

UNIVERSIDAD
NACIONAL
DE COLOMBIA

TALLER NO 2 -PARCIAL:ASOCIACIÓN DE LOS ROBOS EN BUENOS AIRES ARGENTINA
CON EL NÚMERO DE HOGAR CON NECESIDADES BÁSICAS INSATISFECHAS
,HOGARES Y VIVIENDAS.

EDGAR JOSÉ CORREA DIAZGRANADOS

ESTADÍSTICA ESPACIAL.
DEPARTAMENTO ESTADÍSTICA

DOCENTE MARTHA PATRICIA BOHORQUEZ CASTAÑEDA.

BOGOTÁ
JULIO 2020

Índice

1. Introducción	3
2. Planteamiento del problema	3
3. Objetivos	3
4. Metodología	4
4.1. Base de datos	4
4.1.1. Variables	4
4.2. Datos	5
4.2.1. Superficie	5
4.2.2. Puntos	5
4.2.3. Mapa por barrios	6
4.3. Modelo sin tener en cuenta autocorrelación espacial	7
4.3.1. Modelo Poisson	8
4.4. Matriz de Pesos Espaciales	9
4.4.1. Vecinos físicos	9
4.4.2. Gráficas vecinas	10
4.4.3. Vecinos a distancia o K-nearest	11
4.5. Moran's I	11
4.6. Clusters	14
5. Modelos	15
5.1. Vectores de Moran ajustados	15
5.2. Modelo de Retardo Espacial	15
5.3. Modelo de Durbin	16
5.4. Comparaciones	17
6. Conclusiones	17

1. Introducción

Los datos por área , se refieren a los datos que se obtienen sobre lugares geográficos, es decir, lugares en la superficie terrestre. El uso de este tipo de datos reconoce “el rol clave que conceptos espaciales como la distancia, la ubicación, proximidad, vecindario y región juegan en la sociedad humana”, permitiendo enfrentar fenómenos desde una perspectiva multivariada y multidimensional. Este tipo de datos permite responder preguntas tales como si existe influencia en términos de proximidad de los barrios vecinos sobre la adopción de ciertas políticas, pero también resultan ser una herramienta clave en la exploración de los datos para observar la existencia de patrones territoriales en los fenómenos estudiados.

2. Planteamiento del problema

Un delito es una acción que va en contra de lo establecido por la ley y que es castigada por ella con una pena grave. En la ciudad de Buenos Aires desde el año 2017 se creó la estrategia del mapa de delitos , El Mapa del Delito es una herramienta online con estadísticas fiables de criminalidad en la Ciudad (ver figura 3.a) . Gracias a estos datos, es posible dimensionar el problema de la inseguridad, establecer patrones, realizar diagnósticos certeros y evaluar la efectividad de las respuestas.

En el marco de las políticas de cooperación internacional aplicadas durante los años setenta en los países del entonces llamado Tercer Mundo, la noción de pobreza basada en el enfoque de la subsistencia fue ampliada a partir de una serie de trabajos efectuados por diferentes organismos de Naciones Unidas.

Con la aparición del paradigma de “necesidades básicas”, el centro de interés comenzó a trasladarse hacia un conjunto más vasto de satisfactores, no reducidos a la dimensión económica. Se logra así una ampliación del campo de referencia del concepto de pobreza.[1] .

La definición de pobreza mediante este enfoque implica la aplicación de un método normativo de carácter multidimensional de medición directa de las condiciones de vida, el cual puede ubicarse entre los métodos que aceptan una visión absoluta de la pobreza, y dentro de ellos, en los denominados “necesidades básicas” –considerando tanto las necesidades biológicas satisfechas de consumo privado así como también aquellas provistas a través de servicios como los de agua potable, educación y salud-. De esta manera, el método elegido es el de Necesidades Básicas Insatisfechas (NBI).

A lo largo de 2019, la Policía de la Ciudad realizó más de 28 mil detenciones, con 19 mil policías afectados a la calle, en una fuerza total de 26.500 efectivos. Es interesante buscar soluciones a problemas locales como la inseguridad para esto es importante entender el contexto de la ciudad y como los hogares con necesidades básicas insatisfechas se involucran en este problema, se debe analizar y plantear estrategias.

3. Objetivos

- Identificar variables que son significativas en el número de robos en la ciudad de Buenos Aires
- Identificar variables que son significativas en el número de Hogares con necesidades básicas insatisfechas en la ciudad de Buenos Aires

- Identificar en que Barrios ocurren la mayoría de robos
- Identificar en que Barrios existen mayores Hogares con necesidades básicas insatisfechas
- Ver la relación que poseen entre robos y tratar de establecer si existe una relación entre ellos
- Ver la relación que poseen los Hogares con necesidades básicas insatisfechas y tratar de establecer si existe una relación entre ellos

4. Metodología

4.1. Base de datos

En este caso se toma el registro de los primeros 5 días de delitos en 2019 de la base de datos del gobierno argentino [2] y la base de datos de los hogares [3], donde se encuentra la información del número de hogares, viviendas, área km², hogares con necesidades básicas insatisfechas.

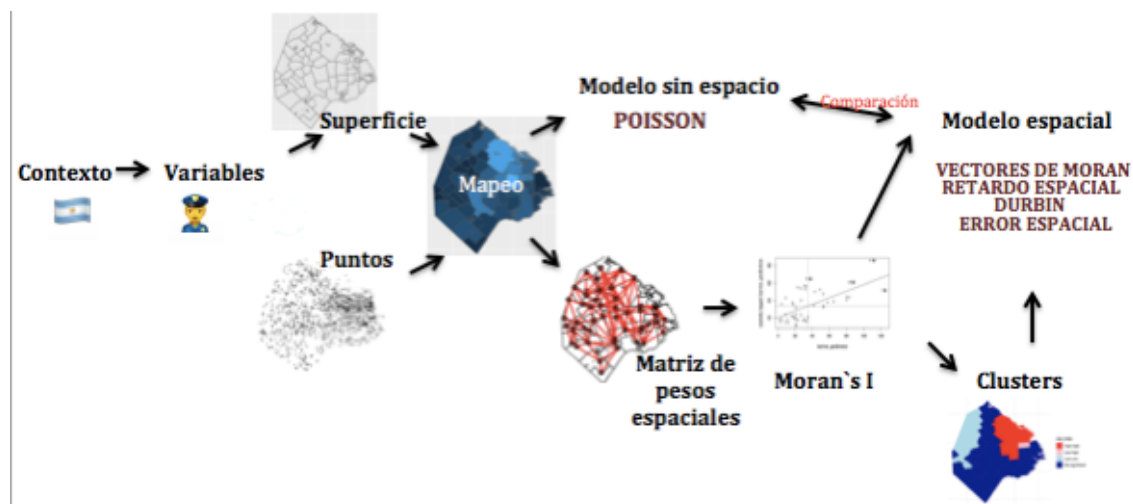


Figura 1: Resumen metodología

A partir de estas bases se busca establecer un modelo donde se pueda evidenciar que variables hacen más probable la ocurrencia de un delito y como este se evidencia en los barrios de la capital Argentina, para realizar dicho modelo se debe seguir la metodología planteada en la Figura 1.

4.1.1. Variables

Las variables que se tienen en cuenta son:

- Robos: es un delito contra el patrimonio, consistente en el apoderamiento de bienes ajenos de otras personas de manera fraudulenta, empleando para ello fuerza en las cosas o bien violencia o intimidación en las personas, en este caso se presentan los robos registrados los primeros 5 días de año en Buenos Aires
- Hogares NBI: El concepto de necesidades básicas insatisfechas (NBI) permite la delimitación de grupos de pobreza estructural y representa una alternativa a la identificación de la pobreza considerada únicamente como insuficiencia de ingresos.

- Hogares: se usa para designar a un lugar donde un individuo o grupo de individuos habita.
- Viviendas: Es todo aquello que se ha construido, adaptado o dispuesto para el alojamiento de personas.
- Franja Horario: Hora en la que el robo se realizó
- población: Número de habitantes
- Área: área geográfica en km²

4.2. Datos

4.2.1. Superficie

Buenos Aires es la capital y ciudad más poblada de la República Argentina. Es una ciudad autónoma que constituye uno de los 24 distritos, que conforman el país. Tiene sus propios poderes ejecutivo, legislativo y judicial. Está situada en la región centro-este del país, sobre la orilla sur del Río de la Plata, en la región pampeana. La Ciudad de Buenos Aires fue cedida por la provincia de Buenos Aires para que fuera la capital federal del país; pero en virtud de la Constitución Nacional es una ciudad autónoma. Se toma el mapa dividido por calles (ver Figura 2), se trabaja con los datos en GEOJSON, por su facilidad de obtención.

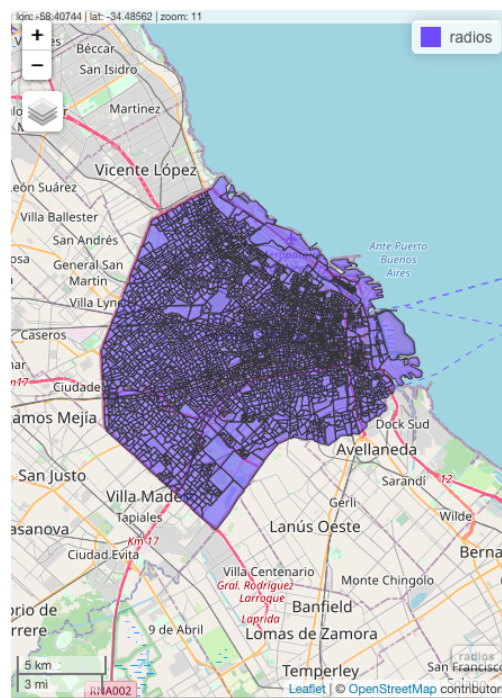


Figura 2: Mapa de la Ciudad de Buenos Aires, calles

4.2.2. Puntos

Los puntos amarillos son los robos ocurridos en la ciudad de Buenos Aires los primeros 5 días del año 2019

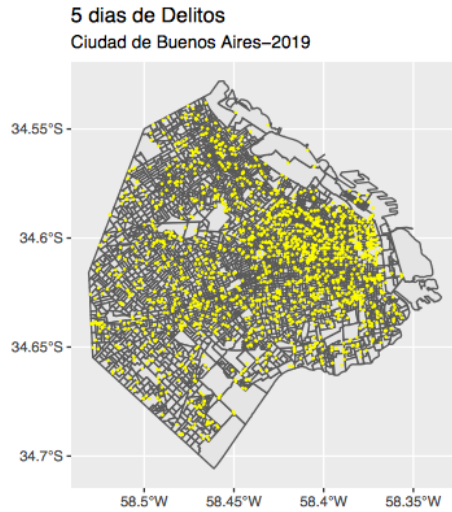


Figura 3: Robos primeros 5 días 2019

Para entender mejor el comportamiento de los datos, se hace una comparación entre el mapa del gobierno de Buenos Aires y un POWER BI propio(ver Figura 4).

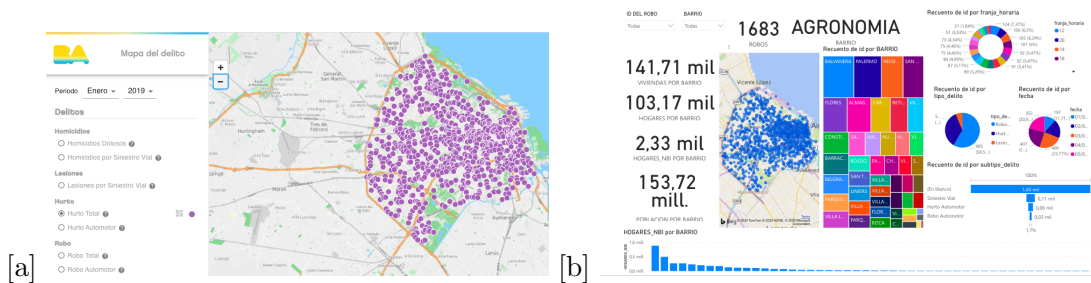


Figura 4: Mapas interactivos

[a.] Plataforma digital control de los robos [9] [b.] POWER BI,elaboracion propia[10]

4.2.3. Mapa por barrios

Los barrios de Buenos Aires son las unidades territoriales en las que está dividida, legalmente y en tercer orden, la mencionada ciudad. Oficialmente, la misma se encuentra dividida en 48 barrios. Los más tradicionales derivan de las antiguas parroquias establecidas en el siglo XIX. Cada barrio tiene su propia historia y características poblacionales que le imprimen color, estilo y costumbres únicas; y son un reflejo de la variedad cultural que subyace en la ciudad.

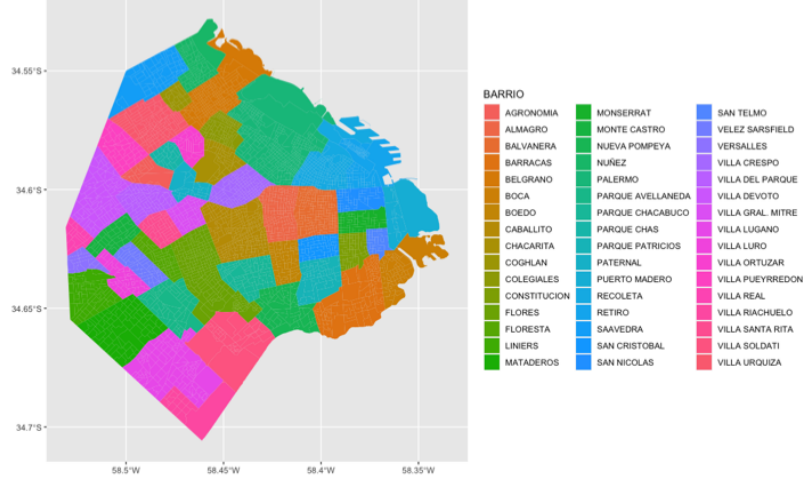


Figura 5: División por barrios Buenos Aires

En la figura 5a se evidencian los barrios con mayor número de robos registrados y en la figura 5b los barrios con mayor número de hogares NBI.

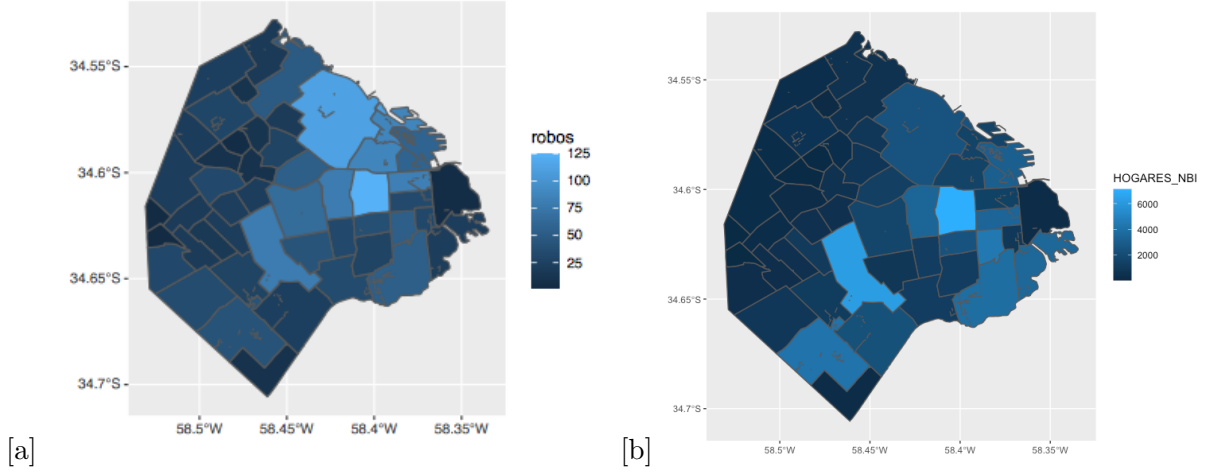


Figura 6: Mapas por barrio

4.3. Modelo sin tener en cuenta autocorrelación espacial

Más allá del análisis exploratorio y descriptivo que se puede realizar con los datos espaciales, estos también son de gran utilidad para realizar inferencia sobre la relación entre distintos fenómenos. La inferencia en base a datos espaciales parte de la base de reconocer que las observaciones en el espacio no pueden ser asumidas todo el tiempo como mutuamente independientes, debido a que las observaciones que están cerca la una de la otra son la mayoría del tiempo similares, por lo que se debe prestar atención a los diferentes patrones de asociación espacial que existen en los fenómenos que estudiamos.

Estos patrones espaciales (que llamamos autocorrelación espacial), miden cómo la distancia influye una variable en particular, pudiendo ser utilizados como información relevante sobre

tipos de influencia que no han sido observadas o consideradas. (Bivand, Pebesma Gomez-Rubio, 2008: 11).

4.3.1. Modelo Poisson

Se trata de un modelo discreto, pero en el que el conjunto de valores con probabilidad no nula no es finito, sino numerable. Se dice que una variable aleatoria X sigue la distribución de Poisson si su función de densidad viene dada por:

$$f(k) = P[X = k] = \begin{cases} e^{-\lambda} \frac{\lambda^k}{k!} & \text{si } k = 0, 1, 2, \dots \\ 0 & \text{en caso contrario} \end{cases}$$

Esta distribución suele utilizarse para contajes del tipo número de individuos por unidad de tiempo, de espacio, etc.

Se realizan 2 modelos Poisson, en el primero la variable respuesta es el número de robos teniendo en cuenta los hogares con necesidades básicas insatisfechas y en el segundo la variable respuesta son hogares con necesidades básicas insatisfechas con variables independientes el número de viviendas y hogares por barrio.

```
Call:
glm(formula = robos ~ HOGARES_NBI, family = poisson, data = barrios_geo)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-5.6190  -2.6673  -0.6383   0.8714   9.2142

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  3.1097090  0.0355216   87.54  <2e-16 ***
HOGARES_NBI  0.0002438  0.0000111    21.96  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 927.53  on 47  degrees of freedom
Residual deviance: 523.27  on 46  degrees of freedom
AIC: 770.92

Number of Fisher Scoring iterations: 5
```

En el primer modelo se observa que el intercepto y el número de Hogares NBI son significativos.


```

Call:
glm(formula = HOGARES_NBI ~ VIVIENDAS + HOGARES, family = poisson,
    data = barrios_geo)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-46.730  -28.376  -14.989    4.357   75.149

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  6.510e+00  6.805e-03  956.55  <2e-16 ***
VIVIENDAS    -6.261e-05  7.039e-07  -88.94  <2e-16 ***
HOGARES       1.027e-04  9.388e-07  109.37  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 72691  on 47  degrees of freedom
Residual deviance: 47567  on 45  degrees of freedom
AIC: 47978

Number of Fisher Scoring iterations: 5

```

En el segundo modelo se observa que el intercepto, el número de viviendas y el número de hogares son significativos. Se debe observar si los residuos poseen autocorrelación espacial a partir del índice I de Moran. Para ello, primero se debe crear la matriz de ponderaciones espaciales W .

4.4. Matriz de Pesos Espaciales

Para realizar este tipo de análisis se debe determinar el set de vecindarios para cada observación, es decir, identificar los polígonos que comparten fronteras entre ellos. Luego, asignar pesos a cada relación vecina, que permiten definir la fuerza de esta relación en base a cercanía. En las matrices de peso, los vecinos son definidos por un método binario (0,1) en cada fila indicando si existe o no relación.

Existen tres criterios diferentes para calcular los vecindarios:

4.4.1. Vecinos físicos

- Rook: Considera como vecinos a cualquier par de celdas que compartan alguna arista (borde).

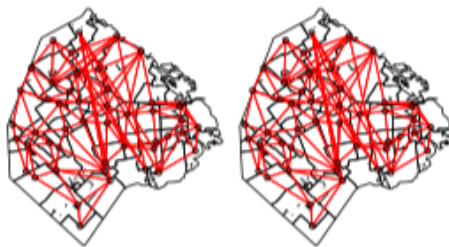


Figura 7: rook B y W

- Queen: Considera como vecinos a cualquier par de celdas que compartan alguna arista o un punto.

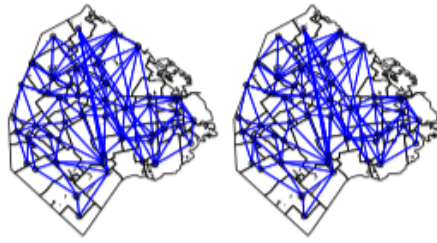


Figura 8: queen B y W

4.4.2. Gráficas vecinas

- NNG: El gráfico de vecina más cercana (nearest neighbor graph -NNG) es un gráfico dirigido con P como su conjunto de vértices y con un borde dirigido de p a q siempre que q sea el vecino más cercano de p .

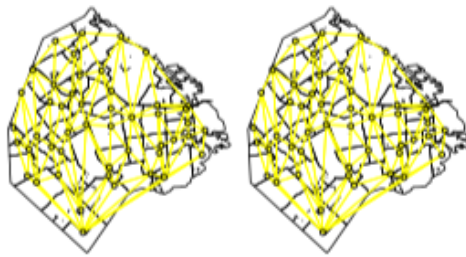


Figura 9: NNG B y W

- SOI: El gráfico de esfera de influencia (sphere-of-influence graph-SOI) utiliza dos círculos donde el diámetro está determinado por los puntos circundantes. En este gráfico, la restricción es que los círculos deben cruzarse.

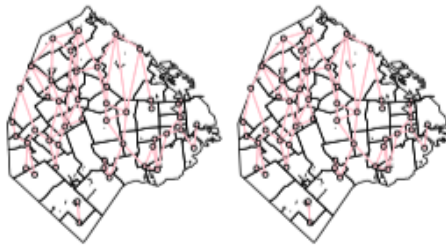


Figura 10: SOI B y W

- Gabriel: El gráfico de Gabriel funciona con un círculo de diámetro mínimo dibujado a través de solo dos puntos

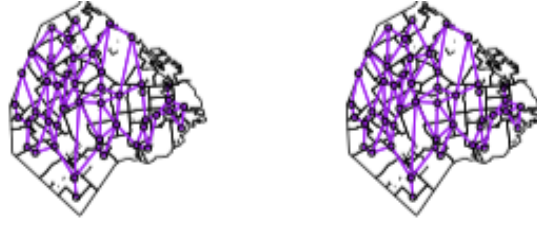


Figura 11: Gabriel B y W

- Relativa: El gráfico de vecino relativo es similar, aunque los diámetros de los círculos están determinados por la distancia entre los dos puntos.



Figura 12: Relativa B y W

4.4.3. Vecinos a distancia o K-nearest

Los vecindarios se generan en base la distancia entre vecinos, donde “k” se refiere al número de vecinos de una determinada locación, calculada como la distancia entre los puntos centrales de los polígonos. Se aplica usualmente cuando las áreas tienen diferentes tamaños para asegurarse que cada lugar tenga el mismo número de vecinos, independiente de cuán grande sean las áreas vecinas. En la figura 12 se muestran k vecinos con k desde 1 hasta 4.

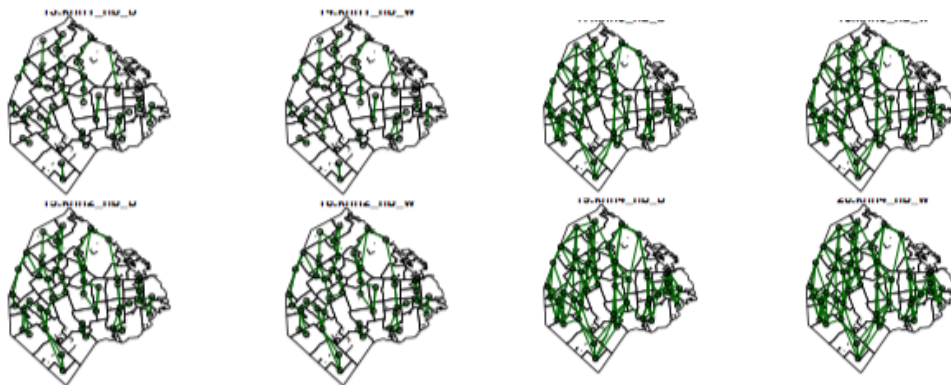


Figura 13: K-nearest

4.5. Moran's I

Para la medición de dependencia espacial se han propuesto numerosos estadísticos, uno de los más utilizados es el índice de Moran (1948), que se define en la fórmula siguiente:

$$I = \frac{n}{\sum_{i=1}^n (y_i - \bar{y})^2} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}}$$

Esta fórmula es una versión expandida de la fórmula para computar el coeficiente de correlación, al cual se le agrega una matriz de pesos espaciales.

Esta estadística es utilizada para realizar un test que permite evaluar la significancia de la correlación espacial. Primero, realizaremos un test a nivel global que crea una sola medida de correlación espacial. Este test de Moran creará una medida de correlación entre -1 y 1 en donde:

- 1 determina una correlación espacial positiva perfecta (que indica que nuestros datos están agrupados en clusters)
- 0 significa que nuestros datos están distribuidos aleatoriamente
- -1 representa autocorrelación espacial negativa (valores disímiles están cercanos entre sí)

Utilizando la simulación de Monte Carlo en donde los valores son asignados aleatoriamente a los polígonos para computar el Moran's I, lo cual es repetido varias veces para establecer una distribución de los valores esperados. Luego de esto, el valor observado del Moran's I es comparado con la distribución simulada para ver cuán probable es que los valores observados puedan ser considerados como aleatorios y así determinar si existe autocorrelación espacial significativa, se selecciona la Matriz W que maximiza el coeficiente de Moran.

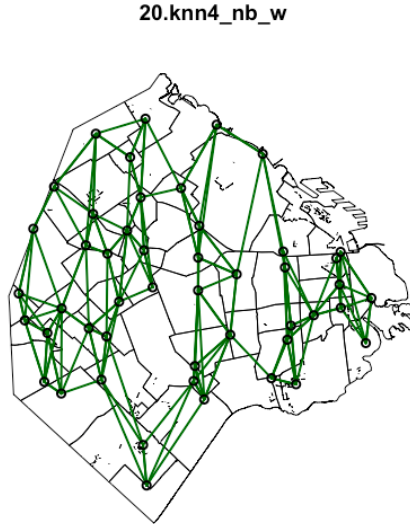


Figura 14: matriz W que maximiza el índice de Moran

La matriz W que maximiza es entonces la matriz con k=4.

Se aplica el test de Moran en los residuales de ambos modelos y se evidencia que se cumple el supuesto de independencia. Como el p-valor es menor a 0,05, se verifica la hipótesis alternativa y, por tanto, los residuos poseen autocorrelación espacial.

```

Moran I test under randomisation

data: barrios_geo$residLRA
weights: mat[[which.max(aux)]]

Moran I statistic standard deviate = 2.8363, p-value = 0.004565
alternative hypothesis: two.sided
sample estimates:
Moran I statistic      Expectation      Variance
      0.229536337      -0.021276596      0.007820066

Moran I test under randomisation

data: barrios_geo$residLRA
weights: mat[[which.max(maux)]]

Moran I statistic standard deviate = 6.5231, p-value = 6.885e-11
alternative hypothesis: two.sided
sample estimates:
Moran I statistic      Expectation      Variance
      0.559213548      -0.021276596      0.007919129

```

Se realiza un scatter plot para visualizar el tipo y fuerza de la correlación espacial en la distribución de nuestros datos.

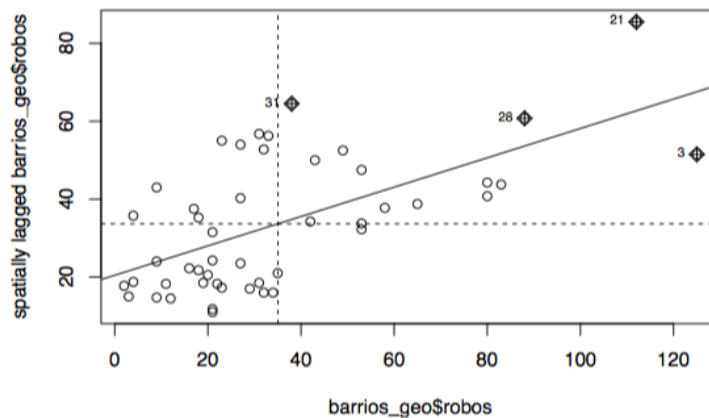


Figura 15: Moran plot

En este mapa la curva del gráfico indica el valor del Moran's I, es decir, la medida global de autocorrelación espacial en los datos. Si bien el test de Moran permite identificar si es que existe clustering a nivel global, no permite identificar si existen clusters locales significantes en la variable que estamos analizando (Lansley Cheschire, 2016, 78). Es por esto que realiza un test de Moran's I a nivel local en donde se calculan los indicadores locales de asociación espacial para cada unidad de nuestros datos y se testea si es que esta relación es estadísticamente significativa, lo que nos entrega datos sobre los patrones geográficos de la relación de los datos espaciales y si existen desviaciones locales de los patrones globales de autocorrelación espacial.

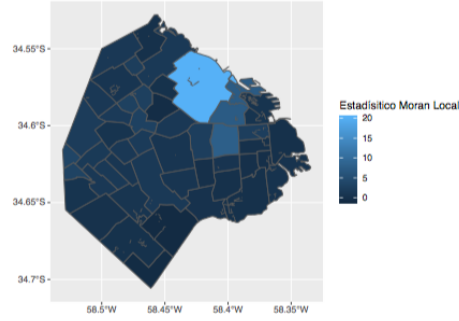


Figura 16: Moran local

Este mapa permite observar la variación en la autocorrelación a lo largo del espacio, pero no es posible identificar si los patrones geográficos de autocorrelación son clusters con valores altos o bajos, lo que permitirá analizar el tipo de autocorrelación espacial que existe y su nivel de significancia. Para esto se crea un mapa de clusters LISA, que creará una etiqueta basada en los tipos de relación que se comparte con sus vecinos (alta-alta, baja-alta, insignificante, etc) en relación a los valores de la variable que estamos analizando (Gini).

4.6. Clusters

Este mapa nos entrega una mayor información sobre los patrones geográficos de autocorrelación espacial. Este mapa nos muestra si existen clusters, es decir, regiones en donde en su núcleo existe autocorrelación espacial positiva, por ende, nos muestra regiones clusterizadas más que lugares individuales.

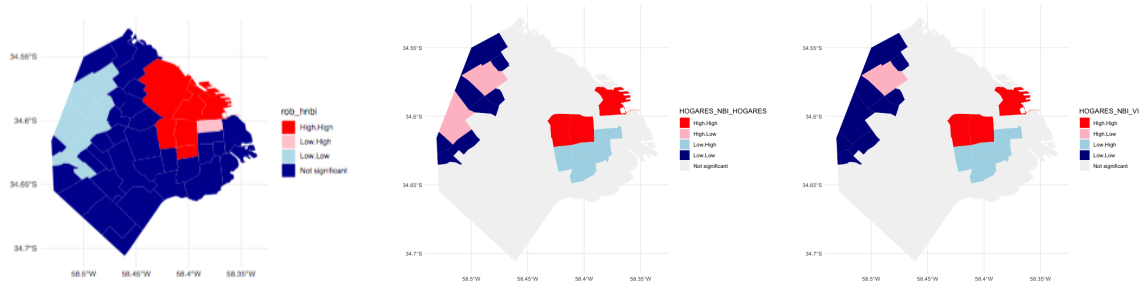


Figura 17: Cluster robos-hogares nbi , hogares nbi-hogares y hogares nbi-viviendas

Se evidencia que en los barrios de Palermo, Recoleta, Retiro, Balvanera, San Cristóbal, y Almagro son donde más se presentan robos, tanto a nivel de barrio como en sus vecinos, San Nicolás presenta pocos robos pero está rodeado de barrios con alta cantidad de robos, barrios como Villa Devoto y Villa del Parque presentan pocos robos y sus vecinos también.

En cuanto a hogares NBI es notorio que Balvanera y Retiro son los que más presentan, barrios como Villa Devoto y Villa del Parque presentan también pocos hogares NBI.

5. Modelos

5.1. Vectores de Moran ajustados

El enfoque del vector propio de Moran [4] involucró los patrones espaciales representados por mapas de vectores propios; Al elegir patrones ortogonales adecuados y agregarlos a un modelo lineal lineal o generalizado, la dependencia espacial presente en los residuos se puede trasladar al modelo.

Utiliza la fuerza bruta para buscar el conjunto de vectores propios de la matriz MWM, donde

$$\mathbf{M} = \mathbf{I} - \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$$

es una matriz de proyección simétrica e idempotente y W son los pesos espaciales. En la forma de retraso espacial de Filtrado espacial y en la forma GLM ME a continuación, X es un vector n de unos, que es solo la intersección.

```
Call:
glm(formula = robos ~ HOGARES_NBI + fitted(MEpoisLR), family = poisson,
     data = barrios_geo)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-5.1917 -1.9381 -0.0201  1.0298  6.2676

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)    3.084e+00  3.755e-02  82.140 < 2e-16 ***
HOGARES_NBI    2.170e-04  1.334e-05  16.266 < 2e-16 ***
fitted(MEpoisLR)vec11  6.829e-01  1.785e-01   3.826  0.00013 ***
fitted(MEpoisLR)vec15 -1.268e+00  1.490e-01  -8.508 < 2e-16 ***
fitted(MEpoisLR)vec5   -3.686e-01  1.829e-01  -2.015  0.04386 *
fitted(MEpoisLR)vec2   -9.707e-01  1.620e-01  -5.993  2.06e-09 ***
fitted(MEpoisLR)vec10 -1.502e+00  1.786e-01  -8.409 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 927.53  on 47  degrees of freedom
Residual deviance: 286.89  on 41  degrees of freedom
AIC: 544.55

Number of Fisher Scoring iterations: 4
```

```
Call:
glm(formula = HOGARES_NBI ~ HOGARES + VIVIENDAS + fitted(MEpoisLMR),
     family = poisson, data = barrios_geo)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-39.281 -14.512  -1.430   9.968  39.402

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)    6.176e+00  8.235e-03  750.0 <2e-16 ***
HOGARES        1.413e-04  9.546e-07  148.0 <2e-16 ***
VIVIENDAS      -9.157e-05  7.143e-07  -128.2 <2e-16 ***
fitted(MEpoisLMR) -4.982e+00  2.911e-02 -171.1 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 72691  on 47  degrees of freedom
Residual deviance: 17845  on 44  degrees of freedom
AIC: 18257

Number of Fisher Scoring iterations: 5
```

En el primer modelo el intercepto, el número de hogares NBI y los vectores 2,5,10,11 y 15 son significativos mientras que en el modelo 2 el intercepto, los hogares las viviendas, el vector 1 son significativos, en este modelo la dependencia de los residuos desaparece.

5.2. Modelo de Retardo Espacial

Estos modelos presentan la correlación espacial (dependencia) en la variable dependiente (Y).

Permiten a las observaciones de la variable dependiente Y en el área i ($i = 1, \dots, n$) depender de observaciones en áreas vecinas. El modelo de retardo espacial básico, llamado modelo autorregresivo espacial de primer orden (SAR), toma la forma muestral:

$$y_i = \rho \sum_{j=1}^n W_{ij} y_j + \sum_{q=1}^Q x_{iq} \beta_q + \varepsilon_i \quad i = 1, \dots, n$$

donde el término error ε_i es independiente e idénticamente distribuido. W_{ij} es el elemento (i, j) th de la matriz de pesos $n \times n$. El escalar ρ es un parámetro a estimar que determinará el

nivel de relación autorregresiva espacial entre y_i y $\sum_j W_{ij}y_j$ (combinación lineal de observaciones espacialmente relacionadas basadas en elementos distintos de cero en la i -ésima fila de \mathbf{W}).

Residuals:	Residuals:								
Min	1Q	Median	3Q	Max	Min	1Q	Median	3Q	Max
-33.0544	-10.7163	-3.4010	9.3159	43.3943	-1720.89	-634.35	-236.99	185.50	4151.59
Type: lag	Type: lag								
Coefficients: (asymptotic standard errors)	Coefficients: (asymptotic standard errors)								
Estimate Std. Error z value Pr(> z)	Estimate Std. Error z value Pr(> z)								
(Intercept) 5.4286130 4.4592991 1.2174 0.2235	(Intercept) -188.484558 283.671327 -0.6644 0.5064043								
HOGARES_NBI 0.0098620 0.0016768 5.8815 4.064e-09	VIVIENDAS -0.133783 0.038200 -3.5022 0.0004614								
	HOGARES 0.207660 0.051424 4.0382 5.387e-05								
Rho: 0.46042, LR test value: 9.316, p-value: 0.0022716	Rho: 0.43984, LR test value: 8.5581, p-value: 0.00344								
Asymptotic standard error: 0.11035	Asymptotic standard error: 0.12581								
z-value: 4.1723, p-value: 3.0149e-05	z-value: 3.4961, p-value: 0.00047219								
Wald statistic: 17.408, p-value: 3.0149e-05	Wald statistic: 12.222, p-value: 0.00047219								
Log likelihood: -206.146 for lag model	Log likelihood: -408.5809 for lag model								
ML residual variance (sigma squared): 300.08, (sigma: 17.323)	ML residual variance (sigma squared): 1388200, (sigma: 1178.2)								
Number of observations: 48	Number of observations: 48								
Number of parameters estimated: 4	Number of parameters estimated: 5								
AIC: 420.29, (AIC for lm: 427.61)	AIC: 827.16, (AIC for lm: 833.72)								
LM test for residual autocorrelation	LM test for residual autocorrelation								
test value: 2.688, p-value: 0.10111	test value: 6.4304, p-value: 0.011218								

Se observan los parámetros estimados necesarios teóricamente para el modelo. Estos son: el parámetro ρ que determina el nivel de relación autorregresiva espacial es estimado por 0.46 y 0.43, los coeficientes β_q en el modelo 1 dados por $\beta_0 = 5.42$, $\beta_1 = 0.009$ y modelo 2 dados por $\beta_0 = -188.48$, $\beta_1 = -0.13$ y $\beta_2 = 0.20$ y por último la desviación típica de los residuos generados por este método, es decir, $\sigma = 17.32$ y $\sigma = 1178.2$. Dado que los p-valor de los test son 0.10 y 0.81 no se rechaza la hipótesis nula de no autocorrelación espacial, es decir, los residuos se pueden considerar independientes y por tanto, tras la aplicación del modelo SAR, la dependencia de los residuos desaparece.

5.3. Modelo de Durbin

En algunos estudios puede sospecharse que la dependencia influye sobre la variable objetivo o dependiente a través de la propia variable objetivo y de las variables predictoras o explicativas. Se puede construir un modelo tipo SAR, aumentado por variables explicativas espacialmente retardadas:

$$Y = \rho \mathbf{W}Y + \mathbf{X}\beta + \mathbf{W}\mathbf{X}\gamma + c$$

donde X es la matriz $n \times (Q - 1)$ no constante de variables explicativas. El modelo puede reducirse como:

$$4Y = (I - \rho \mathbf{W})^1(\mathbf{X}\beta + \mathbf{W}\mathbf{X}\gamma + \varepsilon)$$

con

$$e \sim \mathcal{N}(0, \sigma^2 I)$$

donde γ es un vector $(Q - 1) \times 1$ de parámetros que miden el impacto marginal de las variables explicativas de observaciones vecinas (fuerzas) sobre la variable dependiente Y . $\mathbf{W}\mathbf{X}$ produce variables explicativas retardadas espacialmente que reflejan un promedio de observaciones vecinas.

<pre> Residuals: Min 1Q Median 3Q Max -32.6598 -10.2819 -3.5759 8.7465 44.7626 Type: mixed Coefficients: (asymptotic standard errors) Estimate Std. Error z value Pr(> z) (Intercept) 5.4009661 4.4716505 1.2078 0.2271 HOGARES_NBI 0.0099946 0.0016921 5.9067 3.491e-09 lag.HOGARES_NBI -0.0011787 0.0031886 -0.3697 0.7116 Rho: 0.50504, LR test value: 6.009, p-value: 0.014233 Asymptotic standard error: 0.14295 z-value: 3.533, p-value: 0.00041085 Wald statistic: 12.482, p-value: 0.00041085 Log likelihood: -206.0809 for mixed model ML residual variance (sigma squared): 295.9, (sigma: 17.202) Number of observations: 48 Number of parameters estimated: 5 AIC: 422.16, (AIC for lm: 426.17) LM test for residual autocorrelation test value: 11.547, p-value: 0.00067848 </pre>	<pre> Residuals: Min 1Q Median 3Q Max -2233.32 -491.50 -111.72 327.31 3453.43 Type: mixed Coefficients: (asymptotic standard errors) Estimate Std. Error z value Pr(> z) (Intercept) 149.821847 312.642467 0.4792 0.6317882 VIVIENDAS -0.202811 0.036013 -5.6316 1.785e-08 HOGARES 0.296320 0.046863 6.3230 2.565e-10 lag.VIVIENDAS 0.238682 0.070681 3.3769 0.0007331 lag.HOGARES -0.320516 0.087016 -3.6834 0.0002301 Rho: 0.58256, LR test value: 14.382, p-value: 0.00014919 Asymptotic standard error: 0.12553 z-value: 4.6407, p-value: 3.4721e-06 Wald statistic: 21.536, p-value: 3.4721e-06 Log likelihood: -401.8435 for mixed model ML residual variance (sigma squared): 1007200, (sigma: 1003.6) Number of observations: 48 Number of parameters estimated: 7 AIC: 817.69, (AIC for lm: 830.07) LM test for residual autocorrelation test value: 1.0172, p-value: 0.31319 </pre>
---	--

En los hogares NBI se evidencia un p valor de 0.31 es decir no hay correlación espacial.

5.4. Comparaciones

	df	AIC
Poisson	2.00	615.98
Vectores de Moran	7.00	544.55
Retardo espacial	4.00	420.29
Espacial de Durbin	5.00	422.16

Cuadro 1: Comparación de modelos de Robos

	df	AIC
Poisson	3.00	47977.62
Vectores de Moran	4.00	18257.23
Retardo espacial	5.00	827.16
Espacial de Durbin	7.00	817.69

Cuadro 2: Comparación de modelos de Hogares NBI

A partir de los cuadros 1, 2 se puede comprobar que, siguiendo el criterio AIC, ambos modelos de regresión espacial se ajustan mejor a los datos que el modelo de Poisson.

El mejor modelo para robos es el **Modelo de Retardo Espacial** , y para Hogares NBI el mejor modelo es el **Modelo espacial de Durbin**

6. Conclusiones

En el número de robos y hogares NBI existe influencia en términos de proximidad de los barrios(correlación espacial en robos y hogares NBI).Es evidente que la existencia de hogares NBI influye

en la cantidad de los robos y que a pesar de que 2 barrios (Palermo y Recoleta) tienen mayores recursos económicos al tener proximidad con barrios con menos recursos económicos se ven afectados, adicionalmente se evidencia que el número de hogares superior al número de viviendas genera mayores hogares NBI.

Se proponen estrategias puntuales para disminuir la problemática de robos, la primera es generar campañas y ayudas gubernamentales (subsidios, empleos, etc) para que por vivienda solo exista un hogar es decir generar independencia de hogares, dicha campaña ayudaría a disminuir los hogares NBI. Sería interesante tener en cuenta la educación y hacer políticas de pedagógicas principalmente en barrios como Balvanera y Retiro.

Referencias

- [1] Jennifer Barsky Reese, Stephen J. Lepore, Elizabeth A. Handorf, Jennifer A. Haythornthwaite, Emotional approach coping and depressive symptoms in colorectal cancer patients: The role of the intimate relationship, *Journal of Psychosocial Oncology*, 10.1080/07347332.2017.1331492, 35, 5, (578-596), (2017).
- [2] <https://data.buenosaires.gob.ar/dataset/delitos/archivo/2928d01c-c186-4250-8998-4610abfef95b>
- [3] https://bitsandbricks.github.io/data/CABA_r.c.geojson
- [4] Pedro R. Peres-Neto and Pierre Legendre *Global Ecology and Biogeography* Vol. 19, No. 2 (March 2010), pp. 174-184
- [5] Sarmiento-Barbieri, I., University of Illinois. (2016). An Introduction to Spatial Econometrics in R. Recuperado de http://www.econ.uiuc.edu/lab/workshop/Spatial_in_R.html
- [6] Cliff, Andrew David, and J Keith Ord. 1973. *Spatial Autocorrelation*. Vol. 5. Pion London.
- [7] Anselin, Luc. 2003. "An Introduction to Spatial Regression Analysis in R." Available at: <https://geodacenter.asu.edu/drupal/files/spdepintro.pdf>.
- [8] INDEC. Proyecciones elaboradas en base al Censo Nacional de Población, Hogares y Viviendas 2010. «Población estimada al 1 de julio según año calendario por sexo. Provincia de Buenos Aires. Años 2010-2040».
- [9] <https://mapa.seguridadciudad.gob.ar/>
- [10]