Universidad Nacional Mayor de San Marcos
Universidad del Perú. Decana de América
Facultad de Ciencias Matemáticas
Escuela Profesional de Estadística

Modelo de Box-Jenkins para el pronóstico de la recaudación de ingreso de la genérica 1.3 de una entidad pública, 2017 – 2021

TRABAJO DE SUFICIENCIA PROFESIONAL

Para optar el Título Profesional de Licenciada en Estadística

AUTOR

Averly David ROQUE GARCÍA

ASESOR

Dr. Hefler Joel MOLINA QUIÑONES

Lima, Perú

2021



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

https://creativecommons.org/licenses/by-nc-sa/4.0/

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Roque, A. (2021). *Modelo de Box-Jenkins para el pronóstico de la recaudación de ingreso de la genérica 1.3 de una entidad pública, 2017 – 2021*. [Trabajo de suficiencia profesional de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Unidad de Posgrado]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	Averly David Roque García
Tipo de documento de identidad	DNI
Número de documento de identidad	45477508
URL de ORCID	No aplica
Datos de asesor	
Nombres y apellidos	Helfer Joel Molina Quiñones
Tipo de documento de identidad	DNI
Número de documento de identidad	40014631
URL de ORCID	https://orcid.org/0000-0003-1307-7253
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Carlos Alberto Jaimes Velasquez
Tipo de documento	DNI
Número de documento de identidad	42762905
Miembro del jurado 1	
Nombres y apellidos	Mg. Hugo Marino Rodríguez Orellana
Tipo de documento	DNI
Número de documento de identidad	40162362
Datos de investigación	
Línea de investigación	A.3.2.6. Análisis de Datos y Modelamiento de Problemas de la Sociedad (Empresas, Instituciones, Poblaciones locales, regionales y nacionales).
Grupo de investigación	No aplica.
Agencia de financiamiento	Sin financiamiento.

Ubicación geográfica de la investigación	Edificio: Hospital de Rehabilitación del Callao. País: Perú Departamento: Provincia Constitucional del Callao. Provincia: Provincia Constitucional del Callao. Distrito: Bellavista Calle: Jr. Vigil 535 Latitud: -12.0631753 Longitud: -77.1331101
Año o rango de años en que se realizó la investigación	Enero 2021 - junio 2021
URL de disciplinas OCDE	Estadística, Probabilidades https://purl.org/pe-repo/ocde/ford#1.01.03

ACTA DE SUSTENTACIÓN DE TRABAJO DE SUFICIENCIA PROFESIONAL EN LA MODALIDAD VIRTUAL PARA OBTENCIÓN DEL TÍTULO PROFESIONAL DE LICENCIADO EN ESTADÍSTICA

En Lima, siendo las 17:30 horas del domingo 03 de octubre del 2021, se reunieron los docentes designados como Miembros del Jurado del Trabajo de Suficiencia Profesional: Mg. Carlos Alberto Jaimes Velasquez (PRESIDENTE), Mg. Hugo Marino Rodriguez Orellana (MIEMBRO) y el Dr. Helfer Joel Molina Quiñones (MIEMBRO ASESOR), para la sustentación del Trabajo de Suficiencia Profesional titulado: "**MODELO DE BOX-JENKINS PARA EL PRONOSTICO DE LA RECAUDACION DE INGRESO DE LA GENERICA 1.3 DE UNA ENTIDAD PÚBLICA, 2017 - 2021**", presentado por el señor **Bachiller Averly David Roque García**, para optar el Título Profesional de Licenciado en Estadística.

Luego de la exposición del trabajo de suficiencia, el Presidente invitó al expositor a dar respuesta a las preguntas formuladas.

Realizada la evaluación correspondiente por los miembros del Jurado Evaluador, el expositor mereció la aprobación de **BUENO,** con un calificativo promedio de **QUINCE (15)**.

A continuación, los miembros del Jurado dan manifiesto que el participante **Bachiller Averly David Roque García** en vista de haber aprobado la sustentación del Trabajo de Suficiencia Profesional, será propuesto para que se le otorgue el Título Profesional de Licenciado en Estadística.

Siendo las 18:00 horas se levantó la sesión firmando para constancia la presente Acta.

Mg. Carlos Alberto Jaimes Velasquez
PRESIDENTE

Mg. Hugo Marino Rodriguez Orellana
MIEMBRO

Dr. Helfer Joel Molina Quiñones
MIEMBRO ASESOR

La Vicedecana de la Facultad de Ciencias Matemáticas, Mg. Zoraida Judith Huamán Gutiérrez, certifica virtualmente la participación del Jurado Evaluador, el titulando, el acto de instalación y el inicio, desarrollo y término del acto académico de sustentación, dejando constancia en el acta respectiva.

DEDICATORIA

A Dios por guiarme en esta etapa de la vida y sobre todo porque me brindaste el don de la perseverancia para alcanzar esta meta.

A mi madre y mi padre, porque gracias al sacrificio y esfuerzo de ustedes esto no habría sido posible, ya que me forjaron como persona y sobretodo me motivaron constantemente para alcanzar mis anhelos, por lo que siempre les estaré eternamente agradecido.

AGRADECIMIENTOS

A mi familia por apoyarme en distintas etapas de mi vida los que me permitieron llegar a este logro.

A los docentes de la Escuela Profesional de Estadística de la Universidad Nacional Mayor de San Marcos por su ardua labor como educadores y forjadores profesionales.

A las grandes amistades que pude conocer en las aulas durante la etapa universitaria en mi ahora alma mater UNMSM.

INDICE

Capítulo I. introducción 9
Capítulo II. Información del lugar donde se desarrolló la actividad 10
Institución donde se desarrolló la actividad 10
Periodo de duración de la actividad 10
Finalidad y objetivos de la entidad 10
Razón social 11
Dirección postal 11
Correo electrónico del profesional a cargo 11
Capítulo III. Descripción de la actividad 12
Organización de la actividad 12
Finalidad y objetivos de la actividad 14
Problemática 14
Metodología y procedimientos 16
Resultado de la actividad 27
Capítulo IV. Conclusiones 33
Capítulo V. Recomendaciones 34
Capítulo VI. Bibliografía 34
Capítulo VII. Anexos 345

ÍNDICE DE TABLAS

Tabla 1: Características de una serie estacionaria y no estacionaria.....23

Tabla 2: Características de las autocorrelaciones simples y parciales.....24

Tabla 3: Estadísticos de los errores de los posibles modelos identificados de la serie ingresos mensuales de la genérica de ingreso 1.3 periodo julio 2017-junio2021.....30

Tabla 4: Pronóstico de la genérica de ingreso 1.3 para el periodo de julio a diciembre correspondiente al año 2021.....32

ÍNDICE DE FIGURAS

Figura 1: Flujograma del modelo de Box-Jenkins……13

Figura 2: Comportamiento de los ingresos mensuales de la genérica 1.3 periodo junio 2017-julio2021……27

Figura 3: Serie de tiempo de los ingresos mensuales de la genérica de ingreso 1.3 periodo julio 2017-junio2021…….. ...28

Figura 4: Serie de tiempo estacionaria de los ingresos mensuales de la genérica de ingreso 1.3 periodo julio 2017-junio2021 luego de aplicar una diferencia……..29

Figura 5: Autocorrelaciones simples y parciales de la serie estacionaria de los ingresos mensuales de la genérica de ingreso 1.3 periodo julio 2017-junio2021……29

Figura 6: Diagnóstico gráfico del modelo propuesto ARIMA (1,1,2) de la serie ingresos mensuales de la genérica de ingreso 1.3 periodo julio 2017-junio2021……31

Figura 7: Pronóstico de la genérica de ingreso 1.3 para el periodo de julio a diciembre correspondiente al año 2021……..32

RESUMEN

El presente trabajo tuvo como finalidad determinar un modelo adecuado de serie de tiempo, el cual permita describir y pronosticar los ingresos mensuales que genera la entidad pública, en la genérica de ingreso 1.3 por concepto de venta de servicios, bienes y derechos administrativos, cuyos registros fueron obtenidos del Sistema Integrado de Administración Financiera (SIAF SP). Para tal fin se recurrió a la utilización de la metodología de Box-Jenkins, para identificar el modelo optimo que permitan explicar los registros obtenidos y el pronóstico para los meses de julio a diciembre del año 2021. Finalmente se concluye que el mejor modelo evidenciado es el ARIMA (1,1,2).

Palabras claves: Box-Jenkins, genérica de ingreso 1.3, ARIMA

ABSTRACT

The purpose of this work was to determine an adequate time series model, which allows describing and forecasting the monthly income generated by the public entity, in the generic income 1.3 for the sale of services, goods and administrative rights, whose records They were obtained from the Integrated Financial Administration System (SIAF SP). For this purpose, the Box-Jenkins methodology was used to identify the optimal model to explain the records obtained and the forecast for the months of July to December of the year 2021. Finally, it is concluded that the best evidenced model is the ARIMA (1,1,2).

Keywords: Box-Jenkins, generic entry 1.3, ARIMA

Capítulo I. introducción

De acuerdo al Decreto Legislativo 1440, el presupuesto público es un instrumento de gestión del Estado para el logro de resultados a favor de la población, mediante la prestación de servicios y el logro de las metas de cobertura con eficacia y eficiencia por las Entidades Públicas. Donde se establece los ingresos que los financian y así como también los límites de gasto durante el año fiscal correspondiente, los cuales son distribuidos a nivel de pliegos presupuestarios de las entidades del sector público.

Los ingresos generados por las distintas entidades del estado y que son administradas por esta, se ven reflejados en la fuente de financiamiento Recursos Directamente Recaudados, entre los cuales se puede mencionar la prestación de servicios, venta de bienes, rentas de propiedad, las tasas, entre otros; así como otros ingresos que les corresponde de acuerdo a la normatividad vigente. (MEF)

Para el presente estudio se tomó los ingresos que la entidad ha generado a través de la recaudación en la genérica de ingresos 1.3 “Venta de bienes y servicios y derechos administrativos”, los cuales fueron los recursos financieros provenientes por la venta de bienes y servicios que la entidad brindó, así como los derechos administrativos o tasas que se cobran por la prestación de un servicio individualizado de carácter obligatorio y que es de naturaleza regulatoria; durante los meses de julio del 2017 a junio del 2021. Para lo cual se usó la metodología de Box-Jenkins, a fin de realizar el pronóstico de la recaudación de ingresos de la Entidad.

El trabajo consta de 8 capítulos, en el primer capítulo se trata de la introducción, en el segundo capítulo se aborda la información del lugar donde se desarrolló la actividad, el tercer capítulo trata de la descripción de la actividad, donde se desarrollará la finalidad y objetivos de la actividad, así como la problemática, metodología, procedimientos y resultados de la misma, en el cuarto capítulo van las conclusiones, en el capítulo quinto se detalla las recomendaciones, mientras que en el capítulo sexto se describe las bibliografías y finalmente los anexos y/o ilustraciones.

Capítulo II. Información del lugar donde se desarrolló la actividad

Institución donde se desarrolló la actividad

La realización de la actividad se desarrolló en el Hospital de Rehabilitación del Callao, el cual un 28 de marzo del 2017, mediante Ordenanza Regional Nº 000003 del Gobierno Regional de la Provincia Constitucional del Callao, publicada en el Diario Oficial El Peruano, se aprueba la creación del Hospital de Rehabilitación del Callao como una entidad de nivel de atención II-E, Unidad Ejecutora (especializado en medicina física y rehabilitación). La creación de la nueva Unidad Ejecutora de la Región Callao, tenía como fin contribuir en la mejora de la calidad del servicio público brindado a la población de la Región Callao, teniendo en consideración especial a las personas que sufren de alguna discapacidad.

Periodo de duración de la actividad

Desde el 01 de febrero al 31 de julio del año 2021

Finalidad y objetivos de la entidad

De acuerdo a los instrumentos de gestión del Hospital de Rehabilitación del Callao, se define de la siguiente manera.

Misión

"Somos un Hospital Especializado de Segundo Nivel de atención que brinda servicios de salud en Medicina Física y Rehabilitación, con calidad y oportunidad, con personal identificado en el marco del ejercicio pleno del Derecho a la Salud de nuestros usuarios con discapacidad."

Visión

"En el 2021 seremos un Instituto Especializado en Medicina Física y Rehabilitación, reconocidos por brindar servicios de salud con calidad y oportunidad, claramente percibidos por nuestros usuarios, en cumplimiento de los Derechos en Salud."

Objetivos

- Brindar atención especializada a la población con discapacidad en estado vulnerable con índices de pobreza, exclusión y marginalidad.
- Cubrir las atenciones de rehabilitación integral a la población con discapacidad física, sensorial y mental, con inclusión social.
- Incrementar las certificaciones de atención de personas con discapacidad de la región Callao.
- Mejorar el nivel de satisfacción del usuario externo en el hospital.
- Lograr el abastecimiento y mejoramiento del acceso a medicamentos genéricos de calidad, propiciando su uso racional.
- Fomentar acciones orientadas a mejorar la gestión en la prevención de riesgos frente a emergencias y desastres.
- Garantizar la eficiente gestión económica, material y financiera del hospital.
- Fortalecer el desarrollo del recurso humano a través de la gestión por competencias e impulso de capacidad docente e investigación.

Razón social

RUC: 20601761999 - Unidad Ejecutora Hospital de Rehabilitación del Callao

Dirección postal

Jr. Vigil 535 - Prov. Const. del Callao - Prov. Const. del Callao - Bellavista

Correo electrónico del profesional a cargo

roque.oppm@gmail.com

Capítulo III. Descripción de la actividad

Organización de la actividad

Para el desarrollo del presente trabajo se utilizó la metodología de Box-Jenkins, el cual consta de cuatro pasos:

Paso 1: Identificación

Este primer paso consiste en que una vez obtenido los registros mensuales de los ingresos que ha generado el Hospital de Rehabilitación del Callao, por concepto de "Venta de bienes y servicios y derechos administrativos" correspondiente a los meses de julio del año 2017 a junio del año 2021, se procede a identificar el posible modelo ARIMA que sigue la serie. Donde el objetivo es determinar los ordenes (p,d,q) óptimos que caractericen mejor a la serie.

Paso 2: Estimación

Luego de haber identificado los valores de las ordenes de (p,d,q), se realiza la estimación de los parámetros ($\phi_i \; y \; \theta_i$) los cuales se pueden realizar a través de máxima verosimilitud, donde se obtienen el error estándar y los residuos del modelo planteado.

Paso 3: Validación

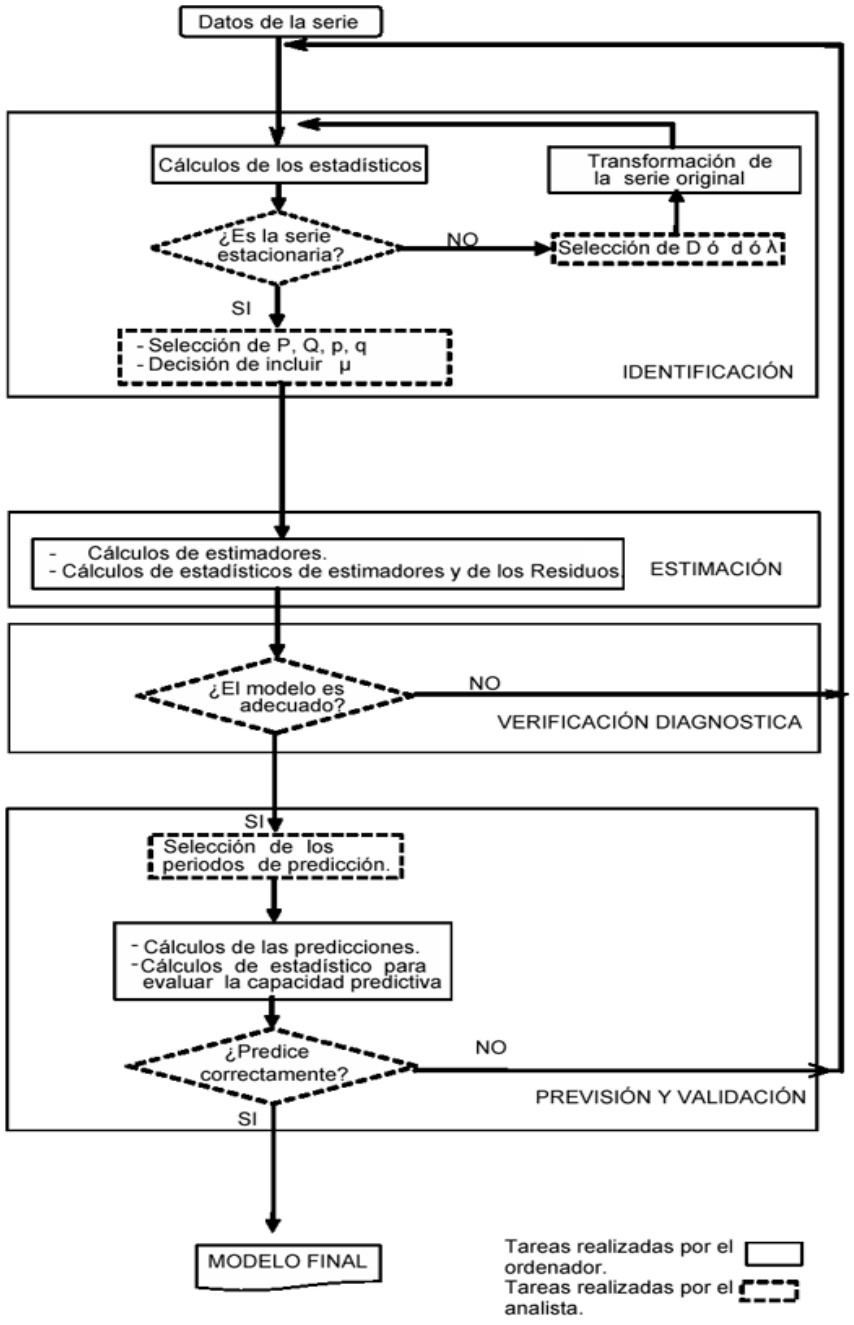
La validación del modelo consiste en realizar los contrastes a fin de comprobar si el modelo elegido se ajusta a los datos, es decir los residuos siguen un proceso de comportamiento de ruido blanco. En caso de no ser así, se repiten los pasos anteriores.

Paso 4: Pronóstico

Luego de haber identificado el modelo optimo, se realiza el pronóstico de los valores futuros de la variable ingresos mensuales del hospital.

Figura 1

Flujograma del modelo de Box-Jenkins



Nota. Adaptado de "Previsión de la temperatura superficial del mar frente a la costa peruana mediante un modelo autoregresivo integrado de media móvil" (p.110), por C. Quispe y S. Purca, 2007, Revista Peruana de Bilogía, 14(1).

Finalidad y objetivos de la actividad

Finalidad

Aplicar la metodología de Box-Jenkins para el periodo de julio del año 2017 a junio del 2021, para pronosticar los ingresos mensuales de la Entidad, de la genérica de ingresos 1.3.

Objetivos

- Describir el comportamiento de los ingresos mensuales de la Entidad, desde julio 2017 hasta junio del 2021.
- Estimar el modelo adecuado de los ingresos mensuales de la genérica 1.3 de la entidad.
- Pronosticar el comportamiento de la serie ingresos mensuales de la genérica de ingreso 1.3, para los meses de julio a diciembre para el año 2021.

Problemática

En la presente investigación se dará a conocer los antecedentes más importantes, en cuanto a la aplicación de la metodología de Box-Jenkins.

Salazar & Saavedra (2014), realizaron un estudio de Pronostico de la ejecución del presupuesto asignado a la Unidad Ejecutora 001 Sede Central perteneciente a Piura, durante el periodo 2013 y 2014, aplicando el modelo de Holt-Winters y los modelos ARIMA y SARIMA. El objetivo de la investigación fue evaluar la capacidad de los modelos Holt-Winters, ARIMA y SARIMA en el pronóstico de la ejecución del presupuesto que fue asignado al Pliego del Gobierno Regional de Piura, en la unidad ejecutora 001 sede central durante el periodo del 2013 y 2014. Los resultados que obtuvieron mediante el modelo de Holt-Winters, fue que le proporcionó un valor de la raíz del error cuadrático medio de 3334077, cuyos parámetros de suavización fueron $\alpha= 0.3$, $\beta=0.16$ y $\gamma=0.000$, mientras que el mejor modelo ARIMA fue ARIMA (0,1,1) y finalmente el mejor modelo SARIMA evidenciado fue SARIMA $(0,1,1)x(0,1,12)_{12}$. El estudio concluye que el mejor modelo fue el SARIMA $(0,1,1)x(0,1,12)_{12}$ reportando así el menor valor para la raíz del error cuadrático medio.

Orosco (2017), desarrolló su investigación referente a la Metodología Bootstrap, para el análisis de los ingresos mensuales del Hotel El Centro, en el Departamento de Ayacucho-Perú durante el periodo del 2001-2014, que tuvo como objetivo analizar el comportamiento serial

de los ingresos mensuales del Hotel El Centro ubicado en el Departamento de Ayacucho durante el periodo de 2011 al 2014, mediante la metodología Bootstrap y de Box-Jenkins. El resultado obtenido muestra que el mejor modelo obtenido para el pronóstico de los ingresos del Hotel El Centro durante el periodo 2011-2014 es el SARIMA $(0,1,1)x(0,1,0)_{12}$, asimismo el autor concluye que, de ambas metodologías utilizadas, la que mejor resultado obtiene es la de Bootstrap con un valor del error absoluto medio porcentual del 7.784% en comparación la de Box-Jenkins que arroja un valor del error absoluto medio porcentual del 7.462%.

Corman (2020), realizó el estudio del comportamiento del gasto público de una entidad pública durante los años 2009-2017, aplicando para ello la metodología de Box-Jenkins. El objetivo del estudio fue la de describir el comportamiento del gasto público de la entidad, durante los años 2009 al 2017 y su relación con las directrices y/o lineamientos de las políticas expansivas emitidas por el Estado, empleando para ello la metodología de Box-Jenkins, con la finalidad de identificar los periodos y las razones en las que hubo disminuciones o aumentos del gasto público de la entidad. Los resultados indican que la aplicación de la metodología de Box-Jenkins, le permitió analizar de forma apropiada los cambios estructurales significativos que presentó la serie del gasto público de la entidad, para lo cual el mejor modelo obtenido es el ARIMA $(2,0,0)x(0,1,1)_{12}$. Finalmente, el autor menciona que el pronóstico realizado para el año 2018, evidencia una disminución en el gasto público de la entidad en comparación con el año anterior, explicándose ello debido a una medida de austeridad planteada por el poder ejecutivo mediante un decreto de urgencia, cuya finalidad era generar una mayor eficiencia en la ejecución presupuestal y el afrontamiento del déficit fiscal.

Lemus, Torres, Cubillos & Camelo (2017), realizaron el estudio sobre el análisis de la ejecución presupuestal en el país de Colombia para los años 1954 al 2013. De acuerdo al estudio el objetivo fue reconocer el comportamiento de la inversión pública en el país de Colombia durante los años 1954 al 2013, y la relación con los lineamientos de política económica contenidos en los planes de desarrollo suscritos por los distintos gobiernos de entonces. Para ello aplicaron la metodología de Box-Jenkins, para ello pudieron evidenciar que el mejor modelo fue el ARIMA (1,1,0), el cual permitió identificar orientaciones presupuestales concretas hacia intereses políticos y gubernamentales definidos en los planes de desarrollo del país de Colombia. Finalmente concluyen que la inversión pública se

encuentra altamente influenciada por los planes de desarrollo de cada gobierno de turno, asimismo se identificaron una tendencia creciente con periodos de estancamiento o decrecimiento debido a las coyunturas económicas surgidas en los respectivos gobiernos.

Asimismo, se realizó un diagnóstico de la problemática que involucra a la variable de investigación, en ese sentido y de acuerdo a todos los antecedentes evidenciados, se requiere realizar el presente estudio.

La Entidad, es una institución especializada en medicina física y rehabilitación de pacientes con discapacidad, cuyo objetivo es atender a las personas con discapacidad del primer puerto y a todas las personas de diferentes procedencias del país sin ninguna excepción. Año tras año la entidad presenta un bajo presupuesto asignado por parte del Gobierno, por lo que es importante generar la mayor recaudación posible de acuerdo a su Presupuesto Institucional Modificado (PIM), a través de los servicios que se brinda como institución. Desde inicios del año 2020 el país está atravesando por un estado de emergencia, a raíz de la pandemia producida por el COVID-19, el cual es un factor determinante en todos los sectores y carteras de servicios del estado peruano, por lo que la recaudación de ingresos se ve afectado. En ese sentido para el presente estudio se empleó la metodología de Box-Jenkins para pronosticar el ingreso de la genérica de ingreso 1.3 “Venta de bienes y servicios y derechos administrativos” que la entidad recauda a través de la fuente de financiamiento Recursos Directamente Recaudados. Dicha metodología aplicada es importante para la Entidad, ya que permitió tomar las medidas del caso para una mejor toma de decisiones.

Metodología y procedimientos

Variable de estudio

Ingreso mensual de la genérica de ingresos 1.3 “Venta de bienes y servicios y derechos administrativos”

Tipo de investigación

Descriptivo

Diseño

No experimental longitudinal

Población

Registro de los montos mensuales de la genérica 1.3 “Venta de bienes y servicios y derechos administrativos” durante los años de julio 2017 a junio 2021.

Series de Tiempo

La definición de una serie de tiempo, podríamos explicarlo como una secuencia cronológica o conjunto de mediciones de observaciones de una variable en particular. Estas observaciones son valores usualmente registrados en instantes de tiempos equidistantes. Muchos son los ejemplos de los sectores donde se podría aplicar las de series de tiempo, algunos de ejemplo son:

- Series Económicas (Precios mensuales de cierto producto, Volumen de exportaciones, etc.)
- Series Demográficas (Cantidad anual de nacimientos, Defunciones mensuales, etc.)
- Series Físicas (Temperaturas promedio mensual, velocidad promedio diaria del viento, etc.)

Toda serie de tiempo proviene de un proceso estocástico, el cual se puede definir como una familia de variables aleatorias relacionadas entre sí. Un proceso estocástico se caracteriza por su función de distribución o por su momento (Gonzales et al., 2009).

En las series de tiempo se evidencian algunos enfoques para el análisis de las mismas, para los cuales existen tres tipos de enfoques para el tratamiento de las series de tiempo, los cuales son:

1) Enfoque de Descomposición: Conocido también como enfoque tradicional, el cual utiliza los modelos de error y descomposición.
2) Enfoque de Dominio del Tiempo: Está basado en la asociación de cada valor de la serie en cuestión, con su pasado inmediato más una variable puramente aleatoria (comúnmente llamado ruido blanco). Este enfoque son los modelos ARIMA o también conocido como modelos de Box - Jenkins.
3) Enfoque de Dominio de Frecuencias: De acuerdo a lo realizado por Schuster (1851-1934), el cual crea una herramienta llamada Periodograma, el cual consiste en descomponer la serie de tiempo en todas las frecuencias posibles, para posteriormente asociarlas a una intensidad en cada una de ellas.

Modelo de Box-Jenkins

Este modelo fue propuesto en 1976 por Box & Jenkins, a través de su publicación Times Series Analysis: forecasting and control. Dicha publicación marcó el inicio de las herramientas de pronóstico, a la metodología de Box-Jenkins también se le conoce como metodología ARIMA (Stellwagen & Tashman, 2013). La metodología consiste en modelar el proceso generador de un grupo de datos, el cual consta de un procedimiento iterativo dividido en cuatro etapas: identificación del modelo, estimación de los parámetros, comprobación y elaboración del pronóstico (Bowerman, O'Connell, & Koehler, 2009; Kumar & Anand, 2015; Stellwagen & Tashman, 2013).

De acuerdo a (Gonzales, 2009) y (De la Fuente Fernández, (2018).) los modelos ARIMA (modelo de Box-Jenkins), se explica de la siguiente forma:

Modelos Autoregresivos AR(p)

Los modelos autoregresivos AR, describen una clase particular del proceso en el cual las observaciones dado un determinado instante, son previsibles a partir de las observaciones preliminares del proceso, más un término de error. Dicho ello, el modelo ARIMA (1,0,0) es el caso más simplificado o llamado proceso autoregresivo de primer orden AR (1). Se define como:

$$\mathrm{AR}(1) \equiv \mathrm{Z_t} = \phi_1 Z_{t-1} + a_t$$

El proceso autoregresivo de orden p, denotado por ARIMA (p,0,0) o simplemente denotado como AR(p), define $\mathrm{Z_t}$ en función a su pasado hasta el retardo $t - p$ mas una innovación contemporánea, denotada como:

$$\mathrm{AR(p)} \equiv \mathrm{Z_t} = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \cdots + \phi_p Z_{t-p} + a_t \qquad a_t \sim RB(0, \sigma^2)$$

Donde:

ϕ_i : Parámetros i = 1, 2, …, p

a_t : Ruido Blanco

En términos del operador de retardos, se puede expresar de la siguiente forma:

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

$$\phi_p(B) Z_t = a_t$$

Donde $\phi_p(B)$ es el polinomio autoregresivo y $(\phi_1, \phi_2, \ldots, \phi_p)$ es el vector de parámetros autoregresivos.

Los procesos autoregresivos finitos AR(p) son estacionarios si y solo si las raíces del polinomio autoregresivo $\phi_p(B)$ caen fuera del círculo unitario (Box, 2008). Esta condición es equivalente decir a que las raíces de la ecuación: $Z^p - \phi_1 Z^{p-1} - \phi_2 Z^{p-2} - \cdots - \phi_{p-1} Z - \phi_p = 0$ todas sean inferiores a uno en modulo.

Asimismo, un proceso autoregresivo siempre será invertible.

Modelos Autoregresivos MA(q)

Los modelos de medias móviles MA, describen una serie de tiempo estacionaria. En estos tipos de modelo, el valor actual se puede predecir a partir del componente aleatorio del momento, pero también se puede predecir a partir del impulso aleatorio anterior. El modelo ARIMA (0,0,1), también conocido como MA (1), se define de la siguiente manera:

$$Z_t = a_t - \theta_1 a_{t-1}$$

Los procesos de medias móviles de orden q, denotado por ARIMA (0, 0, q), o simplemente MA(q), se define como:

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q} \qquad a_t \sim RB(0, \sigma^2)$$

Donde:

θ_i : Parámetros i = 1, 2, …, q

a_t : Ruido Blanco

En relación del operador de retardos es de la siguiente manera:

$$Z_t = (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q) a_t$$

$$Z_t = \theta_q(B) a_t$$

Donde $\theta_q(B)$ hace referencia al polinomio de medias móviles y $(\theta_1, \theta_2, \ldots, \theta_q)$ es el vector parámetros autoregresivos.

Los procesos de medias móviles siempre son estacionarios.

Los procesos de medias móviles MA(q) son invertibles si las raíces del polinomio de medias móviles $\theta_q(B)$ están fuera del círculo unidad (Box, 2008). Dicha condición equivale decir a que las raíces de la ecuación: $Z^q - \phi_1 Z^{q-1} - \phi_2 Z^{q-2} - \cdots - \phi_{q-1} Z - \phi_q = 0$ todas sean inferiores a uno en modulo.

Modelos ARMA (p, q)

Una forma extendida de los modelos AR(p) y MA(q), son aquellos modelos donde se incluyen términos autoregresivos y términos de medias móviles, los cuales se denotan como ARIMA (p, 0, q) y en forma de ecuación se denotan como:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \cdots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}$$

$$a_t \sim RB(0, \sigma^2)$$

En función del operador de retardos es de la siguiente forma:

$$Z_t - \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \cdots + \phi_p Z_{t-p} = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}$$

$$Z_t(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) = a_t(1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q)$$

$$\phi_p(B) Z_t = \theta_q(B) a_t$$

Donde:

$\phi_p(B)$: polinomio autoregresivo

$\theta_q(B)$: polinomio de medias móviles.

Los procesos autoregresivos de medias móviles finitos ARMA(p, q) son estacionarios si y solo si las raíces del polinomio denotado por $\phi_p(B)$, caen fuera del circulo unitario. Dicha condición es equivalente a que las raíces de la ecuación: $Z^p - \phi_1 Z^{p-1} - \phi_2 Z^{p-2} - \cdots - \phi_{p-1} Z - \phi_p = 0$ todas sean inferiores a uno en modulo.

Los procesos autoregresivos de medias móviles finitos ARMA(p, q) se dice que es invertible, si y solo si las raíces del polinomio denotado por $\theta_q(B)$ esten fuera del circulo unidad. Dicha condición es equivalente decir que las raíces de la ecuación: $Z^q - \phi_1 Z^{q-1} - \phi_2 Z^{q-2} - \cdots - \phi_{q-1} Z - \phi_q = 0$ todas sean inferiores a uno en modulo.

Modelos ARIMA (p, d, q)

Los modelos ARIMA (0, d, 0) es una serie de tiempo que se convierte en ruido blanco (un proceso completamente aleatorio) luego de haber sido diferenciados **d** veces.

El modelo ARIMA (0, d, 0) se representa como: $(1 - B)^d Z_t = a_t$

Los modelos ARIMA (p, d, q) o también llamado proceso Autoregresivo Integrado de Medias Móviles de orden (p, d, q), en términos de ecuación se expresa como:

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)(1 - B)^d Z_t = (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q) a_t$$

$$\phi_p(B) \Delta^d Z_t = \delta + \theta_q(B) a_t$$

Donde $\phi_p(B)$: polinomio autoregresivo estacionario y $\theta_q(B)$: polinomio invertible de medias móviles no tienen raíces comunes.

Los modelos ARIMA (p, d, q) describen una serie de observaciones, después de que estas fuesen diferenciadas d veces, con la finalidad de extraer las posibles fuentes de no estacionariedad. Dicha expresión se puede aplicar a cualquier modelo. Si hay p, d, q componentes que son cero, el término correspondiente se elimina de la formula general.

Los modelos estacionales o modelos cíclicos, son aquellos modelos que se caracterizan por la presencia de oscilaciones cíclicas, también llamados variaciones estacionales.

Las series que presentan tendencias y variaciones cíclicas, se representados a través de los modelos ARIMA (p, d, q) (P, D, Q). Donde (p, d, q) hace referencia a la tendencia secular o parte regular de la serie y el (P, D, Q) es la parte cíclica (presencia de ciclos) de la serie temporal.

Modelización ARIMA (p, d, q)

Cuando se habla de modelización de la metodología ARIMA, básicamente se refiere a la construcción de dicho modelo, esto se realiza de forma iterativa a través de un proceso que consta de cuatro etapas, los cuales son: Identificación, Estimación, Validación y Predicción.

Identificación

En esta etapa se busca el mejor modelo ARIMA (p,d,q) que represente la serie, es decir aquel modelo que reproduzca las características de la serie en estudio. Para esta etapa se divide en dos partes:

1) **Análisis de la estacionariedad**: Aquí se define las transformaciones necesarias que se van a realizar para así obtener una serie estacionaria, y para lograr ello debemos señalar que existe dos apartados:
    - **Estacionariedad en varianza**: Se dice que una serie es estacionaria en varianza, cuando la variabilidad de la serie en estudio en torno a su media, se mantiene constante a través del tiempo. Existen casos en las que las series no son estacionarias, en este tipo de casos se hace uso de las transformaciones de Box-Cox.

$$Y_t^{(\lambda)} = \begin{cases} \dfrac{Y_t^{\lambda} - 1}{\lambda} & \lambda \neq 0 \\ lnY_t & \lambda = 0 \end{cases}$$

      De las transformaciones más usadas en la práctica es la logarítmica, asimismo cuando se analiza la estacionariedad en varianza, se hace uso del grafico de la serie original y el grafico de la serie después de haber realizado las trasformaciones correspondientes.
    - **Estacionariedad en media**: Hablamos de estacionariedad en media, cuando la serie oscila alrededor de un nivel constante o no. Para tomar esta decisión se

analizan las características que diferencian entre una serie estacionaria y las no estacionaria, las cuales se presentan en la Tabla 1.

Tabla 1

Características de una serie estacionaria y no estacionaria

Serie Estacionaria	Serie No Estacionaria
Fluctúan alrededor de una media constante.	Existencia de tendencia o tramos con medias diferentes.
La función de autocorrelación decae exponencialmente.	En los procesos con presencia de alguna raíz unitaria, su función de autocorrelación muestral tiene decaimiento muy lento.

Ahora cuando la serie es no estacionaria en media, se procede aplicar d diferencias de manera sucesiva de orden 1, hasta lograr obtener la estacionariedad propiamente dicha.

$$Z_t = (1 - L)^d Y_t$$

En esta primera fase se busca determinar la cantidad menor posible de diferencias d que se realizarán a fin de lograr la estacionariedad de la serie. En ese sentido a fin de determinar el número de diferencias a usar, se debe tener en cuenta las siguientes herramientas:

- Gráfico original de la serie y de las transformaciones realizadas.
- Correlograma estimado de la serie original y así como también de las transformaciones realizadas.
- Contraste de raíces unitarias.

Los contrastes de raíces unitarias permiten hacer una inferencia de la existencia o no de una raíz unitaria en la serie. Estas se basan en la hipótesis nula de existencia de raíz unitaria en Δd−1 Y_t de rechazarse esta, ya no se aplicará más diferencias a la serie, de lo contrario de no ser rechazado la hipótesis nula se aplicará una diferencia más de orden 1. Dentro de los de raíces unitarias que más se usa tenemos al de, Dickey – Fuller.

2) **Identificación del modelo estacionario:** Luego de haber determinado las diferencias d, se tiene que identificar los órdenes p y q del modelo del proceso que puede reproducir las características de la serie estacionaria.

   Para la identificación de las ordenes p y q, se tienen que comparar las funciones de autocorrelaciones muestrales con las FAC teóricas del modelo ARMA. Hay casos en los cuales el correlograma no se anula, pero sin embargo decrece de forma exponencial, es ahí donde la identificación donde surge la complejidad, para ello se usa la función de autocorrelación parcial (FACP). A modo de resumen en la Tabla 2, se presenta las características de ambos.

Tabla 2

Características de las autocorrelaciones simples y parciales

	AR (p)	**MA (q)**	**ARMA (p,q)**
FAC	Mixtura de exponenciales o sinusoidales que decaen rápidamente.	Picos grandes en los q rezagos o corta abruptamente a partir del rezago q, es decir la FACP se anula luego del rezago q.	Decrecimiento exponencial.
FACP	Picos grandes en los p rezagos o corta abruptamente a partir del rezago p, es decir la FACP se anula luego del rezago p.	Dominada por mixtura de exponenciales o sinusoidales que decaen rápidamente.	Decrecimiento exponencial.

Estimación

El modelo Box-Jenkins no solo tiene que ser estacionario, también tiene que ser invertible. Invertible significa que las observaciones recientes están más ponderadas que las observaciones más remotas; los parámetros $(\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q)$ utilizado en el modelo disminuyen desde las observaciones más recientes hasta las observaciones pasadas posteriores (Bowerman, O'Connell y Koehler).

Luego de haber realizado la identificación del modelo, se tienen que estimar los valores de los parámetros ϕ y θ, a través de la suma de los mínimos cuadrados de los errores. Generalmente este proceso es automatizado por los softwares computacionales. Al respecto se analizan si los coeficientes de autocorrelación se deben excluir del modelo. Se utiliza el error estándar y el valor t asociado a cada coeficiente del modelo.

Los t-valor y p-valor aproximados prueban la siguiente hipótesis. Sea θ cualquier parámetro particular en un modelo de Box-Jenkins.

$$H_0: \theta = 0$$

$$H_1: \theta \neq 0$$

Se rechazara H_0 si y solo si se cumple alguna de las siguientes condiciones.

i. $|t| > t_{\propto/2}^{(n-n_p)}$; donde n_p es la cantidad de parámetros en el modelo

ii. $p - valor < \propto$

Tener en cuenta que, si la hipótesis nula se rechaza, más fuerte será la evidencia que indique que el parámetro es importante en el modelo. También observamos la correlación estimada entre los parámetros, los parámetros siempre están correlacionados, pero correlaciones muy altas superiores a 0,9 sugieren estimaciones de parámetros deficientes. Cuando comparamos modelos, el mejor modelo tiene un error estándar más pequeño, para ello se usan el criterio bayesiano de Schwarz (BIC) y el criterio de información de Akaike (AIC). Sin embargo "una de las mejores medidas para seleccionar el mejor modelo de pronóstico es el MSE o RMSE, pero a pesar su efectividad probada ésta no es muy utilizada entre los pronosticadores" Gálvez (2016).

Validación

En esta etapa se procede analizar la adecuación entre el modelo estimado al conjunto de datos. De acuerdo a Box y Jenkins existe varios test para validar el modelo elegido, entre los cuales tenemos:

- Sobreparametrización, este test se basa en ajustar el modelo de orden superior al elegido, para luego corroborar si son significativamente distintos de cero los parámetros.

- Existen test donde se buscan corroborar que los residuos sean consistentes con el supuesto de ruido blanco, al respecto Box y Pierce plantean lo siguiente.

$$Q = \sum_{k=1}^{m} r_k^2 \qquad \text{siendo } r_k = \frac{\sum_{t=k+1}^{n} a_t a_{t-k}}{\sum_{t=1}^{n} a_t^2}$$

Donde:

a_t: son los residuos estimados

n: es la cantidad de observaciones

De acuerdo a Box y Pierce, donde el supuesto es que el valor de m es grande, demuestran que el estadístico Q tiene una distribución Chi-cuadrado con (m-n-p) grados de libertad, donde se rechaza la H_0: $Los\ residuos\ son\ ruido\ blanco$ para valores de Q muy altos. Asimismo, señalan que el valor de m es arbitrario, pero que sin embargo es conveniente tomar el valor más alto posible.

- De lo dicho por Box y Pierce, existen casos en lo que los valores de m no son muy altos, es ahí donde Ljung y Box plantean un estadístico alternativo.

$$Q' = \frac{n(n+2)\sum_{k=1}^{m} r_k^2}{(n-k)} \approx X^2_{m-p-q}$$

Predicción

En esta última etapa se busca pronosticar los valores futuros de la serie en estudio, donde el objetivo es obtener predicciones optimas de Y_t. En ese sentido para un modelo ARIMA, sea $Y_{T+\ell}$ donde T es la última observación que se dispone y ℓ representa el número de periodos, el predictor óptimo de $Y_{T+\ell}$ viene dado por la esperanza condicional de los datos $Y_t(\ell) = E_T[Y_{T+\ell}]$. De acuerdo a (Gonzales, 2009) basta con escribir el modelo en forma de ecuación y obtener las esperanzas condicionales, teniendo en cuenta que:

$$E_T[Y_{T+j}] = \begin{cases} Y_{T+j} & , \quad j \leq 0 \\ Y_T(j) & , \quad j > 0 \end{cases}$$

$$E_T[a_{T+j}] = \begin{cases} a_{T+j} & , \quad j \leq 0 \\ 0 & , \quad j > 0 \end{cases}$$

Posteriormente los intervalos de confianza vienen dados de la siguiente forma.

$$[Y_t(\ell) \pm N\alpha_{/2}\sqrt{V(e_T(\ell))}]$$

Donde: $V(e_T(\ell)) = \sigma \sum_{j=0}^{\ell-1} \varphi_j^2$, siendo φ_j los pesos del modelo

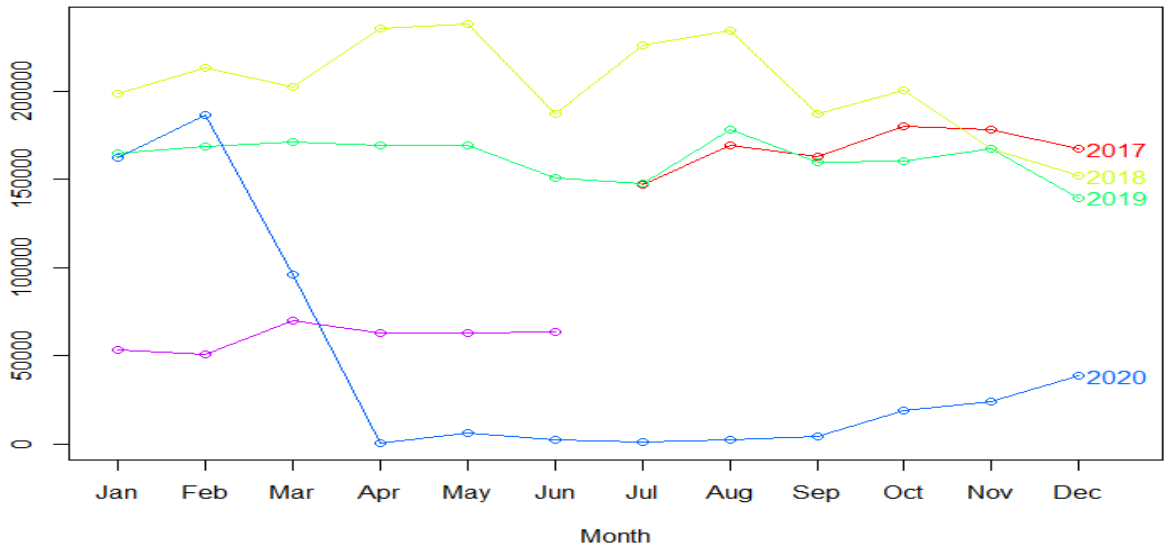
Resultado de la actividad

Mediante la Figura 2, se puede apreciar la evolución histórica de los ingresos que ha tenido mensualmente la entidad para la genérica de ingresos 1.3 por concepto de venta de servicios, derechos administrativos y bienes, como se ha dicho la institución fue creada a mediados del año 2017. De acuerdo a la gráfica se ve una caída bien marcada para el año 2020, esto a raíz de que un 15 de marzo del año 2020, mediante Decreto Supremo N° 044-2020-PCM y sus posteriores ampliaciones, se declara el estado de emergencia nacional a consecuencia del COVID-19.

Al respecto se precisa que la entidad al ser una unidad ejecutora de nivel II-E, no fue considerado un hospital que atienda pacientes con COVID-19, razón por la cual deja de brindar sus atenciones de medicina física y rehabilitación hasta mediados del mes de setiembre, a fin de implementar el protocolo y procedimiento respectivo de acuerdo a las normas técnicas del MINSA. Sin embargo, la entidad durante ese lapso de tiempo mencionado, seguía generando ingresos por concepto de venta de derechos administrativos y de bienes.

Figura 2

Comportamiento de los ingresos mensuales de la genérica 1.3 periodo julio 2017-junio2021



Identificación del modelo

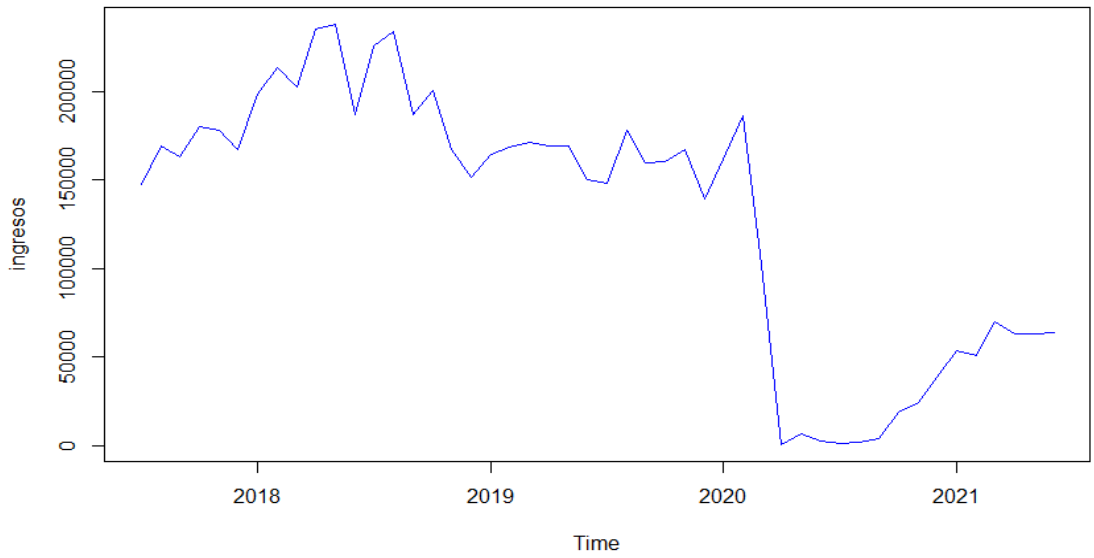
Gráfico de la serie original

De acuerdo a la Figura 3, se muestran los ingresos mensuales de la genérica de ingreso 1.3 “Venta de bienes, servicios y derechos administrativos” desde julio del año 2017 hasta

junio del 2021. Se aprecia que no presenta un comportamiento estacionario, tampoco se puede visualizar una marcada tendencia y la no presencia de ciclos.

Figura 3

Serie de tiempo de los ingresos mensuales de la genérica de ingreso 1.3 periodo julio 2017-junio2021



Verificando la estacionariedad

De acuerdo a lo observado en la Figura 3, se procede a realizar la verificación de la estacionariedad de la serie, para ello se aplica el test de Dickey-Fuller, el cual establece lo siguiente.

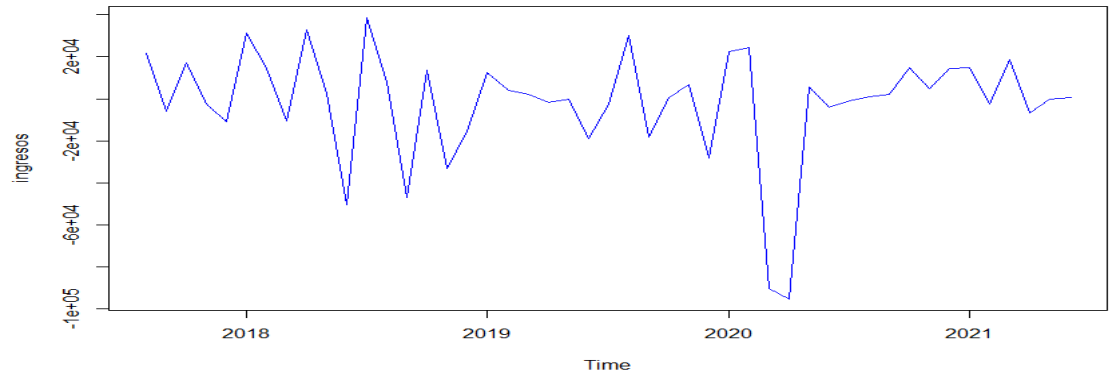
H_0: La serie no es estacionaria (Presencia de raíz unitaria)
H_1: La serie es estacionaria (No existe presencia de raíz unitaria)

Para la prueba de Dickey-Fuller, el criterio de decisión es que el p-valor sea menor a 0.05 para lo cual se tiene que rechazar la hipótesis nula. En ese sentido aplicando el test, en el Software R, sale un p-valor = 0.2858, por lo que la H_0 no es rechazada, lo que significa que la serie no es estacionaria.

A fin de obtener la estacionariedad de la serie, se procede a realizar una diferencia, el cual se puede apreciar en la Figura 4. Luego de haber aplicado el test, nos arroja un p-valor = 0.04727, rechazando asi la hipótesis nula, lo que significa que la serie es estacionaria es primera diferencia.

Figura 4

Serie de tiempo estacionaria de los ingresos mensuales de la genérica de ingreso 1.3 periodo julio 2017-junio2021 luego de aplicar una diferencia

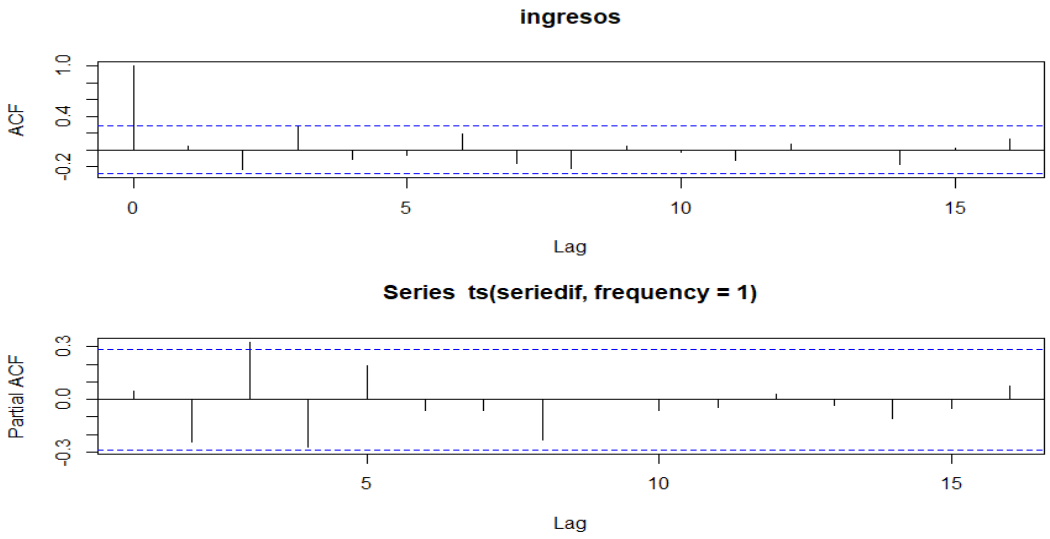


Identificación del modelo estacionario

Habiendo obtenido la estacionariedad de la serie, se procede a identificar el modelo que se va usar, para ello haremos uso de las autocorrelaciones (simples y parciales) de la serie estacionaria. De acuerdo a la Figura 5, se observan rezagos fuera de las bandas de confianza tanto en la parte de los autoregresivos (AR) y las media móviles (MA), los cuales darían indicios de más de un modelo que se puede identificar.

Figura 5

Autocorrelaciones simples y parciales de la serie estacionaria de los ingresos mensuales de la genérica de ingreso 1.3 periodo julio 2017-junio2021



Posteriormente luego de realizar algunas pruebas de acuerdo a lo planteado por la metodología de Box-Jenkins, se identificó 3 posibles modelos ARIMA los cuales son:

- ARIMA (1,1,2)
- ARIMA (2,1,1)
- ARIMA (1,1,1)

Estimación

En base a los posibles modelos identificados, se procedió hallar los estadísticos de error MAE (Error Absoluto Medio), MAPE (Error Porcentual Absoluto Medio) y RMSE (Error Cuadrático Medio), dichos estadísticos permiten determinar el modelo con menor error posible. De acuerdo a la revisión de la literatura el MAPE generalmente presentan sesgos, los cuales, a la hora de usarlo para la selección de algún modelo de pronóstico, no serán tan precisos ya que favorecen que los valores pronosticados estén por debajo de los valores reales. Por lo que se sugiere tomar el RMSE a la hora de seleccionar el posible mejor modelo.

En la Tabla 3, se detallan los estadísticos de los errores de los modelos ARIMA propuestos. Según lo antes señalado se concluye que el mejor modelo es el ARIMA (1,1,2) ya que presenta el menor valor del RMSE.

Tabla 3

Estadísticos de los errores de los posibles modelos identificados de la serie ingresos mensuales de la genérica de ingreso 1.3 periodo julio 2017-junio2021

MODELO	RMSE	MAPE	MAE
ARIMA (1,1,2)	22409.6	163.5	16079.4
ARIMA (2,1,1)	22482.3	220.6	16723.1
ARIMA (1,1,1)	22946.5	142.5	16228.9

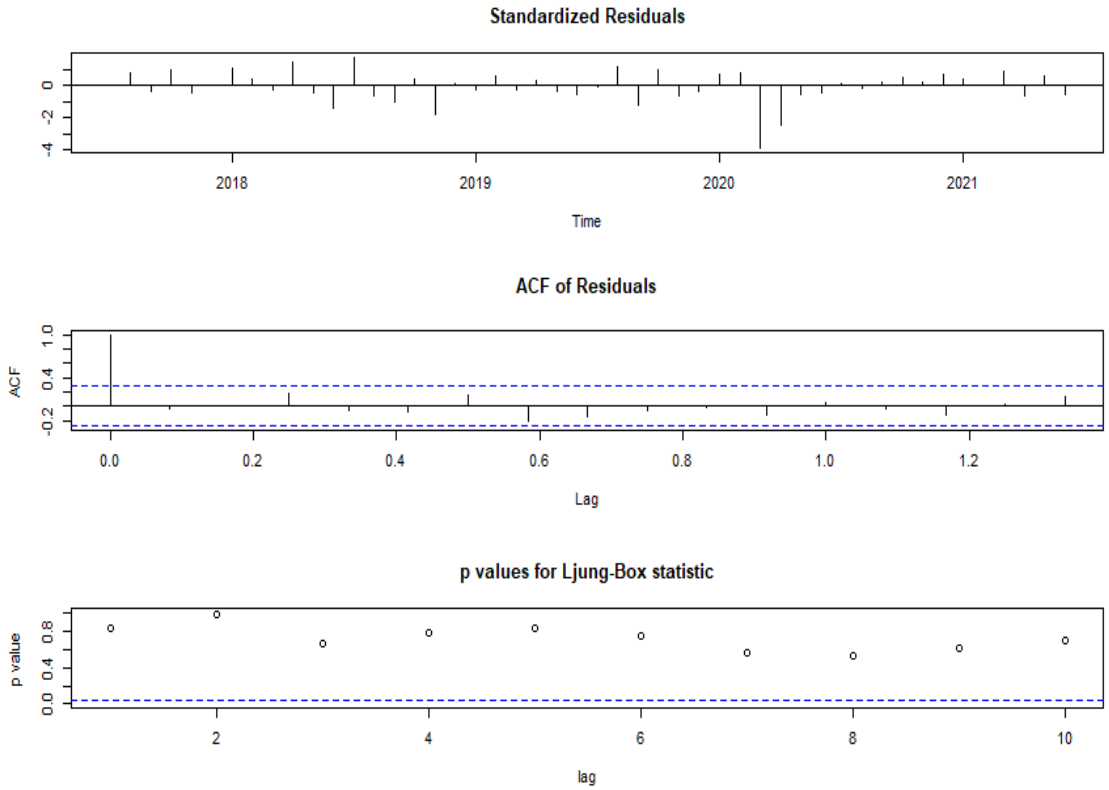
Validación

Luego de haber realizado la estimación de los coeficientes del modelo ARIMA propuesto, se debe realizar la validación, es decir comprobar la eficiencia del modelo. Gráficamente podemos hacer un diagnóstico para comprobar si el modelo elegido es bueno.

En la Figura 6 se puede apreciar los errores estandarizados, los cuales tienen un comportamiento de ruido blanco, asimismo las autocorrelaciones están dentro de las bandas de confianza y finalmente podemos ver los valores del p-valor del estadístico Ljung-Box, el cual identifica la existencia o no del ruido blanco. De acuerdo a lo dicho, los residuales tienden a un ruido blanco, es decir el modelo se ajusta bien.

Figura 6

Diagnóstico gráfico del modelo propuesto ARIMA (1,1,2) de la serie ingresos mensuales de la genérica de ingreso 1.3 periodo julio 2017-junio2021



Lo dicho anteriormente se debe probar mediante el test de Ljung-Box, dicho test sirve para verificar la existencia de ruido blanco o no en la serie, entonces para ello se establece la siguiente prueba de hipótesis.

$$H_0 = \varepsilon_t \sim RB(0, \sigma^2)$$
$$H_1 = \varepsilon_t \, no \, son \, ruido \, blanco$$

Se define el criterio de decisión de la siguiente manera: si el p-valor < 0.05, se tiene que rechazar la hipótesis nula, en caso de ser mayor a 0.05 no se va rechazar la hipótesis nula. Entonces luego de aplicar el test al modelo ARIMA (1,1,2) propuesto, tenemos que el p-valor=0.82 lo que significa que existe ruido blanco, lo que conlleva a decir que el modelo se ajusta adecuadamente y se valida.

Pronostico

Luego de haber sido comprobado la validación de los errores, se comprueba que el modelo estimado ARIMA (1,1,2), presenta un valor pronosticado adecuado, teniendo los siguientes resultados, los cuales son proporcionados en la Tabla 4 y que gráficamente se puede visualizar en la Figura 7.

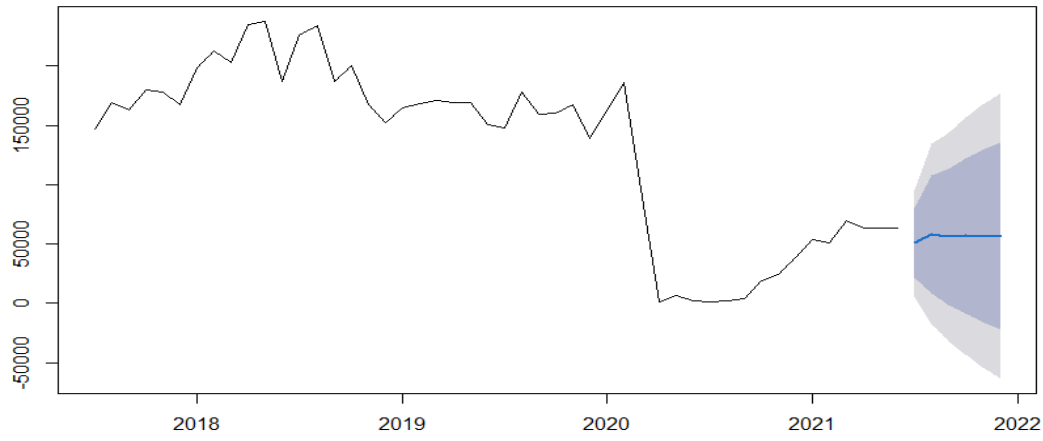
Tabla 4

Pronóstico de la genérica de ingreso 1.3 para el periodo de julio a diciembre correspondiente al año 2021

Periodo	Pronostico ARIMA (1,1,2)
Jul-21	50,793.52
Ago-21	58,502.54
Set-21	56,158.09
Oct-21	56,871.08
Nov-21	56,654.25
Dic-21	56,720.19

Figura 7

Pronóstico de la genérica de ingreso 1.3 para el periodo de julio a diciembre correspondiente al año 2021



Capítulo IV. Conclusiones

En este Trabajo de Suficiencia Profesional, se aplicó la metodología de Box-Jenkins para el periodo de julio 2017 a junio 2021, el cual permitió analizar y pronosticar los ingresos mensuales que ha presentado y presentará la Entidad, en la genérica de ingresos 1.3 por concepto de venta de servicios, derechos administrativos y bienes, por lo que los datos estadísticos evidenciados, otorgaron ciertos atributos que permitieron detectar comportamientos específicos que ha tenido la Entidad en sus ingresos mensuales.

Se evidenciaron factores que incidieron en los ingresos mensuales de la entidad, los cuales fueron a raíz del estado de emergencia sanitaria decretado por el gobierno del estado peruano, a consecuencias del COVID-19. Donde los ingresos registrados para el año 2020 fueron los más afectados para la entidad, ya que se evidenciaron montos muy por debajo de lo habitual.

El pronóstico realizado para el periodo de julio a diciembre correspondiente al año 2021, mediante el modelo ARIMA (1,1,2), evidencian un aumento en comparación con el año 2020, esto se explica que, a través del gobierno del estado peruano, se han ido dictaminando medidas en cuanto al protocolo sanitario, así como también poco a poco se han ido permitiendo las aperturas graduales de los distintos sectores y/o carteras de servicios a nivel nacional. Permitiendo ello una mayor afluencia y concienciación de la salud de la población peruana.

Capítulo V. Recomendaciones

Hacer de conocimiento a la alta dirección de la institución, a fin de que utilicen la metodología de Box-Jenkins, para un adecuado pronóstico de los ingresos mensuales que recauda por los conceptos de venta de servicios, derechos administrativos y bienes.

De acuerdo a los buenos resultados obtenidos por la metodología de Box-Jenkins, se recomienda la aplicación de la misma, como técnica de predicción para la ejecución del presupuesto asignado a la institución.

Utilizar otras técnicas o metodologías de pronóstico a futuro, sobre todo las relacionadas a la inteligencia artificial, a fin de comparar con la técnica utilizada en el presente trabajo.

Capítulo VI. Bibliografía

Abusada Salah, R., Cusato Novelli, A. y Pastor Vargas, C. (2008). *Eficiencia del Gasto en el Perú. Lima-Perú:* Instituto Peruano de Economía.

Avishek Pal y PKS Prakash (2017). *Practical Time Series Analysis*. Packt

Bedoya Luza, S.L. (2018) *Modelamiento univariado del número de defunciones infantiles producidas por infecciones respiratorias agudas, a través de la metodología de Box-Jenkins, Puno 2008-2016* [Tesis de Pregrado, Universidad Nacional del Altiplano]. http://repositorio.unap.edu.pe/handle/UNAP/7788

Bedoya Micolta, F. y Romero Johana, L. (2019). *Pronostico del precio del café colombiano para enero del 2018*[Trabajo de Pregrado, Universidad de Santiago de Cali]. https://repository.usc.edu.co/handle/20.500.12421/2064

Corman Trujillo, J. (2020). *Comportamiento del gasto público de una entidad pública usando el modelo de Box-Jenkins 2009-2017*[Tesis de Pregrado, Universidad Nacional Agraria La Molina]. http://repositorio.lamolina.edu.pe/handle/UNALM/4634

De la Fuente Fernández, S. (2018). Series Temporales: Modelos ARIMA [Archivo PDF]. https://www.estadistica.net/ECONOMETRIA/SERIES-TEMPORALES/modelo-arima.pdf

Gálvez, T. [Tomas Gálvez]. (09 de agosto de 2016). *¿Cómo medir la precisión de los pronósticos?* LinkedIn. https://www.linkedin.com/in/tomas-galvez/

Gonzales Casimiro, M.P. (2009). *Análisis de series temporales: Modelos ARIMA. Universidad del País Vasco.*

INEI (2011). Documento metodológico: Análisis de series cronológicas, aplicadas a series económicas.https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib0985/libro.pdf

Lemus Polania, D.F., Torres Romero A., Cubillos Fonseca J.S. y Camelo Sanchez (2017). *Análisis de la ejecución presupuestal en Colombia durante el periodo 1954-2013*. Revista CIFE, 19(30), 23-45.

López Jiménez, K.J. y Villanueva Vásquez, W.J. (2019) *Modelos Arima univariantes de series temporales para la producción y demanda de agua en el distrito de Lambayeque, periodo*

2002-2017 [Tesis de pregrado, Universidad Nacional Pedro Ruiz Gallo]. https://hdl.handle.net/20.500.12893/8216

Mota López, A. (2016) *Pronostico del precio del crudo de exportación mexicano con la metodología de Box-Jenkins para series de tiempo* [Tesis de pregrado, Benemérita Universidad Autónoma de Puebla]. https://hdl.handle.net/20.500.12371/12812

Orosco Gavilán, J.C. (2017). *Metodología Bootstrap en el análisis serial de los ingresos mensuales. Hotel el Centro, Ayacucho 2011-2014*[Tesis de Pregrado, Universidad Nacional San Cristóbal de Huamanga]. http://repositorio.unsch.edu.pe/handle/UNSCH/1776

Palma Wilfredo (2016). *Time Series Analysis*. Pontificia Universidad Católica de Chile

Quispe Carlos y Purca Sara (2007). Previsión de la temperatura superficial del mar frente a la costa peruana mediante un modelo autoregresivo integrado de media móvil. *Revista Peruana de Bilogía*, 14(1), 110.

Ramos Inga, A.P. (2020). *Pronóstico de los ingresos tributarios mensuales del gobierno central peruano aplicando redes neuronales y modelos SARIMA, en base a los años 2013-2018*[Tesis de Pregrado, Universidad Nacional de Piura]. http://repositorio.unp.edu.pe/handle/20.500.12676/2264

Salazar Ojeda, M., y Saavedra Cuba, G. (2014). *Pronóstico de la Ejecución del Presupuesto Asignado al Gobierno Regional de Piura en la Unid Ejecutora 001 para el periodo 2013 y 2014 Aplicando el modelo Holt-Winter y Modelo ARIMA y SARIMA* [Tesis de Pregrado, Universidad Nacional de Piura]. http://repositorio.unp.edu.pe/handle/UNP/743

Theresa Hoang D.N. (2013). The Box-Jenkins Methodology for Time Series Models. Entertainment Group, Burbank C.A.

Capítulo VII. Anexos

Anexo 1: Ingreso mensual de la genérica 1.3 de la entidad periodo julio 2017 – junio 2021

AÑO / MES	2017	2018	2019	2020	2021
ENERO	-	198,581	164,604	162,000	53,462
FEBRERO	-	213,144	168,754	186,344	51,169
MARZO	-	202,546	171,069	96,024	69,855
ABRIL	-	235,199	169,298	904	63,193
MAYO	-	237,818	169,226	6,598	63,048
JUNIO	-	187,259	150,460	2,438	63,599
JULIO	147,223	225,793	147,884	1,472	-
AGOSTO	168,964	233,843	178,060	2,311	-
SETIEMBRE	163,140	186,892	159,853	4,336	-
OCTUBRE	180,211	200,548	160,571	19,289	-
NOVIEMBRE	177,919	167,556	167,423	24,121	-
DICIEMBRE	167,128	151,911	139,390	38,647	-