

CATEDRA DE ECONOMETRÍA

# **UNA INTRODUCCIÓN A LA ECONOMETRÍA ESPACIAL**

---

**DEPENDENCIA Y HETEROGENEIDAD**

**ALFREDO BARONIO – ANA VIANCO – CRISTIAN RABANAL**

**AÑO 2012**

## Contenido

UNA INTRODUCCIÓN A LA ECONOMETRÍA ESPACIAL.....	2
1. ¿Qué es la econometría espacial?.....	2
2. Los efectos espaciales .....	3
2.1 La heterogeneidad espacial.....	3
2.2 La autocorrelación espacial.....	4
2.3 Fuentes de autocorrelación espacial.....	5
3. La matriz de pesos espaciales .....	6
3.1 Criterios habituales de contigüidad .....	7
3.2 Otras especificaciones para la matriz de contactos .....	8
4. El análisis exploratorio de datos espaciales (AEDE) .....	10
4.1 Técnicas gráficas.....	10
4.2 Medidas de asociación espacial global .....	12
4.3 Medidas de asociación espacial local.....	13
5. El análisis confirmatorio .....	15
5.1 La dependencia espacial sustantiva .....	16
5.2 La dependencia espacial residual.....	17
6. Tratamiento de la heterogeneidad espacial .....	18
6.1 Contrastación de heterogeneidad espacial.....	18
6.2 Especificación de un modelo con heterogeneidad espacial .....	19
CASOS DE ESTUDIO, PREGUNTAS Y PROBLEMAS.....	20
Ejercicio con Geoda: Radios censales de Río Cuarto.....	20
Introducción a OpenGeoDa.....	21
Crear un entorno de trabajo .....	24
Georreferencia de las variables.....	26
Indicadores del espacio.....	27
Herramienta regresión .....	30
REFERENCIAS BIBLIOGRÁFICAS .....	31

# UNA INTRODUCCIÓN A LA ECONOMETRÍA ESPACIAL

## 1. ¿Qué es la econometría espacial?

La econometría espacial es la parte de la econometría que se dedica al estudio de los *fenómenos económicos espaciales*. Aunque si bien todas las actividades económicas se desarrollan en un espacio determinado, los fenómenos económicos espaciales pueden definirse como aquellos en los que la variable espacio, entendida en sentido amplio y definida luego según algún criterio para alguna matriz de contactos, juega un rol tan importante que su exclusión podría dar lugar a modelos econométricos con severos errores de especificación. La importancia del espacio, se manifiesta a menudo en la autocorrelación espacial y/o la heterogeneidad espacial. La creciente importancia de la temática se vio reflejada en la creación del apartado C21 (reservado para Métodos Econométricos, Modelos de Sección Cruzada y Modelos Espaciales) por parte del *Journal Economic Literature* (Paelink *et al*, 2004).

Existen cinco *principios fundamentales* que rigen al análisis espacial (Paelinck y Klaassen, 1979):

- *Interdependencia*: la dependencia recíproca entre las diferentes unidades de análisis deben ser debidamente incorporadas.
- *Asimetría*: refleja la idea de concentración y desconcentración gradual en diferentes áreas.
- *Alotopía*: la causa de un fenómeno espacial debe buscarse en otro lugar.
- *No linealidad*.
- *Topología*: Inclusión de variables de distancias entre dos localizaciones, coordenadas, densidades y otras.

De acuerdo con Corrado y Fingleton (2011), aunque muchos economistas se han resistido a la econometría espacial, por considerar que las variables espaciales se adicionan al análisis sólo por mostrar significatividad estadística sin ninguna justificación teórica, los estudios espaciales dentro de la corriente principal de la

economía han ido incorporando variables del tipo espacial. En particular, la econometría espacial ha proporcionado herramientas valiosas para el estudio de las externalidades en forma de spillovers (efectos indirectos) espaciales.

Existe una gran cantidad de ramas de la economía que han incorporado al análisis a la econometría espacial: la economía urbana, la economía regional y la macroeconomía entre otras. En este sentido, la teoría de los centros de desarrollo constituye uno de los principales fundamentos para el análisis espacial de los fenómenos económicos. En este sentido, los aportes teóricos más significativos han provenido desde la Teoría del Desarrollo Económico y de la denominada Teoría de la Organización Espacial (Posada, 1978).

Habitualmente, en el estudio econométrico de fenómenos económicos espaciales suelen surgir, dos problemas principales, apuntados anteriormente, o al menos uno, y que son inherentes a la propia naturaleza de la estructura de datos a analizar: la heterogeneidad espacial y la autocorrelación espacial –también llamada dependencia espacial por algunos autores (Moreno y Vayá, 2000).

## **2. Los efectos espaciales**

### **2.1 La heterogeneidad espacial**

La heterogeneidad espacial, refiere a la idea de variación en las relaciones que se establecen para los fenómenos económicos espaciales conforme varía el espacio de estudio. Este problema se manifiesta básicamente de dos formas: parámetros que asumen diferentes valores según se incluyan determinados zonas o no, esto es inestabilidad estructural, o errores provocados por especificaciones de modelos incorrectos, que pueden dar lugar a heterocedasticidad. La heterocedasticidad espacial es habitual en los estudios económicos que tienden a analizar un fenómeno determinado con el enfoque centro-periferia, ya que cada región deriva en diferentes valores de los parámetros.

En ambos casos si estos problemas no se resuelven, se verán reflejados en el término de perturbación. En la sección 6 se aborda el tratamiento de modelos con problemas de heterogeneidad espacial.

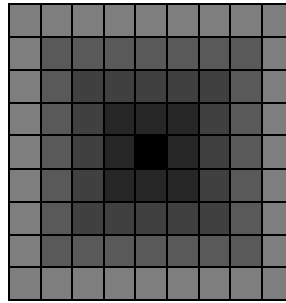
## 2.2 La autocorrelación espacial

La autocorrelación espacial implica que el valor de una variable se encuentra condicionado por el valor que esa variable asume en una región vecina. Como se verá más adelante, la vecindad no necesariamente quedará definida como contigüidad física, sino que existen una gran cantidad de criterios para definirla, a partir de una matriz de contactos. Esta debilidad, ha sido una de las principales objeciones metodológicas a la econometría espacial y la robustez de los resultados que con ésta se pueden alcanzar.

*La autocorrelación espacial podrá ser negativa o positiva.*

Será *positiva* cuando la presencia de un fenómeno económico en una determinada unidad, se extienda a las regiones colindantes. Por ejemplo, piénsese en el valor de las propiedades en una ciudad. En general, ceteris paribus las características propias de cada inmueble, el precio de los mismos depende de factores tales como la ubicación –habitualmente medida en función a la distancia de algún centro importante, como puede ser una plaza central-, la disponibilidad de servicios, la concentración comercial del área en la que se sitúa, etcétera. De esta forma, un inmueble tenderá a tener un valor más o menos similar al inmueble colindante (siempre que no tenga características propias que lo diferencien), pero levemente decreciente en la medida que se encuentren más lejos de un punto central (que podrá ser una plaza importante –plaza central-, o algún otro punto relevante de similar importancia). La *figura 1* refleja esta idea de manera simplificada, ya que lógicamente, en una gran ciudad podría haber una gran cantidad de “puntos centrales”. De esa forma podría pensarse en una sucesión de cuadrados, como el de la figura 1, uno a continuación de otro. Otro ejemplo que podría adaptarse al diagrama está dado por valor de las tierras, de donde precisamente David Ricardo construyó el concepto de renta.

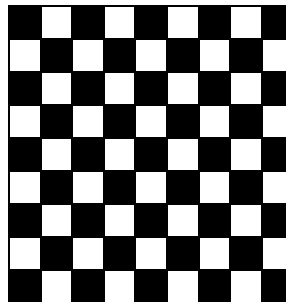
**FIGURA 1: Autocorrelación espacial positiva: valor de las propiedades en una ciudad según la distancia a un punto central**



FUENTE: Elaboración propia

Por el contrario, *será negativa*, cuando la presencia de ese fenómeno haga imposible o disminuya significativamente las posibilidades de que ese mismo fenómeno pueda ocurrir en las regiones colindantes. En tal caso, la representación se aproximará a un tablero de ajedrez, como en la *figura 2*.

**FIGURA 2: Autocorrelación espacial negativa**



FUENTE: Elebaración propia

### 2.3 Fuentes de autocorrelación espacial

Las principales fuentes de autocorrelación espacial pueden ser los errores de medida y la propia interacción espacial de las unidades. En términos económicos, los efectos desbordamiento (spillovers) pueden generar la autocorrelación espacial. Esto se ha visto potenciado con los procesos de integración económica.

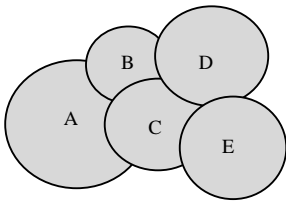
### 3. La matriz de pesos espaciales

En el análisis de series temporales es usual utilizar un operador de rezago, para capturar la influencia de las observaciones pasadas en la dinámica de la serie y el valor de la realización contemporánea. Otra razón por la que resulta relevante tiene que ver con la posibilidad de realizar pronósticos. De esta manera, la influencia del rezago temporal es unidireccional: las realizaciones pasadas afectan a las presentes, y harán lo propio con las futuras, conforme a una estructura dinámica. Ahora bien, en el marco del análisis espacial se establecen relaciones multidireccionales, por lo que resulta necesario construir una matriz que permita incluirlas de manera adecuada al análisis.

La *matriz de pesos espaciales* (también denominada matriz de contactos o matriz de proximidad espacial) y simbolizada con  $W$ , es una matriz cuadrada de  $N \times N$  (siendo  $N$  el número de unidades espaciales), no estocástica cuyos elementos ( $w_{ij}$ ) reflejan la intensidad de la interdependencia entre cada par de regiones  $i, j$  (Moreno y Vayá, 2000).

El valor que asume cada  $w_{ij}$  se basa en las siguientes determinaciones de adyacencias. De manera simplificada  $w_{ij} = 1$ , si dos regiones son contiguas y  $w_{ij} = 0$  en caso contrario. Típicamente los elementos de la diagonal principal son cero, pues ninguna región puede ser vecina de sí misma. La *figura 3* ilustra un ejemplo hipotético:

**FIGURA 3: Ilustración de una matriz de pesos espaciales**

		<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>
	<b>A</b>	0	1	1	0	0
	<b>B</b>	1	0	1	1	0
	<b>C</b>	1	1	0	1	1
	<b>D</b>	0	1	1	0	
	<b>E</b>	0	0	1	1	0

FUENTE: Elaboración propia

En algunos casos suele normalizarse la matriz  $W$  por filas. Esto es, se divide cada elemento  $w_{ij}$  por la suma de fila a la que pertenece. De esta forma, la suma de los pesos asociados a cada área es igual a 1. Por otra parte, de acuerdo con Kapoor et. al. (2007) la matriz  $W$  debe ser delimitada de manera uniforme y en valor absoluto, lo que implica la existencia de una constante  $c < \infty$  tal que

$$\overbrace{1 \leq i \leq N \sum_{j=1}^N |w_{ij}| \leq c}^{\text{máximo}} \text{ y } \overbrace{1 \leq j \leq N \sum_{i=1}^N |w_{ij}| \leq c}^{\text{máximo}} \quad (1)$$

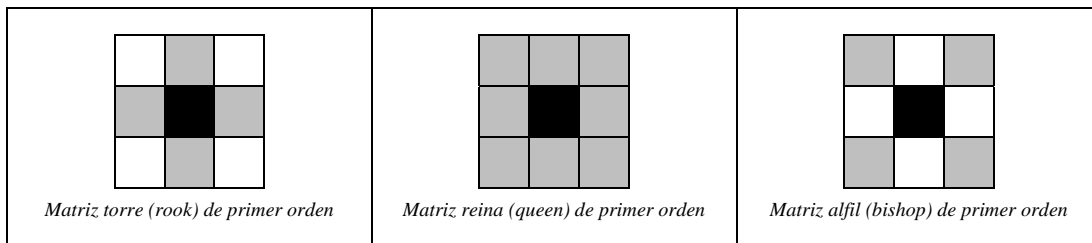
para producir los resultados asintóticos exigidos por una estimación consistente.

La matriz de pesos espaciales desempeña un papel fundamental en la incorporación de las relaciones espaciales de las variables al modelo. En el contexto de una única ecuación, por ejemplo, es posible generar una matriz de rezago espacial de la variable endógena ( $WY$ ), a partir de multiplicar la matriz  $W$  por un vector de variables endógenas  $Y$ , de orden  $N \times 1$ . Otra posibilidad consiste en incorporar otras variables espacialmente rezagadas, desde una matriz de variables  $X$  de orden  $N \times k$ , reflejado en las columnas de la matriz  $WX$ .

### 3.1 Criterios habituales de contigüidad

Usualmente los criterios de contigüidad (o de vecindad) más utilizados son los que se resumen en la *figura 4*, a continuación:

**FIGURA 4: Criterios de contigüidad en la matriz de pesos espaciales. Retardo espacial de primer orden**

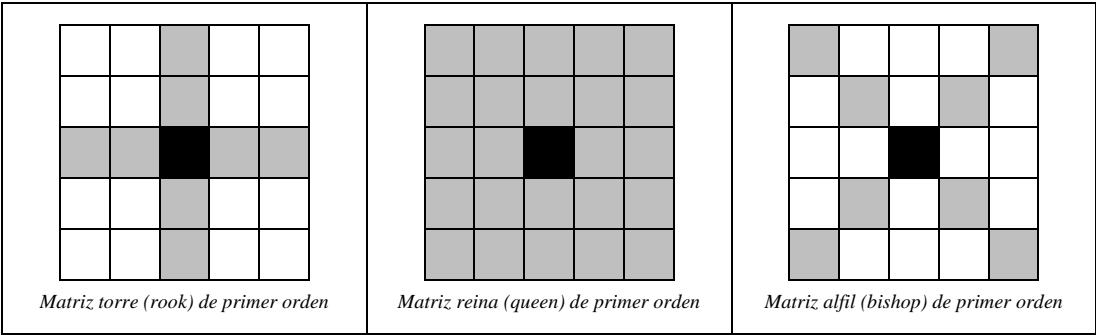


FUENTE: Elaboración propia



Donde los casilleros grises son vecinos del casillero negro, en tanto que los casilleros blancos no son considerados contiguos respecto a los negros. Los criterios presentados ilustran situaciones de contigüidad de primer orden. No obstante, en algunos problemas puede resultar particularmente útil el abordaje a partir de contigüidades de segundo orden.

**FIGURA 3: Criterios de contigüidad en la matriz de pesos espaciales. Retardo espacial de segundo orden**



FUENTE: Elaboración propia

En general, no existen situaciones que requieran la adopción de un criterio de vecindad que vaya más allá del segundo orden. En todo caso, sería necesario explorar otras hipótesis de vecindad que respondan a otras formulaciones sobre la base de especificaciones económicas o geográficas alternativas, como pueden ser: la distancia entre dos unidades, el nivel de intercambio comercial entre diferentes regiones o países, etc. En todos los casos, lo que resultará importante aquí, será escoger variables que sean operativas. En la próxima sección se presentan algunas de las formulaciones alternativas más difundidas.

### 3.2 Otras especificaciones para la matriz de contactos

Existen numerosos criterios para definir cuándo dos regiones pueden considerarse vecinas. La reseña de procedimientos presentada a continuación no pretende ser exhaustiva, pero los mismos constituyen las alternativas más difundidas.

1) Cliff y Ord (1981) construyeron una matriz sobre la base de la distancia que separa a dos regiones  $i$  y  $j$ ,  $(d_{ij})$ , ponderada por el tamaño de la frontera que presenta en común  $(\beta_{ij})$ . En este caso la formulación da lugar a una matriz asimétrica. De esta forma los elementos de  $\mathbf{W}$  se reducen a:

$$w_{ij} = (d_{ij})^{-a} (\beta_{ij})^b \quad (2)$$

Donde  $a$  y  $b$  son parámetros a estimar. Anselin (1980), comenta que dichos parámetros suelen ser dados a priori y no estimados conjuntamente).

2) Bodson y Peeters (1975) han sugerido que la matriz  $\mathbf{W}$  considere un criterio de accesibilidad general. Esto es, que tenga en cuenta y combine los distintos canales de comunicación entre regiones. En consecuencia, esta formulación permite redefinir el concepto de vecindad, ya que para serlo habrá de haber buenos canales de comunicación entre las regiones. La formulación analítica de esta propuesta se expresa como:

$$w_{ij} = \sum_{n=1}^N k_n \left\{ \frac{a}{[1+b*\exp(-c_j d_{ij})]} \right\} \quad (3)$$

Siendo:

$k_n$  la importancia relativa de la vía de comunicación  $n$ ,

$N$  el número de vías de comunicación,

$d_{ij}$  la distancia entre las regiones  $i$  y  $j$ .

$a$ ,  $b$  y  $c_j$  son parámetros a estimar.

3) Case *et al.* (1993) han propuesto que la matriz  $\mathbf{W}$  se sustente en distancias económicas. En este caso, se asumirá que una región es vecina de otra cuando esas distancias, en términos de alguna variable económica operativa, sea pequeña (lo que deberá ser definido por el investigador). A menudo, desde esta perspectiva, la distancia económica es entendida como la reducción de los costos de transacción asociados a la interrelación económica de regiones lejanas entre sí (no contiguas). Esto es presumiblemente provocado por mejoras que se asocian a condiciones económicas estructurales similares y costos de información más bajos.

## 4. El análisis exploratorio de datos espaciales (AEDE)

El análisis exploratorio de datos espaciales –AEDE- (o ESDA por sus siglas en inglés Exploratory Spatial Data Analysis) es comúnmente considerado un subconjunto del análisis exploratorio de datos – AED- (o EDA -Exploratory Data Analysis-) que se ocupa de las características distintivas de datos geográficos, con especial énfasis en los problemas de autocorrelación espacial y heterogeneidad espacial (Anselin, 1998; Haining, 1990).

El AEDE puede definirse como un “una colección de técnicas para describir y visualizar distribuciones espaciales, identificar localizaciones espaciales atípicas o outliers espaciales, descubrir patrones de asociación espacial, clusters o puntos calientes, y sugerir regímenes espaciales u otras formas heterogeneidad espacial” (Anselin, 1998; p. 4).

De esta forma, al igual que el AED, el AEDE procura detectar en los datos posibles patrones y sugerir hipótesis sobre posibles relaciones entre las variables involucradas en el análisis. Dentro de la colección de técnicas incluidas en el AEDE se encuentran:

- Visualización de distribuciones espaciales
- Visualización de asociación espacial
- Indicadores locales de asociación espacial (LISA – Local Indicator of Spatial Association-)
- Indicadores multivariados de asociación espacial

El desarrollo de nuevos paquetes informáticos y el mejoramiento de otros que ya existían, han permitido a dichas técnicas alcanzar su potencial como procedimiento para proponer posibles conexiones entre las variables. En el caso de aplicación, al final del capítulo, se trabajará con el software libre *OpenGeoDa* desarrollado por Luc Anselin en la Arizona State University.

### 4.1 Técnicas gráficas

Existen dos perspectivas para llevar a cabo un AEDE a partir de técnicas gráficas. Por un lado, la aproximación geoestadística y por el otro, la aproximación *lattice*. De acuerdo con Anselin (1998) la principal diferencia entre estas dos técnicas se encuentra en la forma en la que la proximidad espacial es formalizada.

### *Abordaje geoestadístico*

En este el supuesto de procesos espaciales continuos conduce al uso de la distancia métrica como medio para organizar las observaciones.

Al respecto, Anselin (1998, p. 5) sostiene: "Dado que la asociación espacial se asume como una función suave de la distancia, una medida formal de la (no)similitud entre dos observaciones, como la diferencia al cuadrado, es comparada con la distancia que los separa. Un mayor grado de autocorrelación espacial implica pequeñas diferencias en distancias cortas y diferencias crecientes en distancias mayores. La función formal que operacionaliza esta noción es el variograma [ver, e.g., Cressie (1993) para detalles técnicos]. Las técnicas del AEDE desde la geoestadística radican en las formas en las que el variograma puede ser visualizado, resumido y probado para la presencia de no estacionariedades locales u otra conducta atípica. Un importante aspecto de esta visualización es que las entidades en un variograma pertenecen a pares de observaciones (separadas por una distancia dada), y no a localizaciones individuales."

Las técnicas gráficas de esta perspectiva se concentran en:

- La función de distribución acumulativa
- El *Box Plot* del variograma
- La nube del variograma multivariado

### *La aproximación lattice*

En este método las observaciones tienen carácter discreto. En consecuencia, la formalización de las similitudes espaciales se desarrolla de una manera diferente a la empleada por el anterior abordaje.

"El concepto clave aquí es la noción de vecino espacial, que conduce a la construcción de matrices de pesos espaciales y variables espaciales rezagadas" (Anselin, 1998; p. 6).

De esta forma, las herramientas del AEDE en la perspectiva lattice intentan reflejar la asociación entre las variables y sus rezagos espaciales, para diferentes definiciones de valores similares y para diferentes pesos espaciales.

Los instrumentos más útiles son:

- El *box map*
- El histograma regional
- El *scartterplot* de Moran
- Los mapas LISA
- El *scartterplot* multivariante

#### 4.2 Medidas de asociación espacial global

El análisis de autocorrelación espacial global realiza un examen conjunto de todas las unidades que componen la muestra para determinar si las unidades espaciales se encuentran distribuidas aleatoriamente o si, por el contrario, lo hacen conforme a un patrón determinado.

Existen tres contrastes desarrollados para ese propósito, que se describen a continuación: la *I* de Moran, la *C* de Geary y la *G(d)* de Getis y Ord.

La expresión analítica del contraste de Moran es:

$$I = \frac{N}{S_0} \frac{\sum_{i,j=1}^N w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2} \quad (4)$$

con  $i \neq j$ , donde  $x_i$  representa el valor de la variable cuantitativa  $x$  para la región  $i$ ,  $N$  el tamaño muestral,  $w_{ij}$  los pesos espaciales de una matriz de contactos  $\mathbf{W}$  y  $S_0$  la sumatoria de los pesos espaciales. La distribución de la *I* de Moran es asintótica normal  $[N \sim (0,1)]$  cuando  $N$  es grande.

Para el caso de la *C* de Geary, la formulación es:

$$C = \frac{N-1}{2S_0} \frac{\sum_{i,j=1}^N w_{ij}(x_i - x_j)^2}{\sum_{i=1}^N (x_i - \bar{x})^2} \quad (5)$$

con  $i \neq j$ , donde los elementos tienen el mismo significado que en la expresión anterior. La *C* de Geary también es asintótica normal  $[N \sim (0,1)]$  cuando  $N$  es grande.

Finalmente, la expresión de la  $G(d)$  es:

$$G(d) = \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} d(x_i x_j)}{\sum_{i=1}^N \sum_{j=1}^N x_i x_j} \quad (6)$$

con  $i \neq j$ , siendo  $i$  y  $j$  vecinos siempre que se encuentren a una distancia  $d$ . El contraste se distribuye normal  $[N \sim (0,1)]$  para  $N$  grande. Nótese que este contraste es aplicable sólo a variables positivas y naturales.

La formulación de los contrastes descriptos anteriormente puede llevarse a cabo con cualquier especificación de matriz  $\mathbf{W}$ , como las descritas en las secciones 3.1 y 3.2. No obstante, los resultados finales de los contrastes, respecto a la determinación de autocorrelación global o no, son sensibles a esas especificaciones.

Otro factor de sensibilidad para con los resultados de los contrastes son las transformaciones que pueden sufrir las variables. Por estas razones, resulta útil ensayar varias formulaciones alternativas de matriz  $\mathbf{W}$ .

La autocorrelación o dependencia espacial detectada podrá ser negativa, positiva o nula. Será negativa cuando los valores se concentren en los cuadrantes II y IV del scatterplot de Moran (representación en reales de la  $I$  de Moran, dada una configuración determinada de la matriz  $\mathbf{W}$ ). En el caso de valores concentrados en los cuadrantes I y III, habrá autocorrelación positiva. Finalmente, habrá autocorrelación nula, si la nube de puntos se encuentra dispersa en los cuatro cuadrantes.

#### 4.3 Medidas de asociación espacial local

Los contrastes analizados en la sección anterior presentan una fuerte limitación: no son capaces de considerar situaciones de aglomeración (*clusters*) en un área determinada, para la que cabría esperar valores más bajos o más altos si existiese una distribución homogénea (Moreno y Vayá, 2000). Esto implica que un determinado esquema de

autocorrelación espacial, detectado mediante contrastes globales, pueda no mantenerse para toda la muestra.

Para la detección de aglomeraciones (*clusters*) resulta útil el análisis de *autocorrelación espacial local*. Este procedimiento permite medir la autocorrelación espacial para la ubicación de cada observación y está basado en el Indicador Local de Moran.

De acuerdo con Anselin (1995) un indicador LISA (Local Indicator of Spatial Association) es un estadístico que satisface dos requerimientos: por un lado, proporciona una cuantificación del grado de agrupamiento significativo de valores similares alrededor de una observación, y por otro lado, la suma de los LISA's para todas las observaciones es proporcional a un indicador global de asociación espacial, por lo que resulta útil para medir la contribución de cada observación al valor del contraste global (sólo para el caso de la  $I$  de Moran).

Este análisis es también particularmente útil para detectar la posible presencia de localizaciones que muestren valores altos respecto a la media de otros emplazamientos colindantes, y para las variables de interés, o viceversa.

Los contrastes más relevantes para el análisis de la autocorrelación local son: la  $I$  de Moran y la  $G(d)$  de Getis y Ord.

La expresión analítica del contraste de Moran es:

$$I = \frac{Z_i}{\sum_i Z_i / N} \sum_{j \in J_i} w_{ij} Z_j \quad (7)$$

Donde  $Z_i$  representa el valor de la región  $i$  para la variable normalizada y  $J_i$  el conjunto de regiones vecinas a  $i$ .  $N$  es el tamaño muestral. La distribución de la  $I$  de Moran es asintótica normal  $[N \sim (0,1)]$  cuando  $N$  es grande. Finalmente, si el valor de la  $I$  es positivo, denotará la existencia de un cluster de valores similares.

La expresión de la  $G(d)$  es:

$$G(d) = \frac{\sum_{j=1}^N w_{ij} d(x_j)}{\sum_{j=1}^N x_j} \quad (8)$$

con  $i \neq j$ , siendo  $i$  y  $j$  vecinos siempre que se encuentren a una distancia  $d$ ,  $x$  la variable de interés no normalizada. El contraste se distribuye normal  $[N \sim (0,1)]$  para  $N$  grande. Nótese que este contraste es aplicable sólo a variables positivas y naturales.

## 5. El análisis confirmatorio

El análisis confirmatorio se ocupa de las diferentes especificaciones posibles para la inclusión del fenómeno espacial en un modelo. Para ello, existen diferentes contrastes diseñados para detectar la existencia de dependencia espacial, y en tal caso su tipología.

Para el testeo de *dependencia espacial sustantiva* son habituales los contrastes basados en los multiplicadores de Lagrange, el LM-LAG propuesto por Anselin (1988) y el test LM-LE de Bera y Yoon (1992). El LM-LE es robusto frente a la existencia de un término de perturbación correlacionado espacialmente.

Para la *dependencia espacial residual* es frecuente la utilización de los contrastes LM-ERR (Burridge, 1980) y su versión robusta LM-EL (Bera y Yoon, 1992), también basados en los multiplicadores de Lagrange. La ventaja del test LM-EL es que resulta robusto ante posibles especificaciones erróneas locales como la presencia de una variable endógena retardada espacialmente (Moreno y Vayá, 2000: p. 82).

Finalmente, el test SARMA permite contrastar la *existencia conjunta de ambos tipos de dependencia espacial*.

Para comprender luego las diferentes tipologías de dependencia espacial, considérese el siguiente modelo de partida:

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\mu} \\ \boldsymbol{\mu} &\sim N(0, \sigma_{\mu}^2 \mathbf{I}_N) \end{aligned} \tag{9}$$

Donde  $X$  es una matriz de  $N \times K$  con variables explicativas.



### 5.1 La dependencia espacial sustantiva

La dependencia espacial sustantiva (también conocida como *modelo lag*) se presenta como un caso en el que el valor de una región depende del valor de sus regiones vecinas (definidas según una matriz **W** de contactos).

La especificación de un modelo de este tipo viene dado por la siguiente expresión:

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (10)$$

Donde  $\mathbf{y}$  es un vector ( $N \times 1$ ) de las  $N$  observaciones de la variable dependiente, **W** representa la matriz de pesos,  $\rho$  es el parámetro autorregresivo que recoge la intensidad de las interdependencias entre las  $N$  observaciones muestrales (coeficiente de autocorrelación espacial),  $\mathbf{W}\mathbf{y}$  es el retardo espacial de la variable  $\mathbf{y}$ , **X** es una matriz de  $k$  variables exógenas, y  $\boldsymbol{\varepsilon}$  un término de perturbación ruido blanco.

Se torna evidente en la expresión (9) que la omisión del retardo espacial derivaría en un incremento del término de perturbación, que pasaría a reflejar dicha autocorrelación espacial.

La contrastación de la existencia de dependencia espacial sustantiva, en ausencia de dependencia espacial residual, se lleva a cabo mediante el test LM – LAG cuya expresión es la siguiente:

$$LM - LAG = \frac{[(\mathbf{e}'\mathbf{W}\mathbf{y})/S^2]^2}{RJ_{\rho-\beta}}$$

Siendo:

$\mathbf{e}$  un vector de residuos MCO de la regresión representada por (9)

**W** una matriz de contactos definida ex ante,

$S^2$  la estimación de la varianza residual del modelo representado por (9),

$RJ_{\rho-\beta} = T1 + [(\mathbf{W}\mathbf{X}\boldsymbol{\beta})'\mathbf{M}(\mathbf{W}\mathbf{X}\boldsymbol{\beta})]/S^2$ , donde  $T1$  representa la traza de  $(\mathbf{W}'\mathbf{W} + \mathbf{W}^2)$  y **M** es una matriz idempotente,  $\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})\mathbf{X}'$

La versión robusta del LM – LAG frente a la existencia de un término de perturbación correlacionado espacialmente, está dado por el LM – LE, según la siguiente expresión:

$$LM - LE = \frac{[(\mathbf{e}'\mathbf{W}\mathbf{y})/S^2 - (\mathbf{e}'\mathbf{W}\mathbf{e})/S^2]^2}{R_{\rho-\beta} - T1}$$

Teniendo los símbolos los mismos significados que la expresión anterior.

En ambos contrastes, la hipótesis nula es la no existencia de dependencia espacial sustantiva, lo que implica  $\rho = 0$  en la ecuación (10), frente a la alternativa de  $\rho \neq 0$ .

## 5.2 La dependencia espacial residual

La dependencia espacial subyacente puede ser residual cuando su origen no sea la dependencia sustantiva, en cuyo caso la correcta especificación vendría dada por el siguiente modelo de *error espacial*:

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} &= \lambda\mathbf{W}\boldsymbol{\varepsilon} + \boldsymbol{\mu} \\ \boldsymbol{\mu} &\sim N(0, \sigma_{\mu}^2\mathbf{I}_N) \end{aligned} \tag{11}$$

La autocorrelación espacial es recogida aquí por  $\lambda$  a través del término de perturbación del modelo,  $\boldsymbol{\varepsilon}$ . La dependencia espacial residual podría explicarse “por la omisión de variables no cruciales que se hallen correlacionadas espacialmente o por la existencia de errores de medida” (Moreno y Vayá, 2000: p. 69).

La contrastación de dependencia espacial en el término perturbación puede llevarse a cabo mediante los contrastes basados en los multiplicadores de Lagrange LM-ERR y LM-EL. El primero de ellos, utilizado en situaciones de perturbaciones homocedásticas y ausencia de dependencia espacial sustantiva, presenta la siguiente expresión analítica:

$$LM - ERR = \frac{[(\mathbf{e}'\mathbf{W}\mathbf{e})/S^2]^2}{T1}$$

Donde los símbolos tienen el mismo significado que en las expresiones anteriores. Ahora bien, si existe heterocedasticidad, el LM-ERR debe modificarse para considerar esta situación:

$$LM - ERR_{heterocedasticidad} = \frac{\left[ \frac{(\mathbf{e}'\mathbf{\Omega}^{-1}\mathbf{W}\mathbf{e})}{S^2} \right]^2}{T} \sim \chi^2(1)$$

Siendo  $T = tr(\mathbf{W}\mathbf{W} + \mathbf{\Omega}\mathbf{W}'\mathbf{\Omega}^{-1}\mathbf{W})$  y  $\mathbf{\Omega}$  la distribución de la varianza de un modelo como (11) pero con heterocedasticidad.

Por otra parte el test LM-EL, que es una versión robusta del LM-ERR frente a posibles especificaciones erróneas locales como la presencia de una variable endógena retardada espacialmente, presenta la siguiente estructura:

$$LM - EL = \frac{[NI]^2}{T1}$$

donde  $I$  es el contraste de Moran y  $N$  el tamaño muestral.

En ambos casos, la hipótesis nula es la ausencia de dependencia espacial en el término perturbación, lo que implica  $\lambda = 0$  en la ecuación (11), frente a la alternativa de existencia de un esquema de media móvil o autorregresivo en el término de la perturbación,  $\lambda \neq 0$ .

## 6. Tratamiento de la heterogeneidad espacial

### 6.1 Contrastación de heterogeneidad espacial

La heterogeneidad espacial puede exteriorizarse a través de dos formas: la heterocedasticidad o la inestabilidad estructural.

Los habituales contrastes de Breusch-Pagan y White, comúnmente utilizados en el análisis clásico, son también válidos aquí, siempre que no exista alguna otra especificación errónea, como la dependencia espacial o inestabilidad estructural. En estos últimos casos, los contrastes mencionados anteriormente ya no resultarán adecuados y requieren de algunas modificaciones.

La razón por la que la contrastación de la heterogeneidad se realiza en ausencia de algún tipo de dependencia espacial, se debe a que los modelos espaciales (con dependencia sustantiva o espacial) añaden un componente espacial a cambio de imponer homogeneidad en los efectos espaciales. Esto es, se considera que la estimación del

parámetro espacial es válida para todas las regiones y que el residuo de las ecuaciones (10) ó (11) tendrá un carácter aleatorio.

## 6.2 Especificación de un modelo con heterogeneidad espacial

Una de las soluciones más difundidas para modelos con heterogeneidad espacial, ha consistido en especificar y estimar un modelo de *expansión lineal espacial*. Considérese un modelo como el siguiente:

$$y = \alpha + \beta X + \varepsilon \quad (12)$$

La expansión se realiza para la constante, como así también para los  $\beta$ . Las coordenadas podrán actuar como variables de expansión, siendo  $\gamma$  la variable que considera la tendencia norte-sur y  $\psi$  la tendencia este-oeste. Bajo estas consideraciones, se tiene que:

$$\alpha_i = \alpha_0 + \alpha_1 \gamma_i + \alpha_2 \psi_i \quad (13)$$

$$\beta_i = \beta_0 + \beta_1 \gamma_i + \beta_2 \psi_i \quad (14)$$

Donde los valores de ordenada al origen y pendiente cambiarán según la región  $i$ .

Incorporando (13) y (14) en (12), se obtiene que:

$$y = \alpha_0 + \alpha_1 \gamma_1 + \alpha_1 \psi_1 + \alpha_2 \gamma_2 + \alpha_2 \psi_2 + \dots + \alpha_k \gamma_k + \alpha_k \psi_k + \beta_1 \gamma_i + \beta_2 \psi_i \quad (15)$$

En esta última expresión los parámetros varían en el espacio, dando lugar a estimaciones consistentes.

Otra posibilidad, igualmente difundida, ha consistido en realizar la expansión según las regiones formen parte de una región núcleo o una región marginal (o centro vs. periferia en otros términos).

## CASOS DE ESTUDIO, PREGUNTAS Y PROBLEMAS

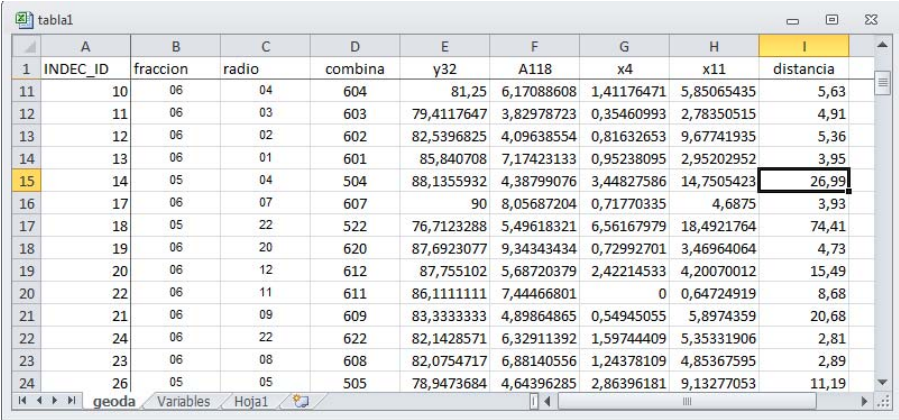
### Ejercicio con Geoda: Radios censales de Río Cuarto

La base de datos del Censo de Población de la Provincia de Córdoba, realizado en 2008, está publicada en <http://estadistica.cba.gov.ar> en dos versiones: por localidades y por radio censal. La tabla organizada por localidades fue utilizada para análisis exploratorio.

La tabla organizada por radio censal, que contiene 1789 variables y 156 observaciones, fue sometida a análisis exploratorio, clasificación y segmentación; de acuerdo al resultado obtenido se han seleccionado cuatro variables:

- X11: personas con necesidades básicas insatisfechas
- Y132: tasa de actividad de 30 a 40 años
- X4: hogares con condiciones sanitarias deficitarias
- A118: mujeres con estudios terciarios o superiores completos

Este conjunto de variables está organizado en Tabla1.xlsx donde, además, se encuentran el identificador de filas (INDEC\_ID), la fracción, el radio y la distancia entre los radios.

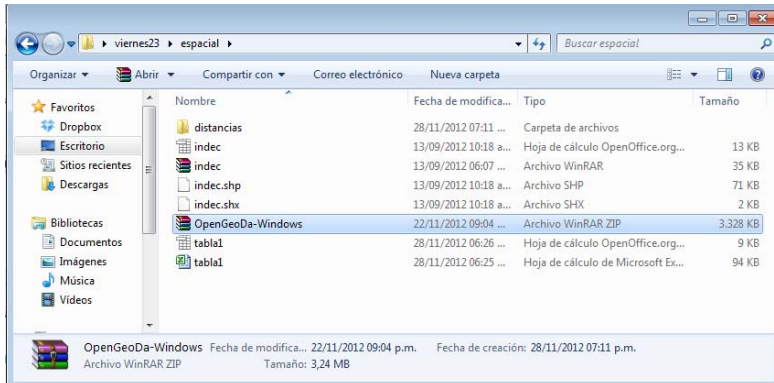


	A	B	C	D	E	F	G	H	I
1	INDEC_ID	fraccion	radio	combina	y32	A118	x4	x11	distancia
11	10	06	04	604	81,25	6,17088608	1,41176471	5,85065435	5,63
12	11	06	03	603	79,4117647	3,82978723	0,35460993	2,78350515	4,91
13	12	06	02	602	82,5396825	4,09638554	0,81632653	9,67741935	5,36
14	13	06	01	601	85,840708	7,17423133	0,95238095	2,95202952	3,95
15	14	05	04	504	88,1355932	4,38799076	3,44827586	14,7505423	26,99
16	17	06	07	607	90	8,05687204	0,71770335	4,6875	3,93
17	18	05	22	522	76,7123288	5,49618321	6,56167979	18,4921764	74,41
18	19	06	20	620	87,6923077	9,34343434	0,72992701	3,46964064	4,73
19	20	06	12	612	87,755102	5,68720379	2,42214533	4,20070012	15,49
20	22	06	11	611	86,1111111	7,44466801	0	0,64724919	8,68
21	21	06	09	609	83,3333333	4,89864865	0,54945055	5,8974359	20,68
22	24	06	22	622	82,1428571	6,32911392	1,59744409	5,35331906	2,81
23	23	06	08	608	82,0754717	6,88140556	1,24378109	4,85367595	2,89
24	26	05	05	505	78,9473684	4,64396285	2,86396181	9,13277053	11,19

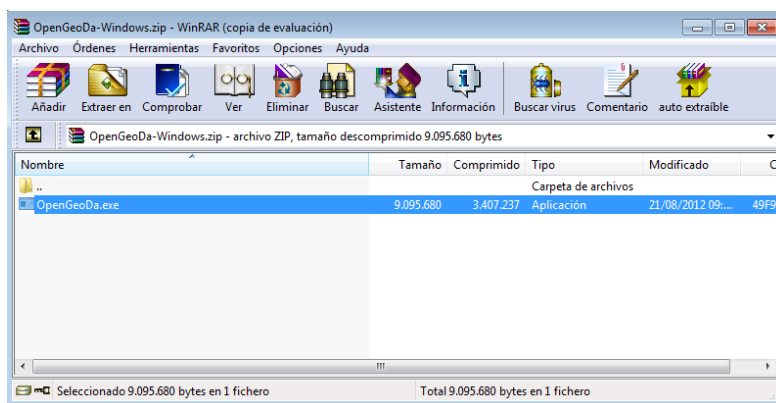
## Introducción a OpenGeoDa

Geoda es un software libre que está disponible en <https://geodacenter.asu.edu/>; es necesario registrarse como usuario para obtener una licencia, siendo este registro gratuito. Este software permite georreferenciar en mapas variables observadas en un territorio y realizar análisis de heterogeneidad y autocorrelación espacial.

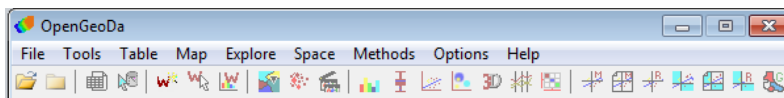
El software consiste en un archivo autoejecutable que se encuentra en el archivo zipeado OpenGeoDa-Windows.



Al clicar en el archivo zipeado se observa el ejecutable



al clicar sobre él se ejecuta el programa, cuya pantalla es la barra de herramientas.



Open GeoDa trabaja con tres archivos:

- Una tabla de datos en formato dbf
- Un mapa en formato shp
- Un vínculo en formato shx

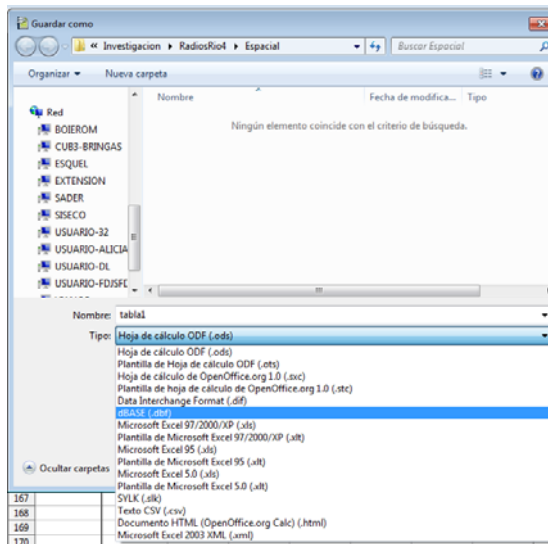
Este último vincula la tabla de datos con el mapa. Los archivos shp y shx se pueden generar en GeoDa, para lo cual se necesitan las coordenadas del espacio a georreferenciar. El archivo dbf se obtiene por construcción a partir de un archivo Excel. Hay varios software que permiten convertir un archivo Excel en dbf, entre ellos se encuentra el paquete OpenOffice –software libre que se obtiene desde [www.openoffice.org](http://www.openoffice.org); particularmente, OpenOfficeCalc quien tiene una panalla similar a Excel

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Identificador	fraccion	radio	combina	y32	A118	x4	x11									
143	143	11	11	02	1102	98,4126984	9,13978495	0	1,05134836								
144	144	11	11	03	1103	88,8888889	5,24861878	1,2295082	5,34801789								
145	144	11	11	04	1104	83,7209302	7,29483283	0,70921986	5,40285177								
146	145	11	11	05	1105	72,8	6,94668821	0	3,04878049								
147	146	11	11	06	1106	94,7368421	8,92018779	0	3,44827586								
148	147	11	11	07	1107	92,1052632	6,36942675	0,60790274	3,36134454								
149	148	11	11	08	1108	80,6818182	4,64716007	1,2987013	5,7913043								
150	149	11	11	09	1109	78,3018968	2,85343709	5,08474576	15,6973108								
151	150	11	11	10	1110	82,8571429	2,10526316	0	5,13784461								
152	151	11	11	11	1111	75	3,91304348	1,37931034	11,1842105								
153	152	11	11	12	1112	81,25	7,37586852	0,4587136	5,6445993								
154	153	11	11	13	1113	86,5853659	6,46464646	0,93457944	6,51955868								
155	154	11	11	14	1114	85	6,54338549	0,23752969	2,18023256								
156	155	11	11	15	1115	81,4814815	6,27240143	0	5,51537071								
157	156	11	11	16	1116	84,8484848	5,55555556	1,5625	7,98226164								
158	157	11	11	17	1117	84	8,40108401	0,93023256	4,28211587								
159	158	11	11	18	1118	89,7727273	5,31309298	1,20845921	7,40740741								
160	159	11	11	19	1119	81,875	4,0553907	0,85324232	10,0094429								
161	160	11	11	20	1120	72,2627737	1,26002291	3,18091451	22,9903537								
162	161	11	11	21	1121	76,4705882	5,01002004	0,68728522	9,10871694								
163	162	11	11	22	1122	89,7959184	6,9124424	0,390625	10,3487064								
164	163	11	11	23	1123	80,4878049	4,83870968	2,2556391	11,1849391								

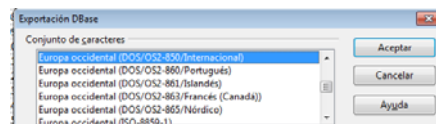
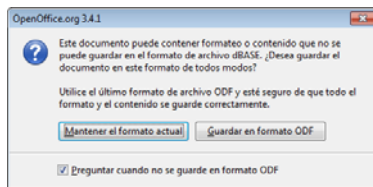
Se debe ir a Archivo-Guardar como...

	D	E	F	A:
	radio	combina	y32	
	02	1102	98,4126984	9
	03	1103	88,8888889	5
	04	1104	83,7209302	7
	05	1105	72,8	6
	06	1106	94,7368421	8
	07	1107	92,1052632	6
	08	1108	80,6818182	4
	09	1109	78,3018968	2
	10	1110	82,8571429	2
	11	1111	75	3
	12	1112	81,25	7
	13	1113	86,5853659	6
	14	1114	85	6
	15	1115	81,4814815	6
	16	1116	84,8484848	5
	17	1117	84	8
	18	1118	89,7727273	5
	19	1119	81,875	1
	20	1120	72,2627737	1
	21	1121	76,4705882	5
	22	1122	89,7959184	4
	23	1123	80,4878049	4

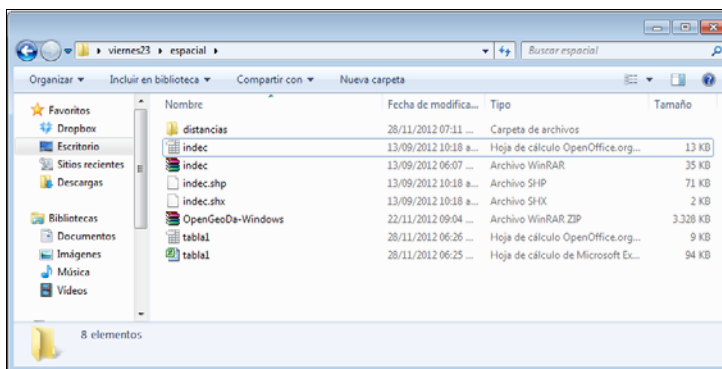
seleccionar la carpeta donde guardar el archivo y el tipo de archivo, aquí es necesario seleccionar dbase (.dbf) y darle un nombre al archivo que se va a generar



Cuando se guarda el archivo, aparece un cuadro de diálogo que permite confirmar el formato, se debe seleccionar *Mantener el formato actual*. Luego se ve otro cuadro de diálogo para seleccionar el conjunto de caracteres, se acepta el que aparece por defecto.



Ahora se tienen



indec.dbf coordenadas del mapa  
indec.shp mapa  
indec.shx vincula coordenadas y mapa  
tabla1.dbf datos a analizar  
tabla1.xls datos a analizar

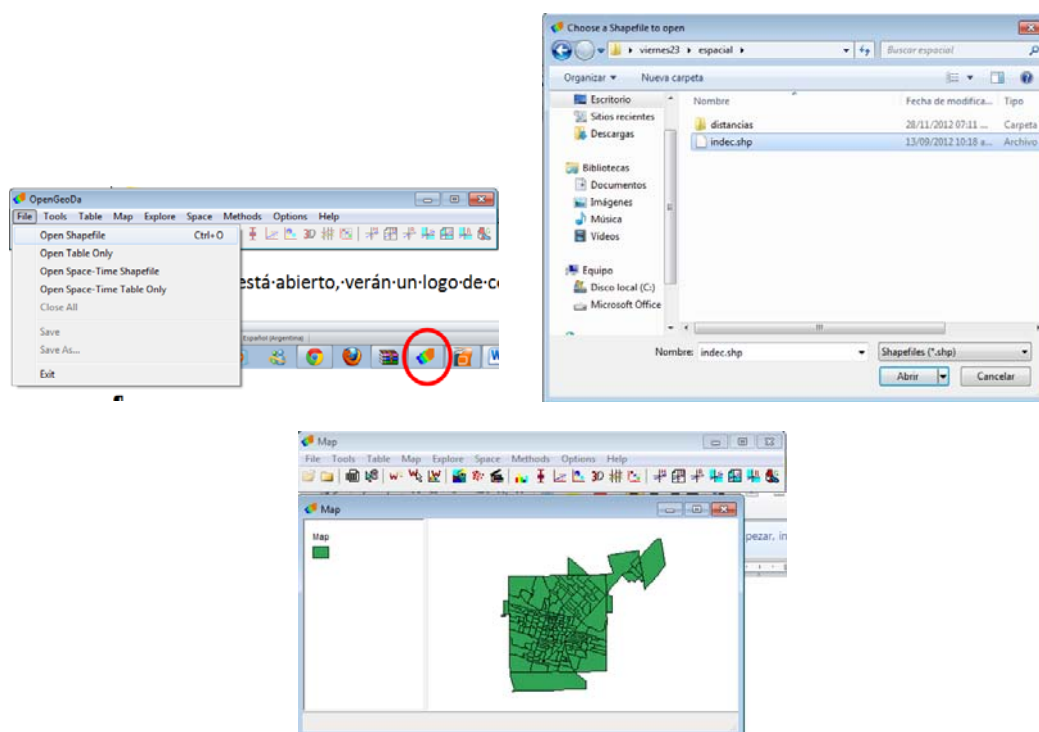
El logo de OpenGeoDa es de colores tipo paleta de pintor





## Crear un entorno de trabajo

Para comenzar a trabajar se debe seleccionar el mapa asociado a la tabla de datos. Para esto se sigue la secuencia *File-Open Shapefile* y se localiza el archivo indec.shp; al hacer click en abrir se observa el mapa de radios censales de Río Cuarto.



Para ver la información que tiene indec.dbf, se debe hacer click en open table. La información contenida en indec.dbf son las coordenadas del mapa (AREA y PERIMETER), INDEC\_ID es la llave (key variable) y permite anexar datos de interés a esta tabla, esta variable va del 1 al 163; la información que sigue hace referencia a la provincia (PRO=14 es Córdoba), departamento (DEP=98 es Río Cuarto), localidad (LOC=170 es la ciudad de Río Cuarto), fracción (FRA) y radio(RAD), estos dos últimos asumen diferentes valores de modo que cada combinación sea única. Es necesario revisar la tabla de datos donde se encuentra la información de interés para que coincida la información; es decir, que la tabla de datos de interés tenga en la primera fila la información correspondiente a la fracción 4 y al radio 7, y así sucesivamente.

	AREA	PERIMETER	INDEC_ID	INDEC_ID	PRO	DEP	LOC	FRA	RAD
1	3889460.09523	8809.45022	2	1	14	98	170	4	7
2	3336416.63199	9775.84021	3	2	14	98	170	4	8
3	293388.68048	2353.15405	4	3	14	98	170	4	9
4	969559.90435	4712.62586	5	5	14	98	230	4	37
5	347168.00343	3471.33722	6	4	14	98	170	4	36
6	719155.13304	2690.81835	7	6	14	98	170	4	10
7	502302.16388	3268.49907	8	7	14	98	170	4	11
8	1145095.41301	4649.79467	9	8	14	98	230	6	6
9	357812.30131	2616.11181	10	9	14	98	230	6	7
10	373530.53888	3174.08935	11	10	14	98	230	6	4
11	107749.39779	1423.77099	12	11	14	98	230	6	3
12	389543.14627	3579.14188	13	12	14	98	230	6	2
13	322743.13823	3265.20078	14	13	14	98	230	6	1
14	853619.04207	4912.91186	15	14	14	98	230	5	4
15	346048.14176	2427.29282	16	17	14	98	230	6	7
16	219571.87247	2579.68428	17	18	14	98	230	5	22
17	182643.82369	2049.80693	18	19	14	98	230	6	20
18	188094.60120	2773.34443	19	20	14	98	230	6	12
19	234728.139182	2201.99372	20	22	14	98	230	6	11
20	252425.26201	2477.61025	21	21	14	98	230	6	9
21	285417.23285	2504.77111	22	24	14	98	230	6	22
22	286427.26830	3025.37906	23	23	14	98	230	6	8
23	386785.13661	3636.81347	24	26	14	98	230	6	6

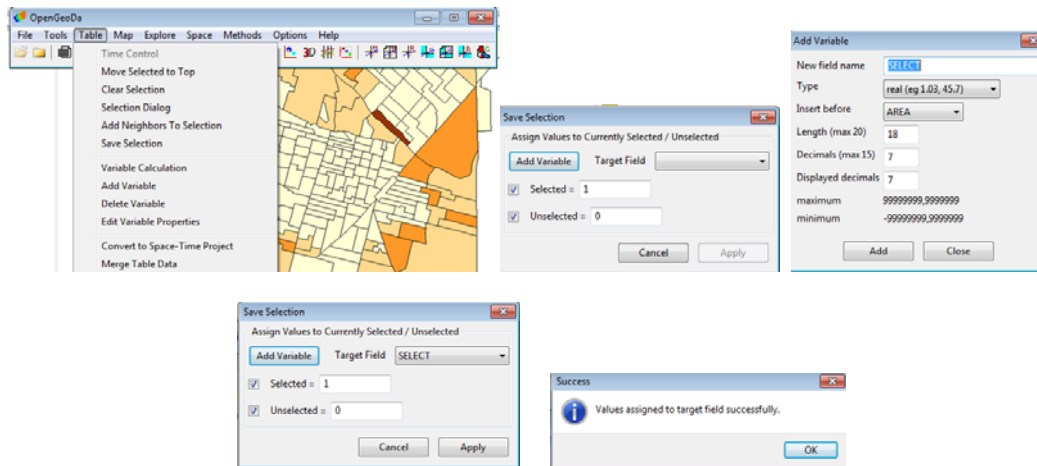
Para empalmar los datos de interés -existentes en Tabla1.dbf- con las coordenadas de indec.dbf, se debe ir a *Table-Merge Table Data*. Se despliega el cuadro de diálogo que permite configurar la fusión de los datos. En *Import file* se debe seleccionar el archivo dbf que contiene la información de interés. La fusión puede hacerse a través de una variable llave o a través del orden dado a las filas; es decir, si hay seguridad de que el orden en el que se encuentran los datos coincide con el orden dado en la tabla que contiene la información para georreferencia, se puede usar la segunda opción *Merge by recorder*. Se observa en la ventana *Exclude* las variables contenidas en el archivo que se está fusionando, las cuales deben adquirir el rol de *Include* para que efectivamente sean empalmadas; al clickear en Merge, aparece la tabla con toda la información.

The image shows the 'Merge - indec' dialog box in the Map software. The 'Input file' is set to 'C:\Users\Ana Vianco\Desktop\viernes23\especial\tab'. The 'Merge by key values' option is selected, with 'current table key' and 'import table key' dropdowns. The 'Merge by record order' option is also visible. The 'Exclude' list contains: INDEC\_ID, FRACCION, RADIO, COMBINA. The 'Include' list contains: Y32, A118, X4, X11, DISTANCIA. The 'Merge' button is highlighted.

Below the dialog box is a screenshot of the resulting data table, which has 22 rows and 14 columns. The columns are: INDEC\_, INDEC\_ID, PRO, DEP, LOC, FRA, RAD, Y32, A118, X4, X11, DISTANCIA. The data is as follows:

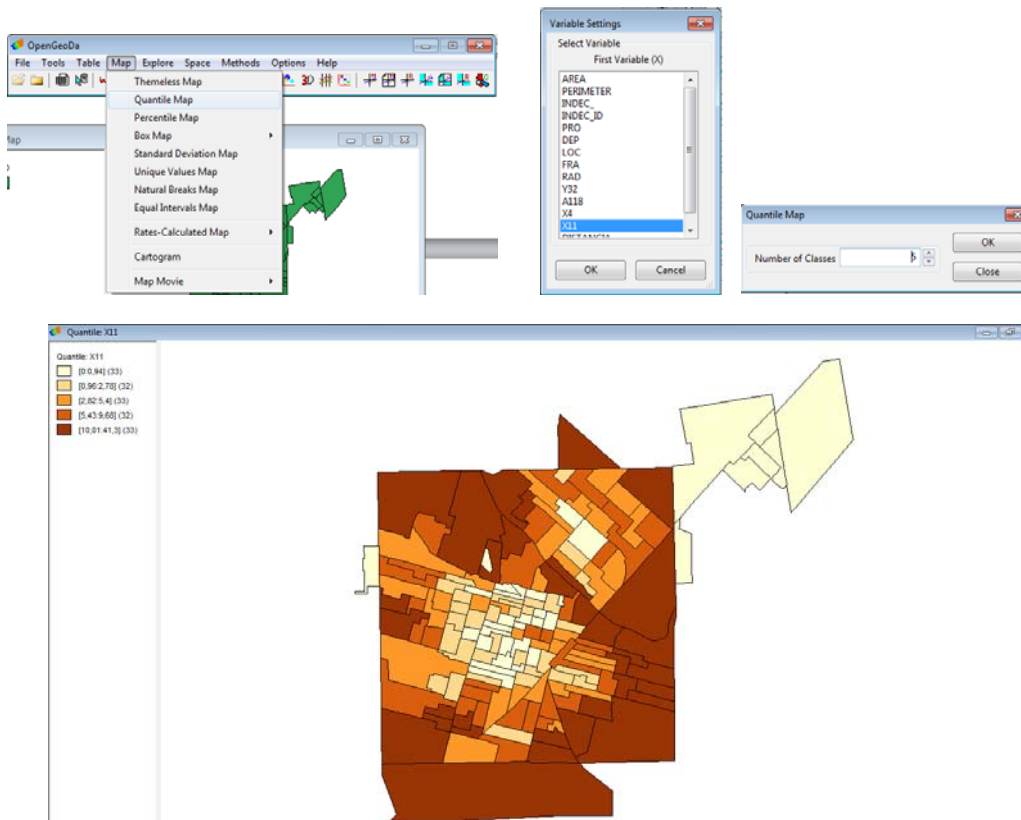
	INDEC_	INDEC_ID	PRO	DEP	LOC	FRA	RAD	Y32	A118	X4	X11	DISTANCIA
1	2	1	14	98	170	4	7	0,00	0,00	0,00	0,00	0,00
2	3	2	14	98	170	4	8	0,00	0,00	0,00	0,00	0,00
3	4	3	14	98	170	4	9	0,00	0,00	0,00	0,00	0,00
4	5	5	14	98	230	4	37	75,68	3,22	1,60	20,39	71,28
5	6	4	14	98	170	4	36	0,00	0,00	0,00	0,00	0,00
6	7	6	14	98	170	4	10	0,00	0,00	0,00	0,00	7,54
7	8	7	14	98	170	4	11	0,00	0,00	0,00	0,00	0,00
8	9	8	14	98	230	6	6	89,62	7,78	6,62	11,44	40,80
9	10	9	14	98	230	6	5	87,63	6,01	1,66	4,36	2,67
10	11	10	14	98	230	6	4	81,25	6,17	1,41	5,85	5,63
11	12	11	14	98	230	6	3	79,41	3,83	0,35	2,78	4,91
12	13	12	14	98	230	6	2	82,54	4,10	0,82	9,68	5,36
13	14	13	14	98	230	6	1	85,84	7,17	0,95	2,95	3,95
14	15	14	14	98	230	5	4	88,14	4,39	3,45	14,75	26,99
15	16	17	14	98	230	6	7	90,00	8,06	0,72	4,69	3,93
16	17	18	14	98	230	5	22	76,71	5,50	6,56	18,49	74,41
17	18	19	14	98	230	6	20	87,69	9,34	0,73	3,47	4,73
18	19	20	14	98	230	6	12	87,76	5,69	2,42	4,20	15,49
19	20	22	14	98	230	6	11	86,11	7,44	0,00	0,65	8,68
20	21	21	14	98	230	6	9	83,33	4,90	0,55	5,90	20,68
21	22	24	14	98	230	6	22	82,14	6,33	1,60	5,35	2,81
22	23	23	14	98	230	6	8	82,08	6,88	1,24	4,85	2,89

Para guardar esta tabla, se sigue la secuencia *Table-Save Selection-Add Variable-Add-Apply-Ok*. El nombre en *New file name* lo ofrece por defecto, si aparece el mensaje de la última imagen, significa que

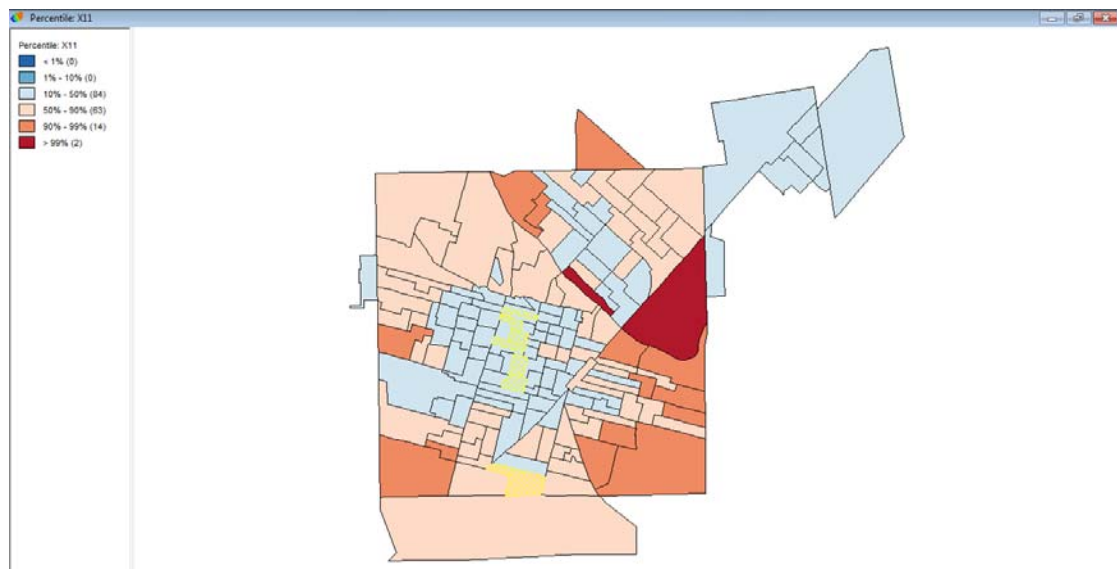


## Georreferencia de las variables

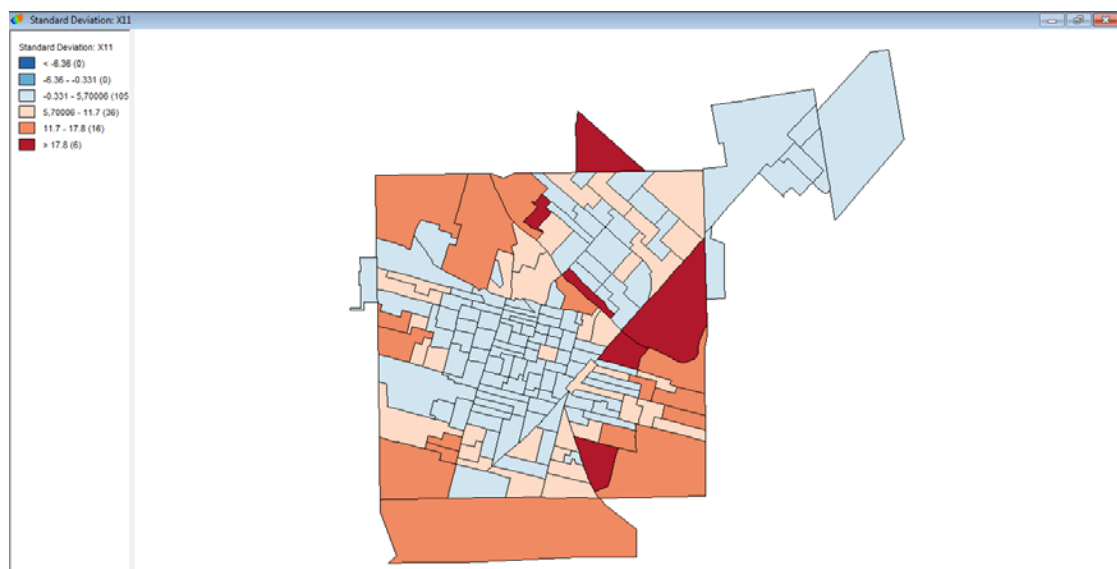
Para hacer georreferencia se trabaja desde la herramienta *Map*. Al seleccionar la opción *Quantile map*, es necesario indicar la variable que se quiere georreferenciar –en este caso X11– el cuadro de diálogo siguiente solicita la cantidad de clases en la que se va a particionar el recorrido de la variable –se opta por 5 clases–; esta configuración realiza el gráfico para los quintiles de la variable X11. Las zonas de color más intenso refieren los lugares donde la variable registra valores más altos.



De igual manera se trabaja con el resto de los mapas. Por ejemplo, el mapa de percentiles (*Map Percentile Map*) genera la imagen



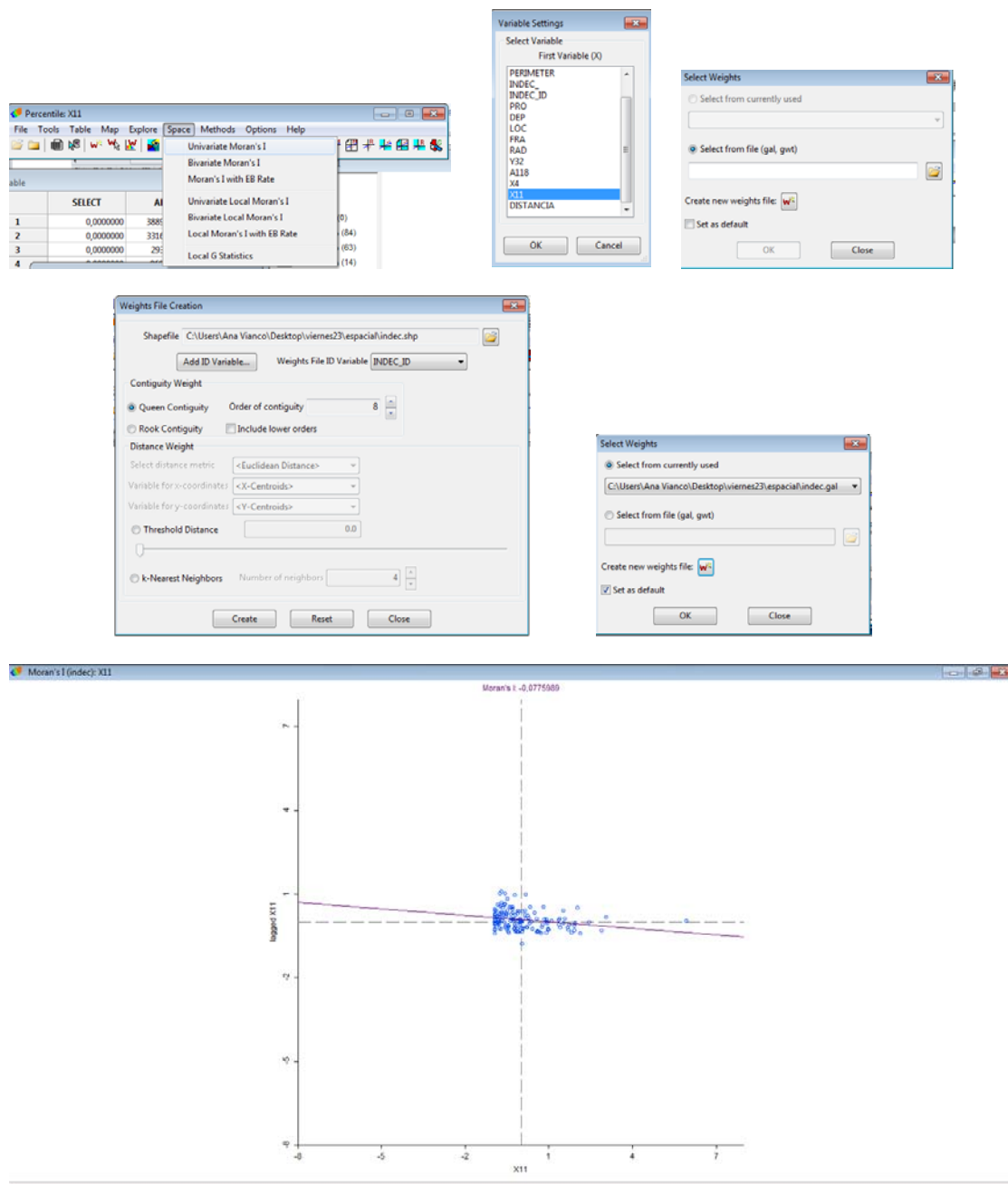
y el mapa de los desvíos *Map Estándar Desviation Map*, presenta los radios censales con mayor variabilidad



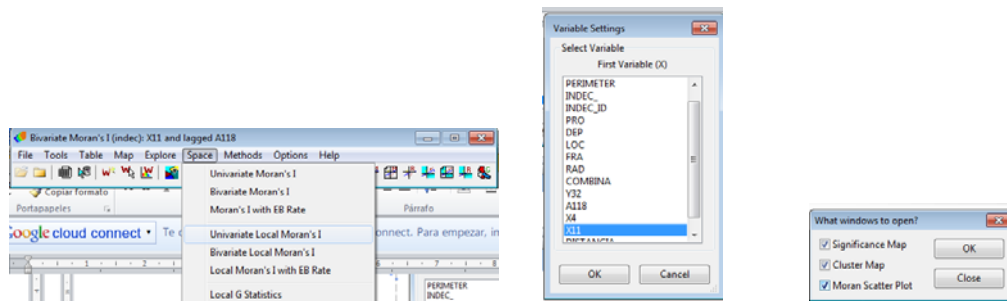
### Indicadores del espacio

En la herramienta *Space* se encuentra el Índice de Moran, tanto la versión para su cálculo global como para el cálculo del local. La secuencia a seguir es *Space-Univariante Moran's I-seleccionar la variable a analizar*; el cálculo de este indicador requiere una matriz de pesos que se construye desde *Create new weights file*. La secuencia es la selección del archivo shp (*Shapefile*), indicar la variable llave (Key variable: INDEC\_ID), seleccionar la matriz de pesos a

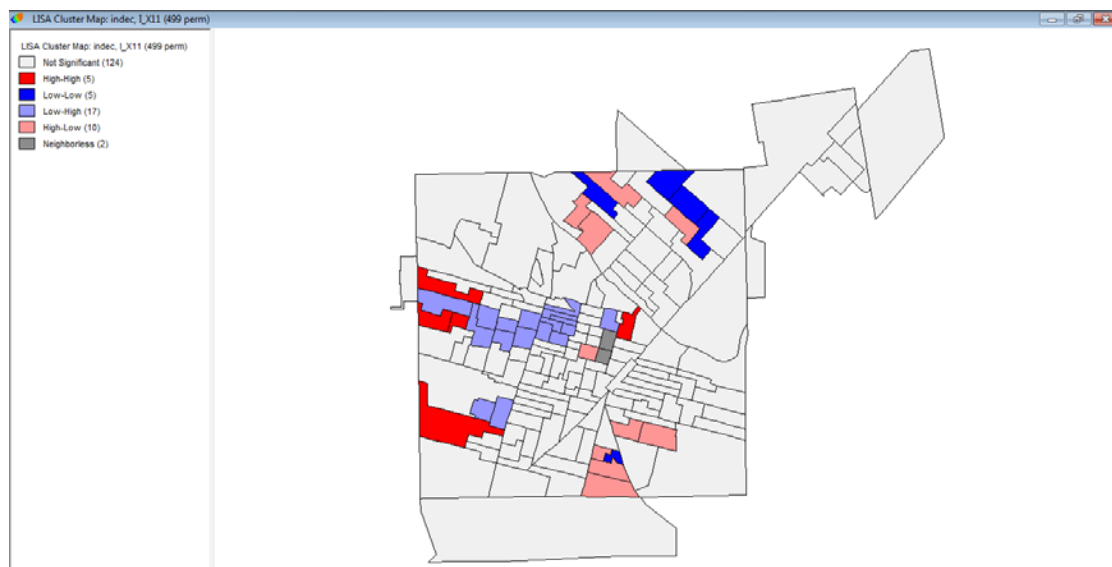
construir (por ejemplo, *Queen contiguity* con *Order of contiguity* igual a 8). Al indicar *Create*, aparece seleccionada por defecto la matriz indec.gal. Al aceptar esta configuración, se observa el Índice Morán univariante y la gráfica de las autorrelaciones espaciales.



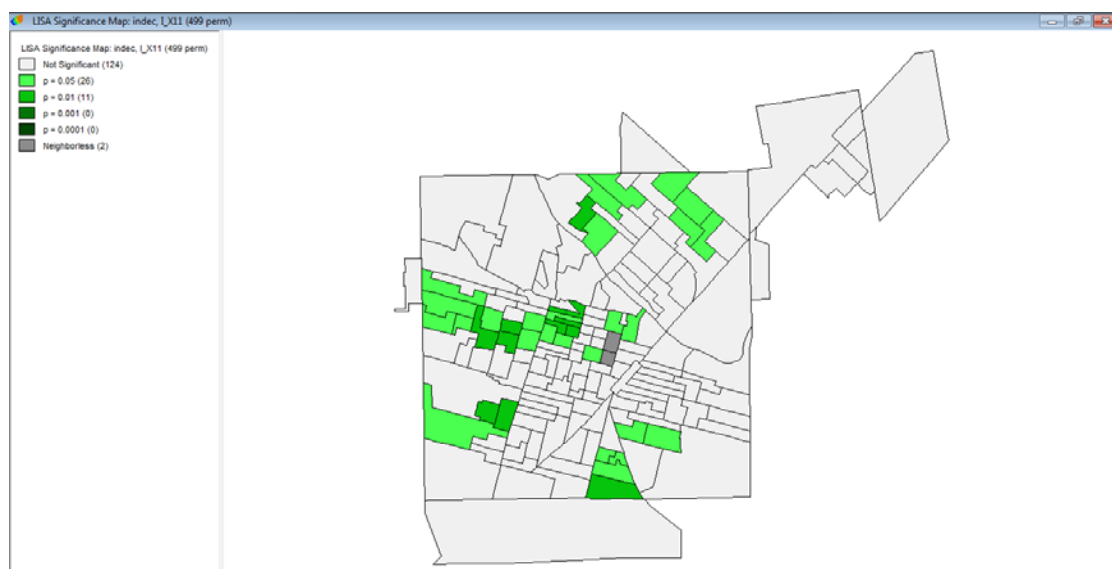
Para conocer el indicador Univariante local de Moran (LISA), la secuencia es *Space – Univariante Local Moran's I (LISA)-selección de variable*, el cuadro de diálogo siguiente permite seleccionar entre los mapas y la gráfica que genera la construcción de este indicador.



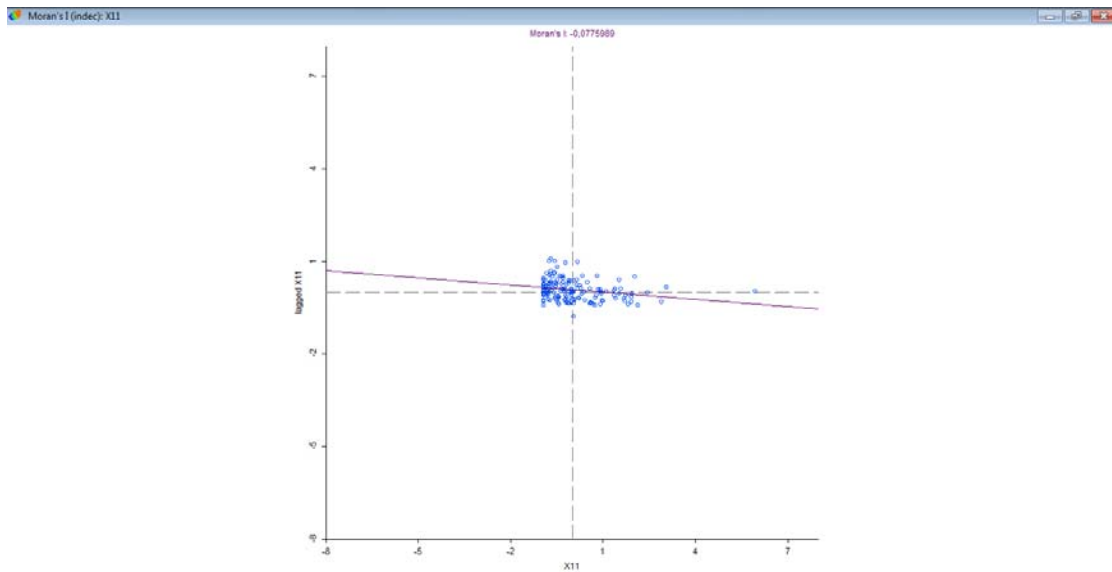
LISA cluster map



LISA Significance Map

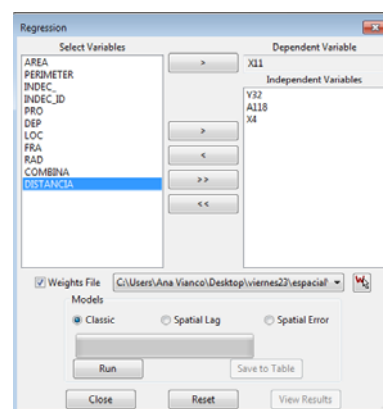
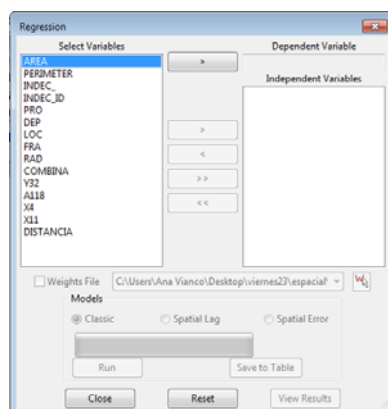
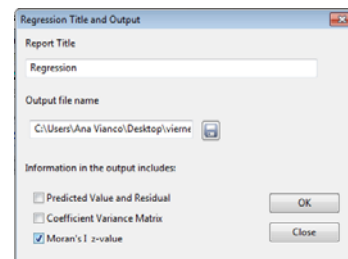
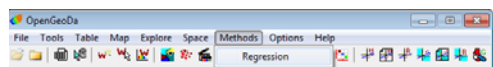


Moran Scatter Plot



## Herramienta regresión

Para hacer regresión se utiliza *Methods-Regresion*, el cuadro de diálogo que se abre permite indicar el título de la salida impresa en pantalla (*Report file*), el archivo donde se quiere guardar el resultado (*Ouput file name*) y la información a incluir (*Information in the ouput includes*). El cuadro de diálogo siguiente permite seleccionar las variables que formarán parte del modelo de regresión; las flechas permiten asignar a las variables un rol en el modelo, se selecciona la matriz de pesos, se opta por el modelo clásico y se ejecuta (*Run*) la configuración que da lugar al resultado (Regression Report).



Regression Report				
Regression SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION				
Data set		indec		
Dependent Variable	X11	Number of Observations:	163	
Mean dependent var	5.70006	Number of Variables	4	
S.D. dependent var	6.01267	Degrees of Freedom	159	
R-squared	0.755528	F-statistic		163.793
Adjusted R-squared	0.750915	Prob(F-statistic)		0
Sum squared residual	1440.63	Log likelihood		-408.882
Sigma-square	9.06056	Akaike info criterion		825.765
S.E. of regression	3.01008	Schwarz criterion		838.14
Sigma-square ML	0.83822			
S.E. of regression ML	2.97291			
Variable	Coefficient	Std Error	t-Statistic	Probability
CONSTANT	1.041295	1.088664	0.9564093	0.3402779
Y22	0.09144167	0.01645413	5.557368	0.0000001
A118	-0.7398503	0.09391159	-7.878157	0.0000000
X4	2.065759	0.1605593	12.86602	0.0000000
REGRESSION DIAGNOSTICS				
MULTICOLLINEARITY CONDITION NUMBER	12.074040			
TEST ON NORMALITY OF ERRORS				
TEST	DF	VALUE	PROB	
Jaarque-Bera	2	44.4106	0.0000000	
DIAGNOSTICS FOR HETEROSKEDASTICITY				
RANDOM COEFFICIENTS				
TEST	DF	VALUE	PROB	
Breusch-Pagan test	3	22.53243	0.0000505	
Koenker-Bassett test	3	11.97509	0.0074662	
SPECIFICATION ROBUST TEST				
TEST	DF	VALUE	PROB	
White	9	42.8777	0.0000023	
DIAGNOSTICS FOR SPATIAL DEPENDENCE				
FOR WEIGHT MATRIX : indec gsl (row-standardized weights)				
TEST	ML-DF	VALUE	PROB	
Moran's I (error)	-0.003986	0.0108800	0.9849368	
Lagrange Multiplier (lag)	1	0.5374507	0.4634912	
Robust LM (lag)	1	0.6170112	0.4321601	
Lagrange Multiplier (error)	1	0.0193900	0.8892518	
Robust LM (error)	1	0.0989514	0.7530917	
Lagrange Multiplier (SARMA)	2	0.6364021	0.7274565	
***** END OF REPORT *****				

## REFERENCIAS BIBLIOGRÁFICAS

ANSELIN, L. (1988): “Lagrange Multiplier test Diagnostic for Spatial Dependence and Spatial Heterogeneity”, *Geographical Analysis*, Vol. 20-1, pp. 1-17.

ANSELIN, L. (1988): *Spatial econometrics: methods and models*, Kluwer Academic Publishers, Boston.

ANSELIN, L. (1995). Local Indicators of Spatial Association — LISA, *Geographical Analysis*, Vol. 27, pp. 93–115.

ANSELIN, L. (1996). *The Moran Scatterplot as an ESDA Tool to Assess Local Instability in Spatial Association*. In M. Fischer, H. Scholten, and D. Unwin (eds.), *Spatial Analytical Perspectives on GIS*. London: Taylor and Francis, pp. 111–125.

ANSELIN, L. (1998): “Exploratory Sapatial Data Analysis in a Geocomputational Environment”. Regional Research Institute and Department of Economics. West Virginia University. Actas de la conferencia de GeoComputación 1998, Bristol (UK).

BERA, A. K. y YOON, M. J. (1992): “*Simple Diagnostic test for spatial dependence*”, University of Illinois, Department of Economics (mimeo). Champaign, IL: Department of Economics, University of Illinois.



- BODSON, P. y PEETERS, D. (1975): "Estimation of the coefficients of a linear regression in the presence of spatial autocorrelation. An application to a Belgian labour-demand function", *Environment and Planning*, pp. 455-472.
- BURRIDGE, P. (1980): "On the Cliff-Ord Test for Spatial Autocorrelation", *Journal of the Royal Statistical Society B*, Vol. 42, pp. 107-108.
- CASE, A., ROSEN, H. y HINES, J. (1993): "Budget spillovers and fiscal policy interdependence: evidence from the states", *Journal of Public Economics*, Vol. 52, pp. 285-307.
- CLIFF, A. y ORD, J. (1981): *Spatial Process. Models and Applications*. Pion. Londres.
- CORRADO, L. y FINGLETON, B. (2011): "Where is the Economics in the Spatial Econometrics", *Discussion Papers in Economics*, N° 11-01, pp. 1-34.
- CRESSIE, N. (1993): *Statistics for Spatial Data* (Revised Version). New York. Wiley.
- HAINING, R. F. (1990): *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge, Cambridge University Press.
- KAPOOR, M.; KELEJIAN, H. H. y PRUCHA, I. R. (2007): Panel Data Models with Spatially Correlated Error Components, en *Journal of Econometrics*, Vol. 140 (1), pp. 97-130.
- MORENO, R. y VAYÁ, E. (2000): "*Técnicas econométricas para el tratamiento de datos espaciales: La econometría espacial*", Edicions Universitat de Barcelona, Barcelona.
- PAELINK, J.; MUR, J. y TRÍVEZ, J. (2004): "Econometría Espacial: más luces que sombras" en *Estudios de Economía Aplicada*, Vol. 22 (3), pp. 383-404.
- PAELINCK, J. y KLAASSEN, L. (1979): "*Spatial Econometrics*". Farnborough, Saxon House.
- POSADA, L. J. (1978): "Fundamentos económicos-espaciales de la teoría de Centros de Desarrollo", *Agricultura y Sociedad*, N° 6, Ministerio de Agricultura y Comercio de España. Pp. 137-180.