

# Problem Set 3: Making Money with ML?

Big Data y Machine Learning para Economía Aplicada

Gustavo Adolfo Castillo Álvarez (201812166),  
Alexander Almeida Ramírez (202225165),  
Jorge Luis Congacha Yunda (201920042) y  
Jaime Orlando Buitrago González (200612390)

2024-05-26

## 1 Introducción

En su *paper*, Rosen (1974) proporciona una aproximación metodológica (econométrica) para estimar precios hedónicos. Más específicamente, explora los equilibrios del mercado y las decisiones de los agentes y los efectos en el bienestar para evidenciar la relación entre los precios de los bienes y sus diferentes características. Esta fue una contribución importante para estimar los precios (Hedónicos) de los bienes a partir de las características o atributos (precios implícitos) a partir de la observación de los precios de bienes heterogéneos sus características asociadas.

Como uno de los *papers* más citados del *Journal of Political Economy* (Greenstone, 2017), se obtiene una aproximación al comportamiento de la oferta y demanda (equilibrio) a partir del proceso generador de datos entre los bienes y sus características. Para el mercado inmobiliario -en el que los comportamientos de los agentes, oferta y demanda, y equilibrios no se pueden explicar en su totalidad con la teoría microeconómica convencional-, esta aproximación ha sido de singular importancia, más aún en el contexto de la ciudad de Bogotá.

Así por ejemplo, Perdomo (2011) utiliza los precios hedónicos como método econométrico para estimar el precio de las viviendas y posteriormente el de la cercanía a Transmilenio; Carriazo et al. (2013) identifican cómo de no incluir la calidad del aire en las estimación a través de precio hedónicos conlleva un sesgo en los precios de las viviendas en Bogotá. Lozano & Anselin (2012) utilizan variables espaciales en los submercados de la viviendas para estimar los precios en Bogotá, identificando una gran capacidad predictiva, pero afectada (sobre estimación) por los estratos socioeconómicos. También se encuentran Medina et al. (2007) quienes demuestran como los subsidios aplicados a las tarifas de los servicios públicos se transfieren al precio de las viviendas.

Esta breve exploración bibliográfica refleja los usos y potencialidades de la contribución de Rosen (1974) al mercado inmobiliario con énfasis en la ciudad de Bogotá. Claro está, como la capital de Colombia, Bogotá es una ciudad heterogénea con dinámicas poblacionales, sociales y económicas distribuidas a por todo el territorio urbano, las cuales pueden afectar los precios del mercado inmobiliario. Más aun, en una localidad como Chapinero, la cual es una de las áreas centrales de la ciudad; con una gran catnidad de población flotante; donde se cuenta con vías que se conectan con el sistema de movilidad de la ciudad; se tienen usos del suelo asociados a actividades comerciales y residenciales; y se tienen parques, canales y corredores ecológicos (Decreto 468, 2006).

Por estas características, tal vez Chapinero sea la localidad con la mayor complejidad de formas de habitar el territorio de la ciudad del país, lo cual la hace un espacio geográfico donde convergen múltiples factores que afectan los precios de las viviendas. Por tal razón, construir uno o varios modelos predictivos de los precios es desafiante y una necesidad, ante una potencial inversión que compre la mayor cantidad de predios, gastando lo menos posible.

Ahora bien, para construir el o los modelos se utilizó una muestra de los anuncios publicados en la página web [Properati](#). Este sitio web de origen argentino y con presencia en varios países de Latinoamérica, le permite a propietarios publicar un bien inmueble con la mayor cantidad de información (precio, características y amenidades, ubicación y descripción), y a personas interesadas en adquirir o arrendar un bien contactarse con el comprador. Es decir, se utilizó una página web que ofrece servicios de intermediación, pero que permite caracterizar la oferta de bienes en Bogotá.

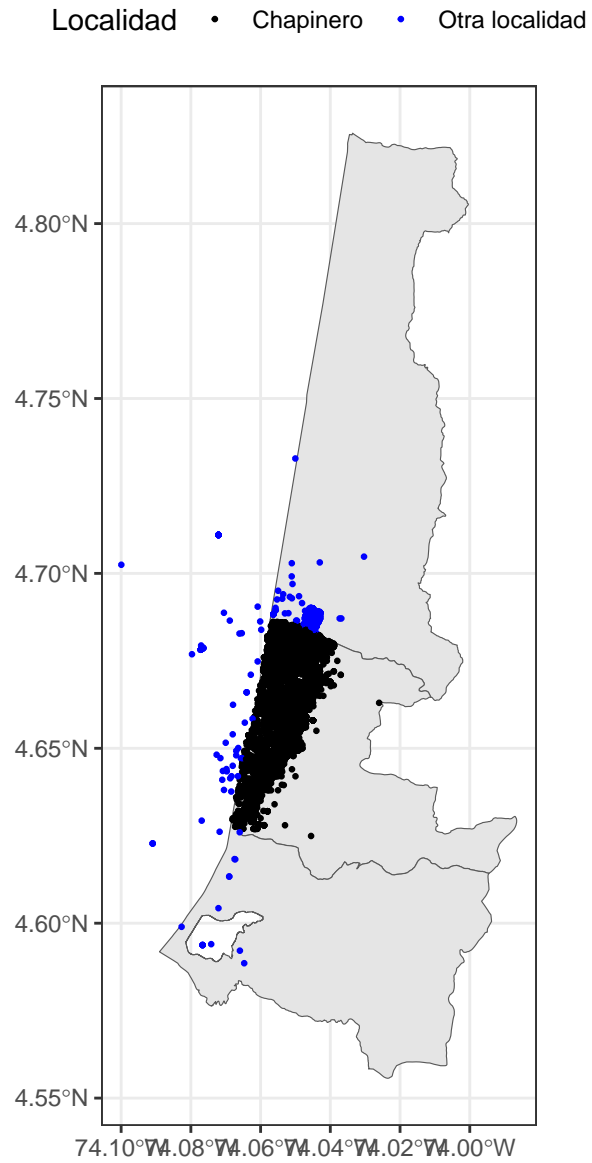
De igual manera, se utilizó la información de la página web Infraestructura de Datos Espaciales de Bogotá (Ideca), un portal de libre acceso y consulta de la Unidad Administrativa Especial de Catastro Distrital (UAECD), la cual permitió asociar información geográfica relevante a la muestra de [Properati](#).

## 2 Datos y Métodos

Los datos principales consistieron en los archivos descargados de la página de datos de la [competencia en Kaggle](#). competencia.

La información espacial de referencia se tomó del Mapa de Referencia para Bogotá D.C. del Ideca disponible [aquí](#). Al cruzar la capa de los polígonos las localidades con la tabla de *test* observamos que 593 propiedades se encuentran fuera de Chapinero (Figura 1), que es la localidad en la cuál se predecirán los precios de las viviendas.

Figura 1: Ubicación datos de validación



Características de la propiedad:

- Precio
- Fecha de publicación de clasificado
- Superficie total y superficie cubierta (80% y 76% valores faltantes)
- Número de habitaciones
- Número de baños
- Tipo de propiedad: apartamento o casa
- Ubicación georreferenciada

Spatial Ammenities: - parques - tiendas o centros comerciales - # Modelo y resultados

### 3 Conclusiones

### 4 Referencias bibliográficas

- Alcaldía Mayor de Bogotá. (20 de noviembre de 2006). *Por el cual se reglamenta la UPZ 99, Chapinero, ubicada en la localidad Chapinero* (Decreto 468 de 2006, Ed.).
- Carriazo, F., Ready, R., & Shortle, J. (2013). Using stochastic frontier models to mitigate omitted variable bias in hedonic pricing models: A case study for air quality in Bogotá, Colombia. *Ecological Economics*, 91, 80-88.
- Greenstone, M. (2017). The Continuing Impact of Sherwin Rosen's «Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition». *Journal of Political Economy*, 125(6), pp. 1891-1902.
- Lozano, n., & Anselin, L. (2012). Is the price right?: Assessing estimates of cadastral values for Bogotá, Colombia. *Regional Science. Policy & Practice*, 4(4), 495-508.
- Medina, C., Morales, L., Bernal, r., & Torero, M. (2007). Stratification and Public Utility Services in Colombia: Subsidies to Households or Distortion of Housing Prices? [with Comments]. *Economía*, 7(2), 41-99.
- Perdomo, J. (2011). A methodological proposal to estimate changes of residential property value: case study developed in Bogotá. *Applied Economics Letters*, 18(16), 1577-1581.
- Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, 82(1), 34-55.

### 5 Repositorio

El repositorio se puede consultar en el siguiente enlace:

- [Problem Set 3: Making Money with ML?](#)