PART I

Basic Concepts of GIS

COPYRIGHTED MATERIAL

CHAPTER **1**

Some Concepts That Underpin GIS

OVERVIEW

IN WHICH you begin to understand the rather large and complex body of ideas and techniques that allow people to use computers to comprehend and design the physical environment. And in which you use ESRI's ArcCatalog to explore geographic data.[1]

You Ask: "What Is GIS About?"

The poem, "The Blind Men and the Elephant," tells the story of six sightless men who approach an elephant, one by one, to satisfy their curiosity.

It was six men of Indostan
To learning much inclined
Who went to see the Elephant
(Though all of them were blind)

That each by observation
Might satisfy his mind.

The First approached the Elephant,
And happening to fall
Against his broad and sturdy side,
At once began to bawl,
"God bless em! but the elephant
Is very like a WALL!"

The Second, feeling of the tusk
Cried: "Ho! what have we here
So very round and smooth and sharp?
This wonder of an Elephant
Is very like a SPEAR!"

The Third approached the animal,
And, happening to take
The squirming trunk within his hands,
Thus boldly up and spake:
"I see," quoth he, "the Elephant,
Is very like a SNAKE!"

[1]ESRI is the Environmental Systems Research Institute. ESRI makes the software, ArcGIS, which you will use in this text to understand the concepts of Geographic Information Systems.

The Fourth reached out an eager hand,
And felt about the knee
"What most this wondrous beast is like
Is mighty plain," quoth he:
" 'Tis clear enough the Elephant
Is very like a TREE!"

The Fifth, who chanced to touch the ear,
Said: "E'en the blindest man
Can tell what this resembles most;
Deny the fact who can,
This marvel of an Elephant
Is very like a FAN!"

The Sixth no sooner had begun
About the beast to grope,
Than seizing on the swinging tail
That fell within his scope,
"I see," quoth he, "the Elephant
Is very like a ROPE!"

And so these men of Indostan
Disputed loud and long,
Each in his own opinion
Exceeding stiff and strong
Though each was partly in the right
And all were in the wrong!

[. . .][2]

Excerpted from "The Blind Men and the Elephant"
(based on a famous Indian legend)
John Godfrey Saxe
American Poet (1816–1887)

You Ask Again: "What Is GIS About?"

Poet Saxe's lines could apply to geographic information systems (GIS) in that relating to the subject may well depend on your point of view. Asking what GIS[3] is about is sort of like asking, "What is a computer about?" The capabilities of GIS are so broad and its uses so pervasive in society, geography, and the technical world in general that a short, meaningful description is impossible. But for starters, here is a generic definition of GIS that you might find in a dictionary:

[2][. . .] indicates an omission.
[3]In this text "GIS" stands for "a geographic information system" or for the plural "geographic information systems" depending on the context.

> *A geographic information system is an organized collection of computer hardware and software, people, money, and organizational infrastructure that makes possible the acquisition and storage of geographic and related attribute data, for purposes of retrieval, analysis, synthesis, and display to promote understanding and assist decision making.*

To better understand one facet of GIS, consider how you might use the technology for a particular application. Solve the following site selection problem:

Exercise 1-1 (Project)

Finding a Site by Manual Means

Wildcat Boat Company is planning to construct a small testing facility and office building to evaluate new designs. They've narrowed the proposed site to a farming area near a large lake and several small towns. The company now needs to select a specific site that meets the following requirements:

- ❑ The site should not have trees (to reduce costs of clearing land). A regional agricultural preservation plan prohibits conversion of farmland. The other land uses (urban, barren, and wetlands) are also out. So the land cover must be "brush land."
- ❑ The building must reside on soils suitable for construction.
- ❑ A local ordinance designed to prevent rampant development allows new construction only within 300 meters of existing sewer lines.
- ❑ Water quality legislation requires that no construction occur within 20 meters of streams.
- ❑ The site must be at least 4000 square meters in size to provide space for building and grounds.

Figure 1-1 is a key to the following maps. It shows the symbols for land cover, soil suitability, streams, and sewers.[4]

Figure 1-2 is a map showing land cover in the area where the site must be chosen. Different crosshatch symbols indicate different types of land cover; the white area in the northern area of the map is water. The land cover codes and categories are as follows:

LC CODE	***LC TYPE***
100	Urban
200	Agriculture
300	Brush land
400	Forest
500	Water
600	Wetlands
700	Barren

[4]These maps are also available on the CD-ROM that accompanies the text. They are the image files: Key_to_maps.jpg, Landcover.jpg, Soil_suitability.jpg, and Streams_&_Sewers.jpg located in the folder IGIS_Arc_AUX.

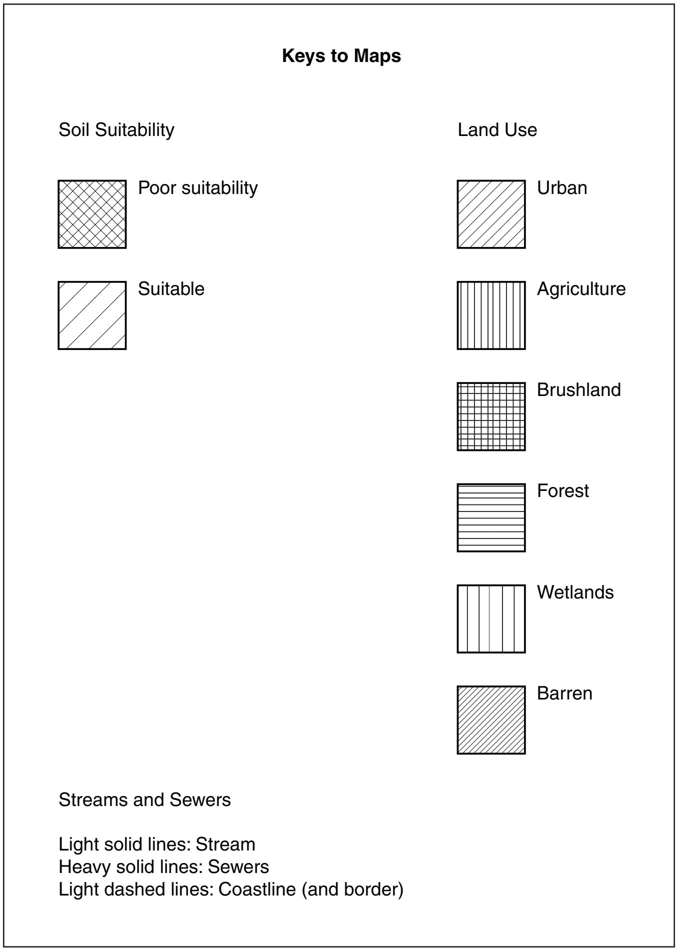


FIGURE 1-1 Key to maps of the Wildcat Boat facility area area maps

Figure 1-3 is a soil suitability map. Lines separate soils of different types. The different soils are categorized as suitable or unsuitable for building. Therefore, you will see the same symbol on both sides of a dividing line, indicating that while such soil types may be different, their suitability is the same.

Figure 1-4 shows natural streams in the area and human-made sewer lines.

You may use scissors, xerography, a computer-based drawing program, a light table, and any other tools to solve the problem.

You are asked to present a map that shows *all the areas* where the company could build while meeting the requirements stated previously. Make your map the same scale and size as those maps provided on the CD-Rom. Outline in red the areas that meet the requirements. You don't need to produce a high-quality

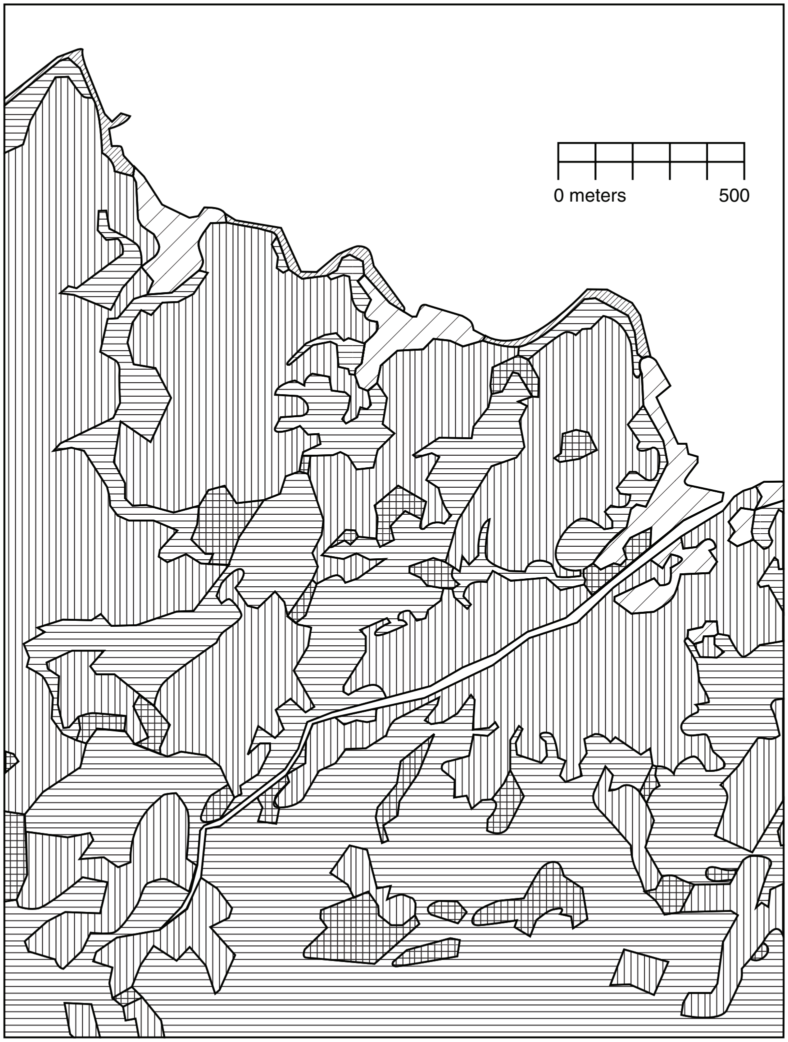


FIGURE 1-2 Land Cover

cartographic product. The main object here—indeed the object of this textbook—is to analyze geographic data. While making maps is important it is not the primary focus of this book.

Write a brief description (100 to 200 words, preferably typed or computer printed) of the procedure you used to make the map.

The problem is much easier than it might otherwise be because the maps provided cover exactly the same area, have the same underlying assumptions regarding the shape and size of Earth, are at the same scale, and use the same projection of Earth's sphere onto the flat plane of the map. These benefits are often not available in the real world, where you frequently need a considerable amount data preparation to solve such a problem.

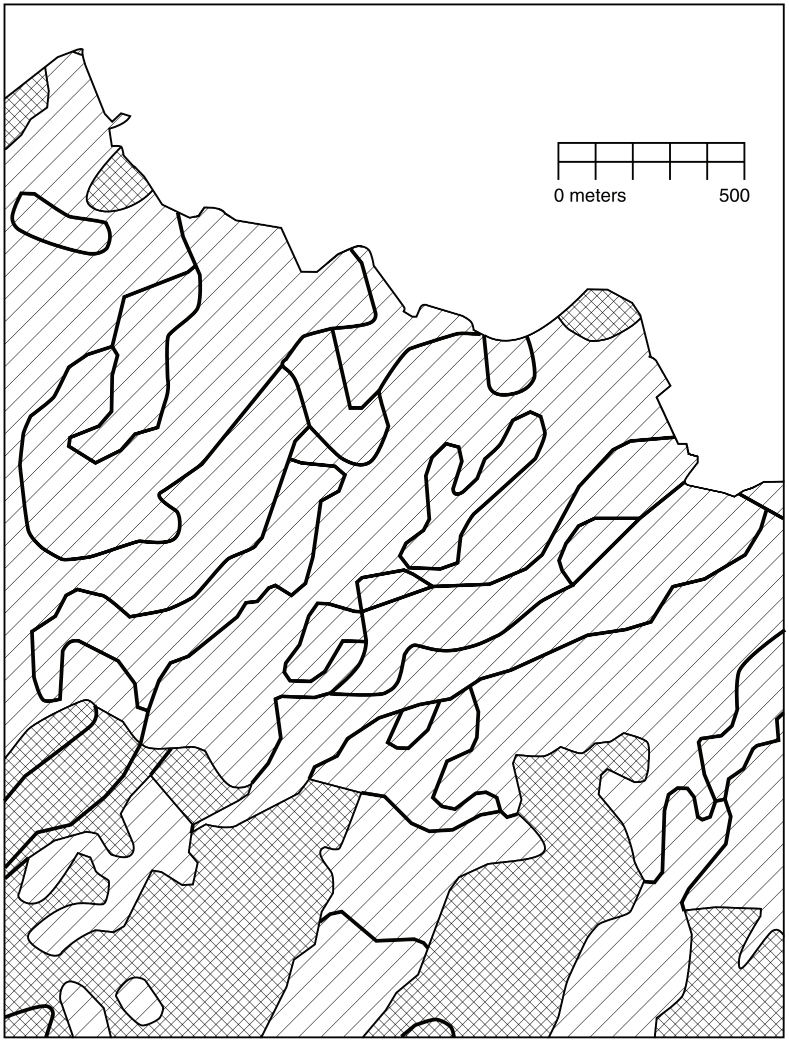


FIGURE 1-3 Soil Suitablilty

More of What GIS Is About

Completing Exercise 1-1 showed how GIS can help you solve one kind of problem. There are many others. Computer-based GIS not only serves the purpose of traditional maps but also helps you perform activities that involve spatial analysis, even without maps. Understanding conditions that occur in the vicinity of the Earth's surface are important in building structures, growing crops, preserving wildlife habitat, protecting ourselves from natural disasters, navigating from one point to the next, and a myriad of other activities.

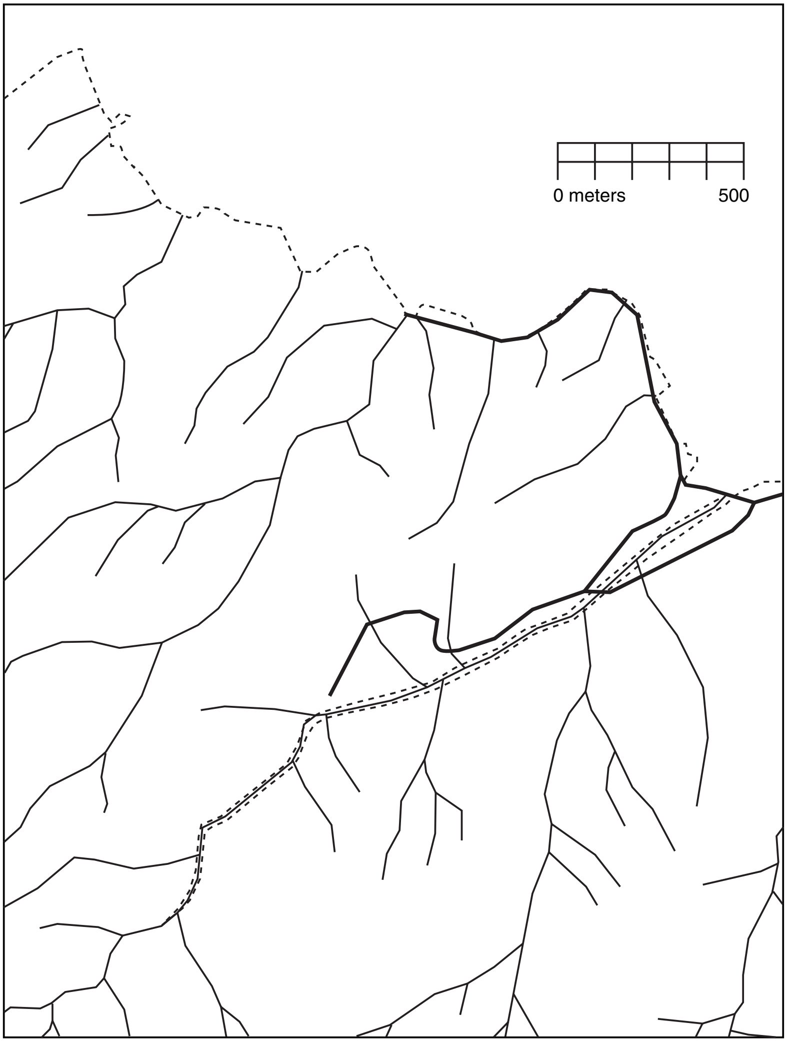


FIGURE 1-4 Streams and Sewers

GIS has many uses:

Land use—Helps determine land cover, zoning, environmental impact analysis, locational analysis, and site analysis.

Natural environment—Identifies, delineates, and manages areas of environmental concern, analyzes land-carrying capacity, and assists in environmental impact statements.

Energy—Examines costs of moving energy, determines remaining available energy reserves, investigates the efficiency of different allocation schemes, reduces waste, reduces heat pollution, identifies areas of danger to humans and animals, assesses environmental impacts, sites new distribution lines and facilities, and develops resource allocation schemes.

Human resources—Plans for mass transit, recreation areas, police unit allocation, and pupil assignment and analyzes migration patterns, population growth, crime patterns, and welfare needs. It also manages public and government services.

Areas of environmental concern—Facilitates identification of unique resources, manages designated areas, and determines the relative importance of various resources.

Water—Determines floodplains, availability of clean water, irrigation schemes, and potential and existing pollution.

Natural resources—Facilitates timber management, preservation of agricultural land, conservation of energy resources, wildlife management, market analysis, resource allocation, resource extraction, resource policy, recycling, and resource use.

Agriculture—Aids in crop management, protection of agricultural lands, conservation practices, and prime agricultural land policy and management.

Crime prevention, law enforcement, criminal justice—Facilitates selection of sites or premises for target-hardening attention, establishment of risk-rating procedures for particular locations, tactical patrol allocation, location selection for crime prevention analysis, crime pattern recognition, and selection of areas or schools for delinquency prevention attention.

Homeland security and civil defense—Assesses alternative disaster relief plans, needs for stockpiling of foods and medical supplies, evacuation plans, and the proper designation of disaster relief areas.

Communications—Facilitates siting of transmission lines, location of cellular equipment, and education.

Transportation—Facilitates alternative transportation plans, locational analysis, mass transit, and energy conservation.

Next Steps: Seemingly Independent Things You Need to Know

Before we launch into the theory and application of GIS, let's look ahead at the remaining text in the Overview of this chapter. You may know some or all of this material already, depending on your background. To use GIS effectively, you should know something about several topics that may seem unrelated at first glance. The next few sections briefly review the relevant aspects of the following:

- ❑ Determining where something is: coordinate systems
- ❑ Determining where something is: latitude and longitude
- ❑ Geodesy, coordinate systems, geographic projections, and scale
- ❑ Projected coordinate systems
- ❑ Geographic vs. projected coordinates: which to use
- ❑ Two projected coordinate systems: UTM and state plane

- Coordinate transformations
- Physical dimensionality
- Global positioning systems
- Remote sensing
- Relational databases
- Another definition of GIS
- Computer software: in general
- Computer software: ArcGIS in particular

Determining Where Something Is: Coordinate Systems

Cartesian Coordinate Systems

A coordinate system is a way of determining where points lie in space. We are interested in two-dimensional (2-D) space and three-dimensional (3-D) space. In general, it takes two numbers to assign a position in 2-D space and three numbers in 3-D space.

Coordinates may be thought of as providing an *index* to the locations of points in space, and hence to the features that these points define.

To make a 2-D Cartesian[5] coordinate system, draw two axes (lines) that cross at right angles on a piece of paper. The point at which they cross is called the origin. Arrange the page on a horizontal table in front of you so that one line points left-right and the other toward and away from you. The part of the line from the origin to your right is called the positive x-axis. The line from the origin away from you is called the positive y-axis. Mark each axis in equal linear units (centimeters, say) starting at the origin, as shown in Figure 1-5. Now, a pair of numbers serves as a reference to any point on the plane. The position $x = 5$ and $y = 3$ (shorthand: (5,3)) is shown. The sheet of paper is the x-y plane.

Create a 3-D Cartesian coordinate system from the 2-D version. Imagine a vertical line passing through the origin; call it the z-axis; the positive direction is up. Now you can reference any point in three-dimensional space. The point $x = 5$, $y = 3$, and $z = 4$ (5,3,4) is shown in Figure 1-6.

This is called a right-hand coordinate system. The thumb, forefinger, and middle finger represent the positive axes x, y, and z, respectively. With your right hand outstretched, arrange those three digits so that they are roughly mutually orthogonal—that is, with 90° angles between each pair. Point your thumb to the right and your forefinger away from you. Now, unless you are in considerable pain, your middle finger will be pointing up.[6]

[5]Rene Descartes, who lived from 1596 to 1650, made major contributions to both mathematics and philosophy. Descartes is credited with integrating algebra and geometry, by inventing the coordinate system that (almost) bears his name.
[6]It's best not to practice this exercise where other people can see you.

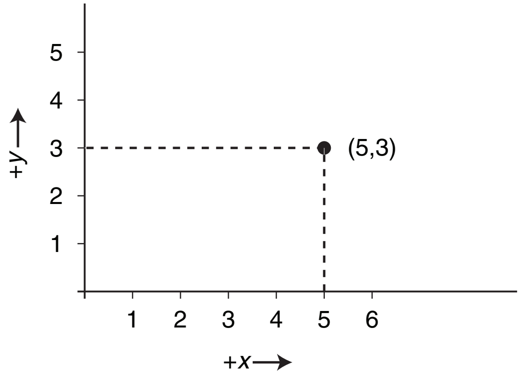


FIGURE 1-5 2-D Cartesian coordinate system

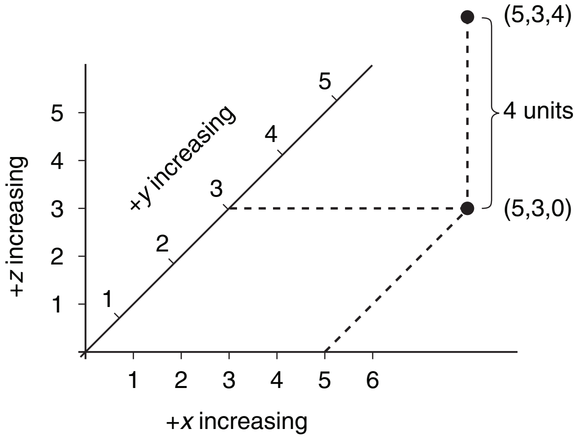


FIGURE 1-6 3-D Cartesian coordinate system

Spherical Coordinate Systems

A spherical coordinate system is another way to reference a point in 3-D space. It also requires three numbers. Two are angles and the third is a distance. Consider a ray (a line) emanating from the origin. The angles determine the direction of the ray. See Figure 1-7.

The latitude-longitude graticule (a gridded reference network of lines encompassing the globe) is based on a spherical coordinate system. As usual, referencing navigation and Earth location issues requires a

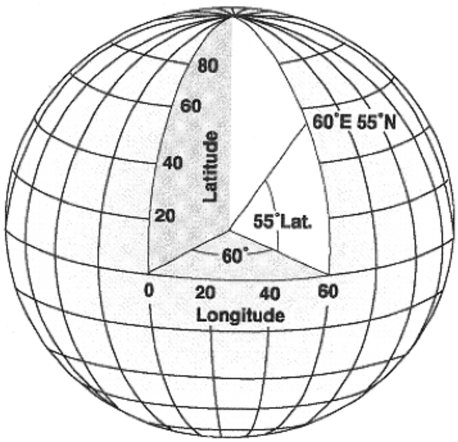


FIGURE 1-7 Spherical coordinate system, showing latitude 55° North and longitude 60° East

different system from the one mathematicians use in more abstract systems.[7] The origin is considered to be the center of the Earth. The equator serves as intersection of the x-y plane and the hypothetical sphere of the Earth. To determine the coordinates of a point, one angle (latitude) is measured from the x-y plane. The other angle (longitude) is contained in the x-y plane and is measured from the meridian that passes through Greenwich, England. The third number in a mathematical spherical coordinate system is the distance along the ray from the origin to the point. When added to the latitude-longitude system, altitude is usually defined instead to be the distance to the point along the ray from mean (average) sea level (MSL) or from a gravity-defined pseudo-ellipsoid used with the NAVSTAR Global Positioning System, to be discussed shortly.

By using three numbers, you can determine and communicate the position of any point on Earth. Of course, a set of parameters must qualify these numbers. Any given point on the surface has probably been addressed by dozens of different sets of numbers, based on the parameters (e.g., units) of the coordinate system chosen. The parameters must match when you combine GIS data.

Determining Where Something Is: Latitude and Longitude

A fundamental principle underlies all geography and GIS: Most things on Earth don't move (or move very slowly) with respect to each other. Therefore, we can talk about the *position* of something embedded in or attached to the ground and know that its position won't change (much). It seems like a straightforward idea, but position confuses a lot of people when it is described as a set of numbers.[8]

---

[7]For example, in the two dimensional Cartesian plane, a mathematician will measure angles starting with the positive x-axis (east) as zero and increasing counterclockwise, to 360 degrees (which is, again, zero). The navigator (think: compass) or cartographer will also use 360 degrees to represent a full circle, but measures clockwise from the positive y-axis (north).

[8]Descriptions of points aren't always just numbers. A possibly apocryphal "meets and bounds" description of a point in Kentucky a couple of centuries ago was "Two tomahawk throws from the double oak in a northerly direction."

Let's suppose in 1955 somewhere in the United States you (or your parents, or their parents) drove a substantial metal stake or pin vertically into solid ground. Now consider that the object, unless disturbed by humans or natural forces such as erosion or earthquake, would not have moved with respect to the planet since then.[9] In other words, it is where it was, and it will stay there. Three numbers—latitude, longitude, and altitude—could identify the location of the object in 1955. But over the last half century, teams of mathematicians and scientists (skilled in geodesy) developed other sets of numbers to describe exactly the same spot where your object now resides. *The actual position of the object didn't change, but additional descriptions of where it is have been created.*

Ignoring the matter of altitude for the moment, suppose the object had been driven into the ground at latitude 38.0000000° (North) and longitude 84.5000000° (West) according to the calculations done before 1955 that indicated the location of the center of the Earth, its shape, and the location of its poles. Most people and organizations in the United States in 1955 used the North American Datum of 1927 (NAD27) to estimate the latitude-longitude graticules, based on parameters of the ellipsoid determined by Clarke in 1866.[10]

The datum described as the World Geodetic System of 1984 (WGS84) offers the most recent, widely accepted view of the center of the Earth, its shape, and the location of its poles. The ellipsoid of WGS84 is virtually identical to the GRS80 ellipsoid.[11] In the coterminous states of the United States, this datum is virtually identical (within millimeters) to the North American Datum of 1983 (NAD83), though they result from different calculations.

According to the latitude-longitude graticule (WGS84), the object previously described would be at latitude 38.00007792° and longitude 84.49993831°. The difference might seem insignificant, but it amounts to about 10 meters on the Earth's surface. Or consider it this way: According to NAD83, a second object placed in the ground at 38°N latitude and 84.5°W would be 10 meters away from the first one. Does that sound like a lot? People have exchanged gunfire in disputes over smaller distances. Given a latitude and longitude, a GIS must know the datum that is the basis for the numbers. Hundreds of datums exist, and many countries have their own.

Geodesy, Coordinate Systems, Geographic Projections, and Scale

First, a disclaimer: This text does not pretend to cover in detail such issues as geodetic datums, projections, coordinate systems, and other terms from the fields of geodesy and surveying. Nor will the text rigorously define most of these terms. Simply knowing the definitions will mean little without a lot of study. Many textbooks and Web pages are available for your perusal. These fields, concepts, and principles may or may not be important in your use of GIS, depending on your project. But the datum, projection, coordination system designations, and measurement units must be identical when you combine GIS or map information. If not, your GIS project may well produce inaccurate results.

---

[9]Well, it hasn't moved much. If it was on the island of Hawaii, it has moved northwest at about 4 inches per year. If you are unfortunate enough to be in a place where there was an earthquake, it might well have moved and not returned to its original position.

[10]Based on a monument on Meades Ranch, in Kansas; the Clarke 1866 ellipsoid was meant to minimize the error between the itself and the geoid, in the United States.

[11]GRS80 is a global geocentric system based on the ellipsoid adopted by the International Union of Geodesy and Geophysics (IUGG) in 1979. GRS80 is the acronym for the Geodetic Reference System of 1980.

How we apply the mathematically perfect latitude-longitude graticule to points on the ground depends partly on human's understanding of the shape of the Earth. This understanding changes the more we learn. Geodesy is the study of the shape of the Earth and the validity of the measurements humans make on it. It deals with such issues as spheroid and datum. You don't have to know much about geodesy to use a GIS effectively, provided your data are all based on the same spheroid and datum (and projection and units, as we will see later). It is the application of geodesic knowledge that caused the differences in the coordinates of the object that was put into the ground 50 years ago. The object hasn't moved. We simply have a better idea of the location of the object relative to the latitude-longitude graticule.

Projected Coordinate Systems

For many reasons, it's often not convenient to use latitude and longitude to describe a set of points (perhaps connected by straight lines to make up a coastline or a country's boundaries) on the Earth's surface. One is that doing calculations using latitude and longitude, say, determining the distance between two points, can involve complex operations such as products of sines and cosines. For a similar distance calculation, if the points are on the Cartesian x-y plane, the worst arithmetic hurdle is a square root.

Latitude and longitude measures for many geographic applications do not work well for several aspects of mapmaking. Suppose you plot many points on the Earth's surface—say, along the coastline of a small island some distance from the equator—on a piece of ordinary graph paper, using the longitude numbers for x-coordinates and latitude numbers for y-coordinates. The shape of the island would look strange on the map compared to how it would appear from an airplane. You would not get useful numbers if you measured distances or angles or areas on the plot. This is due to a characteristic of the spherical coordinate system: The length of an arc of a degree of longitude does not equal the length of an arc of a degree of latitude. They are almost equal near the equator, but the difference grows as you go further north or south from the equator. At the equator a degree of longitude translates to about 69.17 miles. Very near the North Pole a degree of longitude might be 69.17 inches. (A degree of latitude, in contrast, varies only between about 68.71 miles near the equator and 69.40 miles near the poles.)

For relatively small areas, mathematically projecting the spherically defined locations onto a plane provides a good solution to problems associated with calculation and plotting. Geographic projection might be thought of as imagining a process that places a light source inside a transparent globe that has features of the Earth inscribed on it. The light then falls on a flat piece of paper (or one that is curved in only one direction and may be unrolled to become flat).[12] The shadows of the features (say, lines, or areas) will appear on the paper. Applying a Cartesian coordinate system to the paper offers the advantages of easy calculation and more realistic plotting. However, distortions are inherent in any projection process; most of the points on the map will not correspond exactly to their counterparts on the ground. The degree of distortion is greater on maps that display more area. Accuracy will suffer when you flatten a curved surface and convert a three-dimensional coordinate system to a two-dimensional one.

After constructing geographic data sets according to latitude and longitude, based on some agreed-upon spheroid (such as GRS80) and datum (such as WGS84), you decide how to represent them graphically for viewing. At one time, cartographers went directly from the latitude-longitude description of an area

[12]This description is a sort of a cartoon to describe a map projection. A map projection is actually a mathematical transformation that "maps" points on the globe to points on a plane; the process may be quite complex depending on the projection; a single light source at the center of the globe does not suffice to explain it.

or feature to a graphical portrayal on paper. This usually involved "projecting" the data from a three-dimensional (3-D) spherical coordinate system to a two-dimensional (2-D) Cartesian one, and setting a scale: A certain number of units on a map represented that number of units on the ground. Using GIS changes this. As I commented earlier, latitude-longitude is the most fundamental and accurate way to represent spatial data. So large areas—those that are most subject to distortion by being projected—may best be left in latitude-longitude coordinates, and subareas projected to other coordinate systems when needed. Computers are really good at doing the complex computations necessary to project data.

The larger the area projected, the greater the tendency for things not to be where the coordinates say they are.

Four considerations for viewing and analysis are size, shape, distance, and direction. Myriad projections have been invented. Many distort size, shape, distance, or direction, and some preserve one or two of them.

Regarding linear units, GIS differs from cartography in that matters of scale may be left until the very end. The position defining numbers in the database should be real-world coordinates—not scaled coordinates. That is, GIS "maps" are stored in ground units rather than map units. Besides being a more fundamental way to store data, this makes it possible to easily make maps of any desired scale on the computer monitor or on paper. Scale is only a consideration when you measure on a physical map. Computers take the worry out of determining scale accuracy. Because of the vast computational power of a computer, there is no difficulty in scale conversion. The days of: "Let's see, this distance is 5.3 inches on the map, and one inch is 12 miles, so the distance is about 64 miles" are over. Scale is a minor concept in GIS—one used only on final output.

Geographic vs. Projected Coordinates: A Comparison

- Advantages of the spherical coordinate system: You can represent any point on the Earth's surface as accurately as your measurement techniques allow. The system itself does not introduce errors.
- Disadvantages of a spherical coordinate system: You will encounter complex and time-consuming arithmetic calculations in determining the distance between two points or the area surrounded by a set of points. Latitude-longitude numbers plotted directly on paper in a Cartesian coordinate system result in distorted—sometimes greatly distorted—figures.
- Advantages of a projected coordinate system on the Cartesian plane: Calculations of distances between points are trivial. Calculations of areas are relatively easy. Graphic representations are realistic, provided the area covered is not too large.
- Disadvantages of a projected coordinate system on the Cartesian plane: Almost every point is in the wrong place, though maybe not by much. All projections introduce errors. Depending on the projection, these errors are in distances, sizes, shapes, or directions.

Matching parameters is paramount in combining GIS data sets![13]

[13]The parameters of data sets may not always match but may be close enough so that any error introduced is trivial. In many instances and places, for example, NAD 83 and WGS 84 match within centimeters. But you must research carefully to know when you can use data sets that don't match exactly. The ESRI manual *Understanding Map Projections* can help.

Two Projected Coordinate Systems: UTM and State Plane

A coordinate system called Universal Transverse Mercator[14] was developed based on a series of 60 projections onto semi-cylinders that contact the Earth along meridians. (To consider, for example, one of these projections, imagine a sheet of paper bent so that it becomes a half cylinder. Then with the axis of the cylinder oriented in an east-west direction—hence, the term *transverse*—the paper is brought into contact with a globe along the meridian designating 3° longitude. Then the surface of the Earth between 0° and 6° is projected onto the paper). UTM projections are further subdivided into areas, called *zones*, covering 6° of longitude and, for most zones, 8° of latitude. A coordinate system is imposed on the resulting projection such that the numbers

- ❑ Are always positive
- ❑ Always increase from left to right
- ❑ Always increase from bottom to top

The representation of our previously discussed object (at 38°N and 84.5°) in the UTM coordinate system, when that system is based on WGS84, is a "northing" of 4,208,764.4636 meters and an "easting" of 719,510.3358 meters. The northing is the distance to the point, in meters[15], from the equator, measured along the surface of an "Earth" that has no bumps. The easting is somewhat more complicated to explain because it depends on the zone and a coordinate system that excludes negative numbers. Consult a textbook on geodesy or cartography, or review the thousands of Web pages that come up when you type the words

```
UTM "coordinate system" "Transverse Mercator"
```

into an Internet search engine (e.g., www.google.com).

A version of the UTM coordinate system is based on NAD27. In this case, our object would have different coordinates: Northing 4,208,550.0688 and easting 719,510.6393. This produces a difference of about 214 meters. If you combined WGS84 UTM data with NAD27 UTM data, the locations they depict would be close enough to cause trouble.[16]

Each state of the United States has State Plane Coordinate Systems (SPCSs) developed, originally in the 1930s, by the U.S Coast and Geodetic Survey. These systems are based on different projections (usually, Transverse Mercator or Lambert Conformal Conic), depending on whether the state is mostly north-south (like California) or mostly east-west (Tennessee). The units represent international feet, survey feet, or meters, depending on decisions made by the state itself.[17] Zone boundaries frequently follow county

[14]The idea of the Mercator projection was developed in 1568 by Gerhardus Mercator a Flemish geographer, mathematician, and cartographer.

[15]In fact, the meter was originally defined as one ten-millionth of the distance from the equator to the North Pole, along a meridian that passed through Paris.

[16]Comparing these coordinates with the WGS 84 UTM coordinates, you see that virtually all of the difference is in the north-south direction. While true for this particular position, it is not true in general.

[17]The meter is the standard unit of length in most places in the world. Two different lengths of "foot" are defined in terms of the meter. A "U.S. Survey Foot" is one in which a meter is considered to be 39.37 inches, exactly; the other sort of foot is the international foot, where an inch is 0.0254 meters, exactly. The survey foot and the international foot are almost, but not exactly, the same length.

boundaries. The coordinate system(s) used in one state are not applicable in neighboring states. Nor can you apply the SPCS of one zone to areas in the state only a short distance away in another zone. Furthermore, the difference between NAD 27 and NAD 83 (WGS 84) can be startlingly large. In Kentucky, for example, 38.0000000° (North) and longitude 84.5000000° (West) would translate into a northing of 1,568,376.1900 feet and an easting of 182,178.3166 feet when based on WGS84. However, when the basis is NAD27, the coordinates are 1,927,939.8692 and 182,145.9821, which makes a difference of some 68 miles!

Why the large differences in projected coordinate systems based on NAD27 versus those based on WGS84? Because those responsible for the accuracy of other coordinate systems took advantage of the development of WGS84—a worldwide, Earth-centered, latitude-longitude system—to correct or improve those projected coordinate systems.

State plane coordinate systems are generally designed to have a scale error maximum of about 1 unit in 10,000. Suppose you calculated the Cartesian distance (using the Pythagorean theorem) between two points represented in a state plane coordinate system to be exactly 10,000 meters. Then, with a perfect tape measure, pulled tightly across an idealized planet, you would be assured that the measured result would differ by no more than 1 meter from the calculated one. The possible error with the UTM coordinate system may be larger: 1 in 2500.

Coordinate Transformations

Coordinate transformation of a geographic data set is simply taking each coordinate pair of numbers in that data set and changing it to another pair of numbers that indicates exactly the same spot on the Earth's surface, but using a different system of assigning coordinates.

Let's make up a coordinate system for a garden delineated by lines between stakes. Suppose the origin (0,0) is at the southwest corner. Now drive several stakes in the ground so that if you passed a string around them they would outline a polygon. Suppose you use a surveyor's tape calibrated in survey feet to measure the Cartesian coordinate for each stake. Stake 0 is at the origin (0,0). Stake 1 is 33 feet east and 0 feet north. (i.e., 33,0). Stake 2 is at (33,10). And so on. See Figure 1-8

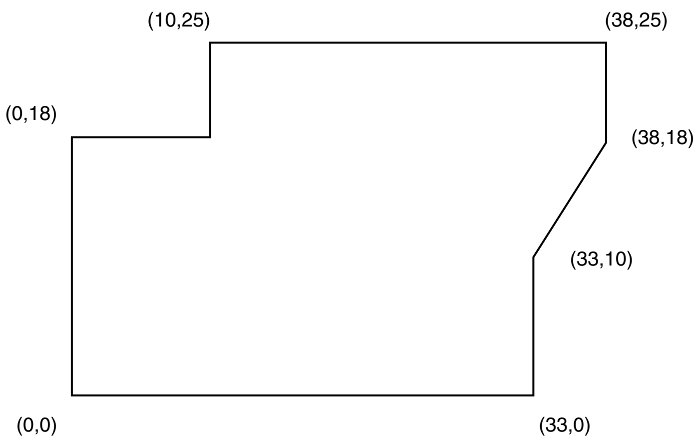


FIGURE 1-8 Locations of garden stakes (coordinates in feet)

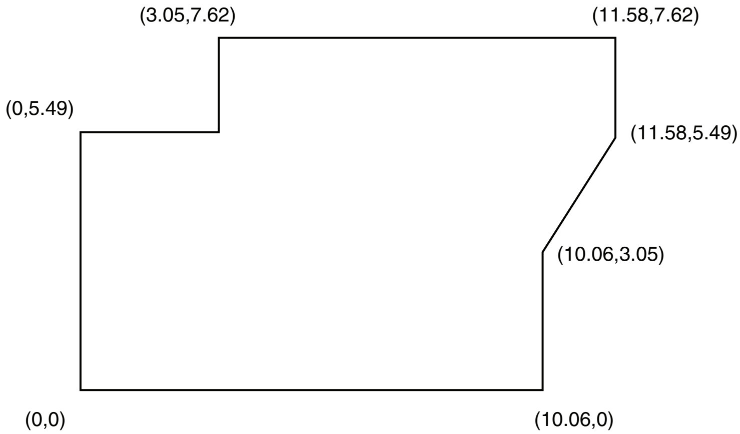


FIGURE 1-9 Locations of garden stakes (coordinates in meters)

The garden grows and becomes overgrown. Next year someone locates the origin and wants to find the locations of the original stakes. He also has a surveyor's tape, but it is calibrated in meters instead of feet. He asks you to provide the data on where the stakes are in meters. In this case, the transformation of coordinate systems is easy: You simply multiply each number by a conversion factor (the number of meters per survey foot, which is 0.3048). So Stake 1's coordinates are (10.06, 0). Those of Stake 2 are (10.06, 3.05). And so on. See Figure 1-9.

The stakes are in the same place that they were before. Their positions are simply referred to by a different set of numbers. Geographic coordinate or datum transformation is nothing more or less than this, except that the mathematical operations applied to each number are usually more involved.

Physical Dimensionality

All matter exists in four dimensions—roughly characterized as left-right, toward-away, up-down, and time. The first three are conceptually clumped together and called spatial dimensions. Time is treated as a dimension here because we want to talk about "position" in space and time.[18]

Any physical object occupies space and persists in time. This is true of the largest object in the universe and the smallest atom. However, when the measure of one or more dimensions of an object is small, or insignificant, with respect to the measure of others or is tiny with respect to its environment, it is useful to describe and represent objects by pretending that they occupy fewer than three spatial dimensions. For example, a sheet of paper can be considered a pseudo-two-dimensional object; it has thickness (up-down), but that dimension is miniscule compared with left-right and toward-away. A parking meter in

[18]In physics, time is treated as the fourth dimension because it is inexorably bound up with the other three according to Einstein's theories (special in 1905 and general in 1916) of relativity. This connection need not concern us except for some esoteric technical matters related to time on global positioning system satellites (for one, time on a satellite runs faster because it is farther from Earth, where there is less gravity).

a city could be considered a pseudo-zero-dimensional object, because the measures of all its dimensions are insignificant with respect to its environment. (The parking meter would certainly not be considered a spatially zero-dimensional object by its designer, manufacturer, or the driver who uses it. Therefore, the pseudo-dimensionality of an object depends on its context.)

It is important to consider the pseudo-dimensionality of both an object and the field (the space) in which it resides. A point on a line is a zero-dimensional object in a one-dimensional space. A line segment on a plane is a one-dimensional object in a two-dimensional space. A polygon is a two-dimensional object in a two-dimensional space. A point in a volume is a zero-dimensional object in a three-dimensional space. The dimensionality of the object must be less than, or equal to, the dimensionality of the space in which it resides.

This issue of dimensionality comes into play here because great economies of computer storage are achievable if an object is considered to have fewer dimensions than it actually does. For example, describing a fire hydrant as a complex three-dimensional spatial object would be an arduous task and involve many numbers and much text. If it is considered a zero-dimensional object, its location can be described (very precisely, but still inexactly) with three numbers, perhaps representing the latitude, longitude, and altitude of some point on the hydrant or, perhaps, its center of mass.

The spatial dimensions of a spatial pseudo-zero-dimensional object are insignificant with respect to the context or environment in which it resides. A theoretical geometric point is the prototypical example. Examples of such zero-dimensional objects on maps or in a GIS could be streetlights, parking meters, oil wells, census tracts, or cities—depending on the spatial extent of the other objects or features in the database.

A spatial pseudo-one-dimensional object or feature has no more than one significant dimension or is made up of essentially one-dimensional objects. A straight-line segment is the prototypical example. In terms of making a one-dimensional object up from component parts, consider the example of a telephone wire strung from one pole to the next to the next and so on. It is considered a one-dimensional object, even though the poles may not be in a straight line, so that the phone line zigzags over a two-dimensional field. (The line itself sags and exists therefore in a three-dimensional context; but these facts notwithstanding, it can be simplified or generalized into a one-dimensional feature—saving us lots of computer storage, processing time, and conceptual complication.) Roads, school district boundaries, pipes for fluid transportation, and contour lines on a topographic map are all examples of one-dimensional entities.

You can disregard the up-down dimension in a spatial-pseudo-two-dimensional object. Plain areas such as voting districts and soybean fields are examples. Such pseudo-areas lack vertical components—variability in altitude—that can be important. The true surface area of hilly terrain may be considerably underrepresented by the plane area within the borders of the plane figure that represents it.

You must consider all dimensions of a spatial three-dimensional object in order to represent it. Three-dimensional features are volumes—say, a coal seam or a building.

Relative to each other, positions of features that we record with a GIS do not move, or do not move quickly, with respect to each other. We want to know how conditions change or objects move. The representation of the objects may be zero-, one-, two-, or three-dimensional. At its most complex, GIS would involve all four dimensions. Most GIS operations and data sets are two-dimensional; the data are assumed to pertain to moments in time, or stable phenomena or conditions over a period of time. Three-dimensional GIS could involve either three spatial dimensions (such as representing the volume of a limestone quarry) or two spatial dimensions and varying time (showing historical changes in a landscape). Rarely will you use four-dimensional GIS.

Results from a GIS are always only an approximation of reality. One of the reasons, among others to be discussed later, is that we simplify objects by reducing their dimensionality.

Global Positioning Systems

A global positioning system (GPS) is a satellite-based system that provides users with accurate and precise location and time information. Using NAVSTAR GPS, you can determine locations on Earth within a few meters, and, with more difficulty and expense, within a few centimeters or better. Timing within 40 billionths of a second (40 nanoseconds) is easily obtained. Timing within 10 nanoseconds is possible.

The U.S. Department of Defense operates NAVSTAR GPS in cooperation with the U.S. Department of Transportation. The acronym stands for "Navigation System Timing and Ranging Global Positioning System." Informally, it is Navigation Star.

The Russian GLONASS (Global Navigation Satellite System) operates similarly. Concerns about U.S. control over NAVSTAR led Europe to begin development of its own independent Galileo system in 2002.

A GPS receiver, which "remembers where it has been," is becoming a primary method of providing data for GIS. For example, if you drive a van with a GPS antenna on its roof along a highway, recording data every, say, 50 feet, you will develop an accurate and precise map of the location of the highway. The NAVSTAR GPS is discussed in Chapter 5.[19]

Remote Sensing

Remote sensing probably started with photographs taken from balloons in the 1840s. The first automated system (not requiring humans to be with the sensors) may have been in the 1890s in Europe when cameras programmed to take pictures at timed intervals were strapped to pigeons!

Evelyn Pruitt probably introduced the modern use of the term "remote sensing" in the mid-1950s when she worked as a geographer/oceanographer for the U.S. Office of Naval Research (ONR). Remote sensing uses instruments or sensors to capture the spectral characteristics and spatial relations of objects and materials observable at a distance, typically from above them. Using that definition, everything we observe is remotely sensed. More practically, something is sensed remotely when it is not possible or convenient to get closer.

We can categorize remote sensing for GIS many ways. Data can be taken from aircraft or satellite "platforms." The energy that the sensor "sees" can come from the objects or areas being examined as a result of radiation emanating from them (caused by the sun or other heat or light sources) or from radiation bounced off them by an energy source associated with the sensor (e.g., radar). The images produced may be developed on film or produced by digital sensors. Satellites in geosynchronous orbits (hanging in a single spot over the equator) or near polar orbits can see different areas of the Earth as it turns. Chapter 2 offers examples of remotely sensed data.

---

[19]See also the author's textbook *The Global Positioning System and GIS*, 2002, Taylor & Francis, 375 pages.

Relational Databases

GIS relies heavily on databases of text and numbers. The relationship between such information and geography is discussed later in the chapter. For now, you simply need to know how text and number data sets are stored.

For this discussion, a *database* is a collection of discrete symbols (numbers, letters, and special characters) located on some physical medium with at least one principal underlying organization or structure. An old-fashioned library card catalog is an example of a database with a single underlying structure: an alphabetical list of authors; the medium is 3×5 index cards; and the data are the symbols on the cards describing books and their locations in the library. Many libraries have substituted computer-based catalogs with the advantages of searching and finding not only authors but titles and subjects as well, so a number of organizing themes may underpin a database.

Another example is a reel of magnetic computer tape that has recorded the most common type of soil found in specific acreage in a county. The tape is the physical medium, the codes assigned to a soil type constitute the data, and the location of each acre—as understood from the position of each datum on the tape—could be the underlying structure.

Existing general purpose databases usually:

- Result from some sort of project; some individual or team constructs it—frequently going to considerable effort.
- Need to be *updated* (modified and corrected as time progresses) if they are to continue to be of value.
- Contain errors *regardless* of size, care of construction, simplicity of data, or quality of physical medium used.
- Serve a function when allied with some process. The function may be as simple as supplying a telephone number (structure—alphabetical by name; medium—cheap bound paper; data—phone numbers, written quite small; process—looking up a name, finding the adjacent number). The function served by the database might also be quite sophisticated—supporting far-reaching land use decisions, for example.

Numerous databases are used to solve problems at all levels of government. Access to these databases is achieved through the use of referencing schemes. Following are some examples:

Referencing Scheme	***Examples of Data Contained***
Names of people	Salary, social security number, medical history, criminal record
Auto license plate number	Color, owner, serial number
Street address	House value, lot size
Job title	Person employed, duties, salary
Transaction number	Money received, paid, transferred, invested
Events	Schedules, orders, crimes, accidents

Many schemes exist for presenting or storing data. Suppose parents want to find a name for a newborn child. A list of potential names in random order might be provided. It would probably be more useful if the list were divided into girl's names and boy's names. Another improvement could be to alphabetize the list, or present it in terms of current popularity of names.

As a second example, consider how you might store telephone numbers associated with names and addresses. You could order this list alphabetically so you could easily find a phone number, knowing the name. Or you could order the list numerically by telephone number so you could find a name, given the number.

A different sort of data structure is "hierarchical." You could use this approach to store the names and positions of people in a corporation or the military according to who reports to whom. Likewise you could store voting districts within cities within counties within states. In the computing world this "tree-structured directory" approach is used to organize the folders and files contained on a hard disk drive on a computer. Any folder on the disk may store other folders and files.

Many schemes exist to store information in the memory of a computer or on its secondary devices, such as disk drives or tapes. The primary method used to store large amounts of information is called the "relational database," or RDB, developed by E. F. Codd.[20] The software is described as a RDBMS (Relational Database Management System). The idea is simple: use a two-dimensional table format in which the rows relate to entities (objects, people, things in general), while the columns relate to attributes (characteristics, properties) of entities. The intersection of a given row and a given column is a cell, containing a value, which defines the particular attribute of the particular entity. See Figure 1-10.

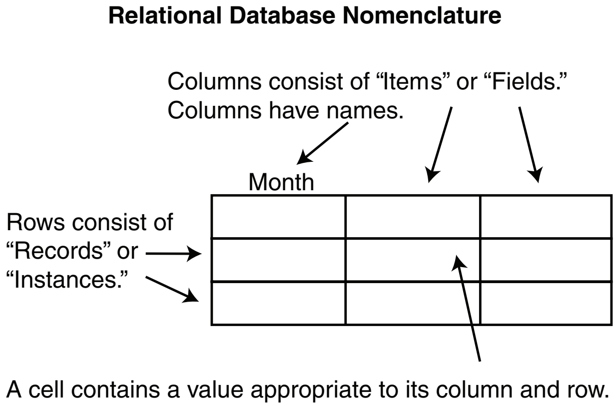


FIGURE 1-10 Components of a relational database table

You can use a database to store names, occupations, and pay schedules of employees, as in Figure 1-11.

Another example shows part of a database of automobiles registered in a state. Each row would represent one car; each column would represent one property of cars. See Figure 1-12.

Here is some terminology: The structure that contains the entity, the row, is also called a *tuple* or *record*. The structure that contains the attributes, the column, is also a *field* or an *item*.

[20]Edgar F. Codd was the originator of the relational approach to database management, beginning in 1970. For this work he was the recipient of the Association for Computing Machinery's (ACM) Turing award in 1981.

Garden-Variety Database

Name	Occup	Pay_Sched
Smith	*Welder*	*Hourly*
Jones	*Accountant*	*Bi-Monthly*
Adams	*Attorney*	*Monthly*

FIGURE 1-11 Trivial example of a relational database table

Mfgr	Num_Doors	VIN	Color	Weight
Porsche	*2*	*123XXX*	*Silver*	*2300*
Porsche	*2*	*887ABC*	*White*	*2100*
Toyota	*4*	*9880123*	*Grey*	*2350*
Honda	*4*	*456789*	*Blue*	*2999*

FIGURE 1-12 A RDB table of automobiles

All the cells of a given column must contain the same sort of value. Some common ones are as follows:

- *Character*—Any valid ASCII characters in a string of almost any length.
- *Short integer*—Can range from slightly less than –32,000 to slightly more than +32,000.
- *Long integer*—Can range from slightly less than negative 2 billion to slightly more that positive 2 billion.
- *Floating-point number*—Can have an exponent as small as 10^{-38} and as large as 10^{38} and you can count on six significant digits.
- *Double-precision floating-point number*—Can have an exponent as small as 10^{-308} and as large as 10^{308} and you can count on 16 significant digits
- *Boolean*[21]—A value that is either true or false.

[21]Named after George Boole, mathematician.

Getting Information from a Relational Database: Queries

Relational databases are designed to give you information. You can obtain the information by selecting a subset of records from the total set by writing an expression that is a mixture of attribute names, arithmetic and logical operators, and values.[22] For a trivial example, suppose research has found that gray cars that weigh less than 2,500 pounds put their occupants at greater risk than average. You want to select those records from the statewide automobile database. You might first get all the records of the cars that are gray.

```
SELECT: COLOR = 'Gray'
```

Given that subset, you might then write

```
SELECT: WEIGHT < 2,500
```

Given this sub-subset of records, you could perhaps write letters to the owners of those cars, making them aware of the danger they face.

Languages to get subsets of records can provide flexibility and efficiency. For example, to do both of the preceding operations with one expression, you might write

```
SELECT: COLOR = 'Gray' AND WEIGHT < 2500
```

Suppose also that the research study showed increased danger to those occupying cars that were built in 1985 or before, regardless of color or weight. You might add to the selection above by saying:

```
ADDSELECT: YEAR <= 1985 (ADDSELECT means add to the current set of selected
records.
```

Or you may use a single query to select all the records you want at once:

```
SELECT: (COLOR = 'Gray' AND WEIGHT < 2500) OR YEAR <= 1985
```

Note the use of parentheses to indicate the order in which operations are done.

Economies in Relational Databases

In theory, the information in a relational database system could reside in a single table. This isn't the best policy and in reality may not even be possible. A relational database usually consists of a set of tables that relate to one another—thus the word "relational."

Each relational database table must contain a column that has a unique identifier for each record in the table. This is known as the key field. If the relational database references people, the key might be social security number, employee number, or student number. In the case of automobiles, the key code might be the vehicle identification number (VIN).

---

[22]There are many relational database software products and several different languages used to query them. What is shown in the following is not a specific query language but rather a generic representation of such a language. ArcGIS may be used with several relational database products, each with a somewhat different query language.

To illustrate why multiple tables might be used, suppose in our relational database it is possible for one person to own more than one automobile. If the entire database is all in one table, then the names and addresses of each multicar owner must be repeated, which increases the amount of storage required. If a multicar owner gets a new street address, several records in the database table would have to be changed. Further, some of the data about owners might be located in another database and might be private. The elegant answer to these problems is to have several tables that contain the information, divided to minimize the repetitions. In the automobile example, it might be that only the owner's identification number is stored in the record of the car itself. This number could be the key field in the database table containing information about the owner. If some of the owner information is private, it could be stored in a separate table as well, with the same key field. The database could be set up so that this table could not be accessed by the automobile table.

Relational database tables that meet certain requirements of efficiency are said to be in First Normal Form (1NF), Second Normal Form (2NF), and so on. The higher the number, the more efficient the database.

In summary, it is useful to partition the information in a relational database into a number of tables. Such partitioning can:

- Reduce the amount of redundant data stored.
- Aid in updating the database.
- Reduce the chances of inconsistency and instability in the database.
- Aid in protecting private or sensitive data.

Languages have been developed for retrieving information from a relational database. One is the Structured Query Language (SQL), developed by the American National Standards Institute (ANSI). You use it to execute queries of a database. Vendors of relational database management systems also have developed proprietary languages.[23] ESRI products interface with a number of RDBMS from various sources.

Relational Databases and Spreadsheets

A relational database table may look a lot like an electronic spreadsheet, such as Microsoft Excel. There are some important differences:

- A relational database is structured, with rows and columns strictly defined. In a spreadsheet you can put anything anywhere.
- In a relational database, the headings of columns are not stored in cells of the relational database. The column headings (attribute names) are known to the database software and are displayed, but are not part of the data. In a spreadsheet, column headings occupy cells.
- A relational database is a logical object about which conjectures can be made and theorems proved. An entire branch of computer science is devoted to work with relational databases.

[23]Some names of past and present relational database management system software products are Access, dBASE, Informix, Ingres, Oracle, SQL Server, and Sybase.

Another Definition of GIS

Each record in a relational database references something: a person, object, idea, equation, feature, subject—some unique entity. In a GIS, the entity is usually a spatial three-dimensional feature such as a parking meter, lake, railroad, and so on.[24] These 3-D features are almost always reduced to abstract objects of fewer dimensions. The parking meter is a point, the railroad is a set of lines, and the lake is an area bounded by a sequence of lines.

Usually, a record in a relational database references a person or object without respect to its current location. A subject of a record in a relational database frequently moves around—like cars or people. But when the subjects of a relational database are fixed in space, the position, or positions, of the feature may be included in the description of the feature along with the attributes. In one sense, the location of the object becomes one of its attributes.

Points

Here, for example, is a map and part of a relational database that together describe parking meters in a town. Each meter has a unique number and is thus the key field. Parking meters (each of which has a latitude and longitude specification) correspond one-to-one with the records in the database. See Figure 1-13.

Hydrant	Flow	Inspected
383	*High*	*1994*
888	*Low*	*1989*
36	*Medium*	*1998*

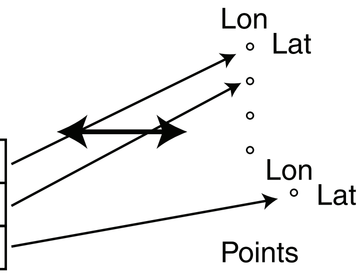


FIGURE 1-13 Marriage of a point geographic database and a relational database table

Let's consider another definition of a GIS: From a computer software point of view, a GIS is the marriage of a (geo)graphical database and an attribute database (frequently a relational database).

Lines

As I mentioned earlier, a GIS can store and analyze pseudo-one-dimensional entities such as roads. Here each RDB table record would refer to segments of a roadway between intersections. The attributes might be number of lanes, highway number, street name, pavement type, length, and so on. See Figure 1-14.

[24]In raster-based GIS, the entity is a set of square areas that share common values. This approach will be discussed extensively in Chapter 8.

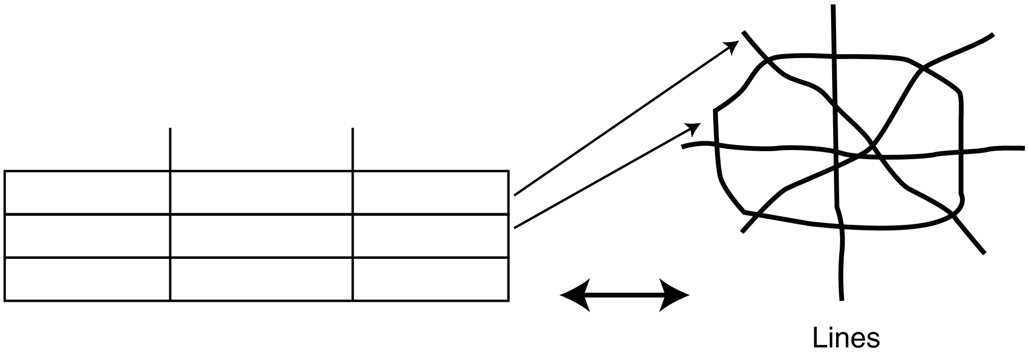


FIGURE 1-14 Marriage of a line geographic database and a relational database table

Again: A GIS is the marriage of a (geo)graphical database and an attribute database.

Areas

A GIS needs to be able to store information about areas, as well as points and lines. For example, such areas might be ownership parcels. Each parcel would be delineated by the lines determined by a surveyor. Those lines define an area. A record would exist in the relational database that would correspond to the area. The attributes in the record might be owner's name, tax identification number, area of the parcel, and perimeter of the parcel. See Figure 1-15.

Did I mention? *A GIS is the marriage of a (geo)graphical database and an attribute database.*

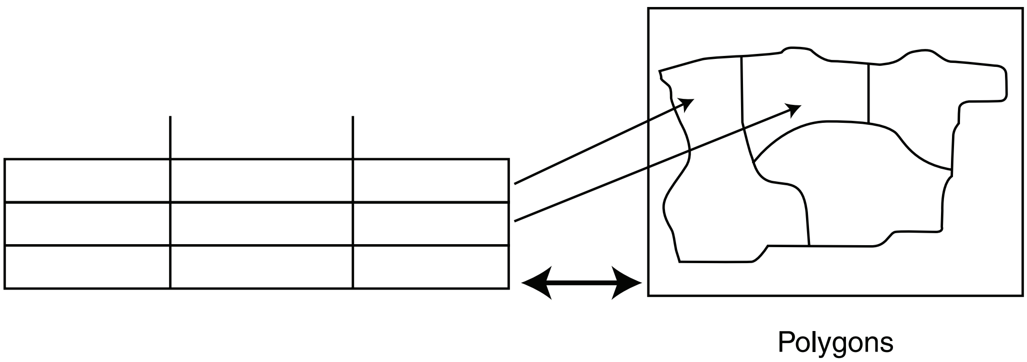


FIGURE 1-15 Marriage of a polygon geographic database and a relational database table

Computer Software: In General

Computers, considered at the most fundamental level, do only three things:[25] get input, manipulate data, and produce output; that is, read bits, stir bits, and write bits. What tells the computer what to do? In the early days of computers, more than a half century ago, each individual instruction to a computer came from outside, one at a time. This process was soon automated so that external media, like perforated paper tape or magnetic tape, contained the instructions. Then several scientists[26] got the idea of placing the instructions in the store of the computer itself. This major breakthrough in computer development allowed the computer to execute one group of instructions and then, on the basis of testing bits in its memory (e.g., is number “a” larger than number “b”?), execute another group of instructions in a different place in the memory. This let the computer simulate “reasoning” and “decision making.” This concept of a *stored program* revolutionized computer use.

The modern computer holds an immense amount of data and stores a large number of programs. If a given program is being used at a given time, it exists in the “fast” electronic memory of the machine. A program or data set not being used usually resides in “slow” disk (usually electromechanical) memory. Dealing with large numbers of diverse elements requires a management scheme. With computers this takes the form of an *operating system* that allows the user to execute specific programs when desired, connects the computer to other computers (e.g., by way of the Internet or other network), makes copies of files, and performs many housekeeping activities.

Writing sets of instructions to be stored in the computer’s memory is known as *programming*. Developing and marketing *programs* (software packages) is an immense business. The operating system that your computer uses is a software package. The word processor used to write this text is a software package. And the GIS you are about to learn, known as ArcGIS from ESRI, is a software package, or rather a suite of software packages.

Computer Software: ArcGIS in Particular

ArcGIS is an integrated GIS that consists of three principal parts:

> *ArcGIS Desktop* software, which is an integrated suite of advanced GIS applications
>
> ArcSDE *Gateway*, which is an interface for managing geodatabases in a database management system (DBMS)
>
> ArcIMS software, an Internet-based GIS for distributing spatial data and services.

ArcGIS has software extensions such as ArcPad, which provides GIS capability for palmtop or handheld computers running under the Windows CE operating system.

ArcGIS is called a *scalable product* because it allows for deployment of GIS of many different “sizes.” A user can choose to have only a system running on a single computer—a *personal GIS*—that allows viewing and simple selection of spatial data without capabilities of editing or extensively analyzing that data. At the other end of the scale, a *multiuser GIS* can serve an entire company or governmental agency. This *enterprise GIS* allows all the capabilities of the software. Options also include systems of intermediate functionality known as *workgroup GIS* and *departmental GIS*.

---

[25]Fundamentals that we’ll discuss further in Chapter 6.
[26]John von Neumann, J. Presper Eckert, John Mauchly, Arthur Burks, Maurice Wilkes, and others.

For many products, the price charged depends on the features and utility that the product provides. When one manufacturer produces many different products, users must learn how to combine and package them to provide the correct degree of utility required (you may not need a school bus if a sedan will do). This is also true of ESRI products.

Software Focus of This Textbook

In this introductory text, we confine ourselves largely to ArcGIS Desktop. The other two products pertain primarily to managing GIS data for large applications or organizations and to distribute GIS capability and data to users over networks. Our emphasis is rather on the analysis, synthesis, and display of spatial and attribute information.

ArcGIS Desktop is itself divided in a number of different ways. It has three main software packages:

- ArcCatalog
- ArcMap
- ArcToolbox

ArcCatalog is mainly an operating system for geographic data. ArcCatalog allows you to set up shortcuts to reach particular files of data quickly and easily and examine those files visually and textually. You can use ArcCatalog to rename and copy spatial data sets.

ArcMap lets you put together graphic and geographic elements to make sophisticated maps and interact with them. It allows you to obtain information from those maps and the attributes associated with their components, by using a variety of processing methods.

You use ArcCatalog or ArcMap to access Arc Toolbox, which provides immense geo-processing and analysis capability. It has tools to do 3-D analysis, spatial analysis, and spatial statistical analysis, to convert from one spatial data paradigm to another, and to convert spatial data into a myriad of geographic projections.

These three packages are fundamental to ArcGIS. But you have probably noticed many other Arc*Xxxxxx* terms floating around. As a scalable product, ArcGIS desktop allows you to purchase what you need. The different levels of capability are bundled into four different packages. In terms of increasing utility (and cost) they are as follows:

ArcReader (free)
ArcView
ArcEditor
ArcInfo

Knowing the capabilities of each is not important for our immediate purposes. You should know that some of these terms are recycled from earlier ESRI product names. In particular, ArcView 9 is a different animal than ArcView 3.x. ArcView 9 is a level of capability of ArcGIS Desktop; ArcView 3.x is an older, but still supported and useful, independent software package.

CHAPTER 1

Some Concepts That Underpin GIS

STEP-BY-STEP[1]

Exercise 1-2 (Project)

Developing a Fast Facts File for the Information You Learn

To make this textbook work well for you, I strongly recommend that you create and maintain a Fast Facts File—a computer text file in which you can record what you learn so that it is at your fingertips. It will serve you whenever you need to know a particular bit of information or perform an operation. Much of what you do you will put into your own fast memory (contained in your cranium), but there will be a number of facts that your fast memory may not contain when you need them. Here's where the Fast Facts File comes in. It is a computer-based equivalent of a loose-leaf notebook that you continually revise and update. There's where you should put procedures and concepts you might forget after a couple of weeks of doing work other than GIS.

For example, you might use the file to note techniques for changing symbology (colors and symbols that represent features on a map), which are addressed at various different points in the text and in the software. You can note the techniques down as you work with them and reorganize them later. The computer-searchable file helps you find what you need when you need it, even if you fall behind in organizing.

Periodically reorganize your notes. Occasionally print out the file and put it in a notebook. Periodically back up the file onto a flash drive, CD-ROM, or other computer. As you progress through this text, you will develop your own little book that will aid you in this course and thereafter. And, because you write it, it will be organized in a way that meets your needs—both as a tutorial and a reference volume.

___ **1.** Use a text editor (e.g., WordPad) or word processor (e.g., Microsoft Word) to create the Fast Facts File. Put your name and other contact information in it. Initially include your computer account name (not your password—put that somewhere else), how to start various software programs, and so on. The file will evolve as you go on. For the moment, start the file and keep it open. It should at least contain the following:

Name: ____

User or Logon Identifier: ____

Password hint or secure location: ____

Do not save the file yet.

[1]The length of time most students will require to complete the exercises in the Step-by-Step sections can be found in the Preface.

Understanding the File Structure for the Exercises

You will be working primarily with two major folders or directories. The first of these folders will be IGIS-Arc. It will contain the data for all the exercises. It will start off as, and will remain, a direct copy of the IGIS-Arc folder that is on the CD-ROM in the back of the book. If you are in a class, operating off a network, your instructor will probably have put this folder somewhere in the network file structure. If you are using this book on your own, you should load the IGIS-Arc folder that is on the CD-ROM directly onto your local hard drive. In either case, the IGIS-Arc directory should be protected so that changes cannot be made to it.

___ **2.** Locate or create the IGIS-Arc folder. It may be in the root folder of a hard drive on your computer, or it may be several levels down in a hierarchy. Whenever it is referred to in this textbook, it will be represented as:

`[__] IGIS-Arc`

The symbols [__] might represent something as simple as C:\

Or it might be a long path such as:

`U:\ABCNet\GIS_Students\GIS401`

Write the path associated with [__], in your computer system, below:

[__] means ____

The second folder you will use to store the work that you do. It will be specifically yours; your initials will be made part of the folder name. For example, if your name were John W. Stephenson, the folder would be called IGIS-Arc_JWS. In the next step, you will create this folder.

___ **3.** Decide where on the computer or network you want to store your work. Use the operating system of the computer to make a new folder by navigating to and selecting the appropriate drive or path, clicking the File menu, and clicking Folder when it shows up under New. Once the folder is created, change its name to

IGIS-Arc_*YourInitialsHere*

(for example, IGIS-Arc_JWS).

In this textbook the simple blank, ___, will be the disk drive or path to the folder where you keep your work.

Write the path associated with ___ below:

___ means ____

___ **4.** Save your Fast Facts File in the folder you have just created, giving it the name `FastFactsFile`. In other words, the full path and filename for your new Fast Facts File will be as follows:

___IGIS-Arc_*YourInitialsHere*\FastFactsFile.someextension

To recap: when I use the designation

[___]IGIS-Arc

(note the square brackets), I mean the place where the data sets on the CD-ROM have been placed. You should not change any of these folders or files. When I use the designation

___IGIS-Arc_*YourInitialsHere*

(note the lack of square brackets and the presence of *YourInitialsHere*), I am referring to the location that contains your work.

Exercise 1-3 (Minor Project)

Getting Set Up with ArcGIS

A click or two of the mouse should give you access to ArcGIS. This should provide easy access to three major components: ArcCatalog, ArcMap, and Arc.

You may reach a component by clicking an icon on the desktop, or by navigating to them by clicking Start, then Programs>, then ArcGIS, then the component—or via some other way prescribed by your instructor.

___ **1.**[2] Find the name or icon for ArcCatalog, *right*-click it and choose Properties. In the ArcCatalog `Properties` window, the `Target` should be something like:

C:\Program Files\ArcGIS\Bin\ArcCatalog.exe Is it? __________. If not, write of the target here: ____

Find and write the targets for ArcMap and Arc (located under ArcInfo Workstation).[3]

____

____

___ **2.** Start ArcCatalog.

ArcCatalog serves as sort of an operating system for GIS. You manage data with it. The Step-by-Step part of this chapter is largely an introduction to Arc Catalog.

___ **3.** Dismiss Arc Catalog by selecting `Exit` from its `File` menu.

___ **4.** Start ArcMap. Agree to start using ArcMap with `A new empty map`.

[2]The line to the left of each step number provides a space where you can put a check mark when you have completed the numbered step.
[3]As you work through the text you will find blanks in which to enter requested information. Being able to do this lets you know you are on the right track. Rather than writing in the book you might want to use notebook paper or type into a document file.

ArcMap is the software package that performs primary and major GIS operations, such as making maps and analyzing spatial data. The Step-by-Step part of Chapter 2 serves as the introduction to ArcMap. ArcMap and ArcCatalog both allow you to get to ArcToolbox, which is used primarily for data conversion and many more advanced GIS operations.

____ **5.** Dismiss ArcMap by clicking the “X” in the red box in the upper right corner of the window.[4]

____ **6.** Start Arc, which is one of the modules within ArcInfo Workstation.

The programs Arc, Arcedit, Arcplot, Grid, and some others are instructed through text commands that are typed by the user after a prompt (such as `Arc:`). Contrast this with most familiar programs in which you “point and click.” Typed input from a command line directed the most powerful ESRI software prior to ArcGIS version 8. The Graphic User Interfaces (GUIs) of ArcCatalog and ArcMap are relatively efficient and elegant ways to invoke Arc and do serious GIS work.

____ **7.** Leave the Arc program by typing `Quit` at the `Arc:` prompt.

Exercise 1-4 (Project)

Looking at the ArcCatalog Program

ESRI has developed an entire product that is basically designed to help you find, select, understand, and manage geographic data files. This product is ArcCatalog. You’ll begin by starting this software and exploring it. Then, considering the definition of GIS—that it is an information system (IS) whose database is a marriage of a geographic database (GDB) and a relational database (RDB)—you’ll use ArcCatalog to look at a ridiculously simple GIS data set. A village has developed a system to help with town planning and maintenance. One part of the GIS is a set of features consisting of fire hydrants. You will find the Village fire hydrants theme, named HYDRANTS and explore that data set.

Anatomy of the ArcCatalog Window

Assuming your computer is on and some version of Microsoft® Windows® (or UNIX or LINUX) is running:

____ **1.** Start the ArcCatalog component of ArcGIS. Make it occupy the full screen by double-clicking the ArcCatalog title bar. Click the word `Catalog` at the top of the left-hand subwindow (pane). If you see any negative (minus) signs in the left pane, click them so they become positive (plus) signs. You should now see a window similar to Figure 1-16. That left pane is called the Catalog Tree. At the top level, it is a view of the slow-speed storage devices (hard drives, network drives,[5] CD-ROMs, and so on) of the computer, plus some other entries to be explained later.

---

[4]You dismissed ArcCatalog in a different way. When there is more than one way to accomplish something (almost always the case), I will usually show you alternatives, just by having you use them. Note them in your Fast Facts File.
[5] The network drives may or may not show up, depending on how your system is configured. Even if they don’t appear here, ArcCatalog has a way to make them accessible to you, which you will learn shortly.

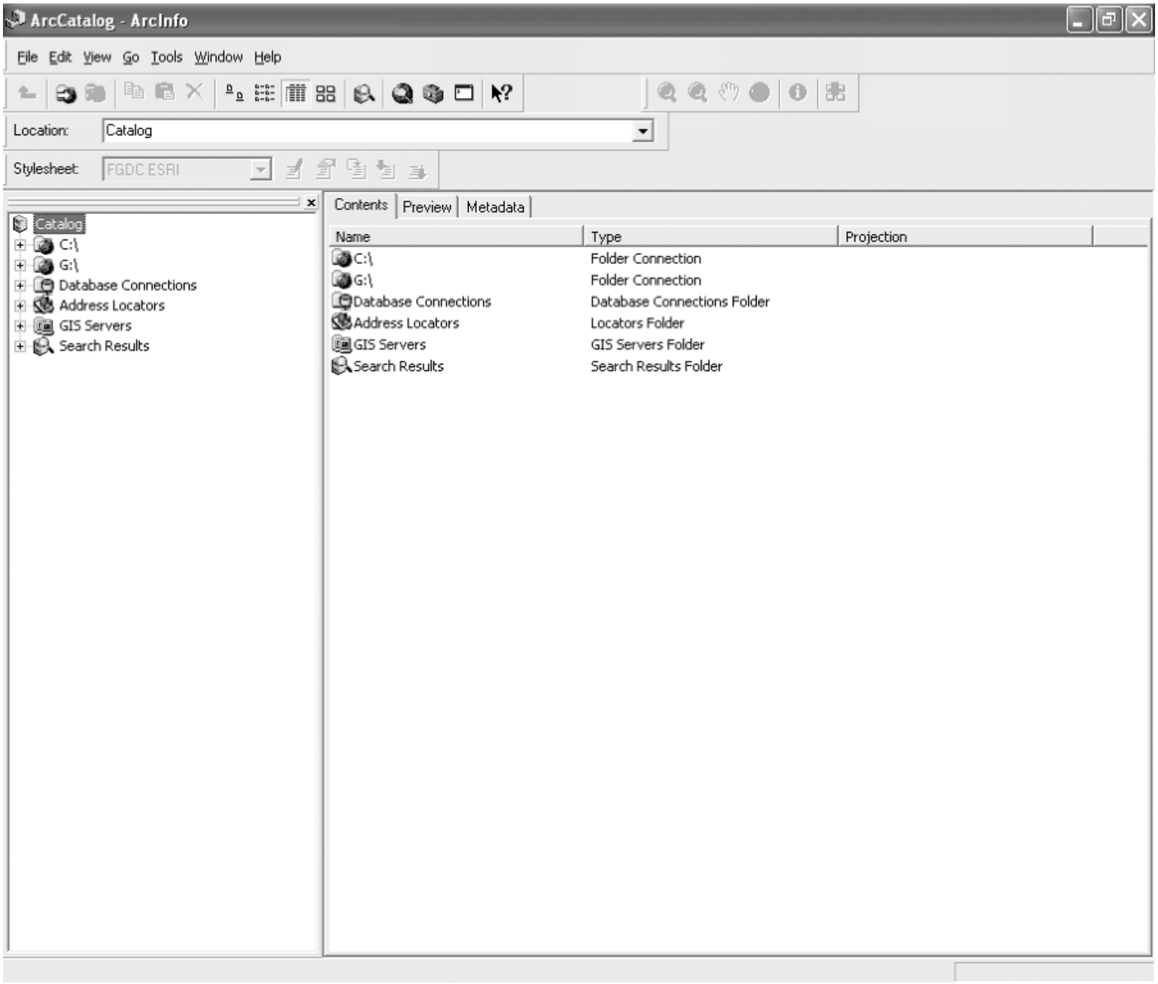


FIGURE 1-16

Near the top of the Catalog Tree you will find the designation for the root folder of the C drive, designated as C:\. Click the folder icon to select it. At the very top left, in the title bar of the overall window, will be some text. What does it say? (Fill in the blank.)

ArcCatalog -__________________ - C:\

Examining the upper left corner of an Arc software window is a way to know (1) which software component you are using (ArcCatalog, ArcMap, and others), and (2) the license your computer is operating under: ArcView (the least expensive with the least capability), ArcEditor, or ArcInfo. Many of the exercises in this text require ArcInfo to run through completely. However, for those of you who have only ArcView or ArcEditor, files that ArcInfo would have developed have been included on the CD-ROM so you can continue past the spots that require ArcInfo.

The title bar of the ArcCatalog window also shows you the path (disk drive and folder) that is currently being referenced by the software (in this case it is the root folder of the C drive, which is C:\). This information also appears in the `Location` text box, which, unless it has been moved or removed, may be found near the top of the window.

2. Now examine the ArcCatalog window.

Many of the menus, icons, and buttons of the ArcCatalog window may be familiar to you from your work with other software. Assuming that your ArcCatalog window has not been customized, you will find the menu headings File, Edit, View, Go, Tools, Window, and Help.

Setting Some Options[6]

3. Be sure that the "C" drive icon is selected. Under `File`, click `Properties`, then click the `General` tab. (From now on I may abbreviate an operation like this by saying `File > Properties > General`.) Since C:\ is selected, you will see some of the properties of the C drive including the amount of disk space—both used and free. See Figure 1-17. Dismiss the `Properties` window by clicking the "X" in its upper right corner.

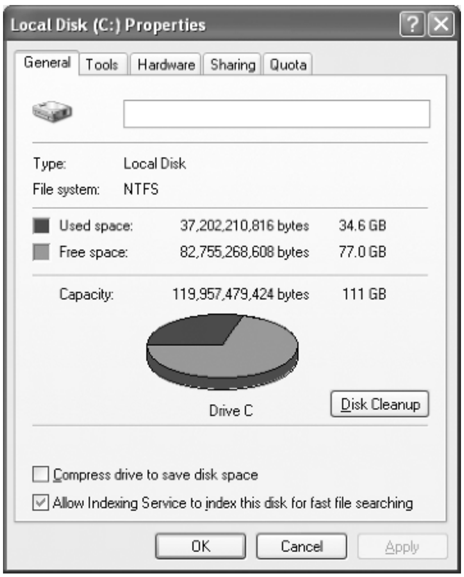


FIGURE 1-17

[6] In general, even though you set options, you cannot count on those remaining set if you terminate ArcCatalog and restart it later. You may have to return to this procedure to reset options.

___ **4.** Click `Catalog` at the top of the left pane. Again click `File > Properties > General`. A bewildering display of possible check boxes and text will show up in an `Options` window with two scrolling panes. These relate to the types of top-level entries and the types of data that the Catalog displays. Since you are at an early-learning stage, you want to have ArcCatalog display everything possible. By clicking the boxes next to the options in the upper and lower panes, you can toggle a check mark on or off. Make sure each option has a checkmark in the box to its left. Below the lower pane, make sure each of the three options, including `Hide File Extensions`, is unchecked. Click `Apply`. (If you made no changes in the pane, `Apply` will be "grayed out"—a standard feature of ArcGIS software if an action is not possible, not needed, or not appropriate.) The resultant window should look approximately like Figure 1-18.

___ **5.** Click the Contents tab in the Options window. Here you may specify the information you want to see about each spatial data set in the catalog when you ask for Details. Read the title over each pane. In the upper pane, put check marks by `Size` and by `Modified`. In the lower pane put a check by `Projection`. The window should look like Figure 1-19. Click Apply. Click OK.

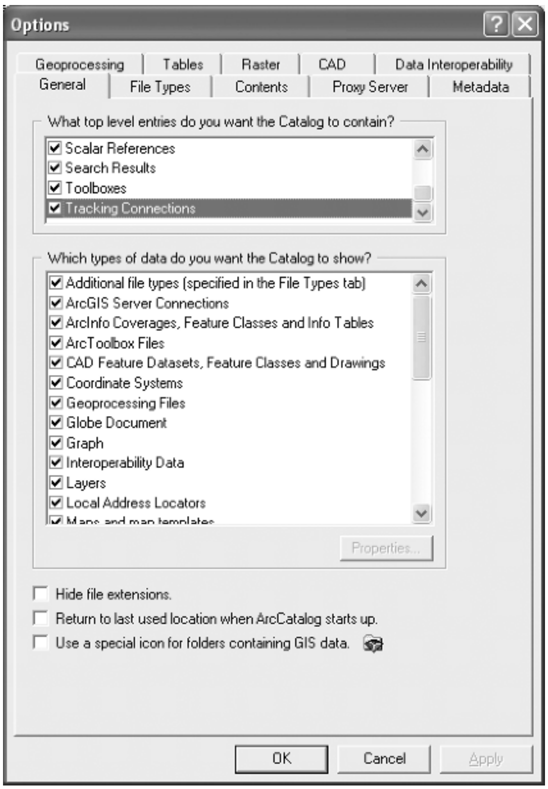


FIGURE 1-18

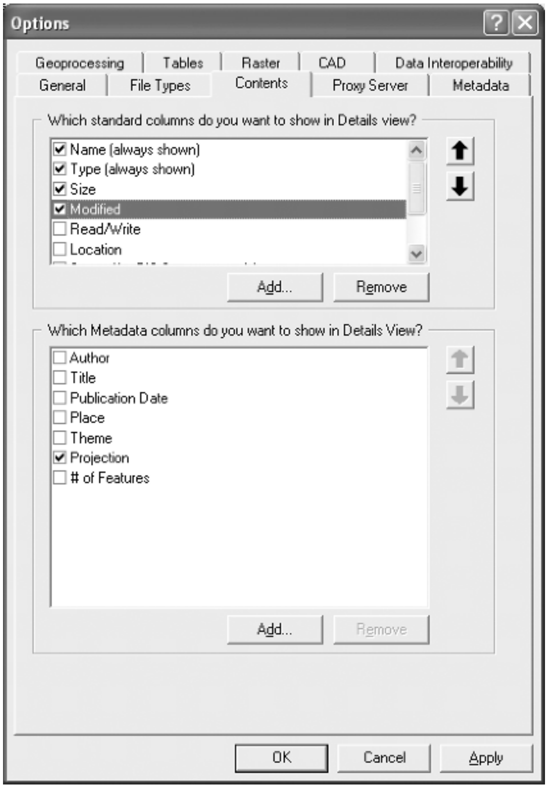


FIGURE 1-19

The Catalog Tree

The main portion of the default ArcCatalog window is usually divided into two panes. On the left you find the Catalog Tree (which you can make disappear and reappear with the Window menu, should you want more visual space for the right-hand pane). The Catalog Tree, when expanded, gives an overall view of the names of your data sets. In this way it resembles what you get when you click My Computer in the Windows Desktop. But you can also look with varying levels of detail at all the data on all the disk drives of the computer by expanding the entries (items)[7] in the tree. (In this way ArcCatalog's presentation is like Windows Explorer.[8] A plus sign (+) indicates that (usually) a given disk drive or folder contains additional folders or files that are hidden from view. A minus sign (–) preceding an entry indicates that any additional

[7]If you are a user of ArcInfo Workstation, or of earlier versions of ArcInfo, you know that the term "item" refers to a column heading indicating a coverage's attributes. Although Arc ESRI manuals refer to a line of text in the Catalog Tree as an "item," we will refer to it as an "entry."

[8]Or like My Computer in Windows XP, when you have checked "Folders" under the Explorer bar in the View menu. That's My Computer > View > Explorer Bar > Folders.

folders or files that are directly contained within an entry are displayed below that entry. A click on a plus sign changes it to a minus and expands the entry; a click on a minus sign collapses the list and shows a plus again.

___ **6.** Press the Contents tab in the right pane, to make it active. In the Catalog Tree, click the hard drive designator of the path [__].[9] The folders in that will appear in the right pane. Click the plus sign next to the entry for that drive. The folders that are in that are displayed, approximately mirroring the contents of the right pane. Click the plus sign in front of the IGIS-Arc entry in the Catalog Tree and observe the results. Click the entry name IGIS-Arc and observe the results.

You may select an entry name in the Catalog Tree by clicking the name. When you do so, the list of contents of the selected entry shows up in the right-hand pane (provided the Contents tab is selected). (That is, the contents of the right-hand pane is what you would get if you expanded the selected entry in the left-hand pane.) This conveniently lets you see your data names at two levels in different ways. You can select a disk drive, folder, or file from the left pane, and the selection is reflected in the right pane.

In the *left-hand pane* (with the Contents tab of the right-hand pane active):

❑ If you click a folder or file, it becomes selected. If it is a folder, its contents are revealed in the right-hand pane. If it is a file, it may be described in the right-hand pane.

❑ If you double-click a folder, it is selected. It is also expanded or collapsed in the left-hand pane, depending on its previous state.

❑ If you double-click a file, the operating system does something with it—executes it, opens it with an appropriate program, or opens a window that displays its properties. For example, if the file is a text file, it may be displayed in the Notepad text editor.

❑ If you right-click an entry, you get a menu of commands that let you do things like copy, paste, delete, rename, make something new (such as a folder, geodatabase, shapefile, coverage,[10] and so on), search, or reveal the properties of the entry.

In the right-hand pane, with the Contents tab active:

❑ If you click a folder or file, it becomes selected. (The parent folder name is semi-highlighted, instead of blue, in the left-hand pane, just for reference.)

❑ If you double-click a folder, its contents are revealed. To return to the parent folder, click the bent up-arrow icon at the extreme left of the Standard toolbar.[11]

❑ If you double-click a file, the operating system does something with it—it executes it, opens it with an appropriate program, or opens a window that displays its properties.

[9]Recall that you recorded the drive and path (if any) of the [__] designator in Step 2 of Exercise 1-2 above. In the event the path to IGIS-Arc does not appear, click File > Connect Folder and navigate to the IGIS-Arc Folder. Select it with a single click. Then OK the Connect to Folder window.
[10]Geodatabases, shapefiles, and coverages are types of GIS data sets, which will be covered extensively shortly.
[11]In the default ArcCatalog window (Figure 1-16), this icon will appear on the line below the Main menu, just under the File keyword, but it may have been moved in your ArcCatalog window.

❑ If you right-click an entry, you get a menu of commands that may let you do things like copy, paste, delete, rename, make a new something (folder, shapefile, coverage, and so on), search, or reveal the properties of the entry.

____ **7.** Experiment with selecting entries in each pane. Also drag the dividing line separating the panes to the left and right to change the space devoted to each. Notice that if there is not sufficient room for the length of the longest entry's text string, a scroll bar appears at the bottom of the window. Further, if you pause the cursor over a text string that is not completely visible, the entire contents of the string will show up in a box that overlays the incomplete string.

The pane on the left side contains, at the least, references to all hard disk drives on the computer,[12] as well as other connections to data. However, you can generate additional references to data sets that you are particularly interested in or want to work with in the future, so you have them right at hand. This is called "connecting to a folder" and is a special feature of ArcCatalog that provides a shortcut to data that you are interested in using. In the next steps, you will make a connection to a folder containing the HYDRANTS data for the village you will explore.

____ **8.** Collapse all the entries in the Catalog Tree completely by clicking each "-". Click `Catalog` in the Catalog Tree.

Connecting to a Folder

____ **9.** From the `File` menu, select `Connect Folder`. A `Connect To Folder` window appears. Using that window, expanding the entries in its tree as necessary, navigate to a folder named

[__] IGIS-Arc\Village_Data\[13]

Click that folder name so that its little folder icon opens and its name is highlighted. See Figure 1-20. Press the OK button. Observe that the Catalog Tree (in the left pane) now contains an entry (currently selected) that gives you access directly to the desired folder. (Also, the path to that folder is shown in the `Location` text box and in the title bar of the ArcCatalog window.) Now whenever you want to get to that folder, you merely need to find this direct shortcut in the Catalog Tree, rather than having to navigate to it through levels of hierarchy.

Basically, connecting to a folder provides a shortcut that you should always set up when you anticipate needing a data set more than once or twice.

____ **10.** Using `File > Disconnect Folder`, remove the entry you just made. Then restore it with the `Connect To Folder` icon, which you will find on the Standard toolbar that is located under the Main menu. The icon resembles the one associated with `File > Connect Folder`. Finally, make a folder connection with [__] IGIS-Arc folder itself. Your Catalog Tree should look something like Figure 1-21; it may be different depending on the [__] path to IGIS-Arc.

[12]If your computer is linked to a network and that network has drive designations, those designations may not show up in the Catalog Tree. A way of forcing the designations to appear will be covered shortly.
[13]Recall that [___] is the path designation of the data for this text. You determined it in Step 2 of Exercise 1-2.

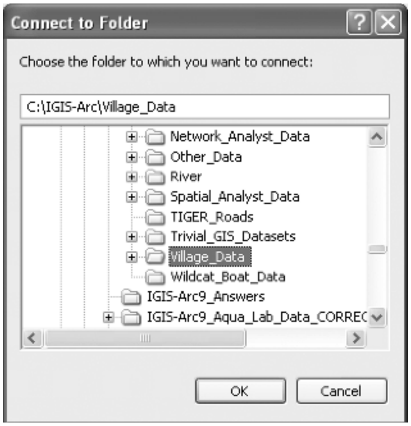


FIGURE 1-20

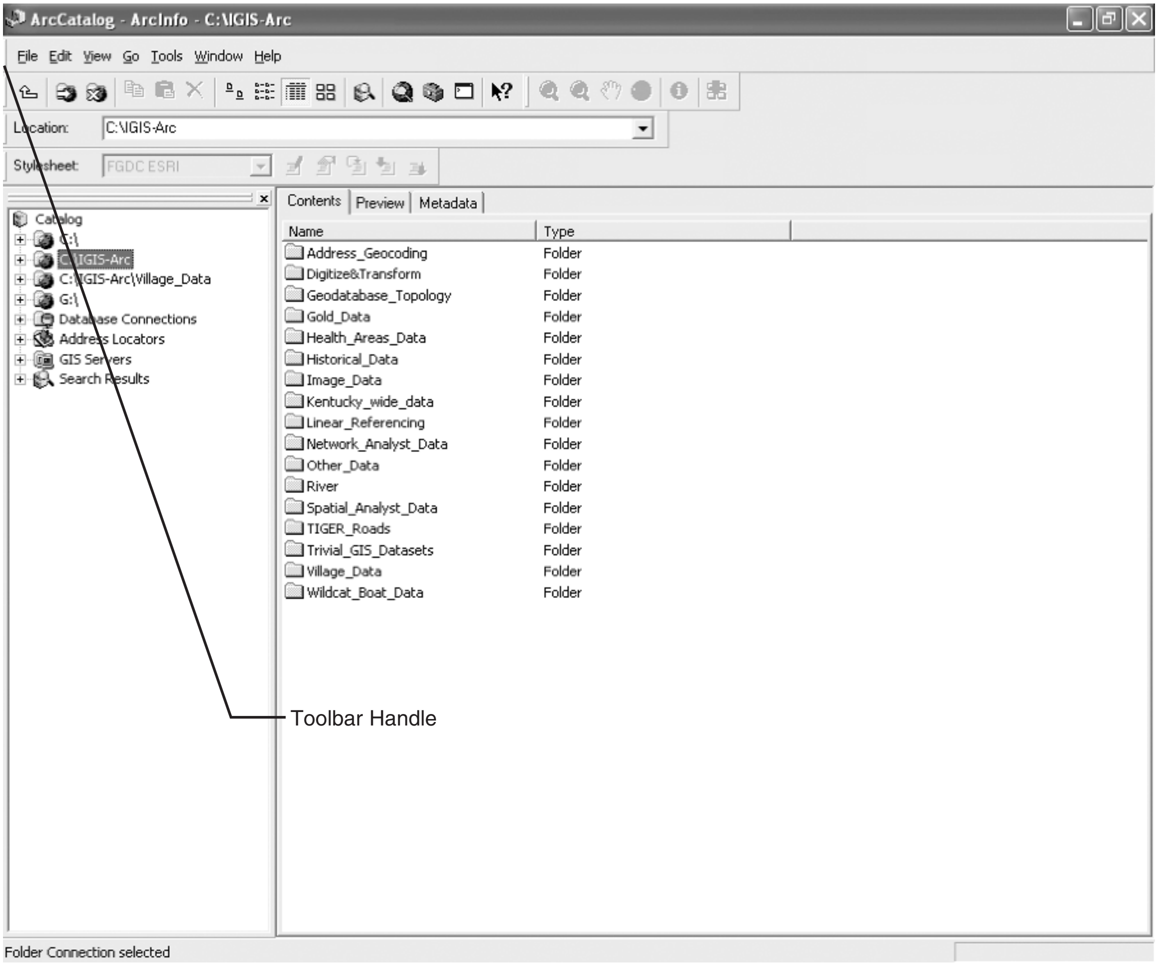


FIGURE 1-21

The Toolbars and the Status Bar

___ **11.** Under `View`, turn off the `Status Bar`. The information bar at the bottom of the screen disappears.

___ **12.** Turn the `Status Bar` back on so that messages are displayed at the bottom of the window. To see how the `status bar` shows information, find a red icon that looks like a tool chest on the Standard toolbar. Move the mouse cursor pointer over to it. The `Status Bar` should read `Show/Hide the ArcToolbox Window`. A yellow label (called a ToolTip) with the same message will appear next to the cursor.[14] Also experiment by clicking files and folders in the Catalog Tree and observing the `status bar`. In particular, look at the `status bar` as you select [__] IGIS-Arc. Expand that entry. Select `Village_Data`. Expand. Select `Hydrants`. Expand. Select `Point`. Press the Preview tab at the top of the window. You will see the points that represent the hydrants in Village Data.

___ **13.** Under `View > Toolbars`, make sure the Standard, Geography, Location, and Metadata toolbars are all turned on. Active toolbars have a check to the left of their names.

___ **14.** Locate an almost-invisible vertical line at the extreme left of the `Main menu bar`. Consult Figure 1-21 to see where this “toolbar handle” is. Using that handle, drag the menu to the upper left of the right pane. Also use this technique to drag each of the other toolbars so the window resembles Figure 1-22. By “undocking” the toolbars in this manner, you can see the name of each. (When the toolbars are docked, the names are hidden to save space.) Also move the Catalog Tree by dragging it by its top. Then turn it off by clicking the “X” in its upper right corner. Turn it on again by checking `Catalog Tree` in the `Window` menu. Re-dock it in its original position by dragging its title bar.

___ **15.** Run the mouse pointer over each of the buttons on the Standard toolbar. Notice that, if you pause for a second or so over an option, a small box containing a ToolTip appears. You can, however, get an instantaneous and more complete description of the tool by looking at the `status bar` at the bottom of the window.

The `Standard` toolbar provides icons that let you do the frequently-used actions explained immediately following. Of course, as with many software packages, there are frequently several ways to perform an action. Some examples of alternative procedures are given.

Go up a level from whatever selection you have made in a pane. (You can get to the top of the Catalog Tree by repeatedly clicking this `Up One Level` icon on the toolbar.)

Make a folder connection (substitutes for `File > Connect Folder`).

Remove a folder connection (instead of `File > Disconnect Folder`).

Copy, paste, or delete a selection (or use Ctrl-C, Ctrl-V, or Delete keys).

[14]Provided ToolTips is set on. To control this: Choose Tools > Customize > Options > Show ToolTips On Toolbars.

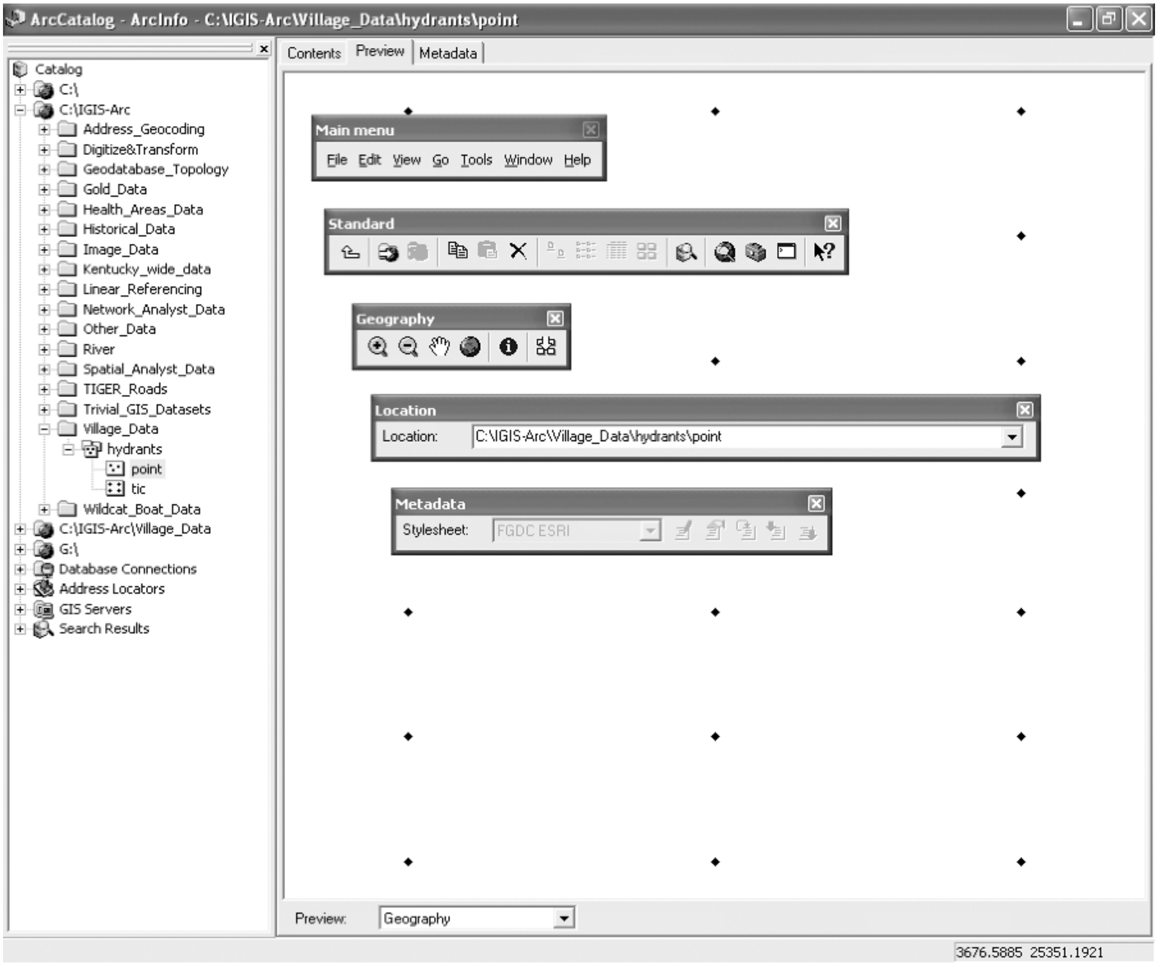
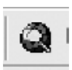


FIGURE 1-22

View folder contents in a variety of ways.

Search for data on the computer, a network, or the Internet, using a variety of criteria.



Launch the ArcMap software package.

Show/hide the ArcToolbox window.

Show/hide the Command Line window.

Get specific help on a tool or command.

____ **16.** Explore the `Geography` toolbar. This toolbar lets you look at (geo)graphic[15] data in a number of ways.

See features in more or less detail by zooming in and out.

Pan around the area of interest.

Look at the entire geographic region of the selection.

Identify, with text, various features indicated by the pointing cursor.

Make thumbnail sketches of the geographic file of interest so you can recognize it when searching for particular data sets.

____ **17.** Explore the `Location` toolbar. The `Location` toolbar displays a text string showing the current selection from the Catalog Tree and the complete path to that selection. Also, you can type in a different path and selection. If you type in a path and selection (or simply select the text string that is shown) and hit Enter, the string will be placed in the drop-down text box menu below the Location toolbar, making it easy for you to select this location next time. This provides another way to quickly get to data you are using. ArcGIS is full of shortcuts!

____ **18.** Explore the `Metadata` toolbar. Metadata is "data about data." It is usually in text form. There are different standards and styles.

The icons of the `Metadata` toolbar are grayed out because the toolbar is not active. If you click the Metadata tab (next to the Contents and Preview tabs), the toolbar will become active. This is a characteristic of the ArcGIS desktop products. A tool is available to the user when, and only when, it can be applied.

Click the Metadata tab of the ArcCatalog window so that the icons of the Metadata toolbar become active.

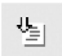
This toolbar lets you choose among the metadata stylesheets (the FGDC ESRI stylesheet is the default). It also allows you to do the following:

[15]What you are looking at is, of course, graphics. But since the image is tied to positions in the real world, I use the nonstandard term (geo)graphic.


Edit metadata.


Look at metadata properties.


Create and update metadata.


Import metadata.


Export metadata.

When you are finished exploring, you will want the toolbars put back to the positions shown in Figure 1-16. This can happen automatically. If you double-click the title of a toolbar, it will align itself horizontally, in the position it was in before you moved it, and its title will be hidden.

____ **19.** Restore the toolbars to their "home" positions.

Exploring Basic GIS Data Storage Models

Large amounts of computer data are usually stored on mechanical devices called disk drives. Disk drives constitute the "slow memory" of a computer. It takes a computer thousands of times longer to retrieve data from slow memory than from fast memory. An analogy with human processes might be that your brain—where you remember your friend's names, your home telephone number, or the route you drive to work—is your "fast memory" or your "electronic memory." On the other hand, the name of your uncle's second cousin, the phone number of the dry cleaner, or how to get to Punxsutawney, Pennsylvania, are things that you probably have to look up; they reside in slow memory—family documents, telephone book, and road map (or a GIS).

Most of the data that exists in the world on computers is, of course, not in use (not being printed, analyzed, or otherwise processed) at any given moment, and is therefore not contained in the "fast" electronic memory of a computer. Data sets in the slow memory of a computer are, in their most basic form, just sequences of 0s and 1s. But beyond that, well beyond that, such data sets are organized by storage paradigms. You are familiar with the idea of folders on a disk drive. Each folder can contain files and other folders. The operating system of the computer is responsible for keeping files and folders straight. Files may consist of binary sequences that result in typed text, music, photographs, and other products when processed by a computer. For a number of reasons, GIS data sets are composed of fairly complex combinations of folders and files. Also, the techniques for storing spatial data sets have evolved over a number of years, so that, even within the single company ESRI, you will find several storage paradigms or formats. Some of these formats have been devised because of the types of data being represented. Other formats come from different inventions based on progress in hardware and software development. In what follows you use the `status bar` to look at the terminology associated with several of them: coverages, shapefiles, geodatabases, rasters, and TINs.

___ **20.** Check that the `Options` you set in the previous Steps 4, and 5 are still as you left them. If not, reset them. Make sure the `Contents` tab is active. Collapse the entries in the Catalog Tree as much as possible; you should see no little minus signs. In the Catalog Tree, click the [__] IGIS-Arc designation. Write here what the `status bar` indicates: ___________________. Expand the entry. Find the entries indicated in the following, expanding them when possible. Write the `status bar` text string for each entry. (Those who are artistically inclined may draw the icon they see next to the entry.)

Village Data ____

HYDRANTS ____

Point ____

River ____

Boat_SP83.shp ____

COLE_DRG.tif ____(say "No" if asked)

COLE_DOQ64.jpg ____(say "No" if asked)

COLE_TIN ____

Wildcat_Boat_Data ____

Wildcat_Boat.mdb ____

Area_Features ____

Soils ____

Close Arc Catalog.

Exercise 1-5 (Major Project)

Exploring Data with ArcCatalog—Fire Hydrants in a Village

Copying Data over to Your Personal Folder

As mentioned earlier, ArcCatalog serves as a sort of operating system for GIS data. One function of an operating system is to make copies of folders and files. When working with GIS data in ESRI formats, you should *always* copy data sets using ArcCatalog—*never* use Windows, UNIX, Linux, or any other primary operating system to copy data sets.

In using this text, you will use your working folder

___IGIS-Arc_*YourInitialsHere*

to do most of the exercises. The idea is to leave the [___] IGIS-Arc folder contents in pristine condition. So you will usually copy over the data sets to your personal folder: ___IGIS-Arc_*YourInitialsHere*. We'll start with the Village Data.

___ **1.** Start ArcCatalog. Expand [__]. Expand IGIS-Arc. Select Village_Data. Right-click the selection and pick Copy. Select

___IGIS-Arc_*YourInitialsHere*.

Right-click the selection and pick Paste.

___ **2.** Make a folder connection with ___IGIS-Arc_*YourInitialsHere*.

___ **3.** Collapse the Catalog Tree as much as possible by clicking all the minus boxes. Then select the folder connection you just made, and expand it as much as possible select HYDRANTS. Click the Contents tab in the right pane. You now see references to both the point and tic components in both panes. Click Point. The Catalog Tree should look something like Figure 1-23. If you see serious discrepancies, use the Delete Key to remove all folders from

___IGIS-Arc_*YourInitialsHere*

and try again. (Be careful not to remove the entire folder IGIS-Arc_*YourInitialsHere* because, if you do, your Fast Facts File will go away.)

FIGURE 1-23

As you determined previously, HYDRANTS is an ArcInfo coverage—coverages are discussed in detail in the next chapter—consisting of 21 points that represent the locations of the fire hydrants in the village. Explore HYDRANTS in the following steps.

4. Select `Point` in the left pane. In the right pane you see the name and type of that component (feature class). Because of settings you made earlier,[16] you also see its size (in megabytes) and the date it was last modified; you would also see its geographic projection, if it had one. You should also see a blank rectangular image. This will later become a miniature image of the coverage called a thumbnail; we will create it shortly. See Figure 1-24. If you now select the `Preview` tab *at the top* of the pane, and click `Geography` in the `Preview` drop-down list *at the bottom* of the pane, you will see an image showing the positions of these hydrants. See Figure 1-25. This is hardly an exciting picture, but there is a lesson here, so please be patient. If you now select `Table` from the Preview drop-down list *at the bottom* of the right-hand pane, you will see a table of 21 rows and several columns. See Figure 1-26.

Each *row* (record) of the table represents one hydrant. Each *column* (field) of the table is an attribute (that is, a property or characteristic) of the data set HYDRANTS. Each *cell* in the table (where a given row and a given column intersect) indicates the particular attribute value for the particular hydrant.

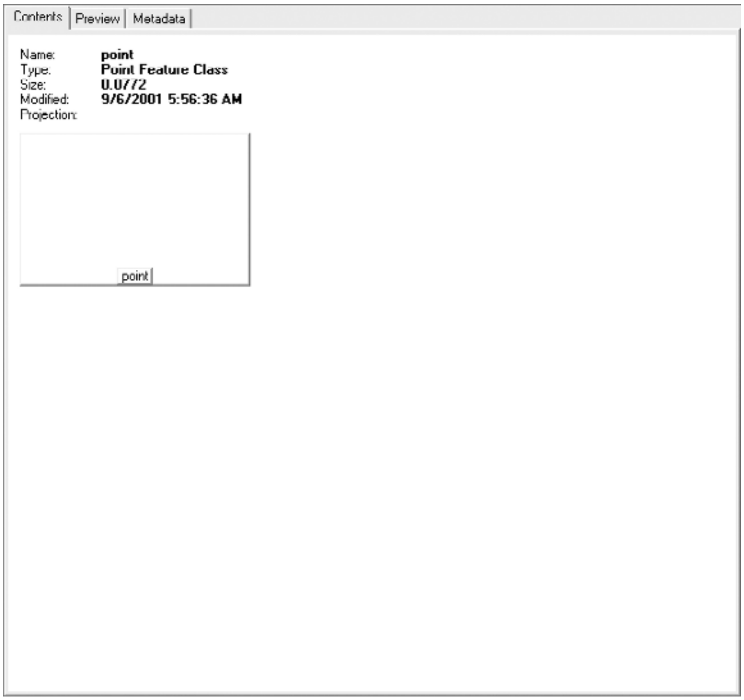


FIGURE 1-24

[16]Repeat Step 5 in Exercise 1-4 if the Options have become unglued.

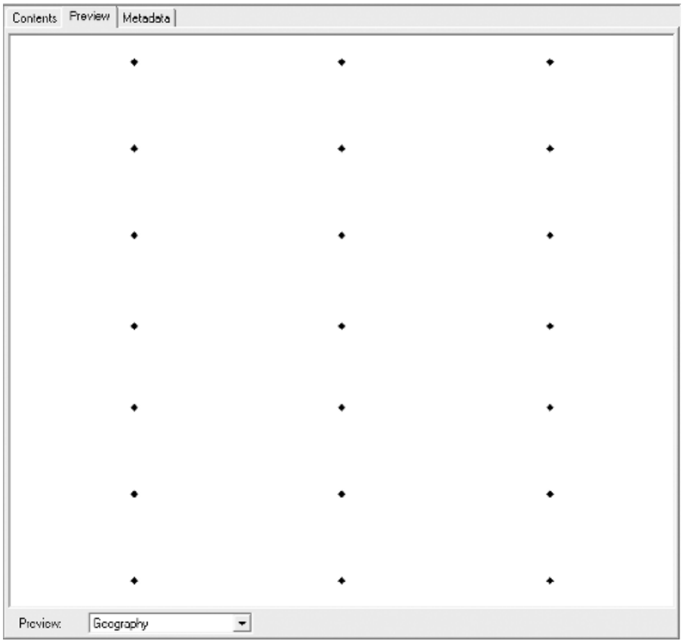
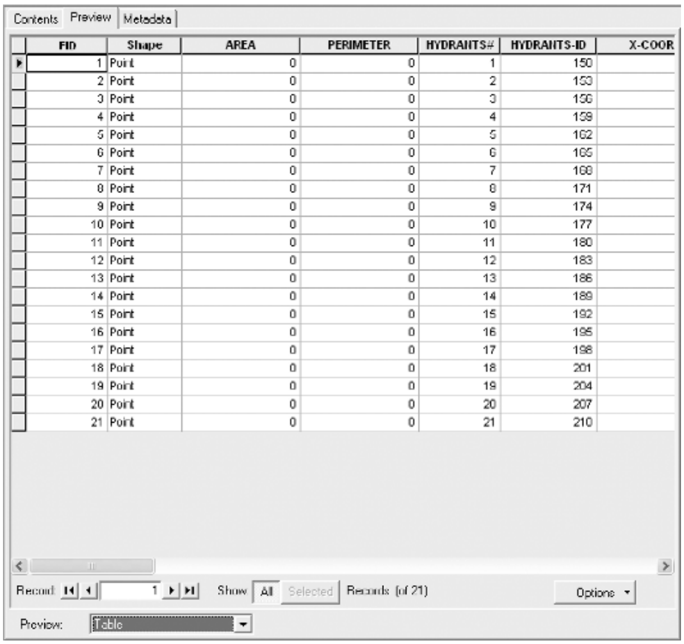


FIGURE 1-25



FID	Shape	AREA	PERIMETER	HYDRANTS#	HYDRANTS-ID	X-COOR
1	Point	0	0	1	150	
2	Point	0	0	2	153	
3	Point	0	0	3	156	
4	Point	0	0	4	159	
5	Point	0	0	5	162	
6	Point	0	0	6	165	
7	Point	0	0	7	168	
8	Point	0	0	8	171	
9	Point	0	0	9	174	
10	Point	0	0	10	177	
11	Point	0	0	11	180	
12	Point	0	0	12	183	
13	Point	0	0	13	186	
14	Point	0	0	14	189	
15	Point	0	0	15	192	
16	Point	0	0	16	195	
17	Point	0	0	17	198	
18	Point	0	0	18	201	
19	Point	0	0	19	204	
20	Point	0	0	20	207	
21	Point	0	0	21	210	

FIGURE 1-26

___ **5.** Locate the record of the hydrant that has a FID (Feature Identifier) of 4. Click the HYDRANTS-ID cell (it has value 159) and then scroll the table horizontally, using the scroll bar at the bottom of the window. You will see that the hydrant COLOR is Red.

Illustrated here is the essence of GIS: the marriage of a geographic database and an attribute database. Each feature in the spatial field has associated with it a row in a relational database table that provides information about the feature. This idea, with variations, is the fundamental underpinning of most of what you will learn about GIS.

___ **6.** Re-read the preceding paragraph.

Examining the Table

___ **7.** The table has a "current record" that, when you first view the table, is the first record. You can change that. Because you clicked on a cell in the fourth row, the current record is now 4. It is marked by a triangle in the box to the left of the record. The number of the current record is also shown in the `Record` text box at the bottom of the window. You also see the total number of records. You can change the current record by

- ❑ Clicking in a box to the left of the FID field,
- ❑ Clicking a cell of a record,
- ❑ Typing the desired record number in the Record text box (press Enter), or
- ❑ Clicking the buttons on either side of the Record text box.

You can scroll the viewable area of the table with the horizontal scroll bar and, if the table has a lot of records, the vertical scroll bar. You can also change the widths of the columns for better viewing by dragging the dividing lines between the column headings. (To see more of the table, you can dismiss the Catalog Tree, but the Tree is probably something you will be using frequently, so that should come into consideration.)

___ **8.** Both the AREA and PERIMETER columns are dull and excessively wide. Place the mouse pointer on the column heading text line and drag the column separators to the left to reduce the widths of these columns. On the other hand, the complete titles of the next two columns—HYDRANTS# and HYDRANTS-ID—may not be visible because the columns are too narrow. Fix that. If you double-click a column separator, the column to the left immediately tailors itself to the size that will show the entire column head with no space wasted.[17] However, the column may not be wide enough to completely show the values in the cells of that column.

A word about identification numbers: In coverages there is a record number (FID) as well as a HYDRANTS# number; they are assigned by the software and are the same. HYDRANTS# is sometimes called the Internal-ID. It is there so the software can keep track of which record goes with which hydrant, regardless of the actions of the user. The HYDRANTS-ID number, on the other hand, is a user-assigned integer; it may be called the User-ID or the External ID.

[17] If an entire column becomes selected while you are doing this, don't worry about it.

You noticed that an attribute for AREA and another for PERIMITER are shown, and that they are both zero. Since it is obvious that the area and perimeter of a point (a zero-dimensional entity) would be zero, it seems a little strange that the software would bother to store or report these values. The reason is that this same table format is used to represent polygons (such as ownership parcels or school districts) that have nonzero areas and parameters. The reason for storing points and polygons with the same format is historical and will be discussed later.

____ **9.** Make record 12 the current record. With the keyboard arrow keys, locate the HYDRANTS-ID value of the record. What is it? ______ What is the FLOW_RATE of the Hydrant? ______ What are the geographic coordinates of the hydrant? __________ __________

Should you wish to do so, you can change the cosmetic appearance of a table through Tools > Options > Tables.

Deriving Information from the Table

Having attribute information in table form allows you to obtain statistics, to search for specific text strings, and to sort the information.

____ **10.** You can find out the average of the values of a column along with the count of the number of values, minimum, maximum, sum, and standard deviation.[18] To do so, you right-click the column heading to bring up a menu of choices from which you choose Statistics. For example, select the column FLOW_RATE by placing the mouse pointer over the heading of the column (the pointer becomes a down-pointing arrow) and right-clicking. (You may have to use the scroll bar at the bottom of the table to find the FLOW_RATE column.) Choose Statistics from the drop-down menu. You get information about the values in the column from the `Statistics of point` window. There are 21 values. The minimum flow rate is 250 and the maximum is 400. You also get a frequency diagram that shows the *number* of hydrants that have each given value (e.g., there are four hydrants with a low flow rate, nine with a moderately high rate, and eight with a high rate). If all the hydrants were turned on at once, how much water would flow? __________. The mean (i.e., average) and standard deviation, although they probably don't mean much in this situation, are about 356 and 54, respectively.

____ **11.** Once you have the `Statistics of point` window up, you can determine the statistics of other columns by selecting the column name in the window's `Field` drop-down menu. What is the average y-coordinate? ___________ Dismiss the `Statistics of point` window.

Sorting the Records

____ **12.** Suppose now that you wanted to have the list ordered so that all the western most hydrants (those with the smallest x-coordinates) appeared at the top of the table. Click the X-COORD column heading to make it active. The column becomes highlighted in blue. Now right-click the column head. Then sort the values in the column by picking `Sort Ascending` from the drop-down menu. Notice that the values of the column are now in order from smallest to largest.[19]

[18]You are probably familiar with these concepts. A discussion of basic statistical measures may be found in the Overview of Chapter 6.
[19]A warning: ArcCatalog will sort the selected column, regardless. You can initiate the drop-down menu by placing the pointer over another column, but this column will not be the one that is sorted.

Note that when a value in a given record is moved, say, during a sort, the entire record moves with it. That is, all the values in a given row stay together (including the FID), *regardless*. (If you are familiar with spreadsheets, you should note that this is different from results you may get with that software.)

If you wanted the Y-COORD column values to be a secondary sort (sorted “within” the X-COORD values), you could select both columns (by holding down the Ctrl key while clicking). (An example to clarify the terminology: In a telephone directory the first names are sorted “within” the last names.) When two or more columns are selected, the leftmost one will contain the primary sort, the next selected on the right will contain the secondary sort, and so on.

___ **13.** Sort ascending the X-COORD and, within it, COLOR. Examine the results.

If having the leftmost column as the primary sort order doesn’t suit you, you must rearrange the columns, demonstrated next.

___ **14.** Select the COLOR column with a click. Hold down the left mouse button. The cursor changes—a little box is attached to it. You can now drag the column. The column will be placed to the left of whatever column the cursor is in when you release the mouse button. Drag the COLOR column to the left of the X-COORD column. Select both columns. Select `Sort Descending` (from the drop-down menu as before) and again note the results.

Finding Values in a Table

___ **15.** Make the COLOR column active by itself. Locate and press the Options button at the lower right of the window. (The button may be hidden or partially hidden; if so, increase the width of the window by dragging the separator between the table and the Catalog Tree.) Find all instances of the text string “Red” using the binoculars (`Find`) icon.[20] Each instance, when found, will have a darker border around it. (You may have to move the Find window to see the COLOR column.) How many records are there with the text string “Red”? _________.

___ **16.** Notice that the Find window gives you considerable flexibility:

- ❑ You can search all the values in a table, or just those in selected columns (fields).
- ❑ The text you type may match only a part of the value or be required to match the whole string.
- ❑ The searched-for string may be required to match the upper- and lowercase of each letter or not, as you choose
- ❑ You may search the entire table or restrict the search to records at or below the current record, or search those at or above the current record.

___ **17.** In the `Search` drop-down window, pick `All`. With `Match Case` checked, search for “yellow.” No records will be found, because of the lowercase “y.” Now click Match Case off. How many instances do you find? ______. Dismiss the `Find` window.

[20] Do not type the quotes in the Find What text box.

___ **18.** Move the COLOR column to the right of the CONDITION column. You have only a limited view of the table. You may not be able to tell which hydrant number (HYDRANT-ID) you are looking at when a given cell in the COLOR column is selected. You can set things up so you can always see a given column or columns. To `Freeze` the HYDRANTS-ID column, select it, right-click, and pick `Freeze/Unfreeze Columns`. It will now appear always as the leftmost column of the table. Scroll the table horizontally to see the effect. (You can thaw a frozen column in much the same way.)

When you terminate the ArcCatalog program, the results of some actions you took are retained, while others are discarded. The folder connections you have added to the Catalog Tree are remembered by the software. So are the positions of toolbars and some other settings. But if you have sorted or rearranged a table using ArcCatalog, these changes will not be retained. Operations in ArcCatalog such as sorting do not change the attribute table that is stored with the data set. You are merely changing the table's appearance during the time ArcCatalog is being used to look at a particular geographic data set. In fact, to restore the view of a table to its original form, you need only select some other folder or feature class and then return to the table.

___ **19.** Click Catalog at the top of the Catalog Tree. Click Contents. Now go back and again preview the HYDRANTS point coverage table. Notice the table is back to its former configuration—no frozen or rearranged columns and no sorted records.

___ **20.** Click the Preview tab at the top right-hand pane. Select Geography from the Preview drop-down menu at the bottom of the pane.

Identifying Geographic Features and Coordinates

Coordinates are a big deal, and frequently a big headache, for those doing GIS work. HYDRANTS is referenced by a local (and, in this case, fictional) coordinate system devised by the village or perhaps by the county. The units are survey feet. Suppose the village enters into an agreement of some sort with a regional water system agency to maintain the pipes and hydrants. If the agency uses a different coordinate system, the question might arise as to what the latitude and longitude values, rather than the local coordinate system values, were for these hydrants. As mentioned previously, the latitude and longitude coordinate system is the primary basis for the most accurate geographic information.

___ **21.** Select the `Identify` tool from the `Geography` toolbar. Write here the description of the tool that you see on the `status bar`. ___________________________. Notice the appearance of the cursor as you move it around the geographic area. Click in a blank area of the window that shows the hydrants. An Identify Results window appears. (You can move the window around by dragging its title bar and resize it by dragging a side or corner to better see the geographic image.)

___ **22.** Click one of the points representing a hydrant. As you click the feature, it is momentarily highlighted. The relational database information for that hydrant appears in the `Identify Results` window. Notice that the `Identify Results` window, in an information box labeled `Location`, gives you a precise value of the Cartesian (x and y) coordinates of the location of the tip of the cursor pointer. These coordinates should agree closely, but probably not exactly, with the X-COORD and Y-COORD values shown as attribute values. (It is not always the case with point data that the coordinates are part of the attribute database; you usually have to take action to put in X-COORD and Y-COORD; how this is done is described later. However, the information in the Location box is always calculated. *It comes from the geographic location specified by the cursor, not the relational database.*)

___ **23.** With the Identify tool, click again in a place where no fire hydrant exists. Attribute information disappears, but coordinate values of the tip of the pointer are still revealed.

___ **24.** Roughly (to the nearest tenth of a foot), what are the coordinates of the lower left-hand corner of the window? ________, ________. How about the upperright-hand corner? _______, _________.

___ **25.** Dismiss the Identify Results window. If you right-click `Point` in the left pane, you get a menu of options. Of interest to us now is `Properties`. Click that. A `Coverage Feature Class Properties` window appears. Under the `General` tab, you see that the number of features is 21. Of more interest is the Items tab, which tells you the names of all the attributes and the types of fields that are used to contain the data. For example, click the row that says `COLOR` to highlight it. Notice that `COLOR` is a Character field of width 9, meaning that it uses nine bytes of memory. Now look at `Y-COORD`—a numeric field taking up 4 bytes of storage and allowing the storage of a floating-point number, with three decimal places. `HYDRANTS-ID` is also a numeric field, but it is an integer and may therefore be stored in a simple binary format. Dismiss the window.

Looking at GeoGraphics

___ **26.** On the Geography toolbar, click the magnifying glass icon that has a "+" in it. Copy the text in the `Status Bar` here. _________________________________. Click the upper left hydrant. Note that it is moved to the center of the window and the distance between it and its neighbors is increased. You have zoomed in on the layer—not that you will see any more detail in this particular image, but you get the idea.

___ **27.** Notice that although you have zoomed up on the image, the symbols representing the features did not get any bigger. So this zooming action is somewhat different than looking at a paper map with a magnifying glass, which would increase both the distance between the points *and* the symbol size.

___ **28.** Click again on the northwest hydrant and observe the results. Click between that hydrant and its neighbor to the south. Click the Full Extent icon (on the Geography toolbar) to restore the view of the entire layer and bring all hydrants back into view. With the `Zoom In` tool active, drag a box around the middle three hydrants in the middle column and observe. The lessons: You can zoom in by clicking a point and also by dragging a box. The image is always recentered, either at the point clicked or the center of the box.

___ **29.** Click the "hand" icon on the Geography toolbar—this is the Pan tool. Check out its function on the `Status Bar`. Move the cursor over the middle point and drag that point to the left side of the pane. When you release the mouse button, you will see that the focus on the image has been shifted ("panned") to the right—in between two hydrants. Click the Full Extent icon.

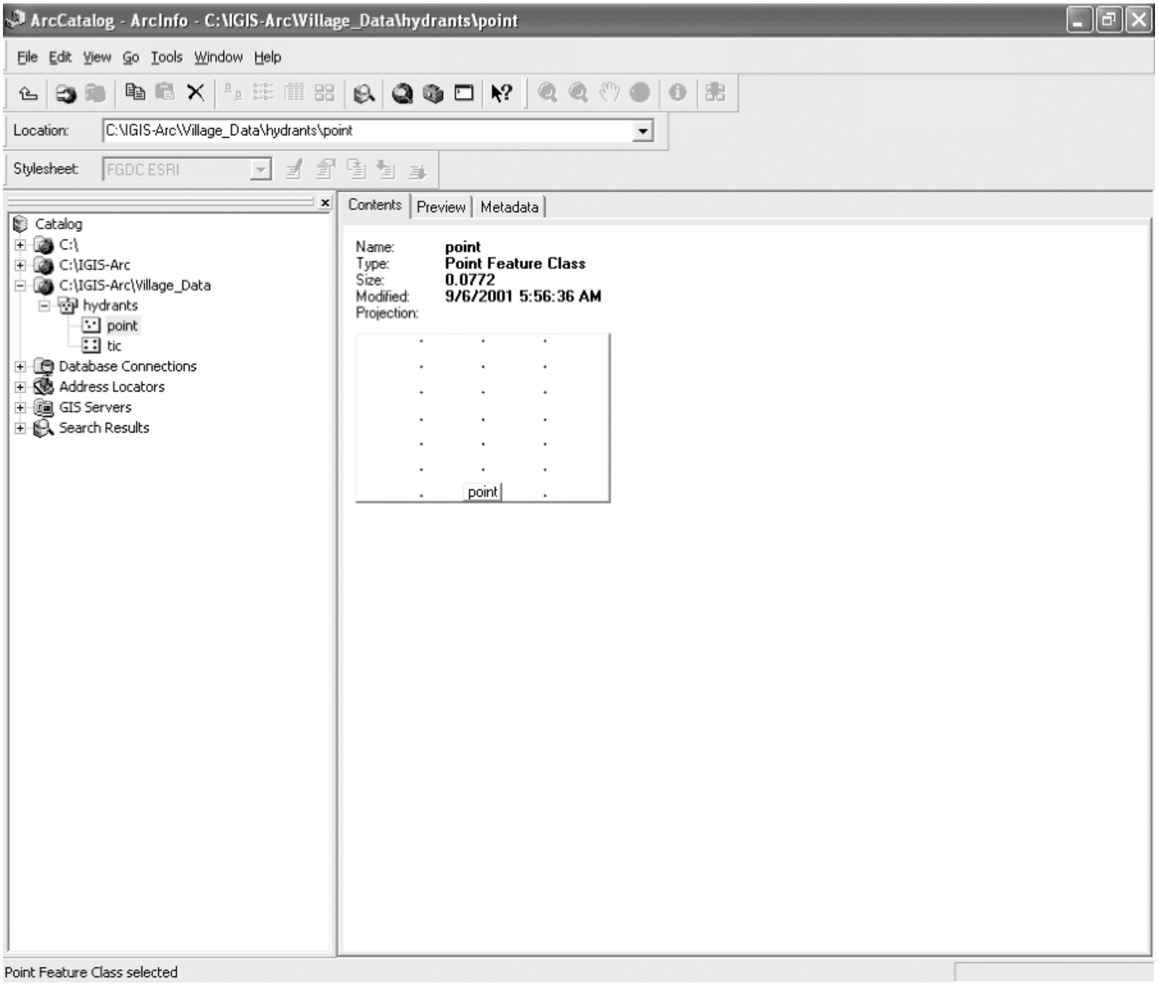


FIGURE 1-27

A "thumbnail" is a small sketch that can aid you in recognizing the contents of a data set. You make a thumbnail sketch in the step that follows.

____ **30.** ***Create a thumbnail of the Hydrants layer:***[21] On the Geography toolbar, find the `Create Thumbnail` icon. Click it, then click the Contents tab and observe. Where there was before a blank rectangle, there is now an image of the layer. See Figure 1-27. Suppose you want only a portion of the coverage to be represented in the thumbnail. Go back to the Preview tab. Use the zoom tool to focus on the six hydrants in the northeast corner. Make a thumbnail sketch of these. Check that it worked by clicking the Contents tab. Now make the thumbnail a third time with all the points shown.

[21] A note on using this text: When ***bold italic*** is used following a step number, it constitutes a general direction. You should read what follows the bold italics to see what specific actions you are to take.

Tics and Ticks:[22] Tying Coverage Geographic Data to the Real World

In ArcGIS parlance, a *tic* is a numbered place that corresponds to a precisely known geographic position. The tic positions tie the coverage's features to accurately known, real-world positions, such as U.S. Coast and Geodetic Survey markers, or perhaps the centers of intersections of major streets whose positions are well known. The coordinates of all the features in a coverage are directly, linearly related to the tic positions.

In the HYDRANTS instance, there are four tics (the minimum for a coverage), and they are at the locations of the four corner hydrants. A given coverage may have a large number of tics. Coverages that describe the same geographic area should have the same tics, if possible. That is, the tic number (and associated location) should be the same, regardless of which data set you are using if the data sets cover the same area. Tics are the key to transforming an ESRI geographic coverage from one coordinate system or projection to another.

____ **31.** Where you had selected `Point` in the Catalog Tree, now select `Tic`. The Contents tab reveals that these are features of the Tic Feature class. (However, the thumbnail will continue to be that of the features of the coverage.) The Preview tab lets you look at the four tics graphically and examine the attribute table. The attributes of the tic component of a coverage are simply the tic number (attribute name: IDTIC) and the x- and y-coordinates of each tic. Using the attribute table, determine the x- and y-coordinates of tic number 3. __________, __________. Now use the `Identify` tool to verify the correctness of what you found. Dismiss the `Identify Results` window.

A First Look at Metadata

As mentioned, "metadata" means data about data. There are so many things to know about geographic data sets that many people who work with them have recognized the need for established standards for describing such data sets. We will look at metadata in more detail later, using some real data sets. The lesson from what follows is just that metadata exists, that it consists of many elements, and that it can be presented in many formats.

____ **32.** Click Hydrants again in the Catalog Tree. Click the `Metadata` tab. The `Metadata` toolbar becomes active. On the `Metadata` toolbar, pick the `Stylesheet FGDC ESRI`, which is the ESRI presentation of the standard metadata format.[23] See Figure 1-28.

Under Description, you see a thumbnail of the data set (if a thumbnail has been created) and textual information that the creator of the data set provided about the general nature of the data. Some of the

---

[22]"Ticks" (spelled like the name of the blood-sucking creatures) are, among other definitions, small marks along the border of a map indicating the positions of grid lines of the map. A given tick mark indicates either a latitude (or northing) or a longitude (or easting), but not both. "Tics" (spelled like the facial spasms) are geographic control points (consisting of both latitude and longitude [northing and easting]). They tie a location in an ArcGIS coverage to a real-world position.

[23]FGDC stands for the Federal Geographic Data Committee of the United States government. The Web site Internet address is www.fgdc.gov. The committee coordinates the development of the National Spatial Data Infrastructure (NSDI).

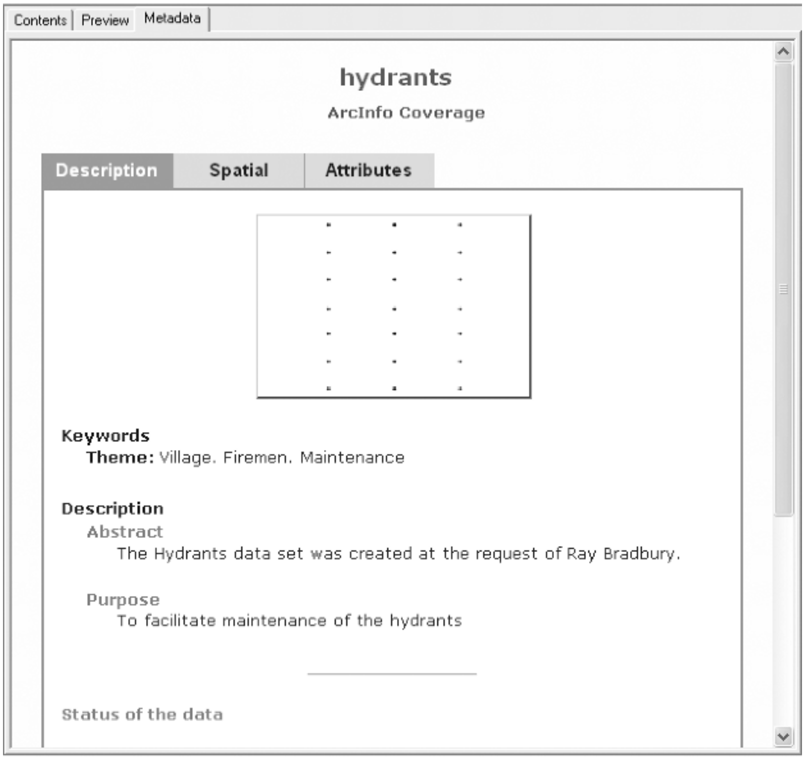


FIGURE 1-28

metadata comes about because ArcCatalog has inspected the data set. For example, under the `Spatial` tab, you could find the bounding coordinates of the features of the data set. These were derived automatically. In the HYDRANTS case, only the local coordinates were available from the data set. Also, generated automatically is most of the information found under `Attributes`: the name of each attribute, along with the data type, precision, and so on.

____ **33.** Explore the `Description` of metadata. You can click hypertext words (in green) to see the details they hide, or whether additional information is required to complete the metadata form. Expand and examine all hypertext. What is the size of the data set? ________________MB.

____ **34.** Explore the Spatial aspect of metadata by clicking the appropriate tab. What is the westernmost (leftmost) bounding local coordinate? ________________.

____ **35.** Explore the Attributes given by the metadata. What is the data type of the FLOW_RATE? ________________. What is the `Output Width` of X-COORD? ________________.

Other stylesheets are available. Several formats are based on the FGDC information. FGDC-FAQ is a "conversational" format in which frequently asked questions are answered. There are other formats, which display varying amounts of metadata with varying degrees of user friendliness.

___ **36.** Using the Stylesheet drop-down menu, look briefly at each of the other stylesheet types to get an idea of its characteristics. For `FGDC Classic`, describe the general form of the data set. ____. When you are through exploring, click the Contents tab.

Using ArcCatalog to Place Data in ArcMap

You have explored ArcCatalog and used it to find and connect to data sets you want and to explore them. As previously described, the other major component of ArcGIS Desktop is ArcMap. ArcMap lets you see, create, examine, query, edit, and develop geographic data and maps. Your first step in using ArcMap is to get the data you have found with ArcCatalog into ArcMap. There are several ways to do this. Three are described in the following. First you need to perform a couple of setup steps.

___ **37.** From the Standard toolbar (or the Tools menu), launch ArcMap and press OK to start with a new, empty map. Make ArcMap occupy the full area of the monitor screen.

___ **38.** Now do either or both Step 39 or Step 40, which follow, to put ArcCatalog and ArcMap on the screen at the same time.

___ **39.** Click the middle icon (named Restore Down) of the three in the far upper right of the ArcMap window. That has the effect of reducing the size of the window. By dragging sides and corners of the ArcMap window, make it occupy approximately the right half of the monitor screen. Make the ArcCatalog window occupy the left half of the monitor screen.

___ **40.** Right-click an empty area of the Windows taskbar to bring up a menu. (The Windows taskbar will be located on one of the four edges of your monitor screen; it contains the Start button.) From this menu, select `Tile Windows Vertically`. (If necessary, minimize any windows besides ArcCatalog or ArcMap that appear, and "re-tile.")

___ **41.** ***Method 1 of inserting data from ArcCatalog into ArcMap:*** Drag and drop the point component of the HYDRANTS coverage from the Catalog Tree of ArcCatalog to the pane in ArcMap that says Layers. The "map" will show up in the right pane of ArcMap. (Disregard any warning message about missing spatial reference information.)

___ **42.** Dismiss ArcCatalog. Maximize the ArcMap window. Right-click the string "Hydrants point." From the menu that appears, select Open Attribute Table. You can see that the coverage HYDRANTS—the combination of the geographical database and the attribute database—is now in ArcMap. (Of course, you have been writing information into your Fast Facts File all along. Record particularly how to open an attribute table in ArcMap.)

___ **43.** Right-click again the reference to the HYDRANTS data. Select `Remove` to take the data set out of ArcMap.

___ **44.** ***Method 2 of inserting data from ArcCatalog into ArcMap:*** In the ArcMap `File` menu, click Add Data . . . . In the Add Data window that appears, use the `Up One Level` icon (repeatedly, if necessary) so that the `Look In` box reads `Catalog`. Notice that the contents of the Catalog

Tree that you made in ArcCatalog is now accessible to you in ArcMap and contains the reference to the Village data. Navigate (you may have to scroll to do it) to the point component of the HYDRANTS coverage by doing the following: Double-click [___]IGIS-Arc\Village_Data, double-click HYDRANTS, click Point, and press `Add`. This, then, is a second way that ArcCatalog facilitates getting data into ArcMap.

___ **45.** Right-click the reference to the HYDRANTS data. `Remove` the data set from ArcMap.

___ **46.** ***Method 3 of inserting data from ArcCatalog into ArcMap:*** To see a third, efficient way of getting data from ArcCatalog into ArcMap, start ArcCatalog (from the ArcMap Tools menu or from the icon on the Standard toolbar). Close ArcMap, without saving changes to `Untitled`.

___ **47.** Make the ArcCatalog window occupy the entire monitor screen. Launch ArcMap. Make it also occupy the full monitor screen, covering up the ArcCatalog window. Note that both ArcCatalog and ArcMap are represented by buttons on the Windows taskbar. (Recall, the Windows taskbar will be located on one of the four edges of your monitor screen. The button for ArcMap will read `Untitled`.) By clicking one button or the other, you can bring up the associated window. (You can also bring up applications by holding down the Alt key and pressing Tab to select the application you want. Flip back and forth a couple of times.) End this step by making the ArcCatalog window active.

___ **48.** Position the cursor over the point component of the HYDRANTS coverage, press and hold down the left mouse button, move the cursor to the ArcMap (`Untitled`) button on the taskbar, wait until the ArcMap window has replaced the ArcCatalog one, move the cursor onto the left pane (labeled `Layers`) of the ArcMap window, and drop (by releasing the mouse button) the HYDRANTS point component. Bingo!

___ **49.** Close ArcMap. Close ArcCatalog.

Exercise 1-6 (Project)

A Look at Some Spatial Data for Finding a Site for the Wildcat Boat Facility

You usually don't get the luxury of working on one GIS project from beginning to end. Life in the GIS world frequently isn't like that. We'll be working on one rather large project (the search for sites for the Wildcat Boat facility that you did manually in Exercise 1-1), but along the way, we'll do several smaller ones.

This is the beginning of the GIS part of the Wildcat Boat project. Using ArcCatalog, you will find data for the site. Then, using the techniques you learned in Projects 1-4 and 1-5, you will use ArcCatalog to explore data sets associated with this project.

Most of the Wildcat Boat data is contained in a *personal geodatabase*. This is the newest data structure developed by ESRI, and it has a number of advantages over the coverage data structure. While most of the data you will encounter in the next few years will probably be in the coverage format (or another format called "shapefile"), the move to geodatabases is proceeding at a fast pace. So most of the work we do in this text will be with geodatabases.

Using the Area on the Disk for Your Own Work

Again, we want to keep the [__] IGIS-Arc data sets intact and unchanged, so you can go back to them at any time. But I want you to be able to modify the Wildcat Boat data. To do that, you need to use your own folder, which you made earlier and in which you are keeping your Fast Facts File and the Village_Data. The purpose of

___IGIS-Arc_*YourInitialsHere*

is to let you store files in a personalized folder, without compromising the information in IGIS-Arc.

Copying Data over to Your New Folder

___ **1.** Start ArcCatalog; make it occupy the full monitor screen. Expand [__]. Expand IGIS-Arc. Click Wildcat_Boat_Data. Right-click the selection and pick Copy. Select

___IGIS-Arc_*YourInitialsHere*.

Right-click the selection and pick Paste.

___ **2.** Using almost the same technique in the paragraph above, copy and paste

[___] IGIS-Arc\Other_Data

into your folder, except this time use Ctrl-C and Ctrl-V for copy and paste.

___ **3.** Collapse the Catalog Tree as much as possible. Then expand the entries (except for Village Data) in

___IGIS-Arc_*YourInitialsHere*

as much as possible. The Catalog Tree should look something like Figure 1-29. If you see serious discrepancies, remove the *folders* from

___IGIS-Arc_*YourInitialsHere*

and try again. (Again, be careful not to remove the folder

IGIS-Arc_*YourInitialsHere*

because, if you do, your Fast Facts File will depart with it.) Come to think of it, now would be a good time to back up your Fast Facts File. Put it on a floppy disk, a Zip drive, a thumb drive, a flash drive, a network drive, or e-mail it to yourself. Maybe you should make a couple of backups, using different methods. *It is rare that a person has too many backups; it is quite common to not have enough.*

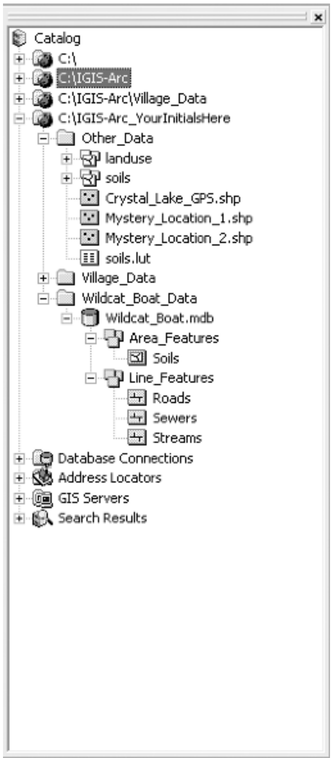


FIGURE 1-29

Searching for GIS Data

In the 1980s and 1990s, the major issue related to spatial data and GIS was how to create the data sets you wanted—either directly from the environment or by converting existing map data. Now the initial emphasis has turned, in many instances, to finding already-existing data—on the Internet and elsewhere. Tremendous stores of spatial data exist, and more comes in every day from satellites and ongoing projects. But you may discover that the data set you want is buried with a lot of other data; it may have an obscure name; the data sets you find may not meet your standards, even if they cover the correct subject and correct geographical area. A major feature of ArcCatalog is the capability to help you discover spatial data sets and, through inspection of their metadata, determine if they meet your needs. Assume that a client or your supervisor asks you to find all the geographic data that could apply to the Wildcat Boat project. You have been told that the data set exists in either a personal geodatabase or coverage. Further, you have been given the geographic coordinates that bound the area of interest.

___ **4.** In ArcCatalog use the Up One Level arrow to get to the top of the Catalog Tree—that's where it says `Catalog`. Click the Metadata tab. Read the paragraphs about the catalog. In addition to the capabilities described there, the catalog can provide the basis for a search for geographic data, which you will now illustrate to yourself. Click the `Contents` tab.

___ **5.** Use `Search` on the `Edit` menu (or the `Search` icon on the `Standard` toolbar) to bring up the `Search–My Search` window. Under the `Name And Location` tab,[24] leave the asterisk (*) in the Name field—since you don't know the name of the data set(s) you are looking for, and the star is a "wildcard" that can stand for any text string. Let's assume you know that the data type will be either a `personal geodatabase feature class`[25] or an `ArcInfo coverage`, so highlight those[26] on the menu—after scrolling down to look at the large number of possible data set types that you could search for. Look now at the Search field: In this case you want to search the Catalog, but you see that you could search just the file system or use the ArcIMS (Internet Map Server) Metadata Service. Under `Look In`, single-click Browse (the yellow folder) to look for

___ IGIS-Arc_*YourInitialsHere*

select it, and click `Add`. Don't start the search yet! In the next step, you will specify the geographic area to search.

___ **6.** Press the Geography tab. Make sure the box next to `Use Geographic Location In Search` is checked. For `Choose A Location`, pick `<None>`. For Map, select US Counties from the dropdown menu. Suppose you happen to know that the site that you want lies in the United States, south of latitude 41.1° and north of latitude 41.0°. Also, it lies east of longitude –74.1° and west of longitude –73.9°. (In what follows, *do not* hit Enter after typing into a box; just move to a new box with the mouse cursor and double-click. Also be sure not to forget the minus signs that are part of the longitude coordinates.) Specify the bounding coordinates by typing in the North, South, East, and West text boxes (e.g., place –74.1 in the West box; if you double-click the text in a text box, it becomes highlighted and you can type new digits to replace those already there.) As you enter information into the text boxes, note the red crosshatched area becomes smaller and smaller. After entering all four coordinates, click `Find Data Entirely Within The Location`. Use the Magnify icon in this window to zoom up on the red crosshatched area, which shows the approximate area for verification purposes. The window should look something like Figure 1-30***.[27]

Under the `Date tab`, make sure `All` is selected. Ignore the `Advanced` tab for now. Return to the `Name & Location` tab. Click `Find Now`. Wait until the magnifying glass stops moving around the Earth Symbol in the `Search–My Search` window. The result should look something like Figure 1-31. Dismiss the `Search–My Search` window.

___ **7.** The ArcCatalog search feature found four personal geodatabase feature classes that fell within the geographic bounds specified: Soils, Streams, Roads, and Sewers, all in the

___IGIS-Arc_*YourInitialsHere*\Wildcat_Boat_Data

folder. It also found a coverage named SOILS in the

[24]By "location" here it is meant "location" on a hard drive of the computer (i.e., path to the data), not spatial location. That comes later.
[25]A feature class in a personal geodatabase is somewhat like a coverage, in that it contains geographic features and an associated table. A detailed explanation will be found in future chapters.
[26]You can, by holding down the Ctrl key and clicking, highlight any number of data types.
[27]Figures indicated in the text with three asterisks (***) following are on the CD-ROM that accompanies the book. They are in the folder IGIS_with_ArcGIS_Selected_Figures. These figures are in color, which may provide you with additional information.

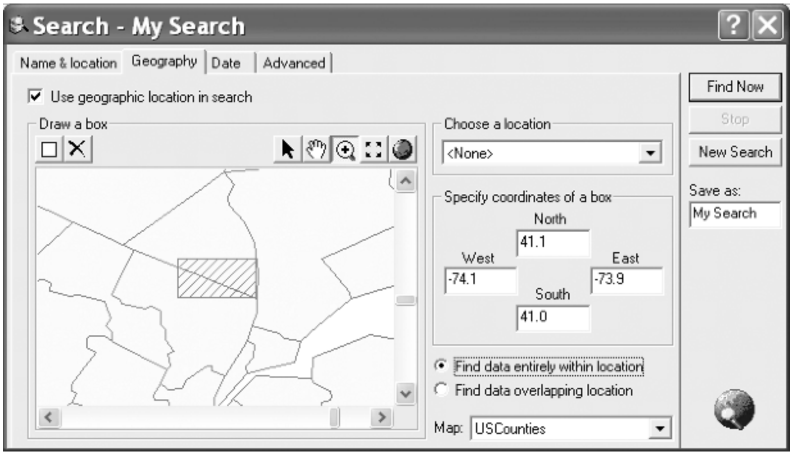


FIGURE 1-30

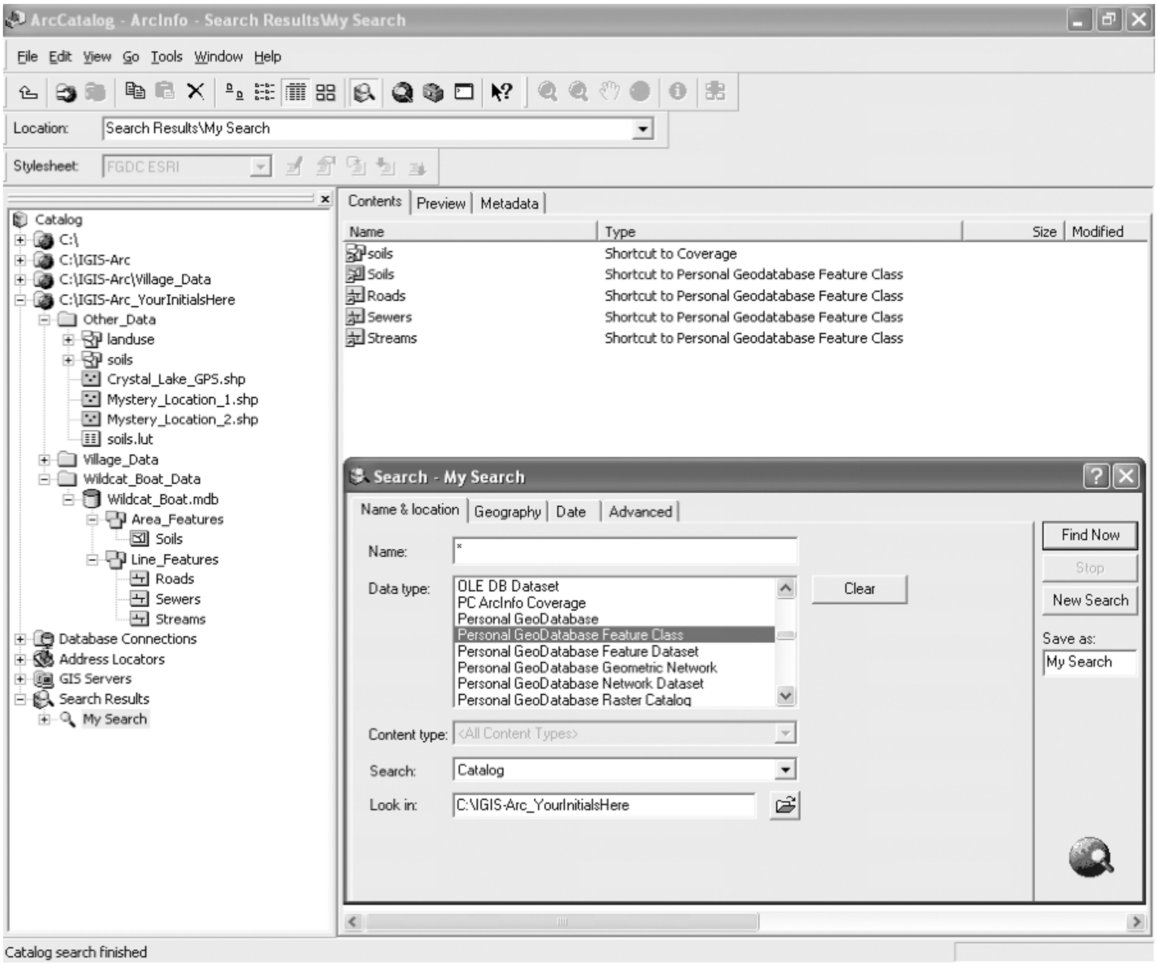


FIGURE 1-31

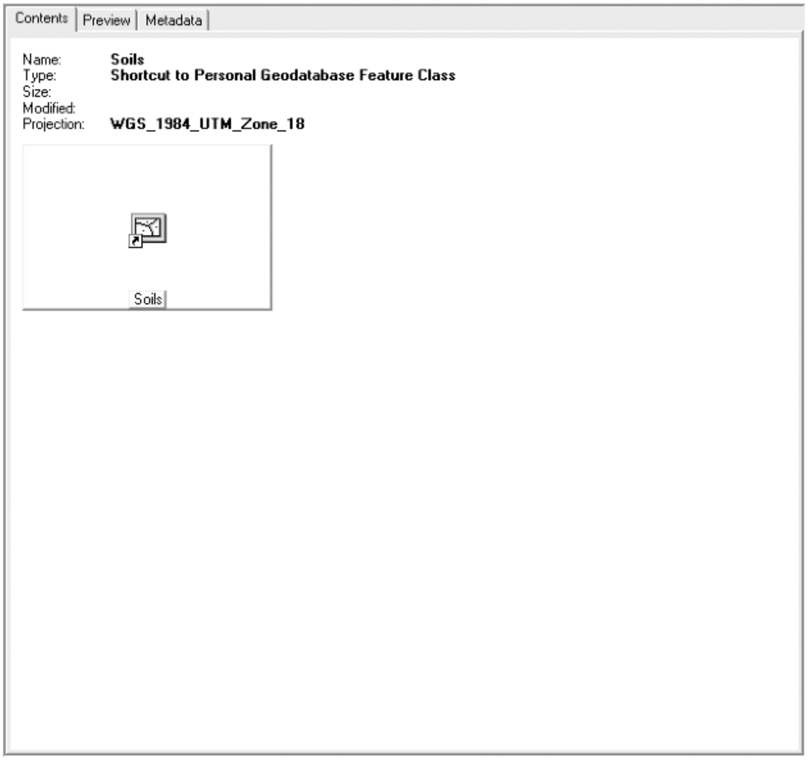


FIGURE 1-32

___IGIS-Arc *YourInitialsHere*\Other_Data folder. (Use the bottom slider bar to see all of the contents.)

___ **8.** In the Catalog Tree, find Search Results. Expand it to find My Search. Expand that. Click each of the five data set shortcuts to determine which are the feature classes and which is the coverage. Click the Soils feature class (not the coverage). With the Contents tab selected, you will find, in the right-hand pane, information about the data set. What is the type of the entry? ____. What is the projection[28] of the data? __________________. See Figure 1-32.

Exploring Soils

___ **9.** Click the Preview tab. You see a map of the Soils you worked with in Assignment 1-1. In the pull-down menu in the `Preview` text box, select `Table`. How many Soils polygons are there? __________.

[28]If Projection doesn't show up, it is probably because the options you set earlier in the chapter are no longer in effect. To set (or reset) the option that shows projection: Choose Tools > Options > Contents > turn on the Projection box.

___ **10.** Examine the table. Notice the fields Shape_Length and Shape_Area. These attributes are the perimeters and areas of polygons. This is information you get "for free" by using a GIS. Sort the Shape_Area column into ascending order. To the nearest tenth of a meter, what is the perimeter of the polygon with the smallest area? __________ meters.

___ **11.** Click the Metadata tab. Make sure that the stylesheet is FGDC ESRI.[29] Click the Spatial tab. Under `Horizontal Coordinate System`, you find

`Projected coordinate system name` ______________________________

`Geographic coordinate system name` ____________________________

Under `Bounding Coordinates`, you find

`Westmost bound: Geographic` ______________ `Projected:` ______________

`Eastmost bound: Geographic` ______________ `Projected:` ______________

`Northmost bound: Geographic` ______________ `Projected:` ______________

`Southmost bound: Geographic` ______________ `Projected:` ______________

Someone suggests to you that these bounding projected coordinates look very small, given the usual size of UTM coordinates. That is, since the northing is supposed to be the number of meters from the equator (that would be in the millions for 40 degrees latitude),[30] a number like 4500 (that's four and a half kilometers) obviously isn't right. So it is necessary to look more closely at the metadata. Click the green `Details` text and record the following:

`False easting:` ___________________

`False northing:` _________________

The question of the small number in the northing is now explained. Each UTM northing coordinate has been adjusted by subtracting 4,540,000 meters from the true northing coordinate to get the local coordinates, which are around 4000 to 6000. Why did the developers of the information decide to do this? To keep the size of the numbers relevant to the problem small. Perhaps, at the time the data sets were developed, the software operated in single precision, using only 4 bytes of storage for each coordinate. Numbers as large as 4 and a half million would contain too many digits to provide sufficient precision.

___ **12.** As you recorded earlier, the type shown in My Search is a Shortcut. While this has let us explore the data somewhat, we would really like to get closer to the real data set. If you right-click Soils in the Catalog Tree (Search Results\My Search), you will get a menu that contains `Go To Target`. Click that, and `Soils` will be selected in a different place in the Catalog Tree. Scroll to find it. It exists within a personal geodatabase feature data set named Area Features. Area Features, along with another data set named Line Features, is located in a personal geodatabase named Wildcat_Boat.mdb, where "mdb" is the extension for a *M*icrosoft *D*ata*B*ase. The

[29] If the Metadata toolbar isn't showing: Choose View > Toolbars > Metadata.
[30] It is about 10 million meters from the equator to the North Pole.

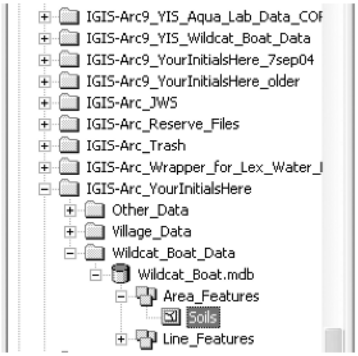


FIGURE 1-33

database is contained in a folder named Wildcat_Boat_Data, in the ___IGIS-Arc_*YourInitialsHere* folder. So you can see that we have quite a hierarchy going here: A folder contains the geographic database; the geographic database contains two geographic data sets; each geographic data set contains one or more geographic feature classes. Visually the hierarchy looks like Figure 1-33. So in the Wildcat_Boat geodatabase we have the Soils, Roads, Sewers, and Streams that we need for the Wildcat Boat facility problem.

But Something Is Missing

Assume at this point that you show these results to your client or supervisor, who insists that there is some land use or land cover data somewhere for this region.[31] Perhaps the name is simply LANDCOVER, so you will try searching for a data set with that name.

___ **13.** Initiate another search. This time, under the `Name And Location` tab, type `LANDCOVER` in the `Name` text box to replace the asterisk. You will search in the `Catalog`. `Look In` should again contain just ___IGIS-Arc_*YourInitialsHere*. And, again, you are looking for a coverage or a personal geodatabase feature class. Under the `Geography` tab, remove any check mark from `Use Geographic Location In Search`, since you already have all the data sets that have metadata indicating the geographic area. Nothing new can show up as long as the area is restricted.

Click `Find Now`. `My Search` now shows a data set named LANDCOVER. To find out more about it, dismiss the Search–My Search window, look at My Search in the Catalog Tree, click the `Landcover icon`, and click the `Contents` tab. You find that the data set is a coverage, located in

___IGIS-Arc_*YourInitialsHere*\Other_Data.

[31]The concepts of land use and land cover have a lot in common—and significant differences. Both describe "what's going on" with a particular piece of real estate. A land cover classification basically tells what sort of surface component is present, such as forest, water, urban, wetland, and so on. In a land use classification, the surface component is considered but so also is the human use of the landscape, so one might find more detailed information such as residential, commercial, and agricultural. Land use might also include information related to prescribed zoning, such as single-family residential or industrial.

Right-click the Landcover shortcut icon in the Catalog Tree and `Go To Target`. Expand LANDCOVER and click Polygon.

___ **14.** Click the Preview tab. In the Preview drop-down menu, make sure Geography is selected. Look at the map. Now select Table in the Preview drop-down menu. What is the FID number of the first record? _________. What is the number of the last record? _________. How many records (and, hence, how many land cover polygons) are there?[32] _______ Sort the records. What is the area of the smallest polygon (record all digits)? _______________ square meters.

___ **15.** Look at the Metadata for the polygon component of LANDCOVER as you did for the Soils geodatabase feature class before. Use the FGDC ESRI stylesheet. Using the information under the Spatial tab, record all of the following that you can. *The information that isn't available is shown by Xs in the following:*

`Projected coordinate system name _____XXX________________`

`Geographic coordinate system name _____XXX________________`

`Westmost bound: Geographic _____XXX_____ Projected: ________________`

`Eastmost bound: Geographic _____XXX_____ Projected: ________________`

`Northmost bound: Geographic _____XXX_____ Projected: ________________`

`Southmost bound: Geographic _____XXX_____ Projected: ________________`

___ **16.** So there is a little problem here. No coordinate system is listed. No geographic coordinates are shown. The bounding coordinates are present in "projected or local coordinates" but not in decimal degrees.

Although the lack of a published coordinate system and the absence of geographic bounding coordinates is an issue we will have to deal with, one mystery has been solved. It was the lack of latitude-longitude coordinates in the metadata that made it so the LANDCOVER coverage could not be found by `Edit > Search` Using `Geography`. This is why the ArcCatalog search was able to find the coverage SOILS but not the coverage LANDCOVER. The specific lesson: *The way ArcCatalog finds appropriate data sets by Geography is by investigating the metadata.* The general lesson: There are lots of ways to look for data—such as over the Internet (e.g., www.geographynetwork.com), through agencies (e.g., www.census.gov), through personal contacts, and so on. *When one method doesn't give you the results you want, try something else.*

Is the Newly Found Data Applicable?

Although the LANDCOVER coverage and the SOILS personal geodatabase probably occupy the same piece of real estate (note that the projected coordinates you wrote down previously are approximately the same for both), you have to make sure. Suppose that, by looking at some paper documentation, you are able to determine that the projected coordinates shown are in fact in meters and the projection of the coverage is

[32]The first record is numbered "2" because an entire polygon ArcInfo coverage an additional, hidden polygon, called the "outside polygon," which is the rest of the Earth—the area not included by the other polygons. It can be confusing. Perhaps for this reason the developers of ArcGIS have decided to leave it out of the table. But it remains polygon "1," so, to maintain consistency with the coverage records, the table starts with FID number "2."

WGS 84 UTM Zone 18, just like your other data. So you should fix up the coverage so it has the same spatial metadata as the other data sets. You can do this with ArcCatalog. Recall that in the original search you also found a *coverage* named SOILS; it was also in the

__:\IGIS-Arc_*YourInitialsHere*\Other_Data folder.

Maybe this SOILS coverage contains spatial metadata that you could assign to LANDCOVER.

___ **17.** In the __IGIS-Arc_*YourInitialsHere*\Other_Data folder, check the spatial metadata for the SOILS coverage against the spatial metadata for the Soils personal geodatabase feature class in the

__IGIS-Arc_*YourInitialsHere*\Wildcat_Boat_Data

folder. Do they have the same coordinates systems? ______

Making a Personal Geodatabase Feature Class from a Coverage

What you want to do is convert the LANDCOVER coverage to a Personal Geodatabase Feature Class so you will have all the data in the same format when you apply analysis tools to it in the future. Also, you need to fix the problem of not having the proper coordinate system associated with the landcover data.

___ **18.** Navigate to

___IGIS-Arc_*YourInitialsHere*\Other_Data\LANDCOVER\polygon

and right-click the icon. Choose Export > To Geodatabase (single). In the `Feature Class To Feature Class` window that appears, the `Input Features` should be set. You want to put the new feature class in ___IGIS-Arc_*YourInitialsHere*, so browse to the `Output Location` of Area_Features so that the window looks like Figure 1-34, making sure Area_Features appears in the Name text box. Click `Add`. For the `Output Feature Class Name`, type `Landcover`. Change the TRUE designations in the `Visible` column to FALSE for AREA, PERIMETER, LANDCOVER#, and LANDCOVER-ID, so that these columns will not show up in the table of the feature class, where they would be useless. See Figure 1-35. Press OK. When you see that the conversion has completed, close the `Feature Class To Feature Class` window.

Looking at the Landcover Personal Geodatabase

___ **19.** Navigate to the Area_Features personal geodatabase feature data set. Click it. Within it you should see Landcover feature class that you created from the LANDCOVER coverage. Select it. Click the Contents tab. Verify that the projection is WGS_1984_UTM_Zone_18.[33] Click the

[33]If projection is not shown, you will have to go back and reset options.

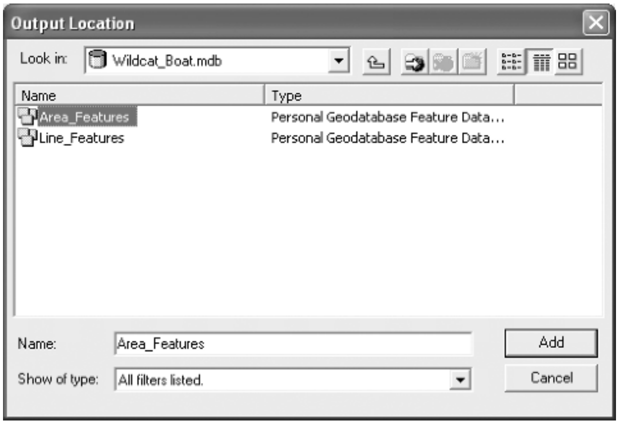


FIGURE 1-34

Preview tab. You see a map of the landcover you worked with in Assignment 1-1. In the drop-down menu in the `Preview` text box, select `Table`. How many land cover polygons are there? ____________ Check this against the number of polygons in the LANDCOVER coverage that you wrote down previously.

____ **20.** Examine the Landcover table. Notice the fields Shape_Length and Shape_Area. These attributes are the perimeters and areas of the polygons. Sort the Shape_Area column into descending order. The largest polygon is a portion of the lake, with a land cover code of 500. What is the perimeter of the polygon with the largest *land* area (to the nearest meter)? _____________ meters. What is the area of the smallest polygon (record all digits)? ______________________

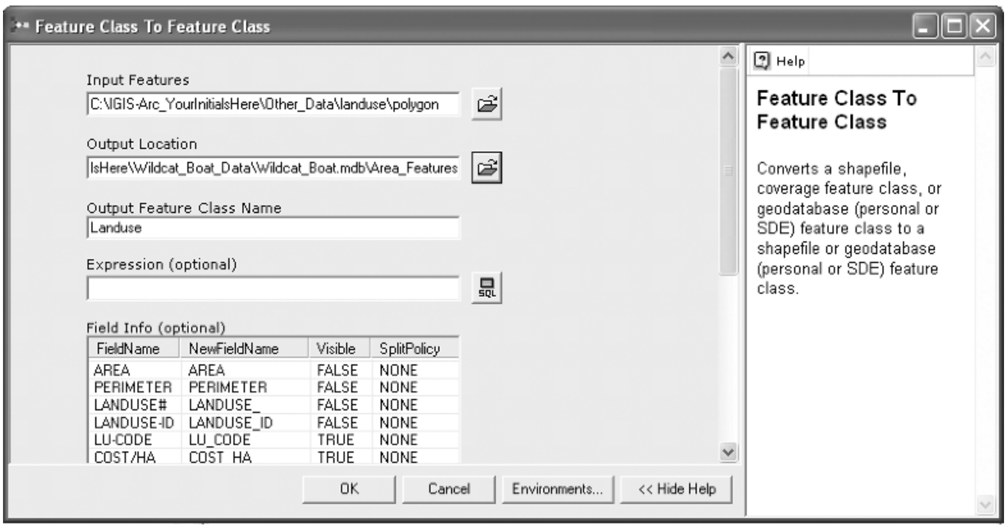


FIGURE 1-35

square meters. Contrast this "double precision" number with the one you previously recorded for the same polygon in the LANDCOVER coverage.

___ **21.** Find the Landcover feature class attributes and their properties with a right-click the Landcover icon. Then select Properties > Fields. Fill out the following table, except for Width.

Field Name	***Data Type***	***Width***
___________	___________	___________
___________	___________	___________
___________	___________	___________
___________	___________	___________
___________	___________	___________
___________	___________	___________

___ **22.** Cancel the personal geodatabase `Feature Class Properties` window. Find the names of the attributes by using the metadata, check them against what you wrote above, and add the Width. (*Hint:* When you click text that is green, you see more text. Clicking again contracts the text.)

Further Examining the Wildcat Boat Facility Area Data Sets

Do the following steps using the data sets in

___IGIS-Arc_*YourInitialsHere*.

___ **23.** In the Catalog Tree, click `Soils` feature class, within the `Area_Features` feature data set. Click the Preview tab.

This feature class uses polylines to enclose areas (polygons). Each of the polygons is considered to be homogeneous in a particular type of soil.

___ **24.** Using the Identify tool, click a polygon near the center of the map. The polygon becomes momentarily highlighted and the Identify Results window appears. You get the plane area of the feature in square meters in the field Shape_Area. You also get, in the field Shape_Length, the sum of lengths of the polylines (in meters) that enclose the polygon. In the attribute field named SOIL_CODE, you are shown a code that identifies the soil type, such as Tn4. There is also an attribute (SUIT) that indicates whether the soil is suitable for the construction of a building: 1 for unsuitable, 2 for moderately suitable, or 3 for suitable. Click some other polygons. Click the large polygon in the north-east (That's water—note the SUITability and the SOIL_CODE: ______, _____). Dismiss the Identify Results table.

___ **25.** Look at the attribute table for the Soils personal geodatabase feature class. Obtain statistics for the Shape_Area column of the table. What is the smallest area? _________ The largest? ___________ Sort the table from smallest to largest and verify the minimum and maximum numbers. What is the area of the largest polygon that is not water? ___________

___ **26.** Find the Soils attributes and their properties with a right-click Soils. Then select Properties > Fields. Fill out the following table, except for Width.

Field Name	***Data Type***	***Width***
___________	___________	___________
___________	___________	___________
___________	___________	___________
___________	___________	___________
___________	___________	___________
___________	___________	___________
___________	___________	___________

___ **27.** Cancel the personal geodatabase feature class `Feature Class Properties` window. Find the names of the attributes by using the metadata, check them against what you wrote above, and add the Width.

___ **28.** Click the Contents tab. Select the Line_Features personal geodatabase feature data set. Check out the presentation using the four display options on the Standard toolbar: Large Icons, List, Details (scroll to see them), and Thumbnails.

___ **29.** In the personal geodatabase feature data set Line_Features, select Roads. Make a thumbnail of the northeastern portion of it. (You may need to check your Fast Facts File to recall how to make a thumbnail.) Click the Contents tab and note the thumbnail and some general information about the feature class. What is its projection?[34] _________________

___ **30.** Click the Metadata tab and, with FGDC ESRI, look at the `description` of the data. Note the thumbnail again. Like much metadata you will find, there is a lot of missing information. Under `Details About This Document`, note the date and time the contents were last updated: Date __________ Time __________ (Cryptically, the date is given as YYYYMMDD and the time as HHMMSS00.)

___ **31.** Look at the Spatial characteristics of the data. What is the geographic coordinate system name? ______________________________ What is the projected coordinate system name? ____________________ To the nearest hundredth of a degree, what are the geographic coordinates of the northwest corner? ________________, ________________ The southeast corner? ________________, ________________.

___ **32.** In the projected coordinates, to the nearest meter, what are the coordinates of the northwest corner? __________, __________ The southeast corner? __________, __________ What are the units of Roads? _________________(*Hint:* Check under Details.)

[34]If the projection is not revealed, it may be that ArcCatalog reverted to its default options. Reset the options. (How to do this should definitely be in your Fast Facts File.)

___ **33.** Look again at the `Preview` of the Roads geography. Zoom to full extent. Then look at the table.

Here again, each row in the table corresponds to a geographic feature. In this case, each feature is a polyline, which is a series of straight-line segments that run between two points. Among the fields of data you see is Shape_Length, which is the sum of the segments that make up each polyline. This is also information that you "get for free" when you use a GIS. (You will see a much more comprehensive discussion of polylines and segments in the Overview of Chapter 4.)

___ **34.** Again look at the geography of Roads. Zoom up on the northernmost road. Notice that it is not a smooth curve but a series of straight lines. These lines begin and end at Cartesian points, which have *x-y* coordinates. Use the `Identify` tool to see the particular attribute values for the polyline that represents this road. Be sure that as you click the polyline, the feature is briefly highlighted. What is the length of the road to the nearest meter? ___________. Each polyline is numbered with an OBJECTID. What is the OBJECTID of this road? _________. The road code for this polyline is ________. Dismiss the `Identify Results` window. Zoom to the full extent of the Roads personal geodatabase feature class.

___ **35.** Find the Roads attributes and their properties. Fill out the following table, except for Width.

Field Name	***Data Type***	***Width***
___________	___________	___________
___________	___________	___________
___________	___________	___________
___________	___________	___________

___ **36.** Dismiss the personal geodatabase feature class `Feature Class Properties` window. Find the names of the attributes by using the metadata, check them against what you wrote above, and add the Width.

___ **37.** Look at the Sewers personal geodatabase feature class. How many polylines are there? __________ What are the diameters of the various sewer pipes (inches)? __________, __________ Make a thumbnail of Sewers and examine it with the Contents tab.

___ **38.** Find the Sewers attributes and their properties. Fill out the table below except for Width.

Field Name	***Data Type***	***Width***
___________	___________	___________
___________	___________	___________
___________	___________	___________
___________	___________	___________
___________	___________	___________

___ **39.** Cancel the personal geodatabase feature class `Feature Class Properties` window. Find the names of the attributes by using the metadata, check them against what you wrote above, and add the Width.

___ **40.** Look at the Streams personal geodatabase feature class. Make a thumbnail. What is the total length in *kilometers* of all the Streams in the study area? __________ What is the longest stream polyline in *meters*? _____

___ **41.** Fill out the following table, except for Width.

Field Name	***Data Type***	***Width***
__________	__________	__________
__________	__________	__________
__________	__________	__________
__________	__________	__________

___ **42.** Cancel the personal geodatabase feature class `Feature Class Properties` window. Find the names of the attributes by using the metadata, check them against what you wrote above, and add the Width.

Looking at Wildcat Boat Data with ArcMap

___ **43.** Launch ArcMap. Tile the ArcMap and ArcCatalog windows vertically. Drag the Sewers personal geodatabase feature class from ArcCatalog into ArcMap.[35] Now drag the Streams personal geodatabase feature class in. Put in the Landcover personal geodatabase feature class. Dismiss ArcCatalog. Make ArcMap occupy the full monitor screen.

Here you see one of the major bonuses of a GIS: the ability to see data of different types and from different sources easily represented on the same map. An additional advantage is that you can immediately choose what is shown and what is not. Notice that the Landcover data set is displayed "under" the Sewers and Streams—both in the Layers pane and, in a different sense, in the map window.

___ **44.** Experiment with displaying and un-displaying the different personal geodatabase feature classes by clicking the boxes next to the data set names.

___ **45.** Remove both Landcover and Streams by right-clicking the data set name and choosing Remove.

___ **46.** Display Sewers. Open the attribute table for Sewers. (Right-click Sewers. From the menu that appears, select Open Attribute Table. Is the way to open an attribute table in your Fast Facts File? If you didn't remember how, it should be.) Shorten the vertical dimension of the attribute table and position it so that you can see it and the graphic image of the Sewers as well. On the Tools toolbar (which resembles the Tools toolbar in ArcCatalog—zoom, pan, etc.), find the Select Features icon (not the Select Elements icon) and press it. Now click one of the sewer lines. It should turn cyan *and the equivalent row of the attribute table should be highlighted*. Pick another arc and observe. Sort the records according to Shape_Length. Click the shortest arc you can find on the map and verify its length in the table.

[35] If you run into "schema lock" problems, dismiss ArcCatalog and load the data set into ArcMap directly, using Add Data.

____ **47.** In the table, click in the small box to the far left of the last record. The selected record becomes highlighted *and so will the equivalent arc*.

We went through the last few steps to reinforce the point (that maybe you now understand so well you don't want to hear it anymore): A GIS is a special case of an information system whose database is the marriage of a (geo)graphic database and an attribute database.

Next you look very briefly at the capacity of ArcGIS to let you add a completely separate database table to the Table of Contents, view such a table, and then join it together with a feature class attribute table so that the information in the separate table becomes part of the attribute table.

____ **48.** Click the New Map File icon on the Standard toolbar. Don't save changes. Add the Landcover feature class from

___IGIS-Arc_*YourInitialsHere*\Wildcat_Boat_Data\Wildcat_Boat.mdb\Area_Features.

____ **49.** Open the Landcover attribute table. Notice that the table contains the column LC_CODE but no indication of what the codes mean. Close the table.

____ **50.** Add, as data, the database table

___IGIS-Arc_*YourInitialsHere*\Other_Data\LC_Code&Type.dbf.

Right-click LC_Code&Type and `Open` the table. What are the columns in Attributes Of LC_Code&Type?

__________________ , __________________, __________________

Dismiss the table window.

____ **51.** Right-click Landcover. Choose Joins and Relates > Join to bring up a Join Data window. You want to join attributes from a table. In box number one, the field in the layer Landcover that the join will be based on should be LC_Code. Such a field is called a *key field*. In box number two, the table you want to join to the layer is, of course, LC_Code&Type. In box number three, the field in the LC_Code&Type table also is named LC_Code.[36] Click OK. (You will be asked if your want to create an index for the join field. When you are processing thousands of records, this can ultimately save time. For our minor demonstration, we can ignore the offer. Select No.)

Seeing the Results of the Join

____ **52.** Using the Identify cursor, click a polygon in the map display. Notice that the amount of information you get is substantially greater than before. Specifically, you see not only the land cover code (LC_CODE) but also the type (LC_TYPE) that is associated with the code. The field names look a little strange, but we'll deal with that issue later. The important thing is that we have added the contents of one table to another. Dismiss the Identify Results window.

[36]The key fields in the two tables do not have to have the same name. In this case they happen to.

____ **53.** Open the attribute table of the Landcover layer. Notice that all the information has been put together. Sort the Landcover type. How may polygons consist of "Barren" land? _______

A note of warning: The product of the preceding process—the table with information gleaned from two tables—will persist only as long as ArcMap is running and Landcover remains as a layer. If you want to save the new table together with the geography of Landcover, you have to take additional steps, which I discuss later.

____ **54.** Dismiss ArcMap without saving Changes To Untitled.

Exercise 1-7 (Project)

Understanding the ArcGIS Help System

The Help system in ArcGIS has many facets. There are several ways to access different kinds of Help. Knowing how to access the various pieces of documentation will make your use of ArcGIS more efficient and less frustrating. You will start with a nifty mechanism for getting help on tools and buttons.

A Button for Instant Help: What's This?

____ **1.** Start ArcCatalog. Click the ^? tool. Then click the Connect To Folder icon on the Standard toolbar and read all about it. Since you know most of this information, I am mainly trying to show you what you can expect from the ^? tool. Click any blank area (or press Esc on the keyboard) to dismiss the information box.

Press Shift-F1. What happens? ____________________. Click the Search tool on the Standard toolbar. Dismiss the text explanation.

Click again on the ^? tool. Now click it again and read about the tool itself.[37] Try using Shift-F1 on a context menu item by first moving the cursor down to View > Refresh, then pressing Shift-F1. Read the box.

In the Catalog Tree, click Catalog. On the Main menu, click View. Notice that Identify Results, while shown as one of the selections, is disabled. To get an explanation of the option, and to find out why it is disabled, move your cursor over it and press Shift-F1. An explanatory note is presented. Frequently at the end of such notes is an explanation of why the tool or option is disabled. This can be extremely useful when you are presented with a grayed-out selection.

The notes obtained with the What's This tool and with Shift-F1 are completely separate from, and frequently more useful than, the regular Help system, which is described next.

[37]The reference to the Table of Contents has to do with ArcMap, which we explore in the next chapter.

The Help System and Documentation

Not only is it unlikely that you will ever learn all of ArcGIS, you may not even learn all there is to know about the Help system. But you should make a major effort, because whatever question you may have about the system probably has an answer somewhere in Help. There is an immense amount of information available. But finding it can seem somewhat like a scavenger hunt.

___ **2.** On the ArcCatalog Main menu, click `Help`. Pick `About ArcCatalog`. From the first line, you can tell what version of ArcCatalog you are running: ____________. From the second line, you can determine the License Type: ____________. The ESRI Web site is: ________________________________. Click OK.

___ **3.** On the ArcCatalog Main menu click Help. At the top of the drop-down menu, you see ArcGIS Desktop Help. We'll come back to that. Next is the help tool you just explored: `What's This`? Depending on which version of ArcGIS you are running, you will find the next menu items in different orders. One is called GIS Dictionary; the symbol next to it indicates that it is a Web page that you get to through the Internet, if you are connected.

___ **4.** Click `GIS Dictionary`. If your computer is connected to the Web and your browser cooperates, a page with the banner ESRI Support Center will appear. In a line that says "You are here:," you will be told that you are at the level of

Home > Knowledge Base > GIS Dictionary

Look up "geodatabase". Look up "shapefile". Look up "coverage". Look up "relational database".

This is but one page of a very rich site. Before leaving it, list the green tabs in the ESRI Support Center banner:

____________, ___________, ____________, ____________,

____________, ___________, ____________

Briefly explore each tab—particularly Home, Knowledge Base, and User Forums. You will get at least two impressions: (1) ESRI has a lot of products and services, and (2) there is a lot of help available. Dismiss the page by clicking the "X" in the upper right corner.

___ **5.** Click Help again on the Main menu. Click the first item in the menu: ArcGIS Desktop Help (which in the future you can invoke just by pressing the F1 key on the keyboard).[38] Make the window occupy the entire monitor screen. Click Contents. Click GIS Glossary and you will be served up with definitions of GIS terms that reside locally with your software. It is also extensive. To access its search capability, press Ctrl-F to bring up a Find window.

In this text I will not define many terms. Nor will I include a glossary. With all the capabilities you have to get definitions on demand, I decided it would be better to save a few trees. If you want a printed dictionary, you can obtain *The ESRI Press Dictionary of GIS Terminology*.[39]

---

[38]In version 9.1 (check to see which version number you wrote down above), you can also get to Help through the Internet by clicking ArcGIS Desktop Help Online. You can choose whether you want your help to come from your local machine or the Web.

[39]ESRI Press, Redlands, CA. Edited by Heather Kennedy. 112 pages.

___ **6.** Dismiss the GIS glossary. The remainder of the Help system under `Contents` is a hierarchical, nested set of "books." When you get down to the bottom of the hierarchy, you will find a help document. Using the "+" icon in front of the ArcCatalog book, open it. Within that, open `Working with metadata`. Click the document icon `About metadata`. A document of that name should appear. Read it. Then click the little triangle in front of `How metadata is organized` to expand that topic. Click it again to collapse the material. At the top of the pane click `Expand All`. Note the effect. Click `Collapse All`. Note that there are other links on this pane. Click `Related Topics`. Click `Learn About Creating Metadata`. To get back to the `About Metadata` pane click the `Back` arrow below the Main menu.

___ **7.** Check out `What's New In ArcGIS Desktop` 9.1, 9.2, or whatever version you are using. Look at the various book icons. Check out What's New In ArcMap.

___ **8.** Right-click any entry in the left pane. Click `Open All`. Scroll down through the pane. It may be daunting to see how much information there is, because of its implication for the size of the software package. That's why they will pay you the big bucks when you know how to use it. Right-click again and `Close all`.

___ **9.** Click the topmost entry: `Welcome To ArcGIS Desktop Help`. Read the `Getting Help` section.

___ **10.** Press the `Index` tab. Type `ADD` as the keyword. As you can see from the list that appears immediately below, there are lots of topics that relate to adding something. Double-click the `Add Rule To Topology` tool/command. An `Add Rule To Topology (Data Management)` document, with an ArcToolbox heading, opens. Notice the `Open Tool` button. Press it and wait. A window called `Add Rule To Topology` opens (you may have to display it by clicking its button on Windows the taskbar). From here you could run that tool (if you had any idea what it did—all things in good time). Cancel the `Add Rule To Topology` window. Now click `Expand All`. The topics—Usage Tips, Command Line Syntax, and Scripting Syntax—will display details about the command. Now click `Collapse All`. Click the triangle in front of Usage tips and note the result.

___ **11.** Look at the Main menu of the help window (File, Edit, View, Go). Check each drop-down menu, looking at the items on those menus. In the View menu, make sure checks are on next to Search and Highlights.[40] Try out `Back` and `Forward` under the Main menu bar.

___ **12.** Press the Search tab. Type

```
catalog tree thumbnail
```

Click List Topics. About how many topics are displayed? ______. Add the phrase `large icons` and press List Topics again. Select `An Overview Of ArcCatalog` and press Display. Also note the three option boxes at the bottom of the left pane that give you options in searching. Close the ArcGIS Desktop Help window.

There are other parts of the help documentation that you will encounter later, when you use ArcToolbox. And finally, ESRI has a number of tutorials on various parts of the software and on the software extensions. The documentation and tutorials consist of PDF[41] files located, probably, in C:\ESRI_Library. Data for the tutorials may, likely, be found at C:\arcgis\ArcTutor.

---

[40]The terms you search for using the search procedure that follows will be highlighted in the document.
[41]Portable Document Format files, readable by Adobe Reader (available free), a basic part of Adobe Acrobat, at www.adobe.com.

Exercise 1-8 (Dull Stuff)

Using ArcCatalog for Mundane Operations

ArcCatalog serves as an operating system (like UNIX or Windows) for geographic data sets. In this brief exercise, you will copy, paste, rename, and delete a coverage. If you tried to do this with the computer's operating system, you would cause errors.

___ **1.** With ArcCatalog running, highlight

___IGIS-Arc_*YourInitialsHere*

in the Catalog Tree. Select `File > New > Folder` and name the folder Housekeeping_Stuff. (Never, ever, accept the proffered name `New Folder`. Because it has a blank in it, ArcGIS may complain (and fail) later if it is used in a path name. Never use a folder name with a blank in it!

___ **2.** Navigate to

___IGIS-Arc_*YourInitialsHere*\Village_Data\HYDRANTS

and highlight it. Select Edit > Copy to place the HYDRANTS onto the ArcCatalog clipboard. (In place of Edit > Copy, you could type Ctrl-C, or you could press the Copy button on the Standard toolbar.)

___ **3.** Highlight the word Catalog in the Catalog Tree, and select View > Refresh (or just press F5). This ensures that the Catalog Tree properly reflects all data sets and puts them in order in the Tree. It is not always necessary to do this, but it doesn't cost much (a bit of time and computer processing). And it will save you on occasion from wondering where something went that was supposed to be there.

___ **4.** Navigate back to Housekeeping_Stuff and highlight it. Select Edit > Paste (or press Ctrl-V, or press the Paste button on the Standard toolbar). Refresh the Catalog Tree again, expand Housekeeping_Stuff, and display the Geography of the point component of this copy of the HYDRANTS coverage.

___ **5.** With HYDRANTS highlighted, right-click the name. Choose Rename. (Alternative ways to start the renaming process: press F2, or left-click the highlighted name). Type `NEW_HYDRANTS`, to change the name of the coverage.

___ **6.** Attempt to change the name again, this time to NEWER_HYDRANTS. This fails because the number of characters exceeds that allowed for coverages. Experiment to determine the number of characters that may exist in a coverage name. What is that number? _______[42]

___ **7.** Delete the coverage using File > Delete, or the Delete button the Standard toolbar, or via Delete on the drop-down context menu that you get by right-clicking the highlighted name. Also delete the folder Housekeeping_Stuff. Close ArcCatalog.

[42]Most geodatasets may have long names. Coverages, and some others, are exceptions.

ArcCatalog operates pretty much like Windows Explorer or the regular operating system windows when it comes to copying, moving, renaming, deleting, and so on. *Always use ArcCatalog when dealing with geodata sets, whether they be coverages, personal geodatabase data sets, or shapefiles.*

___ **8.** To prove the validity of the preceding admonition, use Windows to perform the same operations. Make a folder named Housekeeping_Stuff in the same place. Copy the HYDRANTS folder onto the Windows clipboard and paste it into the Housekeeping_Stuff folder. Then restart ArcCatalog and try to display HYDRANTS. Nothing doing. I'll explain why later, but, again, know this: You can't properly or effectively copy, paste, rename, or delete geodata sets outside of ArcCatalog.

___ **9.** Use ArcCatalog to delete the Housekeeping_Stuff folder. Close ArcCatalog.

Exercise 1-9 (Review)

Checking, Updating, and Organizing Your Fast Facts File

The Fast Facts File that you are developing should contain references to items in the following checklist. The checklist represents the abilities you should have upon completing Chapter 1.

Important note: This checklist is on the CD-ROM that accompanies the book. It is available in Microsoft Word format. Rather than typing or writing by hand the text that follows, you can copy and paste it into your Fast Facts File from the CD-ROM file.

Name: ____

My User or Logon Identifier: ____

My password is written down in this secure location: ____

The hard drive location of this FastFactsFile is ____

The FastFactsFile is backed up on ____

The date of the last update of the FastFactsFile is ____

The date of the last reorganization of the FastFactsFile is ____

The path to IGIS-Arc, associated with [__], is _______________

The path to IGIS-Arc_*YourInitialsHere*, associated with ______, is _______________

Operations using ArcGIS Desktop software:

___ To initiate ArcCatalog

___ To determine the level of ArcGIS (ArcView, ArcEditor, or ArcInfo)

___ To see properties of an entry in the Catalog Tree

___ To set options for ArcCatalog

___ To expand or contract entries in the Catalog Tree

___ To copy, paste, delete, rename, make a new something (folder, shapefile, personal geodatabase feature class, and so on), search, or reveal the properties of the entry in the Catalog Tree

___ To see the entire path of an entry in the Catalog Tree

___ To connect to a folder so that it appears as a single entry in the Catalog Tree

___ To disconnect from a folder

___ To turn toolbars off and on

___ To move toolbars around

___ To turn on ToolTips

___ To get specific help on a tool or command

___ To launch ArcMap or ArcToolbox from ArcCatalog

___ To examine and modify metadata

___ To see the components of a coverage (point, line, polygon)

___ To see the attribute table of a geographic data set

___ Ways of selecting a row in a relational database table as the current record are

___ An Internal-ID is different from a User-ID in that

___ To change the appearance of a table cosmetically

___ To get statistics on the numeric values in a table column

___ To sort the values in a column

___ To arrange the records order using both a primary and secondary sort

___ To move a table column

___ To search for values in a table

___ To make a column always visible

___ To look at the graphics of a geographic data set

___ To look at the items in a table that relate to a graphically represented feature

___ To see the coordinates that the cursor is pointed at

___ To see the characteristics and parameters of the attributes of items in a table

___ To magnify the graphics of the area being examined

___ To move around on a map that is being examined

___ To make thumbnails of geographic data sets

___ To determine the locations of tics in a geographic data set

___ To explore the three areas of metadata: description, spatial, and attributes

___ To select among various metadata stylesheets

___ Three ways to get data sets into ArcMap are

___ Three ways of initiating the data set copying process in ArcCatalog are

___ Three ways of initiating the data set renaming process in ArcCatalog are

___ Three ways of initiating the data set deleting process in ArcCatalog are

___ Ways of searching for geographic data sets with ArcCatalog are

___ Find the properties of a geographic data set

___ In ArcMap, to highlight both a selected feature and the corresponding row in the table

What's Next?

In this chapter, you used ArcCatalog to find and explore geographic data sets and to install those data sets into ArcMap. In the next chapter, you begin working with ArcMap extensively—looking at examples of the wide variety of GIS data available.