

ÍNDICE

I. La información geográfica. Bases de datos	19
1. La matriz geográfica o punto de partida de la estructuración de la información espacial	20
2. Las unidades espaciales de recogida de la información. Los modelos y estructuras de datos espaciales.	22
2.1. La estructura espacial del modelo raster	23
2.2. La estructura espacial del modelo vectorial	25
Lista de coordenadas.....	26
Diccionario de vértices.....	28
La organización arco/nodo	28
Estructura TIN	30
3. La componente temática. Las bases de datos relativas a los atributos	32
3.1. Las escalas de medida	33
3.2. Las bases de datos de atributos	35
3.2.1. Las estructuras raster.....	35
Enumeración exhaustiva	36
Codificación <i>run length</i>	36
Estructuras jerárquicas.....	38
3.2.2. Las bases de datos del modelo vectorial. El modelo de datos relacional	39
La estructura de las bases de datos relacionales. Restricciones básicas en su elaboración.....	40

La manipulación de las bases de datos relacionales .	44
▶ Juntar (<i>Join</i>).....	44
▶ Selección (<i>Select</i>)	45
▶ Proyección (<i>Project</i>).....	46
▶ Unión (<i>Union</i>).....	46
▶ Intersección (<i>Intersection</i>)	47
▶ Diferencia (<i>Diference</i>).....	48
II. Los métodos y técnicas cuantitativas en el análisis geográfico. El desarrollo informático de los mismos.....	49
1. El desarrollo de los métodos y técnicas cuantitativas en el campo geográfico	49
2. Principales técnicas de la estadística clásica, en la resolución de problemas geográficos.....	52
2.1. El análisis exploratorio de datos (EDA).....	53
2.2. Descripción estadística de una variable.....	57
2.3. La medición por agrupación y asociación de fenómenos....	60
2.3.1. Relación entre variables cualitativas.....	61
2.3.2. Relación entre variables cuantitativas	64
2.4. El modelo de regresión lineal simple	65
2.5. Estadística inferencial. Estimación de parámetros poblacionales a partir de muestras.....	68
2.5.1. Técnicas de muestreo.....	68
2.5.2. Estimación de parámetros poblacionales.....	70
2.5.3. Tests de significación estadística.....	72
2.6. El muestreo como técnica de investigación social	75
2.7. Un método de análisis multivariante: el análisis factorial..	78
2.7.1. Introducción	78
2.7.2. Fases de análisis	80
2.7.3. Fundamento matemático.....	82
3. Técnicas estadísticas de análisis espacial	85
3.1. Análisis espacial de un mapa de puntos.....	87
3.1.1. Visualización de un mapa de puntos	87
3.1.2. Descripción de un mapa de puntos	87
3.1.3. Modelización de un mapa de puntos	91

3.2. Análisis espacial de un mapa de áreas	95
3.2.1. Visualización de un mapa de áreas	96
3.2.2. Descripción de un mapa de áreas.....	96
3.2.3. Modelización de un mapa de áreas	98
4. Los modelos espaciales y el análisis geográfico.....	99
5. La informática y su importancia en el desarrollo de las técnicas y métodos cuantitativos en el análisis geográfico	102
5.1. Principales tipos de análisis geografico obtenidos por orde- nador	103
5.2. Los Sistemas de Información Geográfica (SIG), herramien- ta informática imprescindible en el análisis espacial	107
5.2.1. Las funciones analíticas de un SIG	108
5.2.2. La versatilidad y limitaciones de los SIG en el trata- miento y gestión de la problemática espacial.....	112
III. Algunos programas para el tratamiento informático de la información espacial	117
1. Los sistemas de información geográfica vectoriales. El programa ArcView	117
1.1. Entrada de datos en un SIG vectorial	117
1.1.1. Los procedimientos de captación y conversión a for- mato digital de la información espacial.....	117
1.1.2. La entrada de la información temática	121
1.2. Funcionalidades básicas de un SIG vectorial	122
1.2.1. La representación de los datos geográficos	122
Características de los mapas temáticos	123
Elementos gráficos de la representación cartográfica. El simbolismo cartográfico.....	124
Tipos de mapas. Los mapas temáticos.....	125
1.2.2. La selección/recuperación de la información geográ- fica	131
1.2.3. Superposición de mapas	135
1.2.4. Análisis espacial de proximidad, contigüidad y carac- terísticas geométricas.....	137
1.2.5. La agregación de objetos geográficos	139

1.3. Funcionamiento del programa ArcView 3.2	141
1.3.1. La organización de los datos en el programa Arc-View 3.2	141
1.3.2. La interface de ArcView 3.2	142
1.3.3. La visualización de la base espacial	144
1.3.4. El manejo de las bases de datos de los atributos.....	146
Añadir o borrar un campo o variable de nueva creación.....	148
Editar una tabla dBase, correspondiente a un tema concreto	149
Juntar una tabla dBase a una tabla de atributos ya existente	150
Hacer un resumen de los datos, si los registros poseen un campo cuyo valor se repite.....	150
1.3.5. Representación cartográfica de mapas temáticos.....	151
Preparación de los mapas temáticos.....	153
▶ Acceso al cuadro de definición del mapa.....	153
▶ Tipos de mapas definidos	154
▶ Métodos de clasificación para la determinación de intervalos.....	157
▶ Fases de definición de cada mapa temático	157
Añadir textos, gráficos o etiquetas a un mapa determinado	157
▶ Añadir texto o gráficos	158
▶ Añadir etiquetas.....	159
Etiquetado individual.....	159
Etiquetado conjunto.....	160
1.3.6. Representación de gráficos.....	161
Creación de un gráfico	162
Tipos de gráficos.....	162
1.3.7. Salidas cartográficas para impresión.....	165
1.3.8. La gestión de la información geográfica.....	169
Identificar los elementos geográficos con el ratón....	169
Conocer los atributos de elementos geográficos seleccionados	169
Encontrar los elementos geográficos que cumplan determinadas condiciones de sus atributos.....	170

▶ Localización por su nombre	170
▶ Localización por condición algebraica	170
▶ Localización de elementos geográficos usando un gráfico	171
Encontrar los elementos espaciales que cumplan determinadas condiciones de carácter geográfico	172
▶ Localización de los elementos geográficos a una distancia dada de un punto concreto	172
▶ Localización de los elementos geográficos cercanos a otros datos.....	172
▶ Localizar de los elementos geográficos existentes en el interior de un polígono	174
▶ Localización de los elementos geográficos de un tema que interseccionan con los de otro tema	175
Operaciones a realizar con los elementos geográficos seleccionados	176
▶ Localizar los elementos que no cumplan determinadas condiciones	176
▶ Cálculo estadístico con los registros o unidades espaciales seleccionados.....	177
▶ Construir un nuevo tema de los elementos seleccionados	179
▶ Otras operaciones a realizar	179
Agregación de la información.....	179
▶ Agregación de los datos de atributos, cuyos valores coinciden en un determinado campo o variable ...	179
▶ Agregación de unidades espaciales en otras de orden superior	179
1.3.9. Edición de datos espaciales	180
Creación de un tema de puntos	180
▶ Edición de un layer de puntos	180
▶ Edición de los atributos de los nuevos elementos puntuales construidos	181
▶ Modificar las características de los elementos puntuales de una capa ya editada.....	181
Creación de un tema de líneas.....	182
▶ Edición de un layer de líneas.....	182

▶ Edición de líneas con puntos comunes a otras	182
▶ Cortar o unir líneas	184
▶ Edición de atributos de los nuevos elementos lineales construidos	184
▶ Modificación de las características de los elementos lineales de una capa ya editada	186
Creación de un tema de polígonos	186
▶ Edición del layer de polígonos.....	186
▶ Edición de atributos de los nuevos elementos poligonales construidos.....	188
▶ Modificación de las características de los elementos poligonales de una capa ya editada.....	189
1.4. Funcionamiento de ArcView (versión 8.1 NT)	189
1.4.1. La nueva versión de ArcView.....	189
1.4.2. La organización de los datos en ArcMap	190
1.4.3. La interface de ArcMap.....	192
1.4.4. La visualización de la base espacial	194
1.4.5. El manejo de las bases de datos de los atributos.....	197
Añadir o borrar un campo o variable de nueva creación. Edición de la información	198
Juntar una tabla de atributos a otra ya existente	200
Hacer un resumen de los datos, si los registros poseen un campo cuyo valor se repite.....	201
1.4.6. Representación cartográfica de mapas temáticos.....	202
Preparación de los mapas temáticos.....	203
Definición del mapa	203
Tipos de mapas definidos	203
▶ Mapas de representación individualizada de símbolos	203
▶ Mapas de representación de variables categóricas	204
▶ Mapas de representación de variables cuantitativas	205
▶ Mapas de representación multivariable.....	207
▶ La utilización de plantillas de mapas (“map template”)	208
1.4.7. Añadir textos, gráficos o etiquetas a un mapa determinado	209

Añadir texto o gráficos	209
Añadir etiquetas.....	210
1.4.8. Salidas cartográficas para impresión.....	213
1.4.9. La gestión de la información geográfica.....	216
Identificación de los elementos geográficos con el ratón	216
Seleccionar elementos geográficos y sus atributos de forma directa	216
Encontrar los elementos geográficos que cumplan determinadas condiciones de sus atributos.....	217
▶ Localización por su nombre	217
▶ Localización por condición algebraica	218
Encontrar los elementos espaciales que cumplan de- terminadas condiciones de carácter geográfico	219
▶ Localización de los elementos geográficos en el interior de una caja rectangular determinada.....	219
▶ Localización de los elementos geográficos a una distancia de un punto concreto	220
▶ Localización de los elementos geográficos de un layer que interseccionan con los de otro.....	221
▶ Construcción de capas de proximidad o buffers a una cierta distancia de los elementos de un layer.....	222
Operaciones a realizar con los elementos geográficos seleccionados.....	223
▶ Cálculo estadístico con los registros o unidades espaciales seleccionados	223
▶ Construir un nuevo layer de los elementos selec- cionados	224
Agregación de la información.....	224
2. Los paquetes de programas estadísticos. El programa Statgra- phics	226
2.1. La estructura general de los paquetes de programas estadís- ticos	226
2.1.1. Definición de la información y su manipulación an- tes del cálculo	228

2.1.2.	Aplicación de un procedimiento de cálculo estadístico.....	230
2.1.3.	Análisis de los resultados	231
2.2.	Funcionamiento del programa Statgraphics	232
2.2.1.	Organización de los datos y resultados obtenidos	232
2.2.2.	La interface de Statgraphics	234
2.2.3.	El manejo y edición de los ficheros de datos de partida.....	235
	Edición de un fichero de datos en el sistema Statgraphics, en formato .sf3	235
	Importación de ficheros de datos.....	238
2.2.4.	Descripción de la ejecución de algunos de los procedimientos de cálculo estadístico más usuales	239
	Descripción estadística de una variable.....	239
	Relación bivariada entre variables	247
	▶ Relación bivariada entre variables cuantitativas...	248
	▶ Relación bivariada entre variables cualitativas.	254
	Análisis factorial.....	258
2.2.5.	Adaptación de los resultados gráficos a las necesidades o pretensiones del usuario.....	262
	La operación de salvar los gráficos	262
	Las operaciones de modificar un gráfico estandar ...	263
	▶ Definir el estilo y color de las tramas.....	263
	▶ Definir el estilo y color de las líneas.....	264
	▶ Definir el estilo, tamaño y color de los puntos	265
	▶ Definir la dirección, el estilo y color de la rejilla o drícula	265
	▶ Definir el estilo de las marcas de separación que cuantifican los ejes X e Y	266
	▶ Definir el título de la leyenda y los nombres de las variables que la integran	268
	▶ Definir el título y la escala de los ejes X e Y	268
3.	Las hojas de cálculo. El programa Excel	270
3.1.	La estructura general de los programas «hoja de cálculo» ..	271
3.1.1.	La organización de la información en un programa «hoja de cálculo»	271

3.1.2. Edición de la información en una hoja de cálculo....	272
3.1.3. Preparación personal de una hoja de cálculo.....	274
3.1.4. La elaboración de gráficos, mediante una hoja de cálculo.....	275
3.1.5. La gestión de la información en una hoja de cálculo	276
3.2. Funcionamiento del programa Excel.....	277
3.2.1. Organización de los datos de partida.....	277
3.2.2. La interface de Excel.....	278
3.2.3. Apertura y manejo de los ficheros Excel.....	279
3.2.4. Gestión de ventanas y libros de trabajo.....	279
3.2.5. La entrada y modificación de los datos en una hoja de cálculo.....	281
3.2.6. Dar formato a los datos para lograr la apariencia de- seada.....	284
3.2.7. La creación de gráficos en un libro de trabajo.....	285
3.2.8. Gestión de los datos de una hoja de cálculo.....	290
Pautas para la edición y organización de las bases de datos.....	290
Clasificación de la base de datos.....	292
Resumen de los datos de una lista.....	294
Creación de una tabla dinámica.....	295
IV. Ejemplos de la utilización integrada de los programas de tratamiento de la información geográfica.....	299
1. Análisis del proceso de distribución espacial de la población en las periferias metropolitanas españolas (1960-1996).....	299
1.1. Planteamiento del problema.....	299
1.2. Preparación y tratamiento de la información.....	300
1.2.1. Organización de las bases de datos espaciales.....	301
1.2.2. Incorporación de la información temática.....	305
1.2.3. Gestión de la información geográfica y temática, pre- parada con anterioridad.....	306
Realización de los mapas de distribución espacial de la población.....	306
▶ Mapa temático de 1960.....	306
▶ Mapas temáticos de 1970, 1981, 1991 y 1996.....	308

▶ Integración de las diversas imágenes en una salida común	308
La evolución del volumen de población en las coronas próximas a cada ciudad central.....	314
Relación entre la distribución espacial de la población y las principales vías de comunicación por carretera.....	319
1.3. Interpretación de los resultados	322
2. Estudio de los movimientos natural y migratorio en el Área Metropolitana de Madrid	325
2.1. Planteamiento del problema.....	325
2.2. Preparación y tratamiento de la información.....	326
2.2.1. Organización de las bases de datos espaciales	326
2.2.2. Incorporación de la información temática	328
2.2.3. Gestión de la información geográfica y temática, preparada con anterioridad.....	329
Obtención de las tasas de dinámica poblacional.....	329
Realización de los mapas de distribución espacial de las tasas de la dinámica poblacional	332
2.3. Interpretación de los resultados	336
3. Estudio de los desequilibrios del mercado laboral juvenil, a partir de datos obtenidos por encuesta, con personas jóvenes de 16 a 29 años, relativo a características socioprofesionales, laborales y de residencia geográfica	338
3.1. Planteamiento del problema.....	338
3.2. Preparación y tratamiento de la información.....	340
3.2.1. Organización de las bases de datos	340
3.2.2. Tratamiento estadístico de la información.....	343
Relación entre las variables del mercado laboral (sexo, edad, salario, actividad y nivel de estudios)....	343
Elaboración de una nueva base de datos, agregando los datos individuales en los distritos de residencia..	348
Representación cartográfica de las variables derivadas (salario2, salario3, estudio2 y estudio3).....	349
Cálculo de la correlación entre variables	353
3.3. Interpretación de los resultados	355

4. Estudio de la diferenciación regional del espacio peninsular español, a partir de la información relativa a determinados índices indicativos del desarrollo económico provincial	357
4.1. Planteamiento del problema.....	357
4.2. Preparación y tratamiento de la información.....	358
4.2.1. Organización de las bases de datos	358
4.2.2. Tratamiento estadístico de la información.....	359
4.2.3. Representación cartográfica de los resultados obtenidos.....	360
4.3. Interpretación de los resultados	364
5. Estudio de la diferenciación social de la periferia metropolitana madrileña	365
5.1. Planteamiento del problema.....	365
5.2. Preparación y tratamiento de la información.....	366
5.2.1. Organización de las bases de datos	366
5.2.2. Tratamiento estadístico de la información.....	368
Representación cartográfica de los resultados obtenidos.....	370
5.3. Interpretación de los resultados	373
V. Bibliografía.....	375

II

LOS MÉTODOS Y TÉCNICAS CUANTITATIVAS EN EL ANÁLISIS GEOGRÁFICO. EL DESARROLLO INFORMÁTICO DE LOS MISMOS

1. EL DESARROLLO DE LOS MÉTODOS Y TÉCNICAS CUANTITATIVAS EN EL CAMPO GEOGRÁFICO

La Geografía Cuantitativa es una rama de la Geografía preocupada por el desarrollo de una serie de actividades específicas: el análisis numérico de los datos espaciales; la elaboración de la teoría espacial; y la construcción de modelos matemáticos que reconstruyan los procesos naturales y sociales en el espacio (Fotheringham, A. S. y otros, 2000). Esta definición implica, como puede deducirse de lo anterior, tanto actividades de carácter más teórico, como la elaboración de modelos explicativos de la distribución de fenómenos físicos y humanos en el espacio, como la inferencia empírica de regularidades espaciales a partir de las realidades observadas.

Esta dualidad, teórica y práctica, característica de la corriente cuantitativa de nuestra disciplina, ha evolucionado a través de los tiempos, desde su nacimiento hasta el momento presente. La Geografía Cuantitativa, surgida al amparo de la revolución positivista de los años sesenta, del siglo pasado, ha sufrido importantes fluctuaciones y vaivenes, que han transformado, profundamente, tanto los objetivos de sus más recientes investigaciones, como la metodología utilizada. La desilusión generada por la rigidez de los planteamientos neopositivistas, que tuvo como consecuencia, en el campo de la Geografía Humana, el nacimiento de nuevos paradigmas, próximos al estructuralismo, marxismo, humanismo o post-modernismo, ha alejado a la mayoría de los cuantitativistas de posiciones excluyentes y totalizadoras, que pretendían la concreción definitiva de los principios básicos de la ciencia espacial, acercándoles a los problemas locales, lejos de la definición de leyes globales explicativas de la distribución de los hechos físicos y humanos en el espacio.

Esta aproximación de la Geografía Cuantitativa a la resolución de problemas prácticos, dentro del marco de la Geografía Aplicada, ha tenido mucho que ver, también, con el desarrollo de los medios informáticos y su creciente empleo en todas las ciencias, particularmente en el campo geográfico. El tratamiento masivo de la información, que posibilitan los ordenadores, ha supuesto un importante empuje de la línea empírica de nuestra disciplina, frente al desarrollo de teorías de carácter deductivo. Además, el nacimiento de los Sistemas de Información Geográficos (SIG) y su decisiva implantación social, desde mediados de los ochenta, ha incorporado al análisis geográfico una poderosa herramienta de tratamiento de la información espacial, que ha condicionado sobremanera los caminos de la investigación. La disponibilidad de un conjunto de rutinas informáticas, que actúan a modo de funciones analíticas, limitadas por la propia decisión de los fabricantes de estos productos, ha conducido el campo de la exploración geográfica a los planteamientos y problemas que operativamente resuelven los SIG, modelando lo que se conoce como «análisis espacial».

Un capítulo importante de la Geografía Cuantitativa, que nos interesa destacar aquí, hace referencia a las técnicas y procedimientos de análisis que se han desarrollado a lo largo de su período de consolidación como rama específica de la ciencia geográfica. Lo que distingue a la Geografía Cuantitativa, de otras ciencias afines, en el tratamiento masivo de la información numérica, como la Econometría, la Sociología Cuantitativa o la Psicología, es, sin duda, la especificidad de los elementos o entes espaciales de recogida de la información. Como ya tuvimos ocasión de exponer en el capítulo anterior, los datos espaciales combinan la información de los atributos con la más genuinamente espacial, que se corresponde con la forma y posición de las unidades geográficas en el espacio. Esta separación de los atributos temáticos de los espaciales ha originado un doble tratamiento de la información geográfica, que es preciso diferenciar, si deseamos analizar con éxito los diversos métodos y técnicas de análisis cuantitativo existentes.

Hasta hace relativamente poco tiempo, la componente espacial de la información geográfica era casi ignorada de la investigación, por lo que su tratamiento específico podía asimilarse al de cualquier otra disciplina científica. Esta vía metodológica, que muchos autores definen como «aespacial», ha consolidado una parte importante de las técnicas de análisis geográfico. Una gran parte de los manuales de estadística, aplicada al campo geográfico, como los de Duncan y Otros (1961), Gregory (1963), King (1969), Hammond y McCullagh (1974) y Ebdon (1977), formulaban la descripción y relación entre las variables geográficas en el campo del

análisis estadístico clásico, sin referencia alguna a las características geométricas o de localización de los entes espaciales, soporte de la recogida de la información. Este tipo de estudios geográficos se completaba con la cartografía final de los resultados, en la perspectiva de interpretar la distribución espacial de los fenómenos geográficos por vía visual y no analítico-matemática. La especificidad de la información geográfica (no independencia de los datos de partida, distribución no gaussiana de los mismos, alejada de la curva normal, etc.) invalidaba, como expondremos más adelante, muchas de las conclusiones de la estadística clásica, sobre todo en el campo inferencial. Nada más lejos de nuestra intención, el excluir o relegar esta vía de análisis, que, aún, hoy en día, continúa dirigiendo gran parte de la investigación geográfica. Sin embargo, resulta imprescindible tomar conciencia del peso que aún tiene en la metodología cuantitativa, por una doble razón. En primer lugar, por servir de marco de referencia para exponer las técnicas cuantitativas específicas a esta manera de proceder, que ha cristalizado en un cuerpo de doctrina, presente en los principales manuales de estadística geográfica, así como para conocer las limitaciones de su aplicación en un contexto geográfico determinado.

Por su parte, el «análisis espacial» es la denominación empleada para referirse al estudio de los datos geográficos, en los que se considera, no solo la componente temática de los mismos, sino sus características espaciales (Unwin, D. 1981). El desarrollo de procedimientos de análisis, relacionados con la componente espacial de la información geográfica, puede realizarse por una doble vía. Por una parte, considerando, exclusivamente, las características geométricas de los elementos espaciales, lo que origina la descripción interrelacionada de los mismos desde la perspectiva de su distancia y forma. Otra alternativa es la consideración de ambos componentes a la vez: temáticos y espaciales. Desde esta perspectiva, resulta habitual la separación de los objetos espaciales en capas de puntos, líneas o polígonos, que han servido para desarrollar metodologías cuantitativas, que afectan al estudio de la distribución espacial de cada uno de estos entes específicos.

Otra división, en el campo cuantitativo de la Geografía, es la que tiene lugar en la diferenciación entre las técnicas estadísticas, tanto espaciales como aespaciales, de la construcción de modelos matemáticos de base espacial. Esta separación no siempre es fácil de realizar, ya que, en muchas ocasiones, se mezclan ambos procedimientos de análisis. Así, es frecuente que un modelo sea construido desde un conjunto de principios matemáticos, pero calibrado por procedimientos estadísticos. El contenido espacial de los modelos, a los que hacemos referencia, nos aconseja

la inclusión de los mismos dentro del apartado de métodos de análisis espacial.

Este breve preámbulo, acerca del carácter de las diferentes técnicas y métodos cuantitativos actualmente existentes, nos va a servir para organizarlos, de forma sintética, con el objeto de pasar a su descripción sumaria. En primer lugar, expondremos las procedimientos de análisis cuantitativo de carácter aespacial y espacial, que han cristalizado en la literatura geográfica como «métodos y técnicas cuantitativas clásicas» y «análisis espacial», para en una segunda fase adentrarnos en la reflexión sobre la importancia de la informática en el desarrollo de este tipo de procedimientos, al implementar su uso en los ordenadores, con un especial detenimiento en el papel de los SIG y su influencia en el planteamiento y resolución de los problemas espaciales.

2. PRINCIPALES TÉCNICAS DE LA ESTADÍSTICA CLÁSICA, EN LA RESOLUCIÓN DE PROBLEMAS GEOGRÁFICOS

Este planteamiento metodológico se sitúa en el campo de la estadística clásica, al analizar los valores temáticos de las entidades geográficas, al margen de su relación con los objetos espaciales de los que proceden. Supone, en realidad, la consideración exclusiva de la base de datos temática, marginando del análisis la base de datos espacial.

El estudio regional clásico, a partir de la matriz geográfica, a la que hicimos referencia en el apartado uno del primer capítulo, sería el mejor exponente de esta manera de proceder. Así, cualquier problema geográfico se vería reducido a un número concreto de filas (características o atributos) y de columnas (lugares). Elaborada la matriz geográfica podríamos abordar el análisis regional desde diversos puntos de vista (Estébanez, J. y Bradshaw, R., 1979):

- a) Examinar la disposición de las celdas, siguiendo una fila completa o una parte de la misma. Supondría el estudio de la variación espacial de una única variable y podría elaborarse un mapa temático de su distribución espacial.
- b) Examinar la disposición de las celdas dentro de cada columna, lo que reflejaría el inventario locacional o estructura local de un lugar determinado.

- c) Comparar dos filas o todo el conjunto de filas. Se estudiaría, así, la covariación o asociación espacial de dos o más hechos geográficos.
- d) Comparar un par de columnas o toda la serie de columnas de la matriz.
- e) Estudiar una submatriz de la matriz general, matizando las variaciones espaciales en una subregión, referidas a unas determinadas variables geográficas.
- f) Comparar una fila o parte de ella en dos momentos de tiempo diferentes.
- g) Comparar una columna o parte de ella en dos momentos de tiempo diferentes.
- h) Estudiar los cambios en las asociaciones o covariaciones espaciales a través del tiempo.
- i) Estudiar las diferenciaciones regionales a través del tiempo.
- j) Comparar una submatriz a través del tiempo.

La descripción y comparación de variables cuantitativas, en un momento del tiempo, o de forma dinámica, resume, de forma sintética, las diferentes posibilidades expresadas en el esquema de análisis anterior. Este planteamiento se acomete en el campo de la estadística clásica, con todas sus potencialidades, pero, también, con todas sus limitaciones para tratar los datos espaciales (problema de la unidad espacial modificable, existencia de la autocorrelación espacial, inadecuación de la suposición de independencia de las observaciones, falta de ajuste a la distribución normal o gaussiana, etc.). La exposición abreviada de las diversas técnicas de análisis estadístico la realizaremos, de acuerdo con un diseño organizativo que incremente, progresivamente, la complejidad del procedimiento estadístico utilizado, comenzando con la descripción individualizada de una única variable y finalizando con la relación multivariada.

2.1. El análisis exploratorio de datos (EDA)

Aunque los geógrafos hemos utilizado, anteriormente a la revolución cuantitativa, métodos estadísticos y gráficos sencillos, con el objeto de describir las variables geográficas, solo recientemente nos podemos referir a los métodos conocidos como *análisis exploratorio de datos* («*exploratory data analysis*», EDA), en la línea de mostrar la distribución de los valores de una variable de manera simple y eficaz. Desde su aparición,

en los comienzos de la década de los años setenta del pasado siglo (Tukey, 1972), esta técnica ha sido incorporada como método de descripción de variables por un número creciente de ciencias sociales, entre ellas la Geografía.

La EDA se caracteriza por el empleo de procedimientos analíticos y descriptivos, gráficos o semigráficos, con la finalidad de mostrar la distribución de valores de una variable en toda su profundidad, sin esconder la estructura subyacente, lo que acontece, normalmente, al utilizar los estadísticos clásicos (media, desviación típica, etc.).

Una forma de expresar la distribución de una variable es mediante una *lista de valores*, donde, junto al identificador de cada caso, se refleje el valor de la medición. Otra forma sería la síntesis de estos datos en una *tabla de frecuencias*, que agrupara, en intervalos, los resultados de la distribución. La tabla de frecuencias es un procedimiento de descripción de una variable que oculta los valores concretos de la misma, por lo que su empleo puede no ser el más adecuado para una determinada investigación. El gráfico denominado de *tallo y hojas* (stem and leaf diagram) es una representación que muestra todos los valores de la variable en un esquema que destaca las cifras que forman parte del tallo (en el ejemplo de la figura 1, las decenas) y las de las hojas (en este caso las unidades). Así, en el ejemplo citado, el menor valor de la serie sería el 12, repetido dos veces, y el mayor valor el 95. El número total de valores sería de 39.

Tallo	Hojas
1	2234
2	1344556
3	2255667778
4	112238
5	3456
6	239
7	126
8	2
9	5

FIGURA 1. Gráfico de tallo y hojas.

Una representación más completa (figura 2) incluye la columna Prof, que muestra cuantos casos acumulados existen desde el comienzo hasta el final de cada fila. En la fila que contiene la mediana, se indica, solo, entre paréntesis, el número de casos de la misma. El resto de las filas reflejan los valores acumulados, desde el otro extremo.

Prof	Tallo	Hojas
4	1	2234
11	2	1344556
(10)	3	2255667778
18	4	112238
12	5	3456
8	6	239
5	7	126
2	8	2
1	9	5

FIGURA 2. Gráfico de tallo y hojas, completado.

Otra representación muy usual es la denominada *gráfico de caja* (boxplot), que presenta la distribución de los valores de una distribución, reflejando, en forma de caja, los más representativos de la misma (valores extremos, cuartiles inferior y superior y mediana) (figura 3). Esta manera gráfica de mostrar la información resulta bastante sugestiva, ya que permite esbozar las áreas de la distribución donde la concentración de casos es mayor o menor. En el ejemplo anterior, los valores más representativos serían: 12 y 95 de valores extremos, 25 como cuartil inferior (ocuparía la posición 10), 37 de mediana (ocuparía la posición 20) y 55 de cuartil superior (ocuparía la posición 30).



FIGURA 3. Gráfico de caja (Boxplot).

La relación entre variables puede expresarse en la metodología EDA, de forma gráfica, sin tener que aventurar, previamente, la existencia de un tipo determinado de correspondencia (lineal, exponencial, etc.). La manera más habitual de estudiar la relación entre dos variables es la elaboración del llamado *gráfico de dispersión*, formado por dos ejes de coordenadas perpendiculares, en cada uno de los cuales se sitúa una de las dos variables y cada observación se representa por un punto. El estudio de la nube de puntos así originada permite identificar las características de la relación entre las variables. La figura 4 muestra, a modo de ejemplo, la nube de puntos, en cada columna, de los tiempos percibidos en diez re-