

HW 1: Efectos espacio-contextuales en el precio de venta de terrenos

Prof. Marcos Herrera-Gómez
E-mail: mherreragomez@udesa.edu.ar

El objetivo de este trabajo es realizar la incorporación de variables espacio-contextuales bajo un modelo econométrico del precio de venta de terrenos en CABA. Los datos son provenientes de la página <https://data.buenosaires.gob.ar/dataset/terrenos-valor-oferta>. La base de datos es de 2019 y contiene 2321 terrenos ofrecidos para la venta con la siguiente información:

- Variables:
 - M2total: total de metros cuadrados.
 - PrecioUSD: precio en dólares.
 - Barrio.
 - Comuna.

Se solicita que realice las siguientes consignas:

1. Revise la base de datos, cuidadosamente, y construya como variable dependiente el logaritmo neperiano del precio en dólares y especifique un modelo de la siguiente manera:

$$\ln(\text{precioUSD}) = \text{const} + \beta_1 \ln(M2total) + \sum_{j=1}^{J-1} \delta_j D_{\text{barrio}-j} + u, \quad (1)$$

donde $D_{\text{barrio}-j}$ es una variable binaria que toma el valor 1 cuando el barrio es j y 0 si no lo es. Se deben incluir $J - 1$ variables ficticias en el modelo.

- Interpretar el coeficiente estimado relacionado a M2total.

2. Utilizando una matriz de contigüidad de 2do orden (incluyendo el 1er orden), W_{cont}^2 , comparar los resultados del test I de Moran univariado para la variable $\ln(\text{precioUSD})$ y para los residuos del modelo (1). Comentar brevemente los resultados.
3. Calcular las siguientes variables espaciales:
 - a) cantidad de robos y hurtos de 2018 en una zona de influencia (buffer) de 500 metros a cada terreno (se incluye csv con lat-long geográficas).
 - b) distancia a la estación de subte más cercana (se incluye base).
 - c) distancia a la estación de tren más cercana (se incluye base).
 - d) distancia al Obelisco como proxy del centro de negocios de la ciudad (generar un archivo csv con latitud y longitud para reproyectar en QGIS).

Presentar un cuadro con estadísticas descriptivas de las tres variables (media, mediana, sd, mínimo, máximo) y las correlaciones respecto a la variable dependiente.

4. Re-estimar el modelo de la ecuación (1) incorporando las variables del punto 3:
 - a) ¿Se mantienen los resultados obtenidos del punto 1.?,
 - b) Interpretar el coeficiente de cada nueva variable espacial, ¿las variables espaciales tienen el signo esperado?,
 - c) ¿Los residuos son no-autorrelacionados espacialmente bajo W_{cont}^2 ?
5. Realizar un análisis LISA sobre los residuos del modelo del punto 4, usando W_{cont}^2 , con perm=999 y p-valor=0.001, para detectar outliers espaciales.
6. Re-estimar el modelo del punto 4, incorporando (si fuese necesario) las categorías de outliers detectadas en el punto 5 (una variable ficticia para los outliers H-L y otra variable ficticia para L-H). Contrastar la presencia de efectos espaciales usando los tests I de Moran y LM's usando W_{cont}^2 . ¿Puede proponerse algún modelo espacial según los resultados de los tests (no es necesario estimar dicho modelo)?.

El análisis de regresión puede realizarse con GeoDa o R. Usar como proyección Posgar 2017 - 6 (EPSG: 5348).

■ Condiciones de entrega:

- Trabajo individual o grupal (hasta 3 integrantes).
- Extensión máxima de 5 páginas en formato pdf.
- Debe adjuntarse la base final de datos analizada solamente (no incluir las reproyecciones para generar las variables espaciales).
- Fecha límite de entrega: 20 de Agosto de 2025.

Enviar por mail a mherreragomez@udesa.edu.ar todos los documentos del homework en un único archivo comprimido bajo el nombre de los autor: hw1_apellido(s).rar.