



Revista EIA
ISSN 1794-1237
e-ISSN 2463-0950
Año XIX/ Volumen 21/ Edición N.42
Julio - diciembre de 2024
Reia4227 pp. 1-22

Publicación científica semestral
Universidad EIA, Envigado, Colombia

**PARA CITAR ESTE ARTÍCULO /
TO REFERENCE THIS ARTICLE /**

Cifuentes Madrigal, A.; Gómez Méndez, T. S.; Jiménez Zapata, M.
Analítica predictiva como apoyo en la salud pública: Modelos de pronóstico con series de tiempo aplicados a la conducta suicida
Revista EIA, 21(42), Reia4227.
pp. 1-22.
<https://doi.org/10.24050/reia.v21i42.1763>

✉ *Autor de correspondencia:*

Maritza Jiménez Zapata
Ingeniería Administrativa, Maestría en Ingeniería, Doctorado en Ingeniería
mjimenez@unal.edu.co

Recibido: 23-02-2024

Aceptado: 27-05-2024

Disponible online: 01-07-2024

Analítica predictiva como apoyo en la salud pública: Modelos de pronóstico con series de tiempo aplicados a la conducta suicida

ANDREA CIFUENTES MADRIGAL¹

TOMÁS SIMÓN GÓMEZ MÉNDEZ²

✉ MARITZA JIMÉNEZ ZAPATA³

1. Universidad Pontificia Bolivariana, Colombia
2. KU Leuven; University of Brescia, Bélgica
3. Universidad Nacional de Colombia

Resumen

La salud mental ha sido identificada por la Organización Mundial de la Salud (OMS) como un asunto de interés prioritario a nivel internacional. En el caso particular de Colombia, se ha destacado la importancia de la atención a la salud mental incluyendo el fenómeno del suicidio. En este sentido, los entes reguladores han buscado alternativas que solucionen o contribuyan a la mitigación de esta problemática. En dicho contexto, se hace pertinente realizar contribuciones académicas en torno al desarrollo de herramientas que faciliten la toma de decisiones de los entes gubernamentales, por ejemplo, a través de la formulación de modelos de pronóstico que permita identificar tendencias y patrones de comportamiento de los intentos de suicidio.

El objetivo del estudio es formular y seleccionar un modelo de pronóstico regional para los intentos de suicidio mediante analítica predictiva. Este artículo contribuye con la brecha en la literatura para el caso de la ciudad de Medellín, Colombia. Se presenta el ajuste, validación y comparación de tres diferentes modelos de series temporales, bajo criterios de mínimo error de pronóstico. Los modelos comparados incluyen aproximaciones paramétricas, Holt-Winters y Box-Jenkins. Los resultados indican que, para los datos analizados, el modelo paramétrico de componentes aditivos cúbico estacional con errores

ARMA(0,5) es el que presenta menor error de pronóstico. El modelo construido logra capturar las tendencias del fenómeno, y tiene un bajo nivel de error para la proyección de tendencias cercanas, pero no logra dar respuesta a los cambios repentinos en la estructura como los que ocurrieron en la pandemia de COVID-19.

Palabras clave: Analítica de Datos; Analítica Predictiva; Intento de Suicidio; Modelo ARMA; Modelos de Pronóstico; Prueba de Hipótesis; Salud Mental; Salud Pública; Series de Tiempo; Suicidio.

Predictive analytics for policymakers: Time-series forecasting applied to suicide attempts

Abstract

Mental health has been identified by the World Health Organization (WHO) as a matter of concern internationally. In Colombia, specifically, the importance of mental health care, including the phenomenon of suicide, has been highlighted. In this regard, policymakers have sought alternatives to address or mitigate this issue. In this context, research becomes relevant regarding the development of tools that facilitate decision-making for governmental authorities, for example, through the formulation of forecasting models that enable the identification of trends and patterns of behaviour of suicide attempts.

The objective of this study is to formulate and select a regional forecasting model for suicide attempts by means of predictive analytics techniques. This paper contributes to fill the gap in the case of suicide attempts in the city of Medellín, Colombia. The fitting, validation and comparison of three different time series models, under minimum forecast error criteria is presented. The compared models include parametric, Holt-Winters and Box-Jenkins approximations; For the analysed data, the parametric model with cubic, seasonal and ARMA(0.5) components is the one with the lowest forecast error. This model manages to capture the trends of the phenomenon, with a low level of error for the projection for nearby trends, but it does not manage to respond to sudden changes in structure such as those that occurred in the COVID-19 pandemic.

Keywords: Data Analytics; Predictive Analytics; Suicide Attempt; ARMA Model; Forecasting Models; Hypothesis Testing; Mental Health; Public Health; Time Series; Suicide.

1. Introducción

Actualmente el interés por la salud mental en la sociedad está en aumento, reflejado en el incremento de planes y políticas públicas orientados a promover hábitos saludables, y en el establecimiento de espacios y planes para la atención temprana en situaciones de alerta (Organización Mundial de la Salud, 2022a). Lo anterior, como una manera de contrarrestar los casos de suicidio, que en los últimos años han estado en crecimiento (Ministerio de Salud y Protección Social, 2018).

En Colombia, existe evidencia de la presencia de conductas suicidas en sus habitantes, de hecho, se estima que cada 20 minutos hay un intento de suicidio en el país (Vanegas, 2022). Y, en línea con las orientaciones internacionales, el ministerio de salud y protección social de Colombia en 2021 estableció la estrategia para la prevención de la conducta suicida (Ministerio de Salud y Protección Social, 2021).

Al reconocer la problemática de los casos de suicidio en Colombia se hace importante que diversas áreas del conocimiento busquen alternativas para aportar en la gestión de este fenómeno.

En este sentido, desde el área de la analítica de datos se plantea la necesidad de realizar un análisis que permita reconocer cómo ha sido el comportamiento de la cantidad de intentos de suicidios a lo largo del tiempo, con el fin de identificar tendencias y posibles predicciones al número de eventos que se pueden presentar en un futuro; y que sirvan de insumo a las autoridades competentes para planear políticas específicas de intervención social, y para contrastar los resultados de algunas políticas con las tendencias identificadas, permitiendo la evaluación de desempeño de los planes de prevención y acción en salud mental.

El estudio de los casos de suicidio en Colombia se ha dado principalmente desde el área de la salud, pero se encuentran trabajos que complementan este enfoque con elementos de ingeniería.

Algunos enfoques incluyen aproximaciones explicativas al fenómeno, como el trabajo de Tabares et al. (2020), quienes se concentran en el análisis de factores distales y proximales que explican la conducta suicida en universitarios de las ciudades colombianas Manizales y Medellín. En este trabajo los autores encuentran asociaciones estadísticamente significativas entre los factores atención emocional y auto-desprecio, con el riesgo suicida; sin embargo, este estudio no se orienta hacia la formulación de modelos predictivos del fenómeno.

Otras aproximaciones, incluyen trabajos como el de Quito Yanquen y Pinzón Choque (2021), en el cual se elabora un modelo de pronóstico de casos de suicidio en la ciudad de Bogotá, con el propósito de representar escenarios que pudieran ser tomados como referencia para la aplicación de programas de prevención en Bogotá, Colombia. Como resultado de este estudio, los autores seleccionaron un modelo ARIMA (10,1,0) con un comportamiento con medida de ajuste del Criterio de Información Bayesiano (BIC) de 477. Este trabajo, sin embargo, se enfoca en los casos de suicidios consolidados, dejando de lado el estudio de los intentos de suicidio, que pueden dar señales para la prevención oportuna.

La presente investigación desarrolla un modelo de pronóstico a partir de los casos de intento de suicidio reportados en base de datos MEData, incluyendo la información desde enero de 2013 a diciembre de 2020, que permita tener un acercamiento a cómo seguirá siendo el comportamiento de los casos intento de suicidio en Medellín. Para esto se utiliza las series de tiempo, como una de las herramientas de apoyo en la analítica predictiva, se valida la confiabilidad del modelo y se selecciona el modelo con menor nivel de error de acuerdo con las métricas de error absoluto porcentual medio (MAPE), la raíz cuadrada del MSE (RMSE) y error absoluto medio (MAE), que pueda otorgarle a los stakeholders información precisa, oportuna y relevante.

2. Relevancia del estudio de datos de salud mental para la gestión oportuna de los intentos de suicidio

La salud mental es considerada como un estado de bienestar en el cual una persona está en la capacidad de desarrollar su potencial y afrontar las adversidades de la vida (Organización Mundial de la Salud, 2022b). Una de las manifestaciones de que la salud mental de un individuo no está alcanzando el bienestar deseado son los intentos de suicidio, lo cual la Organización Mundial de la Salud (OMS) ha declarado un problema de salud pública.

La OMS hace continuos llamados a los países a ser más atentos frente a esta problemática, por ejemplo, en junio de 2021 en un comunicado de prensa aseguraron que los suicidios siguen siendo una de las principales causas de muerte en todo el mundo (Organización Mundial de la Salud, 2021). Además, la OMS en su plan de acción sobre salud mental 2013-2030, plantea como meta que la tasa de suicidio se reduzca en un tercio para el año 2030 (Organización Mundial de la Salud, 2022a).

Mientras el suicidio se refiere a la muerte autoinfligida (Ministerio de Salud y Protección Social, 2018); los intentos de suicidio se definen como la “conducta potencialmente lesiva autoinfligida y sin resultado fatal, para la que existe evidencia, implícita o explícita, de intencionalidad de provocarse la muerte” (Ministerio de Salud y Protección Social, 2018, p.2). Dado que ambos conceptos se encuentran ligados, se revisan trabajos enfocados en los casos de suicidio o en los intentos de suicidio.

Los autores Cardona Arango et al. (2016) caracterizan los suicidios en Colombia, en el periodo comprendido entre 2000 y 2010, destacando que para este periodo se registró una tendencia a la baja. Ahora, desde el observatorio de salud mental se registra la tasa de intentos de suicidio según los casos por total de habitantes, para el 2015 registró un valor de 4,12, que incrementó en 2016 a un valor de 36,08, y el último reporte se dio para el 2022 con un valor de 67.54 (Ministerio de Salud y Protección Social, n.d.). Estas cifras contrastan con los reportes o tendencias que se habían

tenido en años anteriores, mostrando una clara alerta de la salud mental en los colombianos.

Por otro lado, Aguilar Siguéñas (2022) plantea un modelo predictivo y explicativo de la ideación suicida en internos de medicina, para esto utiliza análisis estadísticos entre diferentes factores que pueden dar indicios de la ideación suicida. Los análisis realizados incluyen correlación de Spearman, análisis de regresión evaluando las variables de interés, y la utilización de un modelo de ecuaciones estructurales. Como resultado, en este artículo el mejor modelo consigue explicar el 88,3% de la ideación suicida.

Adicionalmente, Gómez Tabares et al. (2019), muestra un análisis a los factores de riesgo y predictores asociados a la conducta suicida en estudiantes, para esto se realiza un análisis de regresión binaria, donde tener alguna enfermedad mental o intento de suicidio en la familia explican entre el 41% y el 62% de la varianza en el riesgo suicida. También se analizan ecuaciones estructurales donde el mejor modelo explica en un 62,7% la variación del riesgo suicida.

También, Fernandez Villacreses (2020), analizó mediante técnicas estadísticas los factores que influyen en la conducta suicida en individuos drogodependientes. En el estudio se utilizan para sus análisis las técnicas de regresión logística, árbol de decisión con CHAID y estadísticos de validación. En este trabajo se encontró que los hombres son el grupo más propenso a incurrir en ideas suicidas debido al consumo de determinadas sustancias, mientras que las mujeres son más propensas a intentar suicidarse, con una probabilidad doble a la de los hombres.

Los autores Chen et al. (2020) , desarrollan modelos de aprendizaje automático para predecir el comportamiento suicida en Suecia, se concluye que combinar el aprendizaje automático con datos registrados tiene el potencial de predecir con precisión el comportamiento suicida en el corto plazo.

De forma similar, autores como Nikam et al. (2023) desarrollan modelos basados en aprendizaje automático, incluyendo regresión lineal y árboles de decisión, para pronóstico de suicidio en india; obteniendo como resultados niveles de precisión entre el 85,98% y 90,48%. Por su parte, Shetty M et al. (2023), utilizan

también aprendizaje automático sobre bases de datos médicas implementando regresión logística para pronosticar intentos de suicidio, obteniendo tasas de precisión máximas del 75%.

También en el área del aprendizaje automático, se encontraron trabajos como el de Boonkwang et al. (2018), donde se utilizan árboles de decisión para identificar los intentos de suicidio, y luego procesarlos por medio de SMOTE (Syntetic Minority Over-sampling Technique) para equilibrar el conjunto de datos. Como resultado se obtuvo un nivel de precisión de 91,26%.

Continuando por la línea del aprendizaje automático, cabe resaltar el estudio de Schafer et al. (2021) donde se comparan modelos teóricos con modelos de aprendizaje automático y se evidencia que estos últimos proporcionan una predicción superior en comparación con los modelos teóricos. A pesar de este panorama alentador, otros estudios muestran que la herramienta de aprendizaje automático es aún muy reciente como para confiarle el análisis de situaciones de salud pública, por lo cual se hace necesaria la interdisciplinariedad con áreas sociales (Linthicum et al., 2019).

Adicionalmente, se encuentra un estudio donde se utiliza la Big Data para predecir la tasa de suicidios, a través de herramientas como aprendizaje profundo, árboles de decisión y KNN. El análisis realizado muestra la posibilidad de tener predicciones para detectar individuos que requieran de apoyo emocional. Los autores presentan resultados con un nivel de precisión cercano al 100% (Madyatmadja et al., 2021).

Por último, el uso de inteligencia artificial en el ámbito de la salud mental se presenta como una estrategia para detectar patrones en medio de grandes datos, determinando factores de riesgo y generando alertas sobre la predicción de suicidios. De manera individual ayudaría a identificar personas en crisis para intervenir con apoyo, de manera colectiva se pueden identificar grupos en riesgo o puntos críticos de suicidio (Fonseka et al., 2019).

De acuerdo con lo anterior, se evidencia que en la actualidad no se cuenta con modelos predictivos recientes que permitan estimar la cantidad de los intentos de suicidios que se puedan presentar en la ciudad de Medellín. Es por esto que se hace pertinente desarrollar

y seleccionar un modelo, por medio de analítica predictiva con la herramienta de series de tiempo para elaborar pronósticos con un bajo nivel de error, que les permita a entes reguladores tomar medidas para afrontar este problema de salud mental.

3. Materiales y métodos

Los datos de intentos de suicidio para la ciudad de Medellín, en el periodo de enero de 2013 hasta diciembre de 2020, se obtuvieron del portal de datos abiertos de la alcaldía de Medellín MEData (Alcaldía de Medellín, 2020).

Con base en esta información se construyen modelos de pronóstico con series de tiempo para los datos utilizando un modelo paramétrico de componentes aditivos, un modelo Holt-Winters y un modelo Box-Jenkins. A cada uno se le realizan pruebas de significancia y el respectivo análisis de validación de supuestos sobre los errores.

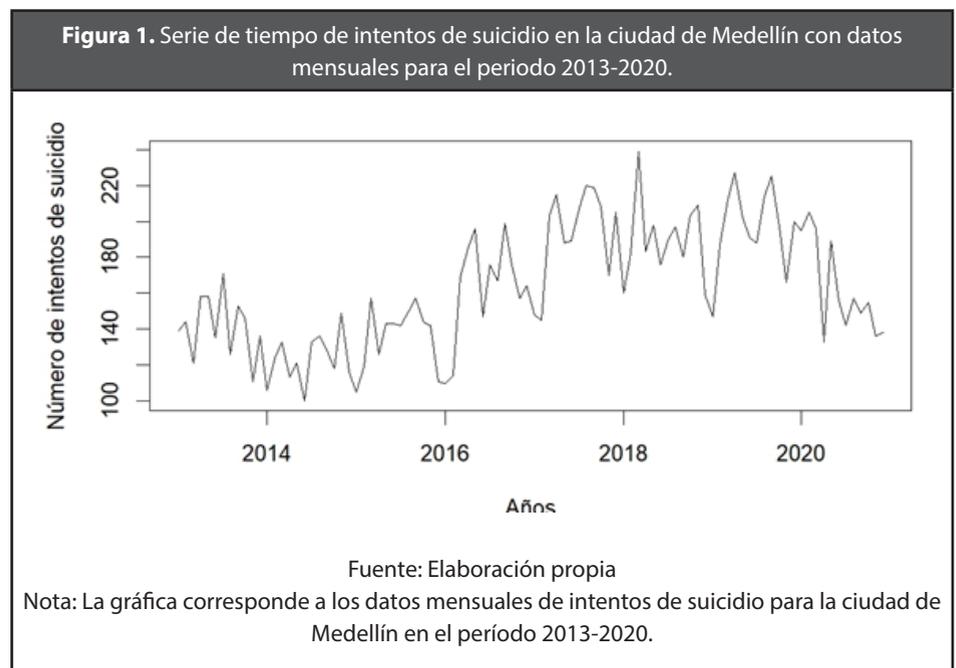
Para validar los supuestos de los errores se verifica que su valor esperado sea 0, y que tengan varianza constante. Posteriormente se valida que los errores sean incorrelacionados utilizando la gráfica de la función de autocorrelación (ACF), donde la mayoría de los rezagos deben estar dentro de las bandas de Bartlett; y, por último, se realiza la prueba de hipótesis Ljung-Box que evalúa la independencia de los errores, donde el valor p debe ser mayor al error permitido que se establece para este trabajo en 5% (Shumway and Stoffer, 2011).

Los modelos se contrastan entre sí, en cuanto a su capacidad de reproducir el comportamiento histórico de los datos (ajuste), y su capacidad de pronóstico. Para analizar el desempeño de los modelos en cuanto al ajuste, los modelos se comparan utilizando el criterio de información de Akaike (AIC) y el criterio de información bayesiano (BIC) de cada modelo. Finalmente, para elegir el modelo que mejor pronostica se comparan buscando los menores valores en las métricas de error absoluto porcentual medio (MAPE), la raíz cuadrada del MSE (RMSE) y error absoluto medio (MAE).

El detalle de la construcción y validación de los modelos, se presentan en la sección 3.1.

3.1. Construcción y validación de los modelos

Para estimar los modelos de la serie de tiempo se ingresaron los datos previamente recolectados de la base de datos MEData al software estadístico RStudio, los cuales se muestran en la Figura 1. En esta se puede observar en el eje vertical el número de casos de intentos de suicidio, mientras que el eje horizontal se ve el registro año a año de esta información.



Durante el periodo 2013-2015 se puede ver una estabilidad en los casos, viendo incluso una disminución a finales del primer año. Posteriormente se ve un incremento a partir del 2016, evidenciando una tendencia creciente desde este año hasta el 2020, periodo en que se dio la pandemia del COVID-19, y se vuelve a ver una disminución en los casos de intentos de suicidio en Medellín.

Para comenzar a modelar la serie de tiempo se define un conjunto de prueba y uno de entrenamiento. El conjunto de entrenamiento permite estimar el comportamiento de los datos, y el conjunto de prueba es utilizado para realizar la verificación de los pronósticos. Los datos del conjunto de entrenamiento son los suicidios en cada mes desde enero de 2013 hasta diciembre de 2019, en total 84 datos, y el conjunto de prueba lo componen los 12 datos restantes desde enero de 2020 hasta diciembre de 2020.

Modelo de componentes

Se ajustó un modelo de componentes aditivo, incluyendo la tendencia, la estacionalidad y los errores. El modelo resultante incluye una tendencia cúbica, estacionalidad modelada con variables trigonométricas con un periodo estacional $s=12$, y errores tipo ARMA(0,5).

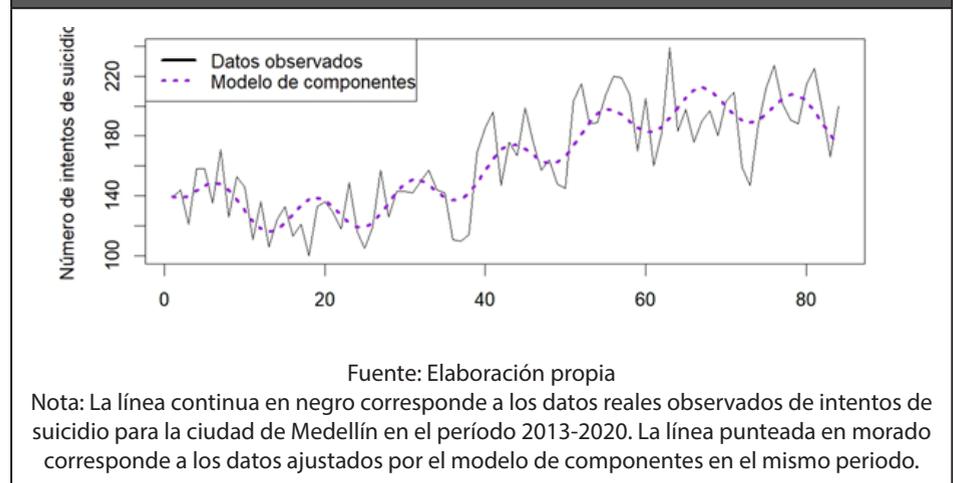
El modelo ajustado se presenta en la ecuación (1)

$$Y_t = 154,9 - 3,583t + 0,129t^2 - 0,00098t^3 - 5,230\sin\left(\frac{2\pi t}{12}\right) - 10,83\cos\left(\frac{2\pi t}{12}\right) + \varepsilon_t$$

Donde $\varepsilon_t = a_t + 0,0587a_{t-1} - 0,0820a_{t-2} - 0,1044a_{t-3} + 0,0811a_{t-4} + 0,2854a_{t-5}$
 con $a_t \sim R.B(0, \sigma^2)$

(1)

Figura 2. Ajuste del modelo de componentes a la serie de tiempo de datos observados.

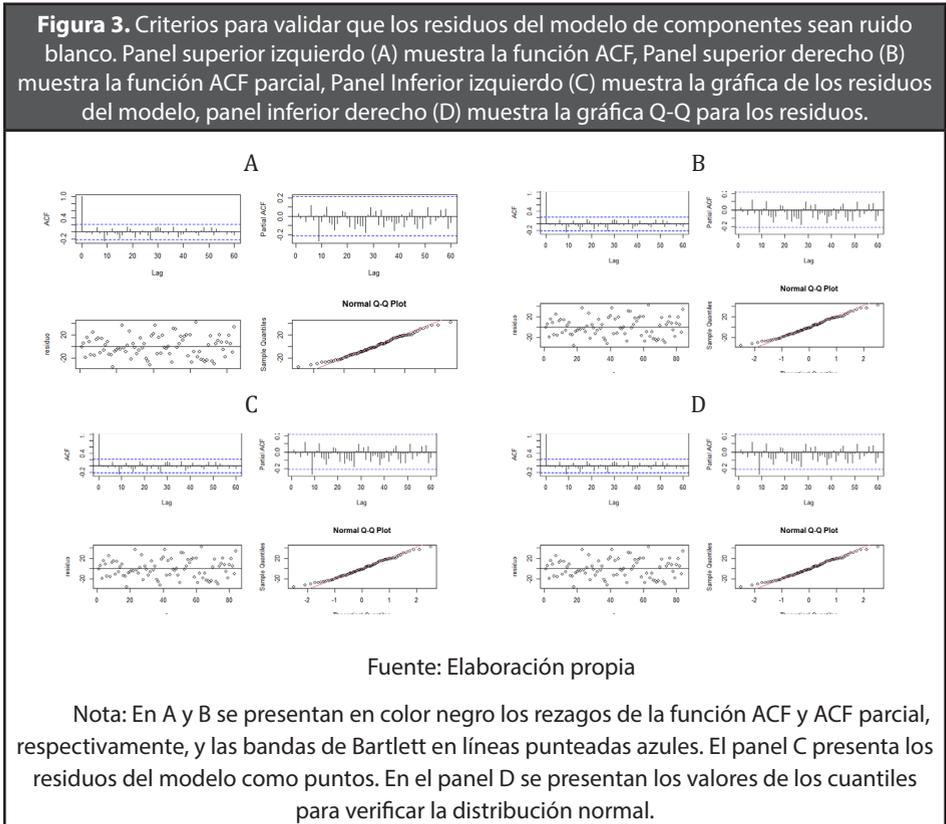


Las métricas de ajuste de este modelo son $R^2 = 0,7046$, R^2 ajustado = 0,6857, AIC= 746,2622 y BIC=763,2779. Por su parte, los valores de las métricas de error en los pronósticos son MAPE = 11,9551, MAE = 19,3690 y RMSE = 22,5247.

Para verificar que el modelo de componentes cumple con los supuestos sobre los errores se rectifica que el valor esperado de los errores sea 0, que la varianza sea homocedástica y que los errores sean independientes. La media cero se puede validar de manera gráfica observando el panel inferior izquierdo (C) de la figura 3, ya que todos los valores oscilan alrededor de 0, y no se evidencia la presencia de alguna tendencia.

Para validar el supuesto de varianza constante, se realiza también una inspección gráfica al panel inferior izquierdo (C) de la Figura 3, donde se observa que los errores están aleatoriamente distribuidos alrededor del 0.

Para verificar que los residuos son independientes se analiza la función de correlación (ACF) presentada en la figura 3 (panel superior izquierdo (A)), donde la mayoría de los rezagos están dentro de las bandas de Bartlett, y se realiza la prueba de hipótesis de Ljung-Box.



La prueba de hipótesis Ljung-Box se especifica de acuerdo con la ecuación (2) y el criterio de rechazo se establece con un valor p (vp) inferior al 5%, como se presenta en la ecuación (3).

Hipótesis: H_0 : Los errores son independientes vs H_a : Los errores no son independientes (2)

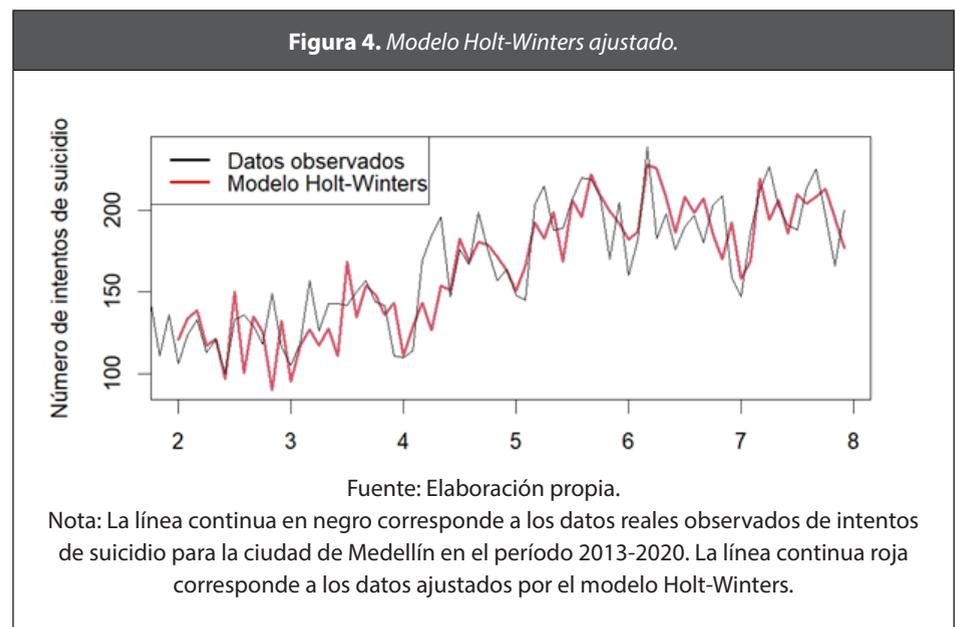
Criterio de rechazo: $vp < 0,05$ (3)

El valor p obtenido fue de 0,4562, por lo cual al considerar un criterio de rechazo para H_0 de $vp < 0,05$ (ecuación (3)), indica que no se puede rechazar la hipótesis nula, es decir que los errores son independientes con un nivel de confianza del 95%. Con esto se determina que los residuos son ruido blanco, y por lo tanto el modelo ajustado puede utilizarse para pronosticar los casos de suicidio en la ciudad de Medellín.

Modelo Holt-Winters

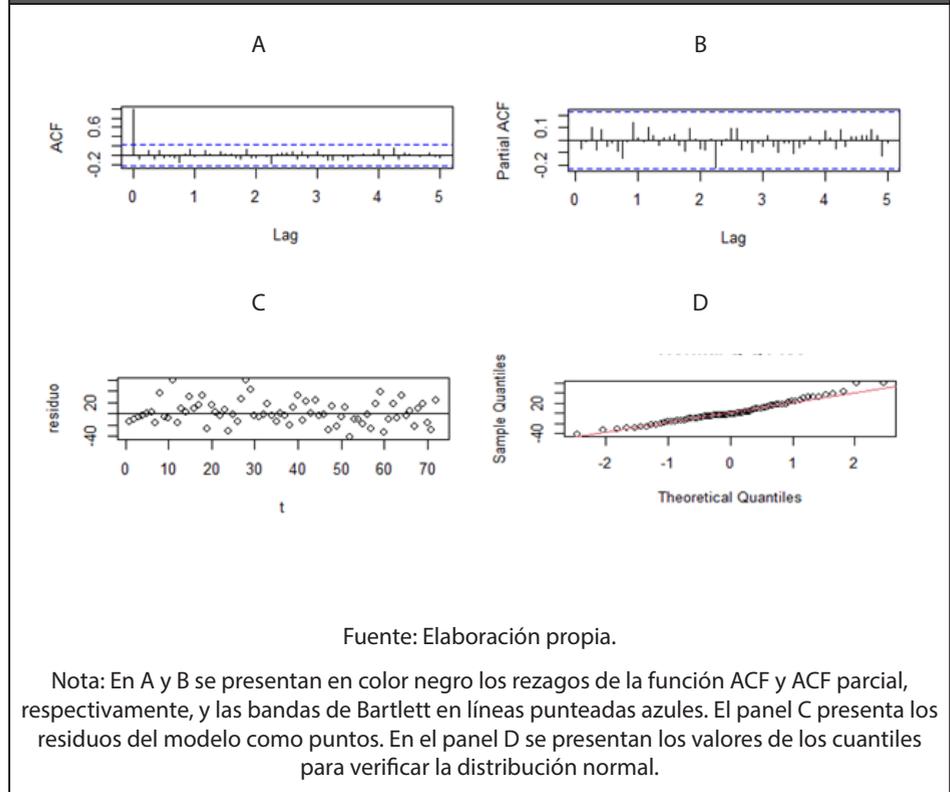
Se modela la serie de tiempo con el método Holt-Winters, el cual se ajustó utilizando los parámetros $\alpha=0,2676$, $\beta=0,0712$ y $\delta=0,4090$. Este modelo arrojó unas métricas de ajuste de AIC de 562,04 y BIC de 569,73, y métricas de pronóstico con MAPE de 31,3684, MAE de 47,2883 y el RMSE de 51,2965.

El comportamiento del modelo Holt-Winters ajustado se presenta en la figura 4.



Al verificar el cumplimiento de los supuestos sobre los errores del modelo Holt-Winters, se evidenció que el valor esperado de los residuos es 0, al no identificar ninguna tendencia en los residuos en el panel inferior izquierdo (C) de la figura 5. También se verifica la varianza constante, donde los errores se encuentran aleatoriamente distribuidos alrededor del 0.

Figura 5. Criterios para validar que los residuos del modelo Holt-Winters sean ruido blanco. Panel superior izquierdo (A) muestra la función ACF, Panel superior derecho (B) muestra la función ACF parcial, Panel Inferior izquierdo (C) muestra la gráfica de los residuos del modelo, panel inferior derecho (D) muestra la gráfica Q-Q para los residuos.



Se verifica la independencia de los errores, utilizando la gráfica de la función de autocorrelación (panel superior izquierdo (A) de la Figura 5), donde la mayoría de los rezagos se encuentran dentro de las bandas de Bartlett.

Finalmente, en la prueba de hipótesis Ljung-Box, usando las ecuaciones (2) y (3), y con un valor p de 0,9871 mayor a 0,05; no se rechaza la hipótesis nula y, por tanto, los errores del modelo Holt-Winters son Ruido Blanco, y el modelo construido puede usarse para pronosticar.

Modelo Box-Jenkins

Se utilizó la metodología de Box-Jenkins para ajustar un modelo de series de tiempo a los intentos de suicidio en la ciudad

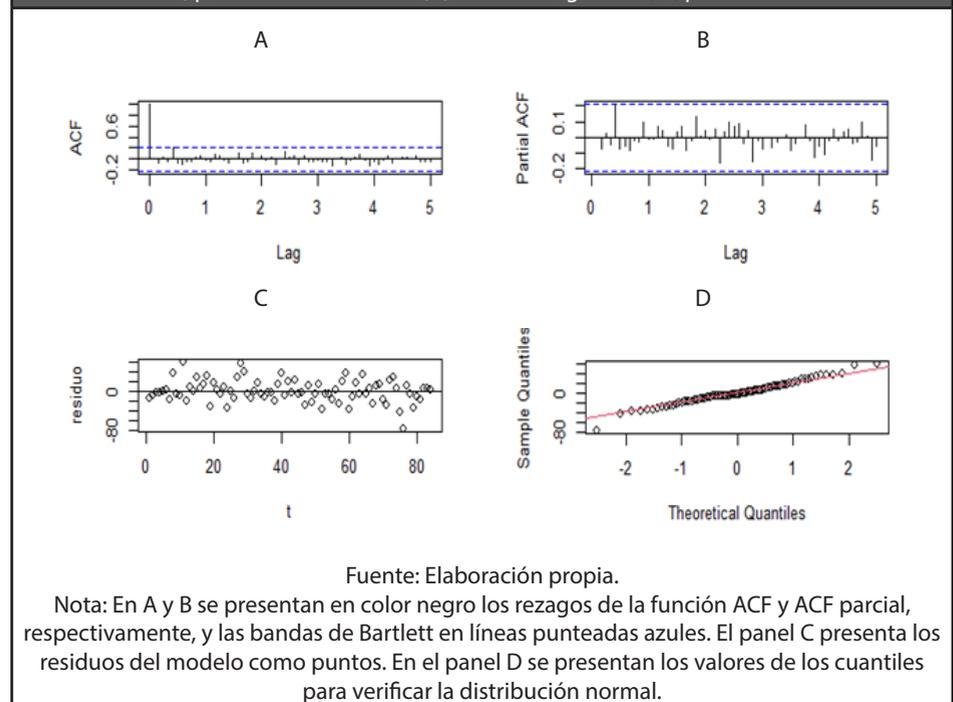
de Medellín, en este caso se utilizó un modelo SARIMA (0,1,1) (2,0,0) [12].

El modelo ajustado presenta un AIC de 748,3265, y un BIC de 758,0019, también tiene unas métricas de error con MAPE de 26,1805, MAE de 38,9156 y RMSE de 43,3434.

Una vez construido el modelo se verifica que los errores sean ruido blanco, para esto se realizan las pruebas gráficas y de hipótesis correspondientes.

En el panel inferior izquierdo (C) de la figura 6 se observa que la mayoría de los valores oscilan alrededor de 0 sin tener una tendencia, cumpliendo así el supuesto de que el valor esperado de los errores sea 0. De manera equivalente se evalúa si el modelo cumple con el supuesto de varianza constante por medio de verificación gráfica, en el panel inferior izquierdo (C) de la figura 6 se observa que los errores se encuentran distribuidos de manera aleatoria alrededor del 0, cumpliendo así el supuesto de varianza constante.

Figura 6. Criterios para validar que los residuos del modelo Box-Jenkins sean ruido blanco. Panel superior izquierdo (A) muestra la función ACF, Panel superior derecho (B) muestra la función ACF parcial, Panel Inferior izquierdo (C) muestra la gráfica de los residuos del modelo, panel inferior derecho (D) muestra la gráfica Q-Q para los residuos.



Adicionalmente, se hace la verificación de que los residuos sean independientes, para esto se analiza la función de correlación (ACF) presentada en el panel superior izquierdo (A) de la Figura 6, donde la mayoría de los rezagos están dentro de las bandas de Bartlett y se realiza a prueba de hipótesis de Ljung-Box.

Para la prueba de hipótesis Ljung-Box, se utilizan las ecuaciones (2) y (3), y considerando un valor $p = 0,8182$, no se rechaza la hipótesis nula, con lo cual se determina que los errores de la serie de tiempo en el modelo Box-Jenkins sí son ruido blanco.

4. Resultados y discusión

Una vez construidos y validados tres modelos de series de tiempo para la serie mensual de intentos de suicidio en la ciudad de Medellín, se realizan comparaciones entre ellos para identificar el modelo que mejor describe y el que mejor pronostica la serie estudiada, de acuerdo con los criterios estadísticos definidos en la sección 3.

Para elegir el modelo que describe de una forma más ajustada el comportamiento histórico de la serie de tiempo, se utilizan los criterios AIC y BIC, donde los valores menores indican un mejor ajuste a los datos. A continuación, se muestra la comparación de los criterios para cada modelo en la Tabla 1.

Tabla 1. Comparación de criterios de ajuste de los modelos construidos.

Modelos/Métrica	AIC	BIC
Modelo de Componentes	746,26	763,27
Modelo Holt- Winters	562,04	569,73
Modelo Box Jenkins	748,32	758,00

Fuente: Elaboración propia

De acuerdo con los valores de la Tabla 1, se identifica que el modelo Holt-Winters es el que cuenta con los mejores valores para el ajuste, es decir, es aquél que puede describir con menor nivel de error el comportamiento histórico de los datos de la serie de tiempo de los intentos de suicidios en Medellín.

Los indicadores obtenidos presentan valores más altos que los de estudios previos como el de Quito Yanquen y Pinzón Choque (2021), trabajo en el que se analizan casos de suicidio en la ciudad de Bogotá, y en donde se obtiene un BIC de 477. A pesar de que en el ajuste los modelos construidos en este trabajo tienen un desempeño inferior, es importante destacar que en ambos casos se estudian regiones diferentes, y que en este artículo se hace énfasis en los intentos de suicidio, lo que marca diferencias en los hallazgos.

Aunque el modelo Holt-Winters, brinda un mejor ajuste a los datos de intentos de suicidio para la ciudad de Medellín, esto no implica que sea el modelo más apropiado para realizar pronósticos, pues los criterios utilizados en este punto sólo permiten identificar la capacidad del modelo para reproducir datos históricos. Por tanto, los modelos construidos se comparan, a continuación, utilizando métricas de error en los pronósticos generados con cada uno de ellos y su contraste con el conjunto de datos de prueba

Para seleccionar el modelo que mejor pronostica, se comparan las medidas de precisión de pronóstico MAPE, MAE y RMSE. Esta comparación se presenta en la Tabla 2; donde se considera que los menores valores son mejores en cada uno de los criterios, pues implican menor diferencia entre los datos reales (conjunto de prueba) y los datos pronosticados (por cada modelo).

Tabla 2. Comparación de criterios de pronóstico de los modelos construidos

Modelos/Criterio	MAPE	MAE	RMSE
Modelo de componentes	11,96	19,37	22,52
Modelo Holt-Winters	31,37	47,29	51,30
Modelo Box-Jenkins	26,18	38,92	43,34

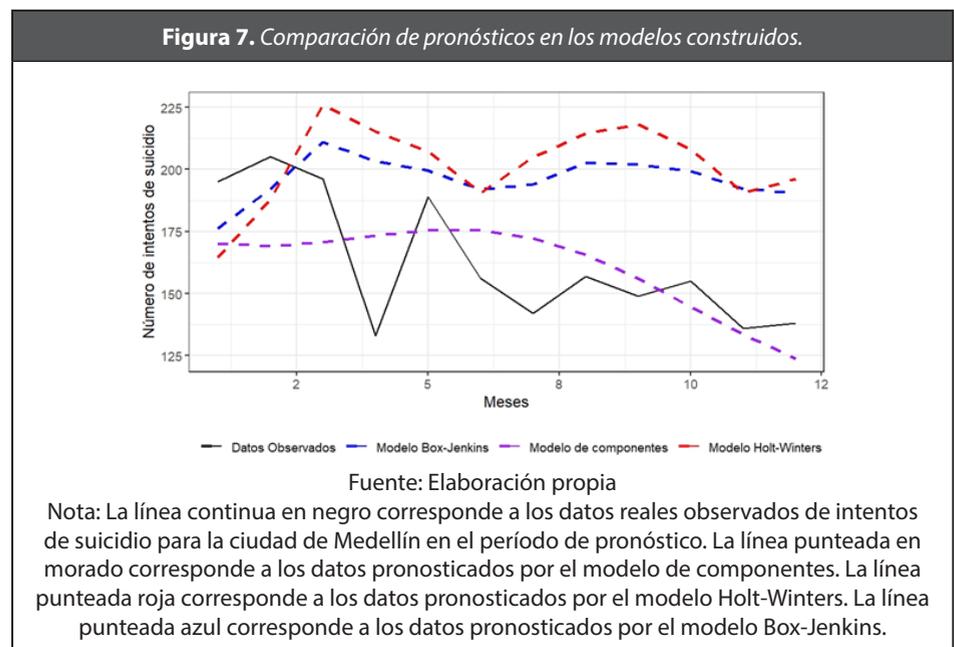
Fuente: Elaboración propia

Como el modelo de componentes cúbico estacional con errores ARMA (0,5) es el que presenta menores valores en MAPE, MAE y RMSE, se concluye que es el mejor para pronosticar datos de esta serie de tiempo.

Estos resultados indican que, para el modelo con mayor precisión, se tiene un porcentaje de error del 12% comparado con las observaciones reales, un error absoluto medio de pronóstico de 19,37 casos de intentos de suicidio, y una desviación estándar de 22,52 casos.

Al comparar los resultados con trabajos previos, se encuentra que los niveles de precisión logrados con otras metodologías son amplios, incluyendo tasas del 75% como en (Shetty M et al., 2023), o valores cercanos al 91% (Boonkwang et al., 2018). El modelo construido, ofrece entonces una aproximación viable para el caso regional de los intentos de suicidio en la ciudad de Medellín.

A continuación, se presenta la figura 7, donde se comparan gráficamente los pronósticos de los tres modelos ajustados contra los datos reales, y se evidencia cómo el modelo paramétrico cúbico estacional con errores ARMA (0,5) es el modelo que presenta una menor distancia entre los pronósticos y los datos reales de los intentos de suicidio en la ciudad de Medellín.



Los resultados de la Figura 7, muestran la existencia de una tendencia decreciente en los casos de intentos de suicidio en Medellín en el periodo utilizado para la validación de los pronósticos. Esta dinámica se encuentra mejor representada por el modelo de componentes cúbico estacional con errores ARMA (0,5). Es importante mencionar, que esta tendencia decreciente, también ha sido identificada en la literatura en otras épocas, como en el trabajo de Cardona Arango et al. (2016) para el caso de Colombia. Mientras que, en años más recientes, estos resultados contrastan con la información nacional que evidencia una tendencia creciente en las tasas de suicidio (Ministerio de Salud y Protección Social, n.d.).

Con los análisis realizados en esta sección, se considera que, la vigencia y calidad de los datos son cruciales para la selección adecuada de modelos de pronóstico y su respectivo ajuste.

Por otro lado, basados en la revisión de literatura se concluye que, usando otras herramientas metodológicas como la inteligencia artificial, el aprendizaje automático, sumado a datos demográficos combinados con información de las personas que presentaron intentos de suicidios, se podría llegar a predecir el riesgo de intento de suicidio a nivel individual con base en factores estructurales, proximales y distales.

5. Conclusión.

Las series de tiempo son de gran utilidad para describir el comportamiento de los datos y realizar pronósticos sobre lo que podría suceder con estos, si las condiciones del entorno no sufren alteraciones importantes. En este trabajo se ajustan, validan y comparan tres diferentes modelos de serie de tiempo utilizando criterios de ajuste y pronóstico, para la serie de intentos de suicidio en la ciudad de Medellín con datos mensuales en el periodo 2013 a 2020.

Para el caso de los intentos de suicidio en Medellín el modelo paramétrico cúbico estacional con errores ARMA (0,5) presentó

mejores criterios de pronóstico al comparar los datos reales con los pronosticados; describiendo un aumento paulatino en los intentos de suicidio en la ciudad de Medellín para los inicios del periodo estudiado y una tendencia a la disminución de casos cuando se está terminando el periodo analizado.

Si bien el modelo elegido captura de mejor manera la tendencia de los casos estudiados, este no logra capturar completamente las variaciones repentinas de casos durante el año 2020, debido a los cambios estructurales que asociados al periodo de cuarentena por el COVID19. En estos casos, podría ser pertinente complementar los análisis estadísticos con modelos que integren perspectivas explicativas y que favorezcan la interdisciplinariedad en los análisis.

Si bien, existen modelos previos que mostraron mejor desempeño tanto en criterios de ajuste como pronóstico, para el caso particular de Medellín no se cuenta con estudios recientes que permitan modelar y comprender el comportamiento de los intentos de suicidio en la ciudad. Por tanto, este trabajo es de gran relevancia, pues contribuye a que entes como el misterio de salud, las EPS, gobernación y alcaldía, puedan realizar estudios que caractericen el comportamiento de los casos e intentos de suicidios, y puedan diseñar planes de contingencia que permitan detener el crecimiento de los casos y reducirlos acorde con las metas nacionales e internacionales.

Trabajos futuros de esta investigación pueden incluir el análisis estadístico de los factores que podrían explicar la conducta suicida con enfoque particular en la población de la ciudad de Medellín. Así como la propuesta de modelos con otros enfoques metodológicos que contribuyan a disminuir el nivel de error obtenido en este trabajo.

6. Referencias

- Aguilar Sigueñas, L.E., 2022. Un Modelo Predictivo de la Ideación Suicida Basado en Depresión, Ansiedad, Estrés y Funcionalidad Familiar en Internos de Medicina, 2021. Trujillo.
- Alcaldía de Medellín, 2020. Unidad de Gestión de la Información y el Conocimiento [WWW Document]. Intentos de suicidio. URL <http://medata.gov.co/dataset/intento-de-suicidio> (accessed 1.14.24).
- Boonkwang, K., Kasemvilas, S., Kaewhao, S., Youdkang, O., 2018. A Comparison of Data Mining Techniques for Suicide Attempt Characteristics Mapping and Prediction, in: 2018 International Seminar on Application for Technology of Information and Communication. IEEE, pp. 488–493. <https://doi.org/10.1109/ISEMANTIC.2018.8549835>
- Cardona Arango, D., Medina-Pérez, Ó.A., Cardona Duque, D.V., 2016. Caracterización del suicidio en Colombia, 2000-2010. *Rev Colomb Psiquiatr* 45, 170–177. <https://doi.org/10.1016/j.rcp.2015.10.002>
- Chen, Q., Zhang-James, Y., Barnett, E.J., Lichtenstein, P., Jokinen, J., D'Onofrio, B.M., Faraone, S. V., Larsson, H., Fazel, S., 2020. Predicting suicide attempt or suicide death following a visit to psychiatric specialty care: A machine learning study using Swedish national registry data. *PLoS Med* 17. <https://doi.org/10.1371/journal.pmed.1003416>
- Fernandez Villacreses, C.P., 2020. Análisis estadístico para identificar factores que influyen en la conducta suicida en drogodependientes. Guayaquil.
- Fonseka, T.M., Bhat, V., Kennedy, S.H., 2019. The utility of artificial intelligence in suicide risk prediction and the management of suicidal behaviors. *Australian & New Zealand Journal of Psychiatry* 53, 954–964. <https://doi.org/10.1177/0004867419864428>
- Gómez Tabares, A.S., Núñez, C., Caballo, V.E., Agudelo Osorio, M.P., Grisales Aguirre, A.M., 2019. Predictores psicológicos del riesgo suicida en estudiantes universitarios. *Behavioral Psychology / Psicología Conductual* 27, 391–413.
- Linthicum, K.P., Schafer, K.M., Ribeiro, J.D., 2019. Machine learning in suicide science: Applications and ethics. *Behavioral Sciences & the Law* 37, 214–222. <https://doi.org/10.1002/bsl.2392>
- Madyatmadja, E.D., Jordan, S.I., Andry, J.F., 2021. Big data analysis using rapidminer studio to predict suicide rate in several countries. *ICIC Express Letters, Part B: Applications* 12, 757–764. <https://doi.org/10.24507/icicelb.12.08.757>
- Ministerio de Salud y Protección Social, 2021. Estrategia Nacional para la Prevención de la Conducta Suicida en Colombia. Bogotá.
- Ministerio de Salud y Protección Social, 2018. Boletín de salud mental Conducta suicida Subdirección de Enfermedades No Transmisibles. Bogotá.
- Ministerio de Salud y Protección Social, n.d. Bodega de Datos SISPRO (SGD) [WWW Document]. URL <https://rssvr2.sispro.gov.co/ObsSaludMental/> (accessed 6.27.24).
- Nikam, O., Singh, S., Patil, S., Nair, S., Raut, A., Lobo, V.B., More, S., Laban, R.M., 2023. A Machine Learning Prediction Model for Envisaging Future Risk of Suicide, in: 2023 5th Biennial International Conference on Nascent Technologies in Engineering (ICNTE). IEEE, pp. 1–6. <https://doi.org/10.1109/ICNTE56631.2023.10146651>
- Organización Mundial de la Salud, 2022a. Plan de acción integral sobre salud mental 2013-2030. Organización Mundial de la Salud.

- Organización Mundial de la Salud, 2022b. Salud mental: fortalecer nuestra respuesta [WWW Document]. URL <https://www.who.int/es/news-room/fact-sheets/detail/mental-health-strengthening-our-response> (accessed 1.14.24).
- Organización Mundial de la Salud, 2021. Una de cada 100 muertes es por suicidio [WWW Document]. URL <https://www.who.int/es/news/item/17-06-2021-one-in-100-deaths-is-by-suicide> (accessed 1.14.24).
- Quito Yanquen, L.C., Pinzón Choque, C.G., 2021. Modelo de pronóstico de casos de suicidios en la ciudad de Bogotá D.C. Fundación Universitaria Los Libertadores, Bogotá D.C.
- Schafer, K.M., Kennedy, G., Gallyer, A., Resnik, P., 2021. A direct comparison of theory-driven and machine learning prediction of suicide: A meta-analysis. *PLoS One* 16, e0249833. <https://doi.org/10.1371/journal.pone.0249833>
- Shetty M, S., Praveen, V., Chinmayi, H.S., K S, B., Kamal, S., Rao, T., 2023. Using Machine Learning to Forecast the Risk of Suicide Attempts Over Time, in: 2023 International Conference on Computational Intelligence for Information, Security and Communication Applications (CIISCA). IEEE, pp. 297–303. <https://doi.org/10.1109/CIISCA59740.2023.00064>
- Shumway, R.H., Stoffer, D.S., 2011. *Time Series Analysis and Its Applications*, 3rd ed, Springer Texts in Statistics. Springer New York, New York, NY. <https://doi.org/10.1007/978-1-4419-7865-3>
- Tabares, A.S., Núñez, C., Agudelo, M.P., Caballo, V.E., 2020. Riesgo suicida y su relación con la inteligencia emocional y la autoestima en estudiantes universitarios. *Terapia psicológica* 38, 403–426. <https://doi.org/10.4067/S0718-48082020000300403>
- Vanegas, G., 2022. Cada 20 minutos hay un intento de suicidio en Colombia: “Derrumbar mitos puede salvar vidas.” EL PAIS.