

Análisis de la mortalidad en la ciudad de Tandil utilizando Sistemas de Información Geográfica

Adela Tisnés*

Resumen

Se propone un abordaje espacial cuantitativo de la distribución de la mortalidad según causas en la ciudad de Tandil, con información obtenida a partir de la sistematización de los certificados de defunción de los registros civiles.

El objetivo es presentar una metodología de análisis espacial, a partir de herramientas de análisis ya existentes, para detectar agrupaciones espaciales que luego se delimitarán y se conformarán zonas que surgen, según las causas, a partir de la distribución de los casos de mortalidad.

Se trabajará tratando inicialmente a las defunciones como un fenómeno puntual. Así, la distribución de los casos de defunción por causas, podrá reflejar un patrón de distribución homogéneo, o, en caso contrario, heterogeneidad espacial o agregación. Se trata pues de analizar cómo se distribuye la variable en el espacio y por qué sucede de esa manera.

Inicialmente se utilizó la aproximación basada en el estudio de la función vecino más próximo, con la utilización de la herramienta de densidad Kernel y con un análisis multivariado Isoclusters, para identificar las zonas vulnerables junto con los factores de riesgo asociados.

Palabras clave: mortalidad, análisis espacial, densidad Kernel, análisis multivariado.

* Becaria de CONICET; CIG- FCH- UNICEN; atisnes@yahoo.com.ar

Mortality analysis in the city of Tandil using Geographic Information Systems

Abstract

We propose a quantitative approach to spatial distribution cause-specific mortality in the city of Tandil according to information collected from the systematization of death certificates at civil registries.

This paper aims at presenting a spatial analysis methodology using pre-existing analysis tools to detect spatial clusters that will later delimit and define areas of cause-specific mortality.

First, we will deal with deaths as an isolated phenomenon. Thus, the distribution of cause-specific deaths may reflect a homogeneous distribution pattern or, otherwise, spatial heterogeneity or aggregation. Therefore, it will be analyzed how the variable is distributed in space and why this happens in such a way.

We have used the method based on the K-nearest neighbor algorithm, kernel density estimation tool and Iso clusters multivariate analysis to identify vulnerable areas which will be associated with risk factors.

Keywords: mortality, spatial analysis, kernel density estimation, multivariate analysis.

Introducción

El objetivo inicial de esta investigación es presentar el uso de una metodología que permita detectar agrupaciones espaciales de mortalidad (partiendo de mapas puntuales de mortalidad por causa), que luego se delimitarán y conformarán zonas que surgen, según las causas, a partir de la distribución de los casos de mortalidad.

Se trabajará tratando inicialmente a las defunciones como un fenómeno puntual, ya que se ubicaron los casos de muerte en un mapa de la ciudad de Tandil según las direcciones de cada uno. La distribución de los casos de defunción por causas, podrá reflejar un patrón de distribución homogéneo, o, en caso contrario, un patrón de distribución heterogéneo en el espacio, lo que es igual a decir que existe agregación de la variable analizada. Se trata pues de analizar cómo se distribuye la variable en el espacio y por qué sucede de esa manera.

Se determinará en primer lugar si las entidades (que corresponden a las defunciones por cada una de las causas), exhiben *clustering*, es decir una dispersión estadísticamente significativa en un rango de distancias, a partir de un análisis llamado 'promedio de vecino más cercano'.

Luego, se realizará un análisis de densidad, que permitirá determinar cuáles son las áreas donde se concentra más o menos el conjunto de puntos representado por los casos de defunciones por causa. Puede calcularse de dos maneras: densidad simple (que considera la ubicación espacial, por ejemplo la ubicación de ciudades de un departamento) y densidad por grupos (*kernel*, que considera además de la densidad simple, el peso de alguna variable asociada, por ejemplo la edad promedio de la población, ya que la edad y la mortalidad están directamente relacionadas).

Por último, se creará una clasificación de la superficie de la ciudad a partir de los mapas de densidad de defunciones por causa, aplicando una metodología de clasificación multivariada, cuyo objetivo es asignarle una clase o categoría a cada celda de un área de estudio. De esta manera, se podrán observar en el espacio las áreas donde mayor densidad conjunta de defunciones existe y que se podrían identificar como las áreas con mayor riesgo de muerte. El análisis siguiente deberá identificar cuáles son los determinantes de la salud y cuáles son los factores de riesgo que se están conjugando para favorecer esta situación.

Desarrollo

El análisis de los patrones espaciales de puntos proviene de la geografía de finales de los años 1950 y principios de 1960, cuando el paradigma del análisis espacial comenzó a tomar importancia dentro de la disciplina. Los investigadores tomaron prestado libremente de la literatura de ecología de plantas, adoptando las

técnicas que habían sido usadas allí para la descripción de modelos espaciales y la aplicación de ellos en otros contextos: por ejemplo en los estudios de la distribución de establecimientos (Dacey, 1962; King, 1962), la distribución espacial de tiendas dentro de áreas urbanas, (Rogers, 1965) y la distribución de drumlins en las áreas de glaciares (Trenhaile, 1971). Los métodos que eran utilizados podrían haber sido clasificados dentro de dos grandes tipos (Hagget *et al.*, 1977). La primera fue la técnica basada en la distancia, utilizando información sobre la distribución espacial de los puntos para caracterizar el modelo (básicamente la significatividad del vecino más cercano). Otras técnicas estaban las basadas en el área, relacionadas con varias características, como la frecuencia de la distribución del número observado de puntos en la regularidad definiendo sub-regiones de las áreas de estudio.

El paso del tiempo y el avance y desarrollo de los sistemas de información geográficos han favorecido el interés por analizar las características de las poblaciones teniendo en cuenta el lugar donde viven. Los datos espaciales consisten en mediciones u observaciones realizadas en localizaciones o en áreas específicas. Además del valor de la medición u observación, los datos espaciales incluyen la localización/posición de los valores observados. Las localizaciones pueden ser referidas a puntos o áreas. Por ejemplo, datos referidos a un punto es una medición de un contaminante en suelo en un lugar con determinada latitud y longitud. Otro ejemplo serían las coordenadas (x, y) en metros del domicilio de un sujeto fallecido por una causa de mortalidad determinada. Los datos referidos a un área son los que se utilizan con más frecuencia en epidemiología y para la geografía de la salud, pues se tratan de observaciones realizadas en una región (municipio, sección censal, provincia). En ambas situaciones, la localización espacial puede ser regular o irregular, aunque lo habitual es que sea irregular ya que raramente se trabaja con rejillas de áreas regulares o mediciones ambientales adquiridas uniformemente de forma equidistante unas de otras.

Si se quiere analizar o reflejar la estructura espacial de una población, comunidad, o cualquier fenómeno de naturaleza discreta, la representación cartográfica de todos los elementos del mismo en una región geográfica concreta es la forma más fiel de reflejar la estructura espacial de la población. En muchas ocasiones, dependiendo de la escala de estudio, tales elementos pueden describirse aceptablemente mediante sus coordenadas geográficas (x, y), generándose así un conjunto de datos que recibe el nombre de patrón espacial de puntos (Diggle, 2003). La metodología habitual en el estudio de estas estructuras asume que el patrón espacial de puntos de una comunidad, población, etc., es una realización concreta de un proceso espacial de puntos subyacente que hay que describir y cuyas propiedades son una buena descripción del patrón concreto.

El análisis de un patrón puntual, que refiere al estudio de la disposición de un conjunto de eventos sobre una región del plano, se enmarca en una de las tres grandes ramas de la estadística espacial: aquella que estudia los procesos puntuales (Giraldo Henao, 2011). Específicamente intenta determinar si dichos eventos pre-

sentan un patrón de agregación (es decir, si existen eventos que se producen cerca de otros eventos), de desagregación (los eventos ahora aparecen diseminados en el espacio) o aleatoriedad completa (los eventos se producen con igual probabilidad en cualquier punto del espacio, con independencia de dónde se hallen los otros eventos) (de la Cruz Rot, 2006; López Abente Ortega e Ibáñez Martí, 2001). También es posible que podamos comparar los patrones de dos conjuntos de eventos y, si el patrón es de agregación o desagregación, puede considerarse su modelización como proceso puntual, lo que permite también realizar análisis estadísticos ricos.

Los procesos puntuales, entonces refieren a la localización de eventos. Un ejemplo sencillo de proceso puntual serían las coordenadas de los domicilios de los casos de una enfermedad. En este tipo de datos se trataría de observar y valorar si existe una tendencia de los eventos a exhibir un patrón sistemático. Especialmente alguna forma de regularidad o de agregación. Se trataría de conocer si la intensidad de los eventos varía sobre la región de estudio y posiblemente buscar elementos que nos ayuden a explicar o comprender el fenómeno (Molina, 2008).

Los Sistemas de Información Geográfica (SIG) son sistemas para la recopilación, almacenamiento, integración, análisis y presentación de datos referenciados en el espacio. A esto último se le denomina georreferenciación. Los SIG permiten la rápida representación gráfica de mapas de patologías y de las exposiciones ambientales. Por sí solos los SIG pueden ser considerados como un potente sistema de gestión de bases de datos. La facilidad de su manejo permite generar colecciones de mapas que pueden sugerir nuevas líneas de trabajo. Sin embargo, hay que estar preparados para recibir críticamente las sugerencias que proporciona la observación de los mapas. El hecho de obtener rápida y fácilmente mapas de distribución, debe ir, necesariamente, seguido de una evaluación sistemática de hipótesis de investigación. La acrítica observación de mapas de agregaciones de datos, pueden generar conclusiones o análisis erróneos, debido a que se tiende a ver agregaciones de datos o simplemente patrones geográficos donde no los hay y los mapas son muy propensos a mostrarlos. Los mapas deben ser sometidos sistemáticamente a pruebas estadísticas que traten de transformar la percepción visual en algo objetivo. Los SIG por sí solos son un instrumento francamente útil, pero es cuando los unimos a las técnicas de análisis espacial, cuando muestran todas sus posibilidades. O quizás mejor, las técnicas de análisis espacial tienen un complemento extraordinario cuando se apoyan en un SIG, dado que en la actualidad, la capacidad de análisis espacial de los SIG son muy limitadas, al menos, en el ámbito de la epidemiología.

Una de las características que es posible analizar es la salud de la población, asociada directamente a la gran variedad de factores de riesgos característicos de cada lugar: medioambientales, demográficos, económicos, etc. La distribución espacial de la mortalidad, depende, en cierta manera, de los patrones geográficos que siguen los factores de riesgo.

Siguiendo este enfoque geográfico y basada en la tecnología de uso cada vez más frecuente de los Sistemas de Información Geográfica (SIG), el trabajo se orienta a la generación de una base de datos espacial que permita almacenar de manera organizada, en un mismo lugar, datos tabulares y cartográficos de fuentes diversas para la ciudad de Tandil, aprovechando los avances en las tecnologías de información. La aplicación desarrollada inicialmente recopila los datos considerados fundamentales para obtener una mirada general pero actualizada de las defunciones ocurridas en los años 2003-2004-2005, obtenidas de los registros de defunciones de los registros civiles de la ciudad, incluyendo edad, sexo, causa de muerte, ocupación, etc. Luego, se obtuvieron datos generales de las condiciones sociales del Censo Nacional de Población, Hogares y Vivienda del año 2001 y del medio físico, que se cargaron en el SIG.

Análisis promedio de vecino más cercano

Este análisis calcula un índice sobre la base de la distancia promedio desde cada entidad hasta la entidad vecina más cercana.

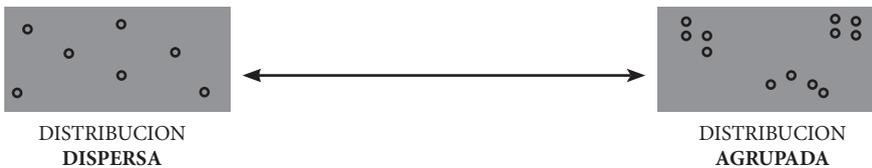


Figura 1. Fuente: Elaborado por Tisnés, 2012.

Se expresa como la relación entre la distancia media observada y la distancia media esperada. La distancia esperada es la distancia promedio que hay entre vecinos en una distribución hipotética aleatoria.

El valor del índice que se obtiene mediante la aplicación de esta herramienta es el resultado de la medición de la distancia entre cada centroide de entidad (punto correspondiente a un fallecimiento) respecto de la ubicación del centroide de su vecino más cercano. Una vez calculado ese valor, lo que se hace es calcular el promedio de todas las distancias de vecinos más próximos. Si esa distancia promedio es más pequeña que el valor promedio de una distribución hipotética aleatoria, entonces, podremos decir que las entidades que se están analizando, se encuentran dispersas (Fig. 1). Esta relación de vecino más cercano promedio, se obtiene calculando la distancia promedio observada por la distancia promedio esperada. Así, si el índice que se obtiene es menor que 1, puede decirse que la distribución posee algún grado de *clustering*. Si en cambio es mayor que 1, la distribución se encuentra dispersa.

El cálculo del índice se hizo utilizando la función *Average nearest neighbor* disponible en el ArcGis 10. Esta herramienta devuelve luego de su aplicación 5

valores: la distancia media observada, la distancia media esperada, el índice del vecino más próximo y las puntuaciones z y p , asociadas al establecimiento de una hipótesis nula, en este caso, relacionada con la distribución espacial de los casos de estudio. Para el análisis de las herramientas de análisis de patrón, es la aleatoriedad espacial completa. Las puntuaciones z y p , entonces, indican si se puede rechazar o no la hipótesis nula¹. Se espera en este caso, rechazar la hipótesis nula, ya que nos estaría indicando un patrón aleatorio. Si rechazamos esa hipótesis, podemos decir que la distribución exhibe *clustering* estadísticamente significativo.

El valor p asociado, corresponde al valor de una probabilidad, que nos orientará acerca de que el patrón espacial observado se haya creado mediante un proceso aleatorio. Si ese valor es muy pequeño, indica que es poco probable (la probabilidad es pequeña) que el patrón observado sea el resultado de procesos aleatorio, por lo tanto podemos rechazar tranquilamente la hipótesis nula. Es interesante analizar cuán pequeño es suficientemente pequeño. Las puntuaciones z son desvíos estándar. Si la aplicación del procedimiento devuelve una puntuación z de +2,5, podría decirse que el resultado son desviaciones estándar de 2,5. Las puntuaciones z como los valores p se corresponden con la siguiente distribución normal estándar (Ebdon, 1985; Mitchell, 2005; Goodchild, 1986):

Las puntuaciones z muy altas o muy bajas (negativas) asociadas con valores p muy pequeños, se encuentran en las colas de la distribución normal. Esto está indicando que es poco probable que el patrón espacial observado refleje el patrón aleatorio teórico representado por su hipótesis nula.

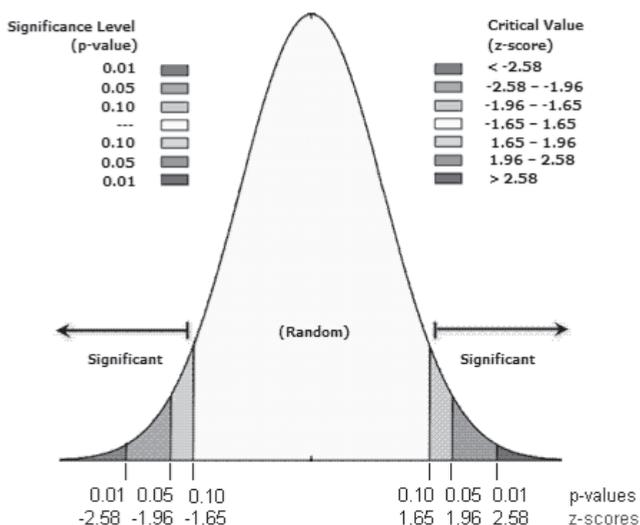


Figura 2. Reporte de Análisis de Vecino más cercano. Fuente: Elaborado por Tisnés (2012) sobre datos de mortalidad Tandil 2003-2005

¹ Los primeros estadísticos (Geary, 1954; Krishna Iyer, 1949; Moran, 1948;1950) probaron la hipótesis nula de no existencia de autocorrelación espacial (esto es de no estructura espacial) frente a hipótesis alternativas no específicas.

Al momento de rechazar la hipótesis nula, debe plantearse una opinión subjetiva respecto del grado de riesgo que se quiere aceptar por estar equivocado (por rechazar la hipótesis nula cuando en realidad no se debía rechazar). Esa decisión, es decir, la elección del nivel de confianza, debe tomarse previa a la ejecución. En este caso se eligió un nivel de confianza del 95 %, que representa una desviación estándar de $-1,96$ y $+1,96$ y un valor p de $0,05$. Si la puntuación z que obtuviéramos se encontrara entre $-1,96$ y $+1,96$ y su valor p fuera mayor que $0,05$, no puede rechazarse la hipótesis nula; el patrón exhibido posiblemente fuera el resultado de procesos espaciales aleatorios. Si la puntuación z obtenida cayera fuera de ese rango (por ejemplo $-3,2$ o $+4,1$), probablemente sería casi imposible que el patrón espacial observado fuera el resultado de la opción aleatoria y el valor p sea pequeño, reflejando eso. Se podrá rechazar la hipótesis nula, para seguir analizando qué puede estar causando ese proceso de clusterización espacial estadísticamente significativa en los datos.

Debido a que el método de vecino más cercano promedio es sensible al valor del área (es decir, pequeños cambios en el valor del parámetro del área pueden ocasionar cambios considerables en los resultados), la herramienta es efectiva cuando se comparan entidades diferentes (en este caso, las diferentes distribuciones de defunciones por causa específica), en un área de estudio fija. Es por eso que para este estudio, se utilizó el área determinada por la suma total de las superficies de cada uno de los 107 radios censales del área urbana de la ciudad.

Como se mencionó anteriormente, en nuestro estudio tratamos las defunciones como proceso puntual. La ubicación de los domicilios de residencia de los casos está estrechamente relacionada con la distribución de la población en el área de estudio.

Es importante identificar los patrones geográficos para describir y comprender cómo se comportan los fenómenos geográficos. Si bien se puede tener una idea del patrón general de las entidades y sus valores asociados al realizar una representación cartográfica de ellos, el cálculo de la estadística cuantifica el patrón. Esto facilita la comparación de patrones para distintas distribuciones o para distintos períodos de tiempo. Generalmente, las herramientas del conjunto de herramientas Análisis de patrones son un punto de inicio para realizar análisis más profundos. Por ejemplo, el uso de la herramienta ‘promedio del vecino más cercano’ para determinar si las distribuciones analizadas promueven el *clustering* espacial o se encuentran aleatoriamente distribuidos.

En cada sociedad existen grupos de individuos, familias o individuos que presentan más posibilidades que otros, de sufrir enfermedades, accidentes, muertes prematuras. Podríamos decir que son individuos o colectivos vulnerables en espacios geográficos determinados. A medida que se incrementan los conocimientos sobre los

diferentes procesos, la evidencia científica demuestra en cada uno de ellos que las enfermedades no se presentan aleatoriamente y que esa “vulnerabilidad” tiene sus razones. La vulnerabilidad se debe a la presencia de cierto número de características de tipo genético, ambiental, biológicas, demográficas, sociales que actuando individualmente o de manera conjunta entre sí desencadenan la presencia de un proceso.

Si se detecta que en una zona determinada de la ciudad un grupo de personas que se sabe asociada con un aumento en la probabilidad de padecer, desarrollar o estar especialmente expuesto a un proceso mórbido (en este caso de padecer), podemos pensar que existen factores (biológicos, ambientales, de comportamiento, socio-culturales, económicos) que, sumándose unos a otros y aumentando el efecto aislado de cada uno de ellos produciendo un fenómeno de interacción, están provocando que en esa zona determinada se observe una cantidad mayor de muertes que en otras zonas. Puede pensarse que la combinación de factores que deberán determinarse (no sólo cuáles son, sino además, en qué medida están contribuyendo al fenómeno) están contribuyendo a que existan zonas con diferencias en la mortalidad.

El primer paso del análisis fue construir mapas donde se pudiera observar cómo se distribuye y cómo varían espacialmente los fenómenos bajo estudio. Para esto, se crearon mapas de distribución de puntos para las defunciones ocurridas en los años seleccionados, por causa de muerte. Luego, se aplicó el método del vecino más cercano, como se mencionó anteriormente.

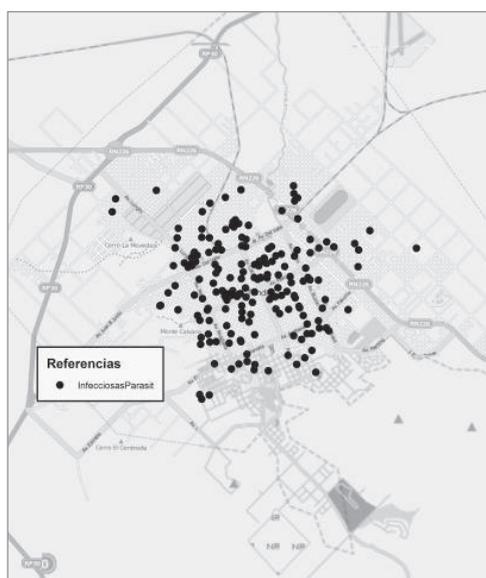


Figura 3. Muerte por enfermedades infecciosas y parasitarias.
Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)

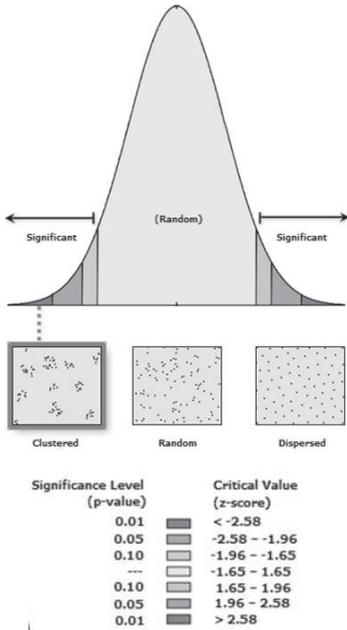


Figura 4. Reporte de análisis de Vecino más cercano. Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)

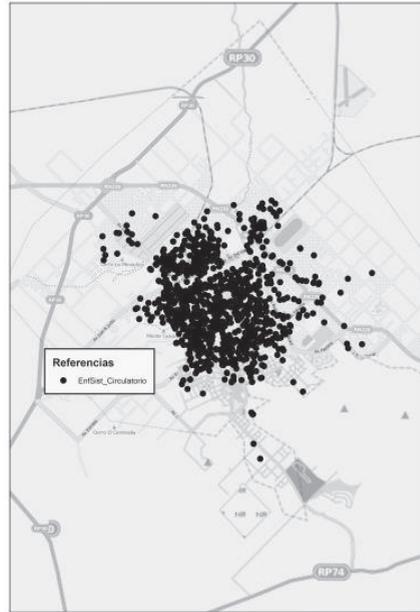


Figura 5. Muerte por enfermedades del sistema circulatorio. Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)

La figura 3 presenta la distribución de defunciones por causas infecciosas y parasitarias en la ciudad. La figura 4 corresponde al análisis del vecino más próximo, aplicado sobre la distribución de las defunciones mencionadas. El análisis del gráfico, indica que la distribución posee algún tipo de *clustering*. Para las enfermedades Infecciosas y Parasitarias, el valor del índice es de 0,438. Nos está indicando entonces que definitivamente existe *clustering* en la distribución. Ahora, si analizamos la validez estadística de esta afirmación y se observa el valor de z, es -14,38 (es menor que -1,96) y el p-valor p que es 0,000 se puede rechazar la hipótesis nula con un 95 % de seguridad.

La figura 5 presenta la distribución de las defunciones por enfermedades del Sistema Circulatorio. La figura 6 por su parte, vuelve a mostrar los resultados obtenidos luego de la aplicación de la herramienta ‘promedio de vecino más cercano’. En este caso, el valor del índice es de 0,3745 mientras que el valor z es -38,88, con lo cual estamos seguros de no caer en el intervalo comprendido por los valores -1,96 y +1,96. Luego, el p valor es 0,00, lo que indica con un 95 % de certeza, que podemos rechazar la hipótesis nula.

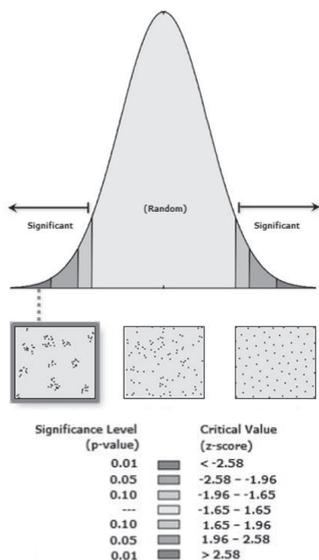


Figura 6. Reporte de análisis de vecino más cercano. Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)

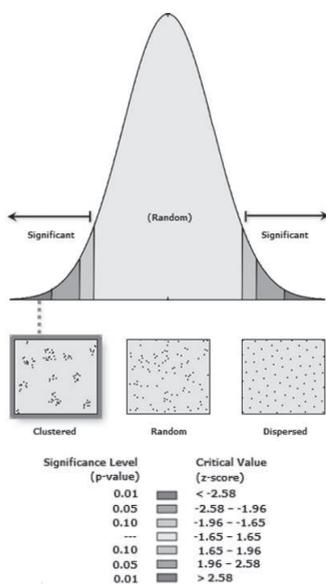


Figura 8. Reporte de análisis de vecino más cercano. Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)

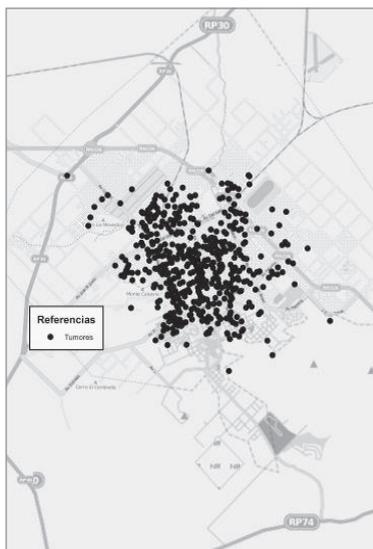


Figura 7. Muerte por tumores. Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)

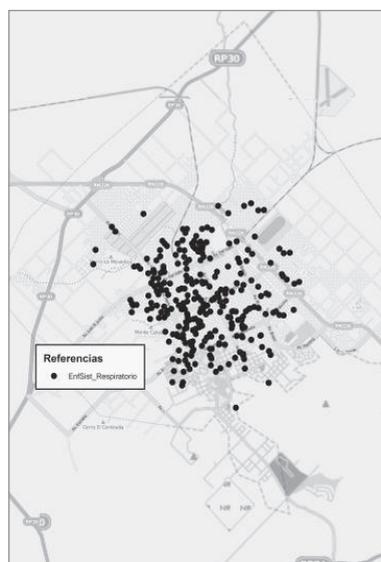


Figura 9. Muertes por enfermedades del sistema respiratorio. Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)

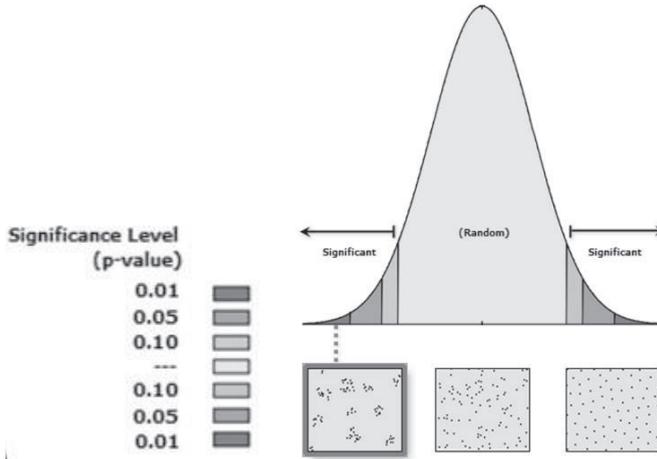


Figura 10. Reporte de análisis de vecino más cercano. Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)

En la figura 7 se puede observar la distribución de las defunciones por Tumores. Luego, la figura 8 muestra los resultados obtenidos para el índice de vecino más cercano. Aquí, obtuvimos un valor del índice de 0,449, mientras que el valor de z es de $-24,274$, con lo cual podemos estar seguros de no caer en el intervalo comprendido entre los valores $-1,96$ y $+1,96$. Luego, el p valor es $0,00$, lo que indica con un 95% de certeza, que podemos rechazar la hipótesis nula.

En la figura 9 se puede observar la distribución de las defunciones por enfermedades del Sistema Respiratorio. Luego, la figura 10 muestra los resultados obtenidos para el índice de vecino más cercano. Aquí, obtuvimos un valor del índice de $0,422$, mientras que el valor de z es de $-18,785$ con lo cual podemos estar seguros de no caer en el intervalo comprendido entre los valores $-1,96$ y $+1,96$. Luego, el p valor es $0,00$ lo que indica con un 95% de certeza, que podemos rechazar la hipótesis nula.

Análisis de Densidad Kernel

Una manera de observar los aspectos de la realidad es examinar cómo se distribuye espacialmente y con qué intensidad eso sucede. El concepto de densidad espacial alude a una relación (cociente) entre el nivel de presencia de un fenómeno en un lugar (fallecimientos por causa) y la superficie de ese lugar. El dividendo de este cociente es inmediato y definido, sin embargo, el divisor, no tiene una obtención tan directa. Esto, debido a que es complicado definir cuál es el ámbito apropiado para ser tomado como base. Debe establecerse el ámbito de referencia de acuerdo con criterios relacionados con la naturaleza del fenómeno.



Figura 11. Enfermedades infecciosas y parasitarias. Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)

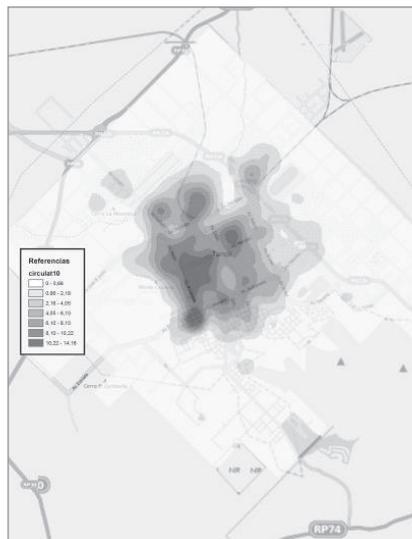


Figura 12. Enfermedades del sistema circulatorio. Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)



Figura 13. Tumores. Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)



Figura 14. Enfermedades del sistema respiratorio. Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)

Los clásicos mapas de densidad, adoptaban circunscripciones poligonales derivadas de divisiones administrativas. Pero, en los mapas que en este trabajo se realizan, debe definirse para cada píxel de la capa ráster un entorno de carácter circular que será el utilizado como base de referencia para calcular la densidad. El centroide de cada píxel se convierte en el centro del círculo, el área que se define a partir del círculo y los puntos que estén comprendidos dentro del mismo se usará para conformar el valor asociado al dividendo. Cada punto puede tener además, un valor o un peso (referido por ejemplo a la cifra de habitantes de ese lugar).

En el caso de la densidad Kernel los puntos del interior de los círculos son ponderados de manera desigual, según su proximidad al centroide del píxel (los más cercanos pesarán más y los más alejados pesarán menos). La densidad Kernel se ajusta a una superficie curva uniforme sobre cada punto. El valor de superficie es más alto en la ubicación del punto y disminuye a medida que aumenta la distancia desde el punto y alcanza valor cero en la distancia máxima especificada, establecido desde el punto. Para calcular la densidad de cada celda raster de salida, se agregan los valores de todas las superficies de kernel en donde se superponen con el centro de la celda ráster. La función kernel se basa en la función kernel cuadrática que describió Silverman (1986, p.76 ecuación 4.5).

Básicamente, el análisis de densidad toma cantidades conocidas de un fenómeno (en este caso los puntos que representan fallecidos por diferentes causas en la ciudad de Tandil) y las expande a través del paisaje basándose en la cantidad que se mide en cada ubicación y la relación espacial de las ubicaciones de las cantidades medidas (Moreno Jimenez, 1998).

Las superficies de densidad muestran donde se concentran las entidades de punto. En este caso, como se mencionaba, se puede tener un valor de punto para cada persona fallecida por alguna causa, pero se desea conocer más sobre la expansión de la población en la región. Si se quiere observar cómo es la distribución espacial de las defunciones por causa, se crea una superficie a partir de esos puntos. La herramienta densidad considerará dónde se encuentra cada punto en relación con el resto de los puntos. Luego, las celdas más cercanas a los puntos, reciben proporciones más altas de la cantidad medida que aquellas que se encuentran más alejadas.

La tabla I presenta los valores máximos, mínimos, el valor medio con su respectivo desvío estándar para cada capa raster de densidad de mortalidad Kernel. Es en el caso del mapa de densidades correspondiente a las defunciones por enfermedades del Sistema Circulatorio en que se pueden observar densidades más elevadas. Esto indica que en las áreas de intensidad más elevada (las de coloración más oscura en los mapas), hay alrededor de 1,5 personas fallecidas por enfermedades del Sistema Circulatorio en un radio de 300 metros.

Layer	STDMEANMAXMIN
Enfermedades infecciosas y parasitarias	0,04610,02040,33810,0000
Enfermedades sistema circulatorio	0,24310,11041,41640,0000
Tumores	0,11850,05540,67650,0000
Enfermedades sistema respiratorio	0,06720,03020,45840,0000

Tabla I. Estadísticas de las capas individuales. Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)

Como el paso siguiente es la aplicación de un análisis multivariado aplicado sobre los mapas de densidad, con la intención de agrupar en *clústers* aquellas regiones donde las densidades sean similares, lo que es necesario hacer en este punto es estandarizar las densidades, para igualar las escalas con las que se está trabajando.

Análisis multivariado

Una vez que se analizó la distribución desigual en el espacio de las defunciones según causa, lo que sigue es utilizar los mapas de densidad para crear un mapa que identifique las zonas donde existe alta densidad para todas las causas simultáneamente y, por el contrario, aquellas zonas de la ciudad, donde existe baja densidad de fallecimientos para todas las causas.

Esto se lleva a cabo a partir de un análisis multivariado de clasificación, cuyo objetivo es asignarle una clase o categoría a cada celda de un área de estudio. En primer lugar, se utiliza un algoritmo de agrupamiento de datos para determinar las características de los grupos de las celdas en un espacio de atributos multidimensional y se almacenan los resultados en un archivo de firma ASCII de salida.

Este procedimiento utiliza un proceso de *clustering* de optimización iterativo modificado (conocido como la técnica de valor medio). El algoritmo separa todas las celdas del número de grupos unimodales distintos especificado por el usuario en el espacio multidimensional de las bandas de entrada. Se utiliza generalmente como preparación para la clasificación no supervisada. El algoritmo de Cluster ISO es un proceso iterativo para calcular la distancia euclidiana mínima cuando se asigna cada celda candidata a un *cluster*. El proceso comienza con la asignación de valores medios arbitrarios por parte del soft, para cada *cluster* (en este caso se designaron 5). Cada celda se asigna lo más cercana a los valores medios. Los nuevos valores medios se vuelven a calcular para cada cluster en base a las distancias de

los atributos de las celdas que pertenecen al cluster después de la primera iteración. Luego, el proceso se repite: cada celda se asigna al valor medio más cercano en el espacio de atributos multidimensional y los nuevos valores medios se vuelven a calcular para cada cluster en base a la pertenencia de las celdas de la iteración. Puede especificar la cantidad de iteraciones del proceso mediante el número de iteraciones. Este valor debe ser lo suficientemente grande como para garantizar que, después de ejecutar el número de iteraciones especificado, la migración de celdas de un *cluster* a otro sea mínima. De esta manera, los *clusters* se volverán estables. Si se aumentara la cantidad de *clusters*, deberían aumentarse también la cantidad de iteraciones.

Una vez corroborada la existencia de una distribución desigual de las enfermedades en el espacio, se busca delimitar mediante la ejecución de la herramienta IsoCluster, aquellas zonas de la ciudad que presentan densidades de mortalidad semejantes. Es decir, se delimitarán regiones que tendrán similares valores de densidad: la magnitud se calcula a partir de la cantidad de defunciones que han ocurrido en una misma unidad de superficie. De esta manera, quedarán, por ejemplo detectadas, aquellas áreas donde existe mayor densidad de defunciones en todas las causas de muerte. Como se indicaba anteriormente, esas áreas de la ciudad, serán en las que vivan grupos de individuos, familias o individuos que presentan más

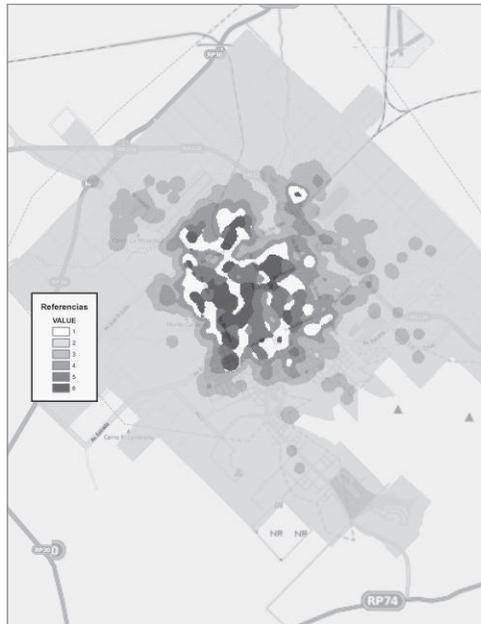


Figura 15. Análisis de IsoCluster. Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)

posibilidades que otros, de sufrir enfermedades, accidentes, muertes prematuras, con algún grado de vulnerabilidad y que esa “vulnerabilidad”, no está siendo una resultante aleatoria, sino que tiene sus razones. Esas razones tienen que ver con factores ambientales, biológicos, demográficos, sociales que actuando individualmente o de manera conjunta entre sí desencadenan la presencia de ese proceso de diferenciación espacial. En este sentido, las áreas aquí encontradas con altos niveles de densidad.

El mapa que resultó de la aplicación de la herramienta IsoCluster (Fig. 15) muestra en color gris oscuro aquellas áreas con mayores niveles promedio de densidad de defunciones en todas las causas de muerte y a medida que los tonos grises se aclaran, descenden los valores promedio de densidad de defunciones.

El siguiente paso fue analizar qué características socioeconómicas y demográficas posee la población que habita cada zona definida según las diferentes densidades de mortalidad. Para esto, se extrajo (de una capa con esa información de población perteneciente al Censo de Población y Viviendas del 2001) la información socio-demográfica referida a la población que habita cada una de las zonas.

Se encuentra que la población aproximada que vive en las áreas de más alta densidad de mortalidad (Fig. 16) es de 4.000 personas. En promedio el 70 % de esas personas poseen cobertura de salud y alrededor del 20 % de estas personas completaron la educación superior o la educación universitaria. Además un 42 % de esta población está ocupada y el 35 % de los hogares que allí viven, no poseen necesidades básicas insatisfechas y alrededor del 30 % no posee privaciones². Abarca una superficie de 2.230.898,51 metros cuadrados. De todas las regiones ésta es la que abarca menos superficie.

En las áreas de densidad media-alta (Fig. 17) viven aproximadamente 15.000 personas, de las cuales el 57 % posee cobertura de salud y el 11 % completó los niveles superior y universitario de educación. Respecto a la población ocupada se observa que asciende al 44 %. Un 26 % de la población no poseen necesidades básicas insatisfechas y el 20 % no posee privaciones. Esta área abarca una superficie total de 2.674.099,34 metros cuadrados.

Asociada a las superficies con densidad media de mortalidad (Fig. 18) hay una población aproximada de 20.000 personas. De ellas, el 65 % posee cobertura de salud, mientras que un 15 % ha terminado sus estudios superiores y universita-

² La incapacidad de los hogares para proveerse de uno u otro tipo de recursos es lo que distingue entre los hogares con privación o sin ella. El primer aspecto, se vincula a la privación patrimonial que afecta a los hogares en forma más estable y dada su característica de persistencia se la considera de tipo estructural o crónico. En cambio, la privación de recursos corrientes puede variar considerablemente en el corto plazo y está ligada más directamente a las fluctuaciones de la economía. Una primer distinción se da entre los hogares con privación y sin privación; estos últimos ostentan condiciones suficientes en ambas dimensiones, mientras que los hogares con al menos insuficiente en alguna dimensión se consideran con privación.



Figura 16. Densidad alta. Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)



Figura 17. Densidad media alta. Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)



Figura 18. Densidad media. Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)



Figura 19. Densidad media baja. Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)

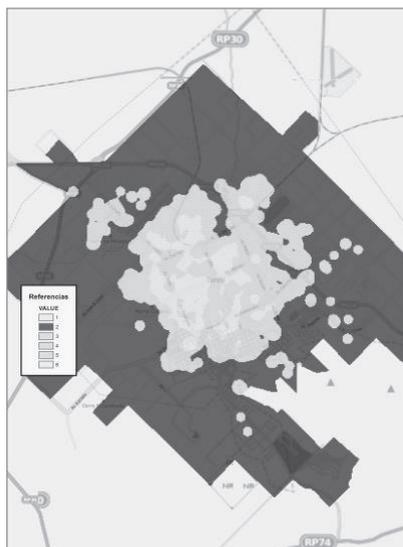


Figura 20. Densidad baja. Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)



Figura 21. Densidad baja baja. Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)

rios. La población ocupada alcanza un 42 %, la que no posee necesidades básicas insatisfechas es casi un 30 % y por último, la población que no posee privaciones llega a ser de un 23 %. Abarca una superficie de 12.310.372,88 metros cuadrados.

En aquellas superficies que se detectó una densidad media- baja de mortalidad (Fig. 19) la población aproximada es de 16.000 personas. El 57 % posee cobertura de salud, mientras que un 11 % ha terminado sus estudios superiores y universitarios. La población ocupada alcanza un 44%, la que no posee necesidades básicas insatisfechas es casi un 26 %. Luego, la población que no posee privaciones asciende a un 19 % del total. Su superficie es de 5.497.630,53 metros cuadrados.

En aquellas superficies que se detectaron con densidad baja de mortalidad (Fig. 20) la población aproximada asociada, fue de 20.150 personas. El 48 % posee cobertura de salud, mientras que solo un 8 % ha terminado sus estudios superiores y universitarios. Respecto de la población ocupada, se observa que alcanza un 42 %. Por otro lado, la población que no posee necesidades básicas insatisfechas es casi un 24 %. Luego, la población que no posee privaciones asciende a un 15 %.

Por último, en las superficies con densidad baja-baja de mortalidad (Fig. 21) se encontró una población total de 14.000 personas. El 66 % posee cobertura de salud mientras que sólo un 17 % ha terminado sus estudios superiores y univer-

sitarios. En cuanto a la población ocupada se observa que alcanza un 43 %. Por otro lado, la población que no posee necesidades básicas insatisfechas es casi un 31 %. Finalmente, vemos que la población sin privaciones asciende a un 25 % y la superficie que abarca es de 3.335.616,27 metros cuadrados.

Un resumen de las principales características mencionadas, se presenta a continuación en la tabla II.

Variables / Densidad	Población total	% Población con cobertura Salud	% Población con Educación Universitaria y Superior completa	% Población Ocupada	% Población Sin NBI	% Población Sin Privaciones	Sup que abarca (m2)
Alta	4000	70	20	42	35	30	2230898
Media Alta	15000	57	11	44	26	20	2674099
Media	20000	65	15	42	30	23	12310372
Media Baja	16000	57	11	44	26	19	5497630
Baja	20150	48	8	42	24	15	3205908
Baja Baja	14000	66	17	43	31	25	3335616

Tabla II. Características poblacionales de las zonas obtenidas.

Fuente: Elaborado por Tisnés (2012) sobre la base de datos provenientes de los Registros Civiles de la ciudad (2003-2005)

Debe decirse que, si bien estamos calculando valores promedio en las variables seleccionadas, debería hacerse un análisis algo más profundo de las características poblacionales, ya que la distribución espacial de las áreas con alta densidad, está abarcando espacios geográficos con diferencias poblacionales notables entre sí. De todas maneras, para un análisis inicial y para una descripción general, esta información será válida.

Conclusiones

El planteo de una zonificación partiendo de estas variables relativas a las defunciones por causa es fundamental para el conocimiento y la posterior planificación del territorio.

El hecho de que exista un desequilibrio espacial, indica que existen lugares que poseen mayor vulnerabilidad que otros respecto de la situación sanitaria. La definición de estas áreas geográficas, son de utilidad a la hora de delinear políticas públicas sanitarias, es decir, planificar. La planificación, pretende actuar *a priori* sobre un proceso que es dinámico, que ocurre de manera constante y cuyo comportamiento y, por lo tanto, el resultado, es impredecible o de difícil predicción. El hecho de planificar, creando zonas en este caso, define y explica los comportamientos que las hacen homogéneas a su interior. La planificación debe llevarse a cabo con el objetivo de ejercer una controlabilidad y una disponibilidad de información específica sobre el espacio que permita particularizar decisiones y acciones orientadas a mitigar

problemáticas asociadas a cada una. Una zonificación permite entonces, reconocer diferencias y actuar en el espacio en función de esas cualidades distintas.

La identificación y acción en relación con ciertos determinantes de la salud implica resolver la manera de abordar el estudio del impacto de esos factores sobre la sociedad, pero también la forma de prevenir o mitigar la acción nociva que representan.

Es necesario desarrollar los instrumentos que permitan efectivamente establecer y medir las diferencias existentes y relacionarlas con los factores que las originan y luego analizar la interacción de los factores entre sí y con las variables de la dinámica de las poblaciones que al conocimiento de la acción específica de cada factor aisladamente. Es decir, metodológicamente, más que universos separados de acciones individuales, se marca la relevancia de la acción intersectorial para el abordaje de la multicausalidad social.

La presente zonificación puede ser el punto de partida para el desarrollo de políticas que atiendan estas diferencias entre sectores más desprotegidos o más amenazados respecto de las variables seleccionadas. Es claro que este informe es un paso inicial de un estudio más complejo y que tiene como objetivo final, explicar estas diferencias existentes.

Bibliografía

BARANGER, Denis. Construcción y análisis de datos. Posadas, UNAM, 1992. 142 pp.

BERENSON, Mark, LEVINE, David. Estadística Básica en administración. Conceptos y aplicaciones. México. Prentice Hall Hispanoamericana S.A. 6° Ed. 1996. 944 pp.

BOSQUE, Joaquín. Sistemas de información Geográfica. Madrid, RIALP, 1992. 504 pp.

BUZAI, Gustavo. La exploración geodigital. Buenos Aires, Lugar Editorial, 2000. 180 pp.

BUZAI Gustavo. Mapas Sociales Urbanos. Buenos Aires, Lugar Editorial, 2003. 384 pp.

BUZAI, Gustavo, BAXENDALE, Claudia. Análisis socioespacial con Sistemas de Información Geográfica. Buenos Aires: Lugar Editorial. 2006. 397 pp.

DACEY, M. Analysis of Map Distributions by Nearest Neighbor Methods, Department of Geography, University of Washintong (1958) unpublished discussion paper No 1; B,J,L Berry, statistical Test of value in Grouping Geographic Phenomenon, Annual Meeting of the Association of American Geographers, Pittsburg 1959; P,W Porter, Farnest and rhe Orephagains- a Fable for the Instruction of Young Geographers, Annals of the Association Of American Geographers vol. 5 (1960) p. 297-299; L.J King, A quantitative expression of the Pattern of Urban Settlements in Selected Areas Of the United States, Tijdschrift voor Economische en Sociale Geografe vol 53 (1962), p.1-7.

de la CRUZ ROT, M. Introducción al análisis de datos mapeados o algunas de las (muchas) cosas que puedo hacer si tengo coordenadas [documento en pdf] Asociación Española de Ecología Terrestre, Ecosistemas 2006. Vol. 15, n° 3, 19-39 pp. Disponible en: <http://www.revistaecosistemas.net/articulo.asp?Id=448> (consulta agosto 2012).

DIGGLE, Peter. Statistical Analysis of Spatial Point Patterns. London: Arnold. Segunda edición. 2003. 159 pp.

EBDON, David. Estadística para geógrafos. España, Madrid: Oikos- Tau, s.a Ediciones, 1982. 352 pp.

ESTEBANEZ, José, BRADSHAW, Roy. Técnicas de cuantificación en Geografía. Madrid, Tebar, 1978. 318 pp.

GARCÍA RAMON, María Dolores. Teoría y Método en la Geografía Humana Anglosajona. Ariel Geografía, Barcelona, 1985. 272 pp.

GIRALDO HENAO, Ramón. Estadística Espacial. [documento en pdf] Departamento de Estadística, diciembre 2001, Bogotá, Colombia. Universidad Nacional de Colombia. Disponible en <http://bibliotecadigital.uns.edu.ar/revistas/rug/einstruc.htm> (consulta julio 2012)

GOODCHILD, Michael Frank., Spatial Autocorrelation. Concepts and Techniques in Modern Geography (CAMTOG). Norwich N° 47, Geobooks, 1986. 96 pp.

GOODCHILD, Michael, HAINING, Robert. Sig y análisis especial de datos: perspectivas convergentes. *Investigaciones Regionales*, España. Asociación Española de Ciencia Regional. 2005 n° 6 pp. 175-201.

GRUPO CHADULE, Ariel. Iniciación a los Métodos Estadísticos en Geografía. España. Ed. Ariel. 1980. 296 pp.

HAGGET, Peter, CLIFF Andrew, ORD, John Keith. *Spatial Aspects of Influenza Epidemics*, London Routledge, 1986. 280 pp.

HARVEY, David. *Teorías, leyes y modelos en Geografía*, Alianza Editorial, 1969.

LOPEZ ABENTE ORTEGA, Gonzalo, IBAÑEZ MARTÍ, Consuelo. *Aplicación de técnicas de análisis espacial a la mortalidad por cáncer en Madrid* [documento en pdf] Documentos Técnicos de Salud Pública, 2001 Madrid, España. Disponible en http://bvs.isciii.es/mono/pdf/CNE_03.pdf (consulta en julio 2012)

MITCHELL, Andy. *The ESRI guide to GIS analysis, Vol. 2: spatial Measurements and Statistics*. Esri Press, 2005. 252 pp.

MOLINA, Adriana. *Sistemas de información geográfica para el análisis de la distribución espacial de la malaria en Colombia*. *Revista EIA*. Medellín, Colombia, Escuela de Ingeniería Julio 2008, N° 9, pp. 91-111.

MORENO JIMÉNEZ, Antonio. *Sistemas y Análisis de la Información Geográfica. Manual de autoaprendizaje con AcrGIS*. Madrid: Ra-Ma, 2005. 895 pp.

PEÑA, Daniel, ROMO, Juan. *Introducción a la Estadística para las Ciencias Sociales*. España: McGraw-Hill, 1998.

PICKENHAYN, Jorge A. (Compilador) *Salud y enfermedad en Geografía*. Buenos Aires. Lugar Editorial, 2009. 156 pp.

SILVERMAN, Bernard. *Density estimation for statistics and data analysis*, Londres, Chapman and Hall, 1986. 175 pp.

TRENHAILE, Alan. *Drumlins: their distribution, orientation, and morphology* *Canadian Geographer / Le Géographe canadien*. Canada. Canadian Association of Geographers, June 1971. Vol. 15, Issue 2, pp. 113–126.

Fecha de recepción: 23 de noviembre de 2012

Fecha de aceptación: 26 de abril de 2013