**Problem Set 3 – Grupo 2**

Alejandro García 201810340 - Andrés Torres 202115992 - Ignacio Serrano 201618789

[**https://github.com/AndresT2022/ProblemSet3.git**](https://github.com/AndresT2022/ProblemSet3.git)

**Introducción**

El presente documento pretende realizar la predicción del precio de las viviendas en Chapinero (Bogotá) y El Poblado (Medellín) por medio de datos espaciales suministrados por el docente y datos espaciales obtenidos en OpenStreetMap.

Se realizaron pruebas con 5 modelos (modelo 1, 2, 3,4 y 5) del cual se definió escoge el modelo XXX. Este modelo es el que presento mejor XXXX.

**Datos**

De los datos suministrados por el docente, se filtró la información para El Poblado y Chapinero, obteniendo xxx cantidad de observaciones para El Poblado y xxx para Chapinero.

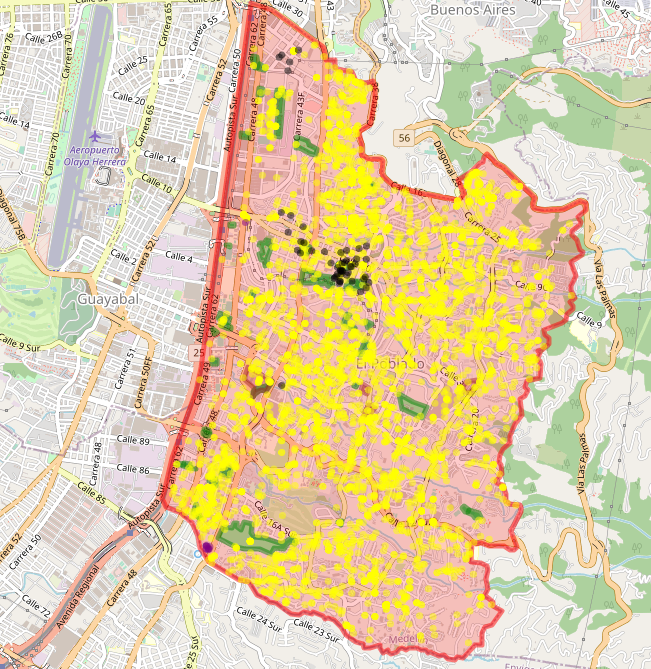
Revisando la literatura asociada a la predicción del precio de viviendas, se encontró que el área de la vivienda y el número de baños son variables que explican muy bien el precio de la vivienda, por lo cual, se realizó una imputación de estas variables con ayuda de la descripción que trae la base de datos y realizando buffer´s, lo anterior debido a que estas variables presentaban muchos missing values. Adicional a las dos variables mencionadas, se utilizó el número de habitaciones, la cual no tenía missing values.

Por medio de OpenStreetMap se incorporaron tres variables al modelo, las cuales son la distancia mas cercana al parque, al bar y a las zonas de comercio. En ese orden de ideas, se utilizaron cinco variables en total para los modelos de predicción propuestos.

*Tabla 1. Estadísticas descriptivas - Base de datos prueba*

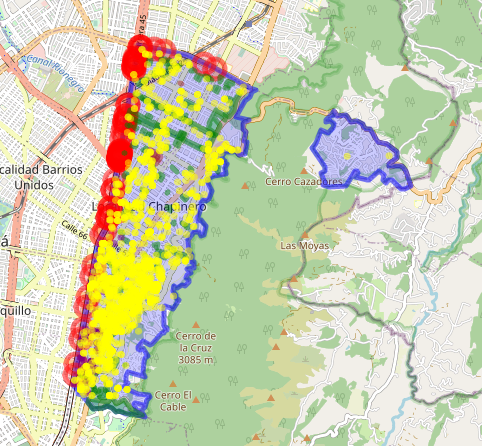
En el siguiente mapa, se muestra la distribución de la muestra (amarillo), los parques (verde), los bares (en negro) y las zonas de comercio (café) para El Poblado.

*Mapa 1. El Poblado*



En el siguiente mapa, se muestra la distribución de la muestra (amarillo), los parques (verde), los bares (en negro), las estaciones de buses (rojo) y las zonas de comercio (café) para Chapinero.

*Mapa 2. Chapinero*



**Modelos y resultados.**

**Conclusión y recomendaciones.**