**Datos**

**Variables externas**

En esta sección se describen las variables que fueron tomadas de fuentes externas a las bases de datos proveídas en el ejercicio. La selección de alguna de estas variables se hizo a partir de un estudio anterior [XXX1], en el que identifican la importancia que tienen el estrato socioeconómico, la distancia a estaciones de TransMilenio y la distancia a parques.

Por otro lado, de acuerdo con [XXX2], los instrumentos de Unidad de Planeamiento Zonal (UPZ) –reemplazadas en el 2022 por las Unidades de Planeamiento Local (UPL)– fueron utilizados para agrupar zonas homogéneas de la ciudad, por lo que pueden ser segmentaciones que capturen parte del precio de la vivienda.

Distancia a estaciones de TransMilenio

Las ubicaciones de las estaciones de TransMilenio fueron tomadas de la página principal de la entidad (disponibles en un archivo geojson), teniendo en cuenta que al utilizar el atributo de paradas de buses en la API de OpenStreetMap salían mezcladas algunas cuantas estaciones para buses que no son BRT. En total, se tienen 149 estaciones de TransMilenio distribuidas en 12 troncales del sistema.

Estratos socioeconómicos

En la plataforma distrital de Datos Abiertos, se tiene, entre otros, las manzanas georreferenciadas según su estrato socioeconómico. Este atributo es asignado por las alcaldías para clasificar a una vivienda según características físicas externas e internas, su entorno inmediato y su contexto habitacional o funcional (DANE, 2023).

Por la georreferenciación inadecuada de algunas viviendas y porque hay áreas de la ciudad que no están incluidas dentro de estas manzanas (vías, áreas públicas, zonas comerciales), se asignó el estrato de la vivienda según su cercanía al centroide del polígono[[1]](#footnote-1), es decir, la vivienda tenía el valor del estrato del polígono cuyo centroide estuviera más cerca. En la siguiente figura se muestra el resultado de la asignación:

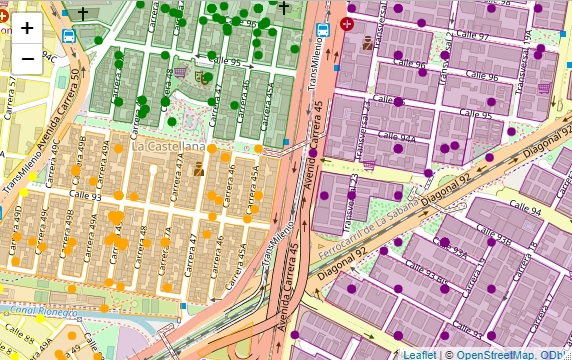


Figura 1. Muestra de asignación de los estratos socioeconómicos a las viviendas. Los puntos señalados en los recuadros rojos son ejemplos de la georreferenciación inadecuada, lo que resulta en que haya viviendas por fuera de las manzanas residenciales.

La zona de la anterior figura es muy interesante porque en un radio de 1000 metros se pueden ver cuatro estratos socioeconómicos: tres (amarillo), cuatro (naranja), cinco (verde) y seis (purpura). Los puntos representan las viviendas de las bases de datos y los polígonos sombreados son los que estaban disponibles en Datos Abiertos

Unidades de Planeamiento Zonal

Los polígonos de la UPZ fueron tomados del Laboratorio Urbano Bogotá (Laboratorio Urbano, 2012) y emparejados con las viviendas de las bases de datos según su ubicación espacial. En total, se tienen 112 UPZ correspondientes 19 localidades dentro de la ciudad.

Bibliografía

[XXX1] Toloza, J., Melo-Marinez, O., Azcarate-Romero, J. (2021). Determinantes del precio de la vivienda nueva en Bogotá para el año 2019: una aproximación a través de un modelo semiparamétrico de regresión espacial. doi:10.17230/ingciencia.17.34.2

[XXX2] Chacón, V. (2018). ZONAS HOMOGENEAS ECONOMICAS – ESTUDIO DE VALOR DE M2 ÁREA PRIVADA CONSTRUIDA DE EDIFICACIONES SOMETIDAS A PROPIEDAD HORIZONTAL EN LA UPZ No. 101 TEUSAQUILLO EN LA LOCALIDAD DE TEUSAQUILLO - BOGOTÁ D.C. <https://repository.udistrital.edu.co/bitstream/handle/11349/14505/ChaconRodriguezVictorDavid2018.pdf?sequence=1>

[XXX3] DANE (25 de octubre del 2023). Preguntas frecuentes estratificación. <https://www.dane.gov.co/files/geoestadistica/Preguntas_frecuentes_estratificacion.pdf>

[XXX4] Laboratorio Urbano Bogotá (22 de octubre del 2023). Población UPZ Bogotá. <https://bogota-laburbano.opendatasoft.com/>

1. Inicialmente se realizó la asignación con el criterio de que la vivienda estuviera dentro del polígono, sin embargo, alrededor del 30% de los datos no estaban dentro de las manzanas delimitadas. Para ilustrar véase los puntos señalados en la Figura 1, que están parques, zonas verdes o vías principales. [↑](#footnote-ref-1)