

Articulo - EDWARD ALEGRIA.docx

 Universidad Peruana Union

Detalles del documento

Identificador de la entrega

trn:oid:::29566:525952741

Fecha de entrega

10 nov 2025, 2:20 p.m. GMT-5

Fecha de descarga

10 nov 2025, 2:36 p.m. GMT-5

Nombre del archivo

Articulo - EDWARD ALEGRIA.docx

Tamaño del archivo

427.6 KB

21 páginas

5905 palabras

33.707 caracteres

14% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...

Filtrado desde el informe

- Bibliografía
- Texto citado

Fuentes principales

- 13%  Fuentes de Internet
- 4%  Publicaciones
- 5%  Trabajos entregados (trabajos del estudiante)

Marcas de integridad

N.º de alertas de integridad para revisión

No se han detectado manipulaciones de texto sospechosas.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.

Fuentes principales

- 13% Fuentes de Internet
- 4% Publicaciones
- 5% Trabajos entregados (trabajos del estudiante)

Fuentes principales

Las fuentes con el mayor número de coincidencias dentro de la entrega. Las fuentes superpuestas no se mostrarán.

1	Internet	docplayer.es	1%
2	Internet	revibiomedica.sld.cu	1%
3	Internet	repositorio.uchile.cl	1%
4	Internet	pt.scribd.com	1%
5	Internet	www.diresahuanuco.gob.pe	<1%
6	Internet	repositorio.unbosque.edu.co	<1%
7	Internet	hdl.handle.net	<1%
8	Internet	www.coursehero.com	<1%
9	Internet	journal.poligran.edu.co	<1%
10	Internet	www.scribd.com	<1%
11	Internet	tools4adhd.eu	<1%

12	Internet	dataonline.gacetajuridica.com.pe	<1%
13	Internet	pesquisa.bvsalud.org	<1%
14	Internet	dokumen.pub	<1%
15	Internet	ruidera.uclm.es	<1%
16	Trabajos entregados	National University College - Online on 2019-05-27	<1%
17	Internet	es.slideshare.net	<1%
18	Internet	www.saber.ula.ve	<1%
19	Trabajos entregados	Bob Jones University on 2025-05-02	<1%
20	Trabajos entregados	Universidad de Montemorelos on 2025-04-07	<1%
21	Internet	bibliotecadigital.udea.edu.co	<1%
22	Trabajos entregados	Pontificia Universidad Catolica del Peru on 2007-04-20	<1%
23	Internet	patents.google.com	<1%
24	Internet	search.bvsalud.org	<1%
25	Internet	www.edx.org	<1%

26	Internet	www.grafiati.com	<1%
27	Internet	www.idreh.gob.pe	<1%
28	Internet	www.yumpu.com	<1%
29	Trabajos entregados	Universidad Nacional Daniel Alcides Carrion on 2025-10-29	<1%
30	Internet	core.ac.uk	<1%
31	Internet	datosabiertos.gob.pe	<1%
32	Internet	link.uaautonoma.cl	<1%
33	Internet	www.researchgate.net	<1%
34	Trabajos entregados	Universidad Carlos III de Madrid - EUR on 2025-06-16	<1%
35	Publicación	Gamboa Domínguez, Luis Elí. "Conocimiento sobre salud bucal y prevalencia de e...	<1%

Título: Análisis de patrones estacionarios de atenciones odontológicas mediante el modelo de series temporales en establecimientos de salud de la región Huánuco

Edward Alegria y Javier Linkolk López-Gonzales

Resumen

Objetivo: Se analizaron diferentes patrones mediante un modelo de series temporales de atenciones odontológicas realizadas en establecimientos de salud públicos de la región Huánuco – Perú durante el período 2018-2023. Metodología: La fuente de información se obtuvo del portal web Plataforma Nacional de Datos Abiertos del Seguro Integral de Salud (SIS). Se preparó un dataset con ayuda de software de inteligencia de negocios Power Bi ® para estructurar las variables de estudio considerándose sólo atenciones odontológicas, donde se aplicaron análisis estadísticos descriptivos y de pronóstico usando software R Studio© mediante librerías aplicadas a base de datos, gráficas y series de tiempo. Culminada la transformación se compararon dos técnicas para el análisis de pronósticos Holt Winters y ARIMA. Resultados: El análisis de autocorrelaciones reveló no estacionariedad por lo que requirió conversión hacia una serie estacionaria. En la evaluación de supuestos del modelo HW obtuvo a estadístico de Breush Pagan ($p = 0.0002$), Ljung-Box ($p = 0.0000$) y Prueba de kolgomorov Smirnov ($p = 0.0012$) todos valores por debajo del nivel de significancia ($p < 0.05$). Para el modelo ARIMA con nivel de significancia ($p = 0.05$) se comprobó valores superiores determinándose homocedasticidad residual ($p = 0.5073$) e independencia de residuos a la prueba de Ljung-Box ($p = 0.78$), por otro lado la prueba de normalidad KS obtuvo un valor $p = 0.0513$ ligeramente superior a nivel de significancia. Los resultados de pronósticos se observaron a horizontes de 12 meses ($h = 12$) con intervalos de confianza 80% y 95% determinándose en modelo HW valores con ciertos límites inferiores negativos lo que demuestra inestabilidad del modelo, frente a ARIMA que presenta mayor estabilidad. Conclusión: el modelo ARIMA presentó mejor desempeño en sus pronósticos frente al modelo de Holt Winters, luego de las transformaciones por la no estacionariedad de los datos considerando un mejor ajuste para las predicciones.

Palabras clave: Odontología en salud pública, estudios de series temporales, predicción, interpretación estadística de datos, ciencia de los datos

Keywords: Public health Dentistry, time series studies, forecasting, data interpretation statistical, data science

Palavras-chave: Odontologia em saúde pública, estudos de séries temporais, previsões, interpretação estatística de dados, ciência de dados

Introducción

Una serie de tiempo es un conjunto de observaciones determinada por valores que toma una variable cuantitativa en diferentes momentos del tiempo. Estas observaciones son registradas con valores numéricos que varían en el tiempo y se puede recolectar en diferentes períodos (días, semanas, meses, trimestres, años, etc.), por lo que un aspecto básico del estudio de las series de tiempo requiere analizar la naturaleza de las variaciones [1].

De acuerdo a Ríos (2008) una serie de tiempo puede descomponerse en cuatro componentes (y en ocasiones un quinto como constante llamada nivel) que no son directamente observables, de los cuales únicamente se pueden obtener estimaciones. Estos cuatro componentes son: Tendencia(T), Ciclo(C), Estacionalidad(E) y Aleatoriedad(A). Los análisis de series de tiempo tienen un enfoque netamente predictivo por lo que la selección de datos para la muestra es una decisión trascendental en la generación de pronósticos ya que determina la cantidad de datos para la construcción del modelo y la cantidad de pronósticos distintos a la muestra que se podrán evaluar. La definición del tamaño y composición de los conjuntos de entrenamiento y evaluación deberá considerar factores tales como el tamaño total de la muestra, el tipo de metodología para el pronóstico a emplear y finalmente la representatividad [2,3], el pronóstico en mención que permite determinar valores futuros de series temporales supone además que existe un modelo de probabilidad que genera dichas observaciones.

Actualmente el uso de series de tiempo en el campo de la investigación destaca por el impacto en diferentes campos de aplicación como pronosticar una demanda [4], monitoreo y mantenimiento de alguna condición de un proceso [5], el uso eficiente de recursos [6] y otros. A raíz del evento pandémico la condición médica en los países del mundo es parte de diferentes análisis en función a estimaciones sobre el comportamiento futuro de las enfermedades ayudando a predecir la propagación o el incremento de las mismas [7, 8] contribuir a gestionar los gastos médicos y predecir problemas de salud a futuro [9]. Cada estimación es crucial, especialmente al considerar las mejores técnicas de pronóstico que faciliten a los gobiernos la planificación y el uso adecuado de recursos, lo que puede ayudar directamente a la protección de la vida humana en materia de salud [10].

Los modelos de series de tiempo proporcionan pronósticos o estimaciones y este valor previsto se compara con un valor observado. Las técnicas de pronóstico de series de tiempo tradicionales basadas en modelos probabilísticos estadísticos han demostrado una eficiencia notable, sin embargo, debido al rápido incremento de volúmenes de datos generados en tiempo real como la corriente del Big Data han creado un nuevo panorama. Estos datos de alto volumen y alta velocidad no son lineales y tienen patrones diversos [11], para ello requieren algoritmos de naturaleza computacional compleja y que finalmente debe integrarse a un sistema existente lo que resulta cada vez más en proceso costosos y uso intensivo de recursos. Los datos de series temporales se recopilan de forma incremental a diario desde dispositivos computacionales portátiles para diversas aplicaciones en atención médica, sin embargo, generalmente no tienen patrones reconocibles por el ser humano y requieren especialistas para su anotación/etiquetado. Por lo tanto, es mucho más difícil etiquetar datos de series temporales que imágenes, y se han etiquetado pocos datos de series temporales en aplicaciones del mundo real [12].

La relación entre políticas sanitarias y el tratamiento de enfermedades también son determinadas en algunas situaciones mediante el uso de series temporales interrumpidas tal como lo plantearon en China mediante un modelo predictivo donde se muestra que casos notificados de VIH/SIDA tenían una relación inversa con una correlación significativa en políticas de tratamiento para COVID-19 [13], en los Estados Unidos de Norteamérica con la construcción de modelos para predecir riesgos de hospitalización en atención médica domiciliaria con valores precisión mayores a 0.8 [14], y en otros países del mundo donde identificaron valores atípicos positivos al evaluar la salud y bienestar de las mujeres mediante un estudio de series de tiempo [15].

Desde la aparición de la COVID-19 la predicción de aparición de casos nuevos fue revisada mediante el uso de series temporales en una comunidad rural del Ecuador revelando resultados una R cuadrado estacionaria de -2,190 y un R cuadrado de 0,658, lo que sugirió una adecuada capacidad del modelo para capturar la tendencia. El Error Cuadrático Medio (RMSE) fue de 8,271, indicando una precisión razonable de las predicciones, mientras que el Error Porcentual Absoluto Medio (MAPE) fue de 18,970, representando una precisión relativa aceptable para dicho escenario [16].

En un estudio similar realizado en el mismo país por Monteiro et al. (2019) describieron las características epidemiológicas y las tendencias temporales de los indicadores de lepra en niños menores de 15 años en Tocantins entre los años 2001 y 2012, mediante usos de formularios de

13 información sociodemográfica y clínica, encontraron que la detección en menores de 15 años aumentó significativamente entre 2001 y 2008 (cambio porcentual anual: APC = 3,8 %; intervalo de confianza del 95 %: IC del 95 %: 0,1 – 7,6) y mostró una disminución significativa entre 2008 y 2012 (APC = -9,4 % ; IC95%: -17,2 – -0,8). La estabilidad de los indicadores apunta a la persistencia de diagnósticos tardíos y demandas reprimidas [17].

Las atenciones de salud a nivel de establecimientos de salud públicos requieren sistemas de información basados en el monitoreo y seguimiento continuo de toda aquella información sanitaria recolectada en determinados períodos de tiempo, y que se registra mediante el uso de formatos físicos o digitales que la autoridad sanitaria del país dispone para dichos fines. Un insumo básico para la programación de los recursos financieros según la demanda de atenciones es aquella atención brindada a cada usuario de salud según la naturaleza de su motivo de consulta, enfermedad o diagnóstico establecido por el profesional de salud, y que como ya se mencionó se almacena en una base de datos dándole una connotación de tiempo.

Para Garay et al., (2011) lo mencionado en el párrafo anterior y los resultados sanitarios en un sistema de salud tiene relación con la manera de adoptar y difundir intervenciones en programas de cobertura específica o sistemas de salud que buscan calcular resultados financieros mediante un análisis de impacto presupuestario [18]. En este mismo contexto la estructura de un sistema de financiamiento de salud es determinante si se pretende ejecutar las intervenciones que se plantean en las políticas sanitarias, las cuales no siempre generan resultados esperados y pueden variar entre diferentes escenarios de cada entorno donde se apliquen [19], esto indica que los aspectos fundamentales del sistema de financiamiento, que incluyen los acuerdos para aumentar, agrupar y adquirir ingresos, pueden afectar la protección financiera, la cual se considera un objetivo esencial del sistema de salud en distintos países [20]

29 El intercambio de la información, así como los sistemas de información, que no tuvieron
25 atención durante decenios actualmente son determinantes en la toma de decisiones para una
22 gestión sanitaria ya que se brinda mejor acceso a la atención de salud y se garantiza la
21 continuidad de la misma, enfocándose de otro modo la concepción de una prestación de
servicios de salud [21].

7 En el Perú la condición de la salud bucal tiene como referencia que la caries dental junto a la
35 enfermedad periodontal y maloclusiones presentan una alta prevalencia de enfermedades de la
4 cavidad oral lo que la constituye como uno de los principales problemas de salud pública en
nuestro país. El último estudio nacional sobre el Perfil epidemiológico realizado entre el año

7 2012 a 2014 refiere que la prevalencia global de caries dental fue de 85.6% siendo más prevalente entre la población infantil a partir de los 5 años de edad, es por ello que está considerada como una de las principales causas de consulta en los establecimientos de salud públicos del país [22].

1 La caries dental es un proceso o enfermedad dinámica crónica que ocurre en la estructura dentaria en contacto con los depósitos microbianos que, debido al desequilibrio entre la sustancia dental y el fluido de la placa circundante, trae como resultado una pérdida del mineral de la superficie dental, cuyo signo clínico es la destrucción localizada de tejidos duros. La probabilidad de desarrollarla y la severidad de las lesiones difieren entre las comunidades [23].

5 De acuerdo a la información de la Dirección Regional de Salud Huánuco (2023) el comportamiento de la morbilidad general presentaba a las enfermedades de la cavidad bucal en el segundo lugar a nivel de la región Huánuco inmediatamente por debajo de las enfermedades respiratorias [24]. Sin embargo, las dotaciones de insumos no son suficientes para coberturar completamente una atención odontológica adecuada que permita reducir la prevalencia de la enfermedad en la región. La disponibilidad de insumos odontológicos no se ha determinado mediante información estadística ni se ha relacionado a la asignación de recursos financieros como lo es el costo per cápita para un usuario que acude a un servicio de Odontología.

27 La información de la autoridad sanitaria regional sobre la salud bucal presentada en indicadores de desempeño y de resultado anuales únicamente brinda una mirada transversal de la situación sanitaria más no considera modelos predictivos sobre el comportamiento de la enfermedad de la caries dental, así como también no predice aquellas necesidades futuras de la atención de salud en los servicios de Odontología. Incorporar una estrategia técnica con aplicación estadística en salud pública puede contribuir a mejorar la gestión sanitaria a través del gasto en un sistema de salud [25], lográndose a nivel local no sólo la incorporación de mayores recursos financieros, sino además un análisis de factores del sistema de salud que permitan incrementar la mayor participación de los dentistas generales en servicios de atención dental primaria [26].

6 El análisis de modelos predictivos de las principales enfermedades mediante el uso de series temporales en nuestro país aún se vienen trabajando mediante el desarrollo del gobierno digital o la política de acceso a datos abiertos, por ello como refiere Montenegro [27] en el año dos mil once explicar por qué aun con los avances tecnológicos, metodológicos y conceptuales de la odontología y sus especializaciones, las enfermedades de la boca siguen siendo unos de los motivos de consulta más apremiantes en las instituciones prestadoras de salud, por lo que es

necesario comenzar a mirar el sistema estomatognático desde una variable de tiempo basado en análisis de sistemas de información con el uso de modelos de series temporales de patrones estacionarios de las atenciones odontológicas.

Métodos

Diseño de investigación

La finalidad del presente estudio pretende aportar nuevo panorama sobre atenciones odontológicas en el sistema de salud público peruano mediante el uso de modelos de series de tiempo.

El presente estudio se plantearía como observacional tipo descriptivo [28] debido a que dentro de sus utilidades el uso de series temporales podría describir algún tipo de comportamiento de las enfermedades. La información ha sido recolectada por el investigador sin intervención en el desarrollo de los acontecimientos o alguna forma de control directo sobre las variables parte del estudio, propio de estudios de diseño no experimental [29] puesto que no hay influencia sobre las mismas en el sentido que los hechos ya sucedieron al igual que sus efectos [30].

Finalmente, relacionado al uso de series de tiempo el diseño será del tipo longitudinal retrospectivo para observar la estacionalidad y tendencia de los datos de atenciones odontológicas registradas en el período 2018-2023.

Selección de la base de datos

Para este estudio se utilizó la base de datos del portal web Plataforma Nacional de Datos Abiertos cuya última actualización corresponde al mes de Agosto del año 2024, la cual contiene todas las atenciones realizadas a los afiliados al Seguro Integral de Salud (SIS) de Perú durante los años 2018 al 2023 [31]. Los datos de las atenciones a los asegurados al SIS se recolectan de manera semestral y son comprimidos para su disponibilidad de descarga en la plataforma web. Se preseleccionó a la región Huánuco que es una de las veinticinco regiones de país (véase figura 01), la cual se sitúa en la parte central del Perú entre la cordillera occidental y la cuenca hidrográfica del río Pachitea, afluente del río Ucayali, según su latitud se ubica entre los 9°55'50", latitud sur y 76°14'32 y 77°99', contando con una población estimada para el año 2022 de 755,280 personas de acuerdo a los datos de análisis situacional de salud brindado de la autoridad sanitaria regional [32].

Figura 01. Ubicación de la región Huánuco-Perú.

de servicio (*cod_serv*), descripción del servicio(*desc_serv*), provincia(*prov*), descripción unidad ejecutora (*desc_UE*) y finalmente nombre del establecimiento de salud donde se realizó la atención (*nomb_estab*)

Análisis estadísticos

La información obtenida en el dataset será analizada utilizando software R a través de su plataforma RStudio© versión 2024.12.1+563 liberado el 13/02/2025. La primera fase consistirá en realizar un Análisis Exploratorio de Datos (EDA) identificado las principales categorías o variables del estudio, mediante estadística descriptiva apoyándonos en la construcción de tablas de frecuencias y resumen de los valores en la serie de tiempo. Para los pronósticos se realizó la comprobación de los supuestos con los estadísticos de Breusch Pagan (BP), Ljung Box test (LB) y Kolgomorov-Smirnov(KS). De acuerdo al resultado de la pruebas en el dataset creado se digitaron scripts utilizando diferentes librerías y funciones en RStudio© relacionadas a base de datos (*readxl dplyr, lubridate*), gráficos (*ggplot2*) y de series de tiempo (*lmtest, forecast, fpp, fpp2, fpp3*).

En base a los resultados el análisis de patrones estacionarios será evaluado por todos los componentes de manera conjunta, así como mediante descomposición de la serie de tiempo, transformaciones logarítmicas, diferenciación regular y/o estacional, de tal forma que se propone el uso de dos métodos de pronóstico: Holt-Winters (HW) y el Método de Medias Móviles Integrada Autorregresivo o del nombre en inglés *Autoregressive Integrated Moving Average* (ARIMA), de tal manera que se busque valores predictivos que sugieran el mejor rendimiento de uno de los modelos elegidos para este estudio.

Los resultados serán exportados hacia tablas y gráficos para la visualización de los componentes de tendencia, cíclicos y estacionales.

Resultados

La Tabla 1 muestra el número de atenciones odontológicas en la región Huánuco realizados durante el período 2018 al 2023 según el código de servicio, representada en su distribución de frecuencias. Para el código de servicio 911 éste fue incorporado por el Seguro Integral de Salud a partir del año 2021 en adelante.

Tabla 01. Cantidad de atenciones por código servicio odontológico (cod_serv) región Huánuco 2018-2023.

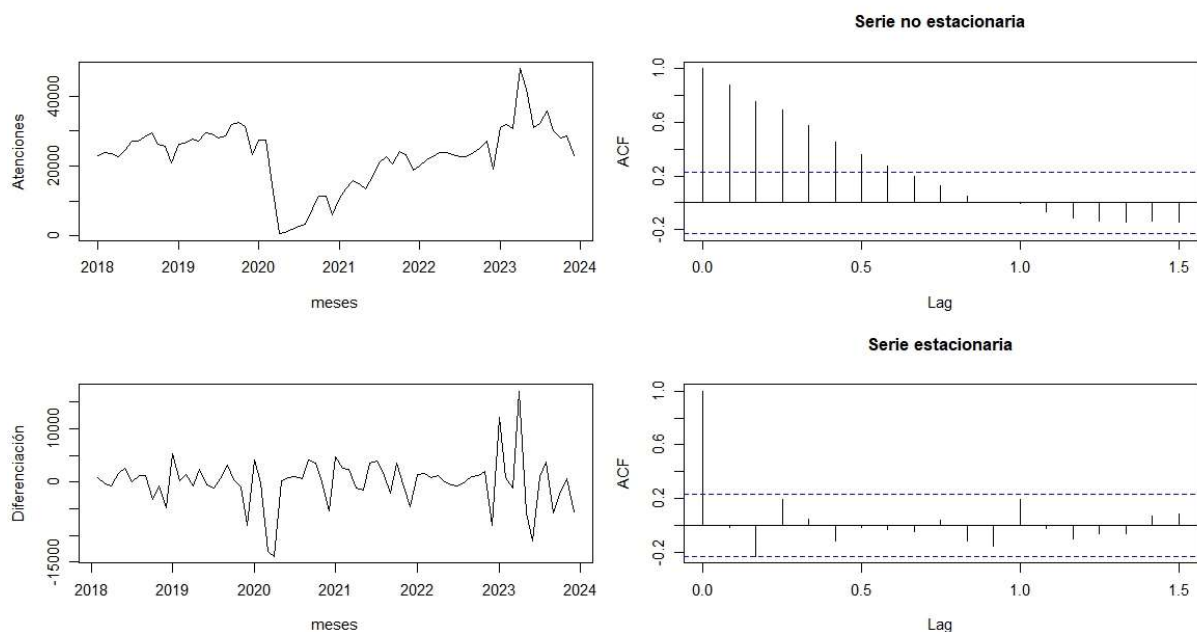
cod_serv	Año												Subtotal código servicio	%
	2018	%	2019	%	2020	%	2021	%	2022	%	2023	%		
20	90963	5.5%	98841	6.0%	40954	2.5%	57339	3.5%	72607	4.4%	101722	6.2%	462426	28.1%
21	118670	7.2%	137611	8.4%	40511	2.5%	67073	4.1%	82520	5.0%	112033	6.8%	558418	34.0%
57	37390	2.3%	39275	2.4%	10136	0.6%	15285	0.9%	22341	1.4%	38642	2.4%	163069	9.9%
58	27068	1.6%	34668	2.1%	10489	0.6%	12016	0.7%	18669	1.1%	33101	2.0%	136011	8.3%
59	26911	1.6%	29188	1.8%	12968	0.8%	19238	1.2%	24339	1.5%	31237	1.9%	143881	8.8%
70	1635	0.1%	1939	0.1%	570	0.0%	524	0.0%	942	0.1%	3230	0.2%	8840	0.5%
911(*)		0.0%		0.0%		0.0%	44135	2.7%	54553	3.3%	72019	4.4%	170707	10.4%
Total	302637	18.4%	341522	20.8%	115628	7%	215610	13.1%	275971	17%	391984	23.9%	1643352	100%

Durante el período de estudio 2018-2023 a nivel de la región Huánuco los códigos de servicio con mayor distribución corresponden al 20 y 21 (34.0%) del tipo de atenciones preventivas, y el menor al código del servicio 70 (0.5%) que corresponde al tipo especializado. Los otros códigos de servicio guardan correspondencia con atenciones del tipo recuperativo.

El comportamiento de la serie de tiempo del número de atenciones odontológicas durante el período de estudio se muestra en la figura 02. Las imágenes superiores representan la serie original y su función de autocorrelación(ACF) identificando tendencia positiva con presencia de estacionalidad observando además un incremento de atenciones en los meses de medio año para llegar a un descenso en el último mes de cada año. Se denota un valle a inicios del año 2020 que refleja un descenso bastante significativo del número de atenciones relacionados con el inicio de la pandemia en nuestro país lo que marco la restricción de atenciones odontológicas dispuesto por la autoridad sanitaria regional debido al alto riesgo de exposición por aerosoles. El análisis ACF de la serie original indica que es una serie del tipo no estacionaria por la caída constante de cada observación, siendo la varianza no constante por lo que no es adecuada como serie capaz de predecir comportamientos, por ello fue necesario realizar una descomposición de los componentes de la serie de tiempo y aplicar la diferenciación en busca de estacionariedad. Se requirió una diferenciación de la serie mediante la función *ndiffs* del paquete *forecast*

mostrándose en las imágenes inferiores de la misma figura la transformación de la serie tipo no estacionaria hacia una serie estacionaria sin el componente de tendencia y según el nuevo análisis ACF podría permitir un mejor tratamiento del modelo para realizar predicciones.

Figura 02. Patrones de serie de tiempo de atenciones odontológicas región Huánuco 2018-2023

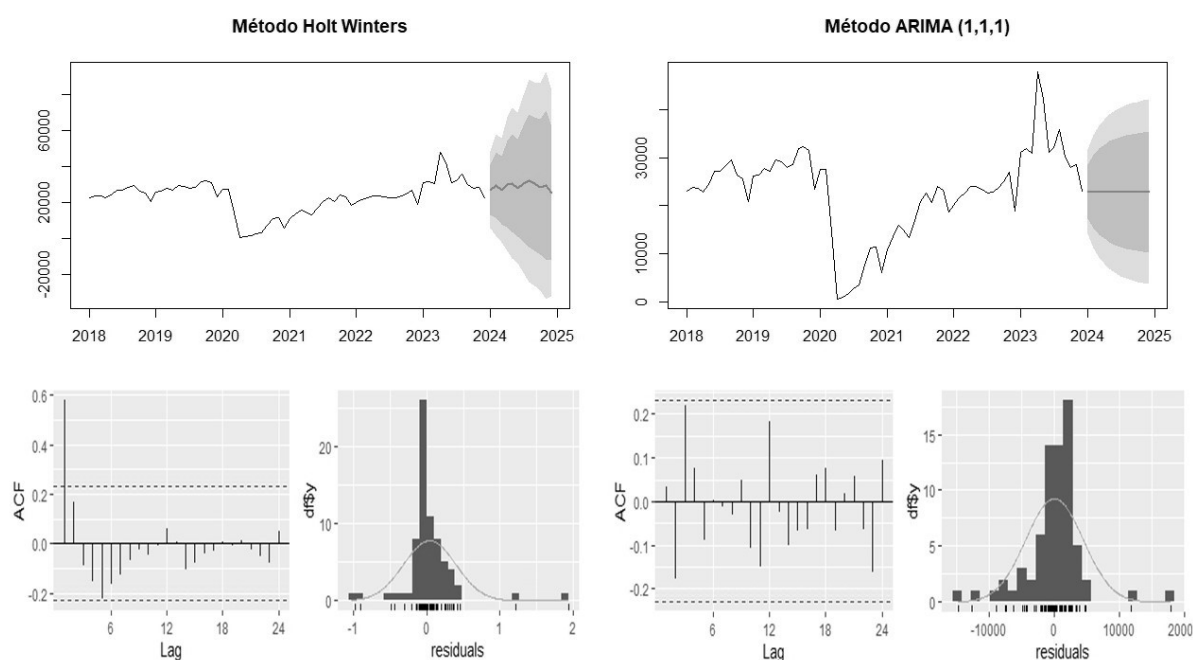


La figura 03 es una composición de imágenes que gráficamente indican determinados patrones del comportamiento de ambos métodos Holt-Winters en la columna izquierda y ARIMA en la columna derecha de la figura.

En el método HW la serie muestra una leve tendencia creciente con posible estacionalidad. El pronóstico se expande con intervalos de confianza amplios. El estadístico de BP con p-valor = 0.000252 (< 0.05) indica que la variabilidad no es constante por lo que no es el mejor modelo para realizar predicciones a futuro. El ACF del modelo muestra autocorrelación significativa en el primer lag o rezago, verificando en los resultados de la prueba de LB un p-valor = 0.0000054, muy por debajo del nivel de significancia (< 0.05) lo que indica que el modelo no es ruido blanco y no podría capturar completamente la estructura temporal de la serie. El análisis de los residuales del método HW en la imagen no muestra claramente una distribución sesgada de los valores, aunque con la prueba de KS con p-valor = 0.001218 (< 0.05) se observa que el modelo no presenta una distribución normal de los errores.

El modelo de ARIMA en la imagen refleja una tendencia general con algunos patrones autorregresivos que podrían estar influenciados por valores previos acorde a la estacionalidad de la serie, así mismo es observable que a diferencia del método anterior los intervalos de confianza son más estrechos. La homocedasticidad residual se comprueba con el p-valor = 0.5073 mayor que el valor de significancia (0.05) permitiendo al modelo tener un buen ajuste para realizar pronósticos. La prueba LB toma como p-valor = 0.78 mayor al nivel de significancia (0.05) lo que indica que los residuos son independientes y no existe autocorrelación significativa sugiriendo un mejor ajuste. En ARIMA, los residuos parecen mostrar una distribución normal a nivel gráfico y aunque el estadístico KS con p-valor = 0.0513 es ligeramente superior al nivel de significancia (0.05), los supuestos analizados muestran un mejor desempeño de éste modelo frente al modelo HW.

Figura 03. Comparativo de pronóstico de métodos Holt Winters y ARIMA 2018-2023



Los datos observados en la tabla 02 fueron el resultado de los pronósticos para el período siguiente de los meses del año 2024 utilizando dos métodos Holt Winters (HW) y ARIMA. Ambos modelos pronostican un horizonte de 12 meses adelante ($h = 12$), con intervalos de confianza del 80% y 95%. El método HW produce predicciones más altas en comparación con el método ARIMA, así como los intervalos de confianza son más amplios y, en determinadas estimaciones se observan los límites inferiores con valores negativos, lo que podría indicar un grado de inestabilidad o sobreajuste del modelo. Para el caso de ARIMA se observa que sus

predicciones son más estables y dentro de un rango más controlado, y a diferencia del método de HW no se muestran valores negativos

Tabla 02. Pronósticos de series de tiempo con método HW vs ARIMA para el año 2024

Mes	Pronóstico (h = 12)	Lo 80	Hi 80	Lo 95	Hi 95
Holt Winters					
Ene-24	27104.95	13089.43	41120.48	5670.06	48539.85
Feb-24	29769.06	11400.97	48137.15	1677.49	57860.63
Mar-24	26811.74	7779.06	45844.42	-2296.23	55919.72
Abr-24	30334.36	6109.72	54559.00	-6714.03	67382.75
May-24	30903.15	3556.56	58249.73	-10919.85	72726.14
Jun-24	28273.76	849.50	55698.02	-13668.03	70215.55
Jul-24	30585.77	-1666.77	62838.32	-18740.23	79911.78
Ago-24	32306.56	-4494.42	69107.55	-23975.68	88588.81
Set-24	30392.50	-6816.45	67601.45	-26513.68	87298.68
Oct-24	28767.51	-8928.23	66463.24	-28883.14	86418.16
Nov-24	29647.74	-11789.93	71085.41	-33725.71	93021.19
Dic-24	25129.42	-12225.10	62483.95	-31999.39	82258.24
ARIMA					
Ene-24	22934.36	17264.19	28604.53	14262.59	31606.13
Feb-24	22924.92	15301.79	30548.04	11266.36	34583.48
Mar-24	22916.54	14021.22	31811.85	9312.33	36520.75
Abr-24	22909.10	13101.88	32716.32	7910.25	37907.96
May-24	22902.50	12411.48	33393.52	6857.87	38947.13
Jun-24	22896.65	11878.92	33914.38	6046.48	39746.81
Jul-24	22891.45	11460.55	34322.35	5409.39	40373.50
Ago-24	22886.84	11127.44	34646.24	4902.39	40871.29
Set-24	22882.75	10859.39	34906.11	4494.61	41270.89
Oct-24	22879.11	10641.80	35116.43	4163.75	41594.48
Nov-24	22875.89	10463.84	35287.95	3893.29	41858.49
Dic-24	22873.03	10317.33	35428.73	3670.75	42075.31

En la tabla 03 se muestra la evaluación de la precisión de los modelos HW y ARIMA. Las métricas de precisión del tipo absolutas como son el promedio de errores(ME), raíz del error cuadrático medio (RMSE) y el error absoluto medio (MAE) de forma general muestran al modelo ARIMA con menores valores frente al modelo HW, lo que sugiere mejor desempeño de la serie temporal al cometer menores errores. Por otro lado, las métricas del tipo relativas como son el error porcentual medio(MPE) y el error porcentual absoluto medio (MAPE) de igual manera muestran al modelo ARIMA con mejor ajuste que el modelo HW, aunque los

valores obtenidos del MAPE en ambos modelos se encuentran dentro del intervalo de 20 – 50% considerándose como una precisión moderada.

Tabla 03. Evaluación de la precisión de modelos HW y ARIMA mediante validación cruzada

	ME	RMSE	MAE	MPE	MAPE
Holt Winters	-333.90	13619.77	9795.87	-0.92	33.54
ARIMA	-187.53	11214.19	7936.90	-0.44	27.01

Discusión

3 Los análisis de series de tiempo tienen por finalidad la búsqueda de predicciones futuras del comportamiento de los datos, para ello los conjuntos de entrenamiento y de prueba van afines al tamaño total de la muestra, el tipo de metodología y el análisis de cada componente de la serie como algo determinante para elaborar el modelo predictivo, por ello que tamaños de muestra pequeños podrían no ser representativos e influir en el rendimiento comparativo de diferentes modelos [2,34], aunque por otro lado se presentan análisis de series temporales de períodos cortos en ciencias ambientales con modelos de predicción con enfoques más modernos como Machine Learning [35] o modelos de regresión de proceso gaussianos (GPR) [36]. La información recolectada corresponde a un período de seis años (2018-2023) extraída de una base de datos abiertos de ahí que se desprenda que la limitación de datos y la focalización del presente estudio en una sola región del Perú lo que no permite ampliar el tamaño de dicha muestra, porque se están considerando todas las prestaciones odontológicas ya realizadas en la región Huánuco en retrospectiva.

Al comparar el total de atenciones sólo del año 2020 considerado como atípico (115 628) frente a la media de años considerados normales 2018–2019 y 2021–2023 (305 545), se observa una disminución del 62.16% lo que advierte una evidente caída que representa un sesgo significativo que podría distorsionar los resultados de los modelos de pronóstico si no se trata adecuadamente.

En la serie de estudio se ha identificado no estacionariedad debido a su componente de tendencia positiva, muy ligado al inicio de la pandemia y de comportamiento similar en otros estudios [37]. En el mismo análisis preliminar se ha identificado estacionalidad en el período con un horizonte de datos mensual. Sin embargo, estudios revelan que la granularidad de los datos como cambiar los intervalos de tiempo podría facilitar la visualización de estacionalidad y la identificación de relaciones entre diferentes patrones de series temporales, mejorando el rendimiento del método [38].

Debido a que la serie no cuenta con un patrón estacionario adecuado para el análisis, la descomposición de la serie temporal fue necesaria con el objetivo de analizar de manera independiente cada uno de sus componentes. Dicho patrón no estacionario según el análisis ACF indica que no es muy adecuado para realizar predicciones, de tal manera que la diferenciación de la serie también se utilizó como alternativa ante la falta de estacionariedad.

El método ARIMA es el modelo con mejor desempeño para esta serie temporal frente al método HW, dado que tiene mayor estabilidad pronóstica frente a este último ya que presenta intervalos de confianza muy amplios lo que podría sugerir cierta incertidumbre en la predicción. Un rendimiento superior de un método sobre otro se atribuye a diversos factores como por ejemplo la adaptabilidad de ARIMA a patrones de series temporales complejas, que incluye captura y respuesta eficaz a fluctuaciones impredecibles de una demanda [39] ; además otra ventaja significativa radica en el manejo de variaciones sin estacionalidad permitiendo al modelo hacer pronóstico precisos incluso aún en comportamiento errático que pueden observarse en atenciones sanitarias de características muy cambiantes [40]. Asimismo, estudios sobre comportamiento de enfermedades digestivas de vigilancia mensual indican que el analizar los componentes del modelo y a su confiabilidad en el manejo de los patrones de demanda erráticos identificando cambios en la tendencia y la estacionalidad contribuirán a proporcionar una información temprana para tomar medidas sanitarias en base a decisiones informadas que puedan reducir algún impacto negativo en la salud [41].

Como se mencionó ARIMA es el modelo con mejor desempeño para el presente estudio. Sin embargo, mejorar la precisión podría tomar en cuenta ver si los datos presenten estacionalidad fuerte y probar modelos SARIMA, así como mejorar el ACF con ARIMAX para capturar dinámicas en la serie inclusive si la estacionalidad no fuera evidente. Finalmente, el uso de técnicas avanzadas como redes neuronales en adecuación a series temporales como son Long Short-Term Memory LSTM debido a que, en el caso de series no estacionarias, estas redes pueden aprender la relación entre las entradas y las salidas sin necesidad de diferenciar la serie o realizar transformaciones previas para estacionarizarla [42].

Referencias

1. Moreno Sarmiento E. Predicción con series de tiempo y regresión. Panorama. 2013;2(4):36–58.
2. Ríos G. Series de Tiempo (Tópicos de minería de datos, Universidad de Chile). 2008;
3. Tonekaboni S, Eytan D, Goldengerg A. UNSUPERVISED REPRESENTATION LEARNING FOR TIME SERIES WITH TEMPORAL NEIGHBORHOOD CODING. ICLR 2021 - 9th Int Conf Learn Represent. 2021;
4. Lu H, Ma X, Ma M. A hybrid multi-objective optimizer-based model for daily electricity demand prediction considering COVID-19. Energy. 2021 Mar 15;219:119568.
5. Arena S, Florian E, Zennaro I, Orrù PF, Sgarbossa F. A novel decision support system for managing predictive maintenance strategies based on machine learning approaches. Saf Sci. 2022 Feb 1;146:105529.
6. Xiang X, Li Q, Khan S, Khalaf OI. Urban water resource management for sustainable environment planning using artificial intelligence techniques. Environ Impact Assess Rev. 2021 Jan 1;86:106515.
7. Maleki M, Mahmoudi MR, Wraith D, Pho KH. Time series modelling to forecast the confirmed and recovered cases of COVID-19. Travel Med Infect Dis. 2020 Sep 1;37:101742.
8. Papastefanopoulos V, Linardatos P, Kotsiantis S. COVID-19: A comparison of time series methods to forecast percentage of active cases per population. Appl Sci. 2020;10(11):1–15.
9. Ahsan MM, Luna SA, Siddique Z. Machine-Learning-Based Disease Diagnosis : A. Healthcare. 2022;1–30.

10. Chattha MA, Malik MI, Dengel A, Ahmed S. KRNN: A hybrid data and knowledge oriented time series forecasting approach for health care applications. *Expert Syst Appl* [Internet]. 2025;268(July 2022):126148. Available from: <https://doi.org/10.1016/j.eswa.2024.126148>
11. Shah D, Thaker M. A Review of Time Series Forecasting Methods. *Int J Res Anal Rev*. 2024;11(April):749.
12. Eldele E, Ragab M, Chen Z, Wu M, Kwoh CK, Li X, et al. Time-Series Representation Learning via Temporal and Contextual Contrasting. *IJCAI Int Jt Conf Artif Intell*. 2021;2352–9.
13. Zhao T, Liu H, Bulloch G, Jiang Z, Cao Z, Wu Z. The influence of the COVID-19 pandemic on identifying HIV/AIDS cases in China: an interrupted time series study. *Lancet Reg Heal - West Pacific* [Internet]. 2023;36:100755. Available from: <https://doi.org/10.1016/j.lanwpc.2023.100755>
14. Topaz M, Davoudi A, Evans L, Sridharan S, Song J, Chae S, et al. Building a Time-Series Model to Predict Hospitalization Risks in Home Health Care: Insights Into Development, Accuracy, and Fairness. *J Am Med Dir Assoc*. 2025 Feb 1;26(2):105417.
15. Martopullo I, Neves PA, Baird S, Liang M, Keats EC, Cherkas A, et al. Understanding progress and challenges in women's health and wellbeing in exemplar countries: a time-series study identifying positive outliers. *Lancet Glob Heal* [Internet]. 2024;12(12):e2012–23. Available from: [http://dx.doi.org/10.1016/S2214-109X\(24\)00364-4](http://dx.doi.org/10.1016/S2214-109X(24)00364-4)
16. Cueva Moncayo MF, Jami Carrera JE, Betancourt Rubio E. COVID-19 prediction using time series modelling in patients in a rural health clinic. *Rev Cuba Investig Biomed* [Internet]. 2023;42. Available from: <https://revibiomedica.sld.cu/index.php/ibi/article/view/3127/1396>
17. Monteiro LD, Mello FRM, Miranda TP, Heukelbach J. Hansen's disease in children under 15 years old in the state of Tocantins, Brazil, 2001-2012: Epidemiological patterns and temporal trends. *Rev Bras Epidemiol*. 2019;22.
18. Garay OU, Caporale JE, Pichón-Riviere A, Martí SG, Mullen M Mac, Augustovski F. Budgetary Impact Analysis in health: Update with a model using a generic approach.

- Rev Peru Med Exp Salud Publica [Internet]. 2011 [cited 2024 May 3];28(3):540–7. Available from: http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1726-46342011000300021#:~:text=En otras palabras%2C es la,cobertura en una determinada intervenci3n.
19. Hsu J, Jowett M, Mills A. Jo u rn a l P ro of. SSM - Heal Syst [Internet]. 2025;100055. Available from: <https://doi.org/10.1016/j.ssmhs.2025.100055>
 20. Annex of the 2030 Agenda. Annex: Global Indicator Framework for the Sustainable Development Goals and Targets of the 2030 Agenda for Sustainable Development. Work Stat Comm Pertain to 2030 Agenda Sustain Dev [Internet]. 2017;1–21. Available from: https://unstats.un.org/sdgs/indicators/Global Indicator Framework_A.RES.71.313 Annex.pdf
 21. Organizaci3n Panamericana de la Salud. Los sistemas de informaci3n y la salud digital en la pandemiade COVID-19: Examen posteriora la acci3n de los 100 primerosd3as de cuarentena. Organ Panam la Salud [Internet]. 2020; Available from: <https://iris.paho.org/handle/10665.2/52487>
 22. Ministerio de Salud Per3. Ministerio de Salud Per3. 2020. p. 23 Cuidado Integral en Salud Bucal. Available from: https://www.minsa.gob.pe/ensap/resoluciones/d_profam/2020/RD-068-2020_ENSAP-MINSA.pdf
 23. Casta3eda Z. prevalencia de caries dental y necesidad de tratamiento en los alumnos de la I.E.P Sagrado Ignacio de Loyola 2015. Univ Se3or Sipan [Internet]. 2016;101. Available from: <http://repositorio.uss.edu.pe/bitstream/handle/uss/132/TESIS.PREVALENCIA DE CARIES DENTAL Y NECESIDAD TRATAMIENTO.ZABDIEL CASTA3EDA MORENO.CORREGIDO.pdf?sequence=1&isAllowed=y>
 24. Direcci3n Regional de Salud Hu3nuco. Oficina de Inform3tica, Telecomunicaciones y Estadística OITE-DIRESA-HCO. 2023. Morbilidad - DIRESA Hu3nuco. Available from: <https://weboite.diresahuanuco.gob.pe/Morbilidad>
 25. Kaushik S, Choudhury A, Sheron PK, Dasgupta N, Natarajan S, Pickett LA, et al. AI in Healthcare: Time-Series Forecasting Using Statistical, Neural, and Ensemble Architectures. Front Big Data. 2020;3(January).
 26. Leavy P, Symmons SM, Mockler D, Fleming P, Daly B, Ford J, et al. How and why do

- health system factors influence general dentists' participation in publicly funded, contracted primary dental care services: A realist review. *Health Policy* (New York). 2025;153(January):3–4.
27. Montenegro MG. A New Approach to Oral Health: A Public Health Perspective. 2011;30(64):101–8. Available from: <https://dialnet.unirioja.es/descarga/articulo/3670238.pdf>
 28. Manterola C, Otzen T. Estudios observacionales. Los diseños utilizados con mayor frecuencia en investigación clínica. *Int J Morphol*. 2014;32(2):634–45.
 29. Manterola C, Quiroz G, Salazar P, García N. Methodology of study designs most frequently used in clinical research. *Rev Medica Clin Las Condes* [Internet]. 2019;30(1):36–49. Available from: <https://doi.org/10.1016/j.rmclc.2018.11.005>
 30. Hernandez Sampieri, Roberto-Fernandez Collado, Carlos-Baptisa Lucio P. *Metodología de la Investigación*. 2014.
 31. Seguro Integral de Salud - SIS. Plataforma Nacional de Datos Abiertos. 2024 [cited 2024 Sep 25]. Datos de Atenciones realizadas a los asegurados. Available from: <https://www.datosabiertos.gob.pe/dataset/datos-de-atenciones-realizadas-los-asegurados-sis>
 32. Dirección Regional de Salud Huánuco. Análisis de situación de salud de la región Huánuco [Internet]. 2023. Available from: <https://publicaciones.diresahuanuco.gob.pe/index.php/s/S5PDICBfYIzfNAz>
 33. Salud OP de la. Clasificación Estadística Internacional de Enfermedades y problemas relacionados con la salud. 1995;10(1).
 34. Hentati Isacsson N, Zantvoort K, Forsell E, Boman M, Kaldo V. Making the most out of timeseries symptom data: A machine learning study on symptom predictions of internet-based CBT. *Internet Interv* [Internet]. 2024;38(September):100773. Available from: <https://doi.org/10.1016/j.invent.2024.100773>
 35. Van Kwan Zhi Koh, Ye Li, Xing Yong Kek, Ehsan Shafiee, Zhiping Lin BW. A review of recent hybridized machine learning methodologies for time series forecasting on water-related variables. *J Hydrol* [Internet]. 2025;35. Available from: <https://doi.org/10.1016/j.jhydrol.2025.132909>

36. Gil-Díaz T, Trumm M. Gaussian process regression as a powerful tool for analysing time series in environmental geochemistry. *Ecol Inform* [Internet]. 2024;84(November):102877. Available from: <https://doi.org/10.1016/j.ecoinf.2024.102877>
37. Shinde GR, Kalamkar AB, Mahalle PN, Dey N, Chaki J, Hassanien AE. Forecasting Models for Coronavirus Disease (COVID-19): A Survey of the State-of-the-Art. *SN Comput Sci*. 2020;1(4).
38. Zhou Y, Ren H, Li Z, Pedrycz W. Anomaly detection based on a granular Markov model. *Expert Syst Appl* [Internet]. 2022;187(August 2021):115744. Available from: <https://doi.org/10.1016/j.eswa.2021.115744>
39. Kumar L, Khedlekar S, Khedlekar UK. A comparative assessment of holt winter exponential smoothing and autoregressive integrated moving average for inventory optimization in supply chains. *Supply Chain Anal* [Internet]. 2024;8(September):100084. Available from: <https://doi.org/10.1016/j.sca.2024.100084>
40. Alzahrani SI, Aljamaan IA, Al-Fakih EA. Forecasting the spread of the COVID-19 pandemic in Saudi Arabia using ARIMA prediction model under current public health interventions. *J Infect Public Health* [Internet]. 2020;13(7):914–9. Available from: <https://doi.org/10.1016/j.jiph.2020.06.001>
41. Tosepu R, Ningsi NY. Forecasting of diarrhea disease using ARIMA model in Kendari City, Southeast Sulawesi Province, Indonesia. *Heliyon*. 2024;10(22).
42. Liu Q, Wang Q, Wang X. RVFL-LSTM: A lightweight model with long-short term memory for time series. *Knowledge-Based Syst* [Internet]. 2025;309(December 2024):112896. Available from: <https://doi.org/10.1016/j.knosys.2024.112896>