“The new source of power is not money in the hands of a few but information in the hands of man.”
John Naisbitt

Chapter 3 GIS

A tool for decision-making

Every day you wake at 6 o’clock in the morning. At 8 o’clock you go to school which is four kilometres south of your house. You return home at 4 o’clock in the afternoon travelling along the same route. Then at 5 o’clock you call your friends and go for a game of football at the nearby playground that is 10 minutes walk from your house. Many of our activities are related to place and time in one way or the other. Planning and decision-making—whether it is planning a new road or finding a suitable location for a health centre—are influenced or dictated by location or a geographic component. The major challenges we face in the world today—over-population, deforestation, natural disasters—have a critical geographic dimension.

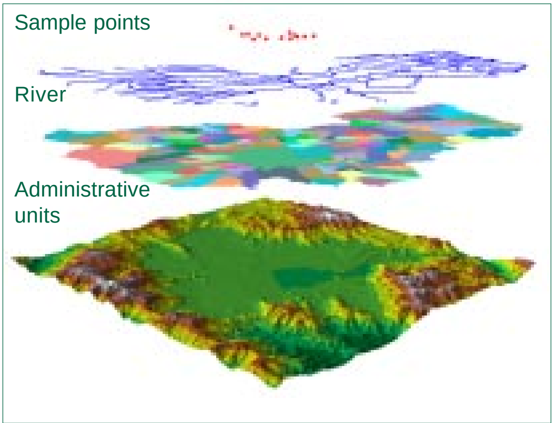

Figure 3.1
Geography in layers

Our geography can be considered as a number of related data layers as illustrated in Figure 3.1. GIS combine layers of information about a place to give an understanding of that place. Which layers of information are combined depends on a purpose: for example, finding the best location for a new supermarket, assessing environmental damage, tracking delivery vehicles or modelling the global environment. A GIS stores information about the world as a collection of thematic layers that can be linked together by geography. This simple but extremely powerful and versatile concept has proven invaluable for solving many real-world problems.

In the strictest sense, GIS are computer systems for collecting, storing, manipulating and displaying geographic information. There are many definitions for GIS. However, their major characteristic is geographic (spatial) analysis functions that provide a means for deriving new information based on location.

GIS functions

There are four basic functions of GIS: data capture, data management, spatial analysis and presenting results.

Data capture

Data used in GIS come from many sources, are of many types and are stored in different ways. A GIS provides tools and methods for the integration of data into a format so that data can be compared and analysed. Data sources are mainly manual digitisation/scanning of aerial photographs, paper maps and existing digital data. Remote-sensing satellite imagery and GPS are also data input sources.

Data management

After data are collected and integrated, a GIS provides facilities that can contain and maintain data. Effective data management includes the following aspects: data security, data integrity, data storage and retrieval, and data maintenance.

Spatial analysis

Spatial analysis is the most important function of a GIS that makes it distinct from other systems such as computer aided design and drafting (CADD). The spatial analysis provides functions such as spatial interpolation, buffering and overlay operations.

Presenting results

One of the most exciting aspects of GIS is the variety of ways in which information can be presented once it has been processed. Traditional methods of tabulating and graphing data can be supplemented by maps and three-dimensional images. These capabilities have given rise to new fields such as exploratory cartography and scientific visualisation. Visual presentation is one of the most remarkable capabilities of GIS that allows for effective communication of results.

Questions GIS can answer

GIS can be distinguished by listing the types of questions it can answer.

Location: What is at…?

This question seeks to find what exists at a particular location. A location can be described in many ways using, for example, a place name, postcode or geographic reference such as longitude/latitude or x and y.

Condition: Where is it…?

This question is the converse of the first and requires spatial data to answer. Instead of identifying what exists at a given location, one may wish to find locations where certain conditions are satisfied (e.g. a non-forest area of at least 2000 m^2 within 100 m of a road and with soils suitable for supporting buildings).

Trends: What has changed since…?

This question might involve both of the first two and seeks to find the differences within an area over time, e.g., changes in forest cover or the extent of urbanisation over the last ten years.

Patterns: What spatial pattern exists…?

This question is more sophisticated. It might be asked to determine whether landslides are occurring mostly near streams or to find out at which traffic points accidents are occurring most frequently. It might be just as important to know how many anomalies there are and where they are located.

Modelling: What if…?

This question is posed to determine what happens if, for example, a new road is added to a network or a toxic substance seeps into the local groundwater supply. Answering this type of question requires both geographic and other information (as well as specific models).

Geographic data

There are two important components of geographic data: geographic position and attributes or properties. In other words, spatial data (where is it?) and attribute data (what is it?). Geographic position specifies the location of a feature or phenomenon by using a coordinate system. The attributes refer to the properties of spatial entities such as identity (e.g. maize, granite, lake), ordinal (e.g. ranking such as class 1, class 2, class 3) and scale (e.g. value such as water depth, elevation, erosion rate). They are often referred to as non-spatial data since they do not in themselves represent location information.

Raster and vector data

Spatial features in a GIS database are stored in either vector or raster form. GIS data structures adhering to a vector format store the position of map features as pairs of x, y (and sometimes z) coordinates. A point is described by a single x-y coordinate pair and by its name or label. A line is described by a set of co-ordinate pairs and by its name or label. In theory, a line is described by an infinite number of points. In practice, this is not feasible. Therefore, a line is built up of straight-line segments. An area, also called a polygon, is described by a set of coordinate pairs and by its name or label with the difference that the coordinate pairs at the beginning and end are the same (Figure 3.2).

A vector format represents the location and shape of features and boundaries precisely. Only the accuracy and scale of the map compilation process, the resolution of input devices and the skill of the data-inputter limit precision.

In contrast, the raster or grid-based format generalises map features as cells or pixels in a grid matrix (Figure 3.3). The space is defined by a matrix of points

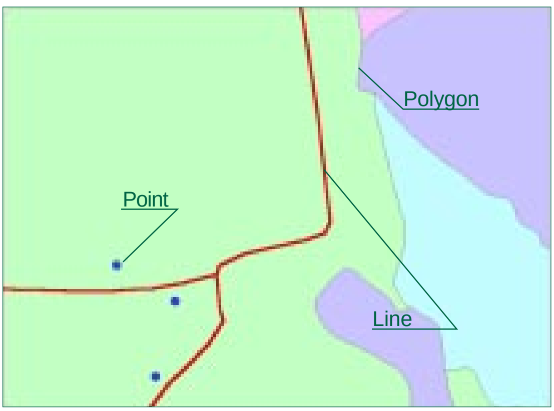

Figure 3.2
Vector format

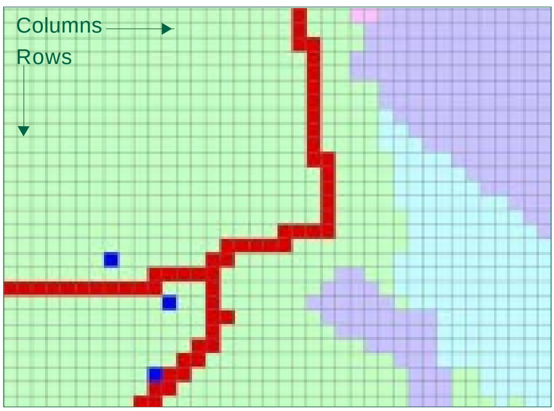

Figure 3.3
Raster format

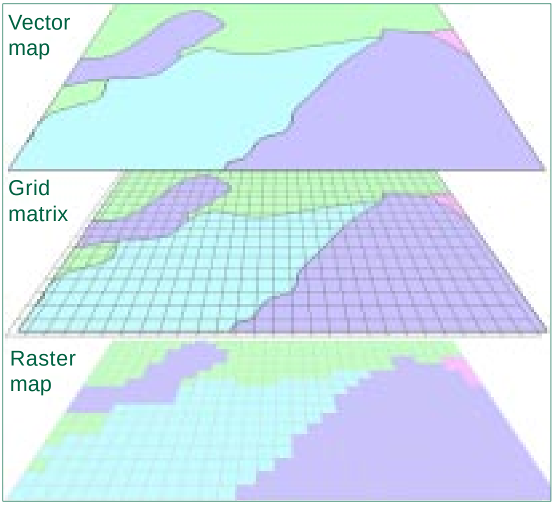

Figure 3.4
Vector–raster relationship

or cells organised into rows and columns. If the rows and columns are numbered, the position of each element can be specified by using column number and row number. These can be linked to coordinate positions through the introduction of a coordinate system. Each cell has an attribute value (a number) that represents a geographic phenomenon or nominal data such as land-use class, rainfall or elevation. The fineness of the grid (in other words, the size of the cells in the grid matrix) will determine the level of detail in which map features can be represented. There are advantages to the raster format for storing and processing some types of data in GIS. The vector–raster relationship is shown in Figure 3.4.

Organising attribute data

GIS use raster and vector representations to model earth features or phenomena. Apart from locations, GIS must also record information about them. For example, the centre line that represents a road on a map does not tell you much about the road except its location. To determine the road's width or pavement type or condition, such information should be stored so that it can be accessed by the system as needed. This means that the GIS must provide a linkage between spatial and non-spatial data. These linkages make the GIS 'intelligent' as the user can store and examine information about where things are and what they are like. The linkage between a map feature and its attributes is established by giving each feature at least one unique means of identification—a name or number usually called its ID. Non-spatial attributes of the feature are then stored, usually in one or more separate files, under this ID number (Figure 3.5)

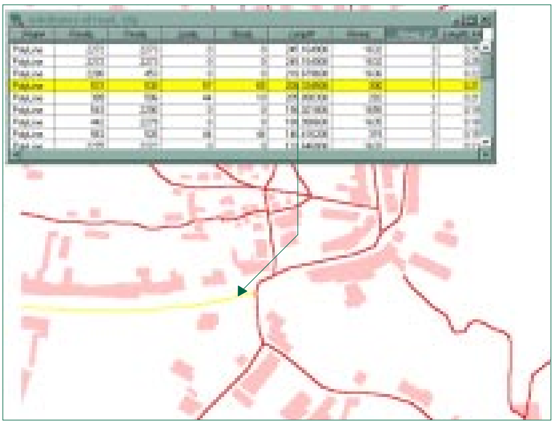

Figure 3.5
Linking attribute data

This non-spatial data can be filed in several forms depending on how it needs to be used and accessed. Many GIS software use the relational database management systems (RDBMS) to handle attribute data.

A relational database is the perception of data as series of tables that are logically associated with each other by shared attributes (Figure 3.6). Any data element in a relationship can be found by knowing the table name, the attribute (column) name and the value of the primary key. The advantage of these systems is that they are flexible and can answer any question formulated with logical and mathematical operators.

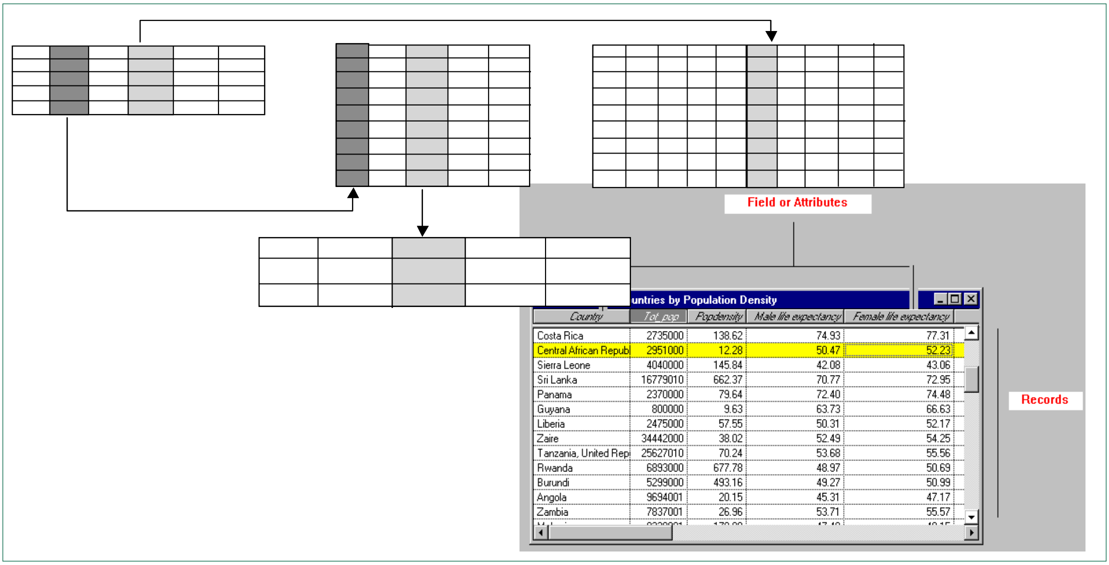

Figure 3.6
Relational database management system

Metadata

Metadata are simply defined as 'data about data'. It gives the information about the content, source, quality, condition and other relevant characteristics of the data (Figure 3.7). For instance, it may describe the content as road or land-use data, the source as where the data have come from, the quality as the level of accuracy, the condition as whether the data are outdated or partial and so on.

Figure 3.7
Metadata