CDC 3E551

Engineering Journeyman

Volume 2. Geographic Information Systems

Air Force Career Development Academy

Air University

Air Education and Training Command

3E551 02 1905, Edit Code 03

AFSC 3E551

Author: MSgt Raymond L. Brown III
368th Training Squadron
United States Air Force Engineering School (AETC)
368 TRS/TDE
6007 Cooley Avenue
Fort Leonard Wood, Missouri 65473
DSN: 581-7950
E-mail address: raymond.l.brown180.mil@mail.mil

With contributions from:

TSgt Robert Jacobs, 368 TRS/Engineering Instructor

Instructional Systems Specialist: Dexter L. King

Editor: Sandra F. Glenn

Air Force Career Development Academy (AFCDA)
The Air University (AETC)
Maxwell AFB, Gunter Annex, Alabama

Material in this volume is reviewed annually for technical accuracy, adequacy, and currency. For Weighted Airman Promotion System (WAPS) Specialty Knowledge Test (SKT) purposes, the examinee should check the *Enlisted Promotions References and Requirements Catalog* to determine the correct references to study.

THIS IS THE SECOND of four volumes for the career development course (CDC) 3E551 Engineering Journeyman.

The material in it pertains to the engineering technology principles that make up the foundation of this career field.

Unit 1 covers basic principles concerning geographic information systems (GIS). Unit 2 covers changing data formats. Unit 3 covers management of databases.

A glossary is included for your use.

Code numbers on figures are for preparing agency identification only.

The use of a name of any specific manufacturer, commercial product, commodity, or service in this publication does not imply endorsement by the Air Force.

To get a response to your questions concerning subject matter in this course, or to point out technical errors in the text, unit review exercises, or course examination, call or write the author using the contact information provided in this volume.

NOTE: Do not use Air Force Instruction (AFI) 38-402, *Airmen Powered by Innovation and Suggestion Program,* to submit corrections for printing or typographical errors. For Air National Guard (ANG) members, do not use Air National Guard Instruction (ANGI) 38-401, *Suggestion Program.*

If you have questions that your supervisor, training manager, or education/training office cannot answer regarding course enrollment, course material, or administrative issues, please contact Air University Educational Support Services at http://www.aueducationsupport.com. Be sure your request includes your name, the last four digits of your social security number, address, and course/volume number.

For Guard and Reserve personnel, this volume is valued at 8 hours and 2 points.

NOTE:

In this volume, the subject matter is divided into self-contained units. A unit menu begins each unit, identifying the lesson headings and numbers. After reading the unit menu page and unit introduction, study the section, answer the self-test questions, and compare your answers with those given at the end of the unit. Then complete the unit review exercises.

Page

Unit 1. Geographic Information Systems Basics

ENGINEERING IS FOCUSED on information of both its collection and storage. The geographic information systems (GIS) we use are built on certain fundamental concepts. Further, we utilize GIS tools for higher headquarters and for installation situational awareness. The installation commander's situational awareness extends into the future development of the installation. The installation development plan (IDP) gives the commander a way to communicate and receive information about each unit's mission. The information on unit missions is maintained and recorded as a part of mission data sets (MDS) that makes up the IDP. The way that the IDP and MDS are stored and organized falls within our realm of expertise. We have a common way to organize the information and a centralized place to store it.

1–1. Geographic Information System Principles

Before we build our example geodatabase, we first need to cover some basic principles. This section discusses the fundamental concepts of geodatabase types, data types, and basic organization. We also cover the central installation product of our GISs—the IDP. When talking about the IDP, it makes sense to discuss what it is made of, which is the MDSs.

201. Fundamentals of geographic information system

There is so much information these days that our greatest task is organizing and making sense out of it. In the Air Force, we have many tools to get a handle on all of our data. Within the engineering specialty, we focus on GISs. GISs are a platform to organize information in relation to a spatial reference. A spatial reference is another way to say coordinate system or location. Giving our data, a place in the world allows us to display, visualize, analyze, and interpret in any way we need. A good example is the combination of security forces crime data, combined with civil engineer facility and infrastructure data. Crimes and facilities share places in the world. If a facility or area of an installation or city has a higher rate of crime than others do, the security forces commander can increase patrols in that area. Another more common example is traffic density overlaid on the road infrastructure. In this case, the usefulness is two-fold. First, engineers can predict that areas of higher traffic need frequent repairs meaning higher maintenance costs. Second, engineers can design the roads to be re-constructed to better handle the traffic and reduce the maintenance costs. Simply put, a GIS allows us to manage our assets and the taxpayer's money more responsibly.

Data organization

The first thing we must understand is that a GIS is data. Knowing how that data is organized is the first step in understanding how a GIS works. We can split data organization into two parts—the map and the data container. Each has different terminology and purposes behind their organizations.

Data container

The data container houses all of the data for an organization. In civil engineering, this means roads, facilities, trees, water systems, and anything that makes up the physical characteristics of the installation. The data container also stores non-physical objects, such as the installation boundary, quantity-distance arches for explosives, and airfield imaginary surfaces. Point, line, and polygon features representing these objects are stored in feature classes, such as road_area. Then, we group feature classes into datasets such as transportation_vehicle (fig. 1–1).

Data containers are most commonly geodatabases. Geodatabases are broken into three types based on use being Enterprise, File, or Personal Geodatabases.

Figure 1–1. Data and feature sets.

Enterprise Geodatabase

An Enterprise Geodatabase is stored on a network server accessible to many users simultaneously. They allow multiple users viewing or reading the data without changing it, and multiple users editing or writing the data at the same time.

File and Personal Geodatabases

File and Personal Geodatabases are stored on a standard computer hard drive. This reduces the number of editors and removes the ability for multiple users to modify the data. A File Geodatabase can have multiple editors but only if they work on different parts of the data. Personal Geodatabases are stored in a Microsoft Access data file format. This means that Personal Geodatabases depend upon the Windows operating system.

Map data

A data container is only useful for organizing the information in a common way. To make use of the information in a geodatabase, we display it on a map. The common installation picture (CIP) is the standard map (fig. 1–2). It is comprised of facilities, physical features, and utilities within the installation boundary. It displays the most basic types of data point, line, polygon, and raster in layers. The layer order puts the imagery behind the roads and parking areas and those behind facilities, and places utilities and the most meaningful data on top.

Vector data

Vector data is the name we apply to points, lines, and polygons as type groups (fig. 1–3). That is points are vector data, lines are vector data, and polygons are vector data. Each type is stored separately. Point data, such as fire hydrants, cannot be stored in the same feature file as line data, such as fences. Polygon data is the same. The file type is a way for the data container to communicate to the mapping software on how to display the information in that feature.

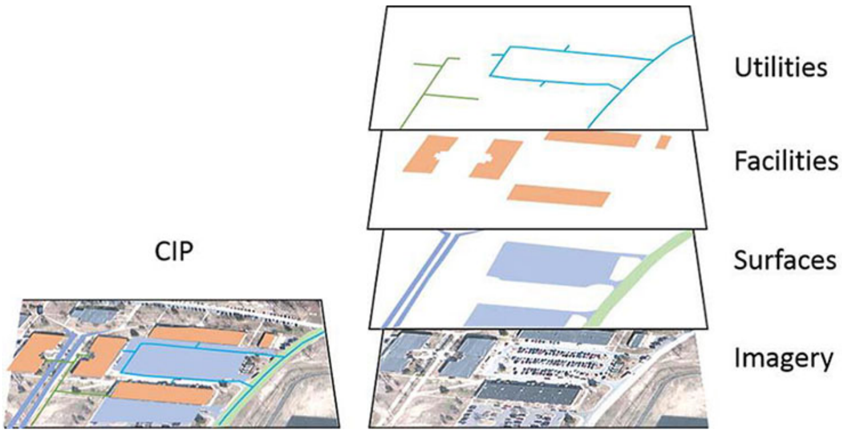

Figure 1–2. Map data layer organization.

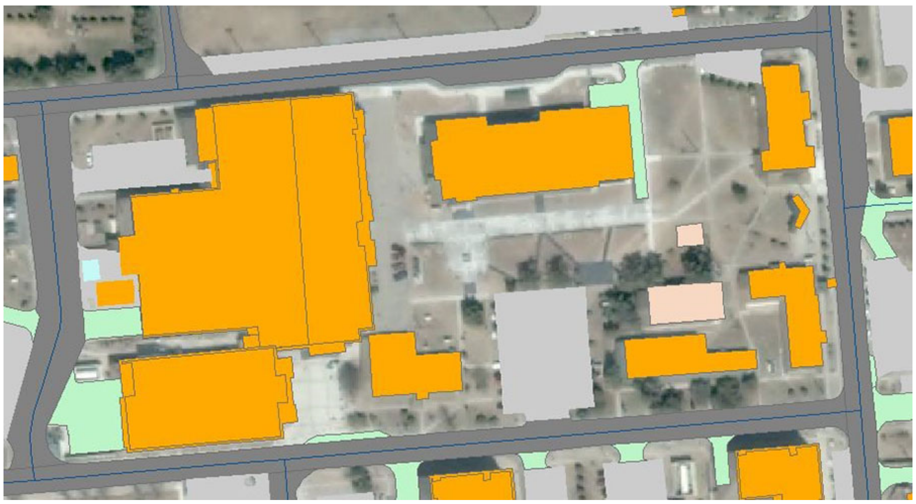

Figure 1–3. Vector data on top of imagery.

Raster data

Raster data refers most often to imagery (fig. 1–4). Raster imagery is a grid of squares each colored differently. Imagery has two unique characteristics, which causes it to be a separate data type. The first is the color of each grid square stored in the raster file in a table. The second is the ratio of the size of the image to the number of grids per unit, called resolution. A 10-kilometer (km) by 10 km image of land with five grids, or pixels, per meter is of a lower resolution than the same image with 50 pixels per meter.

Figure 1–4. Zoomed raster grid.

Attribute data

So far, we have talked about the basics of data organization and display. However, the most powerful part of a GIS is in the attribution. Attribute data gives each feature unique characteristics. Road data is a simple example. First, we have polygon geometry, which describes to the software how to display the shape. Second, their attributes give each polygon a name. Another attribute could further identify them based on whether they are made of asphalt or concrete. Yet, another attribute can tell us when the road was constructed. Another name for attribute data is tabular data. The name tabular refers to the data organized into a table (fig. 1–5). The tables are broken into rows for each feature, giving each one a unique number or feature identification. The columns of the table each refer to a different characteristic of the features, which we call fields. This is the name, material, and date constructed we talked about above.

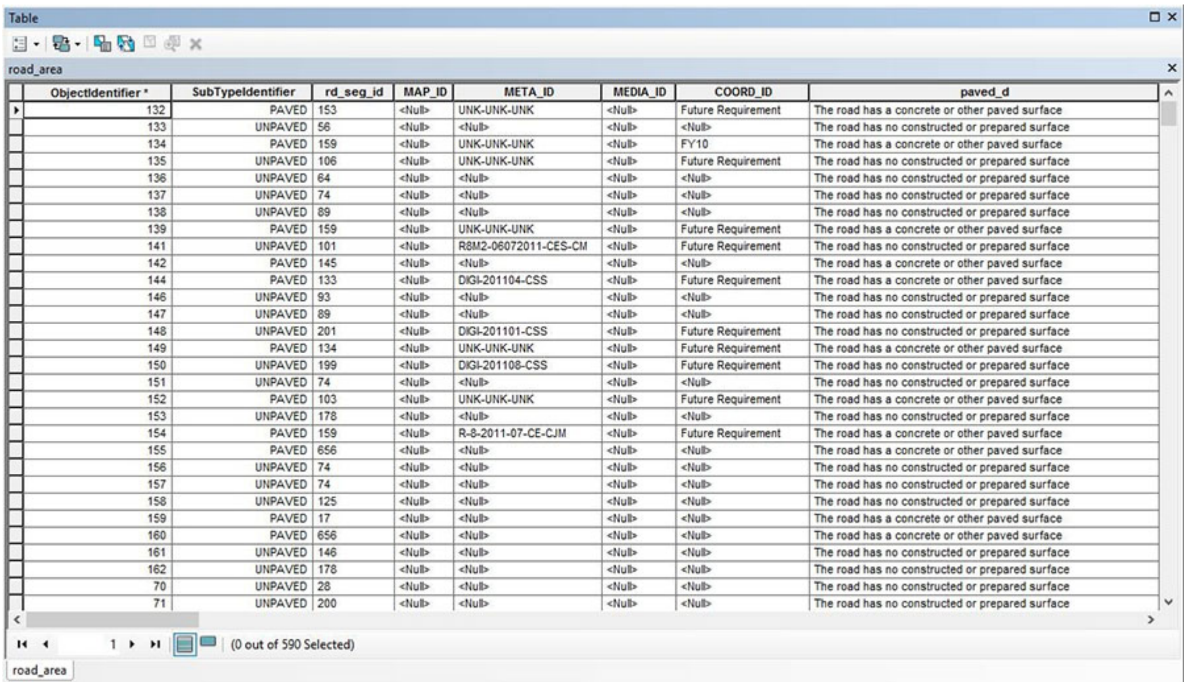
Table
road_area

ObjectIdentifier *	SubTypeIdentifier	rd_seg_id	MAP_ID	META_ID	MEDIA_ID	COORD_ID	paved_d
132	PAVED	153	<Null>	UNK-UNK-UNK	<Null>	Future Requirement	The road has a concrete or other paved surface
133	UNPAVED	56	<Null>	<Null>	<Null>	<Null>	The road has no constructed or prepared surface
134	PAVED	159	<Null>	UNK-UNK-UNK	<Null>	FY10	The road has a concrete or other paved surface
135	UNPAVED	106	<Null>	UNK-UNK-UNK	<Null>	Future Requirement	The road has no constructed or prepared surface
136	UNPAVED	64	<Null>	<Null>	<Null>	<Null>	The road has no constructed or prepared surface
137	UNPAVED	74	<Null>	<Null>	<Null>	<Null>	The road has no constructed or prepared surface
138	UNPAVED	89	<Null>	<Null>	<Null>	<Null>	The road has no constructed or prepared surface
139	PAVED	159	<Null>	UNK-UNK-UNK	<Null>	Future Requirement	The road has a concrete or other paved surface
141	UNPAVED	101	<Null>	R8M2-06072011-CES-CM	<Null>	Future Requirement	The road has no constructed or prepared surface
142	PAVED	145	<Null>	<Null>	<Null>	<Null>	The road has a concrete or other paved surface
144	PAVED	133	<Null>	DIGI-201104-CSS	<Null>	Future Requirement	The road has a concrete or other paved surface
146	UNPAVED	93	<Null>	<Null>	<Null>	<Null>	The road has no constructed or prepared surface
147	UNPAVED	89	<Null>	<Null>	<Null>	<Null>	The road has no constructed or prepared surface
148	UNPAVED	201	<Null>	DIGI-201101-CSS	<Null>	Future Requirement	The road has no constructed or prepared surface
149	PAVED	134	<Null>	UNK-UNK-UNK	<Null>	Future Requirement	The road has a concrete or other paved surface
150	UNPAVED	199	<Null>	DIGI-201108-CSS	<Null>	Future Requirement	The road has no constructed or prepared surface
151	UNPAVED	74	<Null>	<Null>	<Null>	<Null>	The road has no constructed or prepared surface
152	PAVED	103	<Null>	UNK-UNK-UNK	<Null>	Future Requirement	The road has a concrete or other paved surface
153	UNPAVED	178	<Null>	<Null>	<Null>	<Null>	The road has no constructed or prepared surface
154	PAVED	159	<Null>	R-8-2011-07-CE-CJM	<Null>	Future Requirement	The road has a concrete or other paved surface
155	PAVED	656	<Null>	<Null>	<Null>	<Null>	The road has a concrete or other paved surface
156	UNPAVED	74	<Null>	<Null>	<Null>	<Null>	The road has no constructed or prepared surface
157	UNPAVED	74	<Null>	<Null>	<Null>	<Null>	The road has no constructed or prepared surface
158	UNPAVED	125	<Null>	<Null>	<Null>	<Null>	The road has no constructed or prepared surface
159	PAVED	17	<Null>	<Null>	<Null>	<Null>	The road has a concrete or other paved surface
160	PAVED	656	<Null>	<Null>	<Null>	<Null>	The road has a concrete or other paved surface
161	UNPAVED	146	<Null>	<Null>	<Null>	<Null>	The road has no constructed or prepared surface
162	UNPAVED	178	<Null>	<Null>	<Null>	<Null>	The road has no constructed or prepared surface
70	UNPAVED	28	<Null>	<Null>	<Null>	<Null>	The road has no constructed or prepared surface
71	UNPAVED	200	<Null>	<Null>	<Null>	<Null>	The road has no constructed or prepared surface

1 (0 out of 590 Selected)
road_area

Figure 1–5. Attribute table.

202. Installation development plan

One of the greatest responsibilities installation commanders have is to plan for the future of their base. Whether deployed or in garrison, wing commanders are the governors of their area of responsibility. To assist the installation/wing commander in planning, each base has an IDP. The IDP shows present and future expansion, constraints, and effects the installation has on its environment. While many of the parts that make up the IDP are the responsibility of other organizations, the representation of the information and storage in a GIS is the responsibility of every Engineering flight. We accomplish this by using the structure of the IDP to break it into parts by subject matter. We call these parts MDSs. Each of these sets are broken down further into the specific database layers that make them up, referred to as mission data layers (MDL).

Mission data sets

According to Air Force Instruction (AFI) 32–7062, *Comprehensive Planning*, there are nine primary MDSs. These are the CIP, land use, airfield operations, transportation, constraints, future development plan, utilities, communications, and energy.

Common installation picture

The CIP gives the location of the installation in the world referenced to the nearest major city. It consists of the roads, buildings, airfield pavements, and other basic infrastructure that make up the base. We, then, place these layers on aerial photography. Categories for the MDL are installation layout, geographically separated unit (GSU), regional location, vicinity location, and aerial photography.

Land use

The land use section represents the planning districts of the installation to include GSUs. Present and future use descriptions accompany each planning district. This includes areas in just outside the installation boundary. Categories include planning districts, existing land use, future land use, vicinity existing land use, vicinity existing zoning, developable opportunity areas, and real estate.

Airfield operations

The airfield operations maps and plans include the airfield and the surrounding airspace. The categories are on-base obstruction to airfield and airspace criteria, approach and departure zone (obstructions to 10,000 feet), approach and departure zone (obstructions beyond 10,000 feet), airspace obstructions (vicinity, airfield pavement plan), aircraft parking plan, and airfield lighting systems.

Transportation

Transportation includes everything from roads to bike lanes, interstate systems around the base, and sidewalks that run throughout the base. This section has descriptions and graphics for every fabricated transportation route and access point in and around the installation. The categories are community network access to base, on-base street network, and future transportation plan.

Constraints

Constraints are limitations both physical and regulatory. Native-American burial grounds, endangered species habitats, flood plains, rivers, and lakes are examples of physical constraints. A good example of a regulatory constraint is the local airspace clearance requirements, meaning how close things can be constructed or placed to the airfield. The categories for constraints are cultural resources, natural resources, installation restoration program, wastewater discharge, storm water discharge, fuel/chemical storage tanks, drinking water supply sources, electromagnetic and radiation sources, airfield and airspace clearance criteria, antiterrorism/force protection, and explosive safety quantity-distance.

Future development plan

Future development plan is a review of the current desired improvements and additions to the installation. The plans are organized by short (1-5 years), medium (6-10 years), and long (20 years) periods. This section also includes alternatives to future development. Categories are short-range development plan, medium-range development plan, long-range development plan, and alternative scenario plans.

Utilities

The utilities part of the IDP involves all the systems required to sustain life and industrial mission operations—everything from drinking water to electricity and natural gas. Categories include water distribution system, sanitary sewerage system, storm drainage system, electrical distribution system, central heating/cooling system, natural gas distribution system, liquid fuels system, and industrial waste and drain system.

Communications

These plans cover an installation's physical communications infrastructure. Everything from cables, manholes, and antennae to equipment rooms, vaults, and repeaters. The only category is base-wide communications, navigational aids (NAVAID), and weather systems.

Energy

Energy deals with power generation and any associated electrical transmission lines or piping systems. Categories are existing generation facilities and future generation facilities.

Self-Test Questions

After you complete these questions, you may check your answers at the end of the unit.

201. Fundamentals of geographic information system

1. What is the difference between the Enterprise, File, and Personal geodatabases?

2. List the three types of vector data.

3. Define the term resolution in reference to raster data.

4. Give some examples of attribute fields for roads.

202. Installation development plan

1. According to AFI 32–7062, *Comprehensive Planning*, what are the nine primary mission data sets?

2. What are the categories of MDL?

1–2. Basic Geodatabase Construction

This unit's final section covers the steps to take to assemble a basic geodatabase from beginning to end. We start by creating and customizing our schema using the online schema builder. Then, we create an empty container to which we attach the schema. Finally, we will end this section by discussing the basics in creating vector data.

203. Utilize browser

By now, you have a basic understanding of a GIS. You also understand what we use the information for such as IDPs. The next step is to learn how to construct and maintain a GIS. The best way to learn that is to create one ourselves.

The first thing we need is a framework for our GIS, or a schema. The framework will allow our GIS to match those of others and to ease the sharing of data between our database and theirs. In the Department of Defense, we use the Spatial Data Standards for Facilities, Infrastructure, and Environment (SDSFIE) Website located at www.sdsfieonline.org to build our schema. The SDSFIE Website is a well-maintained industry standard for organizing schema.

SDSFIE online

When you navigate to the Web site, a Web page similar to figure 1–6 displays. The first step in using this system is to register an account. This may take some time because the site moderators will review your need to access and make a decision. Once you have an account and login, you can begin to build your own GIS database. The tool for this is under the “Models & Workflows” dropdown list and labeled “Browse Generate.”

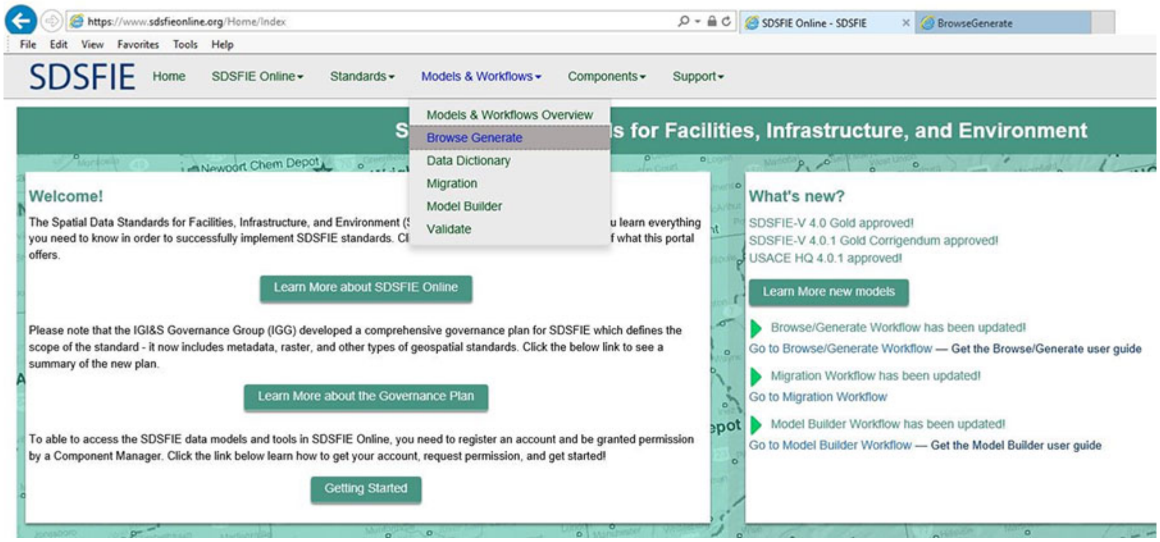

Figure 1–6. SDSFIE online.

Browse generate

The left pane of the Webpage is where we start (fig. 1–7). The first things we see are two radio buttons marked “Approved” and “My Drafts,” respectively. The “Approved” button allows you to select those data models approved for your service component—the Air Force. “My Drafts” lists any data models that you have created and saved. We select “Approved,” then expand the data model list under the title “Select a data model.” The list contains all the data models that the Air Force has approved for our use. A data model is a style of organization for a database and is subject to change. The most obvious difference between data models is the naming convention used for feature classes. Older models have a layer called “structure_existing_area,” but newer models simply have a “buildings” layer for simplicity. The data model we select for our example is “SDSFIE 4.0 Gold.”

In the “Model Elements,” section of the pane is a search bar. When left blank, the Webpage will list all feature classes that are part of the selected data model (fig. 1–8). These lists are very long, and when we know exactly which features we need to store in our database, we can search for them for ease of access.

When a feature class is selected, a list of the attributes associated with it displays to the right. Each attribute has a short description (fig. 1–9). In our example, there is an attribute called “roadName.” The description indicates that this attribute is to contain the road’s name. The next column describes the type of information that the attribute accepts. The “roadName” attribute takes a string data type. A string is a grouping of characters such as “Jefferson” or “engineer01.” The “Model Name” column gives us the name of the attribute and the “Symbols” column indicates limitations and specifics of each attribute. Placing the cursor over any of these symbols will display a brief explanation.

For our database, we have selected the “Building,” “PavementSectionRoadway,” and “RoadCenterline” feature classes. Satisfied with our selection, we now need to generate the extensible markup language (XML) file that we will use to implement our framework. To do that, we select the

“Generate” button in the left pane. This displays a dropdown list asking which type of model we want to generate physical, logical, or metadata.

Physical model

A physical data model differs from the other data models by use. It can be implemented and set up for use. It is also ready to be filled with information and to create relationships between that information.

Logical model

Logical models are general data models used to design physical models. If the physical model is a constructed house with furniture and people, the method of construction is the logical model. An example is a physical model with a “Buildings” layer. A logical model connects that layer to other layers and explains how each attribute behaves. The specifics, like whether an attribute is for road names or construction date information, are part of the physical model.

Metadata model

Metadata is the data about the data. The simplest example is survey information. The data collected is the location and elevation data. The metadata is obtained, we surveyed the information, and, then, stored in the database. The metadata has its own model and is a database of information about the database.

We do not need a logical model because we do not need to create a new physical model. We need a physical structure for our database example so we will hover over “Physical Model” and select “ESRI XML Workspace Document (Geodatabase XML),” which opens the “Generate XML Dialog.”

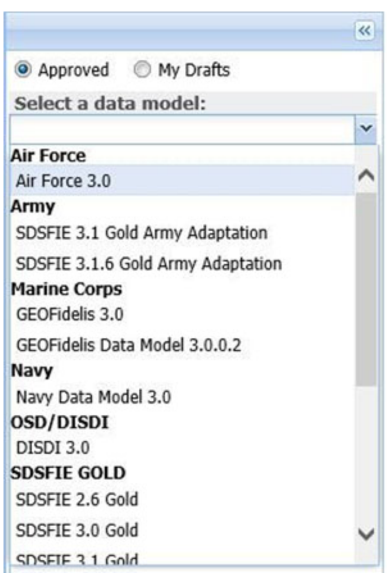

Figure 1–7. Data model list.

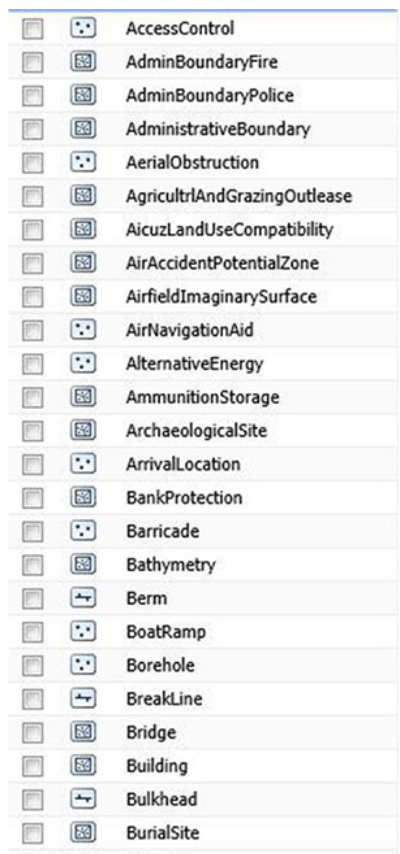

Figure 1–8. Feature class list.

Attributes | View Requirements

Symbols	Model Name	Definition	Data Type	Len...
	addressNumber	The number of the building used in the address.	String	10
	addressNumberSuffix	The extension of the number of the building used in the address.	String	4
	buildingIdpk	Primary Key. A unique, user defined identifier for each record or instance of an entity.	String	40
	cityOrCommunityName	Valid service community name as identified by the MSAG.	String	32
	facilityNumber	The locally developed asset number used for identification of the feature.	String	10
	featureDescription	A narrative describing the feature.	String(MAX)	
	featureName	The common name of the feature.	String	80
	installationId	The official code assigned by the Military Service (includes Washington Headquarters Services) to identify the site or group of sites that make up an installation.	String	11
	isHeritageAsset	A yes or no indicator of whether the building has historical, cultural, or archaeological value and/or significance.	YesOrNo	3
	mediaId	Used to link the record to associated multimedia records that reference data.	String	40
	metadataId	Used to represent or link to feature level metadata.	String	80
	roadName	The name of the street or road without prefix or suffix.	String	255
	roadPostDirectional	The directional suffix associated with the road name.	String	3
	roadPrefixDirectional	The directional prefix associated with the road name.	String	3
	roadSuffix	Valid Street abbreviation, as defined by the US Postal Service Publication 28. (e.g. AVE)	String	10
	rpaConstructionType	The type of construction, with respect to permanence, for the real property facility.	String	4
	rpaPredomCurrentUseCatCode	The real property category code (CATCODE) that classifies the current use of the asset and corresponds to the largest quantity of the asset that is utilized for a single purpose.	String	5
	rpInterest	A designator used to identify the type of legal interest that DoD holds in a real property asset.	String	4
	rpnid	The designator that distinguishes one real property network (RPN) from another within a RPI database.	String	4
	rpsuid	A non-intelligent code used to permanently and uniquely identify a DoD real property site.	Integer	
	rpuid	A non-intelligent code used to permanently and uniquely identify a DoD real property asset.	String	18
	sdsId	A unique identifier for all entities in the SDSFIE.	GUID	
	stateName	Alpha state abbreviation.	String	5

Figure 1–9. Attribute list.

Generate XML

This window has many options for configuring our data framework. The most important one is the "ESRI Software Version" dropdown list. This makes the XML file compatible with the version you use. The version we use in our example is 10.4. Since we are setting our data up to be compatible with any other Air Force geodatabase, we are going to leave the defaults settings for the other options. Finally, we will select "Export to XML" (fig. 1–10).

The export creates a compressed folder containing three files. One file is an instruction "readme.doc" explaining how to import the file into a database. The next is a .tbx file, which adds the programs needed to import the XML file into a database, if we do not have them already, which we do. The last file is the XML file "import–1.xml" containing the framework, or schema, for our geodatabase.

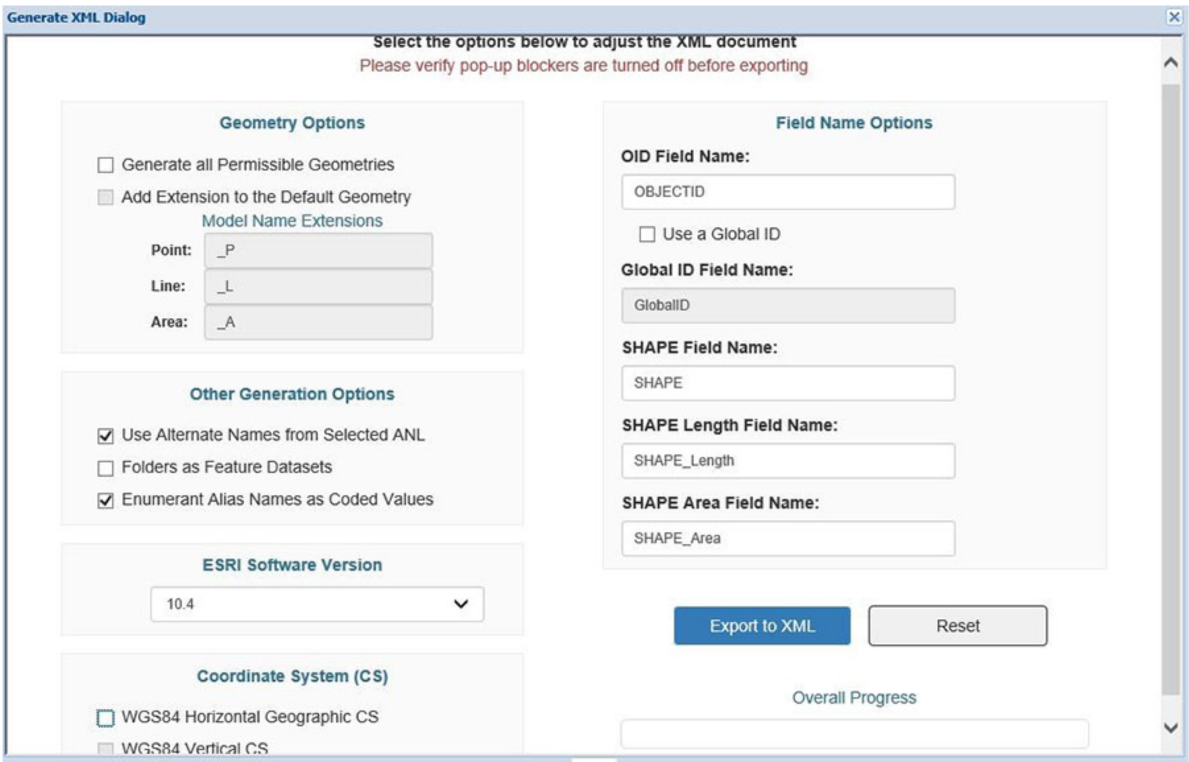

Figure 1–10. Generate XML dialog.

204. Creating a geodatabase

So far, we have a framework for our database, but we have not created a database. Our GIS software is broken into three parts—map, catalog, and toolbox. Map displays the database as graphics, catalog is an administrative tool similar to the File Explorer program that Windows uses, and toolbox is a set of specialized functions for manipulating and editing data. To create our database and add our physical model schema to it, we use catalog.

Create a data container and apply schema

Unlike File Explorer, catalog connects to folders instead of being a part of the file system. The first thing we have to do is connect to the folder that houses our database. The second icon from the left called "Connect to Folder" opens a file navigation window allowing you to select and connect to the appropriate folder (fig. 1–11).

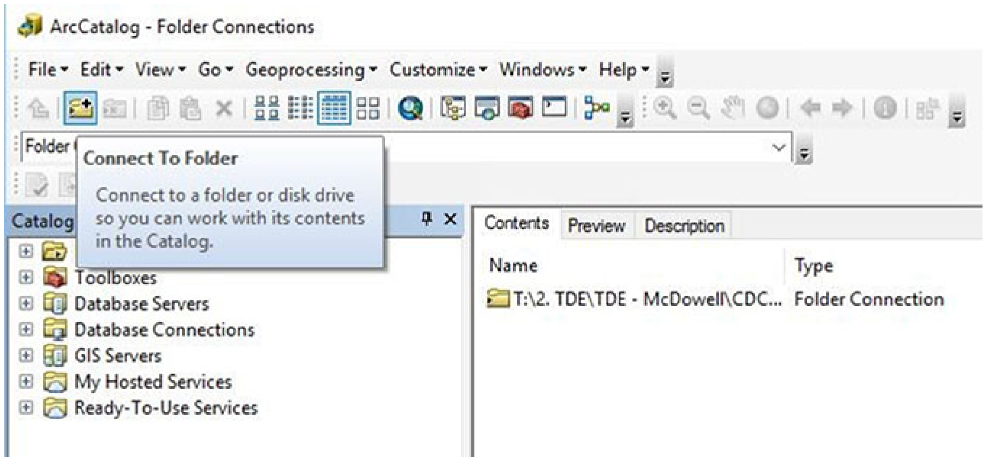

Figure 1–11. Connect to Folder.

The folder connection appears under the “Contents” tab of the right pane. We can now access the folder and its contents through Catalog. To create a file or personal geodatabase, enter the folder, and right-click in the contents view pane (fig. 1–12). We are using the file geodatabase because we need more than one person at a time editing the data, and we, therefore, select “file geodatabase.”

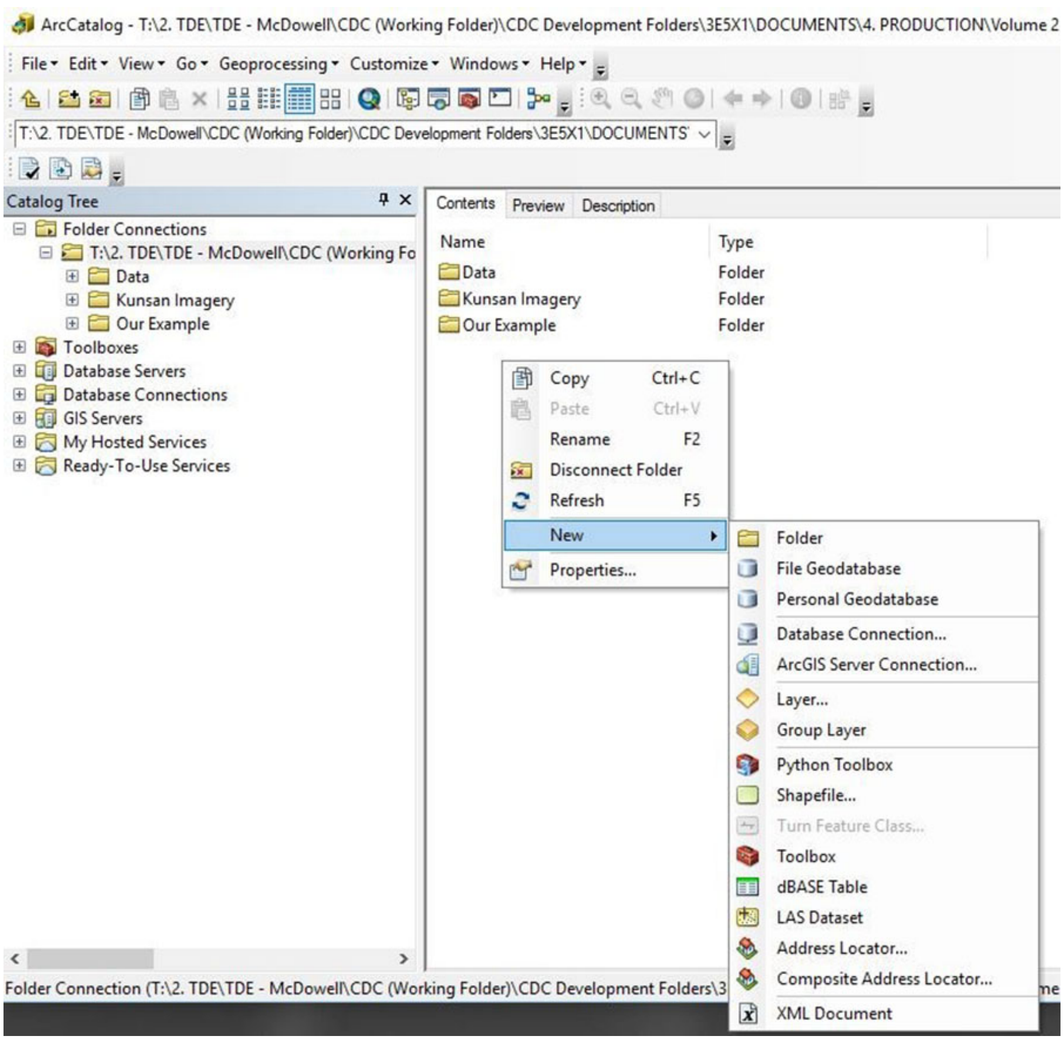

Figure 1–12. Create a file geodatabase.

The last step we need to finish setup of our geodatabase is to add the schema, or framework, to it. To do this, right-click on the geodatabase we just created in the contents pane and hover over “Import.” When an expanded list of import options appears at the bottom of the list, select the item labeled “XML Workspace Document…,” (fig. 1–13).

An import wizard will open to guide us through the import process (fig. 1–14). Always double-check everything. Start with the file path listed next to “Importing data to:” ensuring that it is what you want. Then, under “What do you want to import:” select either “Data” or “Schema.” XML files are able to contain both schema and data. Say we exported an XML copy of a database from a personal geodatabase and wanted to add it to our own. Choosing the “Data” option tells the wizard to add both the schema, or framework, and the data contained in the document with it. “Schema Only” adds only the empty framework to the database. For our purposes, we set it to “Data.” Since the XML file we are importing has no data contained in it, it makes no difference which selection we make. Finally, we need to tell the program were to find the XML file. Select the folder icon next to the “Specify the XML source to import:” textbox and navigate to the XML file location. When we are satisfied that everything is the way we need it, we select “Next.”

The window now displays a review of feature classes and their attributes. The purpose here is to review the rows and ensure there are not conflicts. Ours will not have any because the database is blank. If we had data in this database and attempted to import an XML, this is where we would ensure that the attribution and feature classes do not conflict with existing feature classes or attribution (fig. 1–15). The source name column lists the name of the attribute or feature in the XML file while the target name column refers to items in the database. Since our database is empty, both columns are identical.

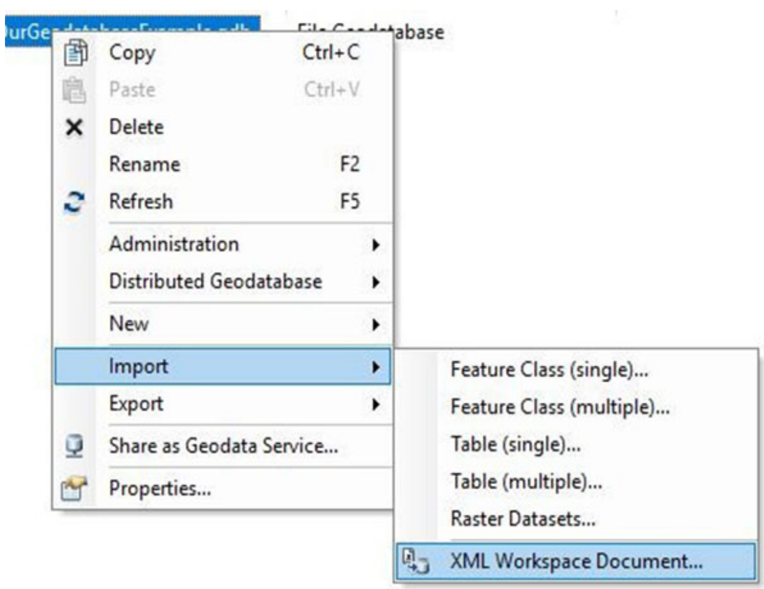

Figure 1–13. Access the XML import wizard.

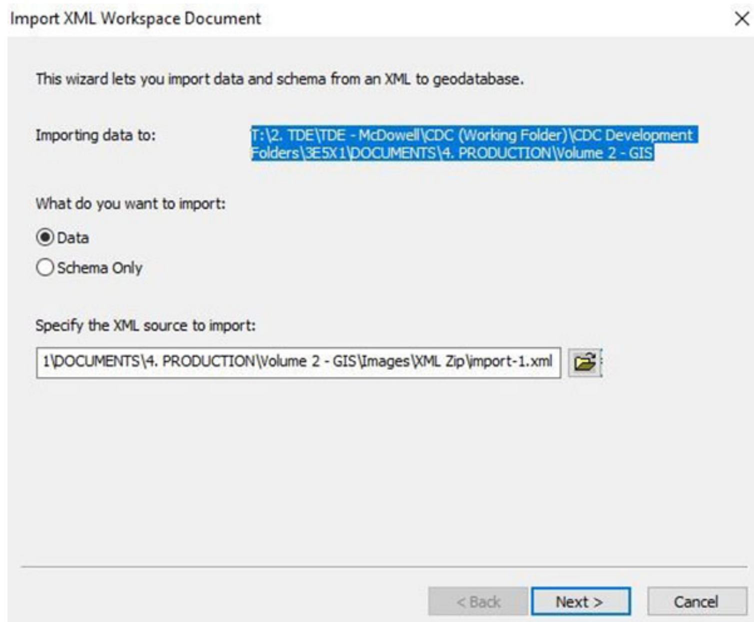

Figure 1–14. Import wizard page 1.

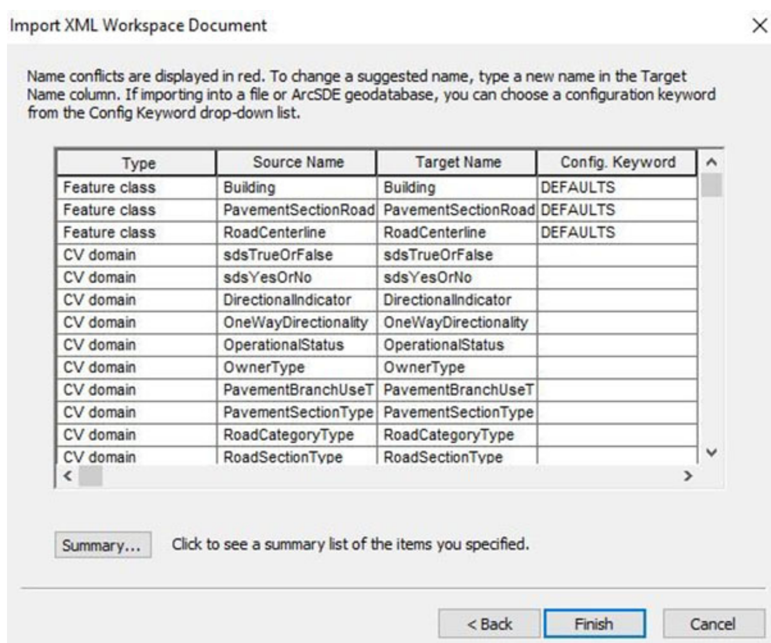

Figure 1–15. Import wizard page 2.

Selecting the "Summary…" button at the bottom left of the window opens the "XML Import Summary." The program lists the import source, where the XML file came from, the type of import, labeled "XML Import," and the feature classes and any associated attribute tables marked "Layers/Tables." This screen is a quick overview of the action to perform (fig. 1–16). Nothing happens until we return to the previous window, by selecting "OK," and selecting the "Finish" button in the lower right-hand part of the window.

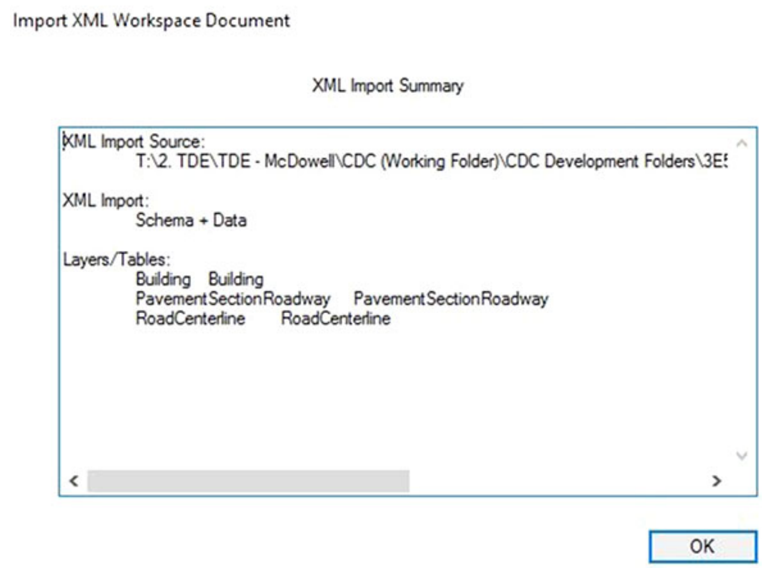

Figure 1–16. XML import summary.

After selecting "Finish," catalog will display a window showing the progress of the import. Once complete, we can enter the geodatabase from the content tab pane in catalog. Before the import, the database was empty, but now it has "places" to put information once we have collected and added it (fig. 1–17). We now have a database and a framework in which we can organize data. The remaining task is to add data to the database.

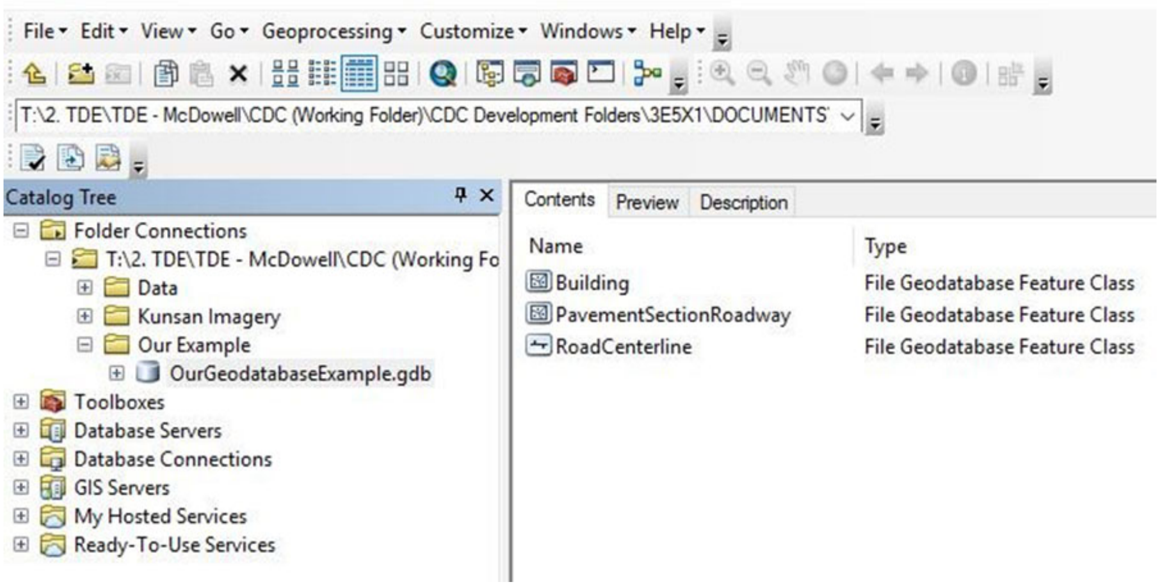

Figure 1–17. Result of the import.

Add data to the data container

Catalog is a good program for creating the containers and shells in which we place data; but a container means little without data. To create and add data into the container it is much more user-friendly to use Map. The Map environment allows us to interact with the data as graphics and tables. So let us begin by opening Map and working through these concepts step-by-step. When we open the program, a window opens labeled "Getting Started (fig. 1–18)." We select "Blank Map" and select "OK." This brings us to our starting point–an empty .mxd file.

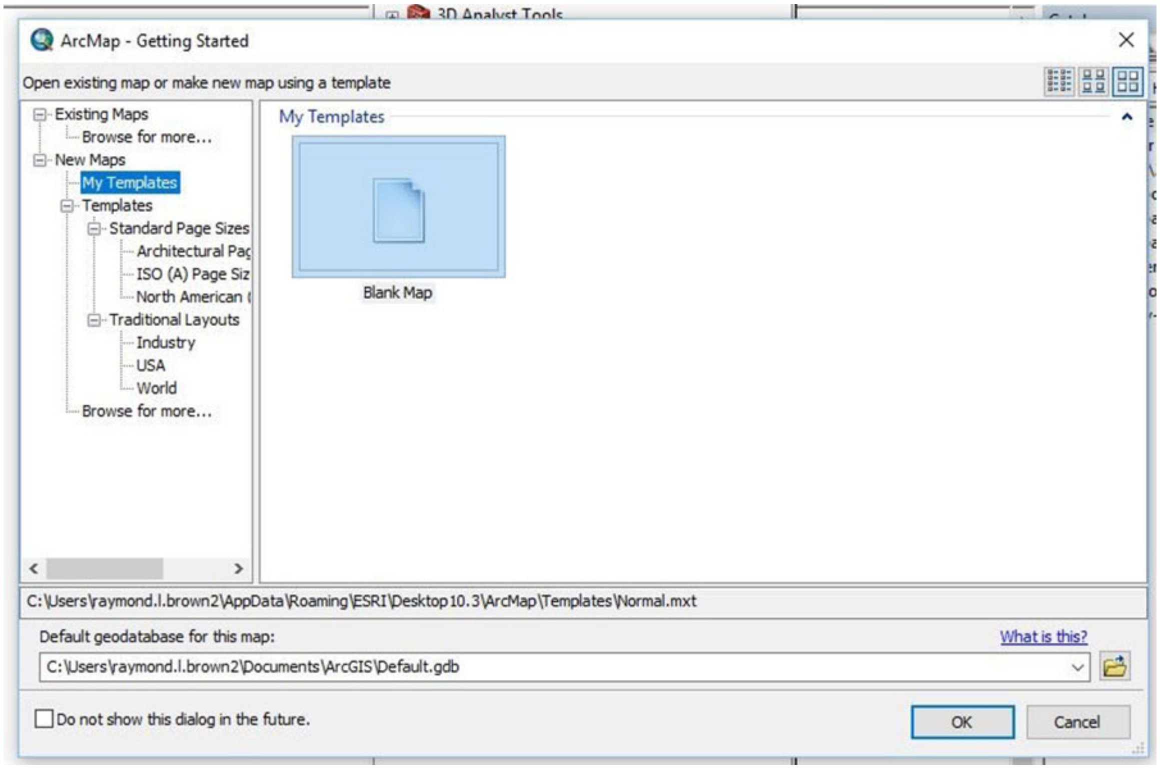

Figure 1–18. Getting started.

Add data function

At the top of the blank map window, below the “Selection” dropdown list, is a black cross over a yellow diamond. This is the “Add Data” function (fig 1–19). The name is a little misleading. What really happens when we select “Add Data” is that the program connects to the database rather than adds the information to the map file. Another way to say this is that the database stores the information while an .mxd map file stores information on how to display the connected data.

Selecting “Add Data” opens a navigation window in the Catalog style asking for data location. Navigate to where we placed our recently created geodatabase and add the feature classes Building, Pavement Section Roadway, Road Centerline, and Water Utility Node to the map.

A warning window will open labeled “Unknown Spatial Reference (fig. 1–20).” This is because the feature classes do not have a coordinate system applied to them. There are a few ways to add a coordinate system to a feature class, let us go over two of them.

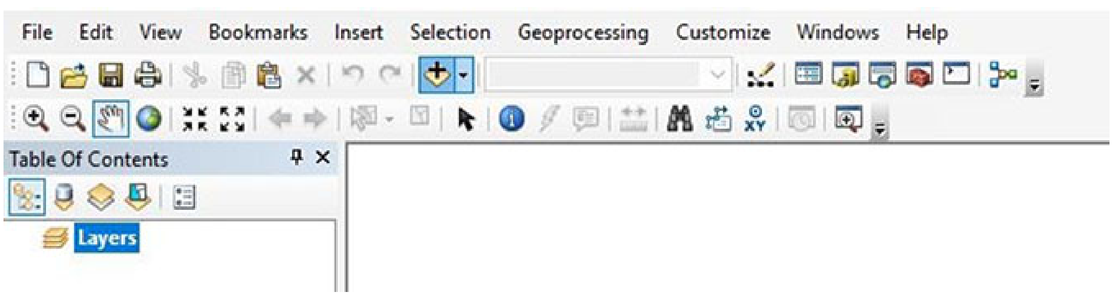

Figure 1–19. Add data function.

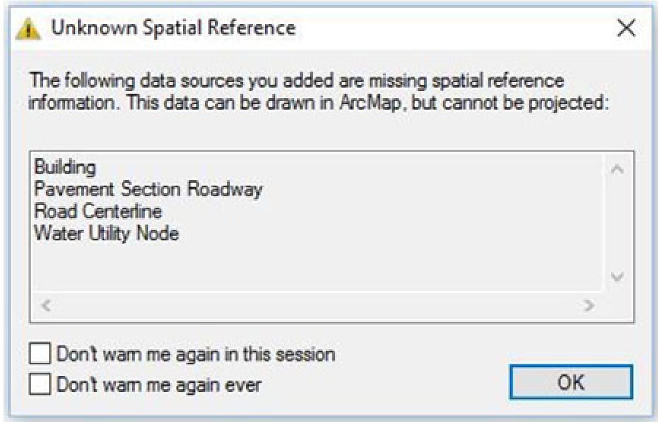

Figure 1–20. Unknown spatial reference.

The first way is to change the coordinate system of the Map file. To change the .mxd file's coordinate system, right-click on the "Layers" group title and select properties at the bottom of the list. Doing so opens the "Data Frame Properties" window (fig. 1–21).

The "Data Frame" refers to the layers listed in a viewport on the Map layout. Generally, a map has only one unless more are required. Another section in this volume covers adding and managing multiple data frames. What is important to understand now is that the change we made to the data frame is applicable only to the Map file. The database is unaltered and has no coordinate system set. Data on this map displays the coordinate values in the database and in the coordinate system the map, or data frame.

The second way to apply a coordinate system is by changing it in Catalog. You do this by navigating to the feature classes in the database through Catalog, right-click on them, select properties, and select the "XY Coordinate System" tab (fig. 1–22). The coordinate system we will use is a projected coordinate system under UTM, WGS84, listed as zone 15N. Since any change here in the database will affect any maps that connect to this database, all Maps displaying this data will use the database coordinate system by default.

Shapefiles

Most of the data that you receive for input into a geodatabase will come in other formats. The most common format is the "shapefile," which is another kind of data container. A shapefile is similar to a feature class in that it is of either a point, line, or polygon data type. The difference is that it exists independent from a database. The name "shapefile" is a bit of a misnomer. A single shapefile is actually four different files. The four files are the .shp, .shx, .dbf, and .prj files. The .shp file carries the geometry data including the coordinates of the features' positions. The .shx file is an index for relating the .shp file to the other files in the shapefile group, specifically the .dbf file. A .dbf file holds the attribute table for the shapefile with an index reading back to the .shx file, which translates again back to the .shp file thereby assigning features in the .dbf table to the .shp geometry. Finally, the .prj file is the coordinate system that the .shp uses.

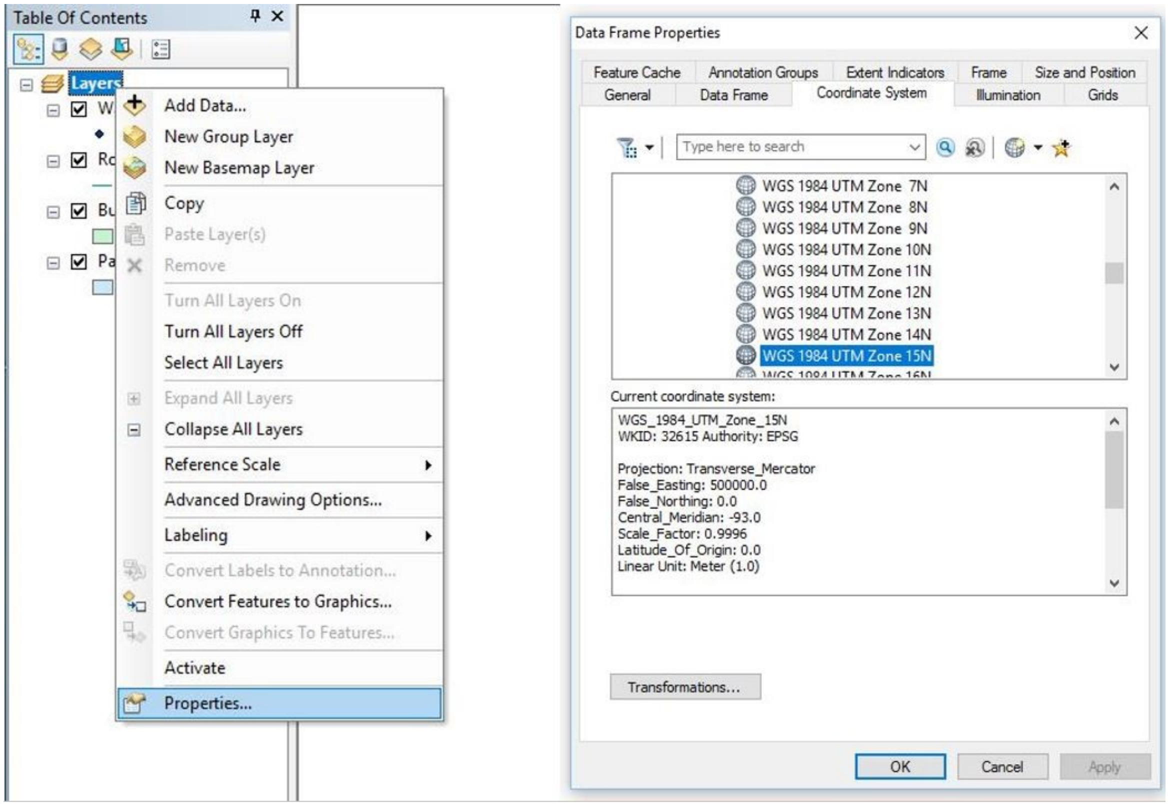

Figure 1–21. Data frame properties.

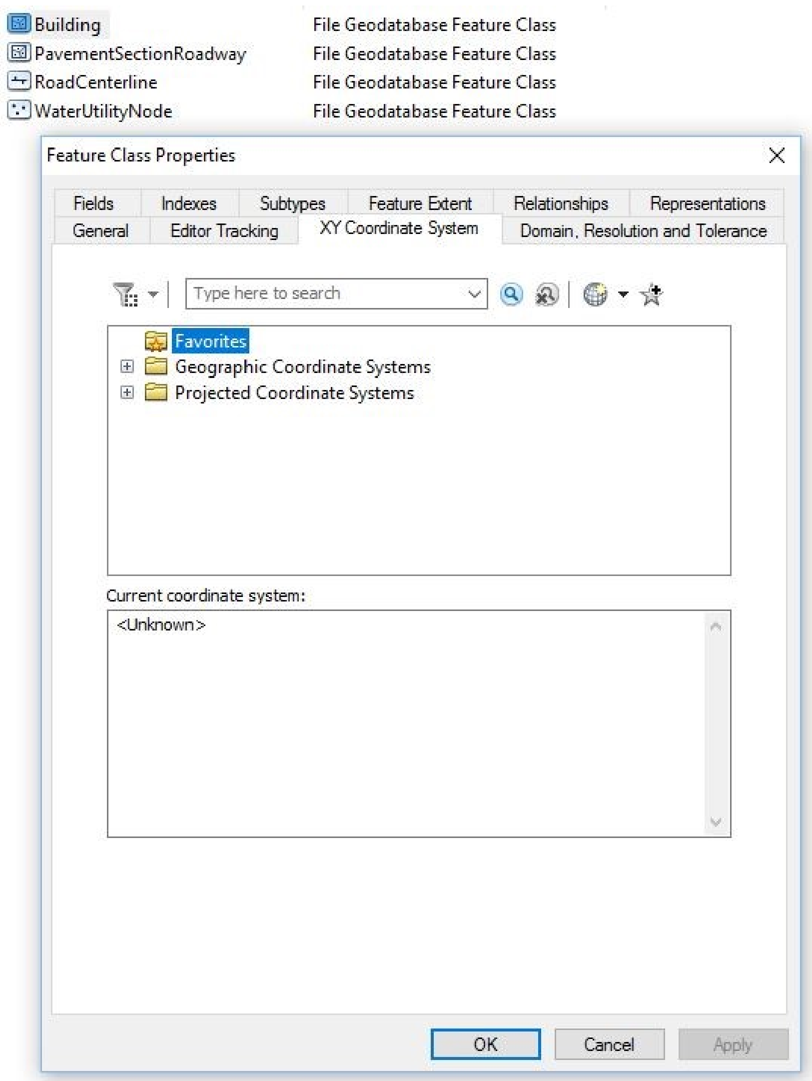

Figure 1–22. Feature class properties.

205. Creating data in a geodatabase

At this point, we have connected to an empty data container. We use the same process to connect to any data container. Our container is empty and a data container is only as useful as the data contained inside. In this lesson, you learn to create data using the Map software. Pay attention to how we manage and organize the data as well.

Editor toolbar

The first thing we need to do a job is the right tools. The Editor Toolbar is a series of tools used to create and modify data. Add it to the map by right clicking on the empty space to the right of the default toolbars above the map space and below the Map program title bar (fig. 1–23). A list of available toolbars appears, select “Editor.” The toolbar appears and can be dragged anywhere on the screen and docked in various places.

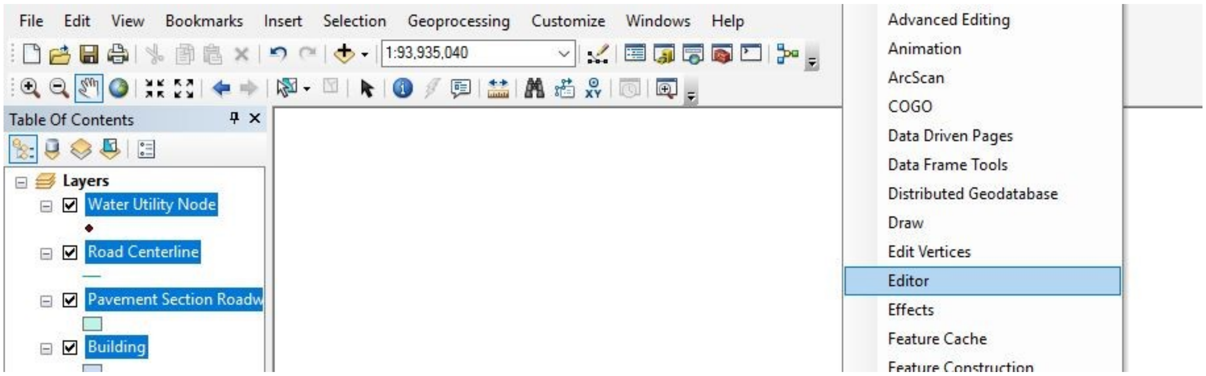

Figure 1–23. Toolbar list.

When we select "Editor" on the toolbar, a dropdown appears that has "Start Editing" listed at the top (fig. 1–24). This creates a "buffer" between the database and the map. We start the editing process and make any changes, but the change does not become part of the database until we save those changes by selecting "Save Edits." This "buffer" protects the database from unintentional changes and allows us to be certain of the changes before we commit them to the data container.

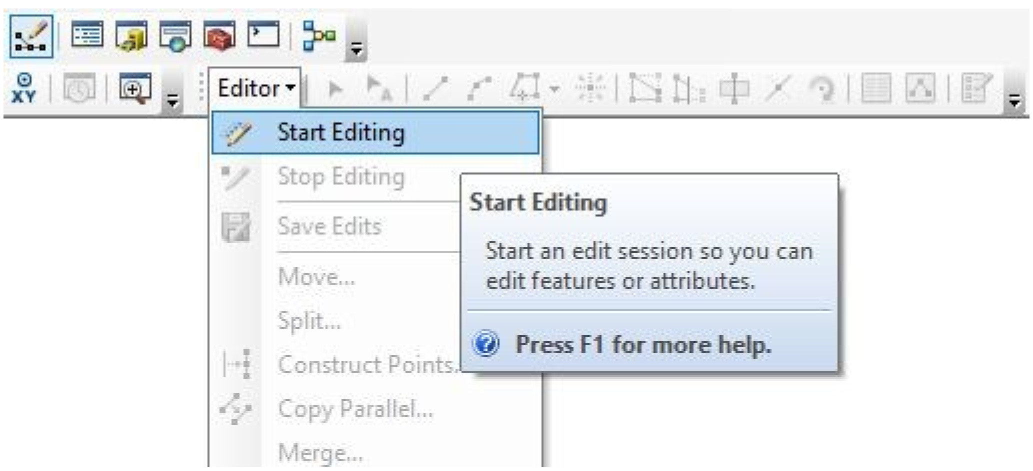

Figure 1–24. Start editing.

After starting an editing session, the "Create Features" panel appears on the left side of the screen. From this panel, we select which layers to add to or modify. Below this is the "Construction Tools" box, which, after selecting a layer, lists the available creation tools for that layer type either point, line, or polygon (fig. 1–25). Our example will be a polygon type on the "Building" layer. It is important to mention that we are not drawing in the same sense as computer-aided drafting (CAD). We are creating geometric entities, each unique and defined by many different attributes, as we will soon see.

Our cursor has changed from a pointer to a small black cross. This indicates that we have a layer and construction tool selected for editing. After selecting a starting point for our polygon, we can right click on the screen to see the list of editor functions (fig. 1–26). Each one has a different purpose and functionality; they are the direction, deflection, length, absolute XY, delta XY, and direction/length.

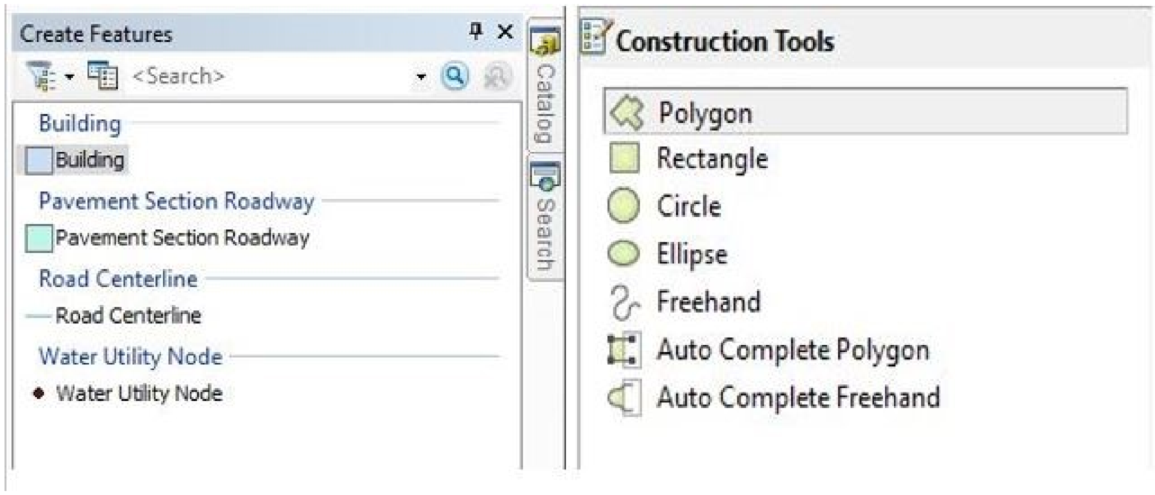

Figure 1–25. Feature layers and tools.

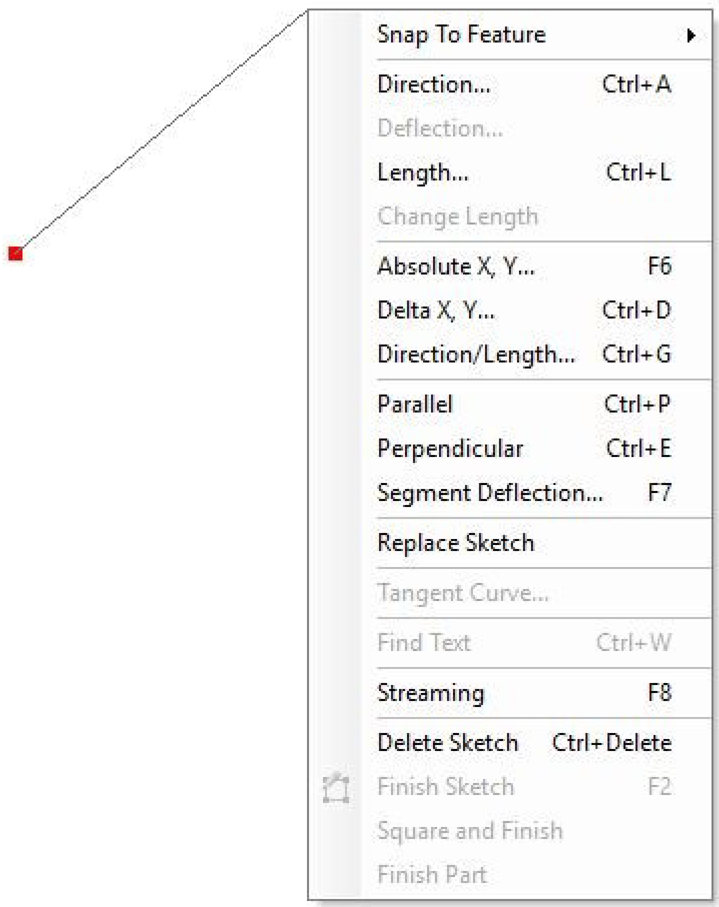

Figure 1–26. Tool functions.

Direction

Direction locks the line to one direction by compass degrees. When selected, a dialog box appears asking for the angle. Once provided, no matter where we place the cursor on the screen, the line will only grow shorter or long while remaining on the same angle.

Deflection

Deflection is another angular function. This time it uses the relative angle compared to the angle of the previous length of line drawn. Think of manual surveying and measuring angles. This is the same concept but on paper instead of with survey equipment. The same dialog box requesting an angle will appear but instead of an absolute compass angle, it takes the value as a turned angle.

Length

Length performs in the same way but opposite the direction function. After providing a value for the length of line, the program allows the user to rotate the line around its starting point without changing its length.

Absolute XY

Absolute XY is most useful when plotting points. The tool allows you to designate which units to use. The small down arrow icon to the right of the "X:" coordinate box in the dialogue allows you to use anything from meters or northings and eastings to latitude and longitude (fig. 1–27). It places a vertex, or point, at the designated coordinates.

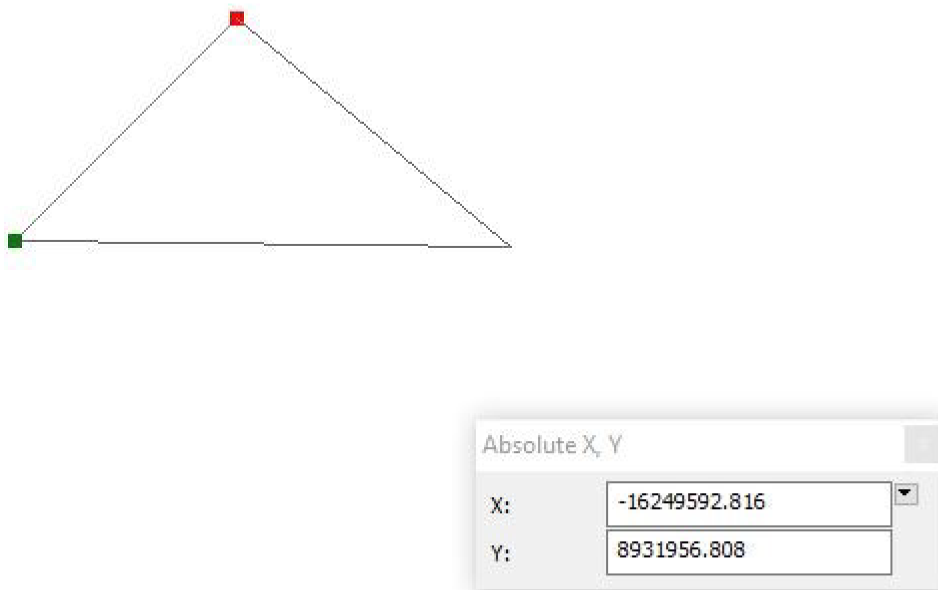

Figure 1–27. Absolute XY dialog box.

Delta XY

Delta XY is similar to the absolute XY function. The difference is that while absolute XY places a point at the specified coordinates the delta XY places a point a specified number of units away from the previous point placed. If we place a point at 1000, 1000, and select a delta XY of 10, 10, then the second point is at the coordinates 1010, 1010.

Direction/length

Direction/length asks for an angle, direction, and a distance (or length), for the position of the next point. This is the most used tool because it allows faster feature creation, especially for polygons. After selecting an initial starting point, an angle of 45 degrees and a length of 200 draws a side of the feature that direction and length and adds a vertex (fig. 1–28).

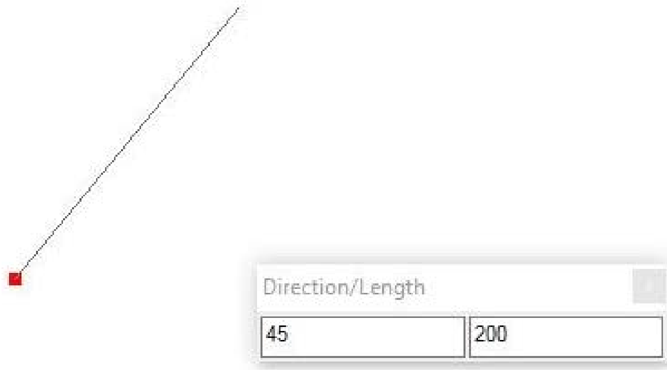

Figure 1–28. Direction and length.

Attribution

The features we drew are only geometric in nature and have no defining characteristics beyond the physical. We give features further descriptive information by modifying their attributes. Attributes are stored in a table format by layer. Each layer table has features listed by the row, and each column is a different attribute. Attribute tables are accessed by right clicking on the desired layer in the table of contents and selecting "Open Attribute Table." The table will display in a separate window (fig. 1–29).

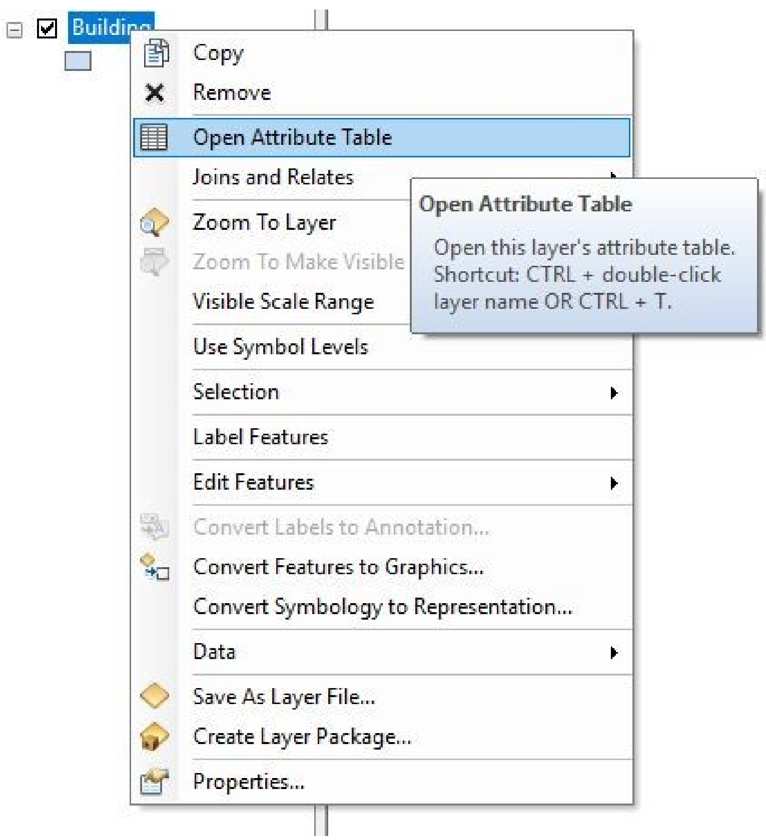

Figure 1–29. Open attribute table.

When we select an item in the attribute table by highlighting a row, the corresponding graphic in the map highlights as well (fig. 1–30). The important thing to understand about attributes is that they further define features in useful ways. For example, the figure shows that our feature can have an address number. This number makes it unique. Later, we will see these attributes used to visualize relationships between features. For now, understand that these attributes are important to describing each feature.

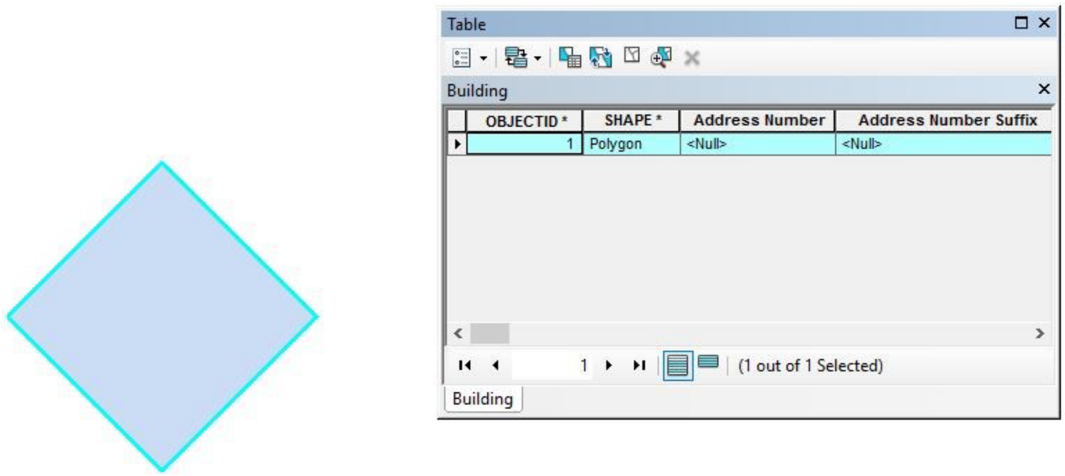

Figure 1–30. Attribute table.

Edit attributes

To edit an attribute, an edit session must be open for the layer in which the feature resides. Our polygon is part of the building layer; we want to add the address number “527” to its attribute table. After we open an editing session and open the attribute table, we double click in the desired cell of the table and add our information (fig. 1–31). We can save the modification we made by selecting “Save Edits.”

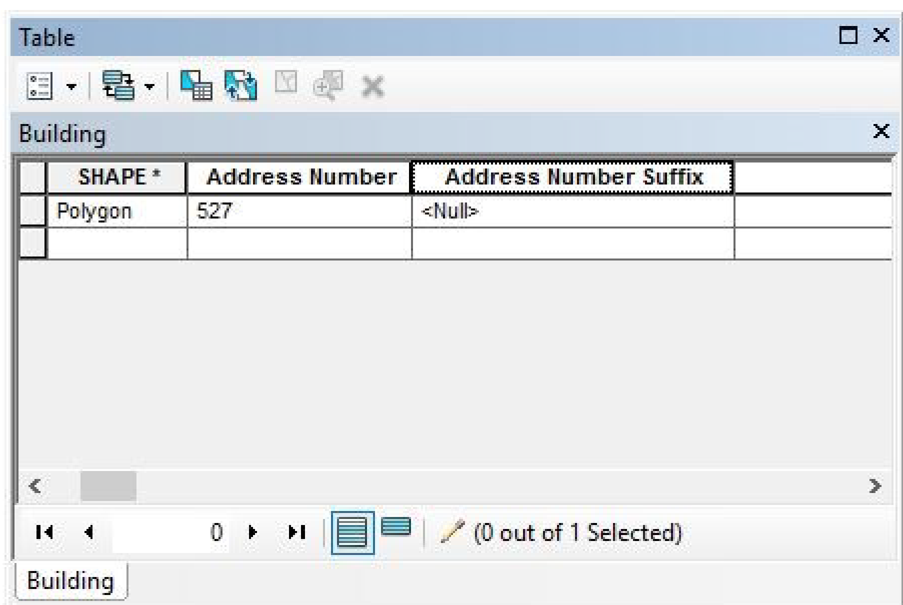

Figure 1–31. Table change.

Add a field

Sometimes it may be necessary to add fields or columns to the table. This is usually because the schema you selected is not sufficient to store data unique to your mission or installation. Therefore, it may be necessary to expand the attribute table to make the data more useful for your needs. To create a new field, you must <u>NOT</u> be in an edit session. The first icon on the left, at the top of the attribute table window, is a dropdown that has a function called “Add Field…” This will open another window that asks for the name of the new field and the data type that will be required to fill it. The data type corresponds to integers, floats (numbers with decimals), text, date, and so forth. It will also have a section called “Field Properties” that has “Alias,” “Allow NULL Values,” and “Default Value” settings (fig. 1–32).

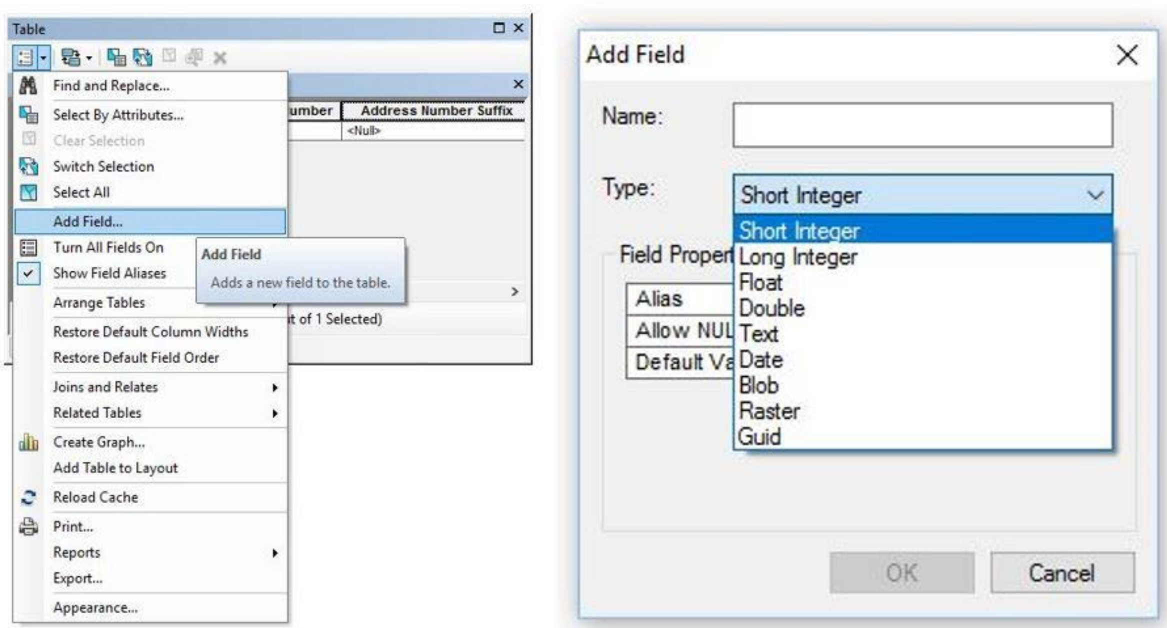

Figure 1–32. Add field.

The “Alias” setting allows you to create a sort of nickname for the field that is easier to remember when searching for it. “Allow NULL Values” is either Yes or No and designates whether this field can be left blank or NULL. A “Default Value” is one that is added to the field for each new feature. The best way to show the usefulness of this is by explaining enumeration.

Enumeration

Enumeration is the way we differentiate between types of features within a feature class or layer. The Water Utility Node feature class in our data is a nice example. Water nodes are points. These points can be fire hydrants, valves, drains, and so forth. How do we know which is which? An attribute in the attribute table describes whether the point in question is a valve or fire hydrant. This way of differentiating features within the same feature class is enumeration. Other layers work the same way. The Pavement Section Roadway layer enumerates by type of pavement the section or feature, which is usually either asphalt or concrete.

Self-Test Questions

After you complete these questions, you may check your answers at the end of the unit.

203. Utilize browser

1. A GIS framework, or schema, allows what two things to happen?

2. What is a data model and what is the most common difference between data models?

3. When generating an XML, what are the three data models we can choose?

204. Creating a geodatabase

1. What is the last step in creating a geodatabase?

2. When importing an XML into a geodatabase what kind of information does the “Summary” display?

3. What function do we use to connect to data sources to display in a Map file (.mxd)?

4. Which properties window do we access to change a Map file’s coordinate system and how do we get to it?

5. What are the two ways to set up a coordinate system and what is the difference?

205. Creating data in a geodatabase

1. Describe how the “buffer” created by the editing process protects the database.

2. What are the six editor functions discussed in this lesson?

3. Give a brief description of enumeration.

Answers to Self-Test Questions

201

1. Enterprises—Stored on a network and can have multiple users simultaneously; Files—Stored on a PC hard drive and can have multiple editors but not simultaneously; Personal—Stored as Microsoft Access databases and are Windows dependent.
2. (1) Points.
   (2) Lines.
   (3) Polygons.
3. The ratio of an image size to the number of grids per unit.
4. Name, what they are made of, date of construction, etc.

202

1. (1) CIP.
   (2) Land use.
   (3) Airfield operations.
   (4) Transportation.
   (5) Constraints.
   (6) Future development plan.
   (7) Utilities.
   (8) Communications.
   (9) Energy.
2. Installation layout, geographically separate unit, regional location, vicinity location, and aerial photography.

203

1. Our GIS to match those of others and to ease the sharing of data between our database and others.
2. A style of organization for a database that is subject to change and the most common difference is the naming conventions.
3. (1) Physical.
   (2) Logical.
   (3) Metadata.

204

1. Add the schema to the geodatabase.
2. Were XML came from, type of import, feature classes, and associated attribute tables. It is an overview of our action.
3. Add Data.
4. Data Frame Properties.
5. The Map file change only applies to that .mxd, while a change in Catalog applies to all .mxds connected to the data.

205

1. Prevents unintentional changes and allows us to be certain of any changes before they are applied to the database.
2. (1) Direction.
   (2) Deflection.
   (3) Length.
   (4) Absolute XY.
   (5) Delta XY.
   (6) Direction/length.
3. A way of differentiating features within the same feature class.

Complete the unit review exercises before going to the next unit.

Unit Review Exercises

Note to Student: Consider all choices carefully, select the *best* answer to each question, and *circle* the corresponding letter. When you have completed all unit review exercises, transfer your answers to the Field-Scoring Answer Sheet.

Do not return your answer sheet to the Air Force Career Development Academy (AFCDA).

1. (201) Which type of geodatabase can have multiple editors but only if they work on different parts of the data?
   a. File geodatabase.
   b. Personal geodatabase.
   c. Microsoft geodatabase.
   d. Enterprise geodatabase.

2. (201) What map data comprises facilities, physical features, and utilities within the installation boundary?
   a. Mission data set (MDS).
   b. Mission data layer (MDL).
   c. Common installation picture (CIP).
   d. Installation development plan (IDP).

3. (201) What are the three type groups of vector data?
   a. Points, lines, and imagery.
   b. Points, lines, and polygons.
   c. Points, polygons, and imagery.
   d. Points, lines, polygons, and imagery.

4. (201) What do we call the ratio of the size of an image to the number of grids per unit?
   a. Resolution.
   b. Image size.
   c. Dots per inch (DPI).
   d. Pixels per meter.

5. (201) Which tables give data unique characteristics?
   a. Data.
   b. Excel.
   c. Raster.
   d. Attribute.

6. (202) Which mission data set (MDS) represents the planning districts of the installation?
   a. Land use.
   b. Constraints.
   c. Future development plan.
   d. Common installation picture.

7. (203) What are the three types of models we can create when generating a schema?
   a. Physical, logical, or rational.
   b. Logical, rational, or metadata.
   c. Physical, logical, or metadata.
   d. Physical, rational, or metadata.

8. (203) Three files generate when creating a schema; which file contains instructions on importing the schema?
   a. .tbx file.
   b. .sch file.
   c. .xml file.
   d. readme.doc.

9. (204) In the geodatabase context menu, under “Import,” what tool do we use to import a schema?
   a. Raster datasets.
   b. Table (multiple).
   c. Feature class (multiple).
   d. XML workspace document.

10. (204) The import wizard for extensible markup language (XML) files has three windows; which window lists name conflicts?
    a. Page 1.
    b. Page 2.
    c. Page 3.
    d. Summary page.

11. (204) According to figure 1–21, what tab, under the Data Frame Properties, allows us to choose a coordinate system for the map?
    a. Grids.
    b. Data frame.
    c. Size and position.
    d. Coordinate system.

12. (204) Which of the four files that make up a shapefile contains the coordinate system?
    a. .prj.
    b. .shx.
    c. .shp.
    d. .dbf.

13. (205) Which editor function allows us to specify an angle and a length to create a line or side of a polygon?
    a. Length.
    b. Direction.
    c. Deflection.
    d. Direction/length.

14. (205) What do we call a way of differentiating objects within the same feature class?
    a. Emulation.
    b. Explanation.
    c. Type coding.
    d. Enumeration.

Please read the unit menu for unit 2 and continue ➔

Unit 2. Exporting and Converting Data

DATA COMES IN MANY shapes and sizes. Understanding how to change data from one format to another is a key skill for every engineering professional. Whether we are giving data away to another federal organization or receiving data from the local government, we need to customize the data into a usable format. Our entire profession revolves around information, and the way it improves decision making. Providing leaders with accessible information means the difference between the right decision and a bad decision.

2–1. Conversion

This section covers ways to take one format of data and turn it into another. This includes the properties of that data, such as their coordinate systems. Try to think in terms of adding survey data to a database. Survey data is the most common data you convert, especially because of the number of software platforms in which you need to be familiar.

206. Convert to shapefiles and rasters

There are times when you have information in one format, and you need to make it portable. Not only do you need it to be portable, but you also need to make it compatible with its final destination. CAD files in the .dwg format are a great source of geometry. Unfortunately, drawing objects cannot store attribute data, which makes analysis impossible. This lesson covers how you convert a .dwg format file into a .shp format file allowing us to add the information to a geodatabase. Once in the geodatabase, you can add attribution, which enables analysis of the features.

Drawing file to shapefile

Anytime you manipulate files, you use catalog. The terms “Convert” and “Export,” within the context of GIS data, are interchangeable. The conversion occurs by “importing” the file into catalog and “exporting” it as another file type. Catalog is, among other things, a machine for converting file types from one kind to another. You use different functions, or tools, to convert files. Here, you use the “Feature Class to Shapefile” export tool.

To access the tool, you right click on the .dwg file to open the context menu (fig. 2–1). Then, you hover over “Export” to reveal two options, “To Geodatabase,” and “To Shapefile.” You select “To Shapefile” to open the conversion, or export, tool.

Selecting “To Shapefile (multiple),” opens the tool’s window. There are two areas here—”Input Features,” and “Output Folder” (fig. 2–2).” The input area has two elements. The first is a single-line text box that takes a file path to the file you want to convert. The second area lists the file path to the file and the output type. Another way to read this is which elements in the file you want to make into a shapefile. For example, if the file had only lines and no points or polygons, then you could remove everything except the “Polyline” file path line. The “Output” text box is where you place the file path for the folder you want the shapefiles to go into. As discussed earlier, the reason you specify a folder is because shapefiles have multiple files.

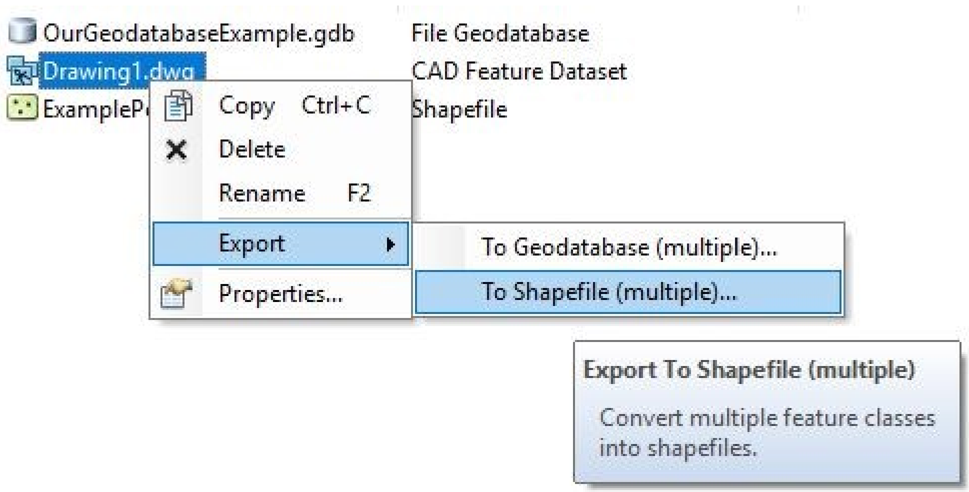

Figure 2–1. Conversion/export tool location.

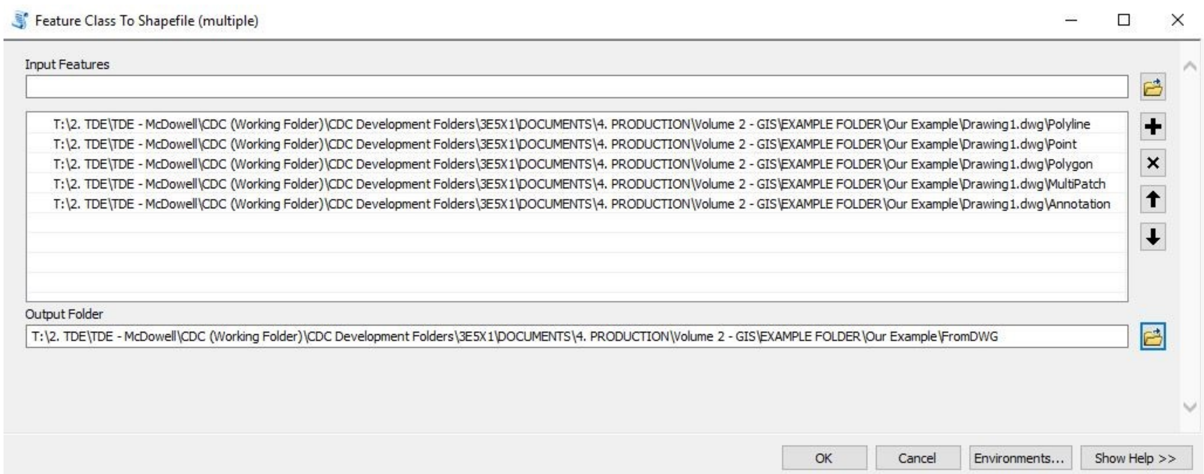

Figure 2–2. Feature class to shapefile window.

Once you have added the file and file types you want to export, select "OK" at the bottom of the window. This completes the process and places the files in the specified directory or folder (fig. 2–3). In catalog, each of the shapefiles lists as a single file. When you navigate to the same folder, using File Explorer, you find that each of these individual files has four additional files. This is important to understand because any modification you make to a single shapefile in catalog extends to the group of files. For example, if you move a shapefile in catalog to another folder, catalog moves its accompanying files as a group.

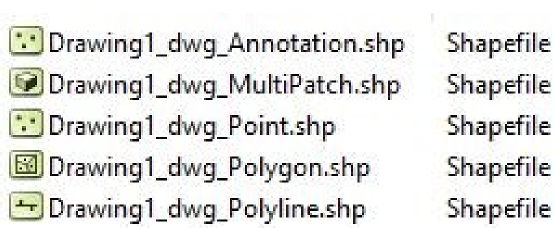

Figure 2–3. Output shapefiles.

Raster file types

Rasters come in many different formats .jpg, .gif, .png, and so forth. Every image file type is a raster. Converting between these files is as easy as changing the suffix of the file (fig. 2–4). For example, to change a .tiff file type to a .sid file type, simply rename the file adding the new suffix. A few things make certain raster types unique. The best example is the multiresolution seamless image (MrSID) database, which has the suffix .sid. You will see the MrSID file type a lot because of its two key characteristics. First, it is georeferenced; meaning it is orthorectified and has coordinates. Second, it uses pyramids to improve rendering speed.

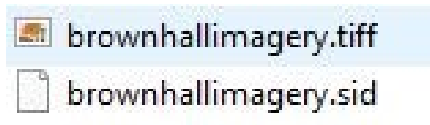

Figure 2–4. Renamed image file.

Pyramids are an important concept when talking about rasters. They are a way for very large rasters to render, or display, on the screen efficiently. Rasters that have no pyramids will render every pixel no matter the scale to which you are zoomed. Pyramids create “samples” of the raster by grouping grids together at larger scales (fig. 2–5).

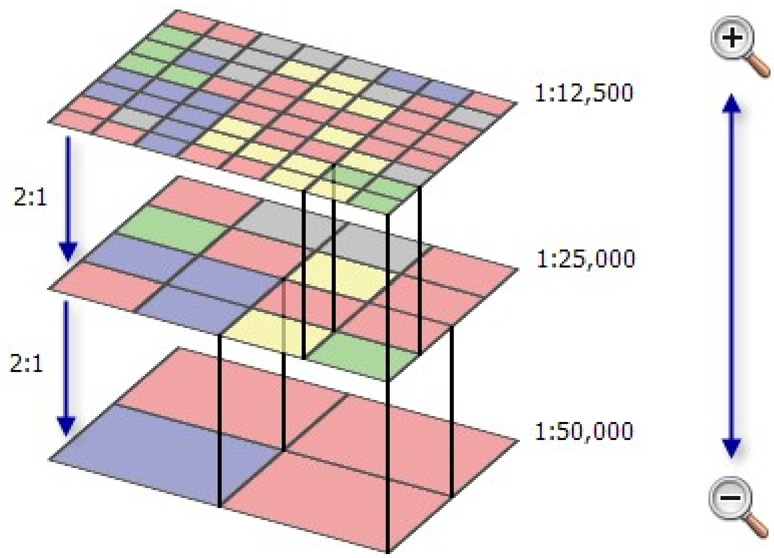

Figure 2–5. Raster pyramids.

The idea behind pyramids is that the raster renders at different resolutions depending upon the zoom factor or scale. This is important to understand because when you create MrSID files by renaming and add it to a Map file; the program prompts you to create pyramids for it. Depending upon the size of the raster file, you can make a judgement call on whether this is necessary.

207. Import into and extract from geodatabases

One way to understand geodatabases is as a group of tables or spreadsheets. At the most basic level, the geometry of points, lines, and polygons, as well as the attribution, that defines them is stored in the form of tables. Because of the table-like form, you can import any information into a geodatabase as long as you can make it conform to the GIS table format. A good example is survey data. The .job file from a survey controller is not compatible with a GIS. However, you can export the survey data within the controller software as a .csv file type.

Importing spreadsheets

Catalog has the functionality to convert any table type, such as a .csv, into different file formats. When you open catalog and select our spreadsheet, you can see the data it contains by selecting the "Preview" tab (fig. 2–6).

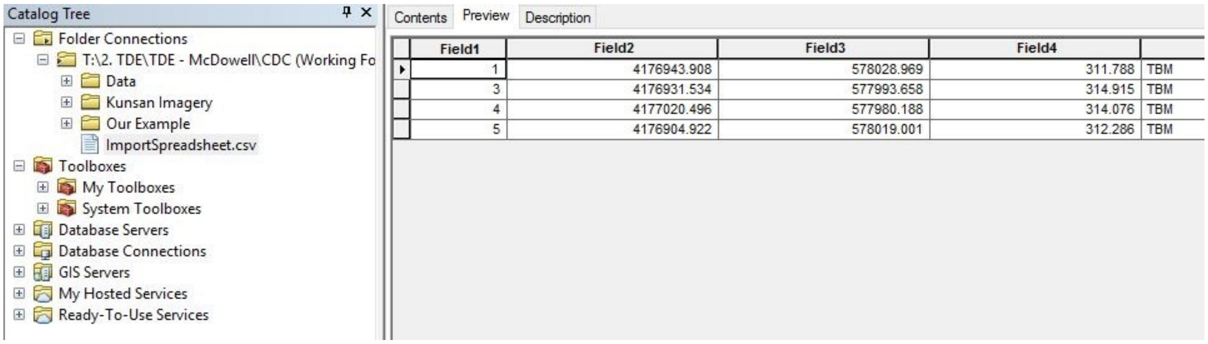

Figure 2–6. Preview a spreadsheet in catalog.

Our spreadsheet is the product of a survey. You can see that the table contains point data. "Field1" is the point number, "Field2" contains the northing, "Field3" contains the easting, "Field4" contains the elevation, and "Field5" contains the point code. This is the result of a survey that was conducted in which the point data was exported to a .csv file in the point format of northing, easting, elevation, and, then, code. To convert the spreadsheet, you right click on it in the "Catalog Tree" to access the context menu. In the context menu, you hover over "Create Feature Class" and select "From XY Table." Doing so opens the "Create Feature Class from XY Table" window (fig. 2–7). Notice from the figure that the "X Field," "Y Field," and "Z Field" were changed to correspond with their respective fields in the spreadsheet—y-field (northing), x-field (easting), and z-field (elevation).

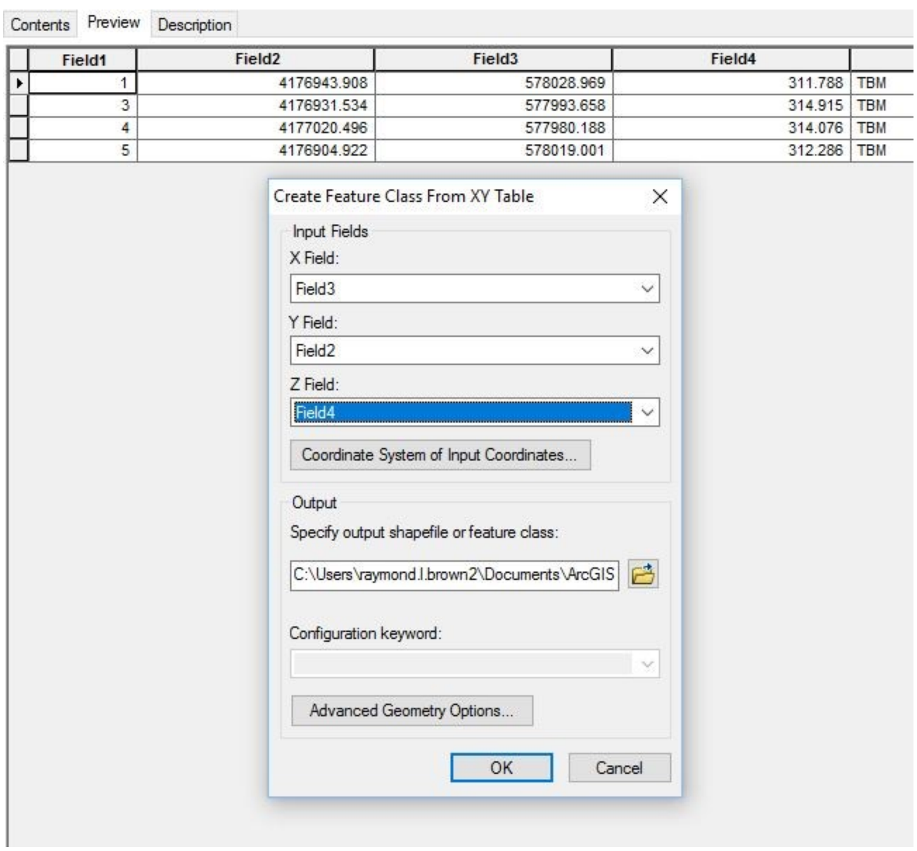

Figure 2–7. Feature Class from XY Table window.

Before you can proceed with the conversion, you need to specify the file path for the output file. The output can be in two general formats—feature class or shapefile (fig. 2–8). If you are making a feature class, the file path must be within a geodatabase; otherwise, the program will return an error. You can store shapefiles wherever you want. Once satisfied with our settings, you select the “OK” button to execute the conversion, which may take a while depending on the amount of information in the spreadsheet.

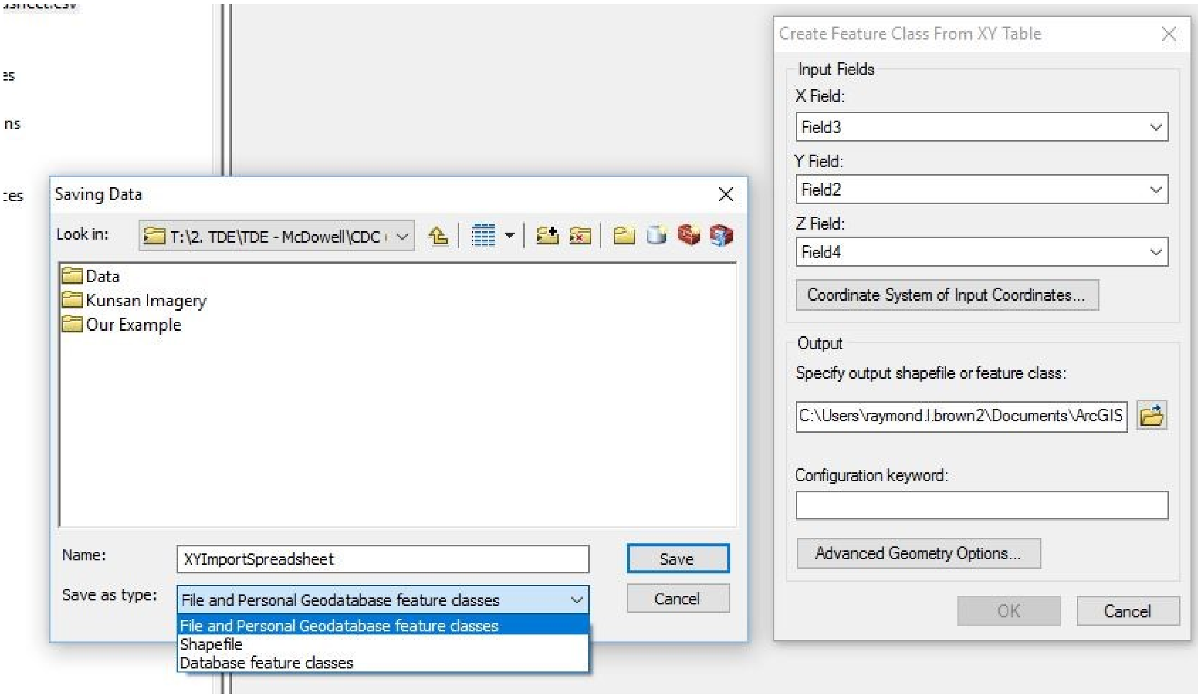

Figure 2–8. File path and file types.

In our example, the spreadsheet was converted to a feature class and added to a personal geodatabase. It now has its own place in the geodatabase and uses the geometry from the table to display the points in the Map application. You can now configure its attribution to suit our needs.

Extracting from a database

Moving data from one place to another is an important part of managing a GIS. Let’s say that you have a Personal Geodatabase, and you need to move the data you have collected into a File or Enterprise Geodatabase. The best way to do this is to use the “Extract Data Wizard.” The Extract Data Wizard is a tool used to transform the data or the schema of a database into a portable .xml format file. Once we generate this file, we can append or add it to any other database. The Extract Data Wizard is located on the “Distributed Geodatabase” toolbar, which is not part of the default toolbars when you open Map. You will add it by way of the context menu (fig. 2–9).

Open the wizard by selecting the icon on the far-right end of the Distributed Geodatabase toolbar. Selecting it opens the “Extract Data Wizard” window (fig. 2–10). The top of the window has a short description of the purpose of the tool followed by the file path of the geodatabase from which you are extracting. Then, the program asks whether you are extracting the data, including the schema, or just the schema by itself. You, then, specify which format—either geodatabase or .xml—for the extracted data and the file path in which to place it. Once you have these minimum options set, you can select next to move on to the second window in the wizard.

The next window shows a list of data in the geodatabase from which you are extracting. When you want all of the data exported, you check “The full extent of the data” checkbox. Otherwise, you check or uncheck data in the list depending on what you need. Keep in mind that some of the data is dependent on other data. This means that unchecking one item may automatically uncheck its dependencies. Always double check to make sure the items you are exporting remain checked.

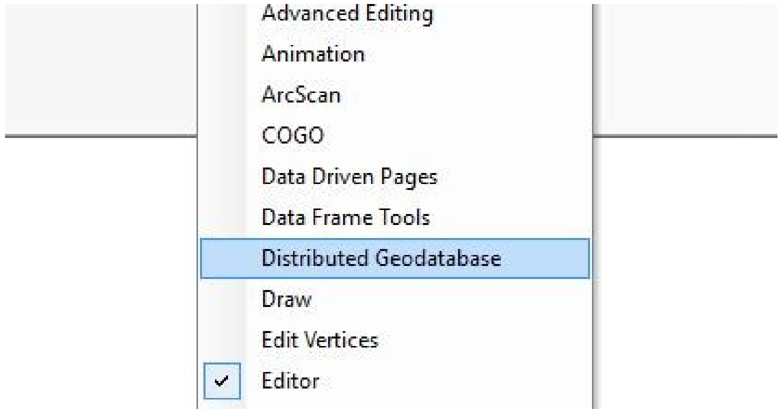

Figure 2–9. Distributed database location.

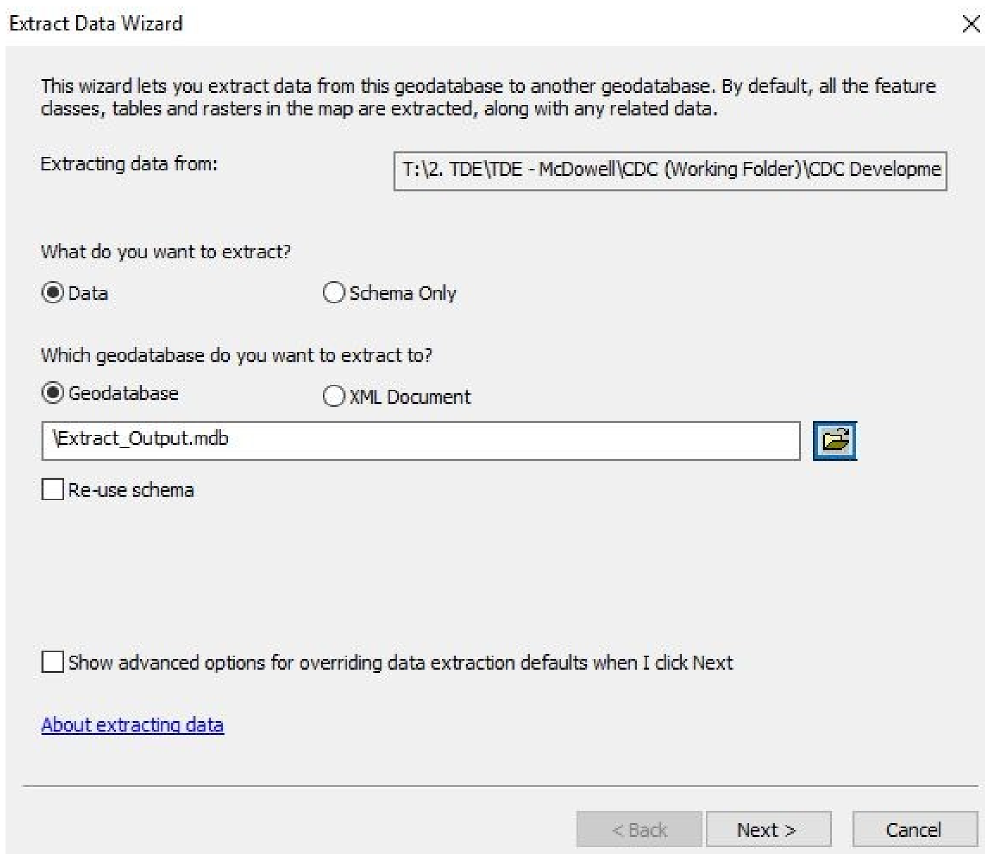

Figure 2–10. Data extract wizard.

208. Convert coordinate systems

We use data from many different sources. These sources are usually civilian and, depending on the organization, the details of the data differ. One of the most common differences is the coordinate system. In the Air Force, we use World Geodetic System 1984 (WGS84), but state-level or private companies usually use one of the state plane systems. Several private surveyors use North American Datum 1983 (NAD83). This all means that if we use their data, we need to understand how to convert it to the coordinate system we use. Let us look at how to do this.

Transformations

The first thing we need to cover is transformations. Geographic coordinate system makes up the base of each projected coordinate system. Different geoids, spheroids, and ellipsoids make up each geographic coordinate system. This means that each coordinate system varies dramatically from the others. Transformations are mathematic models that translate between each of these coordinate systems. Our Map software has these transformations built into the software, but we need to set them up and apply them to the map.

To apply a transformation to a map, right-click on the data frame to access the context menu. From here, select properties at the bottom of the list (fig. 2–11). This opens the “Data Frame Properties” window.

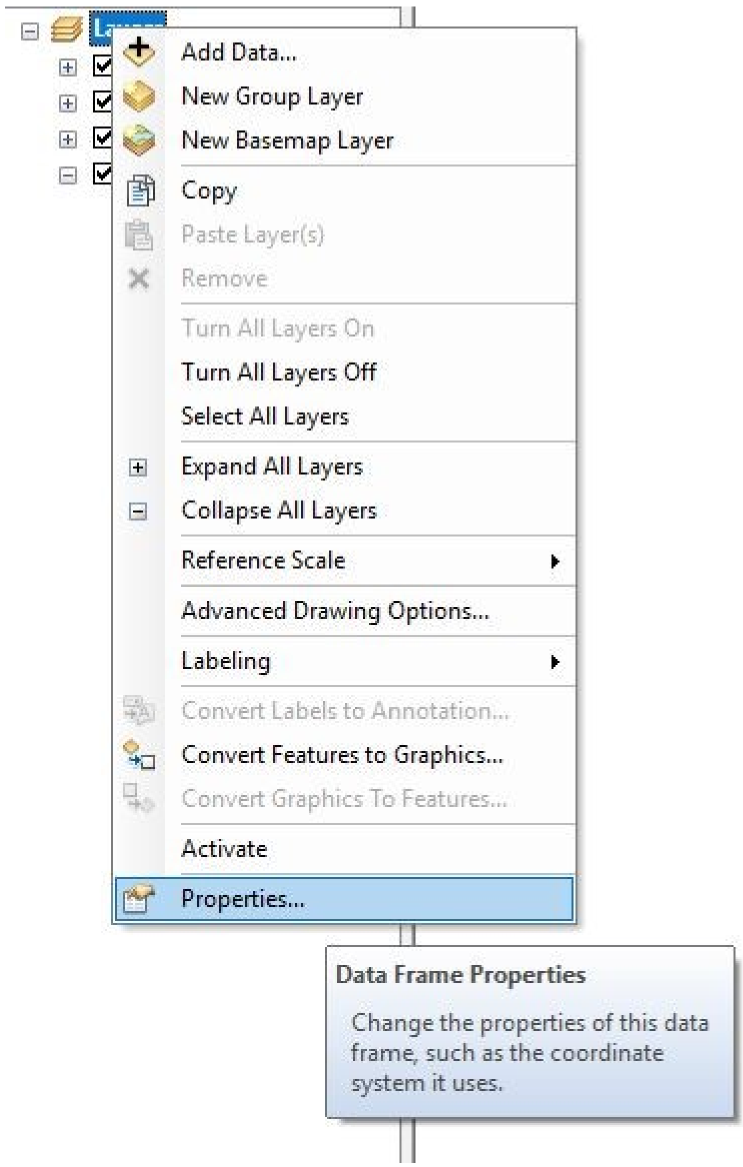

Figure 2–11. Data frame properties location.

Then, you navigate to the “Coordinate System” tab. If you remember, you apply coordinate systems to the map document from this window. In the bottom, right area of the window is the “Transformations…” button (fig. 2–12). Selecting this button opens the “Geographic Coordinate System Transformations” window.

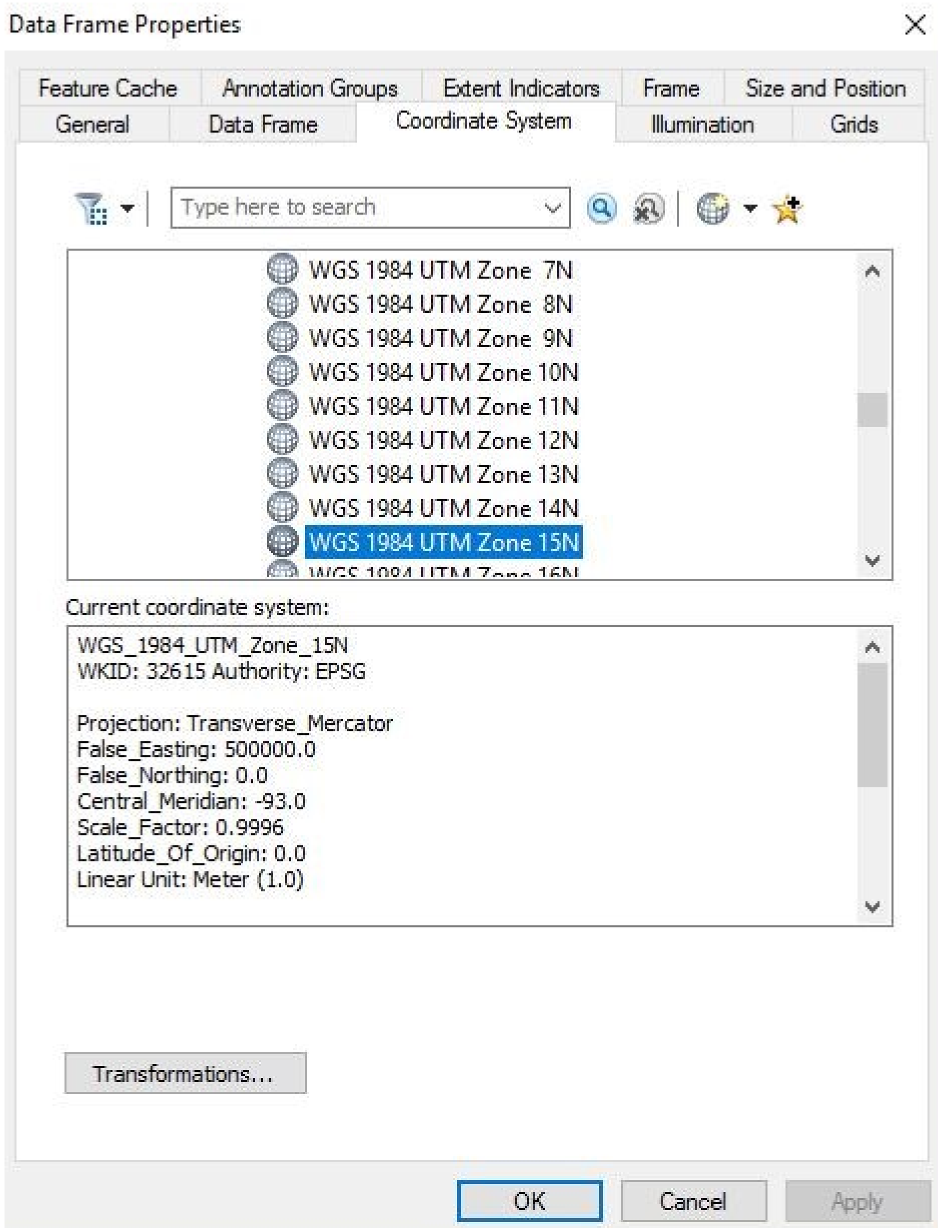

Figure 2–12. Data frame properties – coordinate system tab.

The “Geographic Coordinate System Transformations” window has three primary parts, the “Convert from,” “Into,” and “Using…,” areas.

The “Convert from” area lists each coordinate system that exists in the map document. The first thing to do is select the coordinate system from which you are converting. For example, you want to transform a NAD83 data item into WGS84. The list will show NAD83 for the data you added to be transformed and WGS84 for the data settings of the map. You select NAD83 as the item you transform from because you want it to become WGS84.

The “Into” drop-down list is initially populated with only those coordinates that exist within the Map. To add more, you select the “Add…” button and navigate to the desired system. It is important to double check the system referred to in this box.

Finally, the “Using” drop-down lists the transformations available for the combination of systems you selected. Options rank from most suitable at the top to least suitable at the bottom. In the example, you are converting WGS84 into NAD83 using the ITRF00 (fig. 2–13). Once you are satisfied, select OK. Ensure that when you close the Data Frame Properties window; you select either “Apply” or “OK.” If you do not, then the transformation options are not saved as part of the Map document.

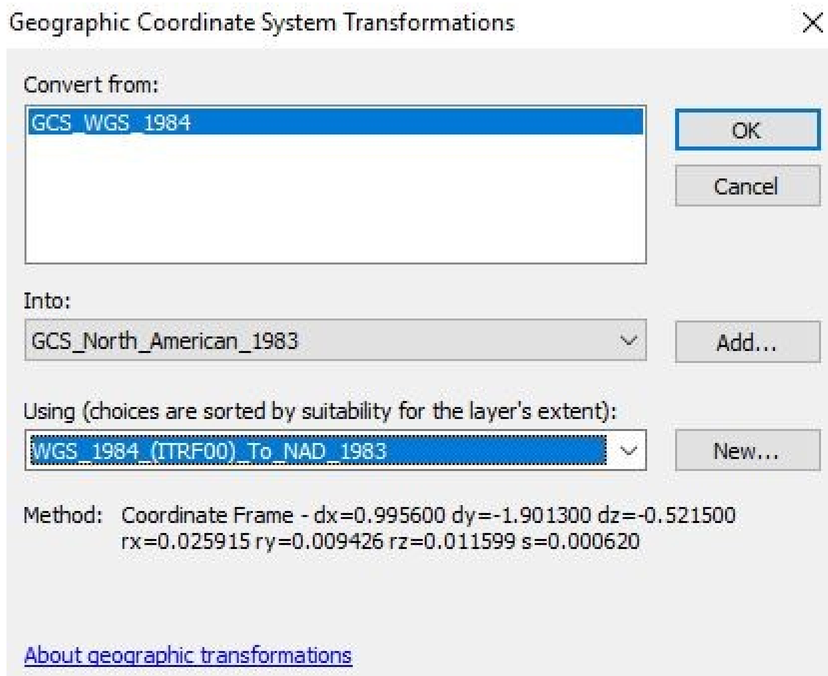

Figure 2–13. Geographic coordinate system transformations.

Self-Test Questions

After you complete these questions, you may check your answers at the end of the unit.

206. Convert to shapefiles and rasters

1. Describe the two input elements in the “Feature Class to Shapefile (multiple)” window.

2. How does catalog treat shapefiles as opposed to how File Explorer treats them?

3. Explain the difference between a raster that uses pyramids and one that does not.

207. Import into and extract from geodatabases

1. Where do we change the output file type when creating a feature class from an XY table?

2. What two parts of a database does the Extract Data Wizard tool transform?

3. What two format options can the Extract Data Wizard tool generate?

208. Convert coordinate systems

1. What makes up a geographic coordinate system?

2. The geographic coordinate system transformations window has three settings from which to select. What are they?

2–2. Export

Most customers and users will not need all data but instead need the data represented in a certain way. More than that, they need it in a certain format normally on paper or in a portable document format (PDF) document.

209. Layout print and plot a map

Map making is a common task for engineering. A map has two basic parts—data and layout. So far, we have talked about the part of a map where we display the data. This includes the labels, vector data, imagery, and so on. In this lesson, we talk about arranging the layout of the map. The layout consists of the tools used in reading a map. These include the north arrow, the scales, text describing who made the map, and so forth. Normally, this task comes last in map creation, right before we print it to paper or output it PDF for delivery.

Navigate the layout

In a Map document, there are two views—data view and layout view. We work with a map most often in the data view. The layout view is accessed either by selecting the very small paper icon on the bottom right of the table of contents or by navigating to it using the "View" dropdown at the top of the Map window.

Moving around the layout view ("paper" space) requires the use of the "Layout" toolbar (fig. 2–14). You need this because the data view navigation toolbar moves the data view around in the viewport. The "Layout" toolbar has tools for zooming in and out, panning, setting the percentage zoom, and so forth. Each of these tools, when hovering the mouse over them, provides a brief explanation of their use.

Figure 2–14. Layout toolbar.

Customize a layout

At the top of the window, to the left of the "Selection" dropdown is the "Insert" drop-down menu (fig. 2–15). In the "Insert" menu, we find common map layout elements. These elements make the map easier to understand. The most common elements are the legend, north arrow, scale bar, and the scale text.

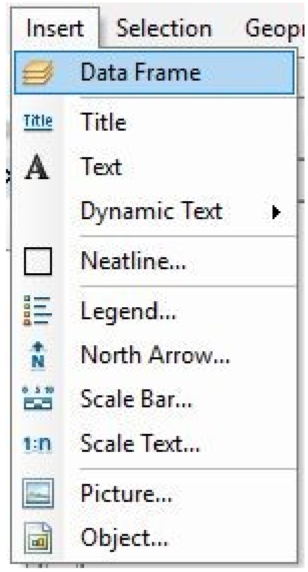

Figure 2–15. Insert menu.

Legend

Selecting "Legend…" from the Insert menu opens the "Legend Wizard (fig. 2–16)." This window is broken up into two columns—one for map layers and the other for legend items. We add and remove map layers to and from the legend using the arrow buttons located between the two columns. Below the columns, there is a setting for the number of columns in the legend. Use the column number control to help fit a legend neatly onto a layout. After you place the legend, item names take the names of the layers in the table of contents (TOC). Changing layer names in the TOC changes their name in the legend.

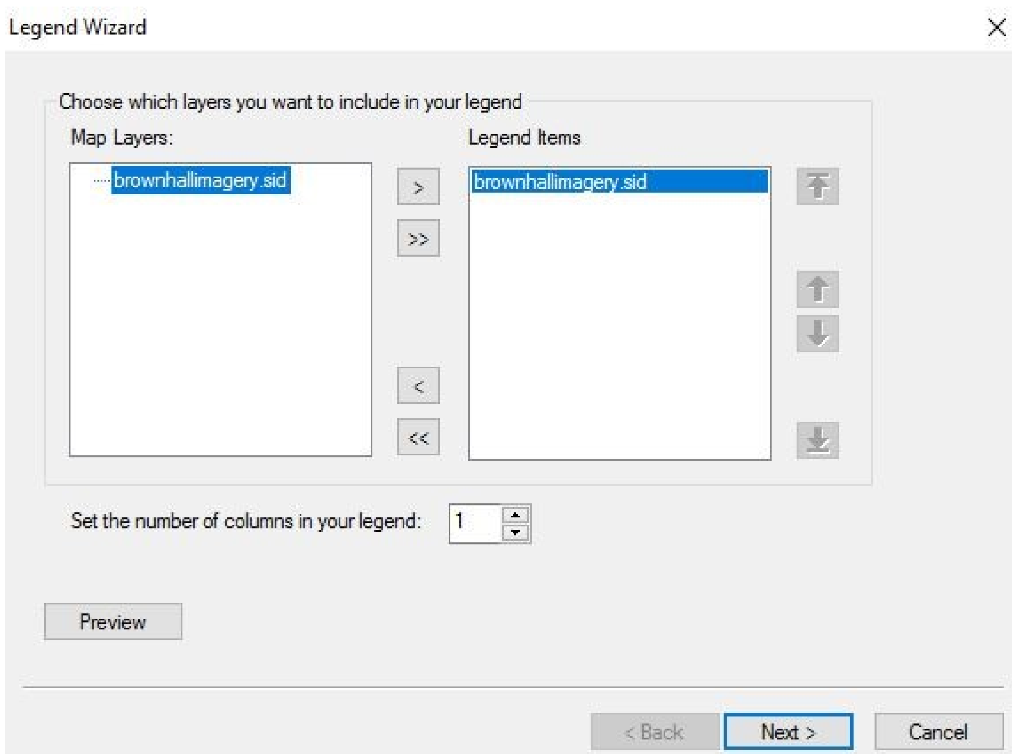

Figure 2–16. Legend wizard.

North arrow

When we select the "North Arrow…" option in the Insert menu, the "North Arrow Selector" opens (fig. 2–17). In this window, we select the style and adjust the properties of our north arrow. When placing the north arrow, the Map documents the coordinate system and rotation auto-rotates the north arrow to match.

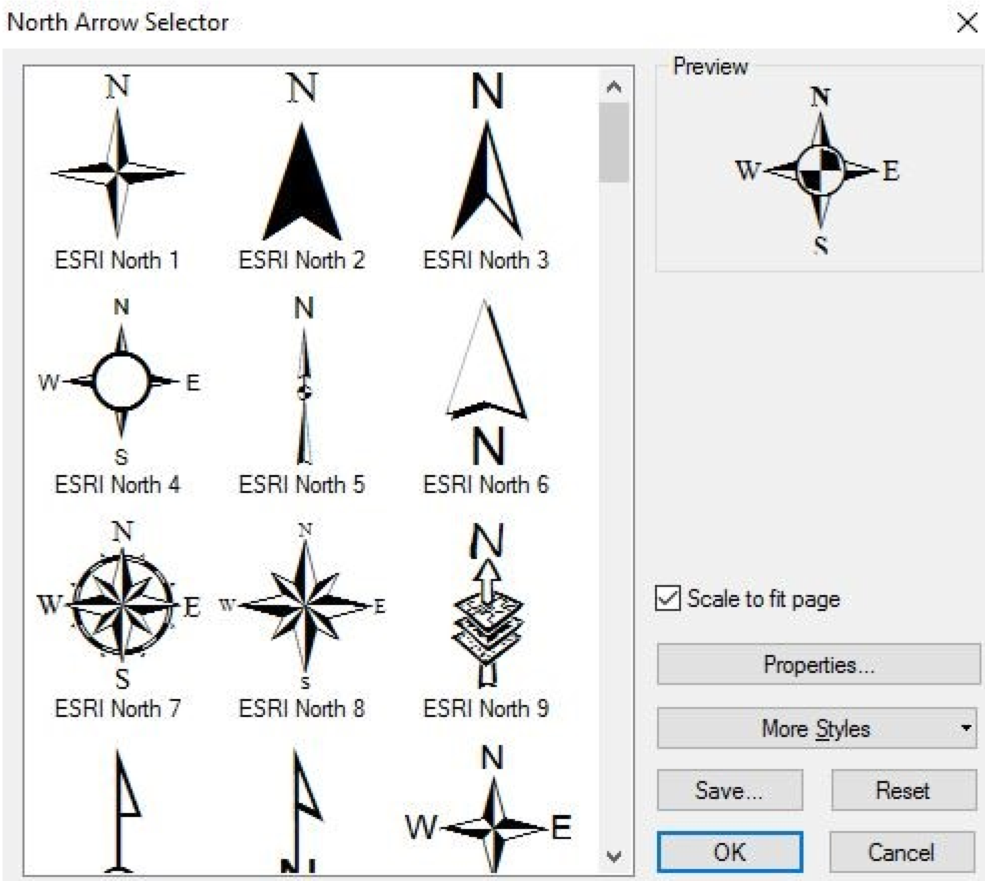

Figure 2–17. Select north arrow.

Scale bar

Selecting "Scale Bar…" in the Insert menu opens the "Scale Bar Selector" window (fig. 2–18). Like in the North Arrow Selector, the style of bar is selected, and more properties are found under the properties menu. The most important properties to be aware of are the division properties. Division settings change the number of divisions in the scale bar. The scale bar sizes automatically to match the scale of the overall Map document. The properties menu also has settings to change the units (miles, kilometers, etc.), which are labeled on the scale.

Scale text

The "Scale Text…" item in the Insert menu opens the "Scale Text Selector" window (fig 2–19). This window allows us to configure the properties of scale text. Scale text is the scale of the map expressed as a ratio. This allows the user to take a ruler, or other graduated scale, and take measurements on the map to a certain degree of accuracy. The properties menu provides options for adjusting the text properties, such as font, bold, italic, font color, font size, and so forth.

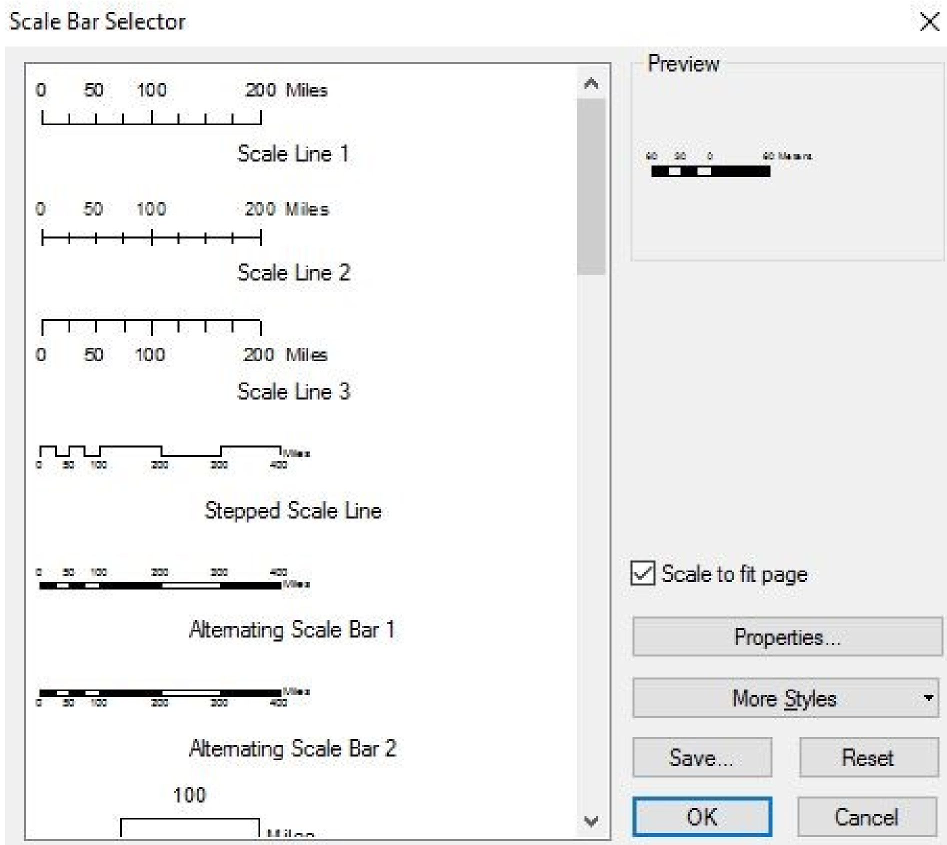

Figure 2–18. Scale bar selector.

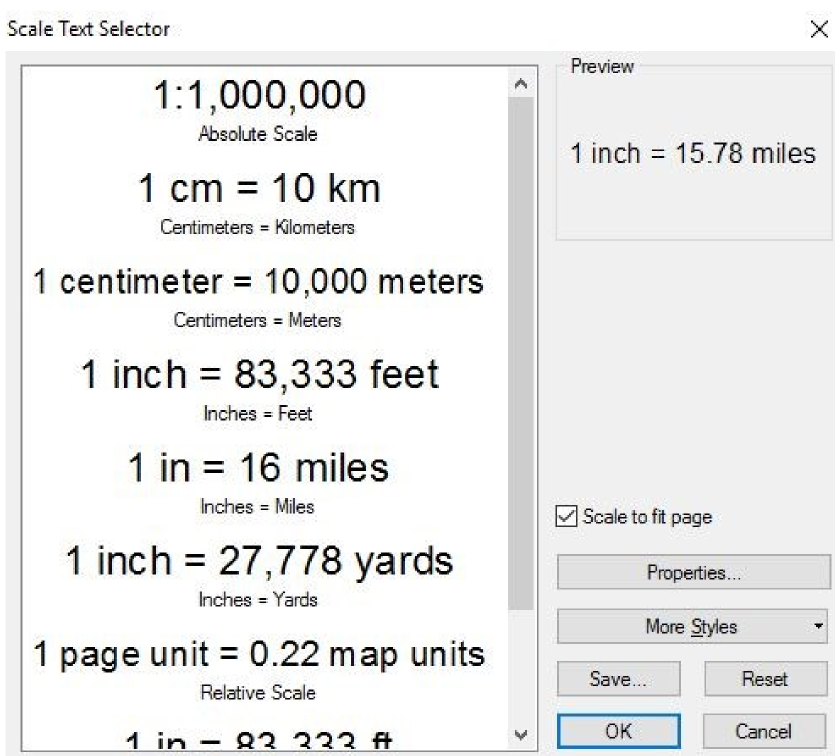

Figure 2–19. Scale text selector.

Draw toolbar

In some cases, you may need to add additional markings to a layout. The “Draw” toolbar has tools for creating graphics as opposed to the vector data in the TOC (fig. 2–20). Add the Draw toolbar to your Map document in the same way you would add any other. Use the Draw tools to add marks and text that do not have associated attribute data or that are unique to the customers’ needs. The tools range from drawing rectangles and polygons, creating text, and providing formatting for both.

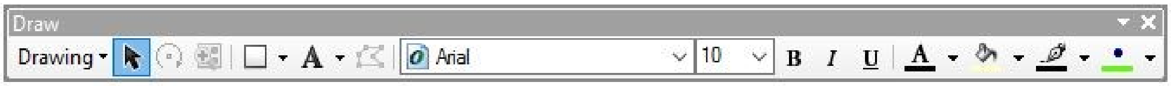

Figure 2–20. Draw toolbar.

Page and print setup

The “Page and Print Setup…” function is located within the File drop-down menu about halfway down. This window provides access to options for printing the map document to paper or creating an image of the map. Here, we adjust the size of the paper in the layout view, which printer to use, and how to orient the page (fig. 2–21). Keep in mind that these adjustments are to the paper area in the layout view and will not automatically adjust elements in the layout. Each item in the layout requires individual attention. You should adjust the “Page and Print Setup…” options before creating or manipulating objects in the layout view.

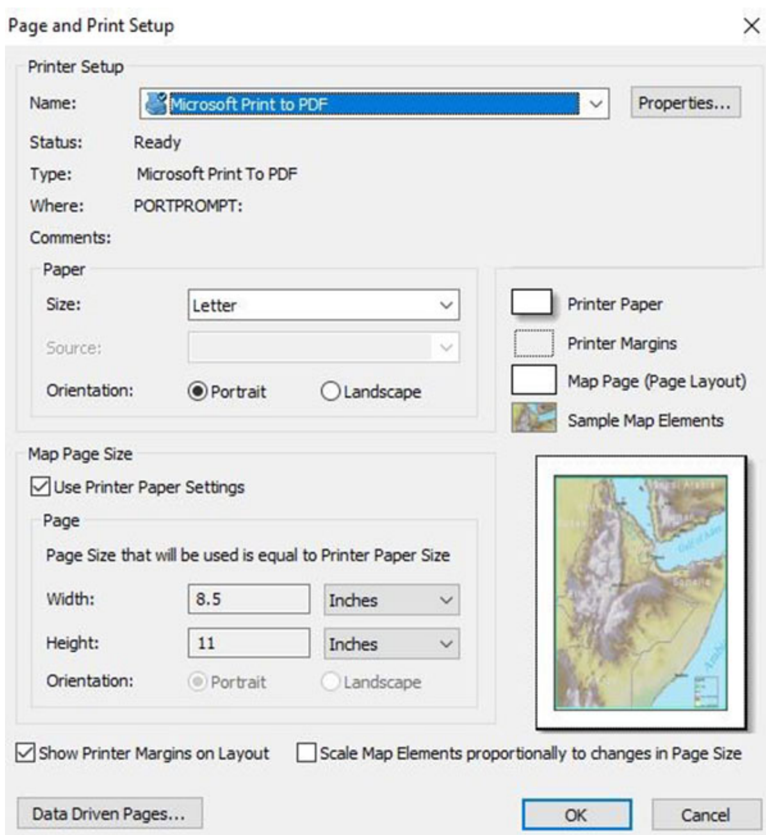

Figure 2–21. Page and print setup window.

Export map

Another way to create an image of your map is by using the “Export Map…” function located in the File menu. Unlike “Page and Print Setup…,” “Export Map…” creates images of the data view as well as the layout view. The “Export Map” window looks much like a normal file navigation window except for the tabs at the bottom. These tabs are different depending upon the file type you set as export. The most common and useful controls are for dots per inch (DPI) and the output image quality slider (fig. 2–22). Either we adjust the slider manually or we use the ratio to set a specific quality between the map we export and the image we create. A ratio of 1:2 means that the image created is half the quality of the original map. DPI can be set higher or lower and has a major effect on the image quality. Keep in mind that a better, quality image equates to a larger file size and may not be deliverable by email.

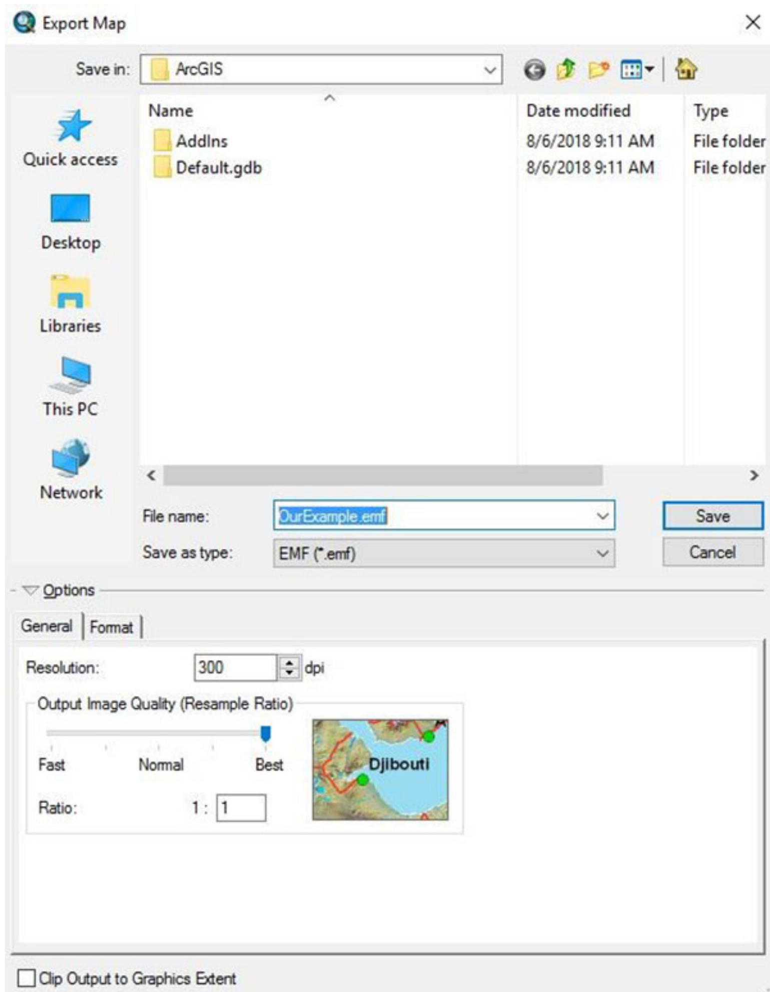

Figure 2–22. Export map window.

210. Export data

There are times when you need to move data from one place to another. Our ability to move the data efficiently is referred to as data portability. Portable data is data that can be easily changed to another format. In a Map document, right clicking on a layer in the TOC opens the context menu. Hovering over the “Data” option in the context menu opens another menu that has a tool called “Export Data (fig. 2–23).” We use the Export Data tool to create a new file of the layer’s data in a different format.

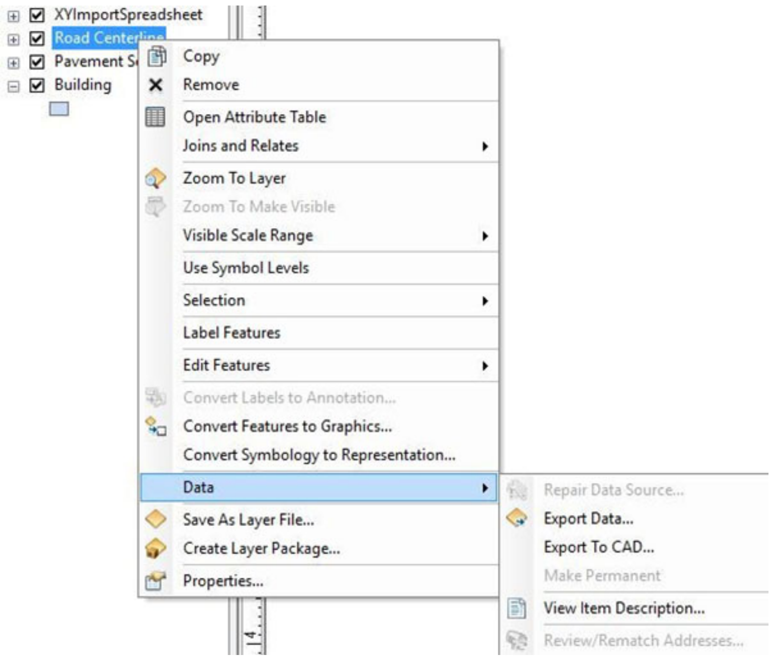

Figure 2–23. Export data location.

Selecting this tool opens the “Export Data” window. Here, the tool asks which features in the layer to export, which coordinate system to use, and the location of the output file. To change the format of the file before export, we select the folder icon to the right of the output file path (fig 2–24).

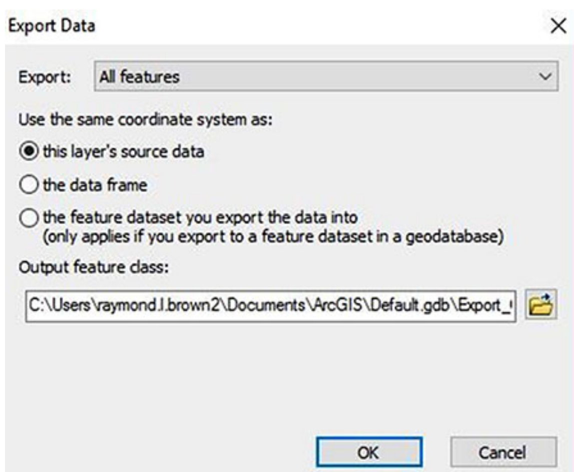

Figure 2–24. Export data window.

The folder icon opens the file navigator. The drop-down list labeled “Save as type:” contains the formats into which the program allows the export (fig 2–25). When satisfied, we select the “Save” button to return to the “Export Data” window. When satisfied with our export settings, selecting “OK” completes the export and saves the output file to the location we specified.

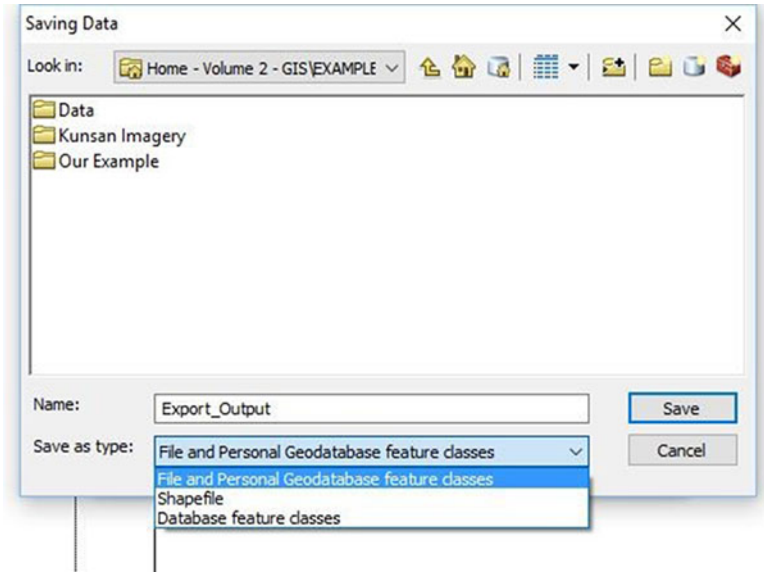

Figure 2–25. Export format dropdown.

211. Quality assurance/quality control

Data quality assurance (QA)/quality control (QC) is the way we implement standards and regulations created at many levels. The cornerstone of which the data has a degree of dependability. The amount of effort we put into meeting standards and regulations reflects our data's dependability. There are standards applied at the national level, within the Office of the Secretary of Defense (OSD), within each service, within each major command (MAJCOM), and finally at each installation. Since data communication of data is so important, it is imperative everyone recognizes it through a set of common data standards. Let us look at a few of these standards and then go over some issues to be aware of in our own data.

Raster data quality assurance/quality control

Raster data is very easy to QA/QC. A poor quality or improperly placed raster file will cause a lot of work later.

The key aspects are:

- The resolution of the data is high enough to suit the user's needs. Typically, for our CIP data, the commercial imagery the Air Force purchases come in six-inches but requires it to be one-meter.
- Like resolution, the spectrum should meet the user's needs. For our CIP data, we use color photos, but there are several other types 'available (panchromative, multispectral, hyper-spectral, etc.).
- The data is properly georeferenced.
- To be able to see all the ground features, ideally the image should have less than 20 percent cloud coverage, the ground should not be snow-covered, there should not be excess standing water from rain, and limited amount of foliage on the trees.
- The image has a nadir angle of less than 25 degrees. This means that if a line were drawn perpendicular to the ground from the aerial sensor, the angle from the sensor should not exceed 25 degrees.

Vector data quality assurance/quality control

Vector data is more complicated to QA/QC due to the multiple data types and the number of potential data stewards.

The key aspects are:

- The data is properly georeferenced. You must consider as the data been arbitrarily been given a coordinate value or was it collected within a coordinate system.
- The data is organized according to the data geometry (point, line, or polygon).
- Adjoining polygon features should be coincident and conflated, without gaps or spaces. Digitized or improperly extracted data is the common cause of this issue.
- The data is stored in the correct data schema. In other words, is everything in the correct feature class?
- We have entered sufficient attribute data. For CIP features, this is within the CIP control document discussed earlier.
- Correct application of topology rules.

Topology rules

There are three simple topology rules to which all geospatial data should adhere. First, features should meet or terminate at proper nodes. A node in a geodatabase is the point representing the beginning or ending point of an edge. In topology, they link, or snap, to all the edges that meet. Second, if a feature is a polygon, it closes on itself. A polygon that does not close on itself is only a line. Last, there should be no duplicates or ghosted features. These happen sometimes when there are multiple copies of a feature like a building polygon placed over each other.

Parsed

Parsed means the data is broken down and looked at in its component parts. Data clearly displays as points, lines, or polygons. You may have a rule that says a polygon feature must have a perimeter larger than five feet. This means that a polygon with a perimeter less than five feet, we draw as a point to represent the data correctly.

Attribute data is cleanly separated. The standard format is for the columns to represent each attribute field. Each field should contain only one data type (that is, integer, text, Boolean, etc.). It is important we identify the primary and foreign keys. The rows of data then represent each discrete feature.

Self-Test Questions

After you complete these questions, you may check your answers at the end of the unit.

209. Layout print and plot a map

1. Why do we need a separate toolbar to navigate the layout view?

2. What are the four most common elements in the layout and under which menu do we find them?

3. What menu do you need to access to modify the text and coloring settings for the north arrow, scale bar, and scale text?

4. What are the two menus used to format a map’s layout space before making an image?

210. Export data

1. What three things does the export data menu ask for?

2. Where do we select the type of output file format we want?

211. Quality assurance/quality control

1. What is the cornerstone of QA/QC, and how do we meet this standard?

2. Regarding QA/QC, what is the usual resolution of our imagery and what is the requirement?

3. What is the QA/QC rule for the data schema of vector data?

4. According to topology rules, what is the definition of a polygon?

Answers to Self-Test Questions

206

1. The first is the file path of the file we plan to convert. The second is a list of output file types.
2. Catalog treats shapefiles as a single file though they are a group of files. File Explorer treats each file in the group as an individual file.
3. Raster files without pyramids render every pixel at every scale or zoom factor. Raster files that use pyramids create lower resolution “samples” to efficiently render raster files as different scales and zoom factors.

207

1. In the “Create Feature Class from XY Table” window, in the output section, when specifying the output file path by selecting “Save as type:” dropdown.
2. The data or the schema.
3. Geodatabase or .xml document.

208

1. Different geoids, spheroids, and ellipsoids.
2. Convert from, into, and using.

209

1. The data view navigation toolbar allows us to navigate the viewport while the Layout toolbar allows us to navigate the “paper” space.
2. Legend, north arrow, scale bar, scale text, found under the “Insert” menu.
3. Properties.
4. (1) Page and print setup.

(2) Export map.

210

1. (1) Which features in the layer to export.
   (2) Which coordinate system to use.
   (3) The location of the output file.
2. In the file navigator under “Export format” dropdown.”

211

1. Degree of dependability, met by the amount of effort we put into meeting regulations.
2. Six-inches but the requirement is one-meter.
3. It needs to be in the proper feature class.
4. A line that closes on itself.

Complete the unit review exercises before going to the next unit.

Unit Review Exercises

Note to Student: Consider all choices carefully, select the *best* answer to each question, and *circle* the corresponding letter. When you have completed all unit review exercises, transfer your answers to the Field-Scoring Answer Sheet.

Do not return your answer sheet to the Air Force Career Development Academy (AFCDA).

15. (206) When exporting a .dwg to a shapefile, how many shapefile types are output by default?
    a. Five.
    b. Four.
    c. Three.
    d. Two.

16. (206) How does catalog view and treat multifile shapefiles?
    a. No differently than file explorer.
    b. Grouped together, as a single file.
    c. As individual parts of the shapefile.
    d. Only the .shp file is viewed or treated.

17. (206) We convert most raster, or image, files by changing the
    a. suffix.
    b. file size.
    c. location.
    d. filename.

18. (206) We group grids together by creating "samples" at larger scales by way of
    a. cylinders.
    b. pyramids.
    c. polygons.
    d. diamonds.

19. (207) What three input fields do we specify when using the "Create Feature Class from XY Table" tool?
    a. Field1, Field2, and Field3.
    b. X field, Y field, and Z field.
    c. northing, easting, and elevation.
    d. point number, point code, and point type.

20. (207) When converting a spreadsheet to a feature class, where must the file output be located?
    a. Geodatabase.
    b. Attribute table.
    c. Compressed folder.
    d. Microsoft access database.

21. (208) What are mathematical models that translate between coordinate systems?
    a. Evolutions.
    b. Translations.
    c. Declinations.
    d. Transformations.

22. (208) In the “Geographic Coordinate System Transformations” window, how does the program rank options in the “Using” dropdown list?
    a. Alphabetically.
    b. Least suitable on top.
    c. Most suitable on top.
    d. Based upon date of creation.

23. (209) The names of items in a legend replicate the names of
    a. layers.
    b. data frame properties.
    c. the legend attribute fields.
    d. feature classes in the database.

24. (209) Which toolbar has tools for creating graphics as opposed to vector data?
    a. Edit.
    b. Draw.
    c. Georeferencing.
    d. Advanced editing.

25. (209) Where do we adjust the size of the layout area and its orientation?
    a. Export Map.
    b. Print Preview.
    c. Printer Properties.
    d. Page and Print Setup.

26. (209) In the “Export Map” window, what two settings adjust the quality of the output image?
    a. General and Format.
    b. Resolution and dots per inch (DPI).
    c. Resample Ratio and Image Quality.
    d. Output Image Quality and Resolution.

27. (210) Which tool do we use to create a new file of a layer’s data in a different format?
    a. Draw.
    b. Export Data.
    c. Edit Features.
    d. Advanced Editor.

28. (210) Which item does the “Export Data” window *not* ask us to provide?
    a. Format.
    b. Output location.
    c. Features to export.
    d. Coordinate system.

29. (211) What is the first topology rule for vector data listed in the text?
    a. Proper georeferencing.
    b. No duplicate or ghost features.
    c. Polygons close on themselves.
    d. Features terminate at proper nodes.

Please read the unit menu for unit 3 and continue ➔

Unit 3. Managing Data

THERE ARE SOME forms of data, or manipulations of data, that are unique. Informally, we can term some forms and manipulations of data as "special." These data forms are unique and the administration of them requires special attention. This unit is all about this kind of data. We are also concerned with data relationships, since that is at the core of every GIS's usefulness. Our ability to create useful connections between data types and their tables makes the information dynamic. Dynamic data changes and evolves with new information. That new information must have some kind of connection to the related data; otherwise, it is just numbers or letters on a page.

3–1. Special Data

At times, you will need to manipulate and convert rarer data formats. Images, for example, come in many formats, and you have to adjust them to fit the data. We accomplish this using a procedure referred to as georeferencing. Another rare form of data is metadata. Metadata describes the data itself. In other words, it is the data about the data. Finally, subtypes allow us to organize and further relate one piece of data to another. These rare data formats are not as common as say, vector data, but they are just as important. Keep in mind that your primary goal is to display and organize data in the most useful way to allow the customer to solve their problem or accomplish their mission.

212. Georeferencing an image

Aerial imagery, most often, does not match data collected by survey methods. If you take measurements of a building's corners, for example, and compare the corners to an aerial image they will not be the same. A number of factors, mostly the angle at which the image collected and the type of lens used, cause the slight difference between the two. To correct this, you use a process called "georeferencing."

To georeference an object, you first need the correct tools. The georeferencing toolbar is located under the context menu when right clicking in the empty space on the top of the window. Once you have the toolbar, the image, and the data you want to reference it to include on the Map document, you can begin.

First, you check that the georeferencing toolbar is "looking" at the image you want to reference. When it is "looking," the image name displays on the toolbar. Second, you find the "Add Control Points" function on the toolbar (fig. 3–1). This function allows you to select a point on the image and stretch it to another point, say, our survey points.

As you select each point and stretch it to the actual points, the program is distorting the image. This may change the relative positions of the points and the image as you georeference each point. It is necessary to double-check and reference some points multiple times before the image fits the way you want it (fig. 3–2). Once you have finished adding control points, you can do two things—save a separate image file or change the image you have added to our map.

To save the georeferenced image as a separate file, you select "Georeferencing" on the toolbar and select "Rectify." This allows you to specify the output location and file name (fig. 3–3). Selecting "Update Georeferencing" instead, applies the modification to the image and ends the georeferencing "session."

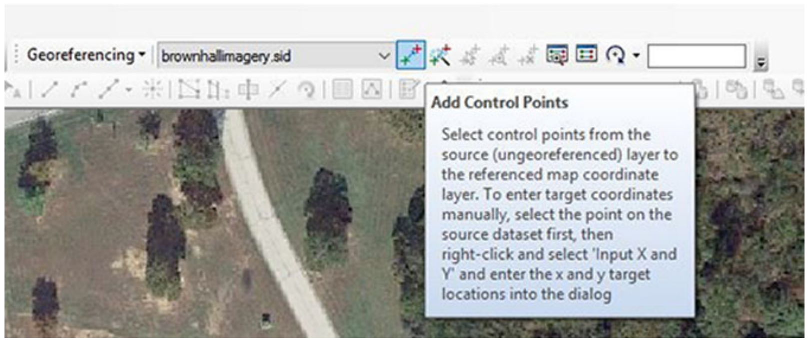

Figure 3–1. Add control points function.

Figure 3–2. Georeferencing building corners.

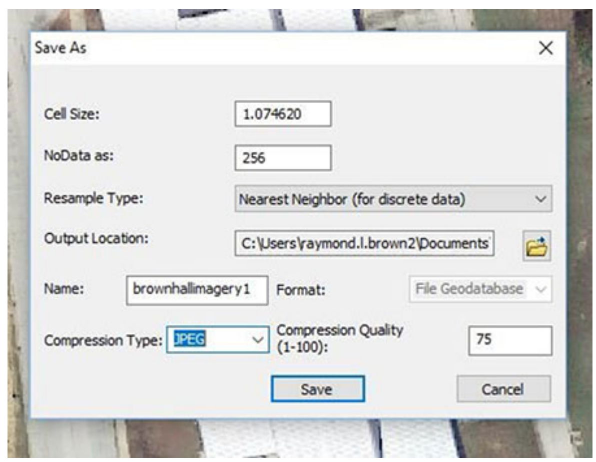

Figure 3–3. Rectify save as window.

213. Using metadata

Metadata is the data… about the data. It is information describing who owns or collects the data, why you need it, what it is for, and anything else a geographic information officer (GIO) may want to explain about the data. The clearest way to view metadata is through catalog.

View metadata

In catalog, select the feature class that you want information about, such as "Building," in the TOC. In the display pane, at the top, are three tabs with the right most labeled "Description." The description tab shows the metadata for the feature class (fig. 3–4). This display can be set to different regulations of information display depending on the "Metadata Style" setting.

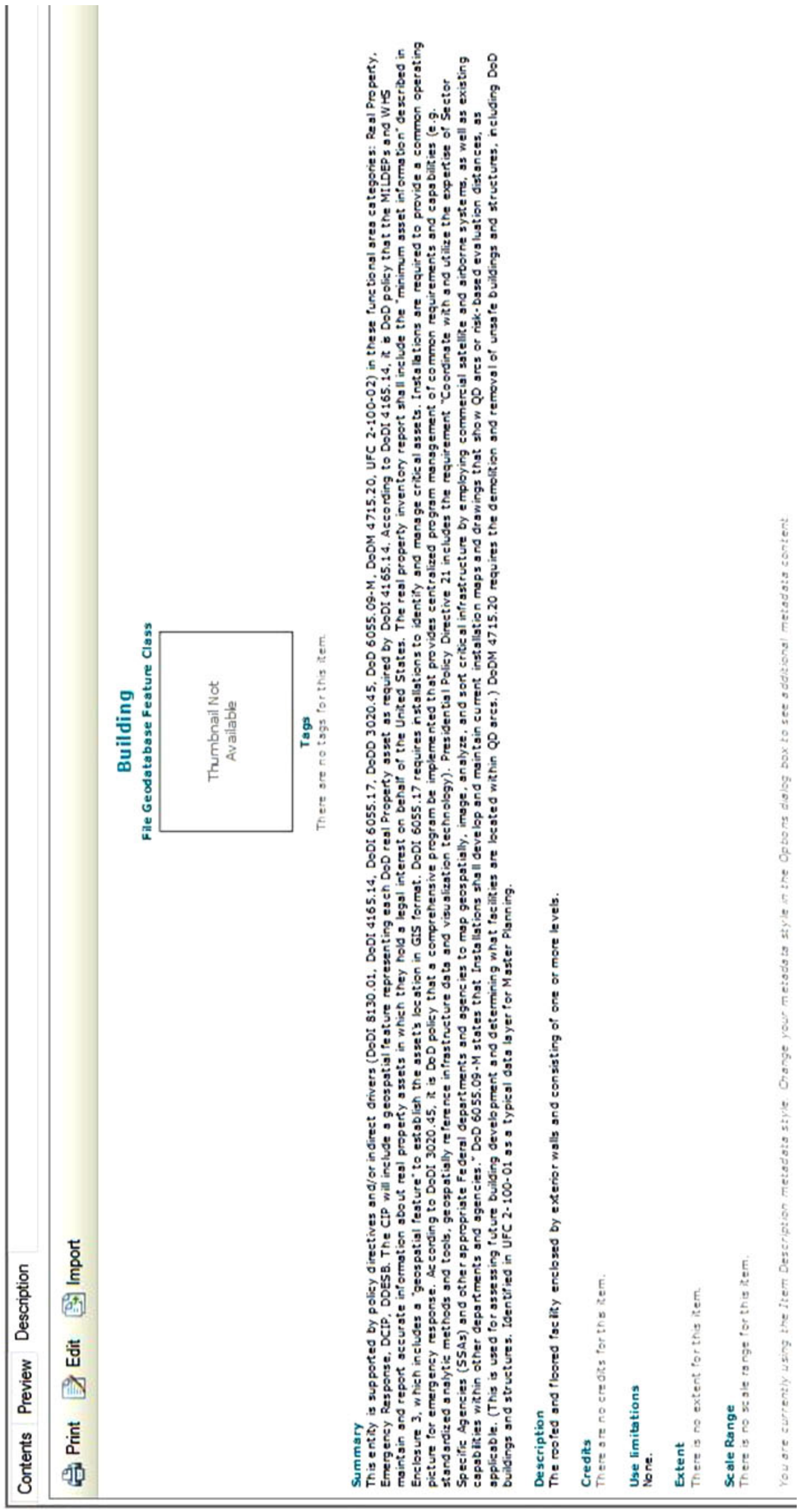

Figure 3–4. Description tab.

The “Metadata Style” setting is in the “Customize” dropdown at the top of the Catalog window. At the bottom of the “Customize” list is “ArcCatalog Options.” Once selected, this opens a window that has the “Metadata Style” setting at the top (fig. 3–5).

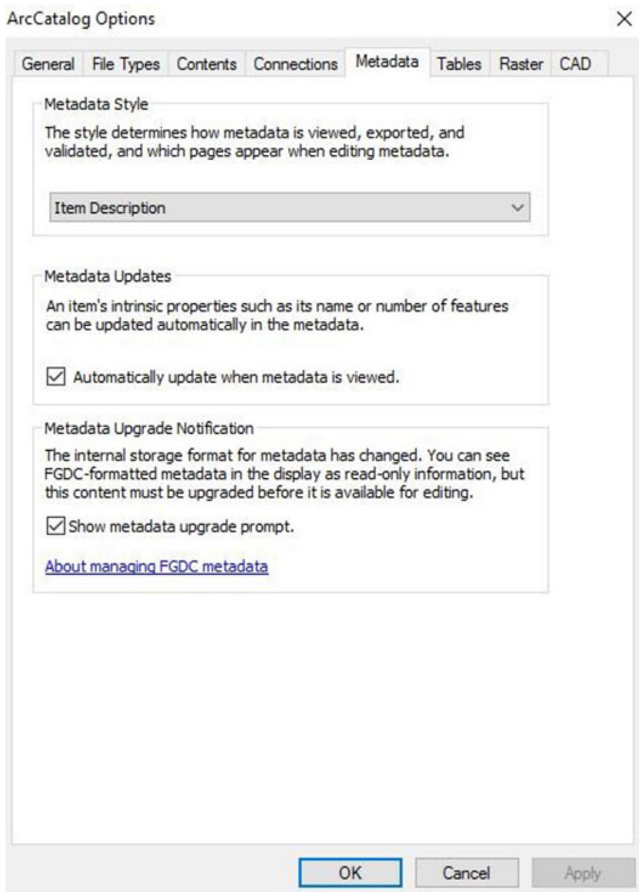

Figure 3–5. ArcCatalog options.

The dropdown list provides different ways the metadata displays the data in the description tab. The default is “Item Description,” which provides a summary of the data. The others are more specific and are dependent upon which standard your GIO maintains or requires updated.

Edit metadata

Editing metadata is simple. Once you have decided which style to display the metadata in, select the “Edit” button below the tab labels at the top of the display pane. Doing so reveals the text editor boxes the descriptions reside in, and you edit them as if they are in a word processing program. When finished, save the edits by selecting the “Edit” button again (fig. 3–6).

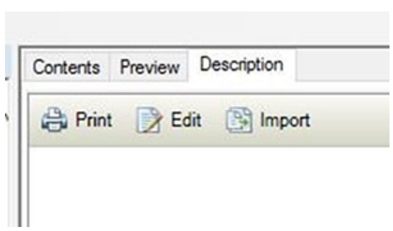

Figure 3–6. Edit metadata

214. Subtypes

The attribute table provides us a way to distinguish data within the same category. That category, the feature class, is usually further divided into subtypes. Subtypes are groups of features that share the same attributes or more commonly share an attribute. The shared attribute is referred to as the subtype field (or attribute).

We use subtypes to improve the efficiency of our data. When you need, for example, to define different kinds of roads, you use subtypes such as a local street, a collector street, and arterial streets defined in the attribution as our subtypes. Further, each subtype can be applied a default value. For example, when adding a new road area and then adding it to the local street subtype, the attribution would default to 35 miles per hour (mph). Subtypes also restrict data entered into a field to only those defined as a subtype. An example would be water mains whose field values are restricted to cast iron or copper. In order to create subtypes, we start with a brand new field of either the long or short integer type (fig. 3–7). This is important because no other data type can be a subtype. Our field is the “Marked_by” field describing how these survey points are marked in the field.

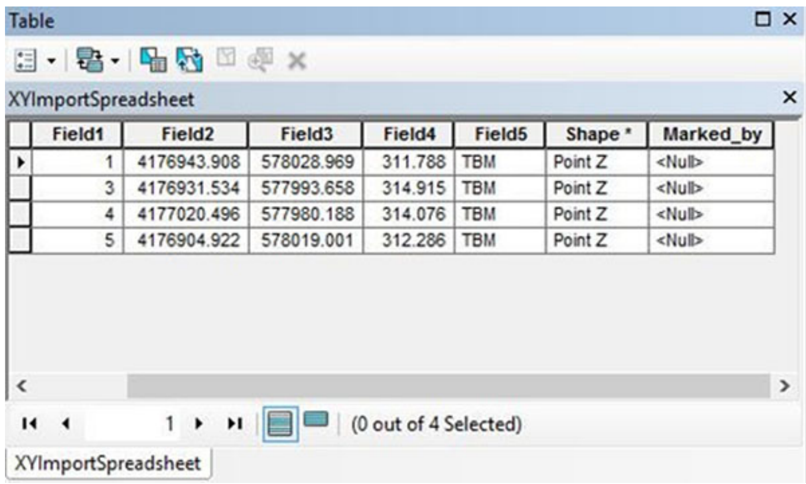

Figure 3–7. New field to start.

Like any task, you need tools with which to do it. The subtype tools are within the Toolbox application under “Subtypes.” There are four tools in the Subtype category “Set Subtype Field,” “Add Subtype,” “Set Default Subtype,” and “Remove Subtype (fig 3–8).”

Set the subtype field

To create a group of subtypes, you first define the attribute that their group is based upon. For example, you define subtype based upon how the survey markers are marked. The “Set Subtype Field” tool asks for two pieces of information—the file path of the table and the field that the subtypes share (fig. 3–9). Our field is “Marked_by.” With settings in place, select “OK.”

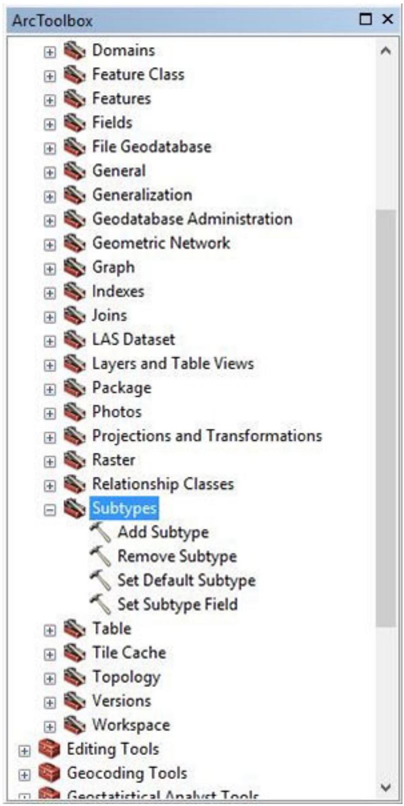

Figure 3–8. Subtype category in Toolbox.

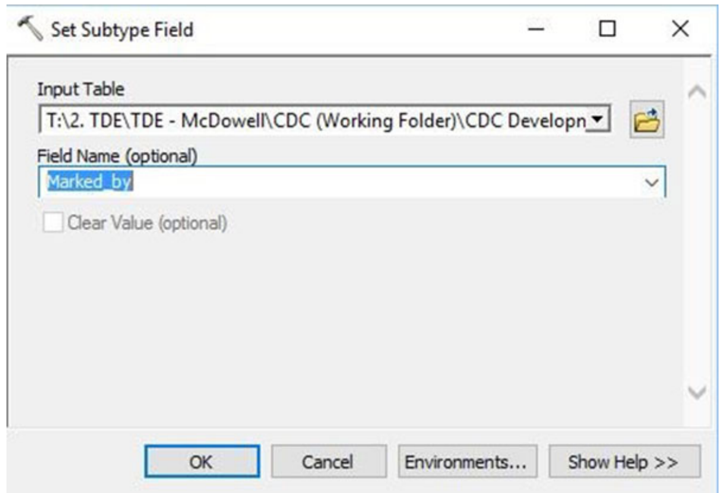

Figure 3–9. Subtype field window.

Adding subtypes

Everything is now set up to add different subtypes. You begin by selecting the “Add Subtype” tool, which opens its respective window. The window asks for the input table’s file path, the subtype code, and the subtype name. The file path is self-explanatory; it is the place where the table is located. The subtype code is the number you assign to the subtype. This integer works behind the scenes to maintain the subtypes and is a connection to the attribute field. The last item is the “Subtype Name,” which is the name of whatever we are defining. In our case, the “Subtype Code” for a marker made of rebar is the integer “1 (fig. 3–10).” Once satisfied, select “OK” and repeat the “Add Subtype” process for each subtype. There are three in our example, “0” for “Unmarked,” “1” for “Rebar,” and “2” for “Stake and Flag.”

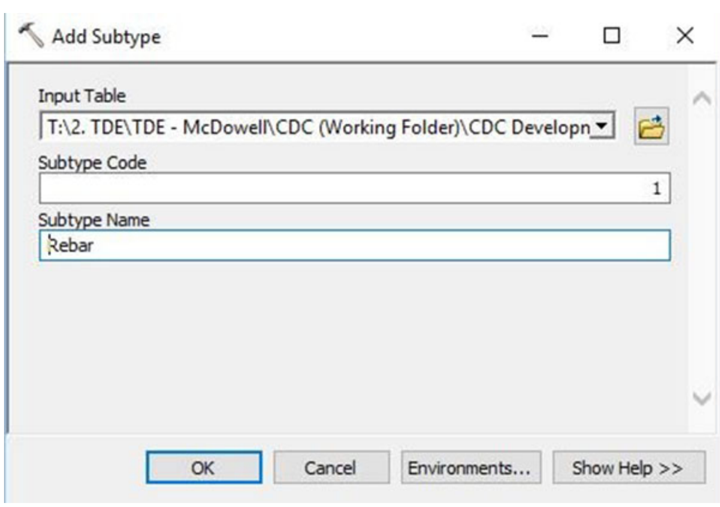

Figure 3–10. Add subtype window.

Setting the default subtype

Let us say that your flight always marks survey point the same way—with rebar. It would be tedious to go into the data after you add it to the database and change each point to the rebar value. It is much easier to have a default value that auto-populates each time a new item is generated or added to the feature class. The “Set Default Subtype” tool does just that. To use it, open the tool, specify the table, and specify the default subtype code (fig. 3–11). After selecting “OK,” each item added to the feature class will start with the specified subtype code value. This does not apply to existing values in the attribute table.

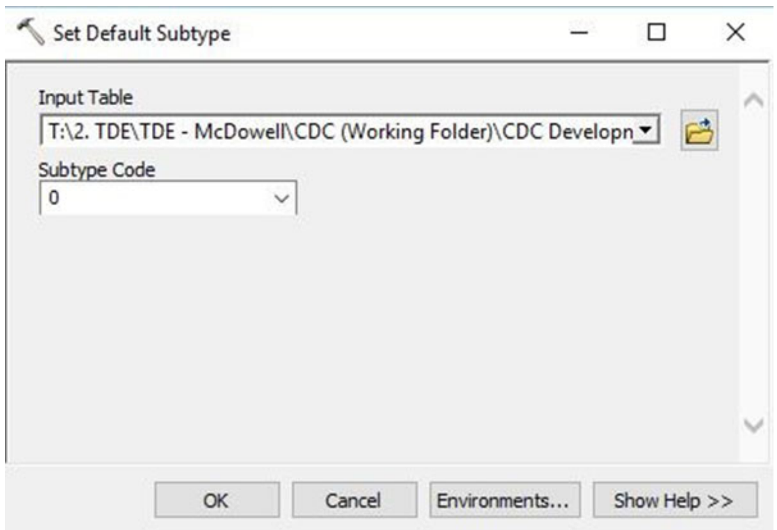

Figure 3–11. Set default subtype.

Removing subtypes

When a subtype is no longer needed, use the “Remove Subtype” tool to get rid of it. This tool asks for the table in question then provides a list allowing you to select which subtypes you want to remove (fig. 3–12). Once you have selected the subtypes to remove, select “OK” and they will no longer be available for selection.

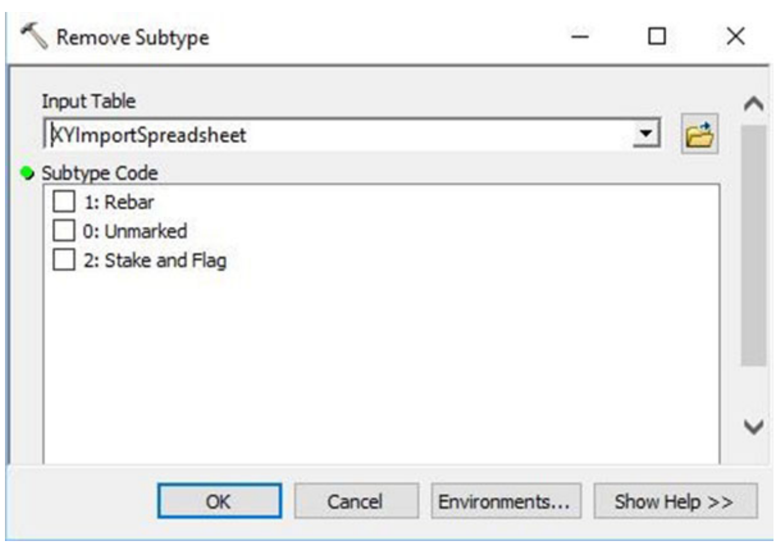

Figure 3–12. Remove subtypes.

There are some benefits of using subtypes. The first is that each subtype can have topology rules applied to it. In other words, a feature class with topology rules can have another set of topology rules, compounding the behavior of the subtypes. For example, if you used “paved” and “unpaved” subtypes for a road feature class, you would set the “paved” subtype to be required to connect to other “paved” subtypes within a certain distance. For the “unpaved” subtype, this topology rule does not have to apply.

Another benefit is the relationship rules. This means subtypes for telephone poles made of either “steel” or “wood” can be restricted in which type of transformer can be specified in the attribute table. You can set “steel” to carry Class A transformers and “wood” only Class B. Basically, anything that applies to a feature class can apply to a subtype. The attribute table now distinguishes between each type of marker. You can be more specific about which survey points you want to look for (fig. 3–13).

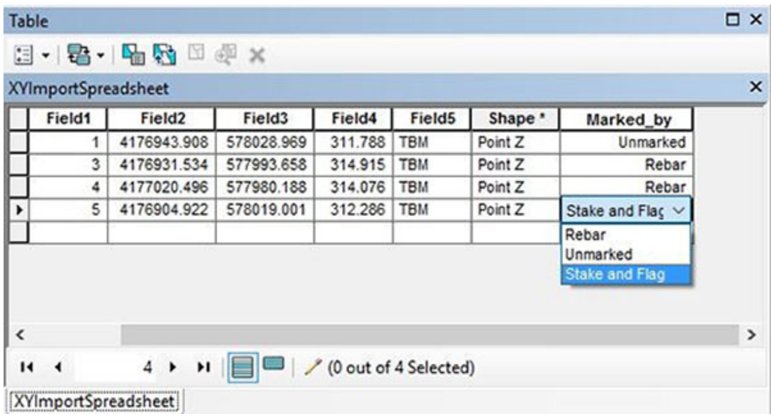

Figure 3–13. Subtypes in table.

Self-Test Questions

After you complete these questions, you may check your answers at the end of the unit.

212. Georeferencing an image

1. What do you check on the georeferencing toolbar before you use the Add Control Points tool?

2. As you select data points and stretch the image to actual points, what is the program doing to the image?

213. Using metadata

1. What kinds of things do metadata describe?

2. From the description tab in catalog, how do we start and stop editing metadata?

214. Subtypes

1. What is a subtype?

2. List the four subtype tools.

3. When adding subtypes, what two values do you need besides the input table?

4. Provide a description of a default subtype.

3–2. Relating Data

A GIS is all about data, more specifically it is about the relationship between pieces of data. Merging, joining, and relating data provide us a way to enhance our data's usefulness. We accomplish this in a number of ways. Merging is most useful when combining similar data together. Joining creates an addition to data tables of one feature class to another. Relating data creates a relationship so that two data tables can refer to each other so that we can get the most out of our data. It is important to remember that our goal is useful data. The more useful the data presented to the customer, the better decisions they can make toward the accomplishment of their respective missions.

215. Merge data

When we talk about merging data, we are taking two or more sets of data of the same type (point, line, or polygon) and creating a new set of data. This new dataset will include all the data from each input set in one file. For example, let's say that a group of five Airmen completes a survey and export their point data as shapefiles. All of the shapefiles are in the same format, meaning that their fields are identical. Instead of selecting each feature and either editing it into another feature class or shapefile, or copying the data into a new container, merge creates a new combined file.

To merge data sets together, navigate to the toolbox and expand the "Data Management Tools" list. Under that, we expand the "General" list revealing a tool called "Merge (fig. 3–14)."

Selecting this tool opens the "Merge" window. This window is broken into three parts—the input dataset, output dataset, and the field map (fig. 3–15). The input datasets are the sets of data (shapefiles, feature classes, tables, etc.) that we plan to combine. You use the second part, output dataset, to specify the place in the file system you want to place the finished product. Finally, the field map options list the fields created for the output dataset. Here, you have an opportunity to remove, add, and modify fields. Additionally, you see how the merge tool treats the input datasets through the properties of the fields.

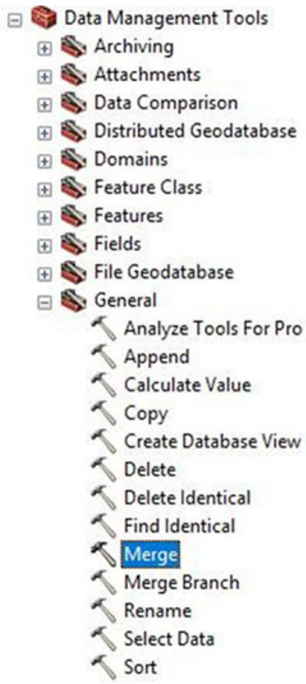

Figure 3–14. Merge tool location.

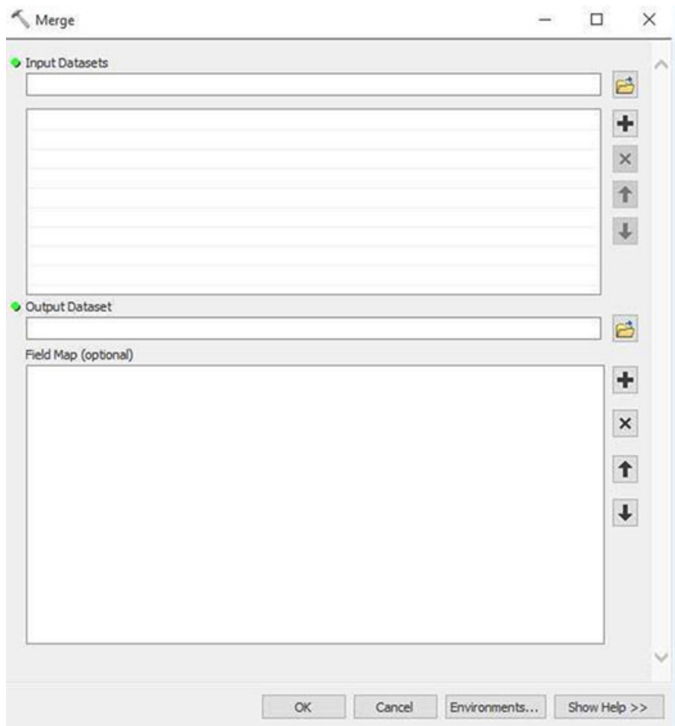

Figure 3–15. Merge tool window.

216. Joining and relating data

After the data, the largest part of a GIS is the data's relationship to other data. Let us say that you have a field in the building feature class that is populated with the name of the owning unit. Each building has a building manager that maintains it. The operations flight, in the civil engineering (CE) squadron, maintains the list of building managers. This list is updated as members make a permanent change of station (PCS) in, out, or move in and out of the position. It would not be practical for the operations flight to notify us each time a person on the list changes, there are hundreds! What would be better is if the GIS could "pull" information from the spreadsheet maintained on the network drive and update the list of facility managers in the database as it changes. Our GIS has this functionality in the form of the Joins and Relates tools. We access both tools by right clicking on the layer we want to join or relate and accessing the context menu (fig. 3–16).

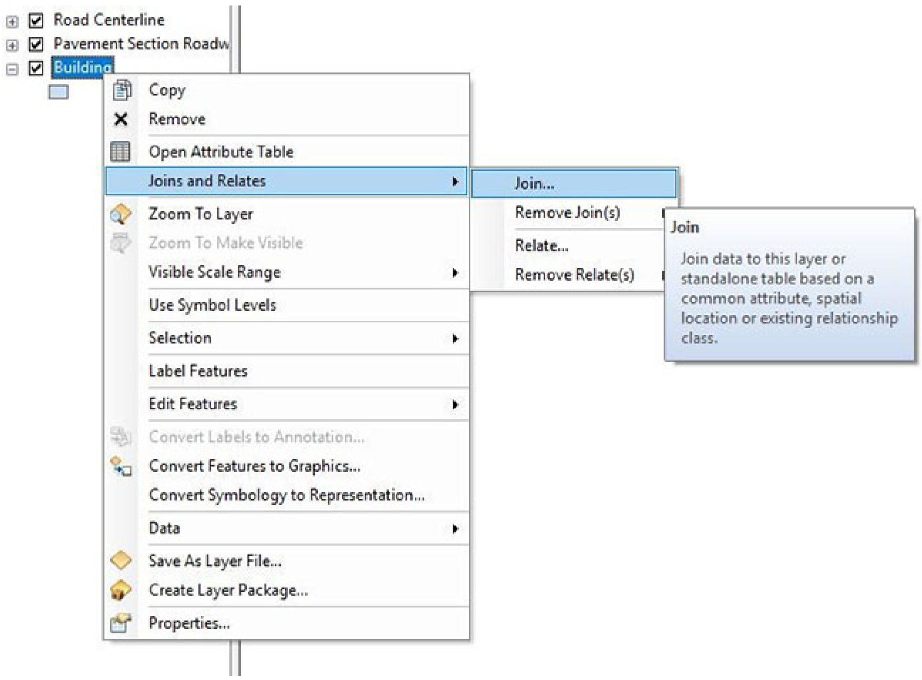

Figure 3–16. Joins and relates tools' location.

Joins

The best way to understand joined data is by example. We have a field for building number in our "Buildings" feature class. A customer wants us to analyze the number of occupants in each building. A join allows us to connect the data layer in our map to the spreadsheet or database the user has with the number of occupants per building. This allows us to mark each building and graphically display them based upon total number of occupants. To create a join, we find it in the context menu and select it to open the "Join Data" window (fig. 3–17).

This window asks you a series of questions. The first is "What do you want to join to this layer?," which has two options. The first option is "Join attributes from a table," which tells the program to join layers or tables based upon the attribute tables instead of the spatial location. The second option is "Join data from another layer based on spatial location," which allows you to select individual or groups of layers based on how close or how far away they are from the layer to which you are joining them.

The next question asks you for the field, in the layer or table you are joining to, that the join is based on. Usually this "ID" field identifies individual points, lines, or polygons. What is important to keep in mind here is that the field on which you base the join needs to have the same kind on information, text, numbers, characters, and so forth. However, the field names are not required to be the same. A dropdown list of all the layers attached to the map file is next. You select the one you want to join to the base layer/table, or navigate to another table using the folder icon on the right. Next, choose the field to match to the field in the base layer. This field needs to share information types and entries to make it useful.

Finally, choose to keep all the records after the join or to remove the records that do not have values that match between the joining fields. The second choice will exclude the fields that have no representation in the base layer's table.

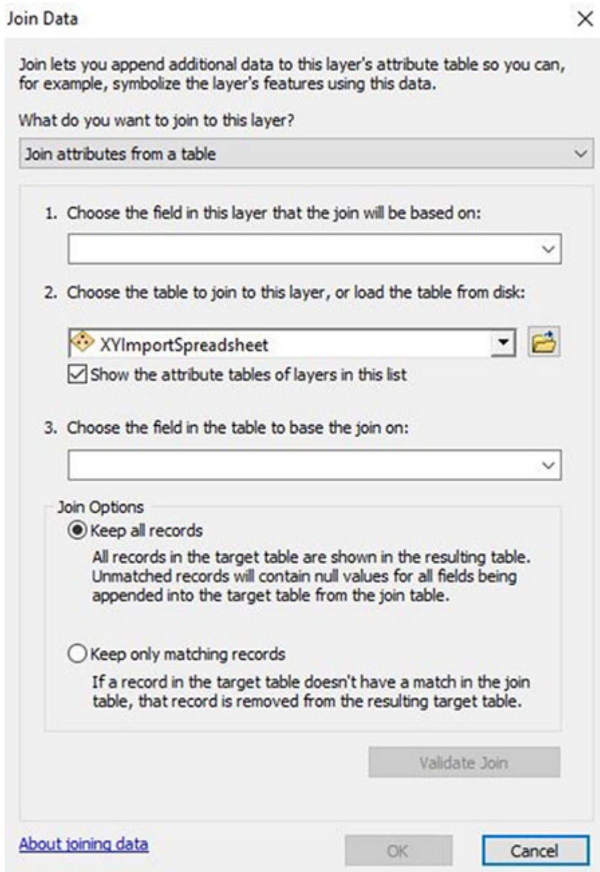

Figure 3–17. Join data window.

Relates

A relate is a two-way relationship between tables. When we relate tables, they refer to each other's values without appending data to the tables. That is the biggest difference between a join and relate. Joins append, or add as an attachment or supplement, one table to another by a common field. Relates use a common field to refer to values within the related table without attaching anything to the attribute table. The effect is that when searching values in a table that relates to another, every value that meets the search criteria displays. In a join, only the first value found displays.

Like joins, the best way to understand relates is by example. Let us say that we have a "Building" layer that joins to a table listing the names of each building occupants. With a join, a search only displays the first occupant in the table. When data is related searching for a building, it yields a list of occupants and searching for an occupant reveals the building they occupy.

Creating a relationship between data tables is similar to creating a join. First, access the context menu by right clicking on the layer you want to base the relationship on, then select "Relate…" This opens the "Relate" window (fig. 3–18). The four pieces of information in the window are identical to a join, making the procedure for completing a relationship the same. The key thing to understand is the difference between a join and a relate. There are four different types of joins or relates that we create based upon the association between the data—one-to-one, many-to-one, one-to-many, and many-to-many.

One-to-one

One-to-one associations are best for joins. They are literally one field value in the first table to one field value in the second table. For example, we have a table that lists the names of the counties as part of a polygon feature class; then, we get a table from another organization that provides the populations change over the past year. Creating an ID field on both tables and joining the tables using that field creates a one-to-one relationship because there is only one county name per population change record.

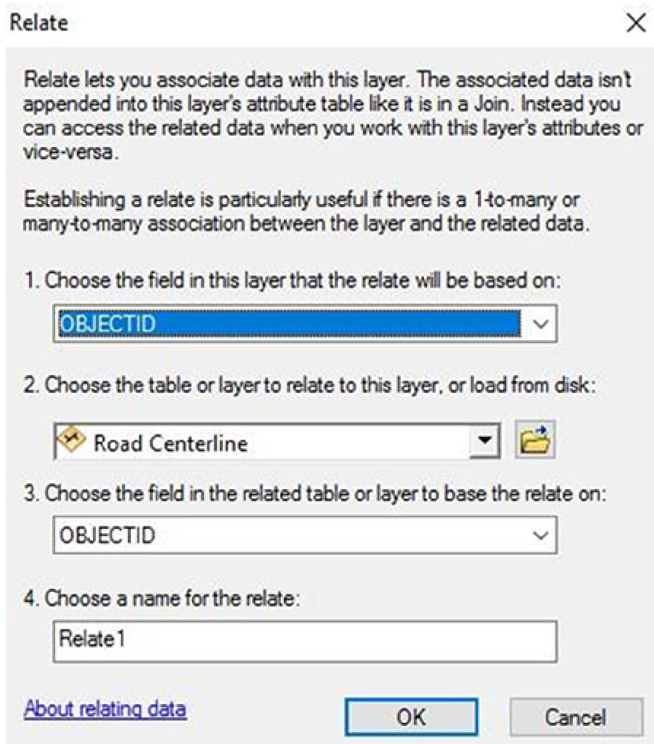

Figure 3–18. Relate data window.

Many-to-one

Many-to-one associations are also best when creating a join. Most commonly, we create many-to-one associations when a field describes a type for the associated feature, and we join a table that describes the type. For example, in our "Building" feature class, we have a field that stores the squadron that owns the building by number (364 TRS, 320 AMXS, etc.). If we join this field with a table that has a description of that number (364 Training Squadron, 320 Aircraft Maintenance Squadron, etc.), we can display the description as a more user-friendly map without changing the feature class fields.

One-to-many and many-to-many

The other two types of association occur whenever there is not a one-to-one or many-to-one association. Relates better utilize these kinds of associations. This is because relates work in both directions reading information from one table to another in both directions. This means that a one-to-many is also a many-to-one depending from which table you are reading.

217. Conducting spatial and attribute analysis

Storing information in a database and displaying that information graphically is only a part of the usefulness of a GIS. The real utility of a GIS comes from its ability to analyze relationships between different elements in the database. There are two primary ways that we conduct analysis—spatial relationship and attribute table information. Spatial analysis covers the relationship between graphic elements and their geometry. Attribute analysis identifies similarities or differences by reading the unique characteristics within their attribute table.

Spatial analysis

There are three types of spatial analyses: overlay, proximity, and network. This lesson describes each and summarizes their processes.

Overlay analysis

A basic question asked in GISs is “what is on top of what?” This could mean, “Does the Air Force base installation boundary cross into the limits of the neighboring town?” or, “Are these buildings within an environmentally protected area?” By overlaying two sets of data, we can analyze them to see where the areas are the same and where they differ.

The layer created by this overlay is comprised of data either intersected or unioned from the input layers. In performing both intersect and union overlay analyses, two or more layers are required. For instance, the land use layer will provide you the acreage for each land use such as residential, commercial, and so forth. The animal habitat layer will provide you the acreage for each habitat range. Combining these will allow you to determine how the habitats relate to the land use zones.

To perform these kinds of analyses, we use functions found in the toolbox application. We open the toolbox application within either a map or catalog session. If you perform this analysis from map, the output becomes an independent layer in your data frame. Performing the analysis from catalog creates a separate layer file. Within toolbox, expand the toolset for “Analysis Tools” and then expand the toolset for “Overlay.”

Intersect overlay analysis

Double clicking the tool for intersect opens the dialog box (fig. 3–19). The first option is to choose “Input Features” referring to the layers we intend to compare. If using in catalog mode, you will need to navigate to the layer by choosing the browse folder. In map mode, you can simply use the drop-down menu if the desired layers are within the map document. The next step is to verify that the correct layers were loaded in the window below. The delete button on the right removes the selected undesired layer. The arrows move the layers up and down the list, thus changing their priority. Next, the “Output Feature Class” section lets us choose the location of the output of the resulting combination of layers. If shapefiles are used, it will locate the output in the same folder where the first shapefile resides. The optional “Join Attributes” drop down allows you to choose whether to keep all attributes, only the feature identifier, or all attributes except the feature identifier. The XY tolerance is the minimum distance separating the features coordinates. Adjust this if some items that should be together are not due to decreased accuracy. Finally, the output type will be either input (meaning the same geometry as the input layers), line, or point. The point output could be valuable if you want to know the point location in which separate layers cross. Selecting “OK” completes the analysis process.

Union overlay analysis

Double clicking the union tool opens a dialog box (fig. 3–20). The first option is to choose “Input Features” referring to individual layers. If using in catalog mode, you will need to navigate to the layer by choosing the browse folder. In map mode, you can simply use the drop down if the desired layers to be unioned are both within the map document. Next, verify that the correct layers were loaded in the window below. Next, the output feature class section lets us choose the location to output the resulting combination of layers. If shapefiles are used, it will locate the output in the same folder where the first shapefile resides. The optional “Join Attributes” drop down allows you to choose whether to keep all attributes, only the feature identifier or all attributes except the feature identifier. Finally, selecting “OK” completes the analysis process. The analysis saves features common to all layers to the output file. Some of the features split to reflect separation caused by joining another feature class.

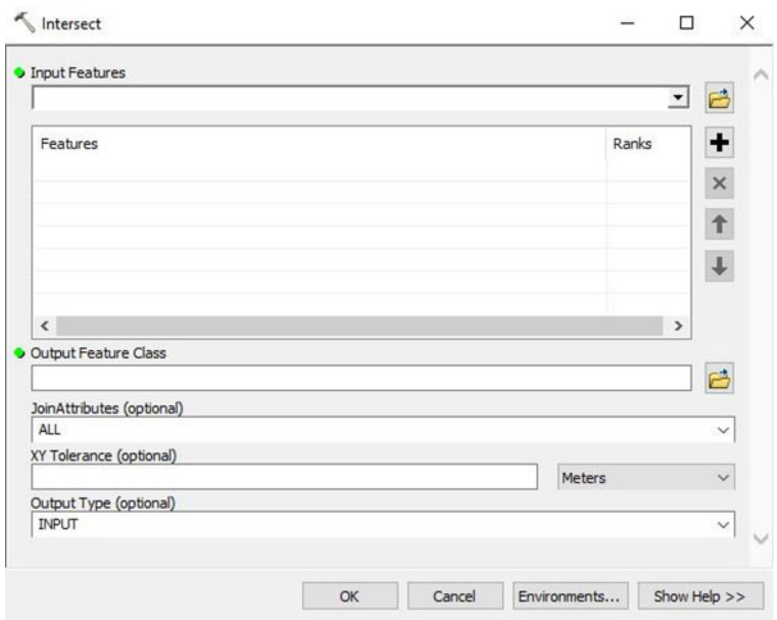

Figure 3–19. Intersect overlay.

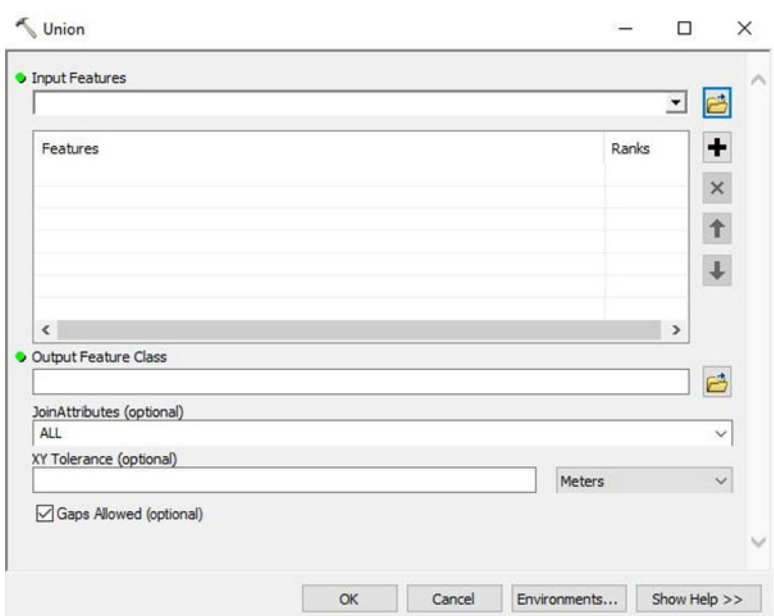

Figure 3–20. Union overlay.

Proximity analysis

A proximity analysis selects features within a given distance of a selected feature. It does this by either selecting them for future analysis or by creating a cordon. A cordon is an imaginary boundary around a feature typically for safety purposes. Think about explosive ordnance disposal (EOD) requiring 2,000-foot cordon around an unexploded ordnance (UXO). In GISs, we call cordons "buffers." Buffers work in two-dimensional space, so the width will be the same regardless of the coordinate system. It will not reflect the curvature of the earth. Generally, this is not a concern to us since we typically are working with areas small enough to negate this. However, be aware of the potential for errors when you are dealing with large buffers.

Buffer types

There are three types of buffers—point, true shape, and dissolved (fig. 3–21), each with advantages and disadvantages.

Buffer type	Description
Point buffer	This type of buffer is a circle. The application calculates the distance for the buffer from a single point.
True shape buffer	This type of buffer calculates the required distance around a feature by drawing some distance away from every point on the polygon's perimeter. True shape buffers can be both to the inside and outside of a polygon feature.
Dissolved buffer	If you are creating buffers on multiple objects, you can merge the buffers so only the outer most distance displays. The tool generates each feature buffer and then erases the interior lines, leaving just the exterior boundary.

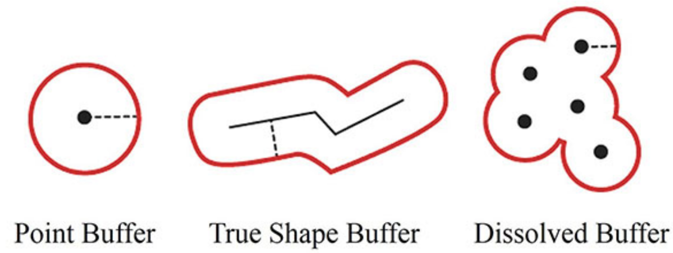

Figure 3–21. Buffer types.

There are two buffer tools that allow you to process the information in a spatial analysis—"Map Buffer" and "Buffer Wizard."

Map buffer

Using the map application, the map buffer tool resides on editor toolbar. Begin by starting an edit session. Then, using the edit tool, select the feature around which you want to create a buffer. Select the editor toolbar menu drop down and choose "Buffer." You will receive a prompt for the buffer's distance. The prompt takes distance in the same units as the map. Type in the distance followed by the "Enter" key and the buffer appears on a new layer under target layer in the TOC.

Buffer wizard

Using the map application, the buffer wizard allows more control over the buffer process. Access this tool by right clicking on any toolbar and selecting "Customize." Select the "Commands" tab and either scroll down the "Categories" list and select the "Tools" category where buffer wizard is an option or type "Buffer Wizard" where it says, "Show commands containing (fig. 3–22)." Once you have found the tool, left click, hold, and drag the tool to an active toolbar. You can dock the tool on the existing toolbar or to a customized one you have created. A bolded uppercase "I" icon indicates the tool's placement. Select the buffer wizard tool and wait for the wizard to load.

The first screen allows you to choose what you want to buffer (fig. 3–23). The features of a layer allow you to choose an entire layer to buffer or features previously selected by checking the box "Use only the selected features." The "Next" button displays the next screen of the wizard.

This screen starts by asking, "How do you want to create a buffer?" Specify buffers by a set distance by typing the distance value and choosing the distance units at the bottom of the screen (fig. 3–24). In addition, you can create buffers defined by an attribute value within a field. A point to make about a buffer defined by attribute, keep an eye on the value of the attribute, changing the attribute does not change the buffer. The buffer is not dynamic. Finally, the buffer wizard has options for creating multiple buffer rings, which is its largest advantage over the standard buffer tool, which can only create a single ring. Simply choose how many rings and the radial distance for each ring. This creates concentric buffer rings. Choosing "Next" takes you to the "Buffer Output" options.

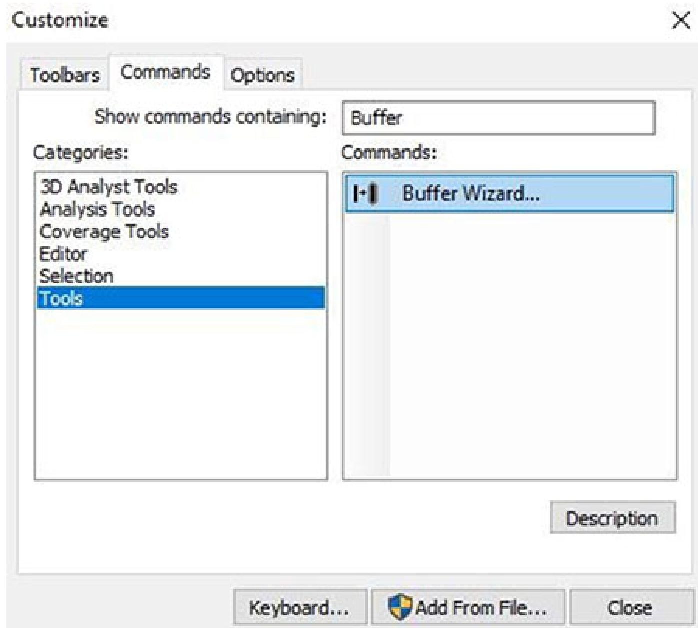

Figure 3–22. Buffer wizard icon.

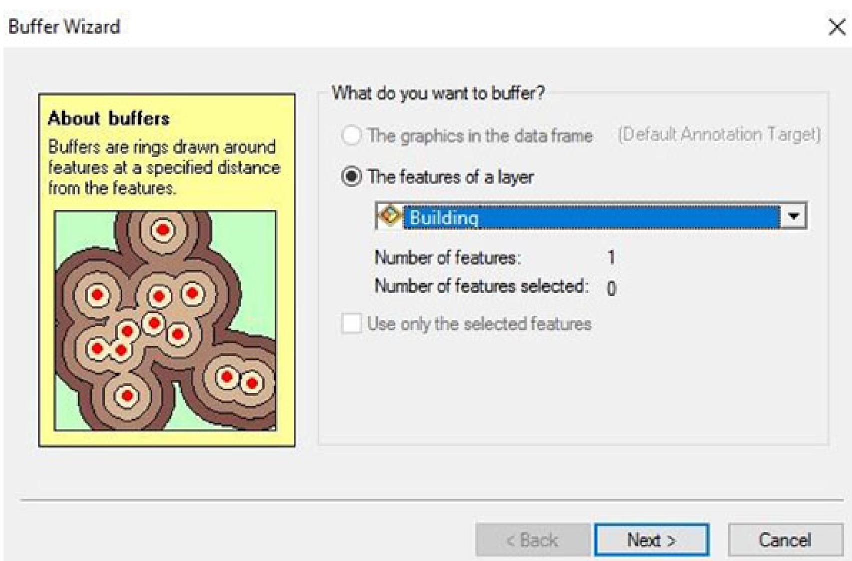

Figure 3–23. Buffer wizard page one.

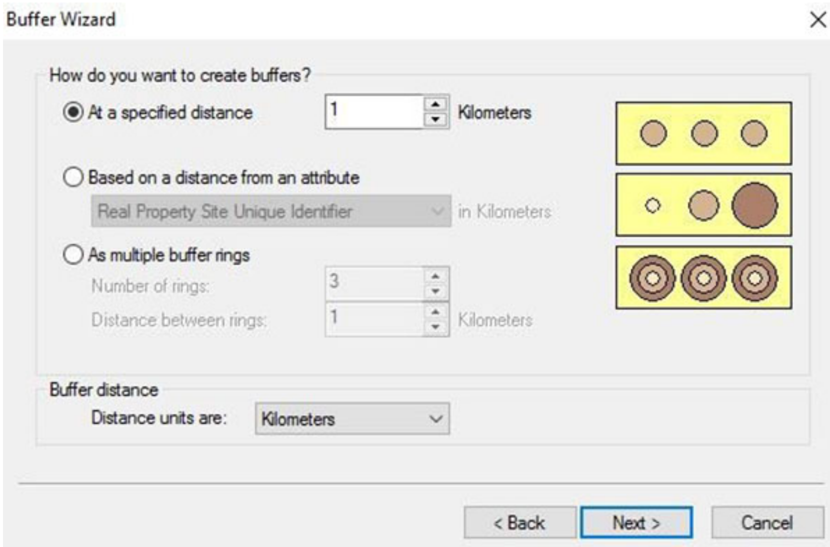

Figure 3–24. Buffer wizard page two.

Start by either dissolving the buffer as we mentioned earlier or not dissolving the buffer. You will not be able to separate dissolved buffers later. The “Create buffers so they are” option lets you choose where to create buffers in regards to the features. The first option of “Inside and outside the polygon(s)” creates the buffer both outward from the feature perimeter and inward. The second option buffers only outward from the feature, and the third option buffers only inward from the feature perimeter. The last option buffers outward and completely encompasses the inside of the feature.

Finally, save buffers in one of three ways (fig. 3–25). The first way saves the buffer as a graphic layer with the map document only. The second option allows you to create the buffers on an existing editable layer. To make this happen, an editing session must be started, and the layer to assign the buffers to must be set as the target layer prior to the buffer wizard being started. The last option allows you to save the buffers as either a shapefile or a feature class. Remember the fields from the buffer are specific to the GIS not the SDSFIE.

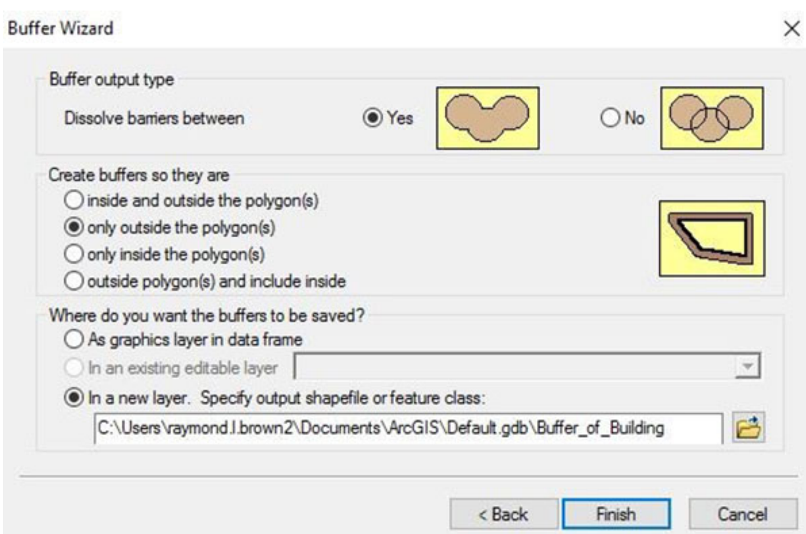

Figure 3–25. Buffer save options.

Network analysis

You may have heard this referred to as least cost routing. It is how to get from point A to point B. Analysis is possible on variety of data points such as the shortest physical route, the quickest (time) route, or other special attributes. It will also provide some analysis of how to go from point A to point C if intersection B is closed. Network analysis requires you to generate a geometric network.

Attribute analysis

Statistics are easy to obtain; however, making sense out of them may become an issue. To perform a statistical attribute analysis, your data must contain a numerical field. Add the layer to your project and open the attribute table by selecting the layer, right clicking, and then clicking "Open Attribute Table."

Once the attribute table opens, select the field you want statistical information about. Right click the field (column) name, and click "Statistics." The statistics dialog opens and is titled statistics of "the layer that was selected (fig. 3–26)." The table below lists output from a statistical analysis:

Field	Output
Count	How many records the analysis on which we performed analysis.
Minimum	Lowest number value encountered for this field.
Maximum	Highest number encountered for this field.
Sum	Total value of all numbers added together for this field.
Mean	Average value for all numbers in this field.
Standard deviation	The square root of the variance. It describes the spread of the data about the mean in the same units as the original measurements.
Frequency distribution	Chart displaying value dispersal.

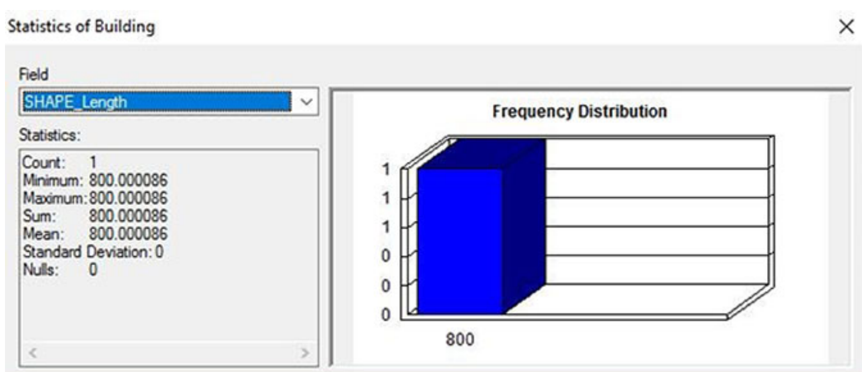

Figure 3–26. Statistics window.

Summarize

By summarizing the data in a table, you can derive various summary statistics—including the count, average, minimum, and maximum value—and get exactly the information you want. If you create a new table containing the summary statistics, you can then join this table to the attribute table of a layer. Doing so lets you symbolize, label, or query the layer's features based on their values for the summary statistics.

The process of attribute summarizing is a little more complex. To access the summary dialog, select the desired layer then open the layer attribute table. Select the desired field for summary, right click, and then click "Summarize."

Geocoding
Geocoding is the process of transforming a description of a location such as a pair of coordinates, an address, or a name of a place to a location on the earth's surface. You can geocode by entering one location description at a time or by inputting many of them at once in a table. The resulting locations are output as geographic features with attributes, which we use for mapping or spatial analysis.

Geostatistics
Geostatistics is a point-pattern analysis of raster data to produce field predictions. It is different from general applications of statistics because it employs graph theories and matrix algebra. Geostatistics create digital elevation models (DEM) and triangulated irregular networks (TIN).

Temporal analysis
A temporal analysis contains tracking features that either change or move over time. We accomplish this with fix-timed data stored with your feature data or it can be done in real-time with the use of a streaming tracking service or Global Positioning System (GPS) information.

Self-Test Questions

After you complete these questions, you may check your answers at the end of the unit.

215. Merge data
1. When we talk about merging data, what do we mean to do?

2. Briefly describe what the Field Map options give us an opportunity to do.

216. Joining and relating data
1. What is the difference between a join and relate?

2. What two association types are best when creating a join?

217. Conducting spatial and attribute analysis
1. List the three types of spatial analysis.

2. Describe the difference between a point buffer and a true shape buffer.

Answers to Self-Test Questions

212
1. Which image to georeference in the dropdown menu.

2. Distorting.

213

1. Who owns or collects the data, why you need it, and for what you use it.
2. Selecting edit when we start and finish editing.

214

1. Groups of features that share the same attributes, or more commonly share a single attribute.
2. (1) Set Subtype Field.
   (2) Add Subtype.
   (3) Set Default Subtype.
   (4) Remove Subtype.
3. Subtype Code and Subtype Name.
4. A value that auto-populates each time a new item is added to the feature class.

215

1. Taking two or more sets of data of the same type and create a new set of data.
2. Remove, add, and modify fields in the output dataset.

216

1. A join appends data to a table while relate creates a reference for each table.
2. One-to-one and one-to-many.

217

1. Overlay, proximity, and network.
2. A point buffer creates a circle a specified distance from a single point while a true shape buffer creates a perimeter line a specified distance from every point on a line or polygon.

Unit Review Exercises

Note to Student: Consider all choices carefully, select the *best* answer to each question, and *circle* the corresponding letter. When you have completed all unit review exercises, transfer your answers to the Field-Scoring Answer Sheet.

Do not return your answer sheet to the Air Force Career Development Academy (AFCDA).

30. (212) To save a georeferenced image as a new file, we select “Georeferencing” on the tool bar then
    a. Save.
    b. Rectify.
    c. Stop Editing.
    d. Update Georeferencing.

31. (213) What tab in Catalog displays the metadata?
    a. Preview.
    b. Contents.
    c. Properties.
    d. Description.

32. (213) Under which menu and tab in Catalog do we find the “Metadata Style” setting?
    a. File, Data.
    b. Edit, Options.
    c. Customize, ArcCatalog Options.
    d. Windows, View Metadata Styles.

33. (214) In which geographic information system program are subtype functions located?
    a. Map.
    b. Office.
    c. Catalog.
    d. Toolbox.

34. (215) In Toolbox, under which list do we find the Merge tool?
    a. Analysis.
    b. Data Management.
    c. Analyze Tools for Pro.
    d. Distributed Geodatabase.

35. (216) A two-way relationship between data tables is a
    a. join.
    b. relate.
    c. union.
    d. representation.

36. (217) What kind of analysis creates a layer comprised of intersected or united data?
    a. Overlay.
    b. Network.
    c. Attribute.
    d. Proximity.

37. (217) Which tool allows the most control over the buffer process?
    a. Map Buffer.
    b. Analysis Tool.
    c. Buffer Wizard.
    d. Proximity Manager.

38. (217) We create buffers based on a specified distance, an attribute distance, or as
    a. averages.
    b. multiple buffer rings.
    c. multiples of map units.
    d. a set of rings based upon the map scale.

39. (217) Which analysis do we also refer to as "least cost routing?"
    a. Spatial.
    b. Integer.
    c. Network.
    d. Attribute.

40. (217) When performing a statistical analysis, which of these is NOT an output field?
    a. Sum.
    b. Count.
    c. Integral deviation.
    d. Frequency distribution.

Student Notes

Glossary

Abbreviations and Acronyms

AFI	Air Force instruction
CAD	computer-aided drafting
CDC	career development course
CE	civil engineer
CIP	common installation picture
DEM	digital elevation model
DPI	dots per inch
EOD	explosive ordnance disposal
GIO	geographic information officer
GIS	geographic information system
GPS	Global Positioning System
GSU	geographically separated unit
IDP	installation development plan
km	kilometer
MAJCOM	major command
MDL	mission data layer
MDS	mission data set
mph	miles per hour
MrSID	multiresolution seamless image
NAD83	North American Datum 1983
NAVAID	navigational aid
OSD	Office of the Secretary of Defense
PCS	permanent change of station
PDF	portable document format
QA	quality assurance
QC	quality control
SDSFIE	Spatial Data Standards for Facilities, Infrastructure, and Environment
TIN	triangulated irregular network
TOC	table of contents
UXO	unexploded ordnance
WGS84	World Geodetic System 1984
XML	extensible markup language

Student Notes

Student Notes

AFSC 3E551
3E551 02 1905
Edit Code 03