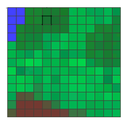
Mapas y cartografía en las ciencias sociales

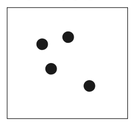
Modelos espaciales

ráster

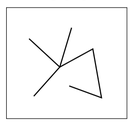


vector

Puntos



Líneas



Polígonos

Shapefiles

archivos

- **.shp** formas
- **.shx** índices
- **.dbf** atributos

} siempre

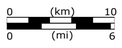
- **.prj** proyección

Composición

layers o capas

Un mapa suele producirse por la superposición de distintos tipos de dato espacial.

Contexto



escala

rosa de los vientos

leyenda

Color (ColorBrewer2.org)

Cualitativa

Datos categóricos. Garantiza la máxima diferencia entre grupos.

Set1
Dark2
Accent
Pastel1

Secuencial

Tipos de datos continuos. Expresa bien diferencias de intensidad.

Reds
Purples
Blues
Greens

Divergente

Datos con un punto medio o de corte claro. Desviación típica, temperatura, etc.

Spectral
RdBu
RdYlGn
PiYG

Programas

ArcGIS
El SIG más completo, pero licencia anual cara y sólo funciona en Windows.

QGIS
La solución SIG abierta, multiplataforma e ideal para la enseñanza.

GeoDa
Foco en geoestadísticas. Mapas sencillos. Ideal para los iniciantes.

R y R-Studio
Integración entre estadísticas, gráficos y SIG. Productividad para investigadores.

Georeferenciación

Shape (shp)

Atributos (.dbf)

cod	nombre
→ 7	Castilla y León

Datos del usuario

cod	nombre	votos
→ 7	Castilla y León	45,0

Proyecciones

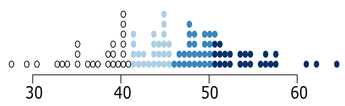
Mercator
Conforme
Mantiene las formas, pero distorsiona las áreas y distancias.

Gall-Peters
Áreas iguales
Mantiene las áreas, pero distorsiona las formas y distancias.

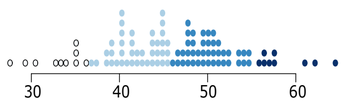
Azimutal
Equidistante
Mantiene las distancias, pero distorciona las áreas y formas.

Clasificación de datos

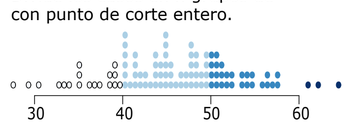
Cuantiles (+ común)
Divide los casos en n grupos con el mismo número de observaciones.



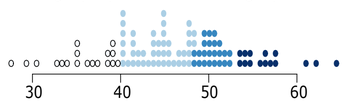
Intervalos iguales
Divide los casos en n grupos de igual valor: 0, 25, 50, 75, 100.



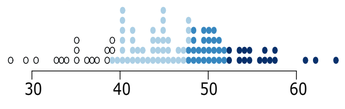
"Pretty"
Divide los casos en n grupos de con punto de corte entero.



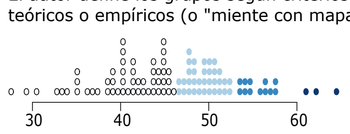
Quiebres naturales (jenks)
N grupos con la mínima diferencia intra-grupo y máxima entre grupos.



Desviación típica
Divide los casos en n grupos de acuerdo con su posición frente a la media.

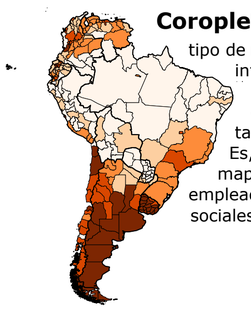


Discrecional
El autor define los grupos según criterios teóricos o empíricos (o "miente con mapas").

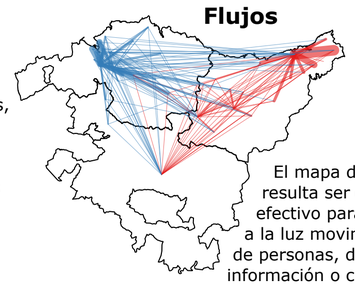


Mapas temáticos

Coropletas
tipo de mapa ideal para informaciones de carácter relativo como porcentajes, tasas e índices. Es, desde lejos, el mapa temático más empleado en las ciencias sociales.

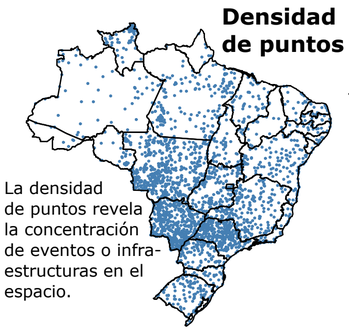


Flujos



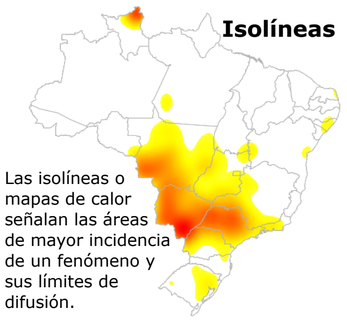
El mapa de flujos resulta ser el más efectivo para sacar a la luz movimientos de personas, dinero, información o cosas.

Densidad de puntos



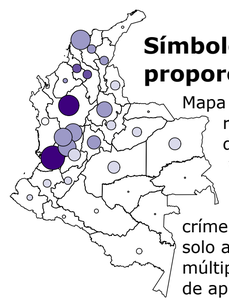
La densidad de puntos revela la concentración de eventos o infra-estructuras en el espacio.

Isolíneas



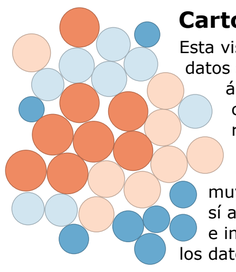
Las isolíneas o mapas de calor señalan las áreas de mayor incidencia de un fenómeno y sus límites de difusión.

Símbolos proporcionales



Mapa excelente para representación de cantidades o valores absolutos. Población, PIB, gasto militar, crímenes. Estas son solo algunas de las múltiples posibilidades de aplicación.

Cartograma



Esta visualización de datos distorsiona las áreas de acuerdo con el valor a ser representado. De ese modo, evita que áreas muy distintas entre sí afecten la lectura e interpretación de los datos.

Operaciones geográficas

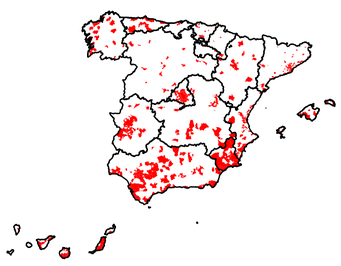
Selección espacial

por atributo

unidades espaciales seleccionadas de acuerdo con un atributo de los datos, como la pertenencia a una misma región, un alto PIB, etc.

por localización

se utiliza para seleccionar atributos espaciales a partir de sua localización relativa a otra unidad espacial. Un ejemplo pueden ser municipios en los que se encuentra un hospital.



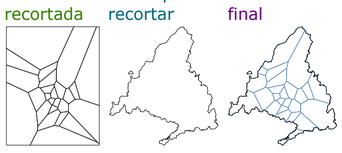
SIG como "cookie cutter"

Algunas operaciones fundamentales de los sistemas de información geográfica consisten en llevar a cabo tareas muy sencillas de agregación de áreas, la unión de diferentes conjuntos de datos, determinar si intersectan entre sí o la diferencia entre ellos.

Intersección

Crea un nuevo objeto espacial a partir de la intersección entre dos ya existentes.

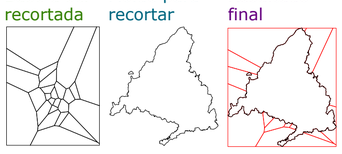
Base a ser recortada | Molde para recortar | Resultado final



Diferencia

Crea un nuevo objeto espacial a partir de la diferencia entre dos ya existentes.

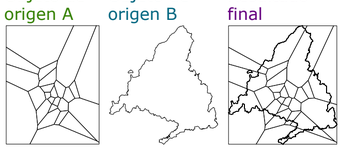
Base a ser recortada | Molde para recortar | Resultado final



Unión (XOR)

Crea un nuevo objeto espacial a partir de la unión entre dos ya existentes.

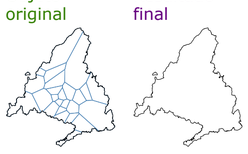
Objeto de origen A | Objeto de origen B | Resultado final



Disolver

Elimina las fronteras internas de un objeto.

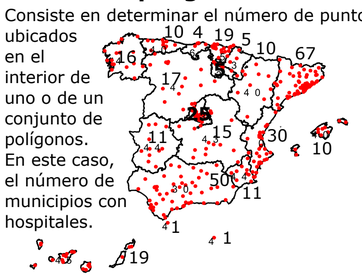
Objeto original | Resultado final



Alcance e influencia

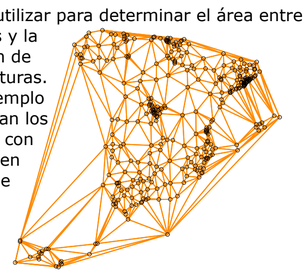
Puntos en polígono

Consiste en determinar el número de puntos ubicados en el interior de uno o de un conjunto de polígonos. En este caso, el número de municipios con hospitales.



Triangulación

Se puede utilizar para determinar el área entre fenómenos y la localización de infraestructuras. En este ejemplo se triangulan los municipios con hospitales en España y se observan muchos espacios vacíos.



Polígonos de Thiessen

También conocidos como diagrama de Voronoi, corresponden al área de influencia o alcance exclusivo de un fenónemo. En el ejemplo abajo utilizamos los municipios de la Comunidad de Madrid que tienen al menos un hospital.

Buffer

Zona alrededor de un objeto espacial definida por una distancia d que permite varios análisis de proximidad, influencia y difusión.

Punto - ¿cuántos niños en edad escolar viven en un radio de 500 metros de una escuela?

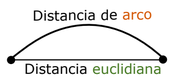
Línea - ¿cuántas personas viven a al menos 100 metros de una línea de metro o autobús?

Polígono - ¿cuál es el nivel de contaminación a una cierta distancia de un basurero?

Cálculos básicos

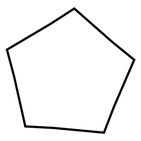
distancia

¿Cuál es la menor distancia entre dos puntos? Algunos contestarían sin dudar: una recta. Un geógrafo podría sugerir un arco dada la curvatura natural de la tierra..



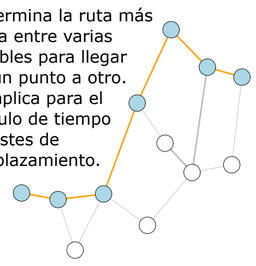
área

Fundamental para el cálculo de algunos indicadores tradicionales como densidad demográfica.



camino más corto

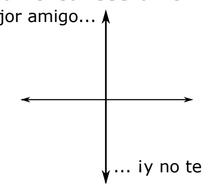
Determina la ruta más corta entre varias posibles para llegar de un punto a otro. Se aplica para el cálculo de tiempo y costes de desplazamiento.



Patrones de puntos

El plano cartesiano

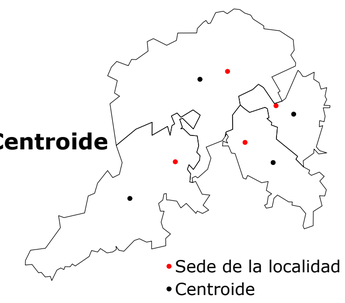
tu mejor amigo...



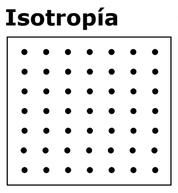
... ¡y no te habías enterado!

La isotropía - hipótesis nula de todo teste estadístico de eventos en el espacio.

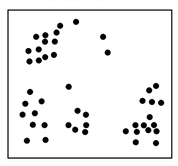
Centroide



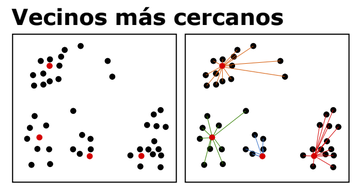
Isotropía



Clusters

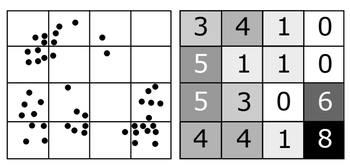


Vecinos más cercanos



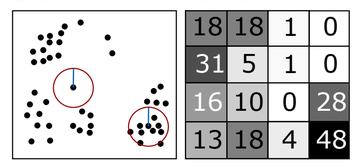
Análisis de cluster

Método cuadrático



3	4	1	0
5	1	1	0
5	3	0	6
4	4	1	8

Densidad de Kernel



18	18	1	0
31	5	1	0
16	10	0	28
13	18	4	48

Geoestadísticas

Distancia entre vecinos

Algoritmos que determinan la probabilidad de asociación espacial (o cluster) a partir de la distancia entre todos los puntos o entre cada punto y su vecino más cercano. Merecen destaque:

- **G** - distancia media entre todos los puntos ubicados a un umbral x de distancia
- **F** - similar a G, pero solo para una muestra aleatoria de los puntos.
- **K** - más exhaustivo, no se limita los vecinos más cercanos, sino todos los puntos.

Kriging

Técnica de interpolación de datos utilizada para estimar la probabilidad de ocurrencia de un evento en una superficie determinada.

Asume que puntos más cercanos tienden a presentar valores más parecidos entre sí (lei de Tobler).

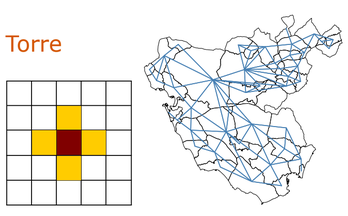
Muy empleado en estudios metereológicos. También puede servir para estimar valores a partir de datos de lugares vecinos.

Matrices de Vecindad

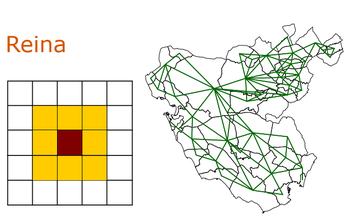
Contigüidad

La vencindad se define por la estructura de contigüidad o fronteras entre unidades de área. Se consideran vecinos todos aquellos que comparten frontera.

Torre



Reina

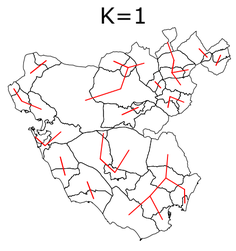


Distancia

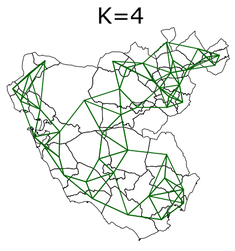
La vinculación se da por la distancia entre unidades espaciales. Ideal para patrones de puntos, pero también aplicable a polígonos o líneas.

K-vecinos más cercanos

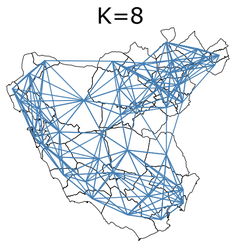
K=1



K=4



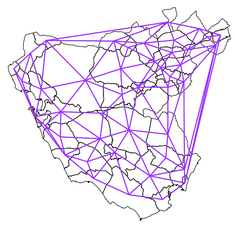
K=8



Conectividad (+ ciencias sociales)

La asociación entre unidades espaciales se basa en su conectividad espacial o en los vínculos de carácter no geográfico existentes entre ellas, especialmente interesantes en las ciencias sociales.

Triangulación

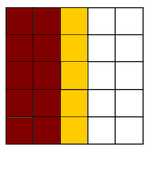


Esfera de influencia

No geográfica

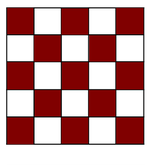
Autocorrelación espacial

La autocorrelación espacial corresponde al modo de medir la dependencia espacial de un fenómeno determinado. Se basa en una idea sencilla: la correspondencia (o disonancia) entre el valor de una variable en un área y los valores correspondientes de sus vecinos.



Positiva

Correspondencia entre el valor de cada unidad de área y el promedio de sus vecinos. Áreas con valores altos vinculadas a vecinos con valores altos.



Negativa

El inverso que lo anterior. Áreas con valores altos conectadas a vecinos con valores bajos y áreas con valores bajos con vecinos con valores altos.

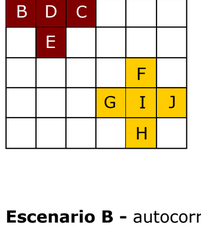
Inexistente

No existe un patrón claro de correspondencia entre el valor de una unidad de área y el valor promedio de sus vecinos.

Variable Spatial-Lag

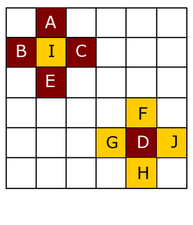
El teste de autocorrelación espacial se basa en una variable spatial lag que consiste en el promedio de los valores de los vecinos de cada unidad de área.

Escenario A - autocorrelación positiva



Id	Var	Vecinos A	Lag A
A	5	D/1	3
B	5	D/1	3
C	2	D/1	3
D	3	(A+B+C+E)/4	4
E	4	D/1	3
F	15	I/1	13
G	18	I/1	13
H	15	I/1	13
I	13	(F+G+H+J)/4	16
J	16	I/1	13

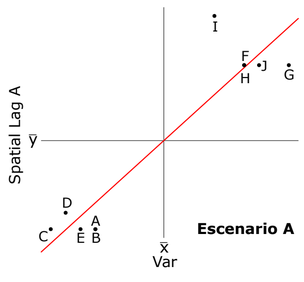
Escenario B - autocorrelación negativa

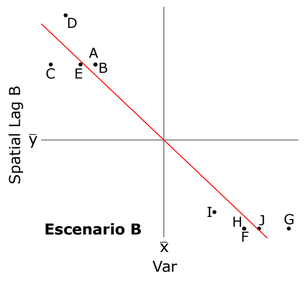


Id	Var	Vecinos B	Lag B
A	5	I/1	13
B	5	I/1	13
C	2	I/1	13
D	3	(F+G+H+J)/4	16
E	4	I/1	13
F	15	D/1	3
G	18	D/1	3
H	15	D/1	3
I	13	(A+B+C+E)/4	4
J	16	D/1	3

El scatterplot de Moran

Corresponde al diagrama de dispersión de los valores estandardizados de la variable de interés y de la spatial lag correspondiente.



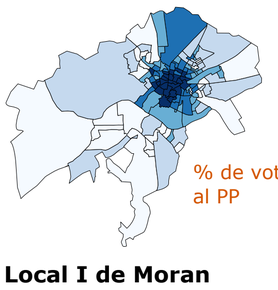


Clusters Espaciales

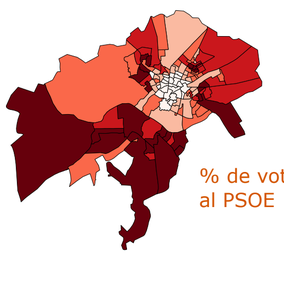
Los clusters espaciales constituyen el resultado del análisis de autocorrelación aplicado a cada unidad espacial que compone una zona o ventana de análisis. El más conocido es el coeficiente I de Moran local, aunque también existen otras opciones como el Gi de Getis y Ord que también revelan patrones subyacentes en los datos.

Patrones Espaciales

Siempre se parte de la misma cuestión: ¿existe algún patrón de aglomeración de datos en el espacio o su distribución es geográficamente aleatoria? El primer paso es verlo en un mapa.



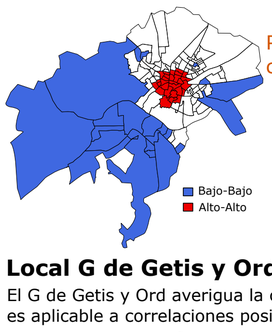
% de voto al PP



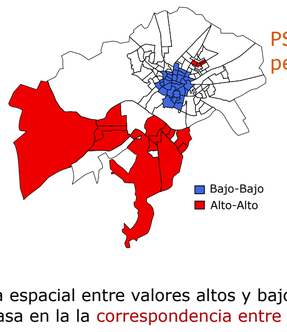
% de voto al PSOE

Local I de Moran

El I de Moran local verifica si los valores de una determinada variable corresponden al promedio de los vecinos de esta variable. Por tanto, se basa en la similitud o diferencia entre valores.



PP - voto céntrico



PSOE - voto periférico

Local G de Getis y Ord

El G de Getis y Ord averigua la coincidencia espacial entre valores altos y bajos. Por esa razón solo es aplicable a correlaciones positivas. Se basa en la la correspondencia entre valores altos y bajos.

Rodrigo Rodrigues-Silveira, Universidad de Salamanca, 2016. rodrodr@gmail.com.