

# *Cook County, IL: Índices de precios de viviendas*

Luis Alejandro Rubiano Guerrero  
202013482

Andres Felipe Rosas Castillo  
0

Carlos Andrés Castillo Cabrera  
202116837

la.rubiano@uniandes.edu.co a.rosasc@uniandes.edu.co ca.castillocl@uniandes.edu.co

## I. INTRODUCCIÓN

A través de este informe se construyen y se comparan **cuatro índices de precios de vivienda** para *Cook County, Illinois* para el periodo comprendido entre 2000 y 2020. Las metodologías correspondientes a cada uno de los índices son las siguientes: (i) y (ii) índices de precios hedónicos, (iii) Índice de precios de la vivienda usada (IPVU), y (iv) estimador de efectos fijos con errores agrupados a nivel de propiedad. El objetivo de este informe es comparar y contrastar los resultados obtenidos a través de cada una de las metodologías, así como discutir las ventajas y desventajas de cada una de ellas.

En la sección II. se describe el conjunto de datos utilizado, así como el proceso de limpieza y tratamiento de los datos. La sección III. detalla la metodología utilizada para cada uno de los índices. En la sección IV. se presentan los resultados obtenidos, así como un análisis comparativo entre los diferentes índices. Finalmente, en la sección V. se presentan las conclusiones del informe y en la sección VI. se proporciona información adicional para la reproducibilidad del análisis.

## II. DATOS Y PREPARACIÓN

### 1. Descripción del conjunto de datos

El conjunto de datos utilizado en este informe es `dataTaller01_PriceIndices.Rds`, el cual contiene información sobre ventas de viviendas en *Cook County, Illinois* entre los años 2000 y 2020. En total, el conjunto de datos contiene 427.649 observaciones y 31 variables. Estas variables incluyen el `pin` (identificador único de la propiedad), `year` (fecha de venta de la vivienda), `sale_price` (precio de venta de la vivienda), `township_code` (código local correspondiente al *township* donde se ubica la propiedad), así como 27 covariables continuas y categóricas de características estructurales y de ubicación de las viviendas.

### 2. Limpieza y tratamiento

Se encontraron valores faltantes para 1060 de las observaciones, en las variables `building_sqft`, `num_bedrooms`, `num_rooms`, `num_full_baths`, `num_half_baths` y `num_fireplaces`. Estas observaciones fueron eliminadas del conjunto de datos.

Adicionalmente, después de eliminar las anteriores observaciones, se encontraron 151.929 observaciones con valores faltantes en la variable `land_sqft`, lo cual representa aproximadamente el 36% del total de las observaciones. Estas fueron

tratadas de manera diferente dependiendo de la metodología utilizada para construir los índices de precios de vivienda.

Adicionalmente se creó la variable `log_sale_price`, la cual corresponde al logaritmo natural del precio de venta de la vivienda, la cual es la variable dependiente de interés en los respectivos índices de precios.

## III. METODOLOGÍA