

## Aplicación del método ARIMA para el pronóstico del avalúo de bienes raíces

### *Application of the ARIMA Method for Real Estate Valuation Forecasting*

**Elba María Boderó Poveda\***  
Universidad Nacional de Chimborazo.  
Riobamba-Ecuador.  
ebodero@unach.edu.ec  
<https://orcid.org/0000-0003-3807-5203>

**Katherine Alexandra Guerrero Morejón**  
Universidad de Sevilla.  
Sevilla-España.  
kguerrero@us.es  
<https://orcid.org/0000-0003-0871-9901>

\*Correspondencia: ebodero@unach.edu.ec

**Cómo citar este artículo:**  
Boderó, E., & Guerrero, K. (2023). Aplicación del método ARIMA para el pronóstico del avalúo de bienes raíces. *Perspectivas Sociales y Administrativas*, 1(1), 26-43.  
<https://doi.org/10.61347/psa.v1i1.56>

**Recibido:** 30 de junio de 2023

**Proceso de evaluación:**

4 de julio al 21 de agosto de 2023

**Aceptado:** 25 de agosto de 2023

**Publicado:** 4 de septiembre de 2023

**Resumen:** El análisis de los datos es uno de los principales aspectos a considerar para la toma de decisiones en el ámbito social y empresarial. El mercado de bienes raíces es un componente decisivo y de soporte para las economías de los países a nivel mundial. Un análisis de la variación de los precios inmobiliarios a través de series temporales puede facilitar la identificación de mejores oportunidades a los inversores. Esta investigación tiene por objetivo aplicar el método ARIMA para el pronóstico del avalúo de bienes raíces. Se utiliza un enfoque cuantitativo con un diseño de investigación no experimental. Para la minería de datos se aplica el proceso KDD conjuntamente con el proceso de ARIMA para el pronóstico. Entre sus principales resultados se obtiene una base de datos limpia, la cual posteriormente es transformada a una serie estacionaria, para de esta manera ajustar el modelo y aplicar al pronóstico al caso de estudio en la ciudad de Riobamba, Ecuador. Se muestran datos con rangos de confiabilidad al 80% y 95%. El procedimiento detallado en este estudio puede aplicarse a cualquier contexto de predicción de series temporales en el sector de bienes raíces.

**Palabras clave:** ARIMA, bienes raíces, minería de datos, pronóstico, series temporales.

**Abstract:** The analysis of data is one of the main aspects to consider for decision-making in social and business contexts. The real estate market is a crucial and supportive component for the economies of countries worldwide. Analyzing the variation of real estate prices through time series can help investors identify better opportunities. This research aims to apply the ARIMA method for real estate appraisal forecasting. A quantitative approach with a non-experimental research design is used. For data mining, the KDD process is applied in conjunction with the ARIMA process for forecasting. The main results include a cleaned database, which is subsequently transformed into a stationary series to adjust the model and apply the forecast to the case study in the city of Riobamba, Ecuador. Data are presented with confidence intervals of 80% and 95%. The procedure detailed in this study can be applied to any time series prediction context in the real estate sector.

**Keywords:** ARIMA, data mining, forecasting, real estate, time series.

**Copyright:** Derechos de autor 2023 Elba María Boderó Poveda, Katherine Alexandra Guerrero Morejón.



Esta obra está bajo una licencia internacional Creative Commons Atribución-NonComercial 4.0.

## 1. Introducción

La previsión inmobiliaria ha evolucionado para convertirse en una parte esencial de la planificación y gestión estratégica en el sector. Inversores, desde fondos de pensiones hasta compañías de seguros y fondos inmobiliarios, realizan asignaciones a largo plazo con el objetivo de anticipar el rendimiento futuro de los activos en diversas áreas y sectores inmobiliarios. La necesidad de conocer proyecciones sobre alquileres, rendimientos y precios en un horizonte temporal específico es crucial para la toma de decisiones informada. La evolución de la previsión inmobiliaria se ha visto impulsada por la acumulación de datos, que abordan factores clave de valoración como ubicaciones, transporte e infraestructura (Al-Marwani, 2014; Papastamos et al., 2015).

En los últimos años, el sector inmobiliario ha implementado procesos innovadores. Las nuevas tecnologías transforman la forma de diseñar, construir, vender, arrendar y administrar bienes inmuebles. Hoy existe un acelerado crecimiento de información en las redes sociales, páginas web propias de inmobiliarias y de terceros, lo cual genera una ingente cantidad de información que aún no ha sido analizada.

Los bienes inmuebles representan activos con heterogeneidad intrínseca, ya que sus características experimentan cambios a lo largo del tiempo. Esta dinámica en la evolución de las propiedades no solo refleja la variabilidad inherente a las estructuras residenciales, sino que también destaca la importancia de comprender y adaptarse a las transformaciones que experimentan los bienes inmuebles con el transcurso de los años (Knoll et al., 2017).

La valoración precisa de bienes raíces se presenta como un desafío fundamental para profesionales y analistas. La capacidad de anticipar y entender las tendencias futuras en los precios de las propiedades no solo es crucial para los agentes inmobiliarios, tasadores e inversores, sino que también impacta directamente en las decisiones estratégicas a nivel económico y social.

La predicción de precios inmobiliarios es crucial para el establecimiento de políticas inmobiliarias y puede ayudar a propietarios de bienes raíces y agentes a tomar decisiones informadas (Pai & Wang, 2020). La idea de obtener nuevas perspectivas sobre los procesos comerciales a partir de los datos está llevando a las organizaciones a invertir en el área de análisis de datos (Pohl et al., 2022). Esto brinda oportunidades para optimizar los recursos e incrementar la competitividad.

Los cambios en los precios de las propiedades afectan aún más a los inversores, responsables políticos y bancos. Los valores de las propiedades se ven afectados principalmente por dos tipos de factores: generales y particulares. Los primeros se conciben como factores macro sociales -como tasas de inflación o de interés- que afectan la tendencia de precios del mercado inmobiliario. Por su parte, los factores particulares, que juegan un rol primordial en el avalúo inmobiliario, afectan a los precios de la vivienda, entre ellos: la ubicación o tamaño de la propiedad, el estilo de construcción, las instalaciones y servicios disponibles, el entorno y la proximidad del área urbana (Shinde & Gawande, 2018). El pronóstico de los precios de bienes raíces permite a los inversionistas tomar mejores decisiones con respecto a la adquisición de inmuebles; así también a entidades gubernamentales en relación con la planificación territorial. Con estimaciones realistas de precios, podrían obtener información sobre las tendencias de precios de la vivienda antes de llevar a cabo cualquier transacción.

La inversión en el sector inmobiliario ha sido históricamente una oportunidad crucial. Dado el variado conjunto de instrumentos financieros relacionados con activos inmobiliarios, su importancia es significativa tanto para los inversores como para los intermediarios (Kabaivanov & Markovska, 2021). En este contexto, la inteligencia artificial emerge como una herramienta estratégica fundamental para potenciar nuestra comprensión de las dinámicas del mercado inmobiliario.

Mittal et al. (2021) en India llevaron a cabo una investigación utilizando una muestra de 414 conjuntos de datos únicos de UCI sobre precios de bienes raíces. La aplicación del método de regresión multivariado OLS se apoyó en variables de control como el costo de la vivienda, la edad, la proximidad a estaciones de MRT, el número de tiendas de alojamiento, las direcciones a pie y la ubicación geográfica. El estudio evidencia que prácticamente todas las variables de control son consideradas fundamentales para predecir el precio de la vivienda. Los autores destacan que, con el empleo de la tecnología en la valoración de propiedades, ahora es viable rastrear transacciones relacionadas con la compra, venta, alquiler y préstamo, y analizar estas operaciones para desarrollar modelos predictivos destinados a la estimación de precios y servicios de tasación electrónica.

En Hong Kong, los investigadores Wang y Li (2019) llevaron a cabo un estudio en el cual seleccionaron 104 artículos para realizar una revisión sistemática de la literatura que abarcara el período de 2000 a 2018, centrándose en modelos y métodos de tasación masiva relacionadas al sector inmobiliario. Como resultado, definen la tendencia futura de la tasación masiva desde la perspectiva del modelo como "tasación masiva 2.0": un proceso que implica el establecimiento, análisis y prueba de modelos para un conjunto de propiedades a partir de una fecha específica. Esta metodología incorpora inteligencia artificial, sistemas de información geográfica y métodos híbridos para mejorar la modelación del valor inmobiliario, integrando datos tanto espaciales como no espaciales.

En Italia, Del Giudice et al. (2017) llevaron a cabo un análisis de la relación entre los precios de alquiler de propiedades y la ubicación geográfica en un área urbana. Para este propósito, emplearon algoritmos genéticos (GA). Con el fin de verificar la confiabilidad de los algoritmos genéticos en tasaciones inmobiliarias y demostrar las potencialidades predictivas de estas técnicas en el análisis de mercados inmobiliarios, aplicaron un análisis de regresión múltiple (ARM) y compararon los resultados obtenidos por GA y MRA. Entre los resultados y conclusiones destacan que el sector inmobiliario demanda valoraciones precisas de las propiedades debido a la diversidad de objetivos perseguidos por diversos actores, como controlar los valores inmobiliarios o actualizar su rentabilidad. Por estas razones, la interpretación del mercado inmobiliario se vuelve crucial junto con su evolución y dinamismo, y al mismo tiempo se requieren con urgencia técnicas adecuadas para analizar de manera apropiada las características del mercado inmobiliario.

Los autores Kang et al. (2020) realizaron un estudio en el cual su objetivo era desarrollar modelos predictivos para los precios de subasta de bienes inmuebles mediante el empleo de inteligencia artificial y enfoques estadísticos. Los modelos de predicción se construyen a través de un modelo de regresión, una red neuronal artificial y un algoritmo genético. En el análisis empírico, utilizaron datos de subastas de apartamentos en Seúl-Corea del Sur entre 2013 y 2017 para anticipar los precios de subasta y comparar la precisión de los pronósticos de los distintos modelos. En los resultados presentan que el rendimiento más destacado se observa en el modelo basado en algoritmo genético, y se resalta que la efectiva segmentación regional, basada en los precios de tasación de subasta, mejora la precisión predictiva de manera significativa.

En Indonesia, Muhammad et al (2022) llevaron a cabo un estudio destinado a prever el índice analítico de precios residenciales utilizando el método ARIMA de Box-Jenkins. Durante el período comprendido entre 2003 y 2020, se aplicó el modelo ARIMA Box-Jenkins al conjunto de datos de precios residenciales (RPI). Los resultados revelaron que los modelos ARIMA más efectivos fueron ARIMA ([4],2,1) para viviendas pequeñas, ARIMA (0.1,[3]) para viviendas medianas y ARIMA (1,1,0) para viviendas grandes. Las proyecciones basadas en el mejor modelo ARIMA arrojaron un RPI en el cuarto trimestre de 2021 de 397,62 para viviendas pequeñas, 357,60 para viviendas medianas y 315,75 para viviendas grandes.

En la ciudad de Riobamba- Ecuador, el sector inmobiliario contribuye significativamente en el desarrollo económico de la ciudad, al satisfacer necesidades básicas de vivienda, también ha sido una fuente de vulnerabilidad y crisis. Los ciclos de auge y declive en este sector han demostrado ser perjudiciales tanto para la estabilidad financiera como para la economía real. Además de cumplir con la necesidad fundamental de vivienda, el sector inmobiliario desempeña un papel crucial como componente importante de la inversión (Cejas & Romero, 2020).

El mercado inmobiliario se distingue de otros mercados financieros por diversas características fundamentales. Su marcada heterogeneidad se debe a la influencia significativa de la ubicación y los atributos físicos de una propiedad. La singularidad de cada bien inmueble añade un componente único a las transacciones, lo que contrasta con la homogeneidad que a menudo se observa en otros ámbitos financieros (Ghysels et al., 2013).

Así, la analítica de datos debe ser tomada en cuenta por aquellas empresas que cambian su modelo de negocio y lo orientan hacia la innovación, la anticipación de riesgos, el incremento de las utilidades y el desarrollo empresarial (Sofía & Adrián, s. f.). Rahmer et al. (2022) aducen que la aplicación de técnicas de analítica de datos para el pronóstico de variables económicas y financieras resulta de vital importancia debido a la incertidumbre involucrada en los procesos de planificación de ventas, operaciones y la toma de decisiones estratégicas.

Por medio de la utilización de técnicas de Machine Learning se pueden analizar datos enfocados a la plusvalía en bienes raíces, además de características sociodemográficas de la población o el tipo de propiedad que se demanda en cada zona (Mira, 2019). Kok et al. (2017) indican que las estimaciones precisas y oportunas de los valores de las propiedades resultan fundamentales para que los inversores o las instituciones que brindan financiamiento tomen decisiones informadas y definan valores de tasación y evaluación de los rendimientos en el mercado de bienes raíces comerciales.

Las series temporales como técnica de aprendizaje automático pueden predecir el comportamiento de variables o tendencias basadas en la previsión del tiempo de valores continuos (Rojas & Gomez, 2014). Una serie temporal analiza datos recopilados cronológicamente. En tanto, la minería de datos de series temporales consta de varias tareas como indexación de contenidos, detección de anomalías, clasificación, análisis y preprocesamiento, segmentación, agrupamiento y predicción (Durán, 2019).

Los datos almacenados en forma de series temporales son susceptibles a contener información valiosa para su dominio de procedencia. El análisis de series temporales presenta características distintivas que lo diferencian de otros tipos de datos. Entre estas características se encuentran alta numerosidad y dimensionalidad de las series a analizar y la constante actualización de los datos (González, 2016).

El modelo autorregresivo de media móvil (ARMA) propuesto por Box-Jenkins combina elementos de los modelos autorregresivos (AR) y de media móvil (MA). Por otro lado, un modelo ARIMA se presenta como un enfoque univariado que pretende describir una variable única como un proceso integrado autorregresivo de media móvil, donde la serie se caracteriza por completo mediante  $p$  (orden del componente AR),  $q$  (orden del componente MA) y  $d$  (orden de integración de la serie) (Awe et al., 2020).

Jenkins y Box (1976) introdujeron por primera vez el método de ajuste e identificación sistemática de los procesos (ARIMA). La siguiente sección ofrece una visión general de los modelos autoregresivos integrados de media móvil (ARIMA) y del procedimiento de Box-Jenkins, además de presentar los datos, conjuntamente organizados con lo mencionado por (Racine, 2019). El modelo ARIMA (Media Móvil Integrada Autoregresiva) es un método ampliamente conocido para la predicción de series

temporales. Se asume que los valores pasados de la serie, junto con los términos de error anteriores, contienen información para el propósito de la predicción. Las principales ventajas del modelo ARIMA incluyen que solo requiere datos de la serie temporal en cuestión y tiene una buena capacidad de predicción a corto plazo (Wang et al., 2019).

Stellwagen & Tashman (2013) destacan que bajo la condición de que se satisfagan ciertos supuestos, los modelos ARIMA generan pronósticos óptimos, lo que implica esencialmente que los errores del modelo carecen de información que pueda mejorar las predicciones. Estos modelos buscan identificar patrones en los datos históricos con el objetivo de discernir el proceso subyacente que genera e influye en el patrón histórico, conocido como el proceso de generación de datos. Se reconoce que los modelos ARIMA destacan por su robustez y eficiencia al prever series temporales financieras, especialmente en pronósticos a corto plazo, superando incluso a las técnicas de redes neuronales artificiales más populares. Estos modelos han sido extensamente aplicados en el ámbito de la economía y las finanzas (Ariyo et al., 2014).

El presente estudio busca responder a la pregunta de investigación de ¿Cómo aplicar el método ARIMA para el pronóstico del avalúo de bienes raíces?, para ello se establecen los objetivos específicos de (1) obtener una base de datos limpia para proceder al análisis a través de la selección, preprocesamiento y transformación de datos, (2) Crear un modelo ajustado mediante el método ARIMA para el pronóstico de los precios de bienes raíces de la ciudad de Riobamba, Ecuador, (3) Analizar los resultados obtenidos de la aplicación del modelo.

## 2. Metodología

El enfoque cuantitativo de la presente investigación permitió el análisis numérico de los resultados. Gracias a su diseño no experimental no fueron manipuladas las variables durante el estudio, y se precisaron valores históricos de los precios de los bienes raíces de la ciudad de Riobamba, Ecuador. En la primera fase se realizó una recopilación y revisión bibliográfica de estudios relacionados con técnicas de minería de datos utilizadas para el análisis de la plusvalía inmobiliaria (Bodero-Poveda et al., 2022). En la segunda se aplicaron procesos sistemáticos como el referido a la minería de datos Knowledge Discovery in Databases (KDD) o el conocido como Extracción de Conocimiento en Bases de Datos, detallados a continuación:

(1) Selección para crear un conjunto objetivo de datos. Se consideró el precio del bien inmueble y el número de metros cuadrados de construcción. La recolección de datos se realizó desde 2015 hasta los dos primeros meses de 2022 desde las fuentes: Nuroa, OLX, Mercadolibre, Remax, Megainmobiliaria, Confianza Inmobiliaria; Publicaciones de Market Place y grupos inmobiliarios de Facebook; fanpage de Facebook de las inmobiliarias Edifika Habitar S. A., Inmobiliaria Avalia, Grupo inmobiliario Horizonte, Gabriela Bermeo Inmobiliaria, Inmobiliaria Yépez, Vida Inmobiliaria, Gestión Inmobiliaria, Ivon Alvarado Pinto Bienes Raíces y Puruwa Home.

(2) Preprocesamiento. Se ejecutó el proceso de calidad de datos, se aplicaron operaciones básicas como la identificación y eliminación de datos atípicos que podían afectar al modelo y se completaron los datos faltantes, necesarios para que existiera una temporalidad secuencial.

(3) Transformación. Se representaron los datos como una serie de tiempo a partir de la división del precio del bien inmueble dividido para el número de metros cuadrados de construcción. Se obtuvo el promedio de los precios de las propiedades por periodo mensual, cuidando que todos los periodos presentarán condiciones iguales en tipo de propiedad y número de bienes.

(4) Minería de datos. Se efectuó a través del proceso para la aplicación de ARIMA (Box & Jenkins, 1976; Racine, 2019).

(5) Interpretación y evaluación. Se consolidó el conocimiento descubierto a partir de la predicción del avalúo inmobiliario.

**Figura 1**

*Proceso metodológico*



Para el desarrollo del proceso, se exponen de manera concisa los conceptos fundamentales que sirvieron como cimiento para la ejecución de las tareas pertinentes. Asimismo, se presentan las fórmulas clave que desempeñaron un papel crucial en el análisis y la resolución de los problemas planteados. Este enfoque permite proporcionar una visión clara y estructurada de los principios teóricos y las herramientas matemáticas empleadas, facilitando así la comprensión integral del proceso llevado a cabo. Estos elementos, debidamente fundamentados, sirven como base sólida para la validez y la coherencia de los resultados obtenidos, contribuyendo a la transparencia y robustez del trabajo realizado.

## 2.1. Series Temporales

### 2.1.1. Datos de series temporales

Para entender matemáticamente las series temporales se debe considerar a una muestra de datos ordenados en el tiempo  $Y_1, Y_2, Y_3, \dots, Y_T$ , que se escribe como  $\{Y_t\}_{t=1}^T$ , para hacer un pronóstico de la serie en algún período futuro. Se denota tal pronóstico como:

$$\hat{Y}_{T+h} = E Y_{T+h} | \{Y_t\}_{t=1}^T \quad (1)$$

donde  $h < 0$  es un número entero que denota el número de días, semanas, años, para los cuales se desea construir una predicción en el tiempo  $T$ ; esto se realiza con  $Y_{T+h}$ . El término  $Y_{T+h} | \{Y_t\}_{t=1}^T$ , representa el valor esperado de la variable aleatoria  $Y$  llevada  $h$  pasos adelante (p. 8).

### 2.1.2. Patrones en series temporales

Al describir datos de series de tiempo resulta importante considerar los términos:

Tendencia: si existe un aumento o disminución constante a lo largo del tiempo.

Estacionalidad: cuando una serie de tiempo presenta una variación periódica por factores temporales como la época del año, trimestre, mes o día de la semana.

Patrón cíclico: cuando los datos exhiben subidas y bajadas que no son de un período fijo. (p. 9)

### 2.1.3. Serie estacionaria versus no estacionaria

La noción de estacionario juega un papel fundamental en el estudio y modelado de datos de series temporales. Se dice que una serie es estacionaria si sus propiedades no dependen del tiempo en que se observa, es decir, una serie que contiene tendencia, no es estacionaria. Por el contrario, un proceso bastante ruidoso es estacionario, porque no importa cuándo se observe, este será similar en cualquier momento. Entonces al modelar datos de una serie de tiempo se debe suponer que el proceso que generó los datos pasados es estable, para que así se pueda estimar significativamente los valores futuros.

Más formalmente, se dice que una serie de tiempo es estacionaria si la distribución de probabilidad conjunta de  $\{Y_t\}_{t=1}^T$  es invariante con respecto al tiempo. En otras palabras, si se desplaza la serie hacia adelante o hacia atrás por  $k$  rezagos, es decir,  $k$  es cualquier número entero positivo o negativo, esto no afectaría a la distribución conjunta de la serie, se denota:

$$f(Y_1, Y_2, Y_3, \dots, Y_T) = f(Y_{k+1}, Y_{k+2}, Y_{k+3}, \dots, Y_{k+T}) \quad (2)$$

donde  $f(Y_1, Y_2, Y_3, \dots, Y_T)$  es la función de densidad de probabilidad conjunta de  $\{Y_t\}_{t=1}^T$ . Se puede pensar en estacionalidad, cuando el proceso estocástico subyacente es, en cierto sentido, estable e invariable a lo largo del tiempo.

Una consecuencia de la estacionariedad es que las medias, varianzas y covarianzas incondicionales de la serie son invariantes en el tiempo, es decir, no dependen ni varían con el índice  $t$ . La media incondicional ( $\mu_Y$ ), la varianza  $\sigma_Y^2$  y la covarianza en el retraso  $k$  ( $\gamma_k$ ) para un proceso estacionario no dependen de  $t$  y se definen como:

$$E[Y_t] = \mu_Y \quad (3)$$

$$Var[Y_t] = E[(Y_t - \mu_Y)^2] = \sigma_Y^2 \quad (4)$$

$$Cov[Y_t, Y_{t+k}] = E[(Y_t - \mu_Y)(Y_{t+k} - \mu_Y)] = \gamma_k \quad (5)$$

(pp.9-12)

### 2.1.4. La función de autocorrelación (ACF)

Como precisa Racine (2019):

Por lo general, es imposible obtener una descripción completa de un proceso estocástico, en otras palabras, especificar realmente la función de densidad de probabilidad conjunta subyacente. Sin embargo, una medida bastante útil de cuánta correlación existe entre puntos de datos vecinos separados por  $k$  períodos de la serie, es la función de autocorrelación que se define como:

$$\begin{aligned} \rho_k &= \frac{E[(Y_t - \mu_Y)(Y_{t+k} - \mu_{Y_{t+k}})]}{\sqrt{E[(Y_t - \mu_Y)^2] E[(Y_{t+k} - \mu_{Y_{t+k}})^2]}} \\ &= \frac{Cov(Y_t, Y_{t+k})}{\sqrt{V(Y_t)} \sqrt{V(Y_{t+k})}} \\ &= \frac{Cov(Y_t, Y_{t+k})}{\sigma_{Y_t} \sigma_{Y_{t+k}}} \end{aligned} \quad (6)$$

Donde  $\sigma_{Y_t}^2$  y  $\sigma_{Y_{t+k}}^2$  son las varianzas incondicionales de  $Y_t$  y  $Y_{t+k}$ , respectivamente, y  $Cov(Y_t, Y_{t+k})$  es la covarianza entre  $Y_t$  y  $Y_{t+k}$  (pp. 14-15).

### 2.1.5. La función de autocorrelación muestral

Por supuesto, el ACF es puramente teórico. En la práctica, solo se tiene un número limitado de observaciones que se puede usar para calcular una estimación del ACF llamada función de autocorrelación muestral dada por:

$$\hat{\rho}_k = \frac{\sum_{t=1}^{T-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^T (Y_t - \bar{Y})^2} \quad (7)$$

(Racine, 2019, p.16)

### 2.1.6. No estacionariedad y diferenciación

Una serie de tiempo no estacionaria debe transformarse en estacionaria antes de que pueda modelarse. Resulta que el ACF de muestra se puede usar para identificar una serie estacionaria, y una forma de transformar una serie no estacionaria en estacionaria, es calcular las diferencias entre observaciones consecutivas, un proceso conocido como diferenciación lineal, y luego examinar el ACF de la serie diferenciada. Para una serie estacionaria, el ACF caerá rápidamente a cero, mientras que el ACF para una serie no estacionaria decaerá muy lentamente.

La diferencia de primer orden se calcula mediante:

$$\Delta Y_t = Y_t - Y_{t-1} \quad (8)$$

mientras que la diferenciación de segundo orden se calcula mediante:

$$\begin{aligned} \Delta^2 Y_t &= \Delta Y_t - \Delta Y_{t-1} \\ &= (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) \\ &= Y_t - 2Y_{t-1} + Y_{t-2} \end{aligned} \quad (9)$$

Una diferencia estacional es la diferencia entre una observación y la observación correspondiente del año anterior, por lo tanto:

$$\Delta Y_t = Y_t - Y_{t-m} \quad (10)$$

Donde  $m$  denota el número de estaciones en un año. Para datos trimestrales  $m = 4$ , mensuales  $m = 12$ , etc. Estos también se conocen como diferencias de retraso-  $m$  y se resta la observación que tiene un retraso de  $m$  períodos. Para distinguir las diferencias estacionales del estándar, es decir, las primeras diferencias, se acostumbra referirse a las diferencias estándar como primeras diferencias, es decir, diferencias en lag=1 (Racine, 2019, pp.16-17).

En el análisis de series de tiempo es habitual recurrir a la prueba Dickey-Fuller (DF) con el propósito de probar si los datos responden a un proceso autorregresivo de raíz unitaria. La prueba DF depende de supuestos sobre el proceso estocástico bajo la hipótesis nula (Lizarazu-Alanez et al., 2007). Presenta una hipótesis que si el valor resultante es menor a 0,05 la serie es estacionaria, en caso contrario la serie no es estacionaria.

$$\nabla Y_t = (p - 1)Y_{t-1} + U_t = \delta Y_{t-1} + U_t \quad (11)$$

### 2.1.7. Modelos de medias móviles integradas autorregresivas no estacionales (ARIMA (p, d, q))

Racine (2019) también especifica lo siguiente:

Si un proceso homogéneo no estacionario de orden  $d$ ,  $w_t = \Delta^d Y_t$  donde  $\Delta^d$  denota la  $d$ -ésima diferencia de una serie, ha sido generado por un proceso de promedio móvil autorregresivo mixto de órdenes  $p$  y  $q$ , entonces se puede expresar como:

$$\Delta^d Y_t = \delta + \phi_1 \Delta^d Y_{t-1} + \dots + \phi_p \Delta^d Y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q} \quad (12)$$

O como:

$$w_t = \delta + \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q} \quad (13)$$

donde  $\epsilon_t \sim (0, \sigma_\epsilon^2)$  es i.i.d. También se puede expresar  $w_t$  usando el operador de desplazamiento hacia atrás  $B$  como  $w_t = (1 - B)^d Y_t$ .

Se dice que  $Y_t$  es  $I(d)$  o integrado de orden  $d$ ; si  $dy/dt = x$ , entonces  $Y$  es la integral de  $x$ . En series de tiempo discretas, si  $\Delta Y_t = x_t$ , entonces  $Y$  también podría verse como la integral, o suma sobre  $t$ , de  $x$ . Usando el operador de desplazamiento hacia atrás se puede expresar el modelo ARIMA ( $p, d, q$ ) como:

$$w_t = \delta + \phi_1 B w_t + \dots + \phi_p B^p w_t + \epsilon_t - \theta_1 B \epsilon_t - \dots - \theta_q B^q \epsilon_t \quad (14)$$

que se puede reorganizar para obtener:

$$(1 - \phi_1 B - \dots - \phi_p B^p) w_t = \delta + (1 - \theta_1 B - \dots - \theta_q B^q) \epsilon_t \quad (15)$$

que a menudo se escribe como:

$$\Phi(B) w_t = \delta + \Theta(B) \epsilon_t \quad (16)$$

o como:

$$w_t = \frac{\delta + \Theta(B) \epsilon_t}{\Phi(B)} \quad (17)$$

donde  $\Phi(B)$  y  $\Theta(B)$  son polinomios de orden  $p$  y  $q$  en el operador de desplazamiento hacia atrás  $B$ . Esta es una forma parsimoniosa de escribir un modelo ARIMA ( $p, d, q$ ) (pp. 55-56).

### 2.1.8. Diagnósticos para modelos ARIMA ( $p, d, q$ )

#### a) Residuos del modelo

Si un modelo proporciona un ajuste adecuado a los datos, los residuos del modelo deberían ser ruido blanco. Hay una función útil de pronóstico del paquete R llamada `checkresiduals`, una función genérica que traza los residuos, la función de autocorrelación de los residuos para todos los retrasos hasta `gof.lag` y un histograma.

**a) Ljung-box**

Es un test que evalúa la autocorrelación en las series teniendo en cuenta el efecto conjunto de la correlación de la misma serie con distintos rezagos. El test evalúa la hipótesis de que el ruido de la serie no muestra correlación para un número de rezagos establecidos (ruido blanco). Para la prueba de Ljung-Box se utiliza comúnmente ARIMA (Bartlett, 2022). Presenta una hipótesis que si el valor encontrado es menor a 0,05 existe ruido blanco, por lo tanto, la serie se puede predecir, indicando una media igual a 0, varianza constante y no estar serialmente correlacionada; en caso contrario el pronóstico de la serie obtendrá valores no confiables.

$$Q = n(n + 2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n-k} \tag{18}$$

**3. Resultados**

**3.1. Selección**

Se recolectaron 2953 registros desde 2015, se consideraron los tipos de propiedades: casas, edificios, departamentos y locales comerciales; los cuales poseen la información del metraje de construcción. Los valores de los precios han sido divididos por el metraje de construcción, para obtener el precio por metro cuadrado de construcción.

**3.2. Preprocesamiento**

Se corrigieron los datos, debido a que la recolección fue ejecutada por varios sujetos. Se evidenciaron problemas de utilización errónea de caracteres como punto y coma en cifras decimales de los precios de los bienes inmuebles, además de errores de digitalización de cifras, entre otros.

Se determinó un promedio del precio del metro cuadrado de construcción por meses, es decir, existieron 12 valores por cada año, lo que significó 84 periodos de estudio, más los dos primeros meses de 2022, para un total de 86 periodos de estudio.

Varios errores fueron corregidos durante esta etapa, principalmente los datos vacíos en los periodos, que representaron el 10 % del total de los meses en estudio; estos fueron completados con el promedio del mes anterior y siguiente. Adicionalmente, se encontró un valor atípico, el cual pertenecía a una propiedad sin precedentes cuyo valor sobrepasaba los límites históricos de venta de propiedades, lo que significó un incremento a más de \$700 el promedio de metro cuadrado en este periodo, por lo que esta propiedad fue retirada del estudio, debido a que provocaba un resultado fuera de la realidad del modelo de series temporales.

**3.3. Transformación**

Una vez definidos los valores de la base de datos limpios y completos, se procedió a transformar los datos a series temporales mediante el software R Studio; se obtuvo la siguiente serie (ver Tabla 1):

**Tabla 1**

*Serie de datos por mes y año*

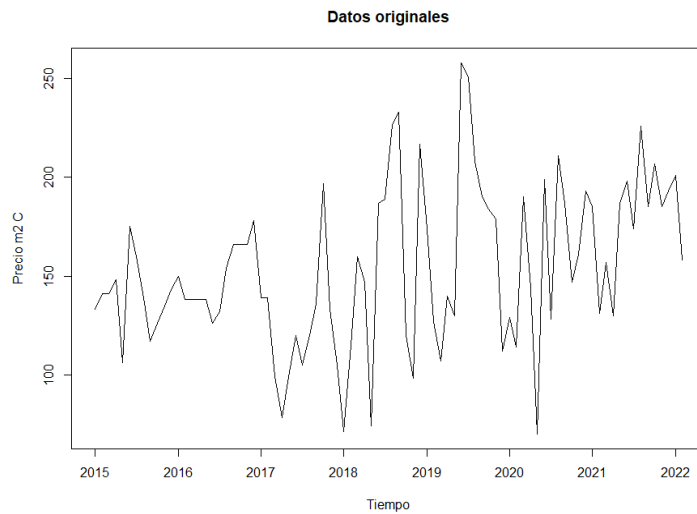
Años	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2015	133	141	141	148	106	175	160	139	117	126	135	143
2016	150	138	138	138	138	126	132	154	166	166	166	178
2017	139	139	99	78	99	120	105	120	136	197	133	105
2018	71	115	160	147	74	187	189	227	233	120	98	217

2019	174	127	107	140	130	258	251	208	190	184	179	112
2020	129	114	190	144	70	199	128	211	187	147	161	193
2021	185	131	157	130	187	198	174	226	185	207	185	193
2022	201	158										

En la Figura 2 se observa la serie de tiempo original de los datos de los bienes inmuebles de la ciudad de Riobamba, en el eje X se encuentra el tiempo desde 2015 hasta los dos primeros meses de 2022 y en el eje Y el "Precio m2 C" que representa el precio por metro cuadrado de construcción.

**Figura 2**

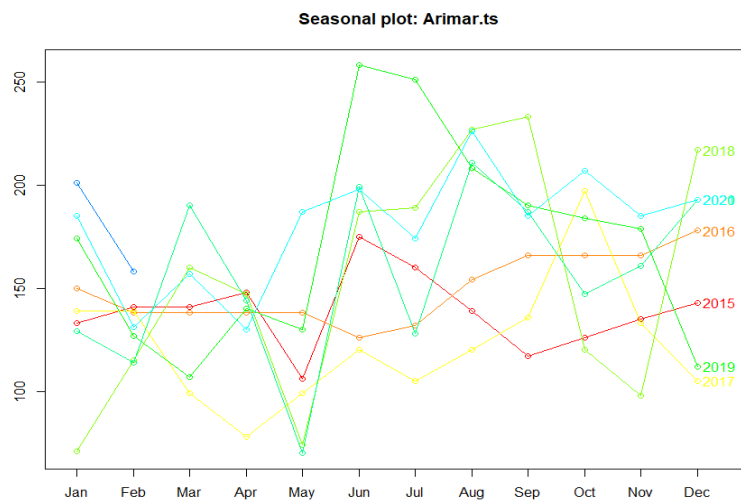
*Serie temporal original de los bienes inmuebles de la ciudad de Riobamba*



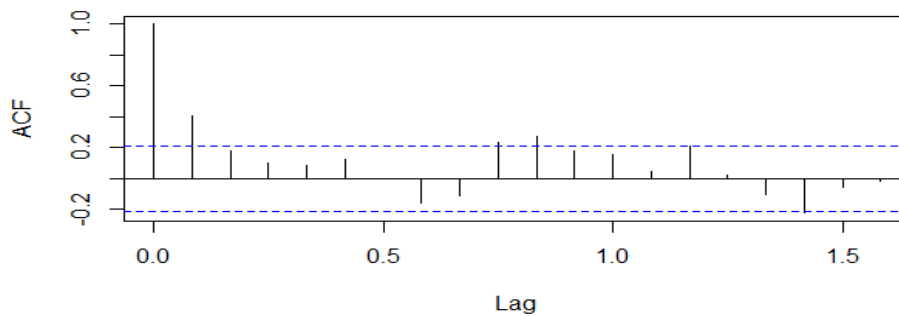
Se realizó una prueba de estacionalidad como requisito para predecir una serie temporal mediante el método de ARIMA. En las Figuras 3 y 4 se aprecia la estacionalidad de las series de tiempo por año. Gráficamente se observa cierta estacionalidad, sin embargo, esta fue comprobada mediante la función de autocorrelación (ACF), la cual indica que debe existir 1 diferencia para que la serie sea estacional.

**Figura 3**

*Estacionalidad por año del precio de los bienes inmuebles de la ciudad de Riobamba*

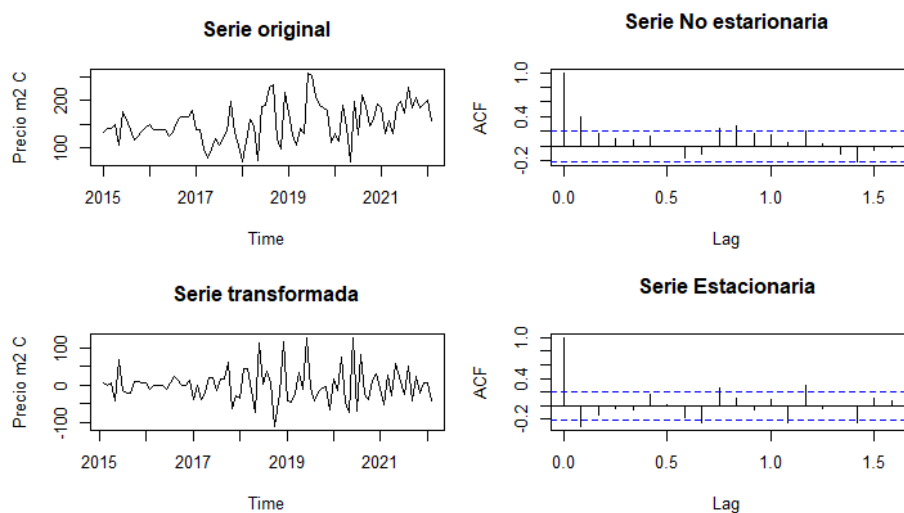


**Figura 4**  
*Estacionalidad con función de autocorrelación*



Se ejecutó la diferenciación mediante el comando `diff` de la serie, para convertirla en estacionaria; una diferencia estacional se establece entre una observación y la observación correspondiente del año anterior. Se aplica en este caso la primera diferencia (diferencia en  $\text{lag}=1$ ), debido a que se comprobó que al aumentar las diferencias el modelo no mejoró. Además, el comando `ndiffs` de R permite encontrar el número de diferencias necesarias para que la serie se convierta en estacionaria. En la Figura 5 se puede observar la serie original y la serie transformada, el análisis de sus retrasos en la serie No estacionaria, y en la estacionaria, respectivamente; esta última indica que sus valores se mueven alrededor de 0. Cabe indicar que al aplicar el logaritmo a la serie esta no se volvió estacionaria, por lo que este método fue descartado en este estudio. Se aplica la prueba de Dickey-Fuller, con un p-valor de 0,01 al ser menor a 0,05 se comprueba la hipótesis de que la serie es estacionaria después de la transformación y permite proceder al ajuste del modelo. Además, se realiza nuevamente la comprobación de las diferencias necesarias dando como resultado 0.

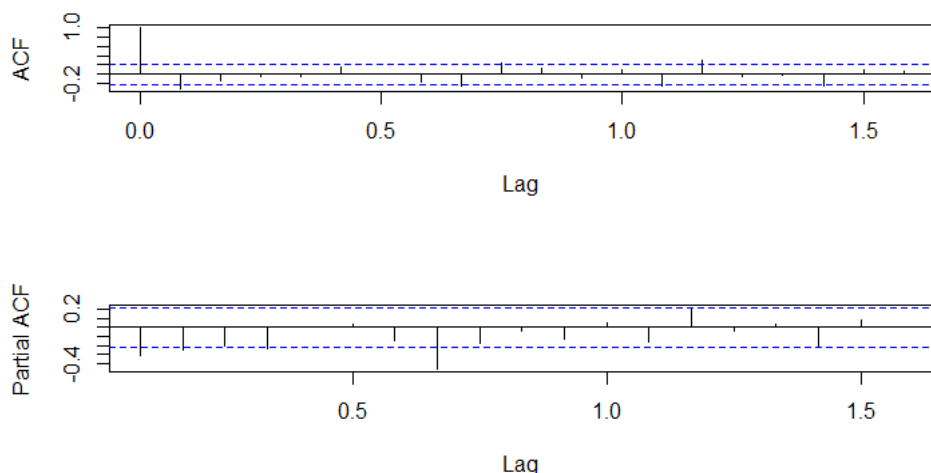
**Figura 5**  
*Comparación de series estacionaria y no estacionaria*



En la Figura 6 puede observarse el número de retrasos por ciclo que debe asignarse al modelo, sus medias móviles y el rezago autorregresivo; además, se utilizó `autoarima` para comprobar estos valores. Una vez realizado este análisis se implementa al modelo 1 media móvil, 1 diferencia y 1 rezago autorregresivo. Se puede observar los residuales estandarizados, la función de autocorrelación residual y el ruido blanco a través de un p-valor de 0,91 (test de Box-Ljung), los cuales indican que el modelo tiene una media igual a cero con una varianza constante y que esta no se encuentra serialmente correlacionada, por lo que el modelo puede utilizarse para la predicción.

**Figura 6**

*Análisis autorregresivo de la serie temporal*

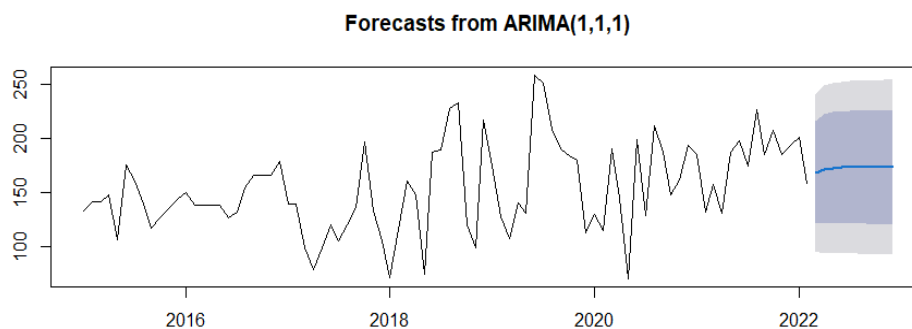


### 3.4. Minería de datos

Se aplica el modelo ARIMA, y se realiza el pronóstico de los precios inmobiliarios por metro cuadrado de construcción. La Figura 7 muestra gráficamente el color azul que representa la predicción y el gris oscuro los límites superior e inferior con una confiabilidad del 80 %, de gris más claro se encuentran los valores para la confiabilidad del 95 %, en los cuales se puede observar que el rango se encuentra más amplio.

**Figura 7**

*Predicción de la plusvalía inmobiliaria en Riobamba, Ecuador*



### 3.5. Interpretación y evaluación

En la Tabla 2 se muestra el pronóstico para 10 periodos después, los valores de predicción indican el precio en dólares promedio de los bienes raíces de la ciudad de Riobamba, Ecuador (pronóstico), de igual manera valores de límite inferior y superior para una confiabilidad del 80 % y 95 %. En el pronóstico el valor más alto se da en los meses de septiembre a diciembre de 2022, mientras que el valor más bajo en el mes de marzo. Así también, el límite inferior más bajo es de USD \$121 y el límite superior más alto es de USD \$226, con una confiabilidad del 80 %; mientras que con una confiabilidad del 95 % el límite inferior baja hasta USD \$92 y el límite superior asciende hasta USD \$254.

**Tabla 2**

*Valores de predicción*

Mes-Año	Pronóstico	Confiabilidad 80 %		Confiabilidad 95 %	
		Límite inferior	Límite superior	Límite inferior	Límite superior
Mar-2022	167,98	120,65	215,31	95,60	240,37
Apr-2022	171,25	120,46	222,04	93,57	248,93
May-2022	172,32	120,74	223,90	93,44	251,20
Jun-2022	172,67	120,76	224,58	93,28	252,06
Jul-2022	172,79	120,65	224,92	93,05	252,52
Aug-2022	172,82	120,49	225,16	92,79	252,86
Sep-2022	172,84	120,32	225,35	92,52	253,15
Oct-2022	172,84	120,14	225,54	92,24	253,44
Nov-2022	172,84	119,96	225,72	91,97	253,71
Dec-2022	172,84	119,78	225,90	91,69	253,99

#### 4. Discusión

Los resultados del estudio, subrayan la importancia de emplear enfoques de series temporales, particularmente haciendo uso del método ARIMA, como una herramienta clave para prever los precios de los bienes raíces (Han et al., 2018). Las series temporales y el método ARIMA son preferidos en la predicción de precios inmobiliarios debido a su capacidad para capturar patrones temporales, modelar tendencias y estacionalidades, manejar datos irregulares y ofrecer herramientas de evaluación robustas (Castán-Lascorz et al., 2022; Verma et al., 2021). Su versatilidad y eficacia generalizada los convierten en herramientas clave para anticipar con precisión las variaciones en los precios del sector inmobiliario.

De manera similar a los estudios previos relacionados se llevaron a cabo varias etapas, desde la recolección de datos hasta la aplicación del modelo de pronóstico, con el objetivo de proporcionar una herramienta útil para la toma de decisiones en el ámbito inmobiliario. Como la investigación realizada por Wang et al. (2019) un aspecto relevante es la corrección de errores en los datos, incluyendo la resolución de problemas de utilización errónea de caracteres y la digitalización incorrecta de cifras. Además, se abordó la presencia de datos vacíos mediante la imputación de valores promedio del mes anterior y siguiente, y se eliminó un valor atípico que afectaba significativamente el promedio del metro cuadrado. Estas acciones fueron cruciales para garantizar la calidad y la coherencia de los datos utilizados en el análisis.

El enfoque en la estacionalidad de la serie temporal se destaca en la transformación de los datos, utilizando el método de diferenciación para lograr la estacionariedad y verificando la estacionariedad mediante la prueba de Dickey-Fuller (Afeef et al., 2018). Esta preparación cuidadosa de los datos es fundamental para la efectividad del modelo ARIMA (Stevenson, 2007). En cuanto a los resultados del pronóstico, se presenta un análisis detallado de la aplicación del modelo, ajustando la función de autocorrelación y utilizando la prueba de Ljung-Box para verificar la presencia de ruido blanco en los residuos (Moffat & Akpan, 2019). Además, se proporciona una proyección a 10 periodos con intervalos de confianza del 80 % y 95 %, lo que brinda información valiosa para la toma de decisiones con diferentes niveles de confianza.

Relacionando estos resultados con la literatura internacional, se observa que en diferentes países se están aplicando diversas metodologías para pronosticar precios inmobiliarios. La inteligencia artificial

emerge como una herramienta estratégica, como se evidencia en estudios en India (Mittal et al., 2021), Hong Kong (D. Wang & Li, 2019), Italia (Del Giudice et al., 2017) e Indonesia (Muhammad et al., 2022). En estos lugares, se han utilizado regresiones multivariadas, algoritmos genéticos, redes neuronales artificiales y métodos ARIMA para abordar la complejidad del mercado inmobiliario.

Este estudio no solo es relevante para la ciudad de Riobamba, sino que también sirve como una guía general para la aplicación de pronósticos en el ámbito inmobiliario. Se destaca la escasez de literatura en este tema en Sudamérica, subrayando la contribución única de este estudio. Además, se menciona la necesidad de futuras investigaciones que aborden la aplicación del pronóstico por sector de ubicación, la comparación con otros métodos de predicción como redes neuronales artificiales y la correlación con variables macroeconómicas como la inflación.

## 5. Conclusiones

En la etapa de selección se realizó la recolección de varias fuentes y creación del conjunto de datos objetivo, se consideró el precio del inmueble y el número de metros cuadrados de construcción. En la etapa de preprocesamiento se ejecutó un proceso de verificación de la calidad de los datos, se obtuvo el promedio por periodo de las propiedades y se completaron los datos faltantes, adicionalmente, se aplicó la eliminación de datos atípicos para que no afectaran al pronóstico del modelo. Por su parte, en la etapa de transformación se verificó que la serie fuera estacionaria de manera gráfica y se analizó cuántas diferencias requirió la serie para que pudiera convertirse en estacionaria. Se aplicó el método de diferenciación y se transformó la serie en estacionaria, se pudo verificar mediante la aplicación de la prueba de Dickey-Fuller.

Una vez que la serie se convierte en estacionaria se creó el modelo de pronóstico, ajustado mediante la función de autocorrelación para inferir en la media móvil y el componente autoregresivo, además mediante autoarima se analizaron y comprobaron los valores de ajuste; se aplicó la prueba de Ljung-Box, que dio como resultado en sus residuos, ruido blanco. El modelo fue aplicado a los precios de bienes raíces promedio de la ciudad de Riobamba, el pronóstico a 10 periodos proporcionó una cifra de predicción del precio del avalúo inmobiliario, se muestran datos del 80 % y 95 % de confianza con un rango con base en su límite inferior y superior.

Esta investigación aplicó una serie temporal mediante el método ARIMA a datos inmobiliarios, existe poca literatura e información en el ámbito y en Sudamérica son prácticamente nulos los estudios científicos en el contexto de bienes raíces. Este estudio funge como una guía para recolectar, procesar, analizar y pronosticar datos inmobiliarios, y así poder tomar decisiones en cuanto al valor que un empresario o comprador puede pagar por un bien inmueble, por lo que resulta de utilidad para la negociación considerar los límites superior e inferior con la confiabilidad dada, dependiendo del perfil del inversionista al 85 % o al 95 %. El procedimiento detallado en esta investigación puede aplicarse para obtener un pronóstico de datos inmobiliarios en cualquier contexto, país o ciudad. Un aspecto clave para futuras investigaciones es la aplicación del pronóstico por sector de ubicación, la comparación con otros métodos de predicción como redes neuronales artificiales y la correlación con variables macroeconómicas como la inflación.

## Referencias

- Afeef, M., Ihsan, A., & Zada, H. (2018). *Forecasting Stock Prices through Univariate ARIMA Modeling*. 13, 130-142.
- Al-Marwani, H. A. (2014). *An approach to modeling and forecasting real estate residential property market* [Doctoral thesis]. <http://bura.brunel.ac.uk/handle/2438/10005>
- Ariyo, A. A., Adewumi, A. O., & Ayo, C. K. (2014). Stock Price Prediction Using the ARIMA Model. *2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation*, 106-112. <https://doi.org/10.1109/UKSim.2014.67>
- Awe, O., Okeyinka, A., & Fatokun, J. O. (2020). An Alternative Algorithm for ARIMA Model Selection. *2020 International Conference in Mathematics, Computer Engineering and Computer Science (ICMCECS)*, 1-4. <https://doi.org/10.1109/ICMCECS47690.2020.246979>
- Bartlett, M. (2022). *Modelización de Series Univariantes: (S)ARIMA | Series de Tiempo*. [https://bookdown.org/victor\\_morales/SeriesdeTiempo/modelizaci%C3%B3n-de-series-univariantes-sarima.html](https://bookdown.org/victor_morales/SeriesdeTiempo/modelizaci%C3%B3n-de-series-univariantes-sarima.html)
- Bodero-Poveda, E., Morales-Alarcón, C., Congacha-Aushay, A. E., & Ramos-Araujo, C. E. (2022). Técnicas de minería de datos para el análisis de la plusvalía inmobiliaria. *Dominio de las Ciencias*, 8(1), Article 1. <https://dominiodelasciencias.com/ojs/index.php/es/article/view/2531>
- Box, G. E. P., & Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*. Holden-Day.
- Castán-Lascorz, M. A., Jiménez-Herrera, P., Troncoso, A., & Asencio-Cortés, G. (2022). A new hybrid method for predicting univariate and multivariate time series based on pattern forecasting. *Information Sciences*, 586, 611-627. <https://doi.org/10.1016/j.ins.2021.12.001>
- Cejas, M., & Romero, M. (2020). *Negociación e inversión: Binomio clave en el sector comercial de Riobamba — Ecuador*. Editorial CIDE. <https://repositorio.cidecuador.org/bitstream/123456789/230/1/negociacion-e-inversion-binomio-clave-en-el-sector-comercial-de-riobamba-ecuador.pdf>
- Del Giudice, V., De Paola, P., & Forte, F. (2017). Using Genetic Algorithms for Real Estate Appraisals. *Buildings*, 7(2), Article 2. <https://doi.org/10.3390/buildings7020031>
- Durán, A. (2019). *Time series data mining: Preprocessing, analysis, segmentation and prediction. Applications* [Universidad de Córdoba, UCOPress]. <http://helvia.uco.es/xmlui/handle/10396/18590>
- Ghysels, E., Plazzi, A., Valkanov, R., & Torous, W. (2013). Chapter 9—Forecasting Real Estate Prices. En G. Elliott & A. Timmermann (Eds.), *Handbook of Economic Forecasting* (Vol. 2, pp. 509-580). Elsevier. <https://doi.org/10.1016/B978-0-444-53683-9.00009-8>
- González, M. (2016). *Clasificación semi-supervisada de series temporales* [Tesis Doctoral, Universidad de Granada]. <https://digibug.ugr.es/handle/10481/44303>
- Han, S., Ko, Y., Kim, J., & Hong, T. (2018). Housing Market Trend Forecasts through Statistical Comparisons based on Big Data Analytic Methods. *Journal of Management in Engineering*, 34(2), 04017054. [https://doi.org/10.1061/\(ASCE\)ME.1943-5479.0000583](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000583)
- Kabaivanov, S., & Markovska, V. (2021). Artificial intelligence in real estate market analysis. *AIP Conference Proceedings*, 2333, 030001. <https://doi.org/10.1063/5.0041806>

- Kang, J., Lee, H. J., Jeong, S. H., Lee, H. S., & Oh, K. J. (2020). Developing a Forecasting Model for Real Estate Auction Prices Using Artificial Intelligence. *Sustainability*, 12(7), Article 7. <https://doi.org/10.3390/su12072899>
- Knoll, K., Schularick, M., & Steger, T. (2017). No Price Like Home: Global House Prices, 1870-2012. *American Economic Review*, 107(2), 331-353. <https://doi.org/10.1257/aer.20150501>
- Kok, N., Koponen, E.-L., & Martínez-Barbosa, C. A. (2017). Big Data in Real Estate? From Manual Appraisal to Automated Valuation. *The Journal of Portfolio Management*, 43(6), 202-211. <https://doi.org/10.3905/jpm.2017.43.6.202>
- Lizarazu-Alanez, E., Villaseñor-Alva, J. A., Lizarazu-Alanez, E., & Villaseñor-Alva, J. A. (2007). Efectos de rompimientos bajo la hipótesis nula de la prueba dickey-fuller para raíz unitaria. *Agrociencia*, 41(2), 193-203.
- Mira, S. (2019). *Digitalización y tendencias del sector inmobiliario* [Master thesis, Universitat Politècnica de Catalunya]. <https://upcommons.upc.edu/handle/2117/172286>
- Mittal, R., Kumar, P., Mittal, A., & Malik, V. (2021). Developing an Evaluation Model for Forecasting of Real Estate Prices. En A. Choudhary, A. P. Agrawal, R. Logeswaran, & B. Unhelkar (Eds.), *Applications of Artificial Intelligence and Machine Learning* (pp. 619-627). Springer. [https://doi.org/10.1007/978-981-16-3067-5\\_46](https://doi.org/10.1007/978-981-16-3067-5_46)
- Moffat, I. U., & Akpan, E. A. (2019). White Noise Analysis: A Measure of Time Series Model Adequacy. *Applied Mathematics*, 10(11), Article 11. <https://doi.org/10.4236/am.2019.1011069>
- Muhammad, A. P., Wibowo, W., & Prastuti, M. (2022). Forecasting of analytic residential price index using ARIMA Box-Jenkins. *AIP Conference Proceedings*, 2668(1), 070019. <https://doi.org/10.1063/5.0116915>
- Pai, P.-F., & Wang, W.-C. (2020). Using Machine Learning Models and Actual Transaction Data for Predicting Real Estate Prices. *Applied Sciences*, 10(17), Article 17. <https://doi.org/10.3390/app10175832>
- Papastamos, D., Matysiak, G., & Stevenson, S. (2015). Assessing the accuracy and dispersion of real estate investment forecasts. *International Review of Financial Analysis*, 42, 141-152. <https://doi.org/10.1016/j.irfa.2015.01.014>
- Pohl, M., Staegemann, D. G., & Turowski, K. (2022). The Performance Benefit of Data Analytics Applications. *Procedia Computer Science*, 201, 679-683. <https://doi.org/10.1016/j.procs.2022.03.090>
- Racine, J. S. (2019). *Reproducible Econometrics Using R*. Oxford University Press. <https://doi.org/10.1093/oso/9780190900663.001.0001>
- Rahmer, B. de J., Garzón Saénz, H., & Solana Garzón, J. (2022). Aplicación de modelos de difusión y de series temporales para pronóstico de demanda agregada. *Revista de ciencias sociales*, 28(3), 142-159.
- Rojas, F. M., & Gomez, C. (2014). Funcionalidades de la minería de datos. *Ingeniería y Región*, 12, 31-40. <https://doi.org/10.25054/22161325.728>
- Shinde, N., & Gawande, K. (2018). Survey on predicting property price. *2018 International Conference on Automation and Computational Engineering (ICACE)*, 1-7. <https://doi.org/10.1109/ICACE.2018.8687080>

- Sofía, F., & Adrián, J. (s. f.). *La analítica de datos como ventaja competitiva en las organizaciones*.
- Stellwagen, E., & Tashman, L. (2013). ARIMA: The Models of Box and Jenkins. *Foresight: The International Journal of Applied Forecasting*, 30, 28-33.
- Stevenson, S. (2007). A comparison of the forecasting ability of ARIMA models. *Journal of Property Investment & Finance*, 25(3), 223-240. <https://doi.org/10.1108/14635780710746902>
- Verma, P., Reddy, S. V., Ragma, L., & Datta, D. (2021). Comparison of Time-Series Forecasting Models. 2021 *International Conference on Intelligent Technologies (CONIT)*, 1-7. <https://doi.org/10.1109/CONIT51480.2021.9498451>
- Wang, D., & Li, V. J. (2019). Mass Appraisal Models of Real Estate in the 21st Century: A Systematic Literature Review. *Sustainability*, 11(24), Article 24. <https://doi.org/10.3390/su11247006>
- Wang, F., Zou, Y., Zhang, H., & Shi, H. (2019). House Price Prediction Approach based on Deep Learning and ARIMA Model. 2019 *IEEE 7th International Conference on Computer Science and Network Technology (ICCSNT)*, 303-307. <https://doi.org/10.1109/ICCSNT47585.2019.8962443>

## Transparencia

### Conflicto de interés

Los autores declaran que no existen conflictos de interés que influyan en la objetividad de este estudio.

### Fuente de financiamiento

No se recibieron fondos financieros de ninguna organización que pudiera tener interés en los resultados presentados.

### Contribución de autoría

Elba María Bodero Poveda: Conceptualización, metodología, software, validación, análisis formal, investigación, gestión de datos, visualización, redacción - preparación del borrador original, redacción - revisión y edición, financiamiento, administración del proyecto, recursos, supervisión.

Katherine Alexandra Guerrero Morejón: Conceptualización, software, validación, análisis formal, investigación, gestión de datos, visualización, redacción - preparación del borrador original, redacción - revisión y edición, financiamiento, recursos, supervisión.

Los autores desempeñaron un papel equitativo en la ejecución de la investigación, involucrándose de manera equitativa en las etapas clave del proceso. Cada uno de ellos contribuyó activamente en las fases de revisión, análisis, procesamiento y síntesis de datos. La participación equitativa de los autores garantizó una colaboración integral y una distribución justa de responsabilidades a lo largo de todas las actividades relacionadas con la investigación.