REPRESENTING THE REAL WORLD AS DATA

How would you create an information system to organize and manage the huge variety of geographic stuff in the world? One approach is to think of all that stuff in terms of discrete objects.

THE DISCRETE-OBJECT VIEW OF THE WORLD

If you conceive of geography in terms of objects, you can sort these objects by similarities. Shape is a fundamental sorting principle: every object can be drawn—in two dimensions—as either a point, a line, or a polygon. Theme, or type, is another principle: every object can be classified as a school, a road, a park, or something else.

Applying these sorting principles of shape and theme, you can come up with collections of things you would recognize on a map: schools represented as points, roads represented as lines, parks represented as polygons, and so on.

Points, lines, and polygons representing geographic objects on a map.

Each object in a collection has a unique location, specified by a pair of spatial coordinates (for points) or a list of coordinate pairs (for lines and polygons).

A polygon (with an *x* indicating its centroid) and the list of coordinate pairs that define its location.

Besides a unique location, every object has a set of facts that pertain to it: a name, a description, or whatever bits of information have been gathered about it. These facts are the object's attributes.

FID	Shape *	PARK NAME	CATEGORY	ADDRESS	TELEPHONE	HOURS
0	Polygon	Los Angeles	State Historic Park	1245 N. Spring Street Los	213-620-6152	8 AM - sunset
1	Polygon	Rio de Los Angeles	State Recreation Area	1900 San Fernando Road	213-620-6152	9 AM - 10:30 PM
2	Polygon	Vista Hermosa Park	Local Park or Recreation Area	100 N. Toluca Street Los	213-250-3578	8 AM - sunset

In ArcGIS, a collection of such objects—with a common shape, common theme, and common attributes—is called a *feature class*. An individual object in the collection is a *feature*. The feature class is the basic storage unit for GIS data created according to the discrete-object view of the world, commonly called the *vector data model*.

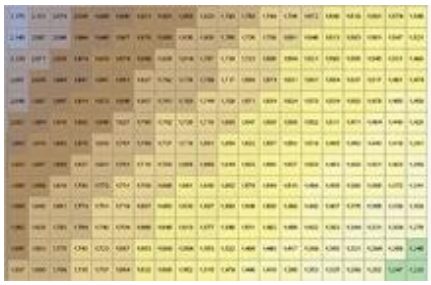
Feature classes can be stored in various file formats, notably the geodatabase and the shapefile. The geodatabase format is newer and more highly developed.

Feature classes in geodatabase format (.gdb) and shapefile format (.shp). Different icons indicate polygon, line, and point data.

THE CONTINUOUS-SURFACE VIEW OF THE WORLD

Although it's a powerful model, the discrete-object view is not an intuitive way to think of certain kinds of geographic information, such as elevation or temperature, that don't have shapes or boundaries and that cover the world everywhere. It's quite possible to represent these phenomena as features (for example, contour lines represent elevation on topographic maps), but a more natural way to think of them is in terms of continuous expanses, or surfaces.

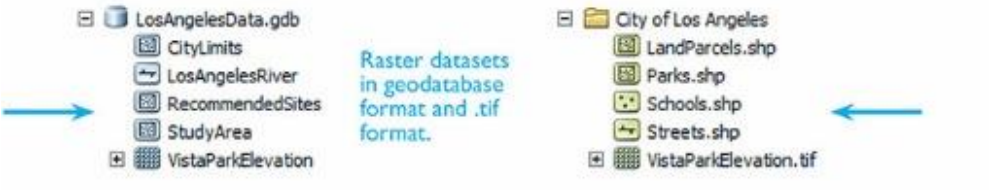
The most common way to model a geographic surface is with a matrix of square cells, or pixels. Each cell represents a unit of area, such as a square meter, and stores a single piece of geographic information—typically, a measured or estimated value—at that location.



In this example, each cell represents 30 square meters of ground and stores a single representative elevation value.

This way of modeling surfaces is called the *raster data model*. It's commonly used for elevation and its derivatives (slope, aspect); for temperature, precipitation, and land cover; for statistical data, such as densities and means; and especially for imagery.

The raster dataset is the basic storage unit for GIS data created according to the continuous-surface view of the world. Raster datasets can be stored in geodatabases or in various standard image file formats, such as TIFF and JPEG.



Feature classes and raster datasets are complementary. In many maps, raster datasets are used for background display, whereas feature classes are used for foreground display and analysis.