAN, 2018, párr. 2); dado que una empresa que no identifica y ejecuta procesos estadísticos se encuentra en dificultades. Es evidente que la estadística da sentido a la información almacenada en las investigaciones y proyectos. Sin embargo, muchos profesionales no se sienten preparados o se llegan a sentir intimidados ante ella porque asumen que se trata de un campo complejo y extraño (Aroca et al., 2009).

En muchas ocasiones, las pequeñas y medianas empresas (PYMES) tienen preconceptos de que el detalle de números y datos solo son competencia de grandes empresas, las cuales mantienen un flujo de caja grande. Sin embargo, este pensamiento es un impedimento para las pequeñas y medianas empresas. De hecho, mediante la aplicación de herramientas estadísticas en sus procesos pueden competir en el mercado de una manera relevante (Universidad EAN, 2018).

Por ejemplo, con la estadística descriptiva obtenemos la orientación sobre cómo resumir de manera clara y simple los datos de una investigación en tablas, gráficos o figuras. Para llevar a cabo un análisis descriptivo, es esencial tener claro los objetivos acerca de lo investigado y también identificar las escalas de medición de las diferentes variables que se pudieron registrar (Rendón-Macías, 2016). Así mismo, incluye métodos que posibilitarán llevar a cabo un análisis básico de las observaciones y de los datos recolectados, con el propósito de llegar a conclusiones sobre cómo es el comportamiento las variables (Navarro Hudiel, 2018).

La evolución dentro del Business Intelligence ha sido significativa a través del tiempo, permitiendo de esta manera a las empresas innovar y mejorar los procesos para los cuáles se requiere la aplicación de esta herramienta. Esta evolución va sujeta por el crecimiento exponencial de la información lo cual genera el aumento por la demanda demás formas de analizar y realizar reporte de datos con la finalidad de optimizar recursos, explotar oportunidades y mejorar la toma de decisiones. Por ello, las empresas han empezado a centrar sus inversiones en varias aplicaciones empresariales; las cuales suelen ser la planeación de recursos (ERP) y administración de la relación con el cliente (CRM). Dentro de la organización estas aplicaciones han dado como resultado una gran y variada cantidad de datos. Cabe destacar que, la finalidad de las empresas con estas inversiones es que las aplicaciones les permitan una mejora en el uso de la información para brindarle la oportunidad de proyectar una eficaz toma de decisiones, agilizar organización interna y comprender cómo dirigir el negocio (Medina Soto, 2005).

No obstante, aún se puede evidenciar diversos errores en las empresas al momento de implementar la inteligencia de mercados. Estos errores no son solo en consecuencia de errores propios, sino también de malas experiencias que van desde objetivos proyectados al inicio pero que no son alcanzados, y expectativas incumplidas lo que demuestra un mal performance. Esto lleva a la poca fiabilidad de volver a plantear un proyecto en base a BI en el futuro. A continuación, se han señalado errores habituales presentes en las organizaciones, con el propósito de mejorar las formas en que se implementa y se invierten recursos económicos y humanos en el ámbito del Business Intelligence. Estos son: Enfoques netamente técnicos, mala selección del proveedor, mala calidad de datos y presupuesto inadecuado.

Justificación

La creciente demanda y accesibilidad de productos en la que nos encontramos actualmente debido a la globalización ha permitido que de igual manera las importaciones dentro del país aumenten, por lo tanto, han dado la oportunidad a los consumidores y a las empresas de elegir una mayor variedad de opciones, capacidad para operar de manera diferente y la posibilidad de acceder a productos de mejor calidad a precios más competitivos en el mercado.

Las importaciones también juegan un papel importante al proporcionar insumos y componentes clave para la producción nacional. Muchas industrias dependen de la importación de materias primas, partes y equipos especializados que no se producen localmente o son más costosos de producir internamente. Esto permite que las empresas nacionales sean más competitivas al utilizar estos insumos importados en su proceso de producción. Dicho esto, se puede afirmar que tienen un impacto de esta manera en los niveles de producción, empleo y calidad de vida de la población. Por lo tanto, es crucial establecer herramientas de política económica que permitan abordar las eventuales dificultades de liquidez (Cárdenas, 2015).

En los últimos años algunos académicos han analizado el comportamiento de las importaciones como un factor clave para el crecimiento de las economías, donde autores como Lee (1995) y Uğur (2008) han demostrado que la introducción de bienes de capital y tecnologías extranjeras incentivan el desarrollo de los países, especialmente para los que se encuentran en vía de desarrollo.

Esto se convierte en una ventaja para convertirse en distribuidor a nivel nacional de los productos importados, adentrándonos ahora en la cadena de suministro, según (Christopher, 1992) citado en (Crespo de Carvalho et al., 2014), una cadena de suministro es una red de organizaciones que están involucradas, en los diferentes procesos y actividades que producen valor en forma de productos y servicios en las manos del cliente final. Para llegar a esto se necesitará contar con una estructura logística eficiente y confiable. Así también, tener centros de distribución estratégicamente ubicados que te permitan abastecer a tus clientes de manera rápida y efectiva en todo el país. Establecer acuerdos sólidos con empresas de transporte y logística para garantizar la entrega oportuna de los productos será de vital importancia.

Alcance

Esta investigación tiene como finalidad obtener resultados que sean positivos para la optimización de los tiempos de importación dentro de la Importadora IMPORGUIDSA, para que de esta manera la proyección de ventas sea beneficiosa para un abastecimiento acorde a la demanda de los clientes a nivel nacional, convirtiéndose en una importadora competitiva dentro del mercado.

El Business Intelligence “permite a las organizaciones acceder a información que es crítica para el éxito de múltiples áreas” (Giraldo-Ríos, 2022) y por ende consiste en el proceso de producir conocimiento analizando con detalles la información abstraída de donde se encuentren los datos almacenados (Ruiz-Quintero, 2014). Por ello, este proyecto está dirigido a empresas que aún no han promovido el uso de la inteligencia artificial en sus operaciones. Para (Leit N et al, 2019) B.I ofrece las herramientas, tecnologías, aplicaciones y prácticas que lograrán satisfacer las necesidades de las empresas para analizar, extraer, visualizar y explotar la información, siendo esta una necesidad crítica para cualquier empresa. Esto sería de gran ayuda y también podría marcar un antes y un después para aquellas empresas que siguen llevando a cabo sus estrategias en base a métodos convencionales; pues la globalización e innovación han provocado una evolución en el mercado. Ante esto, el contar con inteligencia de negocios dentro de las empresas se ha convertido en un valor agregado que promueve el desarrollo competitivo.

Objetivos

Objetivos generales

Analizar mediante modelo ARIMA la proyección de ventas de productos electrónicos para optimizar los tiempos de importación.

Objetivos específicos

- Enfocar la óptima ejecución de la investigación basados en los conceptos del modelo ARIMA.

- Sustentar bajo los diferentes procesos la utilidad del modelo de proyección para las ventas e importaciones.
- Analizar los resultados obtenidos mediante el modelo planteado (ARIMA).

Marco Teórico

Importaciones

Las importaciones son acuerdos comerciales establecidos entre distintos países con el objetivo de intercambiar bienes y/o servicios que no están disponibles localmente y que son requeridos para satisfacer las necesidades de las sociedades. (Delgado Olaya & Yáñez Sarmiento , 2019).

De acuerdo con una investigación llevada a cabo por (Lahura, 2003), las importaciones de Ecuador de productos provenientes de Asia han experimentado un notable aumento a partir de 2012, siguiendo la tendencia observada en el resto de América Latina. En particular, se ha observado un incremento del 14,6% en las importaciones de productos textiles, del 11,7% en electrónica y maquinarias, y del 11,3% en equipos.

Coefficiente de correlación

El coeficiente de correlación es conocido por ser una herramienta fundamental y de mucha relevancia para los estudios econométricos de relaciones lineales bivariados que implican la utilización de datos de corte transversal o series de tiempo. Enfocándonos en el caso de las series de tiempos es necesario hacer uso de un instrumento que permita determinar la fuerza y hallar un sentido de la posible relación lineal que existe entre los pares de variables mencionados. Éste se denomina coeficiente de correlación (Lahura, 2003).

Según lo indicado por (Vinuela, 2016), la correlación es básicamente una medida normalizada que evalúa la asociación o covariación lineal entre dos variables. Esta medida, conocida como coeficiente de correlación (r), puede variar entre -1 y +1,

donde los extremos representan correlaciones perfectas negativas y positivas, respectivamente. Un valor de $r = 0$ indica que no hay una relación lineal entre las dos variables. Una correlación positiva indica que ambas variables varían en la misma dirección, mientras que una correlación negativa implica que ambas variables varían en direcciones opuestas. Lo interesante del coeficiente de correlación es que r en sí mismo proporciona una medida del tamaño del efecto, el cual suele interpretarse de la siguiente manera:

Tabla 1 Tipos de correlación

correlación despreciable: $r < 0.1 $
correlación baja: $ 0.1 < r \leq 0.3 $
correlación mediana: $ 0.3 < r \leq 0.5 $
correlación fuerte o alta: $r > 0.5 $

Series Temporales

Los métodos desarrollados para analizar estos datos también pueden ser utilizados con frecuencia en conjuntos de datos que no se recopilan a lo largo del tiempo, sino que se organizan en una dimensión lineal, como, por ejemplo, la altitud sobre el nivel del mar a lo largo de una línea de latitud. Estos métodos son efectivos en situaciones estáticas, pero la interpretación de los resultados se vuelve más complicada cuando el factor tiempo no está involucrado (Granger y Newbold, 2014).

Villavicencio (Villavicencio, s.f.) describe en su resumen teórico titulado "Introducción a las series de tiempo" que una serie de tiempo se define como un conjunto de observaciones de una variable que se recopilan en momentos específicos, se ordenan de forma cronológica y se espacian uniformemente entre sí. En general, los datos de una serie de tiempo suelen estar correlacionados entre sí. El objetivo principal del análisis de una serie de tiempo X_t , donde $t = 1, 2, \dots, n$, es realizar pronósticos.

Componentes de una serie de tiempo

Los componentes de una serie de tiempo son:

Componente tendencia

La definición de tendencia se refiere a una variación a largo plazo que ocurre en relación al nivel promedio o a un cambio a largo plazo en la media. La tendencia se caracteriza por un movimiento gradual y continuo de la serie a largo plazo.

Componente estacional

Las series de tiempo exhiben una cierta periodicidad, es decir, variaciones que se repiten en ciertos intervalos, como semestrales, mensuales, anuales, semanales, entre otros. Para ilustrar esto de manera más clara, se puede tomar el ejemplo de las ventas de una empresa, que aumentan durante los meses de noviembre y diciembre debido a las festividades navideñas. Estos efectos son fácilmente comprensibles y pueden ser medidos de forma explícita o incluso eliminados de la serie de datos mediante un proceso conocido como desestacionalización.

Componente aleatoria

Este componente no sigue ningún patrón de comportamiento predecible, sino que es el resultado de factores aleatorios o fortuitos que impactan de manera independiente en una serie de tiempo.

De estos tres componentes los dos primeros son componentes determinísticos, mientras que la última es aleatoria. Así se puede denotar la serie de tiempo como:

$$X_t = T_t + E_t + I_t$$

Donde:

$$T_t = \text{Tendencia}; E_t = \text{Componente estacional}; I_t = \text{Componente aleatorio.}$$

Clasificación de las series temporales

Las series de tiempo se clasifican en:

Estacionarias

Una serie se considera estacionaria cuando mantiene su estabilidad a lo largo del tiempo, lo que implica que tanto la media como la varianza se mantienen

constantes. Esto se puede observar gráficamente mediante la oscilación de los valores de la serie alrededor de una media constante, y la variabilidad con respecto a esa media también permanece constante a lo largo del tiempo.

No estacionarias

Las series no estacionarias son aquellas en las que tanto la tendencia como la variabilidad experimentan cambios a lo largo del tiempo. La tendencia se refiere a los cambios de la media a medida que la serie crece o disminuye a largo plazo, lo que implica que la serie no oscila alrededor de un valor constante.

Para que una serie de tiempo sea estacionaria se debe cumplir las siguientes propiedades:

$$\text{Media } E(X_t) = E(X_{t+k}) = \mu$$

$$\text{Varianza } V(X_t) = V(X_{t+k}) = \sigma^2$$

$$\text{Covarianza } \gamma_k = E[(X_t - \mu)(X_{t+k} - \mu)]$$

Características de una serie temporal

De acuerdo con Medina (Medina, y otros, s.f.), las series temporales presentan diversas características principales, entre las cuales se incluyen la estacionariedad, estacionalidad, linealidad y tendencia. Aunque una serie de tiempo puede contener una o más de estas características, se explicarán por separado con el fin de facilitar su comprensión.

Estacionariedad

La estacionariedad es una característica de un proceso aleatorio y se relaciona con la constancia del valor medio y la varianza de los datos observados a lo largo del tiempo. Por otro lado, la covarianza entre las observaciones x_t y x_{t-d} dependerá de la distancia entre las observaciones y no mostrará cambios a lo largo del tiempo. En la práctica, la estacionariedad de una serie de tiempo se verifica mediante la inspección

del patrón de los datos. (Medina, y otros, s.f.)

Linealidad

La linealidad en una serie de tiempo implica que la forma de la serie está determinada por su estado actual. En otras palabras, el patrón de una serie de tiempo lineal depende tanto de los valores presentes como de los valores pasados. Ejemplos de representaciones lineales son los modelos autoregresivos (AR), de media móvil (MA), ARMA y ARIMA, que se basan en técnicas de autorregresión y promedios móviles. Estos modelos utilizan funciones lineales para describir y predecir el comportamiento de la serie de tiempo. (Medina, y otros, s.f.). Una serie de tiempo representada por el modelo lineal se ve de la forma:

$$X_t = \sum_{i=-\infty}^{\infty} Q_i Z_{t-i}$$

Donde:

Q_i Es el conjunto de constantes que satisface la condición.

Z_t Es el ruido blanco, con un valor medio cero y varianza σ^2 .

Estadísticamente la forma multivariable de un proceso lineal se define de la forma:

$$X_t = \sum_{i=0}^n C_i Z_{t-i}$$

Donde:

C_i Representa la serie matricial $n \times n$ con elementos absolutos y

Z_t Representa el ruido blanco con valor medio cero y la matriz de covarianza.

Tendencia

Este componente en una serie de tiempo refleja una característica de largo plazo relacionada con el aumento o disminución de los valores en los datos, ya sea a nivel

local o global, debido a la superposición de los valores de la serie temporal y una perturbación con una tendencia ascendente o descendente. La presencia de este componente perturbador se puede detectar al observar los cambios en los valores medios en intervalos de tiempo sucesivos a través del patrón de la serie de tiempo. El análisis de tendencias desempeña un papel crucial en la predicción de series de tiempo y se realiza utilizando técnicas de regresión lineal y no lineal que ayudan a identificar de manera efectiva el componente de tendencia no monótona en la serie temporal (Medina et al., s.f.). Para determinar el carácter de la tendencia presente en una serie de tiempo, ya sea lineal o siguiendo una relación de un polinomio exponencial, se utiliza el ajuste de los datos recolectados.

$$X_t = \alpha t + \beta + \varepsilon_t$$

$$X_t = \text{Exp}(\alpha t + \beta + \varepsilon_t)$$

$$X_t = \alpha t + \beta t^2 + \gamma + \varepsilon_t$$

Proceso Estocástico

Se refiere a un conjunto de variables aleatorias que representan la misma magnitud en diferentes momentos de tiempo. En general, este proceso se considera lineal, lo que implica que cada variable puede obtenerse como una combinación lineal de las variables anteriores. Cuando se toma una muestra de n datos, se trata de una muestra de un vector compuesto por n variables aleatorias ordenadas en el tiempo. Se denomina proceso estocástico a este conjunto de variables (x_t) , donde $t=1\dots n$. (Carrasco Arroyo, 2018).

Un proceso estocástico es equivalente a la aleatoriedad, lo que nos permite seguir un fenómeno a lo largo del tiempo. Cada valor obtenido mediante una variable aleatoria nos proporciona información sobre cómo evoluciona el fenómeno a medida que pasa el tiempo. Cada posible valor que puede tomar la variable aleatoria se conoce como estado, y los cambios entre estados se denominan transiciones (Flores, s.f.).

Los estados de las variables aleatorias pueden ser continuos o discretos, y el tiempo puede ser discreto o continuo (Estocásticos, s.f.).

El resultado puede variar en función del tiempo y se define como $X_t = t \in T$

Donde X_t representa el número de variables o el estado del proceso en un momento específico, t representa el índice que representa el tiempo, y T se refiere al espacio de parámetros.

Tabla 2 Proceso Estocástico

	t Discreto	t Continuo
X Discreta	Un proceso con estado discreto y tiempo discreto se refiere a una cadena de eventos en la que tanto el estado como el tiempo se representan de forma discreta. Un ejemplo de esto sería el número de unidades producidas mensualmente de un producto.	Proceso de estado discreto y tiempo continuo (Proceso de saltos puros), Ej.: Unidades producidas en el momento t
X Continua	Un proceso con estado continuo y tiempo discreto implica que el estado puede tomar valores continuos, mientras que el tiempo se representa de forma discreta. Un ejemplo de esto sería la cantidad de toneladas de producción diaria de un producto.	Un proceso con estado continuo y tiempo continuo, también conocido como proceso continuo, se refiere a una secuencia de eventos en la que tanto el estado como el tiempo se representan de manera continua. Un ejemplo de esto sería la velocidad de un vehículo en un instante específico, t .

MODELO ARIMA (p, d, q) (P, D, Q) s

El modelo ARIMA es una herramienta que nos permite representar un valor como una combinación lineal de datos previos y errores aleatorios. Además, tiene la capacidad de incorporar un componente cíclico o estacional si es necesario. En otras palabras, el modelo ARIMA busca capturar todos los elementos necesarios para describir de manera adecuada el fenómeno en estudio. Según las recomendaciones de Box y Jenkins, se sugiere tener un mínimo de 50 observaciones en la serie temporal para utilizar este modelo de forma efectiva.

La metodología de Box y Jenkins se resume en cuatro fases:

La primera fase consiste en identificar el posible modelo ARIMA que sigue la serie, lo que requiere:

- Decidir qué transformaciones aplicar para convertir la serie observada en una serie estacionaria.
- Determinar un modelo ARMA para la serie estacionaria, es decir, los órdenes p y q de su estructura autorregresiva y de media móvil.

La segunda fase: Una vez que se ha seleccionado de manera provisional un modelo para la serie estacionaria, se procede a la segunda fase de estimación. En esta etapa, se utilizan métodos de máxima verosimilitud para estimar los parámetros AR (autorregresivos) y MA (media móvil) del modelo. A partir de estos estimadores, también se calculan los errores estándar correspondientes y se obtienen los residuos del modelo.

La tercera fase es el diagnóstico, donde se comprueba que los residuos no tienen estructura de dependencia y siguen un proceso de ruido blanco. Si los residuos muestran estructura se modifica el modelo para incorporarla y se repiten las etapas anteriores hasta obtener un modelo adecuado.

La cuarta fase es la predicción, una vez que se ha obtenido un modelo adecuado se realizan predicciones con el mismo.

Modelos Autorregresivos AR(p)

Un modelo autorregresivo AR es una categoría específica de proceso en el que las observaciones en un momento dado se pueden predecir utilizando las observaciones previas del proceso y un término de error. El caso más básico es el ARIMA (1,0,0) o AR (1), también conocido como modelo de primer orden, que se define matemáticamente de la siguiente manera:

$$AR(1) = X_t = \phi_1 X_{t-1} + a_t$$

En esta expresión, a_t representa el ruido blanco con una varianza de $\sigma^2 w$ y un valor medio de cero. El proceso autorregresivo de orden p se denota como ARIMA(p,0,0) o simplemente AR(p).

$$AR(p) = X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} \cdots \phi_p X_{t-p} + a_t$$

También puede ponerse mediante el operador de cambio retroactivo B , de la forma:

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) X_t = a_t$$

$$B^k(X_t) = X_{t-k}$$

Un proceso autorregresivo AR(p) se considera estacionario cuando las raíces del polinomio B , expresado como $(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)$ se encuentran fuera del círculo unidad. Esta condición es equivalente a que las raíces de la ecuación: $X^p - \phi_1 X^{p-1} - \phi_2 X^{p-2} - \cdots - \phi_{p-1} X - \phi_p = 0$ sean todas inferiores a 1.

Modelo de Medias Móviles (MA)q

Este modelo describe una serie temporal estacionaria en la que el valor actual se puede predecir principalmente a partir de la componente aleatoria en ese momento y, en menor medida, de los impulsos aleatorios anteriores. Este modelo se conoce como ARIMA (0,0,1) o también como MA (1).

Se da por la expresión:

$$X_t = \alpha_t - V_1\alpha_{t-1}$$

El proceso de Medias Móviles de orden q, representado por ARIMA (0,0, q) o simplemente MA(q), se define mediante la siguiente expresión:

$$X_t = \alpha_t - V_1\alpha_{t-1} - v_2\alpha_{t-2} - \dots - v_q\alpha_{t-q}$$

Puede ponerse, mediante el operador de cambio retroactivo B, en la forma:

$$X_t = (1 - V_1B - V_2B^2 - \dots - V_qB^q)\alpha_t$$

Un proceso de Medias Móviles (MA) de orden q se considera invertible si las raíces del polinomio en B definido por $(1 - V_1B - V_2B^2 - \dots - V_qB^q)$ se encuentran fuera del círculo unidad. Esta condición es equivalente a que las raíces de la ecuación $X^q - \phi_1X^{q-1} - \phi_2X^{q-2} - \dots - \phi_{q-1}X - \phi_q = 0$ sean todas menores a 1 en términos de su valor absoluto.

Modelo Autorregresivo de Medias Móviles (ARMA) p, q

El modelo (AR)p y (MA)q es un modelo que combina componentes autorregresivos y de medias móviles. También se puede representar utilizando la notación ARIMA (p,0, q), donde p denota el orden de los términos autorregresivos y q representa el orden de los términos de medias móviles (De la Fuente Fernández, s.f.).

Se presenta por la ecuación:

$$X_t = \phi_1X_{t-1} + \phi_2X_{t-2} + \dots + \phi_pX_{t-p} + \alpha_t - v_1\alpha_{t-1} - v_2\alpha_{t-2} - \dots - v_q\alpha_{t-q}$$

Puede ponerse de la forma:

$$X_t - \phi_1X_{t-1} + \phi_2X_{t-2} + \dots + \phi_pX_{t-p} + \alpha_t - v_1\alpha_{t-1} - v_2\alpha_{t-2} - \dots - v_q\alpha_{t-q}$$

Es decir:

$$X_t(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) = \alpha_t(1 - v_1 B - v_2 B^2 - \dots - v_q B^q)$$

Marco Conceptual

Mediante la implementación del modelo ARIMA se buscará obtener una proyección de las ventas de cada producto para de esta manera optimizar los tiempos de importación, mejorando el nivel de servicio. A continuación, se detallarán los conceptos a utilizar en el desarrollo del proyecto.

Proceso estocástico

Se trata de un modelo que caracteriza la probabilidad de una secuencia de observaciones a medida que se suceden en el tiempo (univariantes, s.f.).

Serie Temporal

Una serie de tiempo es una muestra parcial de un proceso estocástico con parámetro de tiempo discreto. (Estocástico, s.f.).

Media

La media aritmética se usa como índice de centralización en muestras grandes y variables que siguen una distribución normal y es la más utilizada.

Covarianza

En el campo de la probabilidad y la estadística, la covarianza es una medida que indica el nivel de variación conjunta entre dos variables aleatorias. Sirve como base para determinar si existe una dependencia entre las dos variables y también se utiliza para estimar parámetros como el coeficiente de correlación lineal y la recta de regresión.

Correlación

La correlación es una medida de la relación lineal entre dos variables continuas (x , y). Para determinar si dos variables están correlacionadas, se verifica si varían en conjunto, es decir, si covarían. Es importante destacar que la correlación no solo puede deberse a la casualidad, sino que también puede ser fortuita. (Vinuesa, 2016).

Ruido blanco

Se refiere a una señal aleatoria que se distingue por la falta de correlación estadística entre sus valores en diferentes momentos. Esto da como resultado una densidad espectral de potencia constante, lo que se representa gráficamente como una línea plana. Esto implica que la señal contiene todas las frecuencias posibles y cada una de ellas tiene la misma potencia. Este fenómeno se asemeja al de la luz blanca, de ahí su nombre. (Aistec, s.f.)

Ergodicidad

En matemáticas, la ergodicidad se refiere a la idea de que un punto en movimiento de un sistema, ya sea dinámico o estocástico, eventualmente visitará todas las partes del espacio en el que se mueve el sistema de manera uniforme y aleatoria. Esto implica que el comportamiento promedio del sistema se puede deducir observando la trayectoria de un punto "típico". Además, una colección lo suficientemente grande de muestras aleatorias de un proceso puede representar las propiedades estadísticas promedio de todo el proceso. La ergodicidad es una propiedad integral del sistema y significa que no se puede descomponer o factorizar en componentes más pequeños. Los sistemas ergódicos se encuentran en una amplia gama de áreas, incluyendo la física y la geometría. Esto se puede entender de manera general como resultado del movimiento de las partículas, es decir, las trayectorias en una variedad hiperbólica divergen. Cuando esta variedad es compacta, es decir, de tamaño finito, las trayectorias regresan a la misma área general y eventualmente llenan todo el espacio.

Variable aleatoria

Una variable aleatoria es un valor numérico que representa el resultado de un experimento aleatorio. Algunos ejemplos incluyen el número de caras obtenidas al lanzar una moneda seis veces, el número de llamadas recibidas por un teléfono en una hora o el tiempo de falla de un componente eléctrico. En este capítulo, abordaremos el estudio de variables aleatorias de manera similar a como lo hicimos en el capítulo anterior con variables estadísticas. Retomaremos conceptos como la distribución y características numéricas como la media y la varianza. En lugar de la frecuencia relativa, nos centraremos ahora en la probabilidad como medida relevante.

Estacionariedad

El concepto de estacionariedad juega un papel fundamental en la teoría de cointegración. En este contexto, se utiliza el concepto de estacionariedad en un sentido débil o de segundo orden, y se lo denomina simplemente como estacionariedad. Cuando se considera una serie de tiempo como una realización de un proceso estocástico, se dice que este proceso es estacionario en sentido débil si tiene momentos de primer y segundo orden finitos y si estos momentos no varían en función del tiempo.

ARIMA

El modelo ARIMA es una técnica econométrica que se fundamenta en modelos dinámicos y se aplica al análisis de datos de series temporales. (Wolters Kluwer, s.f.)

Marco Legal

Derechos y obligaciones de los consumidores

Según la Ley Orgánica de defensa del consumidor No. 2000-21.

Art 4.- Derechos del consumidor: Son derechos fundamentales del consumidor, a más de los establecido en la Constitución Política de la República, tratados o convenios internacionales, legislación interna,

principios generales de estado y costumbre mercantil, los siguientes:

1. Derecho a la protección de vida, salud y seguridad en el consumo de bienes y servicios, así como a la satisfacción de las necesidades fundamentales y el acceso a los servicios básicos.
2. Derecho a que proveedores públicos y privados oferten bienes y servicios competitivos, e óptima calidad, y a elegirlos con libertad.
3. Derecho a recibir servicios básicos de óptima calidad.
4. Derecho a la información adecuada, veraz, clara, oportuna y completa sobre los bienes y servicios ofrecidos en el mercado, así como sus precios, características, calidad, condiciones de contratación y demás aspectos relevantes de los mismos, incluyendo los riesgos que se pudiesen presentar.
5. Derecho a un trato transparente, equitativo, y no discriminatorio o abusivo por parte de los proveedores de bienes o servicios.
6. Derecho a la protección contra la publicidad engañosa o abusiva, los métodos comerciales coercitivos o desleales.
7. Derecho a la educación del consumidor, orientada al fomento del consumo responsable y a la difusión adecuada de sus derechos.
8. En el capítulo II de la ley Orgánica de regulación y control de poder de mercado.

Sección I

Art 5.- Mercado Relevante: A efecto de aplicar esta Ley la Superintendencia de Control del Poder de Mercado determinará para cada caso el mercado relevante. Para ello, considerará, al menos, el mercado del producto o servicio, el mercado geográfico y las características relevantes de los grupos específicos de vendedores y compradores que participan en dicho mercado.

El mercado del producto o servicio comprende, al menos, el bien o servicio y materia de la conducta investigada y sus sustitutos. Para el análisis de sustitución, la Superintendencia de Control del Poder de Mercado evaluará, entre otros factores, las preferencias de los clientes o consumidores; las características, usos y precios de los posibles sustitutos; los costos de la sustitución; así como las

posibilidades tecnológicas y el tiempo requerido para la sustitución.

Art. 6.- Volumen de negocios: A efectos de la presente Ley, se entiende por volumen de negocios total de uno o varios operadores económicos, la cuantía resultante de la venta de productos y de la prestación de servicios realizados por los mismos, durante el último ejercicio que corresponda a sus actividades ordinarias, previa deducción del impuesto sobre el valor agregado y de otros impuestos al consumidor final directamente relacionados con el negocio.

Sección II

Del poder de Mercado

Art. 7.- Poder de mercado: Es la capacidad de los operadores económicos para influir significativamente en el mercado. Dicha capacidad se puede alcanzar de manera individual o colectiva. Tienen poder de mercado u ostentan posición de dominio los operadores económicos que, por cualquier medio, sean capaces de actuar de modo independiente con prescindencia de sus competidores, compradores, clientes, proveedores, consumidores, usuarios, distribuidores u otros sujetos que participen en el mercado.

Art. 8.- Determinación del Poder de Mercado: Para determinar si un operador económico tiene poder de mercado en un mercado relevante, debe considerarse, entre otros, uno o varios de los siguientes criterios:

- a. Su participación en ese mercado, de forma directa o a través de personas naturales o jurídicas vinculadas, y su posibilidad de fijar precios unilateralmente o de restringir, en forma sustancial, el abastecimiento en el mercado relevante, sin que los demás agentes económicos puedan, en la actualidad o en el futuro, contrarrestar ese poder.
- b. La existencia de barreras a la entrada y salida, de tipo

legal, contractual, económico o estratégico; y, los elementos que, previsiblemente, puedan alterar tanto esas barreras como la oferta de otros competidores.

- c. La existencia de competidores, clientes o proveedores y su respectiva capacidad de ejercer poder de mercado.
- d. Las posibilidades de acceso del operador económico y sus competidores a las fuentes de insumos, información, redes de distribución, crédito o tecnología.
- e. Su comportamiento reciente.
- f. La disputabilidad del mercado.
- g. Las características de la oferta y la demanda de los bienes o servicios.
- h. El grado en que el bien o el servicio de que se trate sea sustituible, por otro de origen nacional o extranjero, considerando las posibilidades tecnológicas y el grado en que los consumidores cuenten con sustitutos y el tiempo requerido para efectuar tal sustitución.

Metodología

Dentro de este capítulo se explicarán paso a paso y a detalle cada uno de los procesos y métodos que serán utilizados dentro de la investigación y análisis de los datos sobre las ventas por parte de la importadora IMPORGUIDSA que busca como objetivo principal optimizar los tiempos de importación desde China de cada uno de los productos que mantienen mayor rotación y consumo en el inventario que son distribuidos a diferentes clientes a nivel nacional. Este proyecto e implementación dentro de la importadora ayudaría de igual manera a crear y estandarizar una estructurada política de inventario y abastecimiento.

Como parte inicial del proceso para la investigación, se extrajo una base de datos de la empresa donde se han colocado como variable las ventas en dólares durante un periodo de tiempo mensual desde el año 2018 hasta el 2023 y los tiempos de importación de los productos que, por contenedor, el shipment lleva un tiempo de tránsito aproximado de 3 meses hasta llegar a las bodegas de la importadora.

Cabe indicar que esta investigación se trata de un estudio enfocado en un aprendizaje automatizado para un análisis de serie de tiempos, esto quiere decir que es un enfoque algorítmico que tiene la capacidad de adquirir conocimiento a través del análisis de datos, sin requerir una programación basada en reglas predefinidas. En este contexto, su objetivo principal es realizar predicciones mediante el aprendizaje supervisado y no supervisado, entre otros métodos. Por el contrario, el modelado estadístico tradicional implica la representación sistemática de las conexiones entre distintas variables en los datos mediante ecuaciones matemáticas. En este contexto, las estadísticas se centran en el análisis de muestras y poblaciones, así como en la formulación y contrastación de hipótesis; concluyendo en que no existe una población o muestra posible de extraer dado que, no se cuenta con los datos históricos de importaciones desde sus inicios, por lo tanto, es posible solo mantenerlo en periodos, en este caso mensuales de los últimos 5 años.

Dentro del ámbito del análisis de series temporales, el enfoque conocido como modelo ARIMA ha demostrado ser una herramienta efectiva para el entendimiento de patrones y la realización de proyecciones futuras. Este modelo se basa en la utilización de información previa, como observaciones pasadas, tendencias y estacionalidad, con el propósito de prever valores en momentos venideros. A diferencia de otros enfoques de investigación, el ARIMA no demanda el concepto de población y muestra, ya que es un modelo univariado. Esto significa que solo toma en cuenta los valores anteriores de la variable en cuestión para realizar sus predicciones, sin incorporar datos relacionados con población o muestra (Beard, 2019)

En lugar de depender de factores externos, como el tamaño de la población o de la muestra, los modelos ARIMA buscan capturar los patrones temporales y la autocorrelación presentes en una única serie temporal (Liu, Liu, & Jiang, 2011). Una limitación importante de los modelos ARIMA es su incapacidad para abordar relaciones no lineales en los datos. Estos modelos asumen la linealidad y, por lo tanto, pueden fallar en predecir con precisión series temporales que presentan patrones o factores no lineales. En tales situaciones, la suposición de linealidad en los modelos ARIMA puede restringir su habilidad para generar pronósticos exactos (Xu, Du , & Lie, 2022).

Métodos de investigación

Normalización

La normalización de bases de datos representa un pilar esencial en la planificación de sistemas relacionales. Este concepto implica la organización metódica de los datos en una base de datos, persiguiendo la eliminación de redundancias para optimizar la integridad y eficiencia de la información. La ejecución del proceso de normalización implica la segmentación de la base de datos en múltiples tablas interconectadas. La finalidad principal radica en minimizar la repetición innecesaria de datos, asegurando que cada fragmento de información se archive en un único punto de referencia (Albarak & Bahsoon, 2018).

La normalización de bases de datos trae consigo numerosas ventajas de gran relevancia. En primer término, se eleva la integridad de los datos al mitigar la redundancia, la cual puede dar origen a inconsistencias y anomalías en la base de datos. Al suprimir la redundancia a través de la normalización, se salvaguarda la integridad de los datos y se disminuye la posibilidad de desajustes (Albarak & Bahsoon, 2018).

En segundo lugar, la normalización acrecienta la cohesión y exactitud de los datos. Cuando los datos están dispersos en diversos puntos, la labor de actualizarlos y mantener su coherencia en toda la base de datos se torna laboriosa. La normalización garantiza la centralización de cada dato, posibilitando una actualización más sencilla y una precisión mejorada (Albarak & Bahsoon, 2018).

En tercer lugar, la normalización optimiza la búsqueda de datos y el rendimiento de las consultas. Al dividir la base de datos en unidades más pequeñas y establecer conexiones entre ellas, las consultas pueden llevarse a cabo con mayor eficacia. La normalización reduce la necesidad de uniones complejas y optimiza el rendimiento general de la base de datos (Albarak & Bahsoon, 2018).

Adicionalmente, la normalización simplifica el mantenimiento y la escalabilidad de la base de datos. A medida que la base de datos crece en tamaño y complejidad, la tarea de administrar y actualizar se vuelve más compleja. La normalización simplifica la estructura de la base de datos, simplificando la gestión y adaptación a medida que cambian las necesidades de la entidad (Albarak & Bahsoon, 2018).

Metodología y de Box y Jenkins

Box y Jenkins (1976) introdujeron una nueva herramienta de pronóstico a través de su publicación "Time Series Analysis: Forecasting and Control", que más tarde se conoció como la metodología Box-Jenkins (BJ) y ahora se reconoce como la metodología ARIMA (Muñoz , Urquijo Vanstrahlengs, Castro Otero, & Lombana , 2017). (Michón Medina & Vizconde Osorio , 2011) destacan que propusieron esta metodología rigurosa con la intención de identificar, estimar y diagnosticar modelos dinámicos para datos de series de tiempo. Este enfoque predictivo se basa en el análisis de las propiedades probabilísticas o estocásticas de los datos de series temporales económicas (Gujatari, 2010).

La metodología Box-Jenkins se emplea para el proceso óptimo de construcción de modelos ARIMA, un factor que contribuye a la popularidad de ARIMA (Ghosh, 2017). Esta metodología implica estimar modelos de media móvil autorregresiva (ARIMA) para lograr el mejor ajuste para los datos longitudinales. Se utiliza ampliamente para el análisis y la previsión de series temporales y se ha convertido en una herramienta fundamental en el campo del análisis de series temporales.

Crear un modelo basado en series de tiempo no es una tarea sencilla, sin embargo, Box y Jenkins lograron desarrollar una metodología práctica para estimar modelos de series de tiempo. El método Box-Jenkins emplea una estrategia iterativa para construir modelos, implicando la selección de un modelo inicial mediante estimación de coeficientes y análisis residual (Muñoz , Urquijo Vanstrahlengs, Castro Otero, & Lombana , 2017). Destacando la necesidad que señala la metodología Box-Jenkins, de utilizar una serie estacionaria o su transformación diferenciada es fundamental cuando no hay estacionariedad (Olivia, 2016).

La metodología Box-Jenkins comprende tres etapas principales:

Identificación y Selección del Modelo:

La primera etapa consiste en identificar y seleccionar un modelo basado en el análisis de diagramas de función de autocorrelación (ACF) y función de autocorrelación parcial (PACF), asegurando que el comportamiento de la variable sea

estacionario y detectando cualquier componente estacional (Guzmán, 2019). Si la secuencia de observaciones carece de estacionariedad, se aplica la diferenciación adecuada para hacer que la serie sea estacionaria (Ghosh, 2017). En particular, este procedimiento depende del criterio personal, dada la ausencia de reglas claras para decidir sobre los componentes AR y MA apropiados. Como resultado, la experiencia juega un papel fundamental en este contexto (Rosales, 2022).

Estimación de parámetros:

La segunda fase implica la estimación de parámetros mediante la aplicación de algoritmos para derivar coeficientes para el modelo lineal que mejor se ajuste a las condiciones variables reales. Esto se puede lograr mediante estimación de máxima verosimilitud o mínimos cuadrados no lineales (Guzmán, 2019).

Validación del modelo:

Por último, la tercera fase consiste en validar y refinar el modelo mediante pruebas de independencia residual, donde sus medias y varianzas deben permanecer constantes en el tiempo o apegarse a una distribución normal (Guzmán, 2019). El modelo construido se puede probar mediante diagnósticos analíticos (Ghosh, 2017). Normalmente, la verificación del diagnóstico se lleva a cabo generando el conjunto residual y probando si presenta las características del ruido blanco.

Ruido Blanco

Ruido Blanco El término "ruido blanco" se refiere a un proceso que involucra una secuencia de variables aleatorias con una media de cero, varianza constante y covarianzas nulas. En el contexto de pronósticos, el ruido blanco representa una señal aleatoria que exhibe una densidad espectral de potencia uniforme en todas las frecuencias (Ade, 2019). Este fenómeno constituye un tipo particular de proceso estocástico, caracterizado por variables aleatorias independientes y no correlacionadas, con una media de cero y una varianza constante (Fryz & Scherbak, 2023). En el análisis de series temporales, el ruido blanco encapsula el componente impredecible y aleatorio de los datos que el modelo no puede explicar ni anticipar.

Además, el ruido blanco se utiliza frecuentemente como un punto de referencia para la evaluación del desempeño de los modelos de pronóstico (Kurov, Sancetta, Strasser, & Wolfe, 2018). Si un modelo de pronóstico no logra superar la precisión de un pronóstico basado en ruido blanco, esto sugiere que el modelo no logra identificar patrones o relaciones significativas en los datos.

En algunos casos, se introduce ruido blanco en un modelo de pronóstico para considerar las fluctuaciones aleatorias e impredecibles inherentes a los datos (Li & Guo, 2021). Esta práctica puede mejorar la capacidad del modelo para capturar la variabilidad y la incertidumbre presentes en las series temporales.

La importancia del ruido blanco en los pronósticos radica en diversas razones. En primer lugar, se emplea como punto de referencia para evaluar la eficacia de los modelos de pronóstico. Si un modelo no puede superar la precisión de un pronóstico basado en ruido blanco, esto sugiere que el modelo no logra detectar pautas o relaciones significativas en los datos (Lawrence, Goodwin, O'connor, & Onkal, 2006).

Además, el ruido blanco simboliza el componente impredecible y aleatorio de los datos que el modelo no puede abordar ni predecir. La inclusión del ruido blanco en los modelos de pronóstico ayuda a considerar la variabilidad y la incertidumbre inherentes a las series temporales (Ragozin, 2022). Esto puede dar lugar a pronósticos más sólidos y precisos, particularmente en situaciones donde los datos exhiben fluctuaciones aleatorias o ruido (Angwin, Wilson, & Arnott).

Así mismo, el ruido blanco se aplica en diversas técnicas y modelos estadísticos para mejorar los resultados de los pronósticos. Por ejemplo, en el ámbito de la econometría, es común asumir que los residuos generados por los modelos de pronóstico siguen un patrón de ruido blanco (Chlebus et al., 2021). Esta suposición contribuye a asegurar que los errores de pronóstico carezcan de sesgos y presenten características aleatorias (Chlebus et al., 2021).

Por último, el ruido blanco cobra relevancia en el contexto de la incertidumbre del modelo y la estimación de parámetros. Los modelos estocásticos que incorporan ruido blanco pueden reflejar con mayor precisión las incertidumbres y fluctuaciones del sistema modelado. Al considerar el ruido blanco como una fuente de incertidumbre, es posible llevar a cabo pronósticos más realistas y confiables (Arnold et al., 2013).

Proceso Autoregresivo

En esta fase se implementa el proceso de auto-regresión con el propósito de discernir si la serie bajo estudio requiere de diferenciación o integración, lo cual indica cuantas veces es necesario transformar la serie en una forma estacionaria. Un modelo se considera auto-regresivo si la variable endógena en un período "t" es explicada por sus propias observaciones correspondientes a períodos anteriores, agregando, al igual que en los modelos estructurales, un término de error. En el contexto de procesos estacionarios con una distribución normal, la teoría estadística de los procesos estocásticos establece que, bajo condiciones específicas, cada Y_t puede expresarse como una combinación lineal de sus valores pasados (parte sistemática), más un término de error (innovación). (Arce & Bahía, 2001, p. 5).

Los procesos de autoregresión presentan una función de autocorrelación parcial abreviada como ACFP, con un número limitado de valores distintos de cero. Un proceso AR(p) tiene los primeros "p" términos de la función de autocorrelación parcial distintos de cero, mientras que los restantes son nulos. No obstante, en la práctica, se considera que una muestra particular proviene de un proceso de auto-regresión de orden "p" si los términos de la función de autocorrelación parcial son cercanos a cero a partir del término que ocupa la posición "p" (Casimiro, 2009).

Los modelos de auto-regresión se abrevian mediante la palabra "AR", seguida por el orden del modelo: AR(1), AR(2), etc. El orden del modelo indica la cantidad de observaciones retrasadas de la serie temporal analizada que se involucran en la ecuación.

Media Móvil

El proceso de promedio móvil constituye una parte integral del modelo ARIMA (Media Móvil Autoregresiva Integrada). En el marco del modelo ARIMA, el proceso de promedio móvil (MA) constituye uno de los tres subprocesos esenciales, conjuntamente con la autorregresión (AR) y la operación de diferenciación, los cuales contribuyen a forjar una serie temporal estacionaria (Chadsuthi et al. , 2015).

La operación de promedio móvil en ARIMA implica la utilización de los términos residuales o de error de periodos anteriores para pronosticar el valor actual de la serie temporal. Esta técnica postula que el valor presente de la serie es una combinación lineal de los términos residuales anteriores y un término constante. El proceso de promedio móvil se emplea para capturar las conexiones a corto plazo y las variaciones aleatorias propias de los datos.

Este método resulta especialmente valioso en la modelización y predicción de series temporales de datos que exhiben fluctuaciones impredecibles y aleatorias. Contribuye a incorporar el ruido o el componente aleatorio de los datos que no pueden ser explicados por las partes autorregresivas o las tendencias.

La determinación y elección adecuada del orden del proceso de promedio móvil en el contexto del modelo ARIMA puede lograrse mediante diversas técnicas estadísticas, como el Criterio de Información de Akaike (AIC) o el Criterio de Información Bayesiana (BIC). Estos criterios colaboran en identificar el orden óptimo que minimiza los errores residuales y brinda el mejor ajuste a los datos.

Estacionariedad

Para aplicar modelo de pronóstico ARIMA, es fundamental determinar si la serie en cuestión es adecuada para este tipo de modelo. En el análisis de series temporales, el objetivo es emplear la teoría de los procesos estocásticos para discernir qué proceso estocástico ha generado la serie temporal en estudio, con el propósito de caracterizar su comportamiento y habilitar predicciones futuras. Para lograr métodos de predicción coherentes, no cualquier proceso estocástico es válido, sino que se requiere que la estructura probabilística sea estable a lo largo del tiempo. La base de la teoría de pronósticos sigue siendo la misma: se extraen patrones del comportamiento histórico de la serie y se extrapolan hacia el futuro. Por lo tanto, Es esencia l que los procesos estocásticos subyacentes a las series temporales exhiban algún nivel de estabilidad. En caso contrario, si presentan un comportamiento cambiante e inestable en cada momento temporal, no resultan idóneos para la predicción. A estas condiciones requeridas para que los procesos estocásticos sean aptos para la predicción se les denomina estacionariedad.

El concepto de estacionariedad se puede caracterizar de dos maneras: en términos de la función de distribución o en relación a los momentos del proceso. La función de distribución de un proceso estocástico engloba todas las funciones de distribución para cualquier subconjunto finito de variables aleatorias del proceso. Los momentos del proceso se refieren al conjunto de medios de todas las variables aleatorias del proceso, la varianza de todas las variables aleatorias y la covarianza entre todas las posibles parejas de variables aleatorias en estudio. En el primer enfoque, se habla de estacionariedad estricta, mientras que en el segundo se hace referencia a la estacionariedad de segundo orden o en términos de covarianza (González, 2009).

Para determinar si una serie está estacionaria, se pueden calcular las medias y variaciones año tras año. Si se observan cambios significativos y fluctuaciones ascendentes y descendentes a lo largo de los años, esto indica la ausencia de estacionariedad. En tal caso, se procede a aplicar logaritmos ya diferenciar la serie original para mitigar la falta de estacionariedad en términos de media y varianza.

La estacionariedad en términos de varianza se logra cuando se puede asumir que existe una única varianza para toda la serie temporal, lo que implica que la variabilidad en torno a la media permanece constante con el tiempo. Si la serie carece de estacionariedad en términos de varianza, se recurre a transformaciones que estabilizan la varianza, como las transformaciones Box-Cox (González, 2009).

La familia de transformaciones Box-Cox engloba una variedad infinita de funciones, como la raíz cuadrada o la inversa. En la práctica económica, la transformación logarítmica es la más empleada. Los métodos utilizados para analizar la estacionariedad en términos de variación incluyen el gráfico de la serie original y de las correspondientes transformaciones. Si la serie no es estacionaria en términos de medios, se logra la estacionariedad mediante la toma sucesiva de diferencias de orden 1 hasta obtener una serie estacionaria.

Estacionalidad

Si la serie de tiempo muestra una tendencia, el primer paso es convertirla en una serie estacionaria a través de una diferenciación de orden d . Una vez que la serie ha sido diferenciada, una estrategia efectiva radica en contrastar los correlogramas de

la función de autocorrelación (ACF) y la función de autocorrelación parcial (ACFP). Este procedimiento combinado proporciona una dirección para la formulación preliminar del modelo.

Raíz Unitaria

La prueba de Dickey-Fuller es una herramienta estadística empleada en la construcción de modelos ARIMA para evaluar la estacionariedad de una serie de tiempo (Wang et al., 2018). La propiedad de estacionariedad es esencial para los modelos ARIMA, ya que estos están formulados para trabajar con datos temporales que mantienen constantes su media, varianza y autocovarianza a lo largo del tiempo. En busca de una señal de no estacionariedad, la prueba Dickey-Fuller identifica una posible raíz unitaria en la serie temporal. En este contexto, una serie temporal se considera no estacionaria si posee una raíz unitaria, que es la hipótesis nula de la prueba. Si el resultado del estadístico de la prueba es considerablemente diferente de cero y el valor p es menor que cierto umbral (generalmente 0,05), se descarta la hipótesis nula,

La versión mejorada de esta prueba, conocida como prueba Dickey-Fuller aumentada (ADF), incorpora factores adicionales como la tendencia y la estacionalidad en el análisis de las series temporales (Wang et al., 2018). Esta prueba ADF se emplea habitualmente en la construcción de modelos ARIMA para examinar la estacionariedad y determinar el nivel de diferenciación (d) necesario para lograr la estacionariedad en la serie temporal.

En el proceso de modelado ARIMA, la prueba Dickey-Fuller se aplica generalmente a la serie de tiempo original. Si se establece que la serie no es estacionaria, se procede a aplicar sucesivas diferenciaciones hasta que se logre la estacionariedad (Wang et al., 2018). El valor de d , que denota el número de diferenciaciones, es un parámetro clave en los modelos ARIMA y se establece con base en los resultados obtenidos de la prueba Dickey-Fuller.

Estas pruebas estadísticas, como la prueba Dickey-Fuller o la prueba Phillips-Perron, son esenciales para determinar si una serie de tiempo contiene una raíz unitaria y, por lo tanto, no es estacionaria. Si la hipótesis nula que sugiere la presencia de una

raíz unitaria se rechaza, es posible emplear métodos adecuados para modelar datos estacionarios. En caso contrario, se pueden aplicar técnicas de diferenciación para transformar la serie temporal en una forma estacionaria antes de llevar a cabo el modelado (Villegas, 2020).

La abreviatura ARIMA, que corresponde a una media móvil integrado autorregresiva, se origina de los componentes clave del modelo:

- AR(p): Autoregresión. Este componente conecta la observación actual con un cierto número de observaciones pasadas.
- I(d): Integración. Consiste en la diferenciación de las observaciones crudas para lograr la estacionariedad temporal.
- MA(q): Medios móviles. Se utilizan para suavizar los datos de series de tiempo y reducir el ruido en los mismos.

En la notación estándar ARIMA(p,d,q), los parámetros p, d y q toman valores enteros que denotan el número de observaciones pasadas, el nivel de diferenciación y el tamaño de la ventana de medios móviles, respectivamente. La selección de los valores adecuados para estos parámetros es crucial en la construcción del modelo ARIMA. Algunos de estos valores pueden ser 0, indicando que no se utiliza ese componente específico en el modelo. Por tanto, el modelo ARIMA puede adaptarse para emular tanto la funcionalidad de un modelo ARMA como de un modelo AR, I o MA básico.

Forecast

Lo que caracteriza el pronóstico en el modelo ARIMA radica en aprovechar la información de los valores pasados y actuales observados para anticipar los valores futuros. La etapa de pronóstico representa la última fase en un enfoque de predicción mediante el modelo ARIMA. Se elige el modelo ARIMA que minimiza la métrica de error (Ayala, 2021), esta métrica utiliza los patrones previos generados por las variables pasadas para producir el pronóstico. En esta etapa, se determina el horizonte de pronósticos, es decir, el número de períodos por adelantado para los cuales se pretende prever, y este horizonte debe ser consistente con los intervalos de tiempo presentes en la serie original.

Se emplea un algoritmo que busca el modelo ARIMA óptimo para los datos de la serie temporal. Este proceso se basa en distintos criterios de selección de modelos, como el Criterio de Información de Akaike (AIC). El algoritmo ajusta diversas combinaciones de modelos ARIMA con diferentes valores de los parámetros p , d y q . Finalmente, se selecciona aquel modelo que minimiza la métrica de error predefinida, lo que garantiza que el pronóstico sea más confiable y preciso (Amarís et al., 2017).

Análisis de resultados

Para dar inicio al análisis y proyección de ventas dentro de la importadora mencionada, hemos procedido a ejecutar el desarrollo en R Studio de la base de datos obtenida. A continuación, empezaremos a describir cada función acorde a lo implementado en nuestro script.

Librerías

library(readxl)

La librería "readxl" es una herramienta en R que se utiliza para leer y cargar datos en formato Excel (.xlsx) en un entorno de trabajo de R. Esta librería facilita la importación de hojas de cálculo y datos almacenados en archivos Excel en tus análisis y proyectos.

library(xts)

La librería "xts" en R (eXtensible Time Series) es una herramienta diseñada específicamente para manejar y analizar series de tiempo en entornos financieros y de datos temporales en general.

library(dplyr)

La librería "dplyr" dplyr es una gramática de manipulación de datos que proporciona un conjunto consistente de verbos para transformar datos de una manera eficiente y sencilla. "dplyr" proporciona un conjunto de funciones que permiten realizar operaciones comunes de manejo de datos de manera intuitiva y coherente.

library(tidyverse)

El tidyverse es un conjunto integrado de paquetes que trabajan en conjunto gracias a que comparten formatos de datos y estructuras de programación similares. El propósito del paquete tidyverse es simplificar la instalación y carga de los paquetes clave incluidos en tidyverse, lo que se logra con una única instrucción.

Al instalarlo se cargarán los paquetes principales de tidyverse:

1. ggplot2, para visualización de datos.
2. dplyr, para la manipulación de datos.
3. tidyr, para ordenar los datos.
4. readr, para la importación de datos.
5. purrr, para programación funcional.
6. tibble, para tibbles, una nueva versión moderna de los marcos de datos.
7. stringr, para cadenas.
8. forcats, para factores.
9. lubridate, para fecha/hora.

library(lubridate)

Lubridate proporciona herramientas que facilitan el análisis sintáctico y la manipulación de fechas. Las funciones de análisis sintáctico de Lubridate leen cadenas en R como objetos fecha-hora. Se debe elegir la función cuyo nombre modele el orden en el que aparecen los elementos año ('y'), mes ('m') y día ('d') en la cadena que se va a analizar.

library(tseries)

La librería "tseries" en R Studio es una herramienta utilizada para el análisis de series de tiempo. Ofrece una variedad de funciones y métodos para trabajar con datos temporales y realizar análisis estadísticos en serie de tiempo.

Se emplea la ecuación general de regresión que incluye una constante y una tendencia lineal, y se determina el valor estadístico "t" para un coeficiente autorregresivo de primer orden igual a uno. El número de rezagos aplicados en el proceso de regresión se denota como "k". El valor predeterminado, que es $\text{trunc}((\text{length}(x)-1) \wedge (1/3))$, representa el límite superior recomendado para el crecimiento de "k" conforme al tamaño de la muestra, en el caso de ARMA(p)-1. Para la configuración general ARMA(p,q), los valores de "p" se derivan mediante interpolación a partir de la Tabla 4.2 de la página 103 de Banerjee et al. (1993). Si el resultado de la estadística calculada excede los valores críticos disponibles en la tabla, se emite una advertencia a modo de mensaje.

library(astsa)

Incluye conjuntos de datos y secuencias de comandos que permiten analizar series temporales en los ámbitos de la frecuencia y el tiempo, abarcando el modelado de espacios de estados, y brinda respaldo a los libros "Time Series Analysis and Its Applications: With R Examples" (4ª ed., 2017) y "Time Series: A Data Analysis Approach Using R" (1ª ed., 2019).

library(forecast)

Esta viñeta del paquete R Forecast es una versión actualizada de Hyndman y Khandakar (2008), publicada en el Journal of Statistical Software. Los pronósticos automáticos de un gran número de series temporales univariantes son a menudo necesarias en el ámbito empresarial y otros contextos. Se describen dos algoritmos de "automatic forecasting" que se han implementado en el paquete de forecast para R. El primero se basa en modelos de espacio de estados innovadores que respaldan métodos de suavizado exponencial. El segundo es un algoritmo de pronóstico escalonado con modelos ARIMA. Los algoritmos son aplicables tanto a datos estacionales como no estacionales, y se comparan e ilustran utilizando cuatro series temporales reales.

También se describen brevemente otras funciones disponibles en el paquete de forecast.

Esta librería también proporciona una amplia gama de funciones y métodos para realizar análisis de series de tiempo y generar pronósticos futuros basados en patrones pasados. Es especialmente útil en el ámbito de la ciencia de datos, la planificación y la toma de decisiones, ya que permite prever cómo se comportarán los datos temporales en el futuro.

Algunas de las principales capacidades de la librería "forecast" incluyen:

1. Modelos de series de tiempo: La librería permite ajustar diversos modelos de series de tiempo, como modelos ARIMA (AutoRegressive Integrated Moving Average), modelos de suavizamiento exponencial y más avanzados, para modelar los patrones y las tendencias en los datos.
2. Generación de pronósticos: Puedes utilizar los modelos ajustados para generar pronósticos y estimaciones de cómo se comportarán los datos en un período futuro.
3. Evaluación de pronósticos: "forecast" proporciona herramientas para evaluar la calidad de los pronósticos generados, lo que es esencial para determinar la precisión del modelo.
4. Visualización: La librería incluye funciones para crear gráficos y visualizaciones de los pronósticos, lo que facilita la interpretación de los resultados.
5. Identificación de patrones estacionales y tendencias: Puedes utilizar "forecast" para analizar las estacionalidades y tendencias presentes en los datos temporales.
6. Predicciones múltiples: La librería permite generar no solo pronósticos puntuales, sino también intervalos de confianza para estimar la incertidumbre asociada a las predicciones.
7. Modelado de incertidumbre: "forecast" también es capaz de modelar la incertidumbre en las predicciones, lo que es importante para tomar decisiones informadas.

library(foreign)

En R Studio esta librería se utiliza para leer y escribir datos en diferentes formatos de archivos que no son nativos de R. Permite a los usuarios importar y exportar datos desde y hacia R en formatos como SPSS, SAS, Stata y otros, que son bastante utilizados en la investigación y análisis de datos en diversos campos.

library(timsac)

El TIMSAC (TIME Series Analysis and Control) representa un conjunto integral de software diseñado para el análisis, predicción y regulación de series temporales. Su creación tuvo lugar en el Institute of Statistical Mathematics. La versión original, conocida como TIMSAC-72, se presentó en el trabajo de Akaike y Nakagawa (1972). Posteriormente, surgieron variantes como TIMSAC-74, TIMSAC-78 y TIMSAC-84, agrupadas en la serie TIMSAC, publicada en Computer Science Monograph 1. Esta serie engloba múltiples programas diseñados para aplicaciones prácticas en el análisis de datos, tales como la gestión óptima de procesos industriales o el estudio de las fluctuaciones económicas.

library(vars)

Se utiliza para trabajar con modelos vectoriales autorregresivos (VAR, por sus siglas en inglés) en el análisis de series temporales multivariadas como estimación, selección de lag, pruebas de diagnóstico, pronóstico, análisis de causalidad, descomposición de la varianza del error de pronóstico y funciones de respuesta al impulso de modelos VAR y estimación de modelos SVAR y SVEC.

library(mFilter)

mFilter es una función genérica que despliega su funcionalidad a través de una metodología versátil para la filtración de información en series temporales. Esta función llama a filtros específicos según el tipo de filtro designado mediante su argumento "filter". Dentro del paquete mFilter, estos filtros han sido diseñados para desempeñar un papel destacado en la suavización y la estimación de componentes relacionados con tendencias y ciclos. Algunos de estos filtros son ampliamente

empleados en los campos de economía y finanzas para calcular el componente cíclico en las series temporales.

Actualmente, `mFilter` sólo se aplica a objetos de series temporales. Sin embargo, existe un método por defecto que debería funcionar para cualquier objeto numérico o vectorial.

library(dynlm)

La estructura y el funcionamiento interno de `dynlm` muestran notables similitudes con los de `lm`. No obstante, en la actualidad, `dynlm` ofrece tres ventajas distintivas en comparación con el uso directo de `lm`:

1. Procesamiento ampliado de fórmulas.
2. Conservación de atributos de series temporales.
3. Regresión de variables instrumentales (mediante mínimos cuadrados en dos etapas).

Para describir el modelo que se ajustará, `dynlm` presenta funciones adicionales que simplifican la tarea. Por ejemplo, se pueden especificar cómodamente la dinámica del modelo utilizando las funciones `d()` y `L()`, o los patrones lineales o cíclicos mediante `tren()`, `season()` y `harmon()`. Todas estas nuevas funciones de fórmula requieren que los argumentos sean objetos que representen series temporales.

library(nlme)

Esta función genérica realiza la estimación de un modelo no lineal con efectos mixtos, siguiendo la estructura propuesta en el trabajo de Lindstrom y Bates (1990), aunque con la capacidad adicional de incorporar efectos aleatorios anidados. Se permite la consideración de correlaciones entre errores dentro de los grupos, así como la posibilidad de que dichos errores tengan variabilidades diferentes.

library(prophet)

"Prophet" en R Studio fue específicamente diseñada para el pronóstico de series temporales, en especial para aquellas que presentan patrones estacionales y

tendencias. Fue desarrollada por Facebook y está diseñada para ser amigable con los usuarios y proporcionar un enfoque simplificado para la creación de pronósticos más precisos y automatizados.

1. Modelado de Estacionalidad y Tendencias: Es sumamente adecuada para modelar series temporales que tienen componentes estacionales y tendencias ya que captura patrones de estacionalidad diaria, semanal y anual, así como tendencias de crecimiento.
2. Manejo de Datos Faltantes: "Prophet" puede lidiar con datos faltantes de manera efectiva, lo que es común en las series temporales.
3. Capacidad de Personalización: Aunque la herramienta automatiza gran parte del proceso de pronóstico, también brinda diversas opciones para ajustar y personalizar modelos según las características específicas de los datos.
4. Gráficos de Visualización: La librería cuenta con herramientas de visualización de datos originales, las tendencias identificadas y los pronósticos generados, lo que ayuda a una mejor comprensión e interpretación de los resultados.

library(gridExtra)

Proporciona una serie de funciones a nivel de usuario para trabajar con gráficos de "cuadrícula", en particular para organizar múltiples gráficos de cuadrícula en una página y dibujar tablas.

Esta librería permite combinar múltiples gráficos en una única trama y controlar la disposición y el formato de estos gráficos de una forma flexible que con las funciones gráficas básicas de R.

library(ggpubr)

La librería "ggpubr" en R Studio es una herramienta que posibilita la creación y personalización de gráficos y tablas empleando el sistema de gráficos "ggplot2". Está diseñada para reducir tareas comunes para visualizar datos y la presentación de resultados estadísticos, permitiendo a los usuarios crear gráficos y tablas de manera más eficiente y estilizada.

library(ggstatsplot)

La librería ggstatsplot es una ampliación del paquete (ggplot2), diseñada para simplificar la creación de gráficos que incorporan detalles de pruebas estadísticas directamente en los gráficos mismos.

Ofrece una interfaz más accesible para generar representaciones visuales enriquecidas en información, específicas para el análisis estadístico de datos continuos (como gráficos de violín, dispersión, histogramas, puntos y bigotes) o categóricos (como gráficos de barras y circulares). Esta herramienta actualmente admite una variedad de pruebas estadísticas comunes, que incluyen pruebas paramétricas, no paramétricas, robustas y versiones bayesianas de pruebas t/ANOVA, análisis de correlación, análisis de tablas de contingencia, metaanálisis y análisis de regresión.

Línea de códigos

```
Impor <- read_excel ("C:/Users/rober/OneDrive/Escritorio/Carpeta Daniel/Base de datos - productos IMPORGUIDSA (2).xlsx")
```

En esta primera línea de código se utilizó la librería “read_excel” para importar la base de datos desde la dirección del escritorio donde se encuentra. Adicional, se le colocó a este primer objeto el nombre de “Impor”. La base de datos está compuesta por 2 columnas: las ventas en dólares y el periodo mensual desde 2018 a 2023.

Ilustración 1 Base de datos importada

	Fecha	Ventas
1	2018-01-01	32509.59
2	2018-02-01	5987.61
3	2018-03-01	22841.15
4	2018-04-01	172742.61
5	2018-05-01	12491.35
6	2018-06-01	3559.54
7	2018-07-01	131851.77
8	2018-08-01	52115.67
9	2018-09-01	695454.15
10	2018-10-01	787095.68
11	2018-11-01	579408.54
12	2018-12-01	335341.39
13	2019-01-01	100110.19
14	2019-02-01	5725.87
15	2019-03-01	8807.02
16	2019-04-01	12980.40
17	2019-05-01	24317.33
18	2019-06-01	2491.56
19	2019-07-01	19228.92
20	2019-08-01	72277.15

Showing 1 to 20 of 65 entries, 2 total columns

```
Impor1 <- Impor
```

Se crea un segundo objeto nombrado “Impor1” que contiene la base de datos del primer objeto “Impor”. En esta línea de código se crea una copia del objeto “Impor”

que se guarda en una nueva variable llamada “Impor1”. Quiere decir que tanto “Impor” como “Impor1” harán mención del mismo conjunto u objeto de datos.

```
Colnames(Impor1) <- c("ds","y")
```

En esta línea de códigos, la función “colnames” en R será implementada para hacer mención a los nombres de las columnas en objetos de datos tabulares como data frames o matrices.

Por lo tanto, “colnames” cambia los nombres de las columnas en el objeto “Impor1”. Los nombres originales de las columnas no se proporcionaron, pero se están reemplazando por "ds" y "y". Esto será útil debido a que más adelante se hará uso la librería "prophet" para el pronóstico de series temporales, ya que "ds" será para representar la columna de los periodos mensuales y la columna “y” que corresponde a las ventas en dólares que se pronosticarán.

Ilustración 2 Aplicación de "colnames"

	ds	y
1	2018-01-01	32509.59
2	2018-02-01	5987.61
3	2018-03-01	22841.15
4	2018-04-01	172742.61
5	2018-05-01	12491.35
6	2018-06-01	3559.54
7	2018-07-01	131851.77
8	2018-08-01	52115.67
9	2018-09-01	695454.15
10	2018-10-01	787095.68
11	2018-11-01	579408.54
12	2018-12-01	335341.39
13	2019-01-01	100110.19
14	2019-02-01	5725.87
15	2019-03-01	8807.02
16	2019-04-01	12980.40
17	2019-05-01	24317.33
18	2019-06-01	2491.56
19	2019-07-01	19228.92
20	2019-08-01	72277.15

Showing 1 to 20 of 65 entries, 2 total columns

Model1 <- prophet(Import1)

En esta línea se crea un modelo de pronóstico utilizando la función “prophet()” junto con el objeto “Import1” que contiene las dos columnas creadas anteriormente, una de fechas llamada "ds" y la de valores a pronosticar llamada "y". El resultado del modelo ajustado se almacenará en la variable “Model1”. El modelo se ajusta de manera exitosa a los datos de las ventas. Adicional, el paquete “prophet” nos deja un mensaje informativo que indica que la estacionalidad semanal y diaria fue desactiva para este modelo.

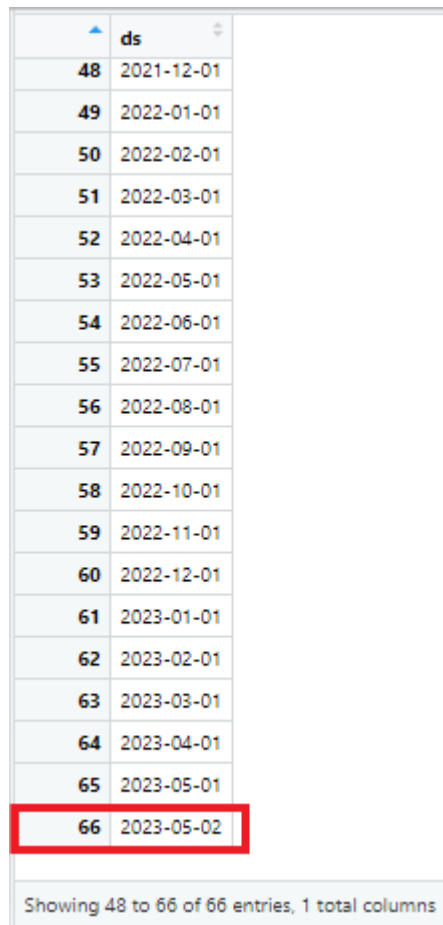
Future1 <- make_future_dataframe(Model1, periods = 1)

En esta línea procedemos a crear un nuevo data frame que nombraremos “Future1” que contiene las fechas futuras adicionales que requerimos para hacer

pronósticos. Luego, la función “make_future_dataframe()” se utiliza para generar estas fechas futuras basadas en la variable antes obtenida “Model1”. El argumento “periods” se decreta en 1, lo que nos indica que se están generando fechas para un solo período futuro adicional.

Ilustración 3 Base de datos con fechas futuras.

Base de datos con fechas futuras.



	ds
48	2021-12-01
49	2022-01-01
50	2022-02-01
51	2022-03-01
52	2022-04-01
53	2022-05-01
54	2022-06-01
55	2022-07-01
56	2022-08-01
57	2022-09-01
58	2022-10-01
59	2022-11-01
60	2022-12-01
61	2023-01-01
62	2023-02-01
63	2023-03-01
64	2023-04-01
65	2023-05-01
66	2023-05-02

Showing 48 to 66 of 66 entries, 1 total columns

Se pasa de tener 65 observaciones en “ds” (fechas) a 66 observaciones que corresponde al periodo futuro.

tail(Future1)

Con esta línea de códigos obtenemos las últimas filas del data frame “Future1”, donde podemos visualizar las fechas futuras generadas para el pronóstico.

```
> tail(Future1)
      ds
61 2023-01-01
62 2023-02-01
63 2023-03-01
64 2023-04-01
65 2023-05-01
66 2023-05-02
```

Forecast1 <- predict(Model1, Future1)

En esta línea se generan pronósticos utilizando “Model1” para las fechas futuras contenidas en “Future1”.

La función “predict()” toma el modelo ajustado y las fechas futuras, para luego devolver las predicciones y otro valores relacionados en el data frame “Forecast1”.

tail(Forecast1[c('ds','yhat','yhat_lower','yhat_upper')])

En esta línea de códigos, se muestran las últimas filas del data frame “Forecast1”, pero se establece mediante “c” únicamente para las columnas “ds”, “yhat”, “yhat_lower” y “yhat_upper”, que representan la fecha, el valor del pronóstico, intervalo de confianza inferior e intervalo de confianza superior del pronóstico, en este orden respectivamente.

```
> tail(Forecast1[c('ds','yhat','yhat_lower','yhat_upper')])
      ds      yhat yhat_lower yhat_upper
61 2023-01-01 72792.98 -130131.16 255775.1
62 2023-02-01 18809.54 -174144.97 204112.9
63 2023-03-01 11746.24 -196217.53 206818.6
64 2023-04-01 106732.97 -82901.66 302192.4
65 2023-05-01 155460.87 -34140.45 352730.1
66 2023-05-02 -106169.46 -291936.20 78328.6
```

En cada fila del resultado, en la columna “yhat”, encontraremos el valor pronosticado para la variable de interés. En este caso, que se pronostican ventas en la fila con “ds” igual a "2023-01-01", el valor pronosticado de ventas en dólares es aproximadamente 72,792.98.

Por consiguiente, en cada fila del resultado, en la columna “yhat_lower”, encontraremos el límite inferior del intervalo de confianza para el valor pronosticado. Esto significa que hay una alta probabilidad de que el valor real esté por encima de

este límite inferior. Siguiendo el mismo ejemplo, en la fila con “ds” igual a "2023-01-01", el límite inferior del intervalo de confianza es aproximadamente -107,810.33.

En la columna “yhat_upper”, se encontrará el límite superior del intervalo de confianza para el valor pronosticado. Similar al límite inferior, esto significa que hay una alta probabilidad de que el valor real esté por debajo de este límite superior. Siguiendo el mismo ejemplo, en la fila con `ds` igual a "2023-01-01", el límite superior del intervalo de confianza es aproximadamente 263,903.68.

Plotting the Model Estimates

Pronostico <- dyplot.prophet(Model1, Forecast1)

Procedemos a realizar un gráfico interactivo del pronóstico obtenido, empleando el modelo “Model1” y los resultados de pronóstico contenidos en “Forecast1”. Este resultado se guarda en la variable “Pronostico”.

Pronostico

En el eje horizontal del gráfico, se reflejarán las fechas o periodos para los cuales se realizó el pronóstico. Se está pronosticando para un período que abarca hasta mayo de 2023. En el eje vertical del gráfico, están los valores de la variable que estás pronosticando. En este caso es la variable de ventas en dólares.

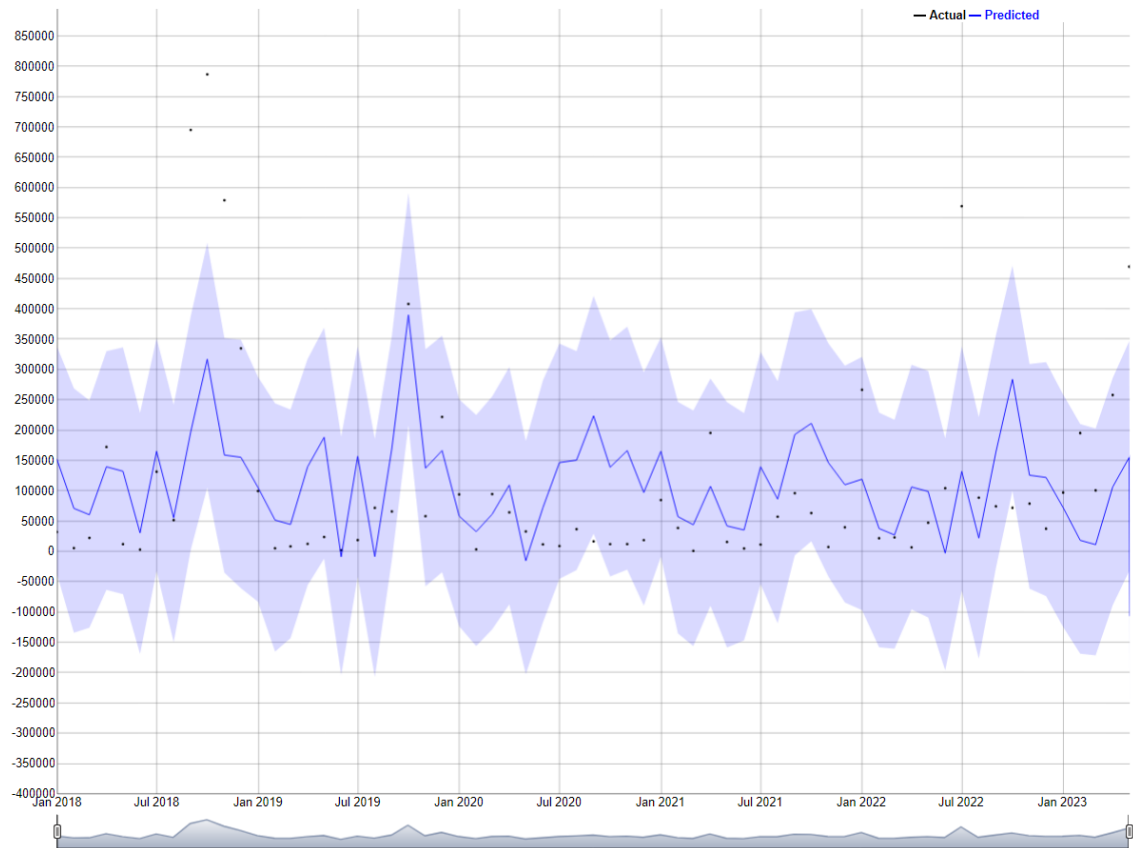
Continuando, la línea que visualizada en el gráfico representa los valores pronosticados para la variable a lo largo del tiempo. Esta línea indicará cómo se espera que la variable cambie en función de las fechas pronosticadas.

Las áreas sombreadas que se pueden observar alrededor de la línea de pronóstico. Estas áreas sombreadas representarán los intervalos de confianza, donde el tono más claro podría indicar el rango de incertidumbre. En este caso notamos que el área sombreada está muy dispersa de lo pronosticado, se podría tratar de valores atípicos o ruido en los datos puede afectar negativamente la precisión del pronóstico, ya que el modelo puede estar tratando de ajustarse a estas variaciones inusuales.

También es posible ver puntos de datos reales observados en el pasado. Estos puntos te permiten comparar las predicciones con los valores reales para evaluar la

precisión del pronóstico que de igual manera que los intervalos se encuentran con una variación considerable respecto a lo pronosticado.

Ilustración 4 Gráfico de valores pronosticados

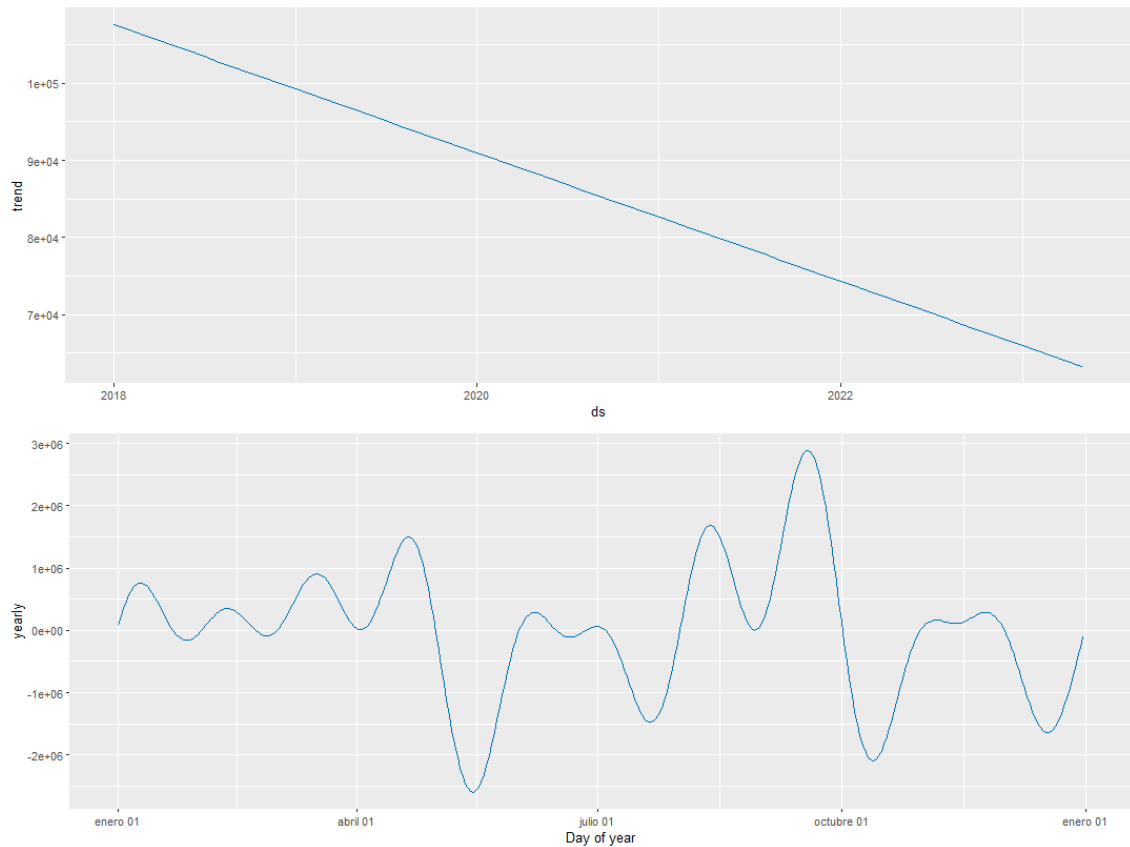


prophet_plot_components(Model1, Forecast1)

Por último, se procede a crear un gráfico que enseña los componentes del modelo de pronóstico, como las tendencias y las estacionalidades.

El gráfico nos muestra la tendencia general de la serie de tiempo a lo largo del período de tiempo cubierto por los datos. En este caso se puede observar que hay una tendencia a la baja. Adicional, también muestra la estacionalidad anual, es decir, los patrones que se repiten cada año en la serie, esta estacionalidad nos confirma que existe mucha variación en los datos con picos muy altos o bajos.

Ilustración 5 Tendencias y estacionalidades



PREDICCIÓN FUTURA

Impor2 <- Impor[,-c(1)]

En esta línea procedemos a crear un nuevo objeto llamado “Impor2”; eliminando la primera columna del objeto “Impor” que corresponde al año mediante [,-c(1)].

Series <- ts(Impor2, start = c(2018,1), end = c(2023,5), frequency = 12)

Luego, procedemos a crear un objeto de series de tiempo “Series”; utilizando la función “ts()” organizamos los datos de “Impor2” en una serie de tiempo que según la función empieza en enero de 2018 y culmina en mayo de 2023, manteniendo una frecuencia mensual (12 observaciones al año). Compruebo visualmente inicios o tendencias de estacionalización.

Plot1 <- boxplot(Series~cycle(Series))

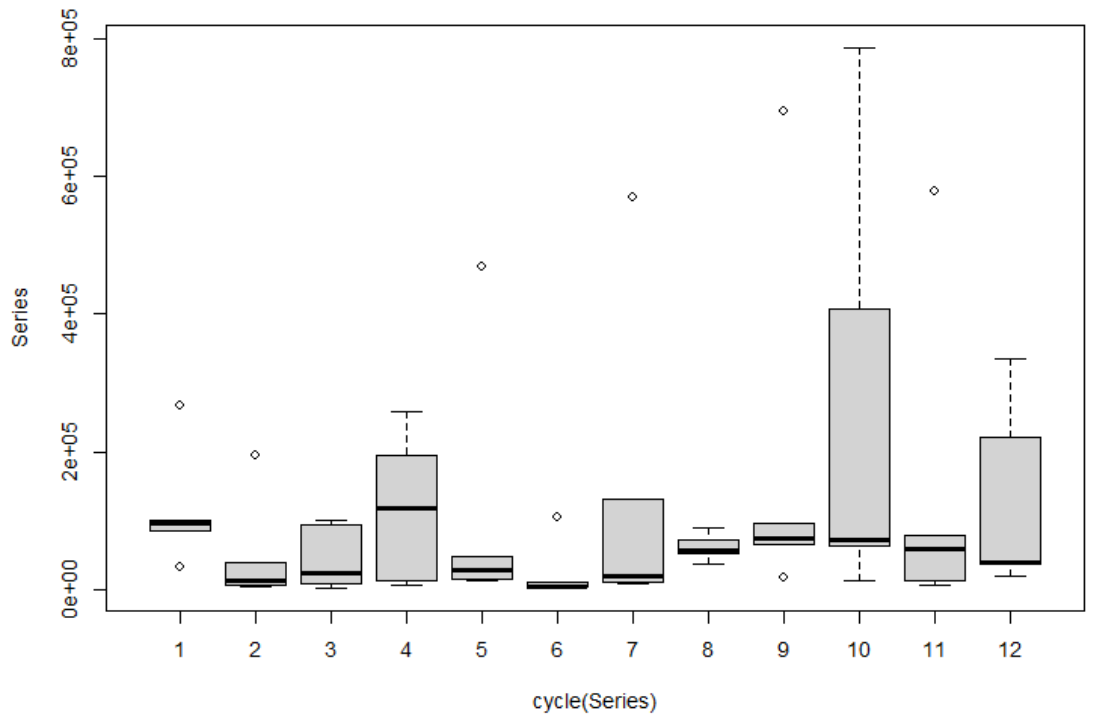
En esta línea de códigos se crea un gráfico de cajas utilizando la función “boxplot()” para enseñar las distribuciones de los datos en “Series” en función del ciclo de la serie de tiempo, que también ayudará a detectar patrones estacionales.

Plot1

Se ejecuta el gráfico de cajas formulado. Aquí estamos graficando “Series” en función del ciclo (una representación de la estacionalidad) de la misma serie. Esto te permitirá visualizar cómo varían los valores de la serie en diferentes momentos del ciclo estacional.

El gráfico resultante muestra la distribución de los valores de la serie de tiempo en cada período del ciclo estacional. Cada caja (box) representa un período del ciclo y nos da una idea de la variabilidad de los datos en ese período. Podemos volver a notar la presencia de datos atípicos en las cajas, lo cual puede ser una señal de la no existencia de ruido blanco.

Ilustración 6 Boxplot para determinar patrones estacionales



```

$stats
      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]      [,7]      [,8]      [,9]      [,10]      [,11]      [,12]
[1,] 85196.89 3906.35 1377.47 7174.02 12491.35 2491.56 9440.60 37170.76 66339.00 12573.15 7832.44 19098.23
[2,] 85196.89 5725.87 8807.02 12980.40 16159.35 3559.54 11844.15 52115.67 66339.00 63738.40 12605.35 37918.75
[3,] 96059.98 14117.88 23244.03 118798.82 28842.05 5379.90 19228.92 57700.53 74853.10 72379.19 58585.00 40278.85
[4,] 100110.19 39100.60 95053.54 195865.90 47727.10 12155.42 131851.77 72277.15 96564.78 408450.44 79347.15 222213.85
[5,] 100110.19 39100.60 101161.35 258392.60 47727.10 12155.42 131851.77 89122.90 96564.78 787095.68 79347.15 335341.39

$sn
[1] 6 6 6 6 6 5 5 5 5 5 5 5

$conf
      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]      [,7]      [,8]      [,9]      [,10]      [,11]      [,12]
[1,] 86440.41 -7409.899 -32387.77 831.7655 8479.827 -693.9272 -65568.16 43454.48 53495.64 -171193.5 11425.41 -89943.6
[2,] 105679.54 35645.659 78875.82 236765.8745 49204.263 11453.7272 104026.00 71946.58 96210.56 315951.8 105744.59 170501.3

$out
[1] 32509.59 266969.90 195714.10 469968.60 104792.00 569749.44 695454.15 17043.10 579408.54

$group
[1] 1 1 2 5 6 7 9 9 11

$names
[1] "1" "2" "3" "4" "5" "6" "7" "8" "9" "10" "11" "12"
    
```

Continuando con el análisis, **stats**: es una matriz que contiene los valores de los estadísticos resumen para cada período del ciclo estacional. Cada columna representa un período y los valores en la columna son los estadísticos resumen para ese período. Los estadísticos incluidos son el valor mínimo, el primer cuartil (Q1), la mediana (Q2), el tercer cuartil (Q3) y el valor máximo. Por ejemplo, en el primer período, los estadísticos son:

- Mínimo: 85196.89
- Q1: 3906.35

- Mediana: 1377.47
- Q3: 7174.02
- Máximo: 12491.35

n: Esta es una lista que muestra la cantidad de observaciones en cada período del ciclo estacional. Hay 6 observaciones para cada uno de los 12 períodos.

conf: Esta matriz contiene los intervalos de confianza para los estadísticos resumen en cada período. Similar a la matriz de “stats”, cada columna corresponde a un período y contiene el intervalo de confianza para el valor mínimo, Q1, mediana, Q3 y valor máximo. Por ejemplo, en el primer período, el intervalo de confianza (80%) para la mediana es [3906.35, 5725.87].

out: Este vector contiene los valores atípicos (outliers) para cada período del ciclo estacional. Los valores atípicos son observaciones que están muy alejadas de los valores típicos en ese período. En este caso, como ya lo mencionamos, hay valores atípicos en varios períodos.

group: Este vector indica a qué grupo pertenece cada período del ciclo estacional. Los grupos son determinados por la función “boxplot” y se utilizan para identificar valores atípicos en relación con los otros períodos en el mismo grupo.

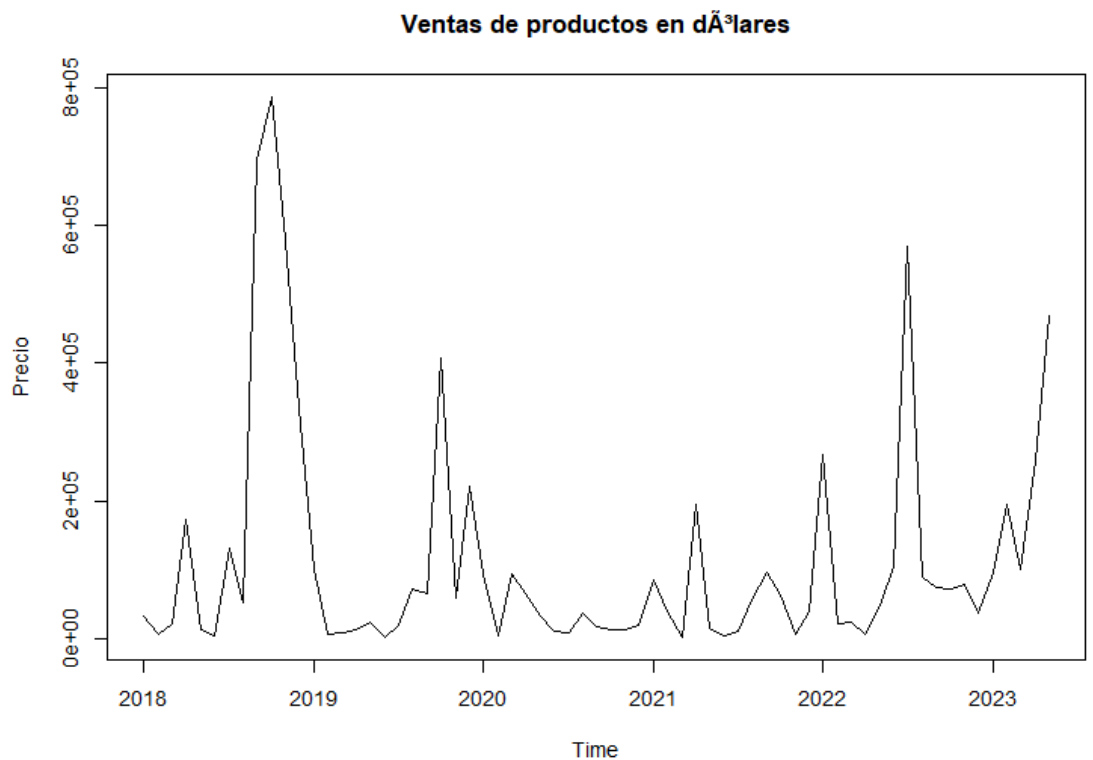
names: Este vector contiene los nombres de cada período del ciclo estacional. En tu caso, los nombres son simplemente los números del 1 al 12.

En resumen, esta información permite comprender la distribución y variabilidad de los valores en cada período del ciclo estacional. Los estadísticos resumen, los intervalos de confianza y los valores atípicos dan una idea de cómo varían los datos en diferentes momentos del ciclo.

```
plot2 <- plot(Series, ylab="Precio", main="Ventas de productos en dólares")
```

Se crea un gráfico básico utilizando la función “plot()”. Este gráfico muestra la serie de tiempo “Series” en el eje y, con etiquetas de eje y y las ventas de los productos en dólares. Se pueden observar patrones de picos altos y picos bajos a lo largo del año, estos nos pueden indicar estacionalidad en los datos de ventas. Este tipo de patrón sugiere que hay una tendencia recurrente en nuestros datos, posiblemente debido a factores estacionales o cíclicos que afectan las ventas.

Ilustración 7 Gráfico de ventas en dólares



Prueba de modelo TM

```
modelo <- auto.arima(Series, seasonal=T, stepwise=F, approximation=T)
```

En esta línea estás ajustando un modelo ARIMA automáticamente a tus datos de ventas “Series”. “auto.arima” es una función en R que busca automáticamente el mejor modelo ARIMA basado en los criterios de información de Akaike (AIC) y Bayesian (BIC). El argumento “seasonal=T” indica que se consideren componentes estacionales en el modelo. El argumento “stepwise=F” desactiva la búsqueda stepwise (paso a paso) en la selección del modelo. El argumento “approximation=T” indica que

se utilice una aproximación en la selección del modelo para mejorar la velocidad de cómputo.

summary(modelo)

Se muestra un resumen del modelo ARIMA ajustado utilizando la función “summary()”. Esto proporciona información sobre los coeficientes del modelo y estadísticas de ajuste.

```
> summary(modelo)
Series: Series
ARIMA(0,0,0)(1,0,0)[12] with non-zero mean

Coefficients:
      sar1      mean
      0.3213 127595.00
s.e.  0.1856  29607.11

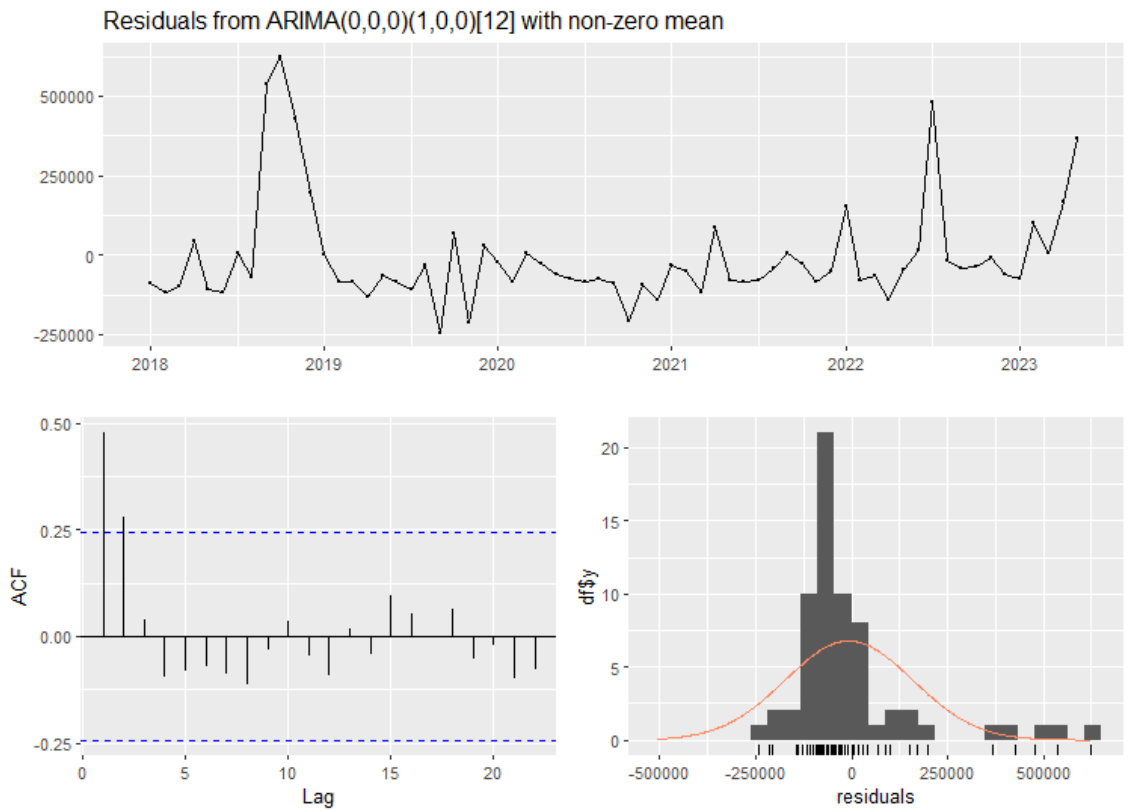
sigma^2 = 2.808e+10:  log likelihood = -873.76
AIC=1753.52  AICc=1753.92  BIC=1760.05

Training set error measures:
              ME  RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set -7489.364 164967 109869.2 -585.2053 608.7512 1.01568 0.4765109
```

Residual <- checkresiduals(modelo)

Esta línea está calculando y almacenando los residuales del modelo ajustado. Los residuales son las diferencias entre los valores observados y los valores pronosticados por el modelo. El resultado `Residual` contendrá información sobre la normalidad y la autocorrelación de los residuales, lo que te ayudará a evaluar la calidad del ajuste del modelo.

Ilustración 8 Medidas de rendimiento del modelo



El modelo ARIMA ajustado es un ARIMA(0,0,0)(1,0,0)[12], lo que significa que no se usaron términos de autoregresión (AR) ni términos de media móvil (MA), pero se usó un término de estacionalidad de orden 1 (sar1) con un período de 12 (indicando una estacionalidad mensual). El valor estimado para sar1 es 0.3213 y el valor estimado para el término de media (mean) es 127595.00.

La varianza residual (σ^2) es aproximadamente $2.808e+10$, lo que refleja la variabilidad no explicada por el modelo. Este valor puede ser útil para evaluar la bondad del ajuste del modelo.

El log likelihood es -873.76. Los criterios de información AIC (1753.52) y AIC corregido (AICc, 1753.92) son medidas de la calidad del ajuste del modelo, donde valores más bajos son mejores. El BIC (1760.05) también es un criterio de información, similar al AIC pero penaliza más los modelos más complejos.

Estas medidas evalúan el rendimiento del modelo en el conjunto de datos de entrenamiento. El error medio (ME) es -7489.364, lo que indica un sesgo negativo. El

error cuadrático medio (RMSE) es 164967, que mide la dispersión de los errores. El error absoluto medio (MAE) es 109869.2. El error porcentual medio (MPE) es -585.2053, lo que indica un sesgo negativo. El error porcentual absoluto medio (MAPE) es 608.7512. El error cuadrático medio absoluto (MASE) es 1.01568, que compara el modelo con un modelo ingenuo. El valor ACF1 es 0.4765109, que es el primer coeficiente de autocorrelación de los residuales.

La prueba Ljung-Box evalúa si los residuales son independientes y sin autocorrelación. En este caso, el valor Q^* es 25.075, y el p-valor asociado es 0.01447. Un p-valor bajo indica que existe autocorrelación en los residuales. Por lo tanto, existe correlación en los residuales. La prueba de Ljung-Box es comúnmente utilizada para verificar si los residuales se asemejan al ruido blanco, es decir, si no hay autocorrelación significativa en los residuales.

En nuestros resultados, se ha obtenido un p-valor de 0.01447 en la prueba de Ljung-Box. Un p-valor menor a un umbral de significancia (como 0.05) sugiere que hay evidencia de autocorrelación en los residuales y que no se pueden considerar como ruido blanco.

Forecast <- forecast(modelo, h=24)

Procedemos a realizar el pronósticos utilizando el modelo ARIMA ajustado para 24 períodos futuros utilizando la función “forecast()”.

Forecast

Ejecutamos el forecast.

> Forecast

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jun 2023	120267.84	-94475.41	335011.1	-208153.61	448689.3
Jul 2023	269670.02	54926.78	484413.3	-58751.43	598091.5
Aug 2023	115232.97	-99510.27	329976.2	-213188.48	443654.4
Sep 2023	110647.74	-104095.51	325391.0	-217773.71	439069.2
Oct 2023	109852.81	-104890.43	324596.1	-218568.64	438274.3
Nov 2023	112091.79	-102651.46	326835.0	-216329.66	440513.2
Dec 2023	98779.83	-115963.41	313523.1	-229641.62	427201.3
Jan 2024	117948.39	-96794.85	332691.6	-210473.06	446369.8
Feb 2024	149483.33	-65259.91	364226.6	-178938.12	477904.8
Mar 2024	119101.22	-95642.02	333844.5	-209320.23	447522.7
Apr 2024	169623.46	-45119.78	384366.7	-158797.99	498044.9
May 2024	237608.00	22864.76	452351.2	-90813.45	566029.5
Jun 2024	125240.60	-100316.42	350797.6	-219719.10	470200.3
Jul 2024	173247.18	-52309.85	398804.2	-171712.53	518206.9
Aug 2024	123622.78	-101934.25	349179.8	-221336.92	468582.5
Sep 2024	122149.43	-103407.59	347706.5	-222810.27	467109.1
Oct 2024	121894.00	-103663.02	347451.0	-223065.70	466853.7
Nov 2024	122613.44	-102943.59	348170.5	-222346.26	467573.1
Dec 2024	118335.98	-107221.04	343893.0	-226623.72	463295.7
Jan 2025	124495.31	-101061.72	350052.3	-220464.39	469455.0
Feb 2025	134628.26	-90928.77	360185.3	-210331.45	479588.0
Mar 2025	124865.74	-100691.28	350422.8	-220093.96	469825.4
Apr 2025	141099.77	-84457.25	366656.8	-203859.93	486059.5
May 2025	162944.87	-62612.16	388501.9	-182014.84	507904.6

Los intervalos indican el rango dentro del cual es probable que caiga el valor real en el futuro. Por ejemplo, para junio de 2023, el intervalo de confianza al 80% va desde -94475.41 hasta 335011.1, y el intervalo al 95% va desde -208153.61 hasta 448689.3. Esto significa que, con un nivel de confianza del 80%, el valor real de junio de 2023 estará dentro del rango -94475.41 a 335011.1.

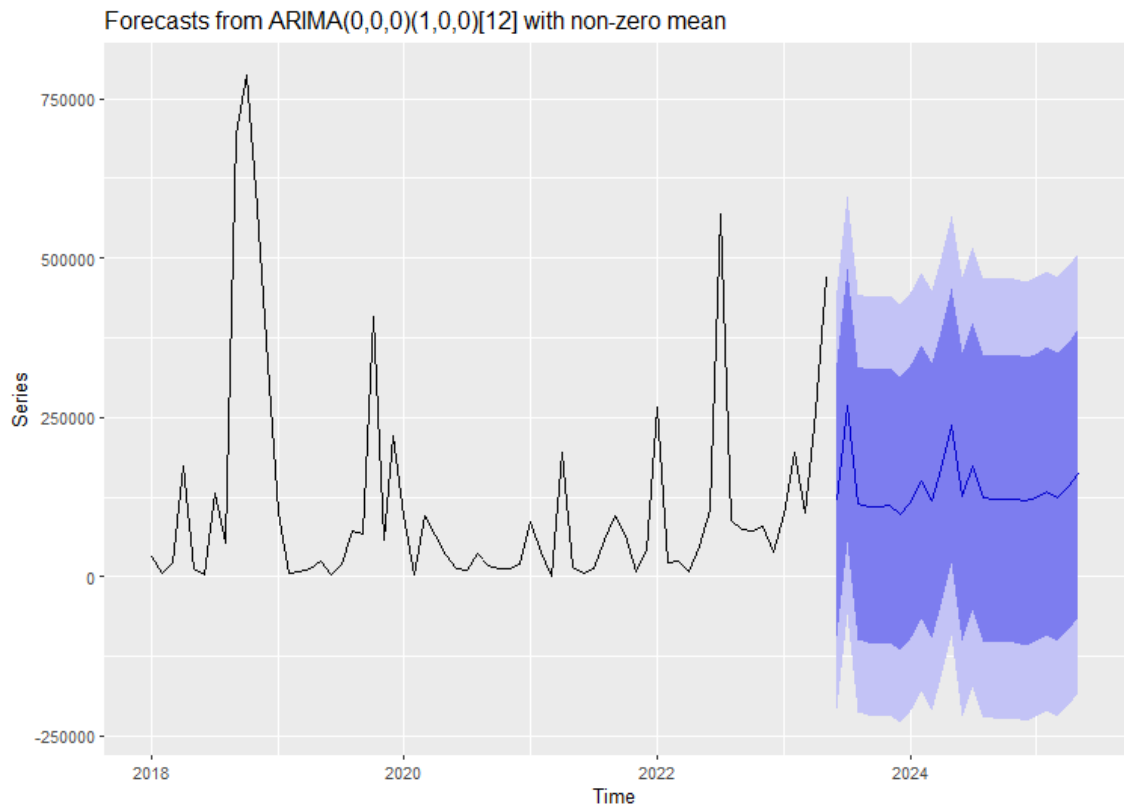
ForecastPlot <- autoplot(Forecast)

Creamos el gráfico de los pronósticos mediante la función “autoplot()” del paquete “forecast” para que de esta manera se genere un gráfico donde se visualizarán los pronósticos con intervalos de confianza.

ForecastPlot

Obtenemos nuestro gráfico visual de pronósticos con los intervalos de confianza respectivos.

Ilustración 9 Gráfico de pronóstico mediante modelo ARIMA.



El gráfico que hemos generado con “autoplot(Forecast)” muestra las proyecciones futuras junto con los intervalos de confianza. En este gráfico, la línea azul representa el pronóstico puntual, y las áreas sombreadas representan los intervalos de confianza del 80% (más oscuro) y del 95% (más claro). Esto te brinda una idea visual de cómo se espera que se desarrollen las ventas en el futuro y qué tan incierto es ese pronóstico. En nuestro caso, se utilizarán los intervalos de confianza del 80% debido a que son los que más se acercan al pronóstico y a cómo la IA ha buscado arreglar los datos que pueden presentar problemas.

Conclusiones del Análisis de resultados

En este trabajo de investigación se tuvo como objetivo realizar un análisis y pronósticos de ventas mediante el modelo ARIMA buscando principalmente la optimización en los tiempos de importación de la importadora (IMPORGUIDSA) pero también canalizando esta información para mejores tomas de decisiones en diferentes aspectos y departamentos de la empresa, como podría ser optimización de procesos y operatividad en compras, control de stock en bodegas, tiempos de importación ya mencionados en la logística y reducción/control de costos en finanzas.

Para empezar con este proceso de análisis e implementación de ARIMA procedimos a extraer una base de datos de las ventas había tenido la importadora durante el periodo de enero del 2018 a mayo del 2023 referente a los productos que se distribuyen a nivel nacional, estos datos fueron puestos a prueba en diferentes procesos con el objetivo de encontrar y saber si existían patrones estacionales, tendencias, picos o caídas notables en las ventas a lo largo del tiempo se utilizó la librería "prophet" y el modelo ARIMA que ambos son útiles para analizar y procesar este tipo de comportamientos.

Luego, de realizar la ejecución de los modelos antes mencionados ARIMA $(0,0,0)(1,0,0)[12]$, se pudo observar que los datos presentaban posible autocorrelación, así como también mucha variación en su pronóstico lo cual infería a que no presentaba una estacionalidad determinada como se muestran en las diferentes gráficas presentadas en el análisis de resultado.

Adicional, estos resultados indican que el modelo ARIMA ajustado podría tener problemas de autocorrelación en los residuales, lo que sugiere que podría haber patrones no capturados por el modelo. También es importante considerar la interpretación del término estacional y el término de media en el contexto nuestros datos y del dominio del problema. La evaluación de si los residuales del modelo son ruido blanco es un paso importante para determinar la calidad del modelo. En nuestro caso, se ha realizado la prueba de Ljung-Box para evaluar la autocorrelación de los residuales del modelo ARIMA. La prueba de Ljung-Box es comúnmente utilizada para verificar si los residuales se asemejan al ruido blanco, es decir, si no hay autocorrelación significativa en los residuales. En nuestro resultado, hemos obtenido

un p-valor de 0.01447 en la prueba de Ljung-Box. Un p-valor menor a un umbral de significancia (como 0.05) sugiere que hay evidencia de autocorrelación en los residuales y que no se pueden considerar como ruido blanco ilustración (8). En este caso, el p-valor bajo indica que los residuales del modelo ARIMA pueden no cumplir con la suposición de ser ruido blanco. Esto puede ser una señal de que el modelo ARIMA puede no ser el más apropiado para tus datos. Podrías explorar opciones alternativas, como probar diferentes configuraciones de parámetros para el modelo ARIMA o considerar otros métodos de pronóstico.

Sin embargo y a pesar de que el modelo genere incertidumbre mediante sus intervalos de confianza y ejecución debido a posibles datos que afectan a sus residuales, no deja de ser un modelo bien estructurado y confiable para la toma de decisiones de la empresa para ventas futuras y planificación de sus importaciones teniendo en cuenta que los picos más altos de ventas se consideran a partir del segundo semestre de todos los años, logrando de esta manera optimizar tiempos y generar un valor adicional para los procesos de la importadora.

Conclusiones

Como modelo inicial es evidente que se ha realizado un proceso exhaustivo y valioso para construir un modelo ARIMA, por lo tanto, es importante reconocer que los datos pueden presentar patrones más complejos que no son capturados adecuadamente por este modelo. La autocorrelación en los residuales sugiere que podría haber información que el modelo no está capturando.

Esto quiere decir, que las ventas pueden estar influenciadas por múltiples factores, como estacionalidad, tendencias, efectos promocionales y otros. Cabe indicar, que la construcción de modelos de pronóstico es un proceso iterativo. A medida que obtengamos más datos y recopilemos más información sobre cómo se desempeñan los pronósticos obtenidos en la realidad, podrás mejorar tus modelos y ajustar tus métodos en consecuencia.

A pesar de las limitaciones del modelo actual, siguen siendo relevantes y de mucha utilidad los pronósticos como guía para la planificación de importaciones en cuanto a tiempo y volúmenes, lo que ayudaría a reducir los costos relacionados con el transporte internacional y mejorar la eficiencia de la cadena de suministro. Sin

embargo, es importante tener en cuenta la incertidumbre asociada con los pronósticos y considerar márgenes de seguridad en las decisiones que se pueden tomar en base a la investigación.

No obstante, existen varios campos donde también podemos utilizar este pronóstico que ayudará de manera óptima y relevante en los procesos y toma de decisiones relacionadas con las ventas y cadena de suministro de la empresa.

Planificación de Inventarios: Con pronósticos más precisos de las ventas, se puede lograr planificar mejores niveles de inventario. Esto puede evitar que haya un exceso de inventario que signifiquen costos adicionales de almacenamiento e impedir en su totalidad las situaciones de agotamiento de stock que al final resultan pérdida de ventas.

Gestión de Recursos: Al tener una visibilidad clara de las tendencias y patrones de ventas, la importadora puede asignar sus recursos de manera más eficiente. Esto incluye la asignación de personal en función de la carga de trabajo prevista y la reducción de operatividad.

Promociones y Estrategias de Ventas: El análisis de las ventas históricas y los pronósticos mediante el modelo ARIMA, puede ayudar al diseño de estrategias de marketing más efectivas e identificar oportunidades para aumentar las ventas en momentos y temporadas clave.

Mejora de la Toma de Decisiones: El modelo ARIMA proporciona una base objetiva para la toma de decisiones. Las decisiones sobre inventario, importaciones y otros aspectos de la gestión se pueden complementar con datos y análisis sólidos, lo que disminuye la desconfianza y el riesgo asociados al sector.

Reducción de Costos: Al eliminar el exceso de inventario, la empresa puede reducir costos en términos de almacenamiento y obsolescencias de productos.

Adaptación a Cambios: El modelo proporcionará información para identificar tendencias emergentes y cambios en la demanda. Esto es muy importante en entornos empresariales cambiantes donde la adaptación rápida es clave.

Mejora en la Experiencia del Cliente: Al tener disponible el producto demandado en el momento adecuado, la empresa puede brindar un mejor servicio

como distribuidor a todos sus clientes a nivel nacional, lo que a su vez puede mejorar la satisfacción y la fidelidad del cliente.

Recomendaciones

En general, nuestro enfoque es sólido y bien estructurado. Se han priorizado la implementación de herramientas estadísticas y de visualización acordes para analizar los datos y generar los debidos pronósticos de ventas futuras. Sin embargo, para una evaluación más exhaustiva y de mayor precisión, tendremos que considerar en realizar análisis adicionales, como la comparación de los pronósticos de ventas con los valores reales una vez que se encuentren disponibles para con esto empezar a contar con datos más concretos y nuevas tendencias. También se podría tomar en consideración métricas de evaluación para el rendimiento del pronóstico, como el error cuadrado medio (MSE) o el error absoluto medio (MAE), con el objetivo se evaluar la precisión de nuestro modelo.

Basándonos en el análisis y resultados que se han obtenido, existen algunas recomendaciones que podrían ayudar con el objetivo de mejorar el modelo de pronóstico de ventas y la optimización para la gestión de los tiempos de importación. Explorar otros modelos podría ser una opción debido a que los residuales del modelo ARIMA parecen tener autocorrelación. Sin embargo, se puede realizar una optimización de parámetros, manteniendo nuestro modelo Arima, podemos experimentar con diferentes configuraciones de los parámetros (como el orden del componente AR, MA y la frecuencia de estacionalidad) para intentar mejorar la calidad de los pronósticos y reducir la autocorrelación de los residuales.

Por último, podemos incorporar variables externas que puedan influir en las ventas (como campañas de marketing, eventos especiales, tendencias económicas, etc.), podemos incorporarlas al modelo para tener una mejor captura de los patrones y las relaciones causales, manteniéndolos en conjuntos de entrenamiento y prueba para verificar cómo se comporta el modelo en datos no vistos.

Anexo 1. Base de datos de ventas IMPORGUIDSA

Fecha	Ventas
1-Jan-18	32509,59
1-Feb-18	5987,61
1-Mar-18	22841,15
1-Apr-18	172742,61
1-May-18	12491,35
1-Jun-18	3559,54
1-Jul-18	131851,77
1-Aug-18	52115,67
1-Sep-18	695454,15
1-Oct-18	787095,68
1-Nov-18	579408,5407
1-Dec-18	335341,39
1-Jan-19	100110,19
1-Feb-19	5725,87
1-Mar-19	8807,02
1-Apr-19	12980,4
1-May-19	24317,33
1-Jun-19	2491,56
1-Jul-19	19228,92
1-Aug-19	72277,15
1-Sep-19	66339
1-Oct-19	408450,44
1-Nov-19	58585
1-Dec-19	222213,85
1-Jan-20	94546,35
1-Feb-20	3906,35
1-Mar-20	95053,54

1-Apr-20	64855,03
1-May-20	33366,76
1-Jun-20	12155,42
1-Jul-20	9440,6
1-Aug-20	37170,76
1-Sep-20	17043,1
1-Oct-20	12573,15
1-Nov-20	12605,35
1-Dec-20	19098,23
1-Jan-21	85196,89
1-Feb-21	39100,6
1-Mar-21	1377,47
1-Apr-21	195865,9
1-May-21	16159,35
1-Jun-21	5379,9
1-Jul-21	11844,15
1-Aug-21	57700,53
1-Sep-21	96564,78
1-Oct-21	63738,4
1-Nov-21	7832,44
1-Dec-21	40278,85

Anexo 2. Script utilizado en R Studio.

```
library(readxl)
library(xts)
library(dplyr)
library(tidyverse)
library(lubridate)
library(tseries)
library(astsa)
library(forecast)
library(foreign)
library(timsac)
library(vars)
library(mFilter)
library(dynlm)
library(nlme)
library(prophet)
library(gridExtra)
library(ggpubr)
library(ggstatsplot)
```

```
Impor <- read_excel("C:/Users/rober/OneDrive/Escritorio/Carpeta Daniel/Base de
datos - productos IMPORGUIDSA (2).xlsx")
```

```
Impor1 <- Impor
colnames(Impor1) <- c("ds","y")
```

```
Model1 <- prophet(Impor1)
Future1 <- make_future_dataframe(Model1, periods = 1)
tail(Future1)
Forecast1 <- predict(Model1, Future1)
```

```

tail(Forecast1[c('ds','yhat','yhat_lower','yhat_upper')])
#Plotting the Model Estimates
Pronostico <- dyplot.prophet(Model1, Forecast1)
Pronostico
prophet_plot_components(Model1, Forecast1)

#FORECAST COMPLETO

#Vamos a quitarle el aÃ±o
Impor2 <- Impor[,-c(1)]

#Vamos a darle formato de series de tiempo (1)
Series <- ts(Impor2, start = c(2018,1), end = c(2023,5), frequency = 12)

#Compruebo visualmente inicios o tendencias de estacionalizaciÃ³n (3)
Plot1 <- boxplot(Series~cycle(Series))
Plot1
plot2 <- plot(Series, ylab="Precio", main="Ventas de productos en dÃ³lares")

#PRUEBA DE MODELO TM
modelo <- auto.arima(Series, seasonal=T, stepwise=F, approximation=T)
summary(modelo)
Residual <- checkresiduals(modelo)
Forecast <- forecast(modelo, h=24)
Forecast
ForecastPlot <- autoplot(Forecast)
ForecastPlot

```

Bibliografía

- Ade, P. A. (2019). The Simons Observatory: Science Goals and Forecasts. *J. Cosmol. Astropart. Phys.*
- Albarak, M., & Bahsoon, R. (2018). Prioritizing Technical Debt In Database Normalization Using Portfolio Theory and Data Quality Metrics. *In Proceedings of the 2018 International Conference on Technical Deb*, 31-40.
- Angwin, A., Wilson, W., & Arnott, W. (n.d.). White Noise Enhances New-word Learning In Healthy Adults. *Sci Rep*, 1(7).
- Aroca et al., P. R. (2009). Estadística descriptiva e inferencial. *Revista el auge de la estadística en el siglo XX.*, 22, 165-176. Retrieved from https://www.researchgate.net/profile/Pedro-Romero-Aroca/publication/275021043_Estadistica_Descriptiva_e_Inferencial/links/55bfd42b08aec0e5f4476a2a/Estadistica-Descriptiva-e-Inferencial.pdf
- Beard, E. M. (2019). 20ding and Using Time Series Analyses in Addiction Research. *Addiction*, 10(114), 1866-1884.
- Corona Armenta et al. (2016). Medición de la innovación en las pymes mediante la aplicación de métodos multicriterio (ELECTRE Y AHP). *Mercados y Negocios*, 14(2), 169-185. Retrieved from <https://doi.org/10.32870/myn.v0i28.5252>
- Crespo de Carvalho et al. (2014). Strategic Fast Supply Demand-Chains in a Network Context: Opportunistic Practices That Can Destroy Supply Chain Systems. *American Journal of Industrial and Business Management*, 4(3), 123-133. Retrieved from 10.4236/ajibm.2014.43019.
- Cueva et al. (2014). Determinación de la línea base para la aplicación de sistemas de BI (Business Intelligence) en el Ecuador, dentro de un programa para el uso de tecnología de avanzada en la administración pública y privada. *ESPE*, 1-9.
- Delgado Olaya, & Yánez Sarmiento . (2019). Impacto de las importaciones no petroleras desde China, en la balanza comercial del Ecuador, período 2013-2017. *INNOVA Research Journal*, 4(3), 205-216. Retrieved from <https://doi.org/10.33890/innova.v4.n3.2.2019.1100>

- Fryz, M., & Scherbak, L. (2023). Properties Of Cyclostationary Random Processes In the Problems Of Energy Informatics. *Sist. dosl. Energ.*
- Ghosh, S. (2017). Forecasting Cotton Exports in India Using the ARIMA model. *Amity Journal of Economics*, 36-52.
- Giraldo-Ríos, L. (2022, marzo 9). *La importancia de la inteligencia de negocios*. Retrieved from Universidad Espiritu Santo: <https://uees.edu.ec/la-importancia-de-la-inteligencia-de-negocios/#:~:text=Un%20BI%20permite%20a%20las,una%20multitud%20de%20otras%20C3%A1reas>.
- Gujatari, D. N. (2010). *Econometría. México: The McGraw-Hill*.
- Guzmán, J. (2019). Desarrollo de un modelo de pronóstico en la metodología de Box-Jenkins para la demanda de estudiantes de pregrado de la fundación universitaria tecnológico COMFENALCO de la ciudad de Cartagena Indias.
- Hallikainen, P., & Tim, Y. (2021). Creative analytics: Towards data-inspired creative decisions. *Wiley*, 32(4), 729-753. Retrieved from <https://doi.org/10.1111/isj.12369>
- Kurov, A., Sancetta, A., Strasser, G., & Wolfe, M. (2018). Price Drift Before U.S. Macroeconomic News. *1(54)*, 449-479.
- Lahura, E. (2003). El coeficiente de correlación y correlaciones espúreas. *Departamento Académico de Economía*, 6-19.
- Lawrence, M., Goodwin, P., O'connor, M., & Onkal, D. (2006). Judgmental Forecasting: a Review Of Progress Over The Last 25 years. *International Journal of Forecasting*, 3(22), 493-518.
- Leit N et al. (2019). Open Source Business Intelligence on a SME: A Case Study using Pentaho. *Iberian Conference on Information Systems and Technologies (CISTI)(14)*, 1-7.
- Li, S., & Guo, S. (2021). Persistence and Extinction Of A Stochastic SIS Epidemic Model With Regime Switching And Lévy Jumps. *DCDS-B*, 9(26).
- Liu, Q., Liu, X., & Jiang, B. (2011). Forecasting incidence of hemorrhagic fever with renal syndrome in China using ARIMA model. *BMC Infect DIS*, 11-218.
- Medina Soto, J. (2005). Business Intelligence: Conceptos y Actualidad. *Diario Gestipolis*, 1-14. Retrieved from <https://www.gestipolis.com/wp-content/uploads/2010/06/business-intelligence-conceptos-y-actualidad-1.pdf>
- Merino, G. (2017). Análisis de los Modelos de Inteligencia de Negocios basados en

- Big Data en las Pymes del Ecuador. *Revista Ciencia y Tecnología*, 18(17), 46-56. Retrieved from <https://doi.org/10.47189/rcct.v18i17.157>
- Michón Medina , C. A., & Vizconde Osorio , T. A. (2011). Modelo SARIMA para la llegada mensual de visitantes extranjeros por el Aeropuerto Internacional "Jorge Chávez ". *Scentia*.
- Muñoz , S., Urquijo Vanstrahlengs, J., Castro Otero, A., & Lombana , J. (2017). Pronóstico del precio de la energía en Colombia utilizando modelos ARIMA con IGARCH. *Revista de Economía del Rosario* , 127-161.
- Muñoz et al. (2016). Inteligencia de los negocios. Clave del Éxito en la era de la información. *Revista Clío América*, 10(20), 194-211.
- Olivia, F. d. (2016). Aproximación a la metodología Box-Jenkins para la predicción de la tasa de cambio EUR/USD. *La Habana*.
- Pourshahid et al. (2011). Toward a Goal-Oriented, Business Intelligence Decision - Making Framework. *Telfer School Management* , 78, 100-115. Retrieved from https://doi.org/10.1007/978-3-642-20862-1_7
- Pupiales, J. (2019). Business Intelligence como soporte en la toma de decisiones de la Secretaría Académica de UNIANDES Ibarra. *UNIANDES*, 1-4. Retrieved from <https://dspace.uniandes.edu.ec/handle/123456789/10219>
- Ragozin , A. (2022). The Application Of Digital Signal Processing Technology To Improve The Accuracy Of Forecasting Time Series Data In Anomaly Detection System In the Observed Processes Of Automated Process Control Systems. *IS*, 4(22).
- Rendón-Macías, M. E.-K.-N. (2016). Estadística descriptiva. *Revista Alergia México*, 63(4), 397-407. Retrieved from <https://revistaalergia.mx/ojs/index.php/ram/article/view/230>
- Rodríguez, & Pedraja. (2016). Innovación y empresas: un análisis de sus características a nivel regional. *IDESIA*, 34(1), 69-75.
- Ruiz-Quintero. (2014). Inteligencia de negocios al proceso de la evaluación docente. *In Vestigium Ire*, 8(1), 208-214. Retrieved from <http://revistas.ustatunja.edu.co/index.php/ivestigium/article/view/1056/1022>
- Silva Solano, L. (2017). Business Intelligence: un balance para su implementación. *InnovaG*(3), 27-36. Retrieved from <https://revistas.pucp.edu.pe/index.php/innovag/article/view/19742>
- Vinueza, P. (2016, Octubre 14). *Correlación: teoría y práctica*. Retrieved from CCG-

UNAM: <http://www.ccg.unam.mx/~vinuesa/>

Xu, X., Du, Y., & Lie, J. (2022). An Improved Arima Model Based On Regularized Gaussian Basis Function and Its Application To Stock Price Forecasting. *BCPBM*(30), 112-121.

DECLARACIÓN Y AUTORIZACIÓN

Yo, **Núñez Ordóñez, Daniel Roberto**, con C.C: # **0932137003** autor/a del trabajo de integración curricular: **Análisis y pronóstico de las ventas para la optimización de tiempos de importación en la empresa IMPORGUIDSA mediante la aplicación del modelo de series de tiempos ARIMA** previo a la obtención del título de **Licenciado en Negocios Internacionales** en la Universidad Católica de Santiago de Guayaquil.

1.- Declaro tener pleno conocimiento de la obligación que tienen las instituciones de educación superior, de conformidad con el Artículo 144 de la Ley Orgánica de Educación Superior, de entregar a la SENESCYT en formato digital una copia del referido trabajo de integración curricular para que sea integrado al Sistema Nacional de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor.

2.- Autorizo a la SENESCYT a tener una copia del referido trabajo de integración curricular, con el propósito de generar un repositorio que democratice la información, respetando las políticas de propiedad intelectual vigentes.

Guayaquil, **04 de septiembre de 2023**

f. 

Nombre: **Núñez Ordóñez, Daniel Roberto**

C.C: **0932137003**

REPOSITORIO NACIONAL EN CIENCIA Y TECNOLOGÍA

FICHA DE REGISTRO DE TRABAJO DE INTEGRACIÓN CURRICULAR

TEMA Y SUBTEMA:	Análisis y pronóstico de las ventas para la optimización de tiempos de importación en la empresa IMPORGUIDSA mediante la aplicación del modelo de series de tiempos ARIMA.		
AUTOR(ES)	Núñez Ordóñez, Daniel Roberto		
REVISOR(ES)/TUTOR(ES)	Carrera Buri, Félix Miguel		
INSTITUCIÓN:	Universidad Católica de Santiago de Guayaquil		
FACULTAD:	Facultad de Economía y Empresa		
CARRERA:	Negocios Internacionales		
TÍTULO OBTENIDO:	Licenciado en Negocios Internacionales		
FECHA DE PUBLICACIÓN:	04 de septiembre de 2023	No. DE PÁGINAS:	#69
ÁREAS TEMÁTICAS:	Área logística, área de ventas, Área de importaciones.		
PALABRAS CLAVES/ KEYWORDS:	ARIMA, Importaciones, Business Intelligence, Inventario, Ventas, Series de tiempos.		
RESUMEN/ABSTRACT (150-250 palabras):	<p>En un entorno empresarial marcado por la creciente globalización y la dinámica cambiante de los mercados, la optimización de la gestión de ventas y la eficiente administración de inventario se han convertido en pilares fundamentales para el éxito y la competitividad de las organizaciones. Este trabajo de investigación se centra precisamente en abordar estos desafíos en el contexto operativo de la importadora IMPORGUIDSA a través de un enfoque analítico basado en el modelo ARIMA (AutoRegressive Integrated Moving Average). El propósito principal de esta investigación es proporcionar una metodología sólida para analizar y proyectar las ventas de IMPORGUIDSA, permitiendo una anticipación precisa de las demandas del mercado. Con la implementación del modelo ARIMA, se busca no solo prever las tendencias futuras de ventas, sino también optimizar los tiempos de importación y la gestión del inventario en consonancia con las necesidades actuales del mercado. Para llevar a cabo este análisis, se ha empleado la herramienta R Studio, que ofrece un entorno eficiente y potente para la manipulación y visualización de datos. Mediante el análisis detallado de los datos obtenidos, esta investigación busca ofrecer una visión informada sobre las pautas de compra de los clientes, los patrones estacionales y las fluctuaciones del mercado. Los resultados obtenidos proporcionarán a IMPORGUIDSA información valiosa para tomar decisiones estratégicas informadas, adaptando su enfoque de importación y gestión de inventario para satisfacer de manera efectiva las demandas cambiantes de los consumidores.</p>		
ADJUNTO PDF:	<input checked="" type="checkbox"/> SI	<input type="checkbox"/> NO	
CONTACTO CON AUTOR/ES:	Teléfono: +593 0995432070	E-mail: daniel.nunez01@cu.ucsg.edu.ec	
CONTACTO CON LA INSTITUCIÓN (COORDINADOR DEL PROCESO UIC):	Nombre: Román Bermeo, Cynthia Lizbeth		
	Teléfono: +593-984228698		
	E-mail: cynthia.roman@cu.ucsg.edu.ec		
SECCIÓN PARA USO DE BIBLIOTECA			
Nº. DE REGISTRO (en base a datos):			
Nº. DE CLASIFICACIÓN:			
DIRECCIÓN URL (tesis en la web):			