

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Property valuation in Buenos Aires, based on online information. A hedonic pricing approach

Emiliano Gutierrez

Instituto de Ciencias e Ingeniería de la Computación (Argentina).
Departamento de Economía, Universidad Nacional del Sur (Argentina)
<https://orcid.org/0000-0002-6424-996X>
emiliano.gutierrez@uns.edu.ar

Laura Lanzarini

Instituto de Investigación en Informática LIDI (Argentina)
Facultad de Informática, Universidad Nacional de La Plata (Argentina)
<https://orcid.org/0000-0001-7027-7564>
laural@lidi.info.unlp.edu.ar

Fernando Delbianco

Instituto de Matemática de Bahía Blanca (Argentina)
Departamento de Economía, Universidad Nacional del Sur (Argentina)
<https://orcid.org/0000-0002-1560-2587>
fernando.delbianco@uns.edu.ar

Rocío Cecchini

Instituto de Ciencias e Ingeniería de la Computación (Argentina).
Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur
(Argentina)
<https://orcid.org/0000-0003-4214-6852>
rlc@cs.uns.edu.ar

RESUMEN

Este estudio realiza una contribución al mercado inmobiliario de Argentina, utilizando información en línea y aplicando modelos estadísticos basados en el enfoque hedónico. A partir de más de 63,000 anuncios de departamentos en venta, se aplican tres métodos estadísticos: Mínimos Cuadrados Ordinarios (MCO), regresión LASSO (Least Absolute Shrinkage and Selection Operator) y árboles de decisión CART (Classification and Regression Trees). Las variables explicativas utilizadas se

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

agrupan en cuatro categorías: características del inmueble, atributos del barrio, proximidad geográfica y movilidad. Los resultados obtenidos revelan que atributos como la superficie, el número de habitaciones y la presencia de amenidades tienen un efecto positivo y significativo en los precios. La ubicación en barrios de alto valor, como Puerto Madero, también es un factor determinante. La cercanía a universidades se asocia con precios más altos, mientras que la proximidad al transporte público tiene un impacto negativo. En términos de movilidad, se observa que una mayor circulación en zonas residenciales está relacionada con precios más elevados, mientras que la movilidad en áreas comerciales y de entretenimiento tiene un efecto contrario. Entre los métodos utilizados, LASSO fue el que mejor performance predictiva presentó para aquellas observaciones fuera de la muestra.

Estos resultados deben considerarse no solo desde lo cuantitativo, sino también desde una perspectiva social, considerando la relevancia de la vivienda como activo de interés, siendo esta información útil para la toma de decisiones de política habitacional.

PALABRAS CLAVE

Precios hedónicos; Vivienda; Argentina; CART; LASSO; MCO.

ABSTRACT

This study makes a contribution to the real estate market in Argentina by utilizing online information and applying statistical models based on the hedonic approach. Using data from over 63,000 apartment listings for sale, three statistical methods are applied: Ordinary Least Squares (OLS), Least Absolute Shrinkage and Selection Operator (LASSO), and Classification and Regression Trees (CART). The explanatory variables used are grouped into four categories: property characteristics, neighborhood attributes, geographic proximity, and mobility.

The results reveal that attributes such as size, number of rooms, and the presence of amenities have a positive and significant effect on prices. Location in high-value neighborhoods, such as Puerto Madero, is also a determining factor. Proximity to green spaces and universities is associated with higher prices, while closeness to public transportation has a negative impact. In terms of mobility, greater circulation in residential areas is linked to higher prices, whereas mobility in commercial and entertainment areas has the opposite effect. Among the methods used, LASSO demonstrated the best predictive performance for out-of-sample observations.

These results should be considered not only from a quantitative perspective but also from a social standpoint, given the relevance of housing as a key asset. This information is useful for decision-making in housing policy.

KEYWORDS

Hedonic prices; Housing; Argentina; CART; LASSO; OLS.

Clasificación JEL: C10, R31.

MSC2010: 91G70, 91B02.

1. INTRODUCCIÓN

Los datos generados diariamente por los usuarios en el entorno digital ofrecen herramientas valiosas para estudiar fenómenos socioeconómicos con mayor precisión. Esta revolución de los datos ha facilitado el seguimiento de una amplia variedad de temas, cuya recopilación tradicional de información hubiera sido considerablemente costosa o incluso inviable. Este cambio es transformador, pues posibilita a los investigadores examinar de forma rigurosa cuestiones relevantes dentro de campos altamente interdisciplinarios, brindando en el caso de la ciencia económica nuevas oportunidades de investigación (Gutiérrez et al., 2018; Mann, 2018; Ng, 2017).

Además, con la expansión del comercio electrónico, surge la oportunidad de investigar mercados específicos a partir de la información en línea (Santos et al., 2023). Un área particularmente interesante es el análisis del mercado inmobiliario, cuya accesibilidad a datos resulta fundamental.

Esta información es de gran valor para los gobiernos nacionales, permitiendo el desarrollo de políticas públicas que mejoren las condiciones habitacionales de los ciudadanos. Esta importancia a nivel global queda reflejada en la Nueva Agenda Urbana 2030, que establece entre sus líneas de acción la necesidad de realizar un “análisis diferenciado de la oferta y la demanda de vivienda, basado en datos desglosados de buena calidad, oportunos y fiables a escala nacional, subnacional y local, considerando las características sociales, económicas, ambientales y culturales específicas” (ONU-HABITAT, 2020). Este enfoque está estrechamente vinculado con el 11° Objetivo de Desarrollo Sostenible de la Agenda 2030, que busca lograr que las ciudades y los asentamientos humanos sean inclusivos, seguros, resilientes y sostenibles (Organización de las Naciones Unidas, 2015).

En el caso de Argentina, el mercado inmobiliario presenta profundas imperfecciones, junto con un significativo grado de desregulación (Baer & Kauw, 2016). Es característico que los precios de la mayoría de las viviendas en venta estén valuados en dólares estadounidenses (Gaggero & Nemina, 2013), con una marcada expansión desde principios del siglo XXI. De esta manera, la evolución del precio de la vivienda en la Ciudad de Buenos Aires es un indicador clave de la riqueza real del sector privado en el país, además de funcionar como cobertura frente a la inflación (Coremberg, 2000).

En este contexto, el estudio del mercado inmobiliario no solo permite una adecuada valoración de los inmuebles, sino que también facilita el conocimiento de la oferta disponible, lo que contribuye a una asignación más eficiente de los recursos. Esto constituye una condición necesaria para alcanzar una asignación óptima del mercado (Akerlof, 1978), favoreciendo la transparencia y la competitividad de las transacciones (Stiglitz, 2000).

Así, el precio final de un inmueble puede ser explicado a partir de la valorización de las distintas características de la vivienda. En este sentido, la utilización de modelos de precios hedónicos, es útil para determinar la valorización de cada uno de los atributos que la componen. Siendo que cada departamento es entendido como un conjunto de características cuyos precios implícitos son obtenidos a partir de las distintas combinaciones de estos determinantes.

De esta forma, el objetivo de este trabajo es generar modelos que permitan determinar el precio de los departamentos para la Ciudad de Buenos Aires (Argentina) y cuantificar el impacto de sus características a partir de un enfoque hedónico.

Concretamente, se busca recuperar los anuncios de venta publicados en línea y combinar esta información proveniente de diversas bases de datos públicas, con el fin de formular modelos que permitan estimar el precio final de un departamento, como también evaluar el desempeño predictivo de estos modelos.

A continuación se describe la forma en que se estructura esta investigación. En un primer apartado describirá los antecedentes relevantes vinculados a aquellos modelos de precios hedónicos focalizándose en el caso argentino. Un segundo apartado describirá tanto los datos como los algoritmos a utilizar. Los resultados derivados de la aplicación de los modelos estimados se-

rán presentados en una tercera sección. Una discusión sobre las mismos se presenta la cuarta sección. Finalmente, en una última sección se presentan las principales conclusiones.

2. ANTECEDENTES

El modelo de precios hedónicos se basa en la idea de que el valor final de un bien puede explicarse por los precios implícitos de sus atributos. Estos precios se revelan a partir de la formulación de una función del precio p , que depende de z atributos de la forma (

$$p(z) = p(z_1, z_2, \dots, z_i).$$

Rosen (1974) es uno de los pioneros en señalar que el precio implícito de un bien puede determinarse a partir de la valoración de sus características, las cuales pueden concebirse como un vector de atributos propios. Estas estimaciones se conocen como modelos hedónicos y suponen una función de precio continua con una forma funcional (típicamente lineal) determinada a priori. Así, el precio de un bien se revela a partir de la combinación de las características que lo componen.

Esta relación entre precio y características es abordada previamente por Lancaster (1966), quien postula que la función entre precio y características es lineal.

En el contexto de la valorización de la vivienda, uno de los primeros aportes sobre el impacto de los precios fue realizado por Ridker y Henning (1967), quienes estimaron el impacto negativo de la contaminación sobre el precio de una vivienda en el área metropolitana de St. Louis (Estados Unidos). Es relevante señalar que los autores no proporcionan una justificación teórica para la elección de este método, ni explicitan la noción de precio implícito en los resultados obtenidos (Chau & Chin, 2003). Este primer estudio dio lugar a una profusa literatura en las décadas siguientes, asociada a la aplicación de modelos de precios hedónicos en el mercado de la vivienda, utilizando técnicas econométricas tradicionales (Malpezzi et al., 2003).

Sin embargo, en el contexto argentino, existen antecedentes de interés que han dado lugar a estimaciones de la valorización hedónica de la vivienda. Uno de los primeros aportes es el trabajo del Laboratorio de Investigaciones del Territorio y el Ambiente (1998). Motivados por la falta de este tipo de análisis en la ciudad de La Plata, los autores plantean un modelo hedónico para dicha localidad, utilizando datos extraídos de planillas de tasación de casas y departamentos de una entidad bancaria, correspondientes al año 1994. Con esta información, desarrollan un modelo de regresión utilizando Mínimos Cuadrados Ordinarios (MCO), donde la variable dependiente es el precio del inmueble.

Se estiman dos modelos: uno que considera únicamente casas, con 48 observaciones, y otro que abarca departamentos, para el cual se recopilaron 25 tasaciones. Las variables explicativas se agrupan en tres categorías: atributos propios de la vivienda, variables ambientales y servicios públicos. Los resultados revelan la existencia de dos submercados claramente diferenciados entre departamentos y casas. En el caso de los departamentos, la variable más relevante es la superficie, mientras que la micro localización, especialmente la proximidad a instituciones educativas y al transporte, también desempeña un papel importante.

Por otro lado, para el modelo de casas, se identifica una mayor variedad de variables influyentes. En este caso, los atributos relacionados con la presencia de espacios de almacenamiento y la cantidad de baños son los que reciben mayor valoración. En cuanto al entorno, tanto la distancia a zonas congestionadas y ruidosas (como la cercanía a escuelas) como la distancia a áreas consideradas inseguras son variables ambientales que afectan significativamente el precio. Además, los espacios verdes públicos, que brindan beneficios recreativos y de calidad ambiental, tienen un impacto notable en la valorización.

Para esta misma ciudad, Rabassa y Zoloa (2016) proponen un modelo espacial de precios hedónicos con el fin de evaluar cuánto estarían dispuestos a pagar los agentes para evitar inun-

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

daciones. Utilizan estimaciones a través de un modelo autorregresivo espacial (SARAR) y MCO, basándose en datos recopilados de inmobiliarias correspondientes a anuncios de venta de viviendas y terrenos. En estas estimaciones, el logaritmo natural del precio del inmueble se considera como variable dependiente. Se desarrollan dos tipos de modelos, el primero de ellos se halla enfocado en determinar el precio de las viviendas y el segundo se focaliza en el valor de los terrenos. En el modelo de vivienda, se consideran variables explicativas como la cantidad de baños, garajes, superficie, habitaciones, antigüedad y distancia al centro de la ciudad. Para el modelo de terrenos, se incluyen factores como el acceso a servicios, el tamaño y la distancia al centro.

Adicionalmente, en ambos tipos de inmuebles, se identifican variables asociadas al riesgo de inundación y a la ubicación en zonas inundables. Todas estas variables mostraron una influencia negativa sobre el precio del inmueble, tanto en el análisis de MCO como en el de SARAR, aportando evidencia sobre la relevancia del riesgo de inundación al momento de adquirir un inmueble.

El aporte de Jansson (2000) examina diversas especificaciones hedónicas para la localidad de Catamarca, utilizando datos del Ministerio de Obras Públicas de la provincia. Este estudio aplica estimaciones de Box-Cox sobre un conjunto de 27 variables explicativas, de las cuales solo 12 logran superar el umbral de significancia.

D'Elía et al. (2020) analizan el impacto del Metrobús en el precio de las viviendas en la Ciudad de Buenos Aires. Con datos recopilados en línea en agosto de 2014, realizan estimaciones hedónicas emparejando propiedades con características similares mediante Propensity Score Matching (PSM). Este análisis abarca tanto barrios cercanos al Metrobús como aquellos alejados. Los resultados sugieren que la proximidad a este servicio no tiene un impacto significativo en el precio de las viviendas. Los autores argumentan que esto se debe a condiciones locales, donde los usuarios del Metrobús suelen residir fuera de la ciudad y utilizan el servicio principalmente por motivos laborales.

Por su parte, Cruces et al. (2008) estudian la calidad de vida en el Área Metropolitana de Buenos Aires, dedicando una sección específica a un modelo hedónico de precios para la ciudad. En este análisis, la variable dependiente es el logaritmo natural del precio del m². Utilizando 5000 observaciones obtenidas de bases de datos de inmobiliarias durante noviembre de 2006, realizan estimaciones mediante Mínimos Cuadrados Ordinarios (MCO) incluyendo tanto departamentos como casas en la misma regresión. Se consideran diversas variables explicativas, que abarcan las características del inmueble (habitaciones, pisos, espacio de garaje, baños, superficie y antigüedad), así como la distancia a avenidas, escuelas, subtes, trenes y espacios verdes. Además, se incorpora una variable binaria para identificar la diferencia de precio entre casas y departamentos.

Los resultados indican que el valor del m² se ve afectado positivamente y con significatividad al 95 % por la superficie total, el espacio del garaje, el número de pisos, la cantidad de baños y la proximidad a avenidas. En contraste, variables como la distancia al subte, los espacios verdes, la superficie cubierta y la antigüedad se asocian negativamente con el valor del inmueble.

Para la ciudad de Resistencia, el trabajo de Arias y Cardozo (2018) analiza el precio del m² en casas, departamentos y terrenos en venta. A partir de un total de 188 datos recolectados de portales web entre 2016 y 2017, plantean un modelo de regresión lineal que considera 14 variables explicativas relacionadas con las características propias del inmueble (como dormitorios, terraza, cocina, baño, lavadero, comedor y living), así como el acceso a diversos servicios (cloacas, pavimento y ripio), la distancia al centro y la distancia a la red de transporte público. En una primera estimación, encuentran que solo cuatro de las variables explicativas superan el umbral de significatividad, por lo que reestiman el modelo utilizando únicamente estas variables (baño, terraza, distancia al centro y distancia a la red de transporte público). Los resultados indican que la existencia de baño, balcón y la cercanía al transporte público tienen un efecto positivo, mientras que la distancia al centro ejerce una influencia negativa sobre el precio del m².

Por otro lado, Meloni y Núñez (2002) evalúan el impacto de los atributos que determinan el precio de los terrenos en la ciudad de San Miguel de Tucumán. Con datos provenientes de la Encuesta de Valores Inmobiliarios realizada en esa localidad, obtienen un total de 607 observaciones. Utilizando 7 variables explicativas –acceso a gas, cloacas, pavimento, ubicación en esquina, localización en zona comercial mixta, presencia en zona comercial y relación frente-fondo–, llevan a cabo regresiones lineales de especificación log-log, donde la variable de interés es el precio por m². Los resultados obtenidos indican que el mayor impacto positivo en el precio se debe a la ubicación y a la disponibilidad de servicios. Asimismo, los autores enfatizan que, en términos generales, los coeficientes asociados a las variables explicativas resultaron estables en todas las regresiones efectuadas.

Fuera del contexto argentino, existen importantes antecedentes respecto de la valoración hedónica aplicada a la valorización de vivienda.

Fuera del contexto argentino, la literatura internacional ofrece múltiples antecedentes que aplican el enfoque hedónico a la valorización inmobiliaria en distintos países, explorando dimensiones como eficiencia energética, factores macroeconómicos, calidad ambiental, seguridad o características urbanas.

En Europa, Cespedes-López et al. (2020) analizan el caso de Alicante (España), evaluando 62.939 viviendas mediante una regresión hedónica donde la variable dependiente es el logaritmo natural del precio de venta y se incluyeron 62 factores explicativos relacionados con características de los inmuebles. Los resultados muestran que las viviendas con certificados energéticos de alta eficiencia no obtienen primas de precio relevantes en este mercado.

Para el caso de Chile, Vergara-Perucich (2023) estudia el mercado de Santiago de Chile con más de 456 mil transacciones, identificando como principales determinantes variables macroeconómicas, la tasa de interés hipotecaria, la antigüedad de la vivienda y el precio internacional del cobre.

Otros trabajos enfatizan el rol de variables ambientales y de seguridad. En este sentido, Marrouch y Sayour 2021 estiman para el mercado inmobiliario libanes el impacto de la contaminación del aire (PM10) sobre los precios de viviendas, mostrando que un incremento marginal en dicha concentración reduce entre 2.88 % y 3.18 % el valor de venta. Por su parte, González-Juárez et al. (2021) analizan seis municipios de Guanajuato (México) y encuentran que tanto la tasa de robos como la de homicidios afectan negativamente el precio de la vivienda.

La evidencia también incorpora determinantes vinculados a la localización y la infraestructura urbana. En China, Luo et al. (2021) muestran que factores como la ubicación central, la calidad de los acabados y la accesibilidad al transporte público (autobuses y metro) incrementan el valor de las viviendas en Dalian. En Turquía, Arslanlı (2020) demuestra que en el área metropolitana de Estambul el emplazamiento en el sector europeo de la ciudad y las amenidades del vecindario generan diferenciales positivos en el precio, mientras que el alquiler y ciertas condiciones de accesibilidad lo reducen.

Por su parte, el trabajo Cohen et al. (2022) evalúa el impacto en el mercado de vivienda inmobiliario de Nueva York de un evento exógena como lo fue la pandemia de COVID-19 utilizando un modelo hedónico lineal para aquellos inmuebles que presentaron ventas repetidas, los autores identifican dos canales principales: un efecto de contagio, relacionado con la exposición percibida de vivir en barrios con altas tasas de infección, y un efecto ingreso, asociado con el aumento del desempleo, señalando entonces que un aumento de casos se traduce en menores precios de venta.

3. METODOLOGÍA Y DATOS

En este trabajo, los datos recolectados fueron extraídos del portal Mercado Libre mediante la implementación de un algoritmo de web scraping. Se consideraron solamente aquellos depar-

tamentos en venta dentro de la ciudad de Buenos Aires. La fecha de extracción de la información corresponde al 11 de agosto de 2023. Tras su preprocesamiento, la cantidad de observaciones válidas fue de 63.017. Los atributos analizados se agruparon en cuatro categorías.

La primera de ellas se refiere a las características del inmueble, incluyendo tanto el precio como los atributos propios de la construcción. En esta categoría se consideran el precio, la superficie, la cantidad de habitaciones y las amenidades disponibles.

Si bien la mayor parte de los avisos estaban valuados en dólares estadounidenses (USD), una proporción inferior al 5 % estaba publicada en pesos argentinos (ARS). Para estas publicaciones, se procedió a la conversión tomando como referencia el valor del dólar informal disponible en el portal de noticias *Ámbito Financiero* (2024) al día de la fecha de la obtención de la información.

Posteriormente, dado que se contaba con la latitud y longitud de cada departamento, fue posible detectar el barrio correspondiente a cada uno de los inmuebles relevados. Esta información fue provista por el gobierno de la Ciudad de Buenos Aires y está disponible en línea Gobierno de la Ciudad de Buenos Aires (2024). Así, este segundo grupo de variables se centra exclusivamente en el barrio donde se localiza cada departamento.

La geolocalización permitió combinar los datos con bases abiertas del Instituto Geográfico Nacional, 2025 en su sitio web. De esta manera, se estimó la distancia más próxima en metros que cada inmueble presenta respecto a universidades, espacios verdes, hospitales, cárceles, estaciones de subte y trenes, y paradas de colectivos. Estas métricas conforman una tercera categoría de atributos considerados. Asimismo, para cada una de estas variables, se consideró un término cuadrático de las mismas con el propósito de evaluar posibles efectos no lineales, para los cuales a partir de determinados umbrales de distancia podría existir un cambio en la influencia sobre el precio de estas predictoras.

El último grupo de variables relevante al momento de valorar un departamento está relacionado con la movilidad en cada una de las 15 comunas de la Ciudad de Buenos Aires, utilizando datos brindados por Google. La medición utilizada corresponde al 15 de octubre de 2022 e incluye seis categorías: áreas residenciales, lugares de trabajo, comercios de bienes de primera necesidad, parques y plazas, comercios minoristas y de esparcimiento, y nodos de transporte. Los indicadores provienen del índice de movilidad de la empresa y fueron estandarizados en el intervalo $[-1,1]$.

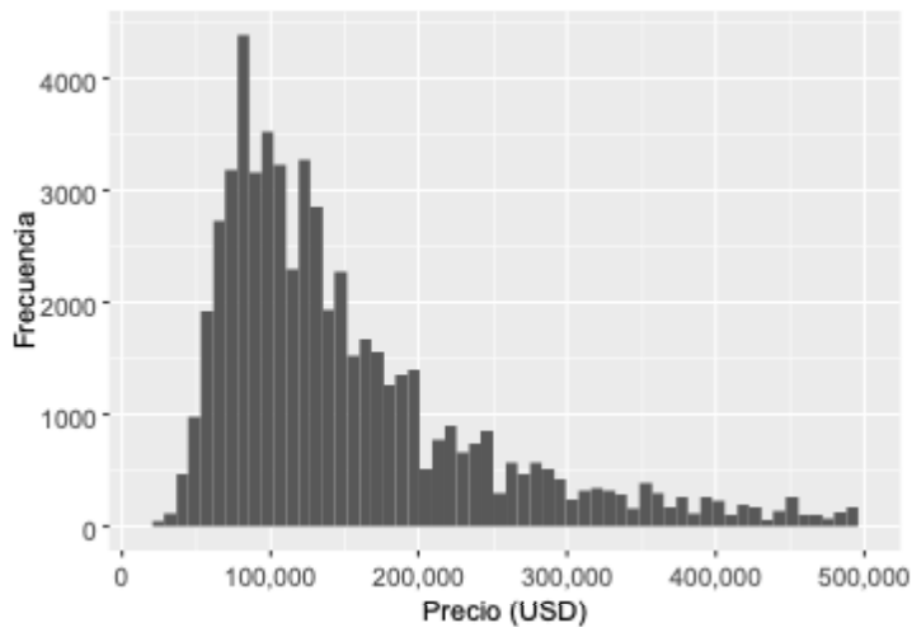
En la Tabla 1 se presenta un resumen de la totalidad de atributos considerados, ascendiendo estos a 107 variables. En el anexo a este artículo, en la tabla 6, se detalla la totalidad de variables utilizadas como también la estadística descriptiva de las mismas se exhibe en la tabla 7. En lo que respecta a la distribución de precios de cada departamento, tal como se presenta en el histograma correspondiente a la figura 1 la mayor frecuencia de valores se encuentra entre los U\$100.000 y U\$200.000, siendo la mediana U\$130.000. No obstante, existe una cantidad significativa de inmuebles con precios superiores, lo que indica una notable dispersión de precios en este mercado.

Tabla 1: Agrupamiento de variables utilizadas.

Categoría	Cantidad de atributos	Descripción	Fuente
Características del inmueble	46	Características del inmueble asociadas a su precio, tamaño, habitaciones y amenities	Avisos de Mercado Libre
Barrio	48	Barrio en donde el departamento se halla localizado	Buenos Aires Data
Proximidad geográfica	14	Distancia en metros a cárceles, hospitales, subte, tren, escuelas y espacios verdes	Instituto Geográfico Nacional
Movilidad	6	Registros de movilidad en sectores de las comunas	Google

Fuente: Elaboración propia.

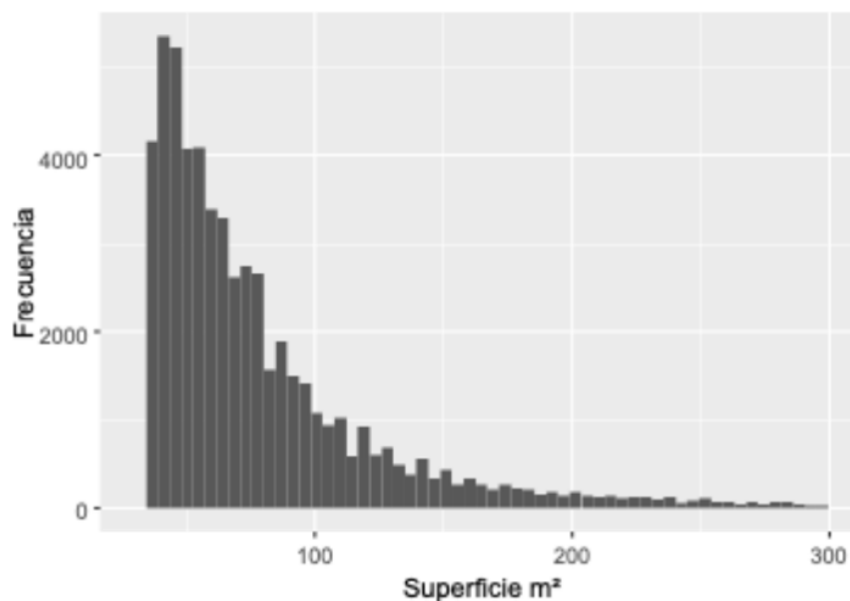
Figura 1: Histograma de precios por departamento



Fuente: Elaboración propia.

El tamaño también presenta una amplia dispersión, tal como se muestra en el histograma correspondiente a la Figura 2. Si bien la mayor parte de la distribución presenta superficies inferiores a 100 m², existe una considerable oferta de inmuebles de mayor tamaño. La mediana se ubicó en los 62 m².

Figura 2: Histograma por superficie



Fuente: Elaboración propia.

Por otra parte, considerando el valor de los departamentos por m², es que se presenta la mediana del mismo por barrio en la Tabla 2, donde se observa que la localización más cara es Puerto Madero, siendo que el precio del m² asciende a U\$5485, seguido de Palermo (U\$2750), Belgrano (U\$2614) y Retiro (U\$2500). En contraste, los barrios con los valores más bajos por m² son Villa Soldati (U\$990), Villa Lugano (U\$1000) y Nueva Pompeya (U\$1352).

Tabla 2: Mediana del precio del m2 por barrio (en dólares estadounidenses).

Barrio	Valor por m2 (USD)	Barrio	Valor por m2 (USD)
PUERTO MADERO	\$ 5.485	PARQUE CHACABUCO	\$ 1.829
PALERMO	\$ 2.750	AGRONOMIA	\$ 1.823
BELGRANO	\$ 2.614	VILLA LURO	\$ 1.818
RETIRO	\$ 2.500	BOEDO	\$ 1.816
COLEGIALES	\$ 2.438	PATERNAL	\$ 1.807
RECOLETA	\$ 2.427	FLORES	\$ 1.786
COGHLAN	\$ 2.391	VILLA SANTA RITA	\$ 1.747
NUÑEZ	\$ 2.350	VILLA REAL	\$ 1.744
VILLA URQUIZA	\$ 2.250	SAN NICOLAS	\$ 1.725
CHACARITA	\$ 2.206	VERSALLES	\$ 1.720
SAAVEDRA	\$ 2.125	MATADEROS	\$ 1.692
VILLA DEVOTO	\$ 2.115	VILLA GRAL. MITRE	\$ 1.669
VILLA CRESPO	\$ 2.053	VELEZ SARSFIELD	\$ 1.638
VILLA ORTUZAR	\$ 2.046	CONSTITUCION	\$ 1.575
CABALLITO	\$ 2.025	SAN CRISTOBAL	\$ 1.563
PARQUE CHAS	\$ 1.952	BALVANERA	\$ 1.557
SAN TELMO	\$ 1.946	FLORESTA	\$ 1.556
VILLA PUEYRREDON	\$ 1.938	VILLA RIACHUELO	\$ 1.512
ALMAGRO	\$ 1.910	PARQUE PATRICIOS	\$ 1.502
BARRACAS	\$ 1.880	PARQUE AVELLANEDA	\$ 1.468
LINIERS	\$ 1.878	BOCA	\$ 1.353
VILLA DEL PARQUE	\$ 1.872	NUEVA POMPEYA	\$ 1.352
MONSERRAT	\$ 1.866	VILLA LUGANO	\$ 1.000
MONTE CASTRO	\$ 1.855	VILLA SOLDATI	\$ 990

Fuente: Elaboración propia.

Considerando entonces que el modelo a desarrollar tiene una función dependiente de cuatro grupos de variables, con la siguiente forma funcional:

$$\log(\text{precio}_i) = f(\text{Características del inmueble, Barrio, Proximidad geográfica, Movilidad}) \quad (1)$$

Siendo expresada en 1 como la variable dependiente el logaritmo natural del precio en dólares del departamento i . Debido a la baja proporción de observaciones en doce barrios, se excluyó la variable binaria asociada a los mismos, siendo descartas las localizaciones de Villa Riachuelo, Santa Rita, Villa del Parque, Villa Devoto, Villa Gral. Mitre, Villa Lugano, Villa Luro, Villa Ortúzar, Villa Pueyrredon, Villa Soldati y Villa Urquiza. Así, se redujo la cantidad de variables explicativas a 101 de las 113 consideradas inicialmente.

Como herramienta de programación, para desarrollar y evaluar el desempeño de los diferentes modelos, se utilizó el lenguaje Python.

Una de las técnicas estadísticas tradicionales refiere al uso de Mínimos Cuadrados Ordinarios (MCO) para modelar mediante relación entre múltiples variables independientes (predictoras) y una variable dependiente continua, con el objetivo de estimar la media poblacional de los valores conocidos de aquel fenómeno que adquiere interés en ser predicho, mediante la forma

$$Y_i = \beta_0 + \beta_1 X_i + u_i \quad (2)$$

Donde para (2), Y corresponde al vector conteniendo el logaritmo natural del precio de cada departamento, X_i la matriz con las variables explicativas correspondiente y u el termino aleatorio de error de cada observación.

Dado que la variable dependiente se encuentra expresada en logaritmos, este modelo se corresponde con una especificación de tipo log-lin, donde los coeficientes β reflejan la semielasticidad de X sobre Y , es decir, β representa el impacto en términos porcentuales de X sobre el valor de un departamento. Para la formulación de la misma se hizo uso del paquete statsmodels (Seabold & Perktold, 2010).

Sin embargo, MCO presenta dos aspectos que pueden suscitar críticas. En primer lugar, su desempeño en predicción, es decir, en datos que no han sido utilizados para armar el modelo, puede mejorarse eliminando algunas variables o reduciendo el valor de sus coeficientes. La segunda razón tiene que ver con la interpretación: cuando el número de variables explicativas es elevado, como es el caso de esta investigación, es conveniente seleccionar solo aquellos predictores relevantes que realmente aporten a la explicación de la variable dependiente. Considerando entonces estos dos aspectos Tibshirani (1996) desarrolla la denominada regresión LASSO (Least Absolute Shrinkage and Selection Operator).

$$\begin{aligned} \min \left\{ \sum_{i=1}^N (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{i,j})^2 \right\} \\ \text{sujeto a } \sum_{j=1}^p |\beta_j| \leq t \end{aligned} \quad (3)$$

Dicho algoritmo parte del supuesto que los predictores se han sido estandarizados previamente, planteándose como función objetivo:

Dada la ecuación (3) parámetro t controla la reducción en el valor de los coeficientes. Si t es tan grande como

$$t_0 = \sum_j |\beta_{j,MCO}|$$

sería entonces una situación similar a la de la regresión lineal mediante MCO. De esta forma valores $t < t_0$ implican que podrían reducir algunos coeficientes a cero.

En el caso de aquellas regresiones LASSO, se pierden los típicos p -valores asociados a la regresión lineal. Esto se debe a que las funciones aplicadas son no lineales y no diferenciables respecto a un valor de t , lo que dificulta la obtención de los errores estándar (Tibshirani, 1996)

Otra forma de plantear (3) es reescribir dicha ecuación como un lagrangiano a resolver, tal como se muestra a continuación:

$$\sum_{i=1}^N (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{i,j})^2 + \lambda \sum_{j=1}^p |\beta_j| \quad (4)$$

Si el valor de λ en 4 es igual a 0 en 4, estaríamos en presencia de una solución similar a MCO. De manera opuesta si $\lambda \rightarrow \infty$, todos los β serían nulos.

La búsqueda del parámetro λ adquiere una profunda relevancia. Para determinar su valor óptimo, se puede llevar a cabo una validación cruzada iterativa sobre el modelo planteado (James et al., 2013; Ranstam & Cook, 2018). La implementación de este modelo, fue efectuada mediante el paquete scikit-learn (Pedregosa et al., 2011).

Un tercer modelo a estimar serán los árboles de decisión CART (Classification and Regression Trees). Desarrollado por Breiman et al. (1984). A diferencia de MCO, este tipo de modelos busca particionar el conjunto de datos de manera recursiva para detectar las relaciones más relevantes con respecto a la variable predictora. El resultado final es un árbol de decisión binario, en el que cada una de las ramas contiene una parte de este conjunto de datos (Varian, 2014).

Siendo el Error Medio Cuadrático (Mean Squared Error, según sus siglas en inglés) entendido como

$$EMC = \frac{\sum_{i=1}^n (y_i - \hat{y})^2}{n} \quad (5)$$

Donde para (5), y_i corresponde al logaritmo natural del precio de cada departamento. El algoritmo comienza dividiendo inicialmente la totalidad de las observaciones en dos regiones ($R1$, $R2$), donde se considera a la media de

$$y \text{ } (\hat{y})$$

para cada una de las regiones. Suponiendo la variable explicativa x_j y el punto de división s , es posible definir un par de planos para los cuales:

$$R1(j, s) = \{X | x_j > s\} \wedge R2(j, s) = \{X | x_j \leq s\} \quad (6)$$

El valor adoptado en s en la ecuación 6 tiene como objetivo minimizar el MSE definido en (5) para cada una de las regiones y puede ser caracterizado como:

$$\min_{j,s} [\min_{c_1} \sum (y_i - c_1)^2 + \min_{c_2} \sum (y_i - c_2)^2] \quad (7)$$

En la ecuación 7 los valores c_1 y c_2 representan las medias de la variable dependiente en cada una de las regiones definidas en (6). Este proceso de división de regiones, minimizando a partir de la selección de una variable explicativa x_j es efectuado iterativamente, dando lugar entonces a un árbol de decisión binario.

Para evaluar que tan bueno es el ajuste de un modelo, es esencial diseñar indicadores específicos que consideren su desempeño predictivo. Estas métricas, pueden utilizarse en los datos de entrenamiento, es decir para aquellos que el modelo haya utilizado para realizar su estimación como también en datos que no hayan sido utilizados en el entrenamiento del modelo. En este sentido, conviene distinguir entonces entre el conjunto de entrenamiento, el cual refiere a las observaciones utilizadas por el algoritmo a estimar, así como también el conjunto de prueba, el cual sirve para poder evaluar el poder predictivo de un modelo con información ajena al mismo. Adicionalmente, como estrategia de remuestreo, se hará uso de la técnica de validación cruzada (cross-validation), la cual consiste en particionar los datos en K grupos, procediendo iterativamente con $K-1$ agrupamientos y utilizando el restante para comprobar su ajuste (Probst, 1998)

Una de las métricas más populares para evaluar el ajuste de un modelo es el Error Medio Cuadrático (MSE), que se define como el promedio de los errores al cuadrado en relación con los valores reales y predichos, tal como se presentó previamente en la ecuación 5. Otra métrica utilizada es la Raíz del Error Cuadrático Medio (RMSE), que se puede entender como la raíz cuadrada del MSE:

$$RMSE = \sqrt{MSE} \quad (8)$$

Una ventaja importante del RMSE señalado 8 en es su interpretación, ya que se puede entender como una media cuadrática de los errores, conservando la escala de los datos que se pretenden predecir.

Otro indicador que se utiliza para observar el ajuste de la predicción es el coeficiente de determinación (R^2), que se calcula como

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (9)$$

En (9), el numerador representa la varianza de la predicción, mientras que el denominador corresponde a la varianza total del conjunto de datos (Hawinkel et al., 2024), siendo

$$\hat{y}_i$$

la media del conjunto de datos, mientras que

$$\bar{y}$$

es el valor precedido para y_i . Esta métrica puede interpretarse como la proporción de la varianza total que es explicada por el modelo. Aunque es común utilizar R^2 para evaluar la bondad de ajuste de la regresión lineal, puede ser utilizado en cualquier otra estimación. Para estos últimos, el rango de valores que puede adoptar es $[-\infty, 1]$, donde un valor próximo a 1 indica un mejor

ajuste. En el caso específico de la regresión lineal mediante MCO y con datos provenientes del conjunto de entrenamiento, es importante señalada que el rango de R^2 se restringe a $[0,1]$.

4. RESULTADOS

En la Tabla 3 se presentan los resultados asociados a MCO conservándose solo los valores de aquellos atributos que superan el umbral de significatividad del 95 %, ascendiendo la cantidad de los mismos a 82. El resultado de esta estimación con la totalidad de las variables explicativas puede verse en la Tabla 8.

Las variables que presentan el mayor impacto positivo están asociadas con la localización, donde el barrio de Puerto Madero (BARRIO_PUERTO.MADERO) implicaría un mayor precio de venta del inmueble. Este barrio es el que presenta el valor más alto del m², tal como se presentó en la Tabla 2. Del mismo modo, barrios como Caballito (BARRIO_CABALLITO), Coghlan (BARRIO_COGLAN) y Recoleta (BARRIO RECOLETA) son otras localizaciones que sobresalen dado su impacto positivo sobre el valor de un departamento.

De manera opuesta, los barrios de San Nicolas (BARRIO_SAN.NICOLAS) y Parque Avellaneda (BARRIO PARQUE.AVELLANEDA) son los que se encuentran asociados con un menor precio final del departamento.

Tabla 3: MCO. Variables significativas al 95 %

Variable	Coeficiente	Error estándar	Valor-t	P-valor
Constante	10,45478	0,014	768,296	0,000
BARRIO_PUERTO.MADERO	0,81574	0,014	59,569	0,000
BARRIO_CABALLITO	0,65007	0,006	103,643	0,000
BARRIO_RECOLETA	0,64000	0,011	60,725	0,000
BARRIO_COGLAN	0,60280	0,012	51,016	0,000
BARRIO_BARRACAS	0,52835	0,012	43,553	0,000
BARRIO_CHACARITA	0,49042	0,013	37,895	0,000
BARRIO_VERSALLES	0,47691	0,028	17,035	0,000
BARRIO_SAAVEDRA	0,47292	0,010	46,460	0,000
BARRIO_PARQUE.CHACABUCO	0,47181	0,010	48,038	0,000
BARRIO_FLORES	0,45658	0,009	53,630	0,000
BARRIO_MONTE.CASTRO	0,42883	0,014	29,631	0,000
BARRIO_BALVANERA	0,40353	0,007	55,912	0,000

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Variable	Coefficiente	Error estándar	Valor-t	P-valor
BARRIO VILLA. CRESPO	0,39190	0,010	40,048	0,000
BARRIO_RETIRO	0,36873	0,010	36,862	0,000
BARRIO_VELEZ. SARSFIELD	0,29676	0,016	18,226	0,000
BARRIO_FLORESTA	0,28838	0,017	17,145	0,000
BARRIO_PARQUE. CHAS	0,28740	0,017	16,723	0,000
BARRIO_SAN. CRISTOBAL	0,28357	0,008	35,931	0,000
BARRIO_ AGRONOMIA	0,26690	0,027	9,943	0,000
BARRIO_LINIERS	0,25141	0,020	12,454	0,000
BARRIO_PARQUE. PATRICIOS	0,23878	0,017	13,844	0,000
BARRIO_PALERMO	0,23467	0,008	28,311	0,000
BARRIO_BOCA	0,22679	0,017	13,697	0,000
BARRIO_PATERNAL	0,19666	0,020	9,744	0,000
BARRIO_BELGRANO	0,18330	0,006	31,346	0,000
BARRIO_ MATADEROS	0,18109	0,018	10,305	0,000
BARRIO_ COLEGIALES	0,14985	0,007	20,829	0,000
HAS_GYM	0,13635	0,005	27,397	0,000
BARRIO_NUNEZ	0,13205	0,008	17,459	0,000
HAS_HALF_BATH	0,12830	0,003	36,773	0,000
BARRIO SAN.TELMO	0,10888	0,013	8,426	0,000
FULL_BATHROOMS	0,10556	0,002	44,826	0,000
HAS_SWIMMING_ POOL	0,10402	0,005	22,922	0,000
BARRIO_NUEVA. POMPEYA	0,10136	0,025	4,033	0,000
HAS_BALCONY	0,09609	0,003	27,851	0,000
BARRIO ALMAGRO	0,07884	0,007	11,516	0,000

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Variable	Coefficiente	Error estándar	Valor-t	P-valor
HAS_DRESSING_ROOM	0,07242	0,005	13,942	0,000
ROOMS	0,06804	0,001	51,167	0,000
HAS_TENNIS_COURT	0,06602	0,045	1,459	0,145
HAS_PLAYROOM	0,06343	0,008	7,518	0,000
BARRIO_BOEDO	0,04662	0,011	4,383	0,000
Disposicion_frente	0,04329	0,003	14,584	0,000
HAS_JACUZZI	0,03806	0,008	4,595	0,000
HAS_GUEST_PARKING	0,03575	0,028	1,278	0,201
HAS_CINEMA_HALL	0,03492	0,043	0,812	0,417
HAS_MULTIPURPOSE_ROOM	0,02861	0,005	6,308	0,000
HAS_SECURITY	0,02798	0,005	5,785	0,000
PARKING_LOTS	0,02446	0,001	21,871	0,000
HAS TELEPHONE LINE	0,02016	0,005	4,217	0,000
HAS HEATING	0,00767	0,004	2,035	0,042
TOTAL_AREA	0,00541	0,000	144,688	0,000
colectivos	0,00051	0,00000	9,382	0,000
trenes	0,00012	0,00000	11,029	0,000
subte	0,00006	0,00000	9,094	0,000
salud_2	0,00002	0,00000	10,523	0,000
salud	0,00001	0,00000	10,523	0,000
espacios_verdes2	0,000001	0,00000	14,398	0,000
subte_2	-0,000001	0,00000	-17,397	0,000
trenes_2	-0,000001	0,00000	-7,384	0,000
colectivo_2	-0,00000	0,00000	-6,963	0,000
universidades	-0,00001	0,00000	-6,143	0,000
universidades_2	-0,00001	0,00000	-6,143	0,000

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Variable	Coeficiente	Error estándar	Valor-t	P-valor
espacios verdes	-0,00014	0,00000	-17,125	0,000
HAS_DINNING_ROOM	-0,01017	0,004	-2,280	0,023
HAS_NATURAL_GAS	-0,01219	0,004	-3,305	0,001
HAS_PLAYGROUND	-0,01224	0,026	-0,468	0,640
HAS_LIFT	-0,01879	0,005	-3,930	0,000
HAS_KITCHEN	-0,02781	0,004	-7,518	0,000
HAS_CLOSET	-0,03359	0,008	-4,325	0,000
transit_stations_percent_change_from_baseline	-0,03833	0,004	-9,850	0,000
HAS_TAP_WATER	-0,04547	0,004	-11,524	0,000
HAS_PATIO	-0,04666	0,006	-8,319	0,000
BARRIO_SAN. NICOLAS	-0,05280	0,010	-5,451	0,000
BARRIO_PARQUE. AVELLANEDA	-0,05525	0,019	-2,933	0,003
HAS_STUDY	-0,06772	0,009	-7,379	0,000
grocery_and_pharmacy_percent_change_from_baseline	-0,07495	0,002	-30,579	0,000
parks_percent_change_from_baseline	-0,07510	0,003	-26,873	0,000
residential_percent_change_from_baseline	-0,11354	0,004	-26,495	0,000
BARRIO_CONSTITUCION	-0,19250	0,011	-17,005	0,000
workplaces_percent_change_from_baseline	-0,20027	0,007	-28,935	0,000
HAS_INDOOR FIREPLACE	-0,23789	0,059	-4,033	0,000
retail_and_recreation_percent_change_from_baseline	-0,40544	0,006	-71,888	0,000

Variable	Coeficiente	Error estándar	Valor-t	P-valor
R ² : 0,817 R ² ajustado: 0,817 AIC: 221,537 BIC: 229,674 F-estadístico: 2513 Prob (F-statistic): 0,00				

Fuente: Elaboración propia.

En lo que respecta a las características del departamento, el acceso a un gimnasio (HAS_GYM) dentro del edificio es el atributo que incrementa en mayor medida el precio final, con un aumento del 13,64 %. Este es seguido por la presencia de medio baño (HAS_HALF_BATH) y la cantidad total de baños (FULL_BATHROOMS). Por el contrario, contar con chimenea (HAS_INDOOR_FIREPLACE) se asocia con un impacto negativo significativo en el valor final del inmueble, disminuyendo su valor en un 23,79 %.

En relación con las variables de proximidad geográfica, la distancia a paradas de colectivo (colectivos, colectivos_2) presenta coeficientes positivos y negativos respectivamente, lo que refleja que el impacto sobre el precio no es constante y varía con la distancia, destacándose que una cercanía en un principio se asocia negativamente con el precio, siendo que, a partir de una lejanía, esta relación se vuelve positiva. Situación similar ocurre en el caso de las distancias a estaciones de tren (tren, tren_2) y de subte (subte, subte_2), como también para las instituciones de salud (salud, salud_2) y espacios verdes (espacios verdes, espacios verdes_2). Por otra parte, la distancia a universidades (universidades, universidades_2) muestra ambos coeficientes con el mismo signo, lo que indica que un mayor alejamiento se asocia con menores precios del inmueble.

En cuanto a la movilidad, se observa que, para todas las variables de este grupo, que una mayor movilidad se asocia negativamente con el precio de un departamento. En particular, la movilidad relacionada con lugares de entretenimiento y esparcimiento (retail_and_recreation_percent_change_from_baseline) presenta el mayor impacto negativo.

Previamente a determinar el valor de los coeficientes, fue necesario un ajuste del modelo en el conjunto de entrenamiento mediante validación cruzada con 5 pliegues, con el objetivo de optimizar el hiperparámetro λ el cual corresponde al definido en la ecuación 4. Se evaluaron 100 valores de λ en una escala logarítmica, que abarco desde 10^{-5} hasta 105.

En la Tabla 4, se presentan los coeficientes que resultaron diferentes de cero, lo que indica las variables que contribuyen de manera efectiva a la predicción del precio del departamento.

Tabla 4: LASSO, Variables con coeficientes no nulos.

Variable	Coeficiente	Variable	Coeficiente
Constante	10,812	BARRIO_PUERTO. MADERO	0,813
BARRIO_RETIRO	0,365	BARRIO_VERSALLES	0,165
BARRIO_CHACARITA	0,161	BARRIO_RECOLETA	0,146
HAS_GYM	0,137	BARRIO_COGHLAN	0,129
HAS_HALF_BATH	0,128	BARRIO_MONTE.CASTRO	0,127
BARRIO_BARRACAS	0,124	FULL_BATHROOMS	0,106
BARRIO_SAN.TELMO	0,105	HAS_SWIMMING_POOL	0,104
HAS_BALCONY	0,096	residential_percent_ change_from_baseline	0,074
HAS_DRESSING_ROOM	0,072	ROOMS	0,068
HAS_PLAYROOM	0,063	BARRIO_VILLA.CRESPO	0,062
HAS_TENNIS_COURT	0,048	disposicion_frente	0,043
HAS_JACUZZI	0,038	BARRIO_LINIERS	0,035
BARRIO_BELGRANO	0,033	HAS_GUEST_PARKING	0,030
HAS_MULTIPURPOSE_ ROOM	0,029	BARRIO_PALERMO	0,028
HAS_SECURITY	0,028	PARKING_LOTS	0,024
HAS_CINEMA_HALL	0,024	HAS_PARTY_ROOM	0,021
HAS_TELEPHONE_LINE	0,020	HAS_BREAKFAST_BAR	0,019
HAS_BEDROOM_SUITE	0,017	BARRIO_PARQUE. CHACABUCO	0,012
HAS_BUSINESS_CENTER	0,012	HAS_MAID_ROOM	0,012
HAS_HEATING	0,009	TOTAL_AREA	0,006
HAS_GRILL	0,005	HAS_LIVING_ROOM	0,004
HAS_AIR_ CONDITIONING	0,004	HAS_BOILER	0,001
colectivos	0,0005	salud	0,00003
trenes	0,00003	subte	0,00003
universidades_2	0,00002	carcel_2	0,000003
subte_2	0,00000001	espacios_verdes_2	0,0000001

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Variable	Coeficiente	Variable	Coeficiente
subte_dist_2	-0,0000001	trenes_2	-0,0000001
colectivo_2	-0,000001	carcel	-0,000001
espacios_verdes	-0,000002	universidades	-0,000004
HAS_LAUNDRY	-0,0007	HAS_INTERNET_ACCESS	-0,002
HAS_GARDEN	-0,004	grocery_and_pharmacy_percent_change_from_baseline	-0,005
HAS_TERRACE	-0,008	HAS_DINNING_ROOM	-0,011
BARRIO_SAN.NICOLAS	-0,013	HAS_NATURAL_GAS	-0,013
HAS_CISTERN	-0,013	HAS_COMMON_LAUNDRY	-0,014
parks_percent_change_from_baseline	-0,019	HAS_LIFT	-0,019
BARRIO_SAAVEDRA	-0,021	BARRIO_AGRONOMIA	-0,021
workplaces_percent_change_from_baseline	-0,025	HAS_KITCHEN	-0,028
HAS_CLOSET	-0,038	HAS_TAP_WATER	-0,047
HAS_PATIO	-0,049	BARRIO_ALMAGRO	-0,067
HAS_STUDY	-0,068	BARRIO_LINIERS	-0,069
BARRIO_SA.,CRISTOBAL	-0,107	BARRIO_BOEDO	-0,121
BARRIO_PATERNAL	-0,125	BARRIO_MATADEROS	-0,143
BARRIO_CONSTITUCION	-0,162	retail_and_recreation_percent_change_from_baseline	-0,422
BARRIO_PARQUE,PATRICIOS	-0,448	BARRIO_BOCA	-0,484
BARRIO_NUEVA.POMPEYA	-0,509	BARRIO_PARQUE.AVELLANEDA	-0,542
HAS_INDOOR_FIREPLACE	-0,616		

Fuente: Elaboración propia.

Una observación importante sobre los resultados obtenidos es que el modelo LASSO excluyó solo 4 variables explicativas. El coeficiente positivo más alto corresponde a los departamentos ubicados en Puerto Madero (BARRIO_PUERTO.MADERO), lo que refuerza los hallazgos previos en MCO. En contraste, el barrio de Parque Avellaneda (BARRIO_PARQUE.AVELLANEDA) presenta el coeficiente más bajo en el análisis de localización. Cabe destacar que este barrio también se caracteriza por tener uno de los precios más bajos por metro cuadrado, como se detalla en la Tabla 2.

En lo que respecta a las características del inmueble, los resultados son similares a los obtenidos mediante MCO. Las variables que tienen un impacto positivo significativo en el precio incluyen el acceso a un gimnasio (HAS_GYM), la cantidad de baños (FULL_BATHROOMS) y la disponibilidad de un medio baño (HAS_HALF_BATH).

En relación con las variables de proximidad, el modelo LASSO identifica la distancia a una cárcel (carcel, carcel_2) como un atributo relevante, donde a medida que aumenta la cercanía a un centro penitenciario, el precio final del inmueble también tiende a incrementarse. En el caso de universidades (universidades, universidades_2) se presenta también una relación distinta observada a la de MCO, donde el término cuadrático invierte su signo. Además, el alejamiento de una estación de subte (subte, subte_2) se asocia con un mayor valor del inmueble. Las demás variables explicativas mantienen un comportamiento similar a MCO.

En el grupo de variables relacionadas con la movilidad, se destaca una asociación positiva entre la movilidad en zonas residenciales (residential_percent_change_from_baseline) y el valor de venta de los departamentos. Este hallazgo contrasta con los resultados obtenidos por MCO, donde la movilidad en áreas residenciales tenía una relación negativa. Las otras variables de este grupo continúan mostrando una asociación negativa con el precio, manteniendo el signo observado en el MCO.

La tercera estimación realizada corresponde a CART. Al igual que en LASSO, el proceso se llevó a cabo utilizando el paquete scikitlearn (Pedregosa et al., 2011), aplicando validación cruzada con 5 pliegues. Para este modelo, se calibraron varios hiper parámetros: la profundidad del árbol, evaluando valores de 1, 3 y 5; la cantidad mínima de observaciones por nodo, con valores de 1000, 2000 y 3000; y la cantidad mínima de observaciones por hoja, considerando valores de 5, 10 y 15. El objetivo de esta calibración fue minimizar el error cuadrático medio (MSE).

El mejor modelo resultante presentó una profundidad de 5. En la Figura 3, se muestra una versión simplificada del mismo y exhibiéndose la totalidad del árbol en el anexo correspondiente a la Figura 4.

La variable que se encuentra en la parte superior del árbol de decisión es la superficie del inmueble (TOTAL_AREA), que clasifica los departamentos según si su tamaño es inferior o superior a 88 m². Esta variable continúa siendo relevante en las ramificaciones subsiguientes.

La primera partición indica, como era de esperar, que los inmuebles de menor tamaño tienden a tener un precio promedio más bajo. Sin embargo, en la segunda ramificación, el tamaño del departamento vuelve a jugar un papel crucial.

Otra variable importante relacionada con las características del inmueble es la cantidad de estacionamientos disponibles. La presencia de al menos un espacio de estacionamiento (PARKING_LOTS) se asocia con un incremento en el precio de venta. Asimismo, la disponibilidad de un gimnasio, que mostro uno de los mayores coeficientes en MCO y LASSO, también tiene un impacto positivo en la valorización final del inmueble.

El modelo también resalta, al igual que en el caso de LASSO, la relación positiva entre la movilidad en zonas residenciales y el precio de venta. Sin embargo, en contraste con otras variables de circulación, como la cercanía a parques (parks_percent_change_from_baseline) y sitios de entretenimiento (retail_and_recreation_percent_change_from_baseline), se observa que un aumento en la circulación en estas áreas se traduce en un precio medio del inmueble más bajo.

Tal como se evidencio en los análisis previos, la localización en el barrio de Puerto Madero (BARRIO_PUERTO_MADERO) continúa destacando en este modelo; los departamentos situados en esta zona experimentan un incremento significativo en su precio de venta.

Un hallazgo relevante de la aplicación de CART es que no incluye en el árbol resultante, las variables relacionadas con la cercanía a cárceles (carcel, carcel_2), instituciones de salud (hospital, hospital_2), universidades (universidad, universidad_2), y estaciones de subte (subte, subte_2), trenes (tren, tren_2) o colectivos (colectivo, colectivo_2).

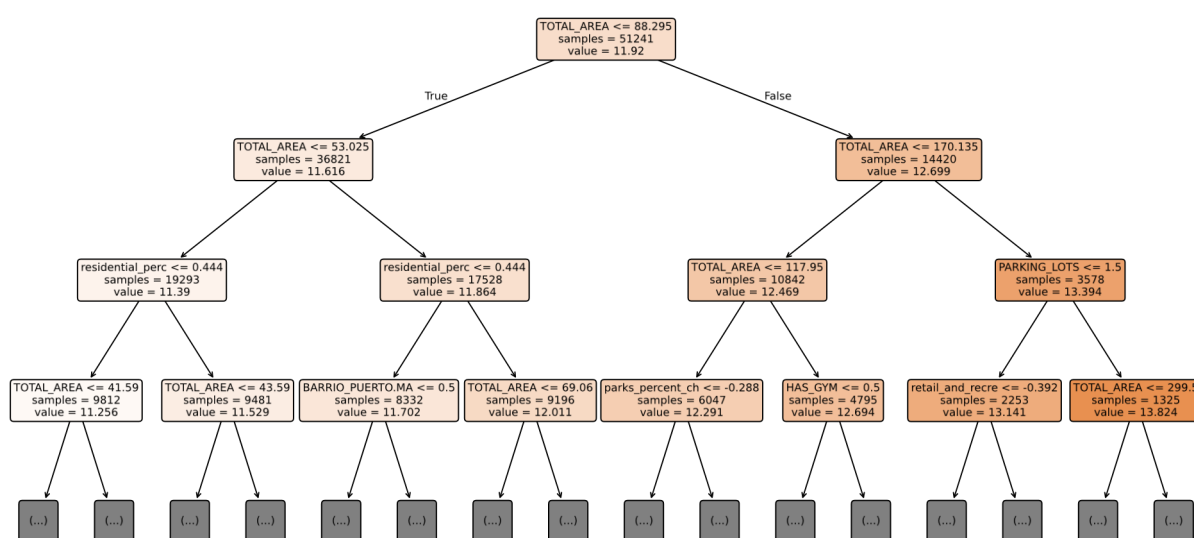
De esta manera, la capacidad predictiva de los modelos desarrollados es un aspecto a evaluar, a fin de observar su ajuste con información que el mismo no ha utilizado en su proceso de entrenamiento.

Además del conjunto de entrenamiento compuesto por el 80 % de la totalidad de los datos, se evalúa entonces el desempeño de los modelos entrenados en el 20 % de datos restantes, siendo detallados estos resultados en la Tabla 5.

Considerando solo los datos utilizados por el modelo en su entrenamiento, CART presento el peor desempeño en cualquier de las métricas consideradas.

Sin embargo, al momento de evaluar con observaciones fuera de la muestra, es decir, el conjunto de prueba, MCO tiene un evidente pobre desempeño, siendo el que mayores errores presenta en sus predicciones. No obstante, sobre el conjunto de entrenamiento los resultados de MCO son prácticamente idénticos a LASSO, donde este último mostró el mejor desempeño en lo relativo a los datos del conjunto de prueba.

Figura 3: CART



Fuente: Elaboración propia.

Tabla 5: Desempeño de los modelos entrenados. Conjunto de entrenamiento y testeo.

	Conjunto de entrenamiento			Conjunto de testeo		
	MSE	RMSE	R2	MSE	RMSE	R2
MCO	0,090	0,300	0,817	0,431	0,657	-0,470
CART	0,101	0,319	0,794	0,115	0,339	0,608
LASSO	0,090	0,301	0,818	0,102	0,319	0,653

Fuente: Elaboración propia

5. DISCUSIÓN

El mercado inmobiliario es uno de los sectores que requiere especial atención en los mecanismos que determinan sus precios, dada la importancia de este bien como uno de los activos de mayor valor que pueden poseer los individuos.

El uso de datos en línea en este contexto permite analizar los determinantes del precio con una mayor cantidad de observaciones en comparación con los métodos tradicionales de relevamiento. Sin embargo, como sucede con cualquier investigación que recurra a información extraída a través de internet, existen problemas derivados de la captura de datos, ya que estos no siempre pueden ser representativos de la totalidad de las propiedades en venta. Además, la publicación de un aviso no garantiza que la operación de venta se lleve a cabo efectivamente.

Se destaca la ausencia de una cantidad significativa de investigaciones previas que aborden específicamente los determinantes del precio de los departamentos en la ciudad de Buenos Aires, y que utilicen información obtenida en línea. Esto dificulta la posibilidad de realizar comparaciones directas con otros estudios. Asimismo, merece resaltarse la naturaleza local de los datos recopilados, basados solamente en la Ciudad de Buenos Aires siendo este un importante limitante al momento de analizar los resultados obtenidos.

Debido a la cantidad de variables explicativas utilizadas, estas fueron agrupadas en cuatro categorías: características propias de la vivienda, características del barrio, proximidad geográfica y movilidad de los individuos.

En cuanto a las variables relacionadas con las características de los inmuebles, diversos estudios previos que emplean modelos hedónicos en Argentina presentan resultados similares, especialmente en lo que respecta a la influencia positiva y significativa de la superficie y la cantidad de habitaciones sobre el precio final. Los resultados obtenidos en los trabajos de D'Elia et al. (2020), Jansson (2000) y Rabassa y Zoloa (2016) confirman la significatividad de algunos de estos atributos en la determinación del precio del inmueble.

Este hallazgo puede explicarse por la idea de que los departamentos, al ser inmuebles de tamaño limitado, verían un impacto marginal considerable en su valor final si cuentan con mayor superficie y/o un mayor número de habitaciones o baños (Laboratorio de Investigaciones del Territorio y el Ambiente, 1998).

En cuanto a los resultados relacionados con el área total, baños y habitaciones, tanto MCO como LASSO revelaron una influencia positiva sobre el precio final de un departamento.

Por otro lado, al analizar el árbol resultante de la aplicación de CART, la superficie se presenta como la variable de mayor jerarquía y está presente en varias ramas. La relevancia de la superficie como uno de los principales predictores es esperable, ya que el tamaño de un inmueble

puede influir en otros atributos y es la primera característica evaluada al determinar el valor de venta de una propiedad.

Entre las amenidades abordadas por la literatura en el contexto argentino, D'Elia et al. (2020) detectan que la disposición del frente, contar con estacionamiento y tener piscina influyen positivamente y superan el umbral de significatividad del 95 % en el precio final del inmueble. La relación positiva entre la cantidad de estacionamientos disponibles y el precio también es señalada en Cruces et al. (2008) y Rabassa y Zoloa (2016), mientras que disponer de piscina como atributo que incrementa el precio es igualmente encontrado en Jansson (2000).

En línea con la evidencia internacional, también se detectan impactos positivos de otras amenidades, como la presencia de seguridad, estacionamientos y balcón, en consonancia con lo reportado por Arslanli (2020) para Estambul, y con los resultados de Marrouch y Sayour (2021), quienes destacan un efecto favorable de disponer de piscina y cantidad de estacionamientos en el caso del Líbano.

Estos resultados son consistentes con los obtenidos en la estimación de la MCO y LASSO, mientras que en el caso de CART, tanto el estacionamiento como la disposición de piscina son seleccionados por el árbol generado.

El segundo grupo de variables explicativas hace referencia al barrio de la ciudad de Buenos Aires donde se localiza el inmueble, siendo las mismas de carácter binario. Sin embargo, el uso de variables binarias, permitiendo controlar por heterogeneidad espacial no observada, tiene como limitación la ausencia de especificación de atributos que explican ese diferencial. La ausencia de datos específicos tales como renta o composición poblacional, permitiría desagregar estos efectos.

Los resultados obtenidos para este grupo de variables mediante LASSO y MCO sugieren que la localización en muchos de los barrios tiene una influencia significativa en el valor final del inmueble. No obstante, en estos dos métodos, al igual que en el caso de CART, el barrio de Puerto Madero es señalado como el más relevante dentro de este grupo de variables explicativas, asociado positivamente con el precio del inmueble. Considerando que dicho barrio posee el mayor valor por metro cuadrado respecto al resto de la ciudad, este resultado es, al menos, esperado.

La proximidad a diferentes servicios, lugares e instituciones es otro grupo de variables abordado en esta investigación. La preferencia por los espacios verdes puede interpretarse como una inclinación hacia áreas abiertas, aunque esto a partir de este impacto se invertiría dado que implicaría un alejamiento de las zonas céntricas.

La relación entre espacios verdes y un mayor precio del inmueble ha sido observada en diversos estudios previos dentro del mercado inmobiliario nacional, como los de Cruces et al. (2008), D'Elia (2020), Jansson (2000), y el Laboratorio de Investigación (1998).

De manera similar a los espacios verdes, la proximidad a universidades también se asocia con una influencia positiva sobre el precio de un departamento, aunque observando el termino cuadrático, podría haber evidencia que este efecto no es totalmente lineal. El impacto positivo de esta institución coincide con los resultados obtenidos en varias investigaciones para el caso de Argentina, donde las instituciones educativas impactan positivamente en el precio del inmueble (Cruces et al., 2008; D'Elia et al., 2020; Laboratorio de Investigaciones del Territorio y el Ambiente, 1998). Fuera del ámbito nacional, el aporte de Luo y Yao (2021), encuentra para China que la proximidad de universidades incrementa el precio de la vivienda.

Respecto a la influencia del transporte público, en este trabajo se analizan tanto la distancia a estaciones de subte, tren y colectivos. Tanto para LASSO como en los resultados obtenidos mediante MCO, la cercanía a cualquiera de estos medios de transporte se traduce en un menor precio de venta, aunque esta situación ocurre manera no lineal, donde la inclusión del término cuadrático revertiría esta influencia, reflejando que una excesiva lejanía tendría como contrapartida un impacto negativo sobre el precio de la vivienda.

En este sentido, uno de los aportes relevantes es el de D'Elia et al. (2020), quienes evalúan específicamente el caso del Metrobús, encontrando no significancia respecto sobre el precio de la vivienda. Sin embargo, hallan evidencia de un impacto positivo en el precio por parte de las estaciones de subte. No obstante, este resultado utiliza una muestra de inmuebles cercanos a la avenida 9 de julio, lo cual representa una diferencia con los datos utilizados en esta investigación. De manera similar, Cruces et al. (2008) encuentran que la proximidad al subte y colectivos presenta una relación positiva con el precio del inmueble. Este hallazgo también es corroborado por Arias y Cardozo (2018), quienes encuentran que la proximidad a las paradas de colectivo aumenta el precio. Del mismo modo, Goytia (2016) observa que la proximidad al subte y tren incrementa el valor de la tierra.

Fuera del plano nacional Arslanian (2020), encuentra que mientras que una mayor distancia al subte se asocia negativamente con el precio, pero el acceso a transportes públicos presentaría un impacto positivo sobre el valor del inmueble.

Si bien estos resultados son diferentes de los obtenidos en esta investigación, se pueden señalar tres aspectos que justificarían esta divergencia. En primer lugar, la mayoría de los antecedentes no abordan específicamente la ciudad de Buenos Aires, o lo hacen de manera parcial. Un segundo aspecto relevante es que en estos estudios no se considera la movilidad dentro de las zonas como una variable explicativa, lo cual podría justificar la influencia positiva observada en el precio. Como tercera cuestión relevante es que ninguno de estos estudios incorpora un término cuadrático para evaluar la existencia de una relación no lineal entre distancia y precio del inmueble.

Por otro lado, en lo que respecta a la proximidad a instituciones penitenciarias, solo LASSO muestra evidencia de que esta variable explicativa tenga algún impacto sobre el precio. Concretamente mayor distancia a este tipo de institución implicaría un menor precio de venta, aunque este impacto es cercano a cero, lo cual también daría cuenta del prácticamente insignificante impacto sobre el precio.

Cabe destacar la ausencia de antecedentes sobre este aspecto en el contexto argentino. Sin embargo, estos resultados, que a priori podrían considerarse inesperados, son encontrados por Stanley (1979) en el caso de viviendas en la prisión de Waupun, ubicados en el estado de Wisconsin, Estados Unidos. El autor encuentra un débil, pero significativo impacto positivo (al 95%) en el precio de las propiedades cercanas a la prisión. Sin embargo, tanto en este último estudio referido como en los resultados aquí obtenidos, no darían cuenta de una relación concluyente en lo que refiere a proximidad y un mayor precio de venta.

Otro tipo de institución que fue evaluada la relación entre cercanía a y precio fue el caso de los hospitales. Tanto LASSO como MCO dieron cuenta de que una menor distancia se asocia con un precio más bajo. Si bien no se registran antecedentes para el caso argentino, el aporte de Luo y Yao (2021), señalan específicamente un resultado opuesto, pero asociado a que los hospitales se localizan en zonas céntricas y comerciales.

La última categoría de variables explicativas refiere a la movilidad de los individuos dependiendo del tipo de zonas, un aspecto sobre el cual no existen antecedentes en el contexto argentino.

La circulación asociada a las estaciones de tren y colectivos presenta un comportamiento interesante. Si bien en la estimación por MCO se encuentra una relación negativa con el precio del inmueble, al observar el árbol generado mediante CART, se detecta que dicha relación no es estrictamente lineal.

El impacto del ruido como factor que tiende a reducir el precio de los inmuebles también podría explicar la relación observada entre la movilidad hacia espacios de trabajo, negocios y farmacias, donde los resultados muestran una relación negativa respecto al precio. Tanto en los métodos LASSO como en MCO, se asigna un signo negativo.

En cuanto a la movilidad asociada a parques, también se observa una relación negativa entre la movilidad y los precios. Este hallazgo es consistente en los métodos LASSO, MCO y CART.

Es importante resaltar que es de esperarse que las zonas con mayor presencia de arbolado y vegetación (Erlwein & Pauleit, 2021) sean aquellas con menor movilidad (Razykova et al., 2014), dado que se encuentran en localizaciones menos concurridas.

Por otro lado, las áreas verdes con mayor circulación se asocian con precios más bajos. En este caso, se podría estar considerando aquellos espacios verdes que son ampliamente concurridos y de fácil acceso. Esta variable podría, por tanto, estar reflejando la influencia de distintas tipologías de espacios verdes. La alta concurrencia en estos espacios puede generar externalidades negativas, como ruido, contaminación visual o una sensación de inseguridad (Chisari et al., 2019), lo que afectaría la valoración de dichas áreas.

La superioridad en el desempeño de LASSO con respecto al conjunto de datos de prueba resalta las implicaciones de este método al trabajar con información real, que el modelo no ha tenido en cuenta al hacer sus estimaciones. Dicha situación es enfatizada por Tibshirani (1996), autor del algoritmo, quien demuestra mediante datos simulados que LASSO ofrece un mejor rendimiento en los datos de prueba justificado porque, aunque el método de MCO presenta un bajo sesgo, su principal desventaja es una alta varianza.

6. CONCLUSIÓN

La disponibilidad de información en línea a partir de la masificación en el uso de internet ha permitido la generación de importantes fuentes de información impensadas hace algunos años. La recuperación de estos datos permite la aparición de nuevas líneas de investigación donde la ciencia económica tiene un vasto camino por desarrollar.

En el caso de este trabajo, se pretendió realizar una contribución al mercado inmobiliario de Argentina, utilizando información en línea y aplicando modelos estadísticos basados en el enfoque hedónico.

Así, se relevaron más de 60000 avisos de venta correspondientes a departamentos dentro de la Ciudad de Buenos Aires, disponibles en un portal de comercio electrónico. Adicionalmente, la utilización de bases complementarias permitió la inclusión de varios atributos relevantes para el estudio de este mercado, donde la variable de interés fue el precio de venta de un departamento.

La elevada cantidad de observaciones disponibles para un estudio de esta índole de más de 100 variables, presenta escasos antecedentes dentro del caso argentino. El agrupamiento de estas variables explicativas en cuatro categorías tuvo como eje central observar como los distintos aspectos de la vivienda podían influenciar el valorización final.

La estimación mediante los tres modelos utilizados (MCO, LASSO y CART) debe entenderse como la aplicación de técnicas complementarias que permiten una mejor interpretación de las relaciones entre las variables explicativas y el valor que presenta un departamento.

Mientras que el impacto de algunas variables vinculadas a las características propias del inmueble como lo son la cantidad de habitaciones, superficie y baños presentaron un impacto positivo asociado al precio en prácticamente todas las estimaciones, otros atributos darían cuenta de un nulo o escaso impacto sobre el valor de venta.

Del mismo modo, para el caso de la localización del inmueble, el barrio de Puerto Madero fue el que mayor relevancia se encontró en gran parte de las estimaciones y asociado con un mayor precio de venta. Para las categorías vinculadas a la movilidad y la proximidad geográfica si bien las mismas presentaron impactos tanto positivos como negativos sobre el precio, en la mayoría de modelos entrenados su influencia fue menor y en algunos casos dependiendo del modelo el impacto podría ser positivo o negativo.

De esta forma, los resultados obtenidos confirman la relevancia de las características estructurales de los inmuebles, pero también muestran que factores urbanos y de movilidad inciden

en la valorización, incluso con efectos no lineales. En este sentido, se aporta evidencia sobre la utilidad de metodologías como LASSO para mejorar la capacidad predictiva frente a los enfoques tradicionales. Asimismo, estos hallazgos, permiten ampliar la discusión teórica al mostrar que además de las características estructurales de los inmuebles, los factores de localización, movilidad y distancia influyen de manera significativa, incluso con efectos no lineales.

Estos resultados deben considerarse no solo desde lo cuantitativo, sino también desde una perspectiva social, considerando la relevancia de la vivienda como activo de interés, siendo esta información relevante para la toma de decisiones de política a fin de potenciar un desarrollo urbano sostenible. En este sentido, las intervenciones urbanas pueden alterar significativamente la dinámica de precios, por lo que deberían considerarse en el momento de diseñar políticas urbanas. Concretamente, proximidad inmediata a espacios verdes y universidades se vincula con precios más altos, mientras que la proximidad a nodos de transporte tiende a reducirlos, reflejando la existencia de externalidades negativas de que debería ser tenidas en cuenta por parte de los tomadores de política urbana.

Adicionalmente, estos resultados pueden ser aprovechados por profesionales del sector inmobiliario y entidades financieras en la estimación de precios, así como en el diseño de herramientas digitales que reduzcan las asimetrías de información en el mercado.

Como limitaciones a este estudio debe señalarse que los datos en línea pueden no ser representativos de todas las transacciones inmobiliarias, como también al ser de corte transversal, es decir, obtenidos en un momento determinado del tiempo, no es posible evaluar las variaciones temporales de los parámetros de cada modelo a lo largo del tiempo. Además, debido a que el análisis se circunscribe a la Ciudad de Buenos Aires, los resultados deben interpretarse como de alcance local.

Asimismo, como líneas de investigación futuras que merecen ser profundizadas a partir de los datos utilizados, sería de interés el estudio de la interacción de las variables explicativas a fin de observar como la combinación de las mismas puede resultar relevante para el desempeño predictivo de los modelos.

Finalmente, se subraya la importancia estimar modelos de valorización de inmuebles enfocados en distintas regiones del país, con el objetivo de determinar si las dinámicas observadas responden a particularidades del mercado inmobiliario propio de la Ciudad de Buenos Aires o si, por el contrario, son tendencias que pueden generalizarse a nivel nacional.

REFERENCIAS

- Akerlof, G. A. (1978). The market for “lemons”: Quality uncertainty and the market mechanism. En *Uncertainty in economics* (pp. 235-251). Elsevier. <https://doi.org/10.2307/1879431>
- Ambito Financiero. (2024). *Dolar informal* [Consultado el 23 de agosto de 2025].
- Arias, F. C., & Cardozo, O. D. (2018). Análisis del valor de la tierra urbana y la vivienda: Estudio de caso en la ciudad de Resistencia, provincia del Chaco, Argentina 2016-2017. *Estudios Socioterritoriales*, 24, 1-17. https://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1853-43922018000200009
- Arslanli, K. Y. (2020). Analysis of house prices: a hedonic model proposal for Istanbul metropolitan area. *Journal of Design for Resilience in Architecture and Planning*, 1(1), 57-68. <https://doi.org/10.47818/DRArch.2020.v1i1004>
- Baer, L., & Kauw, M. (2016). Mercado inmobiliario y acceso a la vivienda formal en la Ciudad de Buenos Aires, y su contexto metropolitano, entre 2003 y 2013. *Eure (Santiago)*, 42(126), 5-25. <https://doi.org/10.4067/S0250-71612016000200001>
- Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Classification and regression trees*. CRC. <https://doi.org/10.1201/9781315139470>
- Céspedes-López, M.-F., Mora-García, R.-T., Pérez-Sánchez, V. R., & Martí-Ciriquian, P. (2020). The influence

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

- of energy certification on housing sales prices in the province of Alicante (Spain). *Applied Sciences*, 10(20). <https://doi.org/10.3390/app10207129>
- Chau, K. W., & Chin, T. (2003). A critical review of literature on the hedonic price model. *International Journal for Housing Science and its applications*, 27(2), 145-165.
- Chisari, O. O., Ramos, M. P., & Leon, S. (2019). Crimen y configuración urbana: una evaluación de precios de propiedades y bienestar en equilibrio general para la ciudad de Buenos Aires. *El Trimestre Económico*, 86(342), 437-466. <https://doi.org/10.20430/ete.v86i342.774>
- Cohen, J. P., Friedt, F. L., & Lautier, J. P. (2022). The impact of the Coronavirus pandemic on New York City real estate: First evidence. *Journal of Regional Science*, 62(3), 858-888. <https://doi.org/10.1111/jors.12591>
- Coremberg, A. A. (2000). El precio de la vivienda en Argentina: un análisis econométrico de sus determinantes fundamentales. *Papeles de población*, 6(23), 93-125.
- Cruces, G., Ham, A., & Tetaz, M. (2008). *Quality of life in Buenos Aires neighborhoods: Hedonic price regressions and the life satisfaction approach*. <https://doi.org/10.18235/0011267>
- D'Elia, V. V., Grand, M. C., & Leon, S. (2020). Bus rapid transit and property values in Buenos Aires: Combined spatial hedonic pricing and propensity score techniques. *Research in Transportation Economics*, 80(1). 1-13. <https://doi.org/10.1016/j.retrec.2020.100814>
- Erlwein, S., & Pauleit, S. (2021). Trade-offs between urban green space and densification: Balancing outdoor thermal comfort, mobility, and housing demand. *Urban Planning*, 6(1), 5-19. <https://doi.org/10.17645/up.v6i1.3481>
- Gaggero, A., & Nemina, P. (2013). El origen de la dolarización inmobiliaria en la Argentina. *Sociales en debate*, 32(60), 136-159. <https://doi.org/10.15446/ede.v32n60.92443>
- Gobierno de la Ciudad de Buenos Aires. (2024). *Barrios de la Ciudad Autónoma de Buenos Aires* [Consultado el 23 de agosto de 2025]. <https://data.buenosaires.gob.ar/dataset/barrios>
- González-Juárez, A., Melo-Guerrero, E., Hernández-Ortiz, J., Martínez-Damian, M.A., Cervantes Luna, J. O., & Valdivia-Alcala, R. (2021). Impacto de la inseguridad en el precio de las viviendas en el estado de Guanajuato: un enfoque de precios hedónicos. *Investigación y Ciencia de la Universidad Autónoma de Aguascalientes*, (83). <https://doi.org/10.33064/iycuaa2021832879>
- Goytia, C., & Dorna, G. (2016). *Big Data and a Spatial Hedonic Approach: Addressing the land market information gap and estimating land prices determinants in metropolitan regions from developing countries*. (inf. tec.). Working Paper, Lincoln Institute of Land Policy.
- Gutierrez, E. M., Larrosa, J. M. C., Delbianco, F., Uriarte, J. I., & Ramírez Muñoz De Toro, G. R. (2018). Aprendizaje automático y métodos hedónicos en el mercado de autos usados en línea de Argentina. LIII Reunión Anual de la Asociación Argentina de Economía Política (La Plata, 14 al 16 de noviembre de 2018).
- Hawinkel, S., Waegeman, W., & Maere, S. (2024). Out-of-sample R²: Estimation and inference. *The American Statistician*, 78(1), 15-25. <https://doi.org/10.1080/00031305.2023.2216252>
- Instituto Geográfico Nacional. (2025). *Capas SIG – Información Geoespacial*. Consultado el 23 de agosto de 2025, desde <https://www.ign.gob.ar/NuestrasActividades/InformacionGeoespacial/CapasSIG>
- James, G., Witten, D., Hastie, T., Tibshirani, R., et al. (2013). *An introduction to statistical learning* (Vol. 112). Springer. <https://doi.org/10.1007/978-1-0716-1418-1>
- Jansson, A.A. (2000). Función de precios hedónicos de viviendas y adaptación del test RESET en modelos no lineales. Aplicación del modelo BOX & COX a los precios de las viviendas de la ciudad de Catamarca, Argentina. *Pharos*, 7(2). 43-59. <https://www.redalyc.org/pdf/208/20807205.pdf>
- Laboratorio de Investigaciones del Territorio y el Ambiente. (1998). *La calidad ambiental como factor explicativo de los precios en el mercado habitacional urbano*.
- Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, 74(2), 132-157. <https://doi.org/10.1086/259131>
- Luo, H., Zhao, S., & Yao, R. (2021). Determinants of housing prices in Dalian city, China: Empirical study based on hedonic price model. *Journal of Urban Planning and Development*, 147(2). [https://doi.org/10.1061/\(ASCE\)UP.1943-5444.0000698](https://doi.org/10.1061/(ASCE)UP.1943-5444.0000698)
- Malpezzi, S., et al. (2003). Hedonic pricing models: a selective and applied review. *Housing Economics and*

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Public Policy, 1 (1), 67–89. <https://doi.org/10.1002/9780470690680.ch5>

Mann, L. (2018). Left to other peoples' devices? A political economy perspective on the big data revolution in development. *Development and Change*, 49(1), 3–36. <https://doi.org/10.1111/dech.12347>

Marrouch, W., & Sayour, N. (2021). Hedonic housing prices and environmental quality in Lebanon. *International Journal of Housing Markets and Analysis*, 14(5), 953–968. <https://doi.org/10.1108/IJHMA-08-2020-0100>

Meloni, O., & Nunez, F. R. (2002). El Precio de los Terrenos y el Valor de sus Atributos. *Económica*, 48, 69–88.

Ng, S. (2017). *Opportunities and challenges: Lessons from analyzing terabytes of scanner data*. Working Paper 23673. National Bureau of Economic Research, <https://doi.org/10.3386/w23673>

ONU-HABITAT. (2020). *La Nueva Agenda Urbana*. <https://unhabitat.org/sites/default/files/2021/10/nueva-agenda-urbana-ilustrada.pdf>

Organización de las Naciones Unidas. (2015). *Transforming our world: the 2030 Agenda for Sustainable Development*. <https://sdgs.un.org/2030agenda>

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of machine learning research*, 12(Oct), 2825–2830. <https://doi.org/10.5555/1953048.2078195>

Provost, F. (1998). Glossary of terms special issue on applications of machine learning and the knowledge discovery process. *Machine Learning*, 30, 271–274.

Rabassa, M. J., & Zoloa, J. I. (2016). Flooding risks and housing markets: a spatial hedonic analysis for La Plata City. *Environment and Development Economics*, 21(4), 464–489. <https://doi.org/10.1017/S1355770X15000376>

Ranstam, J., & Cook, J. A. (2018). LASSO regression. *Journal of British Surgery*, 105(10), 1348–1348. <https://doi.org/10.1002/bjs.10895>

Razykova, V., Bartke, S., & Schwarz, N. (2014). A dynamic model of urban green spaces, residential mobility and real estate market. Social Simulation Conference. https://www.researchgate.net/publication/271585587_Poster_A_dynamic_model_of_urban_green_spaces_residential_mobility_and_real_estate_market

Ridker, R. G., & Henning, J. A. (1967). The determinants of residential property values with special reference to air pollution. *The Review of Economics and Statistics*, 49(2), 246–257. <https://doi.org/10.2307/1928231>

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*, 82(1), 34–55. <https://doi.org/10.1086/260169>

Santos, V., Augusto, T., Vieira, J., Bacalhau, L., Sousa, B. M., & Pontes, D. (2023). E-commerce: issues, opportunities, challenges, and trends. *Promoting organizational performance through 5G and agile marketing*, 224–244. <https://doi.org/10.4018/978-1-6684-5523-4.ch012>

Seabold, S., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. 9th Python in Science Conference. <https://doi.org/10.25080/Majora-92bf1922-011>

Stanley, C. E. (1979). *The impact of prison proximity on property values in Green Bay and Waupun, Wisconsin*. Division of Corrections; Bureau of Facilities Management.

Stiglitz, J. E. (2000). The contributions of the economics of information to twentieth century economics. *The Quarterly Journal of Economics*, 115(4), 1441–1478. <https://doi.org/10.1162/003355300555015>

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1), 267–288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>

Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, 28(2), 3–28. <https://doi.org/10.1257/jep.28.2.3>

Vergara-Perucich, J. F. (2023). Testing housing price drivers in Santiago de Chile: A hedonic price approach. *Critical Housing Analysis*, 10(2), 44–57. <https://doi.org/10.13060/23362839.2023.10.2.558>

ANEXOS

Tabla 6: Variables utilizadas

Variable	Tipo	Descripción
Precio	Numérica	Precio de venta publicado en dólares.
TOTAL_AREA	Numérica	Área total de la vivienda.
ROOMS	Numérica	Cantidad de habitaciones.
FULL_BATHROOMS	Numérica	Cantidad de baños
PARKING_LOTS	Numérica	Cantidad de estacionamientos.
HAS_INTERNET_ACCESS	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_TAP_WATER	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_AIR_CONDITIONING	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_BALCONY	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_BOILER	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_HEATING	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_INDOOR_FIREPLACE	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_CISTERN	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_KITCHEN	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_DINNING_ROOM	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_MAID_ROOM	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_BREAKFAST_BAR	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_BEDROOM_SUITE	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_STUDY	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Variable	Tipo	Descripción
HAS_NATURAL_GAS	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_JACUZZI	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_GARDEN	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_LAUNDRY	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_TELEPHONE_LINE	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_LIVING_ROOM	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_PATIO	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_CLOSET	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_TERRACE	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_HALF_BATH	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_DRESSING_ROOM	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_CINEMA_HALL	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_PLAYGROUND	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_LIFT	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_BUSINESS_CENTER	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_TENNIS_COURT	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_GUEST_PARKING	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_GYM	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_COMMON_LAUNDRY	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_GRILL	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Variable	Tipo	Descripción
HAS_SWIMMING_POOL	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_PLAYROOM	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_ROOF_GARDEN	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_PARTY_ROOM	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_MULTIPURPOSE_ROOM	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
HAS_SECURITY	Binaria	Presencia de la amenidad (1) o ausencia de la misma(0)
DISPOSICION_FRENTE	Binaria	Disposición del inmueble al frente (1) o en otra posición (0)
retail_and_recreation_percent_change_from_baseline	Numérica	Movilidad estandarizada zonas de comercio minoristas y/o esparcimiento
grocery_and_pharmacy_percent_change_from_baseline	Numérica	Movilidad estandarizada comercios de bienes de primera necesidad.
parks_percent_change_from_baseline	Numérica	Movilidad estandarizada en parques y plazas.
transit_stations_percent_change_from_baseline	Numérica	Movilidad estandarizada en nodos de transporte.
workplaces_percent_change_from_baseline	Numérica	Movilidad estandarizada en lugares de trabajo.
residential_percent_change_from_baseline	Numérica	Movilidad estandarizada en zonas residenciales.
carcel	Numérica	Distancia en metros a la cárcel más cercana.
carcel_2	Numérica	Cuadrado de la distancia en metros a la cárcel más cercana.
subte	Numérica	Distancia en metros a la estación de subte más cercana.
subte_2	Numérica	Cuadrado de la distancia en metros a la estación de subte más cercana.
trenes	Numérica	Distancia en metros a la estación de tren más cercana.

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Variable	Tipo	Descripción
trenes_2	Numérica	Cuadrado de la distancia en metros a la estación de tren más cercana.
espacios_verdes	Numérica	Distancia en metros a la espacio verde más próximo
espacios_verdes_2	Numérica	Cuadrado de la distancia en metros a la espacio verde más próximo
universidades	Numérica	Distancia en metros a la universidad más cercana.
universidades_2	Numérica	Cuadrado de la distancia en metros a la universidad más cercana.
salud	Numérica	Distancia en metros al hospital más cercano.
salud_2	Numérica	Distancia en metros al hospital más cercano.
colectivos	Numérica	Distancia en metros a la parada de colectivo más cercana.
colectivos_2	Numérica	Distancia en metros a la parada de colectivo más cercana.
BARRIO_AGRONOMIA	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_ALMAGRO	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_BALVANERA	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_BARRACAS	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_BELGRANO	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_BOCA	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_BOEDO	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_CABALLITO	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_CHACARITA	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_COGLAN	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Variable	Tipo	Descripción
BARRIO_COLEGIALES	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_CONSTITUCION	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_FLORES	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_FLORESTA	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_LINIERS	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_MATADEROS	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_MONSERRAT	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_MONTE.CASTRO	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_NUEVA.POMPEYA	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_NUÑEZ	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_PALERMO	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_PARQUE.AVELLANEDA	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_PARQUE.CHACABUCO	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_PARQUE.CHAS	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_PARQUE.PATRICIOS	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_PATERNAL	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_PUERTO.MADERO	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_RECOLETA	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_RETIRO	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_SAAVEDRA	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Variable	Tipo	Descripción
BARRIO_SAN.CRISTOBAL	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_SAN.NICOLAS	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_SAN.TELMO	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_VELEZ.SARFIELD	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_VERSALLES	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_VILLA.CRESPO	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_VILLA.DEL.PARQUE	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_VILLA.DEVOTO	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_VILLA.GRAL..MITRE	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_VILLA.LUGANO	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_VILLA.LURO	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_VILLA.ORTUZAR	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_VILLA.PUEYREDON	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_VILLA.REAL	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_VILLA.RIACHUELO	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_VILLA.SANTA.RITA	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_VILLA.SOLDATI	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)
BARRIO_VILLA.URQUIZA	Binaria	Ubicación dentro del barrio (1) o fuera de sus límites (0)

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Variable	Media	Mediana	Desvío estandar	Max	Mín	Curtosis
Precio	211650,02	130000,00	3673102,85	9000000000,00	16000,00	59732,74
TOTAL_AREA	79,72	62,00	57,58	1000,00	30,00	20,23
ROOMS	2,61	3,00	1,38	85,00	-1,00	394,57
FULL_BATHROOMS	1,41	1,00	0,79	40,00	0,00	114,31
PARKING_LOTS	0,29	0,00	1,24	100,00	0,00	4177,32
HAS_INTERNET_ACCESS	0,21	0,00	0,41	1,00	0,00	0,02
HAS_TAP_WATER	0,54	1,00	0,50	1,00	0,00	-1,98
HAS_AIR_CONDITIONING	0,32	0,00	0,47	1,00	0,00	-1,43
HAS_BALCONY	0,53	1,00	0,50	1,00	0,00	-1,99
HAS_BOILER	0,04	0,00	0,20	1,00	0,00	18,56
HAS_HEATING	0,19	0,00	0,39	1,00	0,00	0,49
HAS_INDOOR_FIREPLACE	0,00	0,00	0,02	1,00	0,00	2147,89
HAS_CISTERN	0,00	0,00	0,07	1,00	0,00	220,78
HAS_KITCHEN	0,52	1,00	0,50	1,00	0,00	-1,99
HAS_DINNING_ROOM	0,20	0,00	0,40	1,00	0,00	0,29
HAS_MAID_ROOM	0,08	0,00	0,28	1,00	0,00	6,87
HAS_BREAKFAST_BAR	0,01	0,00	0,12	1,00	0,00	65,03
HAS_BEDROOM_SUITE	0,03	0,00	0,16	1,00	0,00	33,32
HAS_STUDY	0,02	0,00	0,15	1,00	0,00	39,28
HAS_NATURAL_GAS	0,39	0,00	0,49	1,00	0,00	-1,79
HAS_JACUZZI	0,03	0,00	0,17	1,00	0,00	27,63
HAS_GARDEN	0,03	0,00	0,18	1,00	0,00	23,81
HAS_LAUNDRY	0,32	0,00	0,47	1,00	0,00	-1,42

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Variable	Media	Mediana	Desvío estandar	Max	Mín	Curtosis
HAS_TELEPHONE_LINE	0,16	0,00	0,36	1,00	0,00	1,58
HAS_LIVING_ROOM	0,21	0,00	0,41	1,00	0,00	-0,02
HAS_PATIO	0,07	0,00	0,25	1,00	0,00	9,96
HAS_CLOSET	0,05	0,00	0,23	1,00	0,00	13,42
HAS_TERRACE	0,12	0,00	0,32	1,00	0,00	3,79
HAS_HALF_BATH	0,26	0,00	0,44	1,00	0,00	-0,75
HAS_DRESSING_ROOM	0,10	0,00	0,30	1,00	0,00	5,38
HAS_CINEMA_HALL	0,00	0,00	0,04	1,00	0,00	739,21
HAS_PLAYGROUND	0,00	0,00	0,07	1,00	0,00	211,84
HAS_LIFT	0,14	0,00	0,34	1,00	0,00	2,53
HAS_BUSINESS_CENTER	0,00	0,00	0,05	1,00	0,00	451,68
HAS_TENNIS_COURT	0,00	0,00	0,03	1,00	0,00	1091,02
HAS_GUEST_PARKING	0,00	0,00	0,06	1,00	0,00	274,08
HAS_GYM	0,14	0,00	0,35	1,00	0,00	2,32
HAS_COMMON_LAUNDRY	0,01	0,00	0,12	1,00	0,00	68,26
HAS_GRILL	0,20	0,00	0,40	1,00	0,00	0,19
HAS_SWIMMING_POOL	0,22	0,00	0,42	1,00	0,00	-0,22
HAS_PLAYROOM	0,03	0,00	0,17	1,00	0,00	27,39
HAS_ROOF_GARDEN	0,00	0,00	0,04	1,00	0,00	543,01
HAS_PARTY_ROOM	0,02	0,00	0,12	1,00	0,00	58,47

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Variable	Media	Mediana	Desvío estandar	Max	Mín	Curtosis
HAS_MULTIPURPOSE_ROOM	0,21	0,00	0,41	1,00	0,00	-0,02
HAS_SECURITY	0,11	0,00	0,31	1,00	0,00	4,55
disposicion_frente	0,38	0,00	0,49	1,00	0,00	-1,76
retail_and_recreation_percent_change_from_baseline	-0,32	-0,50	0,59	3,14	-0,95	12,47
grocery_and_pharmacy_percent_change_from_baseline	-0,36	-0,29	0,90	1,77	-1,88	-0,27
parks_percent_change_from_baseline	-0,23	-0,40	1,05	2,19	-1,73	0,35
transit_stations_percent_change_from_baseline	0,20	-0,12	0,91	1,50	-1,55	-1,17
workplaces_percent_change_from_baseline	-0,17	-0,09	0,78	2,26	-1,98	0,19
residential_percent_change_from_baseline	0,33	1,18	0,98	1,18	-1,77	-0,57
BARRIO_AGRONOMIA	0,00	0,00	0,04	1,00	0,00	505,86
BARRIO_ALMAGRO	0,05	0,00	0,21	1,00	0,00	16,64
BARRIO_BALVANERA	0,04	0,00	0,20	1,00	0,00	19,49
BARRIO_BARRACAS	0,01	0,00	0,11	1,00	0,00	82,38
BARRIO_BELGRANO	0,08	0,00	0,27	1,00	0,00	7,74

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Variable	Media	Mediana	Desvío estandar	Max	Mín	Curtosis
BARRIO_BOCA	0,00	0,00	0,07	1,00	0,00	199,35
BARRIO_BOEDO	0,01	0,00	0,11	1,00	0,00	72,59
BARRIO_CABALLITO	0,07	0,00	0,26	1,00	0,00	8,66
BARRIO_CHACARITA	0,01	0,00	0,11	1,00	0,00	76,36
BARRIO_COGLAN	0,01	0,00	0,10	1,00	0,00	86,21
BARRIO_COLEGIALES	0,03	0,00	0,17	1,00	0,00	27,63
BARRIO_CONSTITUCION	0,01	0,00	0,11	1,00	0,00	82,63
BARRIO_FLORES	0,03	0,00	0,18	1,00	0,00	24,56
BARRIO_FLORESTA	0,01	0,00	0,09	1,00	0,00	126,05
BARRIO_LINIERS	0,02	0,00	0,14	1,00	0,00	43,28
BARRIO_MATADEROS	0,01	0,00	0,11	1,00	0,00	80,28
BARRIO_MONSERRAT	0,02	0,00	0,13	1,00	0,00	48,99
BARRIO_MONTE. CASTRO	0,01	0,00	0,09	1,00	0,00	125,49
BARRIO_NUEVA. POMPEYA	0,00	0,00	0,04	1,00	0,00	501,57
BARRIO_NUÑEZ	0,04	0,00	0,21	1,00	0,00	17,58
BARRIO_PALERMO	0,14	0,00	0,35	1,00	0,00	2,35
BARRIO_PARQUE. AVELLANEDA	0,00	0,00	0,06	1,00	0,00	243,07
BARRIO_PARQUE. CHACABUCO	0,01	0,00	0,11	1,00	0,00	80,03
BARRIO_PARQUE.CHAS	0,01	0,00	0,07	1,00	0,00	186,98
BARRIO_PARQUE. PATRICIOS	0,00	0,00	0,07	1,00	0,00	200,04

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Variable	Media	Mediana	Desvío estandar	Max	Mín	Curtosis
BARRIO_PATERNAL	0,00	0,00	0,06	1,00	0,00	242,06
BARRIO_PUERTO.MADERO	0,01	0,00	0,12	1,00	0,00	68,71
BARRIO_RECOLETA	0,07	0,00	0,25	1,00	0,00	10,30
BARRIO_RETIRO	0,02	0,00	0,13	1,00	0,00	53,32
BARRIO_SAAVEDRA	0,02	0,00	0,15	1,00	0,00	37,12
BARRIO_SAN.CRISTOBAL	0,01	0,00	0,12	1,00	0,00	65,68
BARRIO_SAN.NICOLAS	0,02	0,00	0,13	1,00	0,00	49,92
BARRIO_SAN.TELMO	0,01	0,00	0,09	1,00	0,00	120,07
BARRIO_VELEZ.SARFIELD	0,01	0,00	0,08	1,00	0,00	159,71
BARRIO_VERSALLES	0,00	0,00	0,04	1,00	0,00	622,93
BARRIO_VILLA.CRESPO	0,04	0,00	0,20	1,00	0,00	20,16
BARRIO_VILLA.DEL.PARQUE	0,01	0,00	0,12	1,00	0,00	68,98
BARRIO_VILLA.DEVOTO	0,02	0,00	0,15	1,00	0,00	40,72
BARRIO_VILLA.GRAL..MITRE	0,01	0,00	0,08	1,00	0,00	145,71
BARRIO_VILLA.LUGANO	0,01	0,00	0,10	1,00	0,00	94,48
BARRIO_VILLA.LURO	0,01	0,00	0,11	1,00	0,00	72,69
BARRIO_VILLA.ORTUZAR	0,01	0,00	0,11	1,00	0,00	77,14
BARRIO_VILLA.PUEYRREDON	0,01	0,00	0,08	1,00	0,00	157,05
BARRIO_VILLA.REAL	0,00	0,00	0,03	1,00	0,00	868,64

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Variable	Media	Mediana	Desvío estandar	Max	Mín	Curtosis
BARRIO_VILLA. RIACHUELO	0,00	0,00	0,03	1,00	0,00	881,49
BARRIO_VILLA. SANTA.RITA	0,01	0,00	0,07	1,00	0,00	185,17
BARRIO_VILLA. SOLDATI	0,00	0,00	0,03	1,00	0,00	856,16
BARRIO_VILLA. URQUIZA	0,05	0,00	0,21	1,00	0,00	15,80
carcel	4209,11	4106,81	2211,99	9804,36	17,10	-0,77
subte	1000,62	596,05	1116,98	5891,32	0,92	4,07
trenes	1105,15	1026,38	602,25	3055,40	0,23	-0,73
espacios_ verdes	1580,30	1570,73	842,59	4662,80	0,00	-0,50
universidades	709,69	575,01	508,63	3547,81	0,04	0,97
salud	687,34	648,61	354,80	2254,70	1,98	0,03
colectivos	93,72	81,72	65,63	730,03	0,30	3,19

Fuente: Elaboración propia

Tabla 8: MCO

Variable	Coefficiente	Error estandar	Valor-t	P-valor
Constante	104,710	0,0128	8,188,218	0,000
BARRIO_PUERTO. MADERO	0,8237	0,0121	679,805	0,000
BARRIO_ CABALLITO	0,6567	0,0062	1,058,097	0,000
BARRIO_COGHLAN	0,6083	0,0118	516,500	0,000
BARRIO_RECOLETA	0,5936	0,0098	603,800	0,000
BARRIO_ CHACARITA	0,5256	0,0128	411,410	0,000
BARRIO_ BARRACAS	0,5156	0,0120	429,080	0,000
BARRIO_PARQUE. CHACABUCO	0,4847	0,0096	504,846	0,000
BARRIO_ SAAVEDRA	0,4739	0,0102	465,450	0,000
BARRIO_FLORES	0,4727	0,0080	591,450	0,000
BARRIO_VERSALLES	0,4555	0,0263	173,108	0,000
BARRIO_MONTE. CASTRO	0,4076	0,0145	281,979	0,000
BARRIO_VILLA. CRESPO	0,4040	0,0092	439,441	0,000
BARRIO_ BALVANERA	0,3900	0,0070	553,774	0,000
BARRIO_RETIRO	0,3664	0,0098	372,537	0,000
BARRIO_PARQUE. CHAS	0,3343	0,0170	196,349	0,000
BARRIO_FLORESTA	0,3202	0,0157	203,302	0,000
BARRIO_VELEZ. SARSFIELD	0,3200	0,0154	208,180	0,000
BARRIO_ AGRONOMIA	0,3009	0,0268	112,089	0,000
BARRIO_SAN. CRISTOBAL	0,2802	0,0079	356,668	0,000
BARRIO_PALERMO	0,2484	0,0083	300,674	0,000
BARRIO_PARQUE. PATRICIOS	0,2470	0,0172	143,511	0,000

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Variable	Coefficiente	Error estandar	Valor-t	P-valor
BARRIO_BOCA	0,2110	0,0165	127,846	0,000
BARRIO_BELGRANO	0,2064	0,0056	367,303	0,000
BARRIO_PATERNAL	0,1908	0,0195	98,115	0,000
BARRIO_NUÑEZ	0,1470	0,0073	201,330	0,000
BARRIO_COLEGIALES	0,1469	0,0071	206,266	0,000
BARRIO_LINIERS	0,1448	0,0158	91,709	0,000
HAS_GYM	0,1335	0,0050	267,544	0,000
HAS_HALF_BATH	0,1289	0,0035	368,440	0,000
BARRIO_NUEVA. POMPEYA	0,1078	0,0252	42,843	0,000
FULL_BATHROOMS	0,1064	0,0024	450,533	0,000
HAS_SWIMMING_ POOL	0,1061	0,0045	233,083	0,000
BARRIO_SAN. TELMO	0,1053	0,0128	82,282	0,000
BARRIO_ALMAGRO	0,0985	0,0068	145,835	0,000
HAS_BALCONY	0,0949	0,0035	274,340	0,000
HAS_DRESSING_ ROOM	0,0744	0,0052	142,827	0,000
ROOMS	0,0681	0,0013	510,824	0,000
BARRIO_MATADEROS	0,0664	0,0148	44,935	0,000
HAS_PLAYROOM	0,0648	0,0085	76,594	0,000
BARRIO_BOEDO	0,0522	0,0106	49,039	0,000
disposicion_frente	0,0432	0,0030	144,997	0,000
HAS_JACUZZI	0,0388	0,0083	46,739	0,000
HAS_SECURITY	0,0291	0,0048	60,051	0,000
HAS_MULTIPURPOSE_ ROOM	0,0266	0,0045	58,464	0,000
PARKING_LOTS	0,0245	0,0011	218,221	0,000
TOTAL_AREA	0,0054	37,479	1,449,928	0,000
colectivos	0,0002	21,705	75,900	0,000

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Variable	Coeficiente	Error estandar	Valor-t	P-valor
colectivos_2	-0,0000	0,0000	-69,630	0,000
salud	0,0000	0,0000	81,481	0,000
salud_2	0,00002	0,0000	105,230	0,000
trenes	0,0000	0,0000	96,194	0,000
trenes_2	-0,000001	0,0000	-73,840	0,000
subte	0,0000	0,0000	87,212	0,000
subte_2	-0,000001	0,0000	-173,970	0,000
espacios_verdes	-30,788	0,2858	-107,721	0,000
espacios_verdes_2	0,000001	0,0000	143,980	0,000
universidades	0,0000	0,0000	-83,203	0,000
universidades_2	-0,00001	0,0000	-61,430	0,000
carcel	0,0000	0,2183	-0,8354	0,404
carcel.2	0,000000862	0,000000879	0,9810	0,327
HAS_KITCHEN	-0,0268	0,0037	-72,266	0,000
HAS_CLOSET	-0,0365	0,0078	-46,882	0,000
transit_stations_percent_change_from_baseline	-0,0399	0,0039	-102,532	0,000
BARRIO_SAN. NICOLAS	-0,0433	0,0097	-44,865	0,000
HAS_TAP_ WATER	-0,0456	0,0040	-115,149	0,000
HAS_PATIO	-0,0485	0,0056	-86,311	0,000
HAS_STUDY	-0,0692	0,0092	-75,160	0,000
parks_percent_change_from_baseline	-0,0711	0,0027	-260,543	0,000
grocery_and_pharmacy_percent_change_from_baseline	-0,0771	0,0024	-319,359	0,000
residential_percent_change_from_baseline	-0,1009	0,0042	-241,164	0,000

Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Variable	Coeficiente	Error estandar	Valor-t	P-valor
workplaces_ percent_change_ from_baseline	-0,1830	0,0063	-289,670	0,000
BARRIO_ CONSTITUCION	-0,1920	0,0112	-172,060	0,000
HAS_INDOOR_ FIREPLACE	-0,2442	0,0592	-41,285	0,000
retail_and_ recreation_ percent_change_ from_baseline	-0,4158	0,0056	-744,207	0,000
HAS_LIFT	-0,0188	0,0048	-39,218	0,000
HAS_TELEPHONE_ LINE	0,0175	0,0048	36,596	0,000
HAS_NATURAL_ GAS	-0,0116	0,0037	-31,403	0,002
HAS_DINNING_ ROOM	-0,0108	0,0045	-24,122	0,016
HAS_HEATING	0,0088	0,0038	23,278	0,020
HAS_MAID_ROOM	0,0119	0,0056	21,277	0,033
HAS_TERRACE	-0,0081	0,0045	-17,872	0,074
HAS_PARTY_ ROOM	0,0223	0,0128	17,474	0,081
HAS_BEDROOM_ SUITE	0,0164	0,0098	16,795	0,093
HAS_TENNIS_ COURT	0,0710	0,0454	15,641	0,118
HAS_BREAKFAST_ BAR	0,0204	0,0134	15,208	0,128
HAS_GUEST_ PARKING	0,0420	0,0280	14,988	0,134
HAS_AIR_ CONDITIONING	0,0043	0,0031	13,845	0,166
HAS_COMMON_ LAUNDRY	-0,0187	0,0137	-13,678	0,171
HAS_LIVING_ ROOM	0,0059	0,0044	13,412	0,180
HAS_GRILL	0,0047	0,0040	11,877	0,235

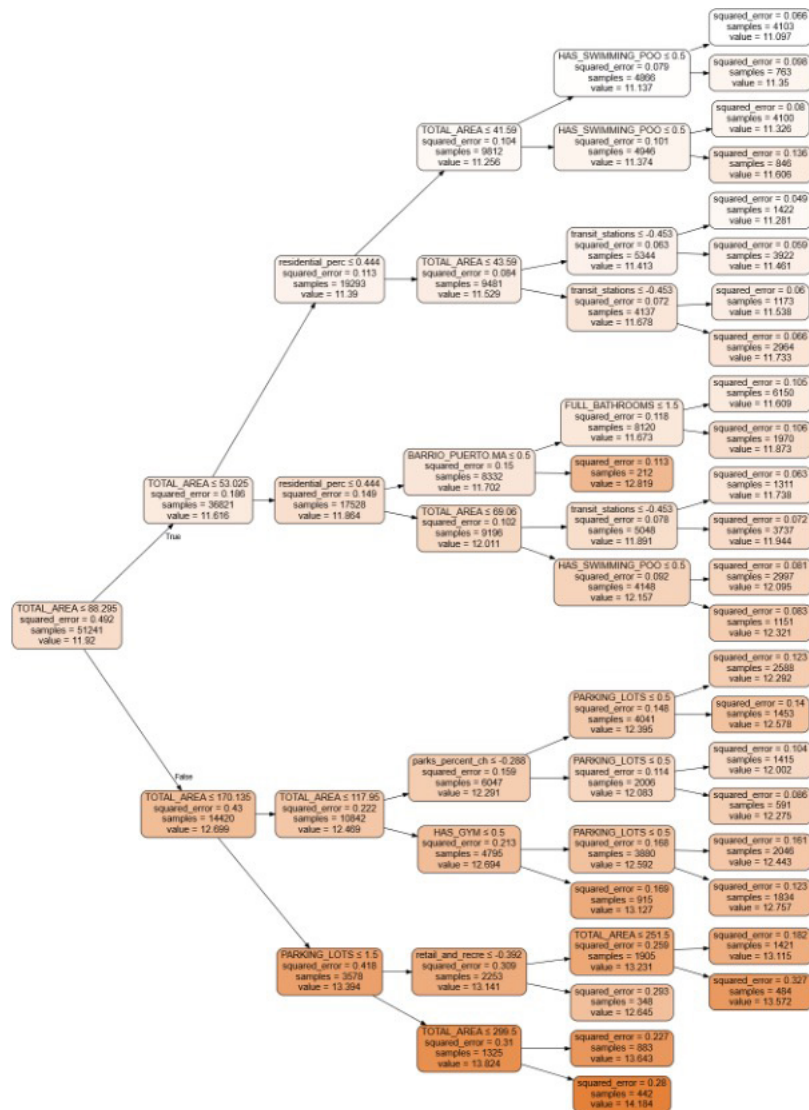
Valorización de inmuebles en Buenos Aires, a partir de información en línea. Un enfoque de precios hedónicos

Emiliano Gutierrez; Laura Lanzarini; Fernando Delbianco; Rocío Cecchini

Variable	Coeficiente	Error estandar	Valor-t	P-valor
HAS_CINEMA_HALL	0,0400	0,0431	0,9284	0,353
HAS_CISTERN	-0,0181	0,0219	-0,8280	0,408
HAS_GARDEN	-0,0048	0,0077	-0,6225	0,534
HAS_INTERNET_ACCESS	-0,0026	0,0045	-0,5790	0,563
BARRIO_MONSERRAT	-0,0048	0,0090	-0,5377	0,591
HAS_BUSINESS_CENTER	0,0191	0,0364	0,5260	0,599
HAS_PLAYGROUND	-0,0121	0,0263	-0,4611	0,645
HAS_BOILER	0,0019	0,0076	0,2507	0,802
HAS_LAUNDRY	-0,0008	0,0033	-0,2488	0,804
BARRIO_PARQUE.AVELLANEDA	0,0046	0,0184	0,2484	0,804
HAS_ROOF_GARDEN	0,0026	0,0352	0,0750	0,940
R ² : 0,817				
R ² ajustado: 0,817				
AIC: 221,537				
BIC: 229,674				
F-estadístico: 2513				
Prob (F-statistic): 0,00				

Fuente: Elaboración propia

Figura 4: CART completo



Fuente: Elaboración propia.