**GIS Notes**

Everything you experience from day to day happens somewhere in geographic space. As a result, you can represent your world and your experiences in it by using maps. You use those maps to find places, save time while traveling, decide where to locate a new store, plan cities, guide the development of wildlife preserves, and satisfy hundreds of other applications.

The computer systems that enable you to store and access all this information are collectively called geographic information systems (GIS).

Examples of people that may use GIS:

* Business owners and marketers
* Urban planners
* Merchandise distributors

Using GIS software, you can put maps and other geographic data into the computer. After you have the data in the computer, you can store, retrieve, and edit that data. You can analyze it (for example, find geographic features, measure distances, or compare patterns) and produce output from it (create new maps from what you find).

A GIS system is comprised of:

* Data
* Computers and software
* Geographic concepts that drive the analysis of data
* People that operate the GIS
* The organizations within which the GIS exists

Primary data are collected firsthand by you, for a particular project. Primary data are usually the best data for the job because you collect them with your specific goals in mind. Secondary data come from others who collect the data for unrelated tasks or gather it with remote sensors.

The process of entering data into a GIS system can be summed up as:

1. Define where, how, and what kind of data to sample
2. Collect that data directly or indirectly
3. Use the software to transform that data

You may need to change some data from hard copy to digital forms; you may need to convert some from uncategorized to categorized data (for example, aerial photo interpretation); and you may need to attach coordinates to digital data so that you can find them in your digital maps.

Hardware used to collect GIS data includes:

* Devices to collect information
* Devices to enter information
* Storage and analysis devices
* Output devices

Grid cells: one method of storing data in squared boxes, that may be utilized by a GIS system.

Organizations that use GIS work best when the organization adapts itself to the technology. If GIS helps the organization perform its tasks, if the employees are adapting to and benefitting from the changes, if the organization provides training, and if GIS enhances the organization's overall goals, that organization can likely incorporate GIS successfully, long-term.

A system designer reviews an organization's structure, products, workflow, and needs. He or she then determines the costs and benefits of GIS for that organization, as well as how the organization might best include GIS in critical operations.

Geographers know that all things are related in geographic space, but close things are more related than far things. This statement describes one aspect of geographic space — closeness — that makes space so important to you as a geographic decision-maker. Listed below are some terms that also relate to geographical space as it is used in GIS.

Density: refers to the measure of how concentrated or dispersed a phenomenon is within a specific geographic area.

Sinuosity: refers to the degree of curvature or winding of a linear feature, such as a river, road, or coastline. Sinuosity is a measure of how much a linear feature deviates from a straight line.

Connectivity: refers to the degree to which geographic features or locations are linked or able to interact with each other. Connectivity can be evaluated in terms of physical connections, such as roads or transportation networks, as well as conceptual connections, such as social or economic linkages between places.

Pattern Change: refers to the observed modifications or transformations in the spatial distribution, arrangement, or characteristics of geographic features over time

Movement: refers to the study and analysis of the spatial trajectories, patterns, and dynamics of objects, people, or phenomena as they change location over time.

Shape: refers to the geometric properties and configuration of geographic features and objects.

Size: refers to the physical dimensions and measurement of geographic features and objects.

Isolation: refers to the degree to which a geographic feature or location is physically or functionally separated from its surrounding environment or other features.

Adjacency: refers to the spatial relationship between geographic features or locations that are next to or bordering one another.

Geographic data come in four basic forms: points, lines, polygons (or areas), and surfaces. A fifth form, related to surfaces, is volumes.

* Points: refer to the simplest form of geographic feature, represented by a single x,y coordinate location on a map or in a spatial dataset.
* Lines: refer to linear geographic features represented by a series of connected x,y coordinate points.
* Polygons: refer to geographic features that are represented by enclosed two-dimensional areas defined by a series of connected x,y coordinate points.
* Surfaces: refer to continuous geographic features or phenomena that can be represented and analyzed as three-dimensional spatial data.
* Volumes: refer to the three-dimensional representation and analysis of geographic features and phenomena that have measurable depth or thickness, not just length and width.

Most GIS systems contain database tables that allow you to store all sorts of descriptive information about the points, lines, areas, and surfaces that you're depicting in your GIS. The nature of database tables requires you to be just as picky about assigning descriptive information to your objects as you are about choosing the right graphics to depict the objects themselves.

Nominal information: Geographic features that have names only. So, you can't compare their descriptive information to any other.

Ordinal information: Geographic features that you can compare by rank. You could have short, medium, and tall trees; dirt roads, paved roads, highways, and superhighways; or large, medium, and small chemical spills.

Interval information: Geographic features that have detailed increments (intervals) that you can measure. One limiting characteristic of interval data is that, although you can get very accurate measurements, you can't form ratios because the starting point is arbitrary.

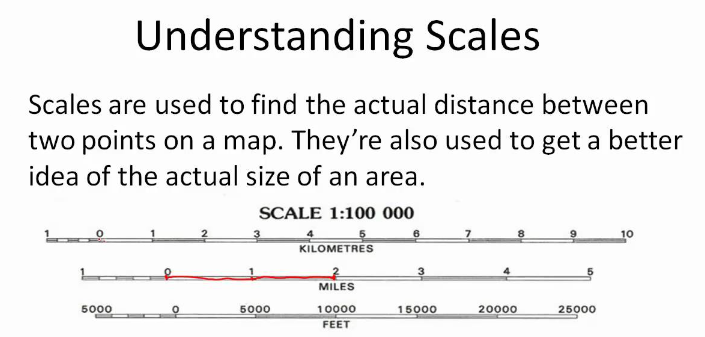
Ratio information: Geographic data that have measurable units, like interval data, but also allow you to make the ratio comparisons that interval data won't. If you own a parcel of land that's worth $20,000 and your neighbor has a parcel worth $10,000, then your parcel isn't just worth $10,000 more, it's also twice as expensive. The key point is that ratio data have an absolute 0.

Scalar information: data are a bit difficult to define, but here's my take: Scalar data have a proprietary measurement system. That is, you create your own scale that applies to only a particular set of data. So, if you're ranking the beauty of a scenic overlook on a scale of 1 to 10, you first have to decide what each number in the scale means. GIS allows you to establish a scalar description for features that you can't really measure any other way.

Map Extent: refers to the geographic area or region that is visible and displayed within the current map view or window.

Map scale refers to the relationship between the distance on a map and the corresponding distance on the earth's surface. Specifically, map scale is the ratio or proportion that expresses how much the map has been reduced or enlarged compared to the actual geographic space it represents.

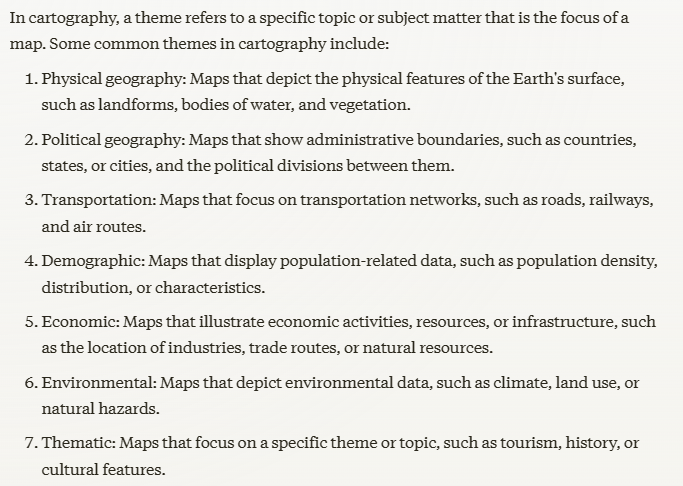
You can often find a map's scale represented by a graphic bar and a fraction that shows the relationship between the size of the map in the numerator (the fraction's top part) and the size of the Earth in the denominator (the fraction's bottom part). Using this mathematical approach, the smaller the fraction (one with a small numerator and a large denominator), the smaller the scale.



Cartographers use symbols that represent point features (such as towns), symbols that represent linear features (such as roads and rivers), and symbols that represent area features (such as lakes and towns).

Cartographers have to carefully consider certain things when they create a map:

* Scale: Determines how many geographic features can be symbolized on a map.
* Data availability: Determines what type of information can be put on the map.
* Limitations of output devices: The cartographer also has to consider how symbols will print.
* Reader characteristics: Not all readers have 20/20 vision, color vision, or any vision at all.



Reference maps offer a great deal of information on a single document. Atlases generally contain reference maps so that many related maps can be contained in the same place. Reference maps often cover very large portions of the Earth.

A thematic map provides as much accurate, detailed information as possible about a particular subject, such as roads or hills, as compared to a reference map, which tries to select the most important information about several subjects.

A road map is one example of a thematic map because it focuses on communicating information about roads.

Thematic maps are the primary kind of maps that you use in your GIS activities.

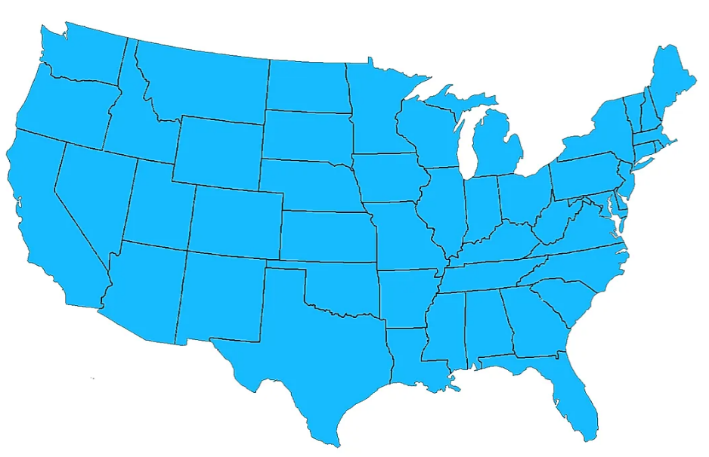
A good rule of thumb is that the larger the map scale, the smaller the area covered and the greater the detail. Larger scale maps are generally better for your GIS activities because they provide the largest amount of detail.

The “weakest link” hypothesis: the success or failure of a complex production process or system depends on the performance of the weakest component or link in the chain. This is mentioned in GIS as a way to remind the GIS operator that quality of a geographic product depends on it’s worst/lowest quality input.

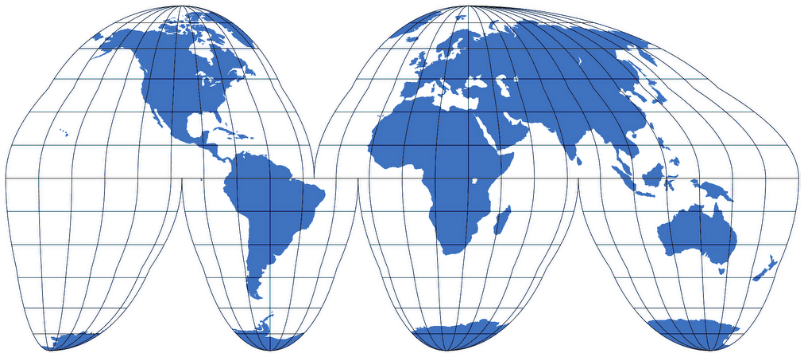
The earth’s spherical shape has some major drawbacks for the mapmaker who's faced with producing a flat map that correctly represents the shapes, angles, distances, and sizes of objects on the Earth. It is expected that some distortion will occur when translating a round object to a flat object.

Map projections —the process of converting the spherical Earth to a flat surface — come in many different types, from contiguous to interrupted, from those that look like photographs of the Earth to those placed on cones or cylinders.

Contiguous map: a type of map in which all the geographic areas or units depicted are connected to each other, with no gaps or disconnected regions. Example shown below:

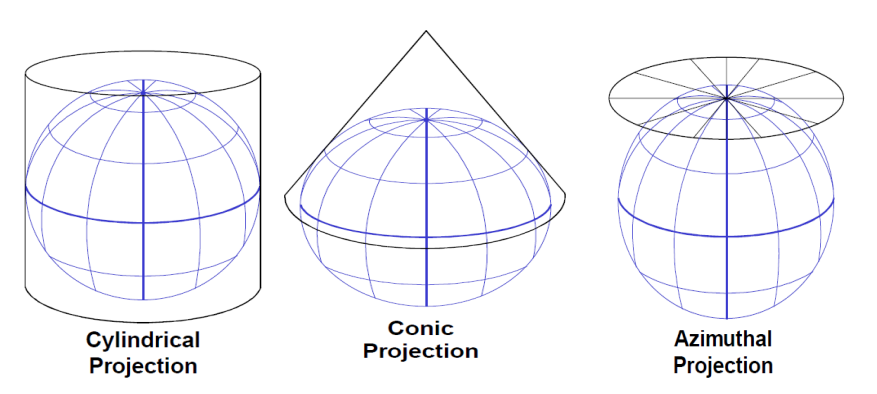


Interrupted map: a type of cartographic projection that intentionally breaks or interrupts the continuity of the map in order to better display certain geographical features or regions. Example shown below:



Another way to describe map projections is:

* Planar or Azimuthal
* Conical
* Cylindrical



When working with GIS, pick the map projection that best represents the properties you want preserved when you create output maps from your analysis. Most high-end GIS software has the capability to convert back and forth from one projection to another. Most have more map projections than you'll ever use.

Having an accurate representation of distance and area measurements in any projected map depends on having accurate measurements of the spherical Earth. The science of geodesy deals specifically with these measurements.

Geodesy is the scientific discipline focused on the measurement and representation of the Earth, including its gravitational field and geometric shape. It encompasses a range of activities and techniques related to accurately determining the size, shape, and position of the Earth.

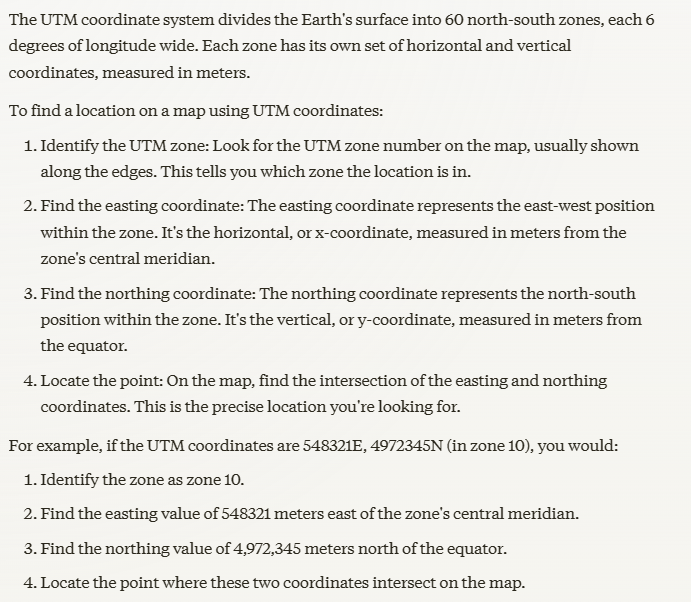
A datum is a set of parameters and control points that define the size, shape, and orientation of the Earth's surface within a particular coordinate system. It provides a frame of reference for accurately locating and positioning geographic features and coordinates.

Your GIS software needs to know what datum you're using for each set of map data that you put into your database. Attaching your coordinates to the wrong datum can result in location and measurement errors.

When loading up your GIS, be sure to use the correct datum for each map source as you add it. Also, convert all your map data to a common datum when you work with more than one source map at a time.

Coordinate systems in GIS (Geographic Information Systems) are reference frameworks used to define and represent the location of geographic features and data on a map or within a digital mapping environment.

One of the many possible coordinate systems that you may encounter in GIS is called the UTM system. UTM stands for Universal Transverse Mercator, which is the most commonly used system. This system divides the Earth from latitude 84° north and 80° south into 60 numbered vertical zones, each 6 degrees of longitude wide.



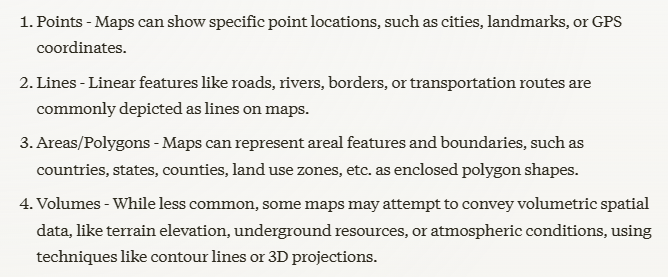
The PLSS, or Public Land Survey System, is a method used in the United States to survey and identify land parcels. Here are some key points about the PLSS:

* Origins: The PLSS was established in the Land Ordinance of 1785 and later refined in the General Land Office Survey Act of 1796. It was created as a way to systematically survey and divide up the public lands in the western United States after the American Revolution.
* Grid System: The PLSS divides land into a grid system based on principal meridians and base lines. The basic unit is the township, which is a 6-mile by 6-mile square. Each township is further divided into 36 one-mile square sections.
* Surveying Methodology: PLSS surveys are conducted using a rectangular system of section lines, rather than following natural features of the land. Surveyors use instruments like theodolites and chains to establish the grid.
* Land Ownership: The PLSS was used to measure and distribute public lands to private owners through mechanisms like the Homestead Act. It also laid the groundwork for the current system of land titles and property ownership in much of the western U.S.
* Significance: The PLSS has been critical for the orderly settlement and development of the American West. It provides a consistent, standardized way to identify and describe land parcels across large geographic areas.
* Continued Use: The PLSS is still in use today and forms the basis for all official land surveys and property descriptions in the 30 states where it was originally implemented.

The baseline is a critical east-west line that serves as the starting point and reference for the PLSS grid system. Surveyors establish the baseline, often along a prominent geographic feature. All township lines run parallel to the baseline, forming a grid of 6-mile by 6-mile townships. The townships are then numbered north and south from this fixed baseline. Baselines provide the foundational reference points that allow for the systematic division and description of land parcels across an entire state or region using the PLSS township and range system.

A principal meridian is a north-south line that, along with a baseline, forms the origin and framework for the Public Land Survey System (PLSS) in the United States.

Symbols are lines, objects, or pictures on the map that represent real objects on the ground.



Those are important concepts related to the measurement scales used in spatial data and maps. Here's a brief explanation of each:

1. Nominals:

* Nominal scale is a basic level of measurement that assigns unique labels or categories to objects or places.
* Nominal data has no inherent order or numerical value, it simply classifies things into discrete groups.
* Examples on a map could include different land use types (residential, commercial, industrial) or political boundaries (countries, states, counties).

1. Ordinals:

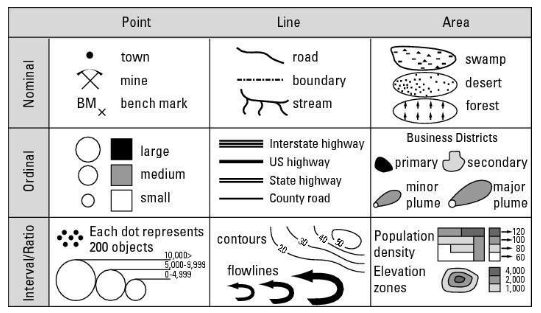
* Ordinal scale indicates a ranked order or hierarchy between different values.
* Ordinal data can be sorted or sequenced, but the differences between values may not be uniform.
* Examples on a map could include elevation levels shown by contour lines or socioeconomic status represented by income quintiles.

1. Intervals:

* Interval scale data has a consistent, measurable distance between values.
* Interval data has a meaningful zero point and the differences between values are equal.
* Examples on a map could include temperature, precipitation, or population density shown using a quantitative color scale.

In GIS, a ratio scale has a meaningful zero point and equal intervals between values, allowing for measurement of absolute magnitudes. Examples include population density, income, and elevation. Ratio data enables powerful quantitative spatial analysis and visualization techniques in GIS.

Geographic features and how they are measured:



In GIS (Geographic Information System), non-comparative data refers to a type of data that represents a single point in time or a snapshot of a particular phenomenon. This type of data is not intended to be compared or analyzed in relation to other data, but rather to provide a standalone representation of a geographic feature or attribute.

Non-comparative data is often used to:

* Describe a single event or occurrence
* Represent a static condition or state
* Provide a baseline or reference point for future comparison

Line features: Streets, rail lines, and boundaries, for example, are unique entities and can't be compared to one another.

Areas: A swamp, the range of a wild species, the land owned by the federal or state government, or the type of zoning for a particular parcel of land.

Volume features: Water aquifers, hills, and buried ore bodies all take up volume and are named on maps.

In GIS (Geographic Information System), comparisons of kind refer to a type of analysis that involves comparing different types of data or phenomena that share a common attribute or characteristic. This type of comparison is used to identify similarities or differences between different datasets or features that are fundamentally distinct, but share a common theme or category.

Graduated symbols are a type of symbol used in mapping and Geographic Information Systems (GIS):

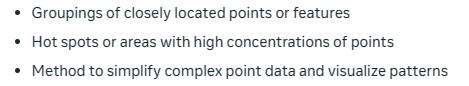
* Symbol size: Graduated symbols vary in size to represent quantitative differences between mapped features.
* Data classification: Data is grouped into classes, with each class assigned a symbol size to represent the range.
* Color: The color of the symbols remains the same, while the size changes to represent different classes.
* Used for: Graduated symbols are used to show differences in magnitude or quantity, such as population density or earthquake magnitude.

In GIS (Geographic Information System), a class refers to a category or group of features or data that share similar characteristics or attributes.

One cool thing about maps is that the symbols represent a scaled-down version of real geography. As much as possible, the map's symbolic objects, features, and background are distributed and located in ways that closely resemble the locations and distributions of real objects.

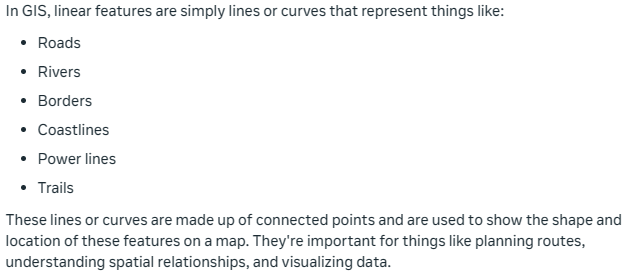
When you identify patterns, you look for a degree of predictability in the arrangement of the objects you're interested in. To make the best use of GIS information, the trick is to notice patterns that you may not be used to seeing, such as patterns of trees, houses, roads, rivers, or any other features you encounter.

Clusters help identify and visualize spatial patterns and concentrations.



Hotspots: refers to clusters of criminal activity in an area.

Uniform distribution in GIS is a probability distribution where all values within a specified interval have the same probability. A good example might be an orchard, where trees are planted in lines that are even to one another.



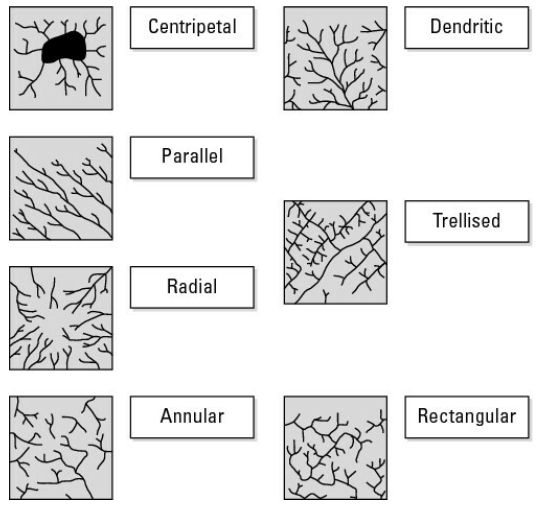
Connectivity: Refers to the physical communication links between devices and the ability of plants and animals to move freely between different wildland areas.

Some road networks and highways allow you to go around obstructions, rather than just through them. These networks form closed loops, or circuits, just like in an electrical circuit. When road networks have circuits, traffic has an alternative because it can flow around obstacles that drivers might want to avoid.

In GIS, linkages are connections between different devices, systems or areas. Linkage Mapper is a GIS tool designed to support regional wildlife habitat connectivity analyses. Linkage Mapper uses GIS maps of core habitat areas and resistances to identify and map linkages between core areas.

Some stream pattern types include:

* Dendritic: The easiest one for me to remember is called dendritic and looks like the branches of a tree. These branches (called tributaries) go out in all directions and seem to have a mind of their own.
* Radial: Another common pattern that streams take is called radial. A radial pattern looks like a dendritic pattern, except that all the streams flow outward, away from a center, like the spokes of a wheel.
* Centripetal: The opposite of radial patterns, centripetal patterns occur when a low spot or depression affects the flow.
* Parallel and sub-parallel streams run along the gentle slopes that result from either natural topography or from land manipulation because of road or mining construction activities.
* Trellis: Trellis stream patterns result from rock strata that's jointed, exposed, and folded from geological forces acting over time. Trellis patterns resemble the street patterns in neighborhoods loosely organized along a grid.
* Rectangular: Rectangular stream patterns show strong right-angle turns and are often the result of cross-cutting joints in the underlying rock.
* Annular: Annular stream patterns occur in areas which have a dome that is eroded.



There are three things that a GIS operator must be able to do with patterns:

1. Accurately describe the patterns in a way that the layperson can understand.
2. Quantitatively compare and contrast your descriptions of the patterns to those of other features and feature sets.
3. Analyze the quantities that you identify to determine a measure of their pattern or shape, and to provide a baseline for changes that might take place over time.

Your GIS software knows which values you need for your analysis because when you put your spatial data into the computer, you tell the software exactly where each point is and how that point relates to the feature's real position on the surface of the Earth.

The nearest neighbor statistic in Geographic Information Systems (GIS) is a way to measure how objects, like points, are distributed across a space. It’s often referred to as the Nearest Neighbor Index (NNI). This statistic helps determine whether the distribution is clustered, random, or uniform:

* Clustered: Points are closer together than expected if randomly distributed.
* Random: Points are neither too close nor too far apart; their distribution is as likely as by chance.
* Uniform (or regular): Points are evenly spaced out, more so than in a random pattern.

The Nearest Neighbor Index is calculated by comparing the average distance from each point to its nearest neighbor with what would be the average distance in a perfectly random distribution. An NNI less than 1 indicates clustering, an NNI greater than 1 indicates a uniform distribution, and an NNI close to 1 suggests a random distribution. This statistic is particularly useful in spatial analysis to understand patterns and their implications in various fields such as ecology, urban planning, and public health.

You use nearest neighbor statistics to make sure that what your eye has already told you is real and not a figment of your imagination.

You can use certain techniques to determine the average direction of linear objects, such as tornado paths, fallen trees, boulder distributions, shelterbelts (protective barriers such as windbreaks), and many other objects. Without going into the trigonometry, it simply gives you a summary statement of the direction in which these events and objects occur. This calculation might, for example, tell you the direction of wind during storms, the movement of a glacier while it leaves debris behind, or the way that farmers line up planted trees to protect their crops from the wind.

After you become familiar with the idea of seeing and recognizing patterns for your own specialty, you next need to use your specialized knowledge to identify the reasons they exist. A process creates every pattern, and every pattern has an effect on the process that created it.

Follow these steps when analyzing GIS data:

1. Recognize and acknowledge the existence of the patterns.
2. Analyze and verify that the patterns are real.
3. Identify the causes and consequences of these patterns based on your knowledge of the study area.
4. Apply this knowledge in your own profession for prediction and planning.

GIS software allows you to create new scenarios and model the consequences of change to help you decide which of your scenarios works best. But the GIS can't tell you what the functional relationships are. It allows you to test different ideas and make pattern comparisons, but you, as the expert, have to decide which factors to examine and which patterns to compare.

GIS specialists can't always apply the software to effectively solve problems because they're often not subject matter experts. So, subject matter specialists need to communicate the possible causes and consequences of patterns to the GIS applications developer.

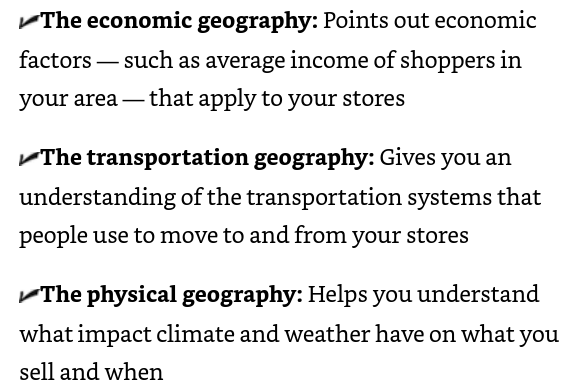
Maps are complex, and human map readers (like you) interpret much of the information that those maps contain. So, to create a GIS that has both complex and useful information, you have to show the computer how to think like you do as a map reader.

A good first step in deciding what to do in your GIS maps is to formulate a conceptual model, or a picture in your mind, of how you plan to tell the computer all the information you gather from a map.

When starting a Geographic Information Systems (GIS) project, the key is first to clearly understand the desired outcome. Begin by envisioning what the final product should look like. Next, break down this final product into the different types of maps or themes that will be needed. Finally, identify the specific elements of each map—like features and their related data—that are necessary for the project. These elements will then be combined to construct the final map. This structured approach ensures that the GIS project is both organized and purposeful.

Sometimes, composing a flowchart before starting is a good way to organize and gain a visual idea of what the GIS will need to look at to get good output.

A few different types of geography for an area include:



By starting at the final product, you can break the final map down into the individual data component that are required to construct it.

Most major GIS decision-making operations deal with the locations or distributions of features.

Outline details of your project and ask the specific questions that tell you exactly what data you need from the thematic maps you've already chosen.

Each digital thematic map is part of a larger group of digital maps that you use to answer your GIS question. Each map has features represented by points, lines, areas, and volumes that contain useful information for you. If you're just starting your current GIS decision-making process, you need to know what data you have in the GIS database in the first place.

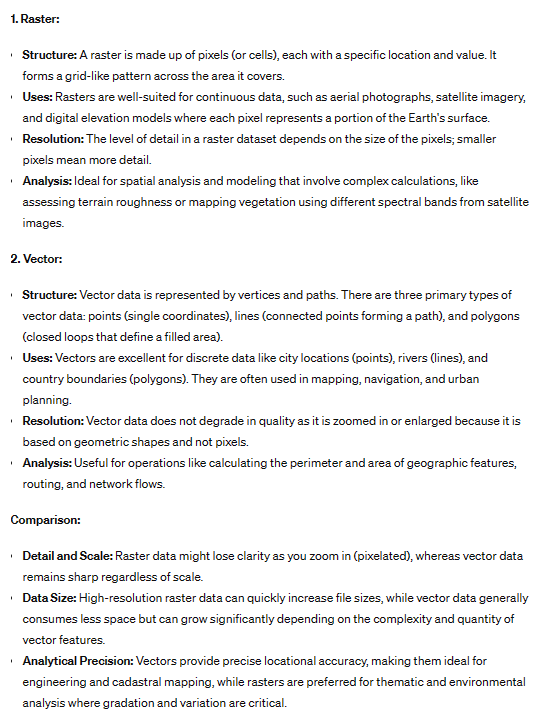
You will need to know characteristics of the data you’re working with. This includes:

* Data measurement- is it nominal, ordinal, interval, or ratio
* Level of detail- what scale is needed from the map to show all the features

A good point to keep in mind is that you want to start off with as much detailed data as necessary. It’s easier to generalize from a detailed dataset, than it is to get specifics from a generalized dataset.

In GIS terms, polygon refers to area, and surface means that a feature has a third dimension (for example, height).

Cartographers use two general data forms that translate the map into digital form: raster (little squares or grid cells) and vector (or points, lines, and polygons).



In a conceptual model of geography, points have no dimensions (length or width).

In grid cell-based GIS, a string of grid cells represents a linear object (such as a road, a railroad, or a walking path). This string of cells can be lined up orthogonally (edge to edge), diagonally (along the corners), or some combination, depending on how curvy the feature is. Line objects are considered to have only one dimension (length), even though they usually have some width in the real world.

Polygons are a fundamental data type used to represent and model spatial features on the Earth's surface or within a defined spatial reference system. Polygons are closed, multi-sided shapes that have no gaps or overlaps, and they are commonly used to represent various geographic entities such as land parcels, administrative boundaries, building footprints, water bodies, and vegetation cover. These cells may be connected (contiguous) or disconnected (noncontiguous). Polygons usually are used to measure area in a GIS.

When you're representing surfaces or their volumes, each grid cell has, in addition to length and width, a number associated with the height or depth of the space. This number may represent elevation above sea level or depth to groundwater for a point at the center of a grid cell. In some GIS, grid-cell values even represent non-physical surfaces, such as population density or land appraisals.

Raster GIS represents points by using a single grid cell, lines by using a line of grid cells, areas by using a group of grid cells, and surfaces by using groups of grid cells that have additional unique values.

In GIS, vectors are a fundamental data model used to represent geographic features as discrete points, lines, and polygons. Vectors are based on coordinate geometry and allow for the precise representation and analysis of spatial data.

A Triangulated Irregular Network (TIN) is a vector-based data structure used in Geographic Information Systems (GIS) to represent continuous surfaces or terrain models. A TIN is particularly useful for modeling irregular or complex terrain surfaces, as it allows for variable resolution and can accurately represent features like ridges, valleys, and steep slopes. The neat thing about the TIN model is that you can use it to predict (interpolate) missing values, create cross sections through surfaces and volumes, draw contour lines, and create 3-D visualizations.

Most modern GIS software has both raster and vector data components, or it can convert from one format to another. You need to determine the most important functionality, accuracy, and storage issues for the work you want to do when you decide on GIS software.

The differences between Raster and Vector data models in GIS are:

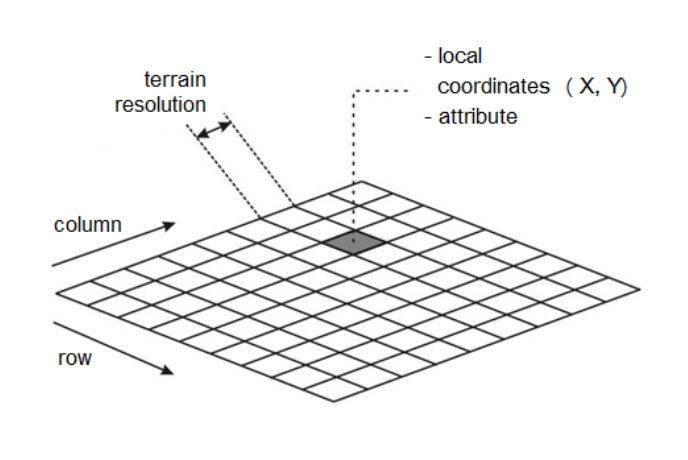
Raster Data Model:

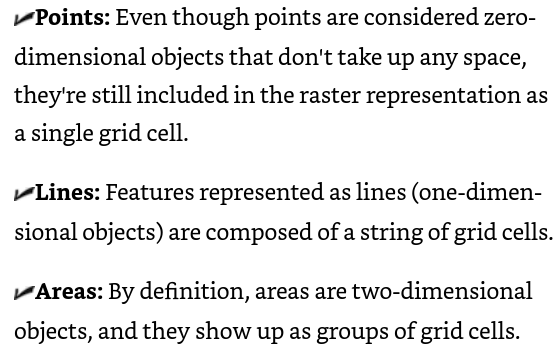
* Represents spatial data as a grid of cells/pixels
* Each cell stores a single value (e.g., elevation, land cover)
* Good for continuous data like satellite imagery, terrain models
* Resolution depends on cell size
* Spatial operations can lead to data degradation

Vector Data Model:

* Represents spatial features as geometric primitives (points, lines, polygons)
* Defined by precise x,y coordinate pairs
* Good for discrete features with well-defined boundaries (roads, parcels)
* Can maintain high precision at any scale
* Efficient representation of topology and spatial relationships

A grid overlay shows how each square represents a portion of real geographic space.





Cell grid resolution: The smaller the dots, the higher the resolution, and the more accurate the representation.

Using layered grids for modeling requires that:

* All the grids represent the same portion of the Earth.
* Grids are co-registered (meaning they lie directly on top of one another).
* Each grid cell is the same size in every map layer.

Raster GIS gives you the power to search the grid in two ways. You can search by coordinates to examine what the grid cell represents, or you can search by grid-cell quality to find out where grid cells with that quality are located.

A map theme typically corresponds to a thematic layer or a collection of related layers that share a common attribute or characteristic. For example, a map theme could represent roads, land use, population density, or vegetation cover.

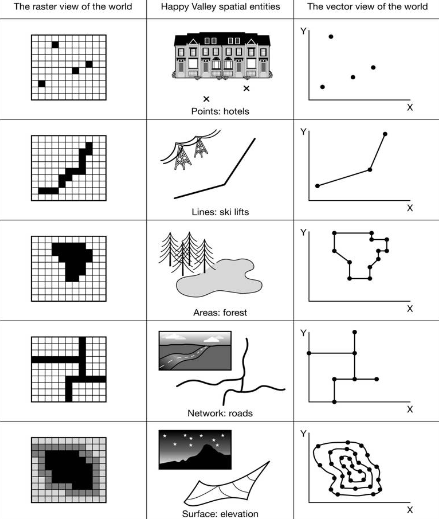
Map data breaks down into these components:

* Theme
* Category
* Value

The Map Analysis Package (MAP) data model makes it easy to find individual categories in a single thematic grid. Map packages are portable files that contain a map document (.mxd), the data referenced by the layers, and the page layout. They can be used to share maps between colleagues, across departments, or with other GIS users. Map packages can also be used to create an archive of a map or a snapshot of its current state.

Extending the raster data model by including a database management system gives you a lot more flexibility. But you need to make sure that you know where to look for the categories and that you give them names you can remember. Always use meaningful and memorable names for categories, ID-codes (codes you use for categories), and maps, if at all possible.

Simple forms of vector representation focus more on the accurate graphic depiction of features and less on the subsequent analysis of geographic information.



In a spaghetti representation of vector data:

1. Linear features are stored as individual line segments without any topological structure or relationships defined between them.
2. Line segments that should be connected (e.g., at intersections or nodes) are not explicitly linked or defined as connected entities.
3. Line segments may overlap, resulting in redundant data and potential inconsistencies.
4. Gaps or undershoots may exist between line segments that should be connected.

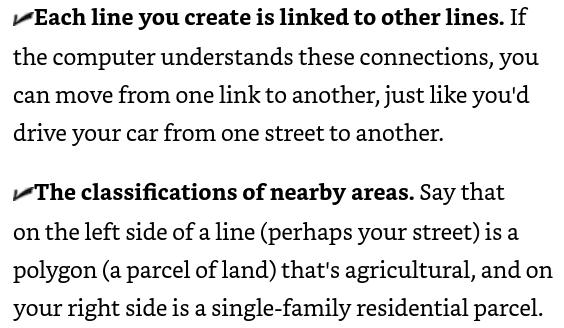
Conceptually, vector GIS software both stores and displays points and lines more accurately, making those points and lines nearly zero- and one-dimensional, respectively, and having polygons more closely represent their own shape. The closer together the points that make up line segments, the more accurately they represent the real lines.

One vector data model created by ESRI gives much more control than the spaghetti model. Instead of limiting storage of points, lines, and polygons to coordinate pairs, this model (called a shapefile) stores the geometry of each feature as a shape that contains the coordinates and links to the attributes. GIS software packages widely use this data model today because of its relatively low processing overhead, low storage requirements, fast drawing speeds, and its ability to handle overlaps and non-contiguous features.

A shapefile isn't a single file. Instead, it's three separate files that allow for the representation of 14 different types of geometric shapes.

A characteristic, called topology, allows the GIS software to recognize and convey some basic spatial relationships that humans express by using an English language construct called the prepositional phrase.

Topology studies the positions and relative locations of geometric figures. It allows the computer to understand how a neighborhood works (including what features are located there, where they are relative to each other, and how to find them). Because computers can't reason or look around at their neighborhoods, you have to tell them everything. And so you explain:



The need for computers to recognize linked lines and area classifications led the bureau to develop two of the more well-known topological data models — the Dual Independent Map Encoding (DIME) and subsequent Topologically Integrated Geographical Encoding and Referencing (TIGER) models.

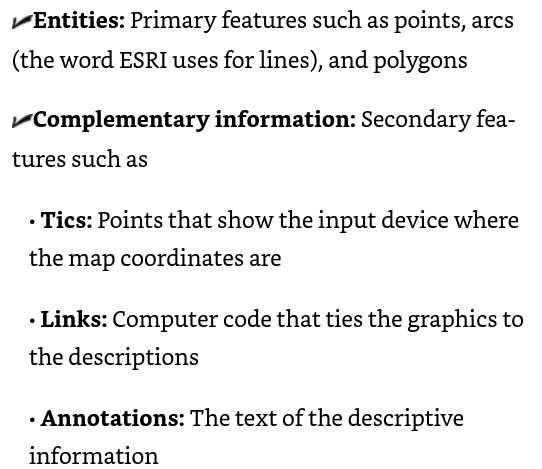
Right and left polygons: Tell you how the links are related to the polygons on either side of them. This is the contiguity property of topology, which identifies the polygon on either side of any given line. By default, it also tells you the direction of the line because it knows which polygon is on the left and which polygon is on the right of that line.

Area, as one of the basic properties of this topological data model, simply means that the polygons are defined by identifiable links, the coordinates of which are stored in the computer.

DIME (Dual Independent Map Encoding) was one of the first really useful datasets that was available to the general public for GIS. It also greatly enhanced the ability of researchers to examine large segments of the population through what's now called census geography. DIME recognizes the three basic components of topology (contiguity, connectivity, and area). The primary purpose of this model is to allow census data users to link the tabulated census information to census geographic units (such as streets, blocks, districts, and so on). The major drawback to the DIME file is that its geography isn't very realistic.

TIGER (Topologically Integrated Geographic Encoding and Referencing) is a data structure developed by the US Census Bureau to store and manage geospatial data. It's a public domain resource that includes comprehensive coverage of the US and its territories, and can be used in GIS applications for mapping and analysis. It improves on data retrieval and improved geography over the DIME structure.

The coverage model was developed by ESRI and has these features:



The development of object-oriented programming (OOP) — brought about a useful adoption for GIS data models. OOP focuses on objects where an object is a set of computer code that can be copied from place to place in a computer program. In GIS, an object represents a type of geographic feature that you can move around and whose properties follow it

The geodatabase (not to be confused with a collection of geocoded data inside a GIS software package that's sometimes referred to as a geodatabase) allows you to store a wide variety of data types (including raster, vector, CAD, and others) in the same system.

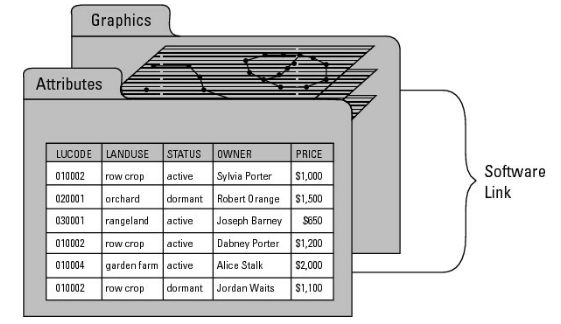
No matter what type of geographic setting or scales of geographic data you use in GIS, you need to link the graphics with meaningful descriptive data. When you perform geographic analysis, you search, manipulate, compare, and analyze this data. Every GIS software package must manage descriptive data and link them to their respective objects on the map. Without this capability, your GIS is nothing more than a digital picture, and you may as well work with a paper map.

Like GIS, computer-assisted cartography (CAC) was created to store, edit, and output maps of the Earth and — crazy as this sounds — other planets. CAC doesn't have the analysis capability that's at the heart of GIS, however. CAC is a straightforward process that moves from input to storage, retrieval, editing, and finally, output (see Figure 6-1). A CAC system just needs to store the attributes, or descriptive data, and have labels attached to the attributes so that those descriptive labels can be applied during output.

Like computer-assisted cartography, computer-aided design (CAD) isn't intended to manipulate or analyze cartographic data. In fact, CAD isn't intended to work with cartographic data at all — but it can. CAD systems — which are designed to store complex drawings in layers — store attribute (descriptive) data as linked lists tied explicitly to the drawing's components (the entities) on a layer-by-layer basis. This layer-by-layer division means that the data are not shared from one layer to another.

Because raster grid cells are uniform in size and identically placed from one map layer to another, you may find them especially useful when you must perform map overlay operations for map analysis. Also, because the raster model is so simple, it makes map overlay faster than in vector systems, which usually have more complicated data structures

Environmental Systems Research Institute (ESRI) first fully implemented the link between graphics and descriptions stored in a database management system in its premier product called ArcInfo. Arc is the portion of the software suite that deals with the graphics, and Info is an existing database management system that's linked to Arc so that changes in entities (the graphics) are reflected in changes in attributes (the data) and vice versa.



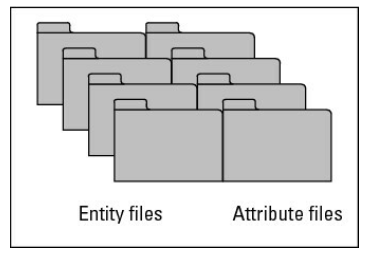
SQL is a fairly standard computer language designed specifically for accessing, querying, and manipulating databases. It allows you to search for very specific information, retrieve it, delete or add to it, update it, and combine queries. SQL searches database tables by using commands such as Select, And, Or, Insert, Delete, Order By, and many more.

An object-oriented database management system (OODBMS) extends the traditional database management system so that it can model and create objects based on those data. Objects are collections of data and computer code that have two primary characteristics — inheritance (a common set of properties that move with them) and hierarchy (a place within larger classes of objects that have common properties).

The biggest advantage of object-oriented systems such as the ESRI geodatabase is that they add these behaviors to the stored objects. So, if you try to put geographic features into a database that violate the known behaviors of the object class they belong to, the software rejects them. Also, the behaviors often translate into how you can and can't link different objects in your GIS.

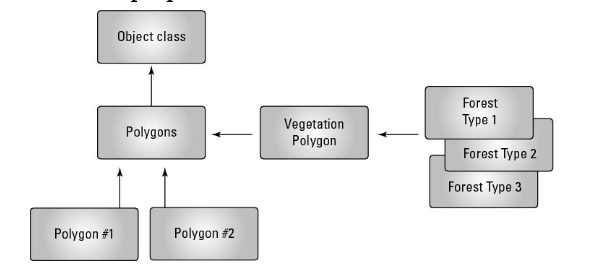
In RDBMS terminology, a general data model that's modified to address the specific data and query needs for a given application is called a schema. So, for example, you might have a general schema for vegetation analysis that you can modify for desert vegetation.

Integrated systems still treat each layer as a separate set of data and provide a naming mechanism so that they can find any layer needed. But the integrated systems don't need links or pointers from entities to attributes because all the records are contained in the same RDBMS files



Objects, the groups of geographic features represented in the database, have a unique property called inheritance, which means:

* Objects can exhibit properties and rules that govern their behavior in your database. These properties and rules make the features you include in your GIS more realistic in their representation of geography.
* Any feature that belongs to an object class (a class is a group of similar objects, such as streets or land-use polygons) shares (inherits) the properties of the class.
* An object class can also belong to larger classes of objects that also share their properties. So, a group of forest polygons would share the properties of all forest polygons, but they would also share the properties of a larger class of polygons — vegetation polygons



* Choosing a system model has much more to do with what you want to accomplish and the nature of your analysis. These three major factors might help you decide on an object-oriented system model: How well you know your data: If your system is based on a well-known geography (for example, on your knowledge of how crime occurs geographically), the object-oriented model allows you to code that logic into your database.
* What depth of analysis you require: If you need to supply additional geographic rules and logics for some existing geographic features that you understand well, then you should use an object-oriented system model that allows you to include those rules.
* How you see your GIS needs evolving: Moving toward the object-oriented GIS can help your system become more adaptable to innovations in system design that will continue to use object-oriented programming and database management techniques.

Select only the best-quality and most appropriate data type for your needs. Collecting more data than you need and gathering data that includes irrelevant details add to the costs and reduce the effectiveness of the resulting GIS.

Sensory data as those data most commonly associated with distant sensing devices, such as the Global Positioning System (GPS), and various forms of imagery, including both aerial photography and digital satellite data. Statistical data include field data and census data, both of which usually rely on some form of direct contact by a person to collect.

Grid cells are specifically related to recorded imagery, they're called picture elements — or pixels, for short. Each time a sensor senses an area, it records a single radiation value in a particular radiation band for each pixel. The more sensors available on a single satellite, the more sensor bands you pick up because each sensor looks for a unique part of the electromagnetic spectrum.

With layered digital images, you can work with the separate spectral (wavelengths of radiation) layers and decide how to put them together to get the exact picture you want.

You can use images gathered from aerial photographs as base maps (a starting point for more complex maps) when you're heads-up digitizing (digitizing off the screen) because these photographs give a nice, relatively current picture of the Earth. Also, enhanced images make heads-up digitizing even easier than the raw images do because enhanced images allow the human eye to see things not visible under normal circumstances.

Supervised classification in GIS is a method of image classification where an algorithm is trained on labeled data to learn the relationship between spectral values and land cover classes. The algorithm is then applied to new images or datasets to assign pixels to specific classes, such as vegetation or urban areas, based on their spectral characteristics. This technique is useful for land cover mapping, crop classification, urban planning, and environmental monitoring, but requires high-quality training data to achieve accurate results.

Unsupervised classification in GIS groups pixels into clusters or categories based on their spectral similarities, without prior knowledge of class labels. The algorithm identifies natural patterns and relationships in the data, assigning pixels to clusters or categories. This technique is useful for exploratory data analysis, discovering unknown patterns, and reducing dataset dimensionality. Unlike supervised classification, it doesn't require labeled training data, but results may require more interpretation to understand the meaning of the identified clusters.

Whichever classification method you choose (supervised or unsupervised), you get what remote sensing scientists call an image map. And you can put that image map into your GIS.

When you use remotely sensed images, you have access to a large number and different types of land-sensing satellites, so you can obtain new images frequently. As a way of updating GIS data, especially information such as land cover and land use, this type of GIS input is invaluable because it covers so much territory, so frequently, and at a relatively minor cost per unit area.

No matter what types of data you intend to collect, you generally can't find and record every single instance of the features you're examining.

Spatial sampling can give you a much better representation of the geography. Whatever method you use, the key is simply to recognize that your data has a spatial distribution and a sampling strategy that ignores that distribution usually results in less than useful data.

Thousands upon thousands of hard-copy maps have yet to make it into computers. You can digitize hard-copy maps in three basic ways: manual or hand digitizing; scanning; and heads-up digitizing from a scanned image, using your computer monitor.

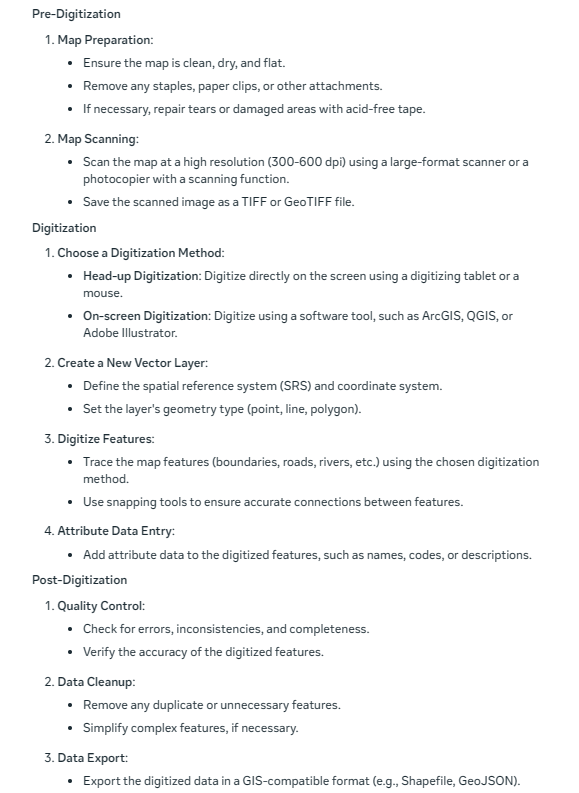
Manual or hand digitizing is perhaps the oldest and still a time-honored, albeit tedious, method. Scanning involves using advanced versions of the technology of the modern digital camera and personal flatbed scanners. Finally, heads-up digitizing is an increasingly common form of input.

Both scanning technology and the software that interprets the results have improved enormously over the past three decades. You may already own or use a simple scanning device, such as a copy machine or small flatbed scanner. Although these types of scanners are a similar technology, for most maps, you want to use a large-format drum scanner

Because much aerial photography and even many maps have already been scanned as graphic images, more GIS data input specialists use those graphic images as a background template against which they do digitizing. You display the image on the computer monitor and then selectively isolate and record the locations of the image features inside the computer

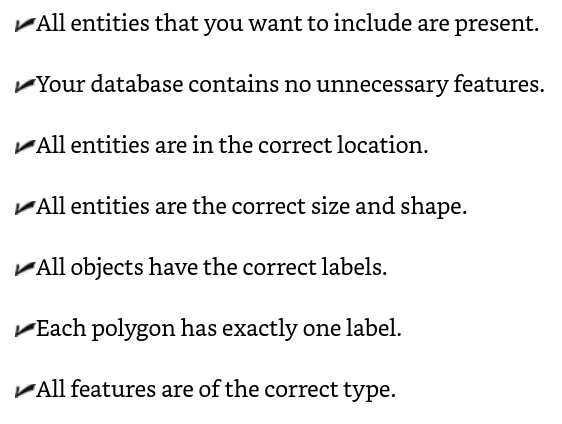
Digitizing some maps takes a very long time, so you can get the most accurate results by digitizing those maps in a series of short sessions because you will be less fatigued while you work.

The general steps for map digitization are:

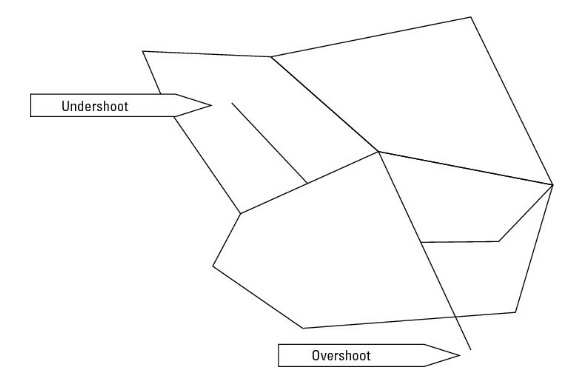


First, decide what you want out of your GIS, formally called the spatial information product (SIP) — meaning what you need your final map to look like. You need to decompose your project — break it into individual bite-sized pieces rather than trying to digitize large complex maps all at one time. Nearly every SIP is a composite of all or pieces of various maps.

You need to look for two general types of errors when deciding whether your GIS database is ready for use: entity (graphical) and attribute (description) errors. Things to check:



Entity errors that relate to the accuracy of location and position, collectively called dangling nodes (dangles, for short), come in two types — undershoots and overshoots. Overshoots and undershoots might cause your GIS to either over-report or under-report the number of polygons in your database, which can also make your labels inaccurate.



Metadata is data about data. It's a complete description of the major facets of quality, source, and accessibility of the data that make them useful not only to you, but to others who might use the data after you do.

There are seven basic entries for geospatial metadata:

* Identification: Defines the database name, developer, area of coverage, categories, themes, collection methods, and data access
* Data quality: How accurate and complete the data are, adaptability for application, positional accuracy, completeness, consistency, and sources for derived data
* Spatial data organization: Data model (Vector/Raster), number of spatial objects, additional non-coordinate methods of location encoding
* Spatial reference: Latitude and longitude coordinates included, projections used, datums, parameters used for coordinate transformations
* Entity and attribute information: Specific attributes, what information is encoded, codes, and code definitions
* Distribution: Where and from whom to obtain the spatial data (you can often access and read the metadata before you actually obtain it), available formats, media type(s), online availability, and cost
* Metadata reference: Compilation date and compiler

Always provide the most current, complete, and accurate metadata you can. By the same token, never buy data for your GIS from a vendor who doesn't supply complete metadata.

In GIS, you can comfortably map many types of information to a geographic grid because all the information can be coded into a computer and selectively retrieved for whatever subset of data you want. You use computer search and retrieval techniques to get just the information you need.

An example of a complex analysis might include combining maps to determine the best place to put a dam based on the amount of water in a stream, the shape and size of the river valley, the number of people and homes that might be impacted, and the cost of putting the dam in at that location. Complex approaches nearly always require fairly simple forms of search strategy.

Whether you're interested in complex analysis and modeling, or the plain vanilla results of a search, you need to understand the search process.

After you decide on a specific search object, you can follow these general steps to find the object and verify the results in most raster GIS:

1. Select the map layer that contains your search object.
2. Specify the names, codes, measurements, and so on that identify your search object.
3. Review the retrieved information and make sure you got what you searched for.
4. Name your new map.

Linear features are common search items. Among the most common linear features are roads and other transportation routes, such as rail lines. In grid GIS, these features appear as a series of connected grid cells within a transportation layer.

Searching for polygons (areas and distributions) in a GIS helps you determine how much of your specified search subject exists (which can be important information), but your search results can provide other insights, as well.

In a grid-based GIS, areas are comprised of groups of grid cells that share common code numbers. So, like with point and line features, the software begins by searching in the appropriate layer. Then, it identifies the groups of grid cells based on how you formulate your search for the existing codes.

In simple raster systems, you look for grid cells to find collections of cells representing points, lines, or areas. The type of search strategy involves selecting the categories that you want to work with within each layer.

Attribute searches in simple grid-based systems find the codes associated with each category or class of data directly. Because no DBMS tables are associated with these grids, the search is simple and often obvious because the categories and classes of data for each map show up in the legend. So by looking at the legend, you can easily see all the information that's available

The primary difference between the database supported type of system and the older, more basic grid-based systems is that the software can store and manage far more complex and detailed categories. The attributes aren't limited by what the map legend can easily display. So, the attribute search is much more powerful with these systems than with their earlier counterparts.

Raster systems that have associated database management systems (DBMS) allow the same approach as vector systems to finding points, lines, and polygons. You can still perform searches based on spatial query, graphic, dimension, data type, and attributes.

To count the number of point objects, you need only count the number of grid cells.

Among the most powerful basic tools of the GIS is its ability to tabulate the numbers you collect and provide you with descriptive statistics. These descriptions include maximum and minimum values for a search, measures of central tendency (the mean, median, and mode), and measures of dispersion or difference around the measure of central tendency (e.g. standard deviation).

You don't need to know the details of basic descriptive statistics, but professional GIS software typically offers basic, as well as more advanced statistical techniques.

When displayed next to your map output, or on the map itself, the statistics have a geographic component. Big features aren't just big — they're big in a particular place, which communicates the importance of space and place to your audience.

The most graphically accurate method of representing maps in a GIS is the vector system. Vector systems represent geographic space by using a series of points, lines, and polygons. This method differs from raster GIS, in which geographic space is stored as a collection of squares called grid cells.

In vector GIS, you use the power of the computer to identify the stored features by name or description. You usually need to find features on maps so that you can use them in more complex tasks, such as finding routes from place to place, comparing different maps, and finding the best place to put a business.

Unlike raster systems, vector GIS systems have location information that's more accurate and visually more like the real world, so you can search geographic space to achieve more accurate results.

If you're often concerned about how certain data types that you retrieve will work together, choose a GIS package that has the object-oriented data model. This model's structure validates the proper relationship of the data during retrieval and helps ensure that your combinations work together to produce legitimate results.

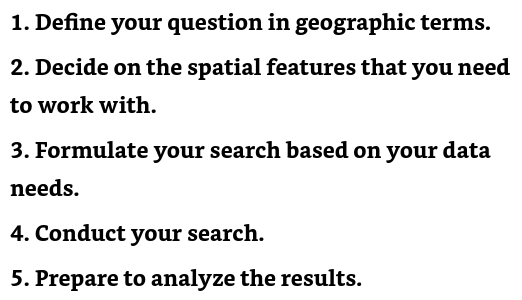
There are a few different ways to query a GIS for particular information:

* Search by attributes, using the SQL interface.
* Search for features by using the graphics interface.

There are three particular criteria that need to be identified prior to running a search:

1. Theme or topic. The search for individual features is first directed to the layer that contains data for that theme.
2. Classification system. Knowing what classification systems identify the information you need can help you decide what specific classes to search.
3. Level of measurement. The level of measurement determines what search values can find the data you want.

Follow these steps to work within the search-drives-analysis concept:



To perform an attribute search using SQL, use these basic steps:

1. Open the database.
2. Highlight the theme you want to search.
3. (Optional- but helpful for accuracy) Open the attribute table for that theme and examine the data.
4. Open the query tool.
5. Formulate the query.
6. Apply the query.
7. View the results.

You can create some really cool layers just by searching for data and then, displaying what you find. And, you save and copy them in case you want to use them again. Each software package works a bit differently, but all enable you to save to the database the layers you create with your searches.

To find out information about an area on the map, graphically, the general steps are:

1. Choose the Select by Location tool or menu.
2. Select the relative location you want to search.
3. Run the Select by Location tool.
4. View the result.

You can sort, group, and perform calculations on the table data by selecting the available options built right into your software, much like your favorite spreadsheet.

After finding data, it’s important to verify:

1. Did I get the correct data?
2. Are the data on the right scale?
3. Are the data timely?
4. Are the data complete?

The vast majority of geographic objects that you encounter every day occur in groups. Some of these groups even have specific names. Birds occur in flocks, cattle in herds, trees in forests or orchards, quail in coveys, geese in gaggles, and the list goes on. Beyond the biological groups, glacial features (called drumlins) form in swarms, streets form networks, streams form watersheds, homes form neighborhoods, and many more. People assign special names to groups to provide a handy way to refer to features that look alike, act alike, occur in similar places, are near certain features, or occur in similar patterns or densities.

Distribution, in GIS usage, can roughly be defined as data spread over a large area, often represented as polygons. Whether polygons represent distributions of individual objects (such as people or animals) or area features (such as land parcels or forests), you can use many of their properties for later analysis. The data inside polygons can exist at any of the measurement scales, including nominal, ordinal, interval, ratio, or scalar.

Polygons are connected to each other, surround other things, and are surrounded by other things. They have a long list of spatial relationships that other objects don't have. To get the most from GIS, you use those properties and relationships to make decisions regarding the data you retrieve when you search your GIS.

Searching for objects in modern GIS requires a logic-based search strategy, which a graphical user interface (GUI) provides. The GUI uses Structured Query Language (SQL), which includes basic AND, OR, and NOT operators, as well as mathematical comparative operators such as =, >=, <, and many others.

Hot spot analysis in GIS identifies areas with high concentrations of spatial phenomena, such as crime incidents or disease outbreaks, by clustering points of occurrence into polygons or converging points. This helps pinpoint areas of high activity or risk, informing targeted interventions or resource allocation.

You can modify polygonal categories, values, or boundaries so that you can control the content of the maps you use for analysis. When you do so, you don't change the data; you just organize it differently.

Some ways that data can be reorganized includes:

* Clumping or aggregating categories
* Changing data measurement scales

You can search for polygons in six ways: by (1) category, (2) data level, (3) size, (4) arrangement, (5) orientation, or (6) position.

Some systems enable you to access nominal categories directly from the map legend, and others give you a GUI based on SQL.

Get to know your GUI search tool as soon as possible — you'll use it often in your GIS work. Start with simple searches, and when you get comfortable with it, move to more complex searches.

Many polygons come in ordinal (ranked) categories. The user might have created those categories for a particular use in which the scale is context-dependent (scalar), or another user might use more universally understood categories, such as predefined values for high, medium, and low housing densities.

Searching for different attributes within a map legend doesn't create a new layer. Most often, searching for attributes allows you to quickly change the display, but that's about it. If you want to save and use the resulting map, you must use the GUI.

Maps can trick you. Before you search for polygons based on value, be sure that you're searching for the correct unit value.

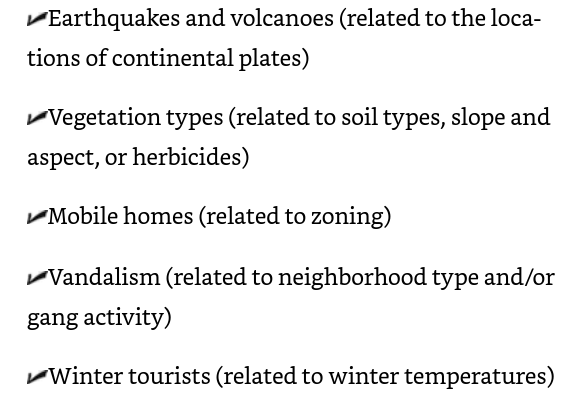
Most GIS software even has a tool you can use to select objects that are touching (adjacent), so you know exactly what parcels of land are next to each other.

Your project goals almost entirely control how you choose the groups of features you want to find. No formula exists for choosing groups, but you can ask yourself some basic questions. These questions are based on the types of data you have, the theme or themes you're interested in, and what feature properties and underlying geography you need to know about.

To find groups of geographic features — whether they're points, lines, polygons, or even surfaces — you first need to know what properties they share. You already do this type of grouping naturally when you decide on the types of movies you like to watch (categories); whether to order a small, medium, or large pizza (rankings); or how much money to spend for various types of entertainment (ranges). You can use the same thought process to select polygons by

* Category
* Rank
* Scale
* Spacing

Here's a quick list of some features that might exhibit patterns related to some form of controls or outside forces:



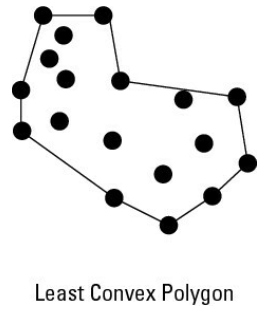
Some striking distribution changes aren't necessarily accurate: They may be a result of lack of sampling.

Don't just assume the patterns that occur on maps are accurate. Check the metadata for information on methods and sampling procedures, as well as any other information that allows you to verify the data's reliability.

Sudden changes most often indicate an equally abrupt change in some other factor that either controls or contributes to the existence of the features. And if your distribution changes gradually, a process is probably controlling that change.

If the distribution changes gradually and in a particular direction, this change may reflect the direction of change in an underlying process.

One basic method of enclosing distributions into a single polygon is almost like using shrink-wrap. You decide which points you want to include, and the software draws a polygon by connecting the outside points, creating the least convex polygon.



Base your grouping strategy on what you need to know plus what you already know, to satisfy many other kinds of GIS projects.

You use map measurements to figure out how far you are from one place or another, how much land you own, how much ore you might expect to get from a new mine, and many other measurements. You can measure heights, widths, depths, lengths, areas, and volumes. You can also compare measurements as ratios — for example, ratios of length to width, perimeter to area, or height to distance.

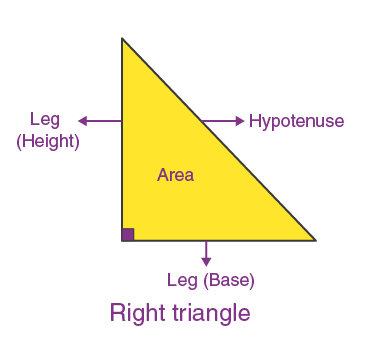
The measurement of absolute distances, whether between objects or along lines, is computed differently for raster and vector data structures.

The Euclidean method makes two assumptions:

1. No obstructions block the path between the two points.

2. You don't need to travel along established paths and roads.

The Pythagorean Theorem is based on the use of a geometric form called a right triangle — any triangle that has one corner that makes a right-angle turn (90 degrees). The longest side of this triangle, the one opposite the right-angle bend, is called the hypotenuse.



When measuring real Earth distances in a GIS, the software calculates the hypotenuse distance using the X and Y values of two points in a modification of the Pythagorean Theorem called the distance formula.

The GIS can also calculate distance between a set of latitude and longitude coordinates on a spherical Earth. You may need to find the shortest distance between two points on the globe. This calculation is called a great circle distance and uses a geographic grid.

Many raster GIS packages allow you to calculate measurements in both flat coordinate systems and the geographic grid (latitude and longitude coordinates). The alternative to measuring by coordinates is counting up grid cells.

Manhattan distance is a way to measure distance in a city with a grid layout. It's like the distance a taxi would drive, only counting horizontal and vertical movements, not diagonal ones. It's useful for estimating walking distances in cities with a grid pattern. To determine the total Manhattan distance, you simply let the software add together each individual line segment's length.

To calculate distance along networks such as footpaths, animal tracks, rail lines, or street patterns, the GIS measures vector segment lengths or grid-cell lengths for vector and raster, respectively.

* Vector: The GIS first retrieves the X and Y coordinates of each line segment's endpoints (stored at input) and then uses the distance equation to determine each line segment's length. Finally, the GIS adds together all the individual line segment lengths.
* Raster: The GIS finds each orthogonal grid cell's width and then adds those to get the total length. When grid cells are diagonal to each other, the lengths are measured as the width multiplied by 1.414 (the diagonal distance of a grid cell).

Nearly all GIS software has a set of special distance functions designed to selectively calculate distances around existing geographic features. These operations are called buffering (no, not like the aspirin), and the result is called a buffer. The software measures outward a certain distance from a point, line, or polygon (even the base of a topographic feature), and it then converts that whole area to a polygon.

By using the GIS variable buffer function, you can select the size of the barrier for each portion of a line or polygon feature, instead of being forced to use the same-sized buffer for the entire feature.

By using the GIS variable buffer function, you can select the size of the barrier for each portion of a line or polygon feature, instead of being forced to use the same-sized buffer for the entire feature.

A setback buffer has that name because it's "set back" in only one direction. You can position a setback in either direction from a line feature (for example, the center of a street). For polygon features, the setback buffer is usually established (set back) from the perimeter to some distance inside the feature.

Bidirectional buffer: Most buffers measure distance outward. Setback buffers can measure distance inward. One advantage of doing a buffer around an area feature is that you can have a buffer going both directions at once. Most GIS software allows you to select this option so if you want to measure a buffer distance both inward and outward from the outside perimeter of an area feature, you can.

In GIS (Geographic Information System), adjacency refers to the spatial relationship between two or more features that share a common boundary or edge. Adjacent features are those that touch each other, but do not necessarily overlap.

Relative measurement: not a formal measurement, but simply knowing whether you are near or far from a point of interest.

When you measure separation and isolation, you compare the absolute coordinates of one geographic feature to several other features so that you can determine the average distance from a feature to other features.

Containment refers to the spatial relationship where one feature is completely enclosed within another feature. In other words, a feature is said to contain another feature if its boundary completely surrounds the other feature, with no gaps or overlaps.

Knowing whether an object is contained in a polygon can help you answer these kinds of questions:

* Does the property I'm interested in fall within a certain school district?
* Does a particular wildlife species live in a certain habitat type?
* Do I live inside a 100-year flood zone?

In GIS, functional distance measures the effort or cost of traveling between two points, considering factors that affect the journey. It's not just about the straight-line distance, but also takes into account barriers, surface distance, mode of transportation, vertical and horizontal factors that can slow or speed up travel. This provides a more realistic and nuanced understanding of distance, useful for planning routes, estimating travel times, and analyzing spatial relationships.

Anisotropy in GIS refers to the variation in spatial relationships and patterns depending on both distance and direction. In other words, the connection between two locations changes not only with how far apart they are, but also with the direction they are from each other.

Isotropy is the property of a totally smooth surface that has no paths, bumps, wrinkles, vegetation, or obstructions of any sort. If you find an isotropic surface, let me know because it's really just a theoretical property. Isotropy doesn't exist in nature, but it does give you a standard against which you can compare all your anisotropic surfaces.

The physical forces associated with impending travel are:

* Slope
* Surface Type
* Surface Features

In GIS, intangibles are subtle and intangible things. This can be something like scenic views and plays a role in the perception of distance when travelling.

A friction layer in GIS is a surface that represents the ease or difficulty of movement across a landscape, typically measured by travel speed or time. It's created by combining datasets such as roads, topography, land cover, and borders to calculate the "friction" or resistance to movement for each location.

In theory, you can quantify friction easily: big friction = big numbers, little friction = little numbers, no friction = zero.

With GIS, surface means much more than the physical topography. You can map and analyze many different statistical surfaces, which include features such as population density, crime rates, and cost of living.

Non-topographic statistical surfaces come in two general forms — those relating to the physical environment and those relating to the human environment. This breakdown doesn't have any intrinsic value; it just provides a simple way to categorize non-topographic statistical surfaces.

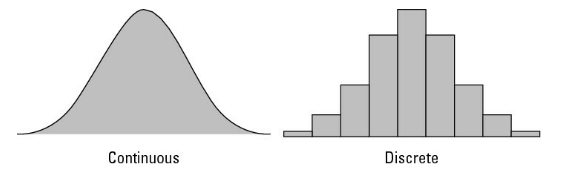
A Triangular Irregular Network (TIN) is a way to represent a continuous surface, like terrain or temperature, in a GIS. It's made up of connected triangles that are formed by drawing lines between known data points. Each triangle has a consistent slope and direction, allowing for a detailed and accurate representation of the surface.

In GIS, a spline is a mathematical tool that creates a smooth curve or surface through a set of points. It helps fill in missing data and creates a continuous representation of the data. Splines are useful for visualizing and analyzing data, creating digital elevation models, and modeling environmental phenomena like temperature and precipitation.

In GIS, a Digital Elevation Model (DEM) is a digital representation of the terrain or landscape, showing the elevation and shape of the land. It's a 3D model that can be used to create maps, visualize terrain, and analyze geographic data. Think of it like a digital version of a topographic map!

In GIS, a discrete surface represents a surface as a collection of distinct points or cells, each with a unique value. It's used for discrete data like land use classifications, soil types, or elevation points. Examples include raster and vector data, and 3D point clouds, which differ from continuous surfaces like TINs or splines.

In GIS, a continuous surface is a smooth, continuous function that represents phenomena that vary gradually over space, such as elevation or temperature. Examples include TINs, DEMs, and continuous raster surfaces, used for spatial analysis, modeling, and visualization to create detailed representations of complex spatial phenomena.



Surfaces can change slowly or abruptly, and you can classify these surfaces as rugged (as in rugged terrain) or smooth. The ruggedness or smoothness of a surface hints at characteristics of that surface.

Statistical surfaces have three properties:

1. Z values across geographic space
2. Measured and recorded at intervals or ratio scales
3. Have continuous distribution

Continuous statistical surfaces require sampling because the data occur everywhere. You can sample a surface by following these steps:

1. Examine what the surface looks like
2. Locate where the surface is smooth and where it’s rugged
3. Define a sampling strategy- how many samples to take and where to take them

Smooth surfaces have less information, and therefore require less sampling than rugged surfaces.

Computer programs don't actually need all the data to perform their interpolations, and you can reach a point of diminishing returns in which you gain little or no benefit by adding more sample points. The software might not even use all the additional points you provide. Not to mention, taking more samples requires more effort, which typically costs money.

Interpolation is the technical term for predicting the Z values of an area based on samples that you've collected for other areas.

For discrete surfaces, you can use two basic ways to locate points for interpolation:

* By using the centroid (geometric center) of each polygon: For a rectangular polygon, the centroid is right at the intersection of two lines drawn diagonally from each corner.
* By using the center of gravity of the distribution: The center of gravity method means that you pick a point that's closest to the location of the highest concentration of your samples.

In GIS, displays can mean any graphical output (mostly on a map).

In GIS, draping means overlaying a 2D image or data over a 3D terrain, creating a interactive 3D visualization. This helps users analyze and understand the relationship between the data and the terrain, and is useful in various fields like urban planning, environmental monitoring, and more.

In GIS, interpolation is a method to estimate values at unknown locations based on nearby measured values. It fills in gaps in data, creating a continuous surface or image, and is used for tasks like mapping terrain, climate, or population density.

Linear interpolation in GIS is a way to fill in missing data by drawing a straight line between known points. It estimates the value of a point between two known points by finding the point on the line that connects them. This helps create a smooth and continuous surface, making it useful for mapping and analyzing data like elevation, temperature, or population density.

When selecting a non-linear interpolation method, first characterize your surface (gradual, steep, or a combination; smooth, rugged, or some combination). Then, become familiar with the nature of the processes at work that cause the surface to exist. With this knowledge, you can match the techniques that best suit the way the surface changes.

Inverse Distance Weighting (IDW) is a method in GIS that uses nearby points to estimate unknown values. The closer a point is to the unknown location, the more influence it has on the estimated value. IDW uses a formula to calculate this influence based on distance, allowing you to create a continuous surface from scattered data points.

Exact interpolation methods in GIS use nearby points to estimate values at new locations without making assumptions. These methods include Nearest Neighbor, Radial Basis Function, and Spline interpolation. They are useful when data is accurate and you want to preserve original values and patterns. They provide a precise representation of the data, making them ideal for applications where accuracy is important.

A trend surface in GIS is a way to create a smooth, gradual surface that shows the overall pattern or trend in a dataset. It's like a "best fit" line, but for a whole surface. It helps identify general tendencies and coarse-scale patterns in the data, rather than exact details. This method doesn't provide an exact prediction of missing values but deviates from exactness to give a general idea of how the slope is trending. You can use this kind of trend surface method if you just want a general idea of a surface — and the accompanying general trend in a distribution — rather than all the gory details.

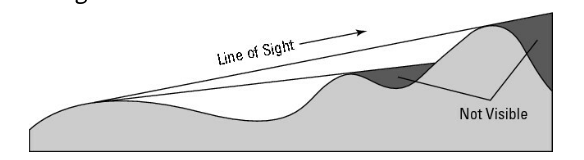
Viewshed analysis in GIS is a tool that shows what areas are visible from a specific location, like a hill or a tower. It helps answer questions like "What can I see from here?" or "What areas will be affected by a new building or structure?" It's useful for planning and understanding visibility in different landscapes.

Point to point: Analyze visibility from one point to another point. You have to know the location of your target point (the place you do or don't want to see) and the location of the observation point where you (the observer) are standing.

Viewshed analysis can also be used to determine visibility between multiple locations, like a hiking trail and a housing development. It helps you see what parts of the development will be visible from different points along the trail, and vice versa. It's like checking the view from multiple points, but with more complex calculations to cover the entire trail and development.

In simple terms, optical geometry is the study of how light travels from our eyes to objects in the landscape, and how it interacts with those objects. It's like drawing imaginary lines from our eyes to the things we see, to understand how we perceive the world around us. This concept is used in physics and optics, but also has applications in GIS and spatial analysis, particularly when studying topography and visibility.

Ray tracing in GIS is a method that simulates how light travels from our eyes to objects in a 3D environment. It helps create detailed and accurate images by tracing the path of light and determining what is visible from a specific viewpoint. This technique is used to generate realistic visualizations and analyze visibility in GIS applications.



You can easily build these little surface features into your viewshed model. So, if you're working with a topographic model and have no vegetation or buildings to get in the way, you need to adjust only for the height of your eyes from the ground. To make that adjustment, you just add that height value — say, 5 feet — to the observer's elevation.

The most accurate viewshed models add the heights of obstructions from data in other map layers, as well as from the observer's height. Most of the time, though, you don't know the observer's height, and your elevation model isn't accurate enough for these minor improvements to make a huge difference. Use these refinements only when you have very good elevation data and a thorough knowledge of the observer and the obstructions.

These three metrics — stream ordering, basin analysis, and water accumulation — are all inextricably linked to one another. You can't really do stream ordering until you can define the stream based on a thorough knowledge of the basin. When you have a complete model of the basin, you can finally model the movement and accumulation of water.

Basins or watersheds are composed of the area upslope from the stream network in such a way that all the water landing on the watershed could potentially flow overland into that stream network.

The accumulation is really a measure of distance based on where the water will go. The deeper the basin, the deeper the potential water accumulation.

You can most effectively model stream basins by using a raster GIS.

The places where the stream grid cells form each sub-basin are called pour points because, at these locations, water pours from one sub-basin to another.

In this simplified version of how the accumulation model works, each grid cell usually has additional attributes that include values such as the amount of precipitation, and absorption rate of the surface over which the water moves.

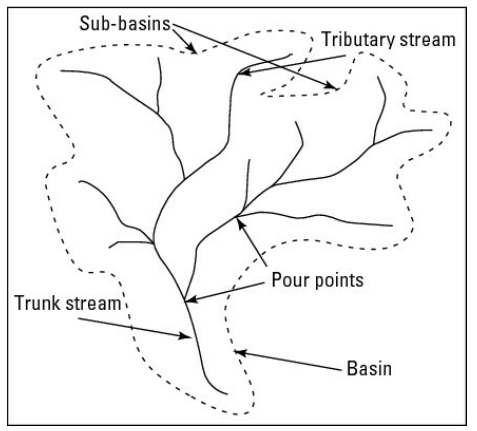
Watersheds are pretty irregular in elevation values, even if you ignore the occasional highs and lows in the surface. These irregularities result in water moving in different directions. When you add differential surfaces to your GIS database, you can build some really powerful models of flow direction.

The GIS software looks for the direction of steepest descent, meaning the one neighboring grid cell that has the steepest drop in elevation from the starting cell.

The stream is a very common land feature. Streams usually don't occur in isolation. They usually have a lot of branching tributaries. The numbers and positions of streams have an impact on stream flow, water accumulation from overland flow, and even the ecological conditions of the streams and the upland corridor (often called a riparian corridor).

To perform stream morphometry (a way to model how tributary streams connect to bigger streams to make stream networks), you first need to define the locations of the streams and their watersheds

Geomorphologists (scientists who study landforms) use this information to understand how the streams will change over time through the processes of erosion and deposition. Most GIS software includes the well-established methods of stream morphometry.



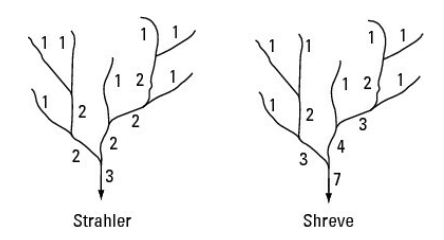
The stream branchings at the pour points allow you to determine stream morphometry by providing the skeleton of your basin.

After locating the stream, you need to perform a stream-order analysis. Sometimes called the bifurcation ratio because the streams tend to break in twos, the stream-order analysis calculates the relative locations and connections of a stream network, characterized by the order in which the streams come together.

You can calculate stream order in several ways, but the two most common are the Strahler method and the Shreve method.

The Strahler method is a technique used in GIS to organize stream networks by assigning a numeric order to links. It starts by labeling all links without tributaries as "first order". When two links of the same order intersect, the resulting link is assigned a higher order (e.g., two first-order links combine to form a second-order link). However, if links of different orders intersect, the order doesn't change. This method helps create a hierarchical structure for stream networks, but has some limitations, such as not accounting for all links and being sensitive to changes in the network.

The Shreve method is a technique used in GIS to organize stream networks by assigning a numeric order to links. Unlike the Strahler method, the Shreve method adds the orders of intersecting links together, rather than increasing the order only when links of the same order intersect. This means that when two links meet, their magnitudes (orders) are added together to determine the order of the resulting link. The Shreve method is particularly useful in hydrodynamic analysis, as it helps estimate discharge volumes and pollution levels by summing the number of sources in each catchment above a stream gauge or outflow.



Special linear objects are called networks. Networks are collections of connected linear objects such as roads, railroads, or rivers that branch from place to place.

Networks that allow for movement along its length is called a corridor.

You can measure connectivity in a network by comparing the number of actual node-to-node links that exist in a given network to the maximum number of nodes that are possible. This measure of connectivity is called the gamma index. Most applications of the gamma index relate to human transportation systems. index. Usually, the index ranges from a value of 0 (which indicates no connected links at all) to 1 (where all possible links are connected).

For any movement, the nature of the network as a corridor has an impact on how fast things move and, in some cases, even whether they move at all. The resistance to movement is called impedance and can be a function of the size of pipeline in a water delivery system, the roughness of dirt or gravel roads, or the number of lanes in highways. The most common use of network impedance is to model street patterns for transportation routing.

Some options in a GIS impedance layer could be:

* Impedance attribute: How long it normally takes to travel a certain distance.
* Default cutoff value: The value at which the computer stops searching for a location. You can override or change this value.
* Accumulation information: A list of possible attributes that accumulate while distance increases, including costs, students riding a bus, and many more.
* Restrictions: Restrictions that you place on the use of links (for example, the permitted types of traffic on portions of your road network). For example, you can choose to force hazardous cargo to use certain streets.
* Hierarchy: A set of rules for travel, regardless of how your software determines the impedance values.

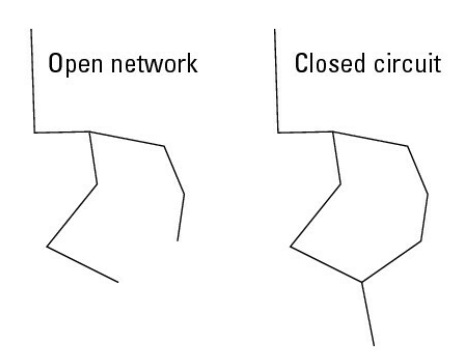
The OD Matrix is a table that shows the travel time and distance between different places. It helps find the best routes for things like delivery trucks. It's made by adding starting and ending points to a map and running an analysis. The resulting table shows the most efficient ways to get from one place to another.

Use your knowledge of traffic patterns and impedances to decide which actual values to include in the OD matrix. For example, your matrix entries will reflect attributes such as speed limits for roads, maximum flow capacity for water in a pipeline network, or the amount of electrical resistance you might get from a particular type of wire in an electrical utility network.

Unidirectional paths in GIS show the direction movement occurs across a network, based on the steepest slope. This helps us understand how something like water moves and can help identify areas at risk of flooding or erosion. There are different ways to calculate these paths, and they're important for managing water resources and predicting water-related hazards. These types of paths are also used in networks used by people, such as motor traffic.

One-way streets are coded in a GIS by including a travel direction restriction in your OD matrix.

One other factor both characterizes the complexity of a network and adds to the robustness of traffic modeling capabilities. That characteristic, called circuitry, is based on the idea of closed loops. Closed loops allow moving objects, fluids, and so forth to travel alternate routes when moving along a network.



After you add to your GIS a complete network that has appropriate road designations and ID codes for your impedance values, the impedance values themselves, plus turn information, you have one of the most powerful networking tools available anywhere.

Many people use street maps to figure out the shortest route from their current location to some specified location. If you search for the shortest path, you use the length of the network links to decide the best route. Because you're not concerned with time, you don't need to include impedance values in your search. Although the GIS ignores most impedance values when you specify that they aren't needed, the software always includes some indication of one-way streets and dead ends so that it doesn't send you on a wild goose chase.

One common method of shortest path search is called the best first algorithm, which takes the shortest path available at each step. Other algorithms try a number of alternatives at predetermined distances and select the best.

The shortest path means shortest in terms of distance traveled. If you want to get there fast, use the fastest-route feature of your GIS.

The fastest-path search allows your GIS to start from one location and accumulate impedance values while it searches through the network to find your destination. Instead of keeping track of just distance, it combines distance and impedance to calculate the route that has the least amount of total impedance. Typically, GIS software measures impedance in time. One aspect of the fastest-path approach is a bit unique because your impedance values are based not just on speed limits, impedances, and typical traffic volumes, but also on the time of day.

A service area identifies places that are within a reasonable distance for use and relies on knowing how many people or houses are located along individual links in a network, much like impedance values, except that the database has to know how many people are along each link.

Service-area analysis allows you to use all the impedance and turning tools that you use for any other analysis. This process of finding service areas, often called allocation by geographers, is a very popular tool among business analysts because they cannot only allocate for their own facility, but also compare themselves to their competition. Allocation is one of the more popular economic placement tools available in GIS today.

Urban and regional planners found comparing different maps essential to the observation, quantification, explanation, and eventual exploitation of multiple patterns. Some of the patterns that occur help identify areas experiencing: thefts, busy shopping areas, where people live according to income, or soil types over a specified distance.

You can analyze the degree to which the patterns correspond, and figure out whether cause and effect relationships exist, much more easily if you can overlay the maps. Initially, overlay functions focused on comparing one set of polygons to another, but they quickly expanded to comparing polygons with point and line features.

All professional GIS software has some form of overlay toolkit that allows you to choose among different methods of overlay and select the layers you want to use. Most GIS toolkits have a separate graphical user interface or offer a button, icon, or tool to click that, in turn, gives you a lot of choices.

Overlaying in GIS compares:

* Points to polygons
* Lines to polygons
* Polygons to polygons

The first two pairings, points to polygons and lines to polygons, use a presence/absence overlay method. In other words, either the points or lines coincide with (occur in or run through) a particular type of polygon or they don’t. Polygon to polygon — introduces some additional potential methods of comparison that, to some degree or another, simulate the intellectual process of visual map overlay.

Creating a new map from your separate layers is important when you want to actually use that map for analysis.

The Point in Polygon (PIP) operation in GIS checks if a point is inside, outside, or on the boundary of a polygon. It works by:

* Defining the polygon's shape with a series of points
* Testing a point to see if it's inside or outside the polygon
* Using a "ray" to count how many times it crosses the polygon's boundary
* Returning the result: inside, outside, or on the boundary

The point-in-polygon operation is a powerful tool for comparing point objects to area objects, but you might not need to perform this operation if the points and polygons are already on the same map.

Line-on-polygon overlay in GIS combines line and polygon features to see how they relate to each other. It helps answer questions like: "Which lines are inside or outside a certain area?" or "What kind of land use is along a road?" This operation is useful for planning, analyzing networks, and studying the environment.

The most powerful, most robust, and probably most common types of overlay operations are those comparing one set of polygons to another. Generally known as polygon overlay, they offer many more options than point-in-polygon and line-on-polygon overlays.

Logical overlay is a method of comparing multiple maps that uses a group of operations based on set theory to search the polygons to determine whether attributes are shared from one map to the other (belong to a common set).

In set theory, three basic set operations compare the objects contained in one set to those of another set. These operations are union (where you combine all the stuff from both sets), intersection (where you select only those things common to both sets), and complement or symmetrical difference (where you identify all the objects that the sets don't have in common).

Union overlay, sometimes called an OR search, collects all the polygons that have any of the attribute search criteria and makes a new map out of them. Union overlay maintains the categories from each of the input layers, so you can use them later to perform still more comparisons of flat agricultural land with other layers such as soil nutrient values.

Use union overlay when you want to broaden your search and combine different attributes of your polygons. All the polygons that have any of the search criteria are included in the output.

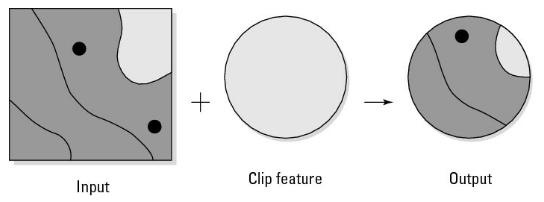
Use intersection overlay when you want to narrow your search to satisfy a number of criteria at the same time. This technique can help you find the most appropriate places for some activity.

Sometimes, when you try to perform a search, you may find it easier to eliminate things that don't match all your criteria, instead of identifying those that do match. This search, called complement or symmetrical difference overlay, computes the geometric intersections of all polygon types and categories that don't have certain attributes in common.

Identity overlay is similar to intersection overlay, but it treats the layers differently. One layer, the identity layer, must be composed of polygons because its attributes determine the identity of everything that coincides with it. But for the other input layer, you can use a point layer (such as wells) or a line layer (such as roads) if you want. When your GIS overlays the two layers, the new shape takes on the attributes of the identity layer. In other words, it takes on the identity of that layer.

Use identity overlay when you want to convert intersecting portions of one layer to the attributes of a more important layer. The more important layer is the layer that often controls what values will be assigned to the new polygons, and you usually determine which layer is more important based on what you want to change and what you want to stay the same after overlay.

In GIS, a clip overlay is a tool that combines two datasets to create a new one. It uses the boundaries of one dataset to "clip" or extract specific areas from the other dataset, creating a new dataset with only the overlapping areas.



Use the clip overlay when you want two or more layers to have the same geometry (size and shape), which allows you to have all the attributes that you want to evaluate, but for a selected portion of the study area.

Map algebra is primarily a raster GIS modeling language that allows a huge amount of flexibility in comparing one map to another. Raster overlay is not only more powerful than its vector equivalent, but also less troublesome.

In GIS, an overlay combines multiple data layers to see how they relate to each other. There are different types of overlays, including ones that show where layers overlap, don't overlap, or combine all areas. Selective overlay lets you choose which areas to combine based on specific criteria.

Use selective overlay when you have specific combinations that you want to examine but don't want to create a separate overlay for each.

Cartographic modeling is an ordered set of map operations designed to simulate some spatial decision-making (making decisions about geographic space) process. The order or sequence of map operations (including creating, combining, and analyzing) is often critical to the success of a cartographic model, and the idea that the result will simulate a spatial decision process is at the heart of a cartographic model.

People use GIS (cartographic) models only as a spatial decision-making support system. This single application seems to limit the utility of cartographic models, so I like to give the whole idea of decision-making a broader definition. By decision-making process, I mean any process that simulates physical or human environments and supports a wide host of possible activities, including

* Immediate action: Activities such as planning, where you decide to immediately employ cartographic cartographic modeling for practical resolution of real-world problems.
* Theoretical studies: Constructing cartographic models that simulate natural or human activities in geographic space, contributing to the body of knowledge by describing how spatial distributions of geographic features relate to other patterns or processes.
* Future planning: Information needed for decision-making support can be exploratory (for example, predicting traffic increases at selected locations over the next ten years) and can often lead to eventual decisions and related actions.

Map Algebra is a way to perform mathematical operations on geographic data (specifically raster data) in a GIS, allowing you to combine, manipulate, and analyze data layers using simple algebra-like expressions. It's like doing math with maps!

Map algebra was developed when GIS software typically used the command-line interface, where you actually typed words to get the software to perform its tasks rather than click icons and pull-down menus. Although today's GIS typically employs a graphical user interface, map algebra functionality is still tied to the idea of written commands.

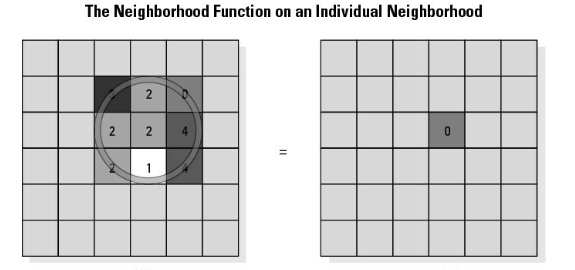
Although these map algebra commands used to be part of a command-line system, they exist as part of the graphical graphical user interface of the modern GIS. And so, you work through the graphical interface to apply the map algebra and make your cartographic models.

Essentially, an interface such as the Raster Calculator provides a means to write out complete and often very sophisticated map algebra expressions based on built-in mathematical capabilities.

The most basic function type is called the local function because it operates on each cell individually. Sometimes called by-cell operations, these functions are among the simplest — and most often used — of the map algebra functions.

Unlike local functions (discussed in the preceding section) — where you zoom in to gain the worm's eye view — focal functions allow you to pull back a bit and begin to see the area around you. This area is sometimes called a neighborhood. Focal functions evaluate a single grid cell based on some characteristic of its surrounding cells, or neighborhood. (Focal functions are sometimes called neighborhood functions.) In general, the focal function allows you to look at a target cell (the cell that you're characterizing or working on), evaluate it based on some relationship it shares with its neighborhood, and return new grid cell value in the same X and Y coordinates but on a new output layer.

To create a simple search pattern, search for the eight neighboring cells around your target cell in a 3-x-3 matrix. That configuration is called an annulus (donut shape) with a search radius of one cell.



Focal Flow is a GIS tool that analyzes the neighboring cells of a raster to determine which ones "flow" into the central cell, based on their values. It helps identify the direction of movement or influence between nearby cells. Think of it like tracing the path water would flow across a landscape!

The adjacent 8 cells around a focal cell are often referred to as the immediate neighborhood.

Zonal functions don't create neighborhoods. They use zones, either from a single grid or from other layers to compare cells by either attribute (description) or shape (geometry). You need to define the zone with which you want to work. A zone is equivalent to what geographers call formal regions. Regions are areas or groups of areas that share common descriptive information. In a raster data model, a region (or zone) is a group of grid cells that have the same attributes.

Formal geographic regions can be contiguous (all in one chunk), perforated (with holes that have different attributes), or fragmented (like islands that share attributes).

In raster GIS, area is calculated by adding up the sizes of all the grid cells included in the zone.

In GIS, "orthogonal" means moving or measuring in a straight line that is perpendicular (at a 90-degree angle) to a reference direction or feature. It's like moving directly across from something, rather than at an angle. Orthogonal directions are used to define spatial relationships, like the direction of water flow or the orientation of a road, and are important in spatial analysis and modeling operations.

In raster, polygon means a collection of grid cells of the same value or attribute.

Sometimes, you need to resample a set of grid cells to a coarser group of grids (bigger grid cells) so that the set matches other raster layers.

The block function takes a uniform, non-overlapping block of grid cell values and changes them based on any of the following operators: mean, majority, maximum, minimum, median, minority, range, standard deviation, sum, and variety — the same operators that are available for zonal functions.

A global function takes the bird's-eye view — it can see and operate on the entire study area at the same time. Global functions are very powerful and complex operations. Some examples include:

* Distance measures: Such as Euclidean distance, Manhattan distance
* Surface functions: Such as finding basins, pour points, and drainage networks
* Interpolation functions: Such as linear, nonlinear, trend surface, and exact
* Hydrology functions: Such as water accumulation and flow direction

Groundwater global functions in GIS are advanced tools that help model and analyze groundwater flow and movement. They're mainly used with raster data and enable you to:

* Simulate liquid flow through different materials and layers
* Account for varying substrate thickness and pressure gradients
* Track the movement of dissolved solids and pollutants
* Model both point and non-point source pollution

These functions are powerful and complex, allowing for detailed analysis and simulation of groundwater systems.

Ordered operations are operations done in a logical sequence, and with GIS this is called a model or cartographic model (since it involves maps).

The key to good cartographic modeling is to formulate the model through the use of flowcharting techniques. Therefore, you should normally create the flowcharts even before you know what's in your database. Such formulations have helped people decide what should go into a model in the first place.

A formulation flowchart starts with the final desired product (the spatial information product), breaks it up into sub-models, and finally outlines the actual map elements that you need.

Allocation, which also works with networks, involves finding all the areas within your sphere of influence, or your territory.

Always take the time to develop a formulation flowchart of your cartographic model that includes everything you want to see the model contain without worrying about whether the data are available.

One solution to data issues in cartographic modeling is to find data that you can substitute for the data that you really need. Such spatial surrogates (replacement data) might include housing in place of population, power lines for missing roads, and fire patterns to indicate wind direction and intensity.

No kind of modeling is an exact science. You might want to include a complete explanation of the methods employed to create the output so that the user knows how reliable the model is.

A true test of modeling often requires a truth set (a known correct answer) against which you test the output.

When you need to figure out whether the software algorithms are working correctly, you can most easily perform an analysis on a very small part of the database and check it against a manual computation of the same area.

Adjusting model constraints takes finesse; consult the user when you make adjustments so that you don't violate the spirit or purpose of the model.

Make sure that you always keep the lines of communication open with the users. If you keep the users involved throughout the process, they're much more likely to accept the final product. A correct model is a useful model only if the user accepts it. Keep the user involved in every phase of your GIS work.

Traditional maps are the most common forms of maps and generally the most easily understood by the general public. Reading and interpreting maps is much easier if the scales, symbology, classes, and graphic design elements are all focused on visual communication.

Five characteristics of good map design are:

* Make sure the map meets user needs
* Make the map easy to understand and use
* Accuracy is essential to good map design
* Map presentation should relate to the data
* Allow the user to review and interact with the map during production

Sometimes, when you create a map, you map qualitative (descriptive or nominal) data, which generally requires that you pick unique symbols for each category.

Most GIS software allows you to use a set of predesigned color or shading schemes, or to create and employ your own. When working in black and white (for publications that don't use color, for example), you can use the shading patterns rather than colors.

Most non-cartographers don't really know how to decide what specific data to map or how to display them. Here are three general rules to help out:

* If you want to see the raw data and relative comparisons, map the counts or amounts themselves
* If you want to control for area or number of features (called normalizing your data), use ratios
* If you want to compare the relative values from place to place or feature to feature, you want to map those values as ranks

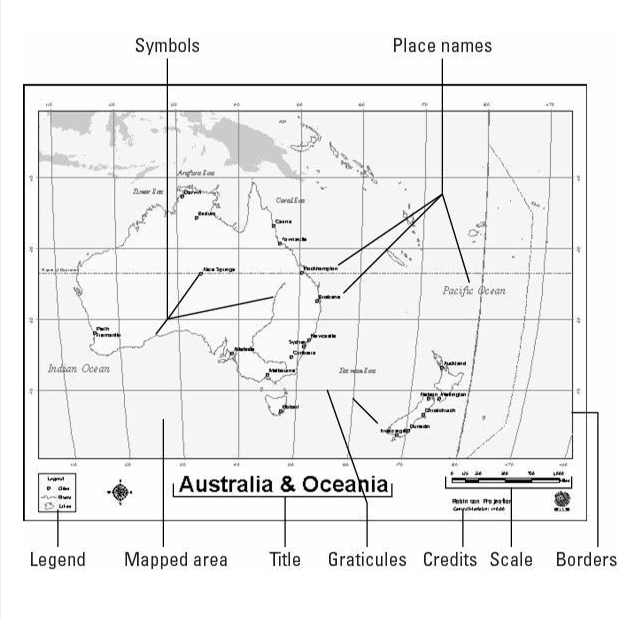
If you have only a few items to map (normally, less than ten), you can display each one with a separate symbol or color. If you have more than ten items (which you usually do), you need to group those items. Otherwise, the reader can't visually separate the many shades and colors needed to display them.

You typically find seven general ways to classify data available in GIS software.

1. Equal Interval
2. Defined Interval
3. Quantiles
4. Natural Breaks
5. Geometrical Intervals
6. Standard Deviations
7. Area-based

Standard map elements include:

1. Title
2. Legend
3. Scale
4. Credits
5. Mapped Area
6. Graticules
7. Borders
8. Symbols
9. Place Names



Another aspect of cartography, called graphic map design, borrows heavily from traditional art and graphic design principles. Here are some considerations:

* Does the map look balanced
* Is the lettering readable
* Do features stand out
* Does the reader know what to focus on
* Does the map make the reader feel dizzy

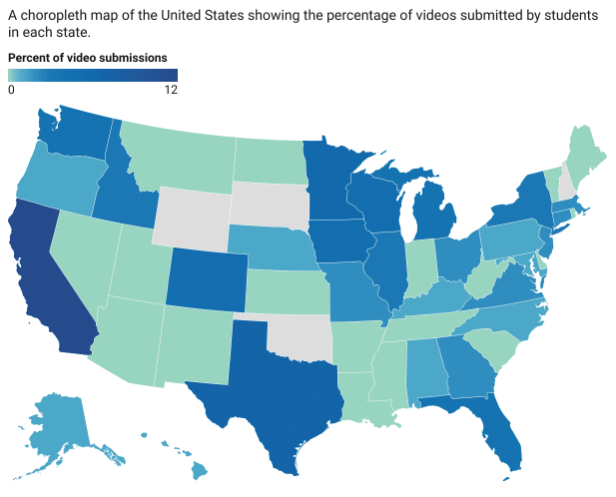
In graphical design, figure-ground refers to the relationship between the main subject (figure) and the background (ground). Effective use of figure-ground principles creates contrast, guides attention, and enhances legibility, making designs more engaging and effective.

Visual hierarchy is a design principle that organizes elements by importance, guiding the viewer's attention and understanding. It's created through size, color, contrast, alignment, and other visual elements, helping designers create clear and effective designs.

A cartogram is a map that exaggerates the size of geographic space based on the descriptive attributes, instead of presenting the space in its accurately scaled size and shape. You use a cartogram to communicate a specific aspect of the mapped information more effectively. You must be aware of the target audience's map reading skills and understanding of the portrayed geography before attempting to use cartograms.

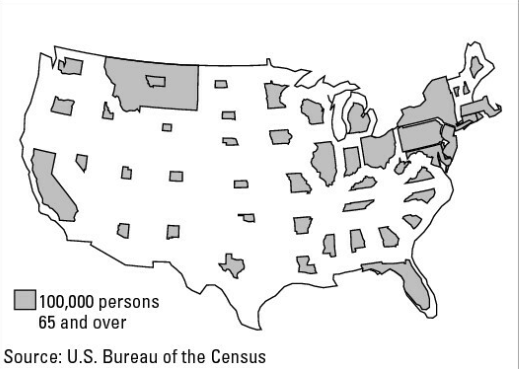
Area cartograms, often called value-by-area maps, distort the sizes of the polygons based on the attribute quantity, rather than its actual physical size.

Choropleth map: a map presenting data on an area-by-area basis.



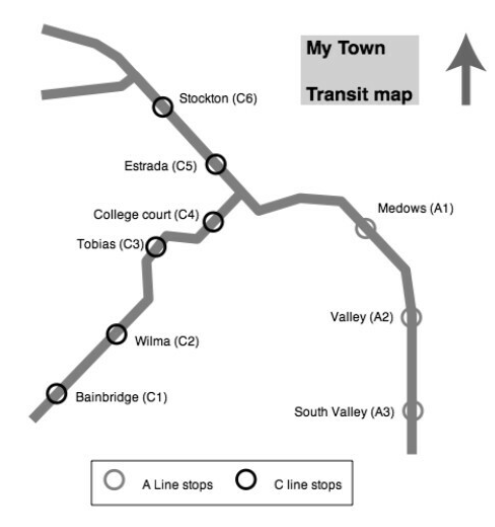
A contiguous cartogram is a type of map that resizes regions to represent a specific data value (e.g. population) while maintaining their adjacency and overall shape.

A non-contiguous cartogram is a map that resizes regions to represent data, but doesn't preserve their geographic connections, often separating or merging them for visual clarity.



Linear cartograms distort distances to make their point.

Perhaps the most used, but often unrecognized, linear cartogram is a routed line cartogram. Busses and mass transit trains display this kind of cartogram to illustrate the stops in their proper sequence.



GIS produces two general types of reports that are suited to business uses:

* Tabular reports: Have a spreadsheet-like layout
* Columnar reports: Have a newspaper column layout

Typically, you can more easily understand summary statistics when they're presented graphically through bar charts, histograms, line graphs, pie charts, scatterplots, and other forms of non-cartographic graphic output. Most GIS software gives you the option of including both tables and graphs directly on the map itself.

You can more easily read and interpret maps if you include a linked table or graph either on or near the map itself. If you use matching colors to show the correspondence between the map symbols and the tabular or graphed data, you can also improve users' understanding.

GIS animations are a form of geographic visualization that shows dynamic changes to a single map layer or even multiple map layers simultaneously to allow visual comparisons. Animations show changes — which might take days or weeks to occur — at a fast pace so that you can see trends you might otherwise miss. Alternatively, if you have a lot of data about changes that occur very quickly, you can use animation to present those changes slowly and, again, identify changes you might not otherwise catch.

The flythrough is a form of animation that happens in three dimensions. Just like the term says, the flythrough allows you to visually fly through a three-dimensional version of your map.

Some different types of professional GIS users are:

* System users: Use GIS to solve problems, provide solutions, and supply decision support for the organization
* System operators: Know the inner workings of the software, can support the system users through troubleshooting and workarounds, and have specialized knowledge about the GIS
* System sponsor: Provides the financial support for the entire system
* GIS supplier: The GIS vendor that supplies the software (and sometimes hardware), technical support, and training
* Data supplier: Two types of data suppliers exist, public and private
* Applications developers: create GIS models and user interfaces
* GIS systems analyst: Systems analysts look at the overall operation of your GIS to ensure its initial and long-term successful integration into the organization

Many GIS designers speak of the spatial information product (SIP) — the planned output from GIS analysis. In most private and commercial business operations, the SIP is a product that provides income to the company. Some commercial companies produce output (SIPs) that they make available for purchase, and others produce output for some other organization that has limited GIS capabilities.

In some cases, simply having the system in place for a long time might satisfy a major need that employees have for flexible access and opportunities for experimentation and modeling.

Some governmental entities that use GIS:

* Management agencies
* Medical agencies
* Meteorological and climate agencies
* Emergency management agencies
* Military and military support organizations

Some non-profit entities that use GIS:

* Relief organizations
* Conservation groups
* University departments
* Philanthropic organizations
* College admissions departments

Some private and commercial entities that use GIS:

* Mapping companies
* Consulting firms
* GIS data providers
* Economic placement companies
* Real estate companies
* Vineyards

To see specific examples of how each of the above entities may use GIS, conduct an online search. Each organization is unique. Each has its own culture, organizational structure, history and traditions, needs and mandates, and vision. These differences have a profound impact on whether the organization can effectively use GIS, or any technology, and whether it can easily integrate that technology into the existing operations. So, you need a thorough knowledge of the structure of an organization — including how decisions flow from place to place, who has what job responsibilities, and how general operations work — to be able to incorporate GIS effectively into the organization.

For any organization utilizing GIS, it is important to continually update organizational diagrams to properly track those internal departments using GIS in some manner. That way, it is easy to track who would be impacted by procedural changes or modifications to systems in place.

Additionally, it is important for any organization using GIS to have a firm training plan in place so that any change in procedure, or new procedure, can be implemented equally across affected departments.

One of the biggest issues with incorporating GIS into organizations is that it introduces a massive increase in technological dependence. This new technology causes changes in how work is done and in the expectations of increased productivity. And both technology and change are accelerating.

Most organizations that begin to incorporate GIS benefit from an initial needs analysis. There are four areas to look at in how it could provide a benefit to an organization:

1. Application design: Requires that you know exactly what maps, graphics, and other decision-support products you want the software to produce for you.
2. Analytical model design: Know what analytical techniques you need your GIS to perform. These techniques differ for ecological, economic, military, and any other type of analysis.
3. Database design: Various software packages use different data models. The application design needs and the analytical model needs you define for your GIS work together to determine your database needs.
4. Institutional/system design: When you implement GIS, you need to consider the people in an organization. You need to answer questions regarding personnel, equipment, space, organizational structure, enterprise-wide goals and objectives, and costs and benefits. Costs and benefits, in particular, often determine whether a GIS is even viable.

Many organizations just starting out with GIS prefer to have a system developed from the ground up so that the design document they receive allows them to create an RFP (request for proposals), which gets GIS companies to beg for your business.

You need to assess the costs and benefits of a GIS implementation if the GIS is designed to bring real financial benefits to the organization. When you perform a cost-benefit analysis (CBA), you determine whether something has more benefit than expense, or vice versa.

A common business practice is to look at the benefit-to-cost ratio for any large investment, including software tools like GIS. In some businesses I've worked with, a ratio of 1.6 to 1 would be a minimum standard needed to justify the expenditure.

In most cases, the first year's costs outweigh the benefits because the initial outlays involve staffing, planning, setup, automation, application development, and training. One pragmatic approach to CBA for GIS implementation is to use a six-year planning horizon.

Costs of implementing a GIS are often:

* Hardware
* Data
* Software
* Application-level development
* Staffing
* Training
* Maintenance

Benefits from GIS are sometimes difficult to identify, usually these come in four groups:

* Direct- improvements to manual procedures that came before GIS implementation
* Agency- increases to company productivity and product quality
* Government- This benefit also applies to partnerships that result in reduced costs or shared resources
* External- Improvements in the quality and quantity of external factors, perhaps including the impact of the products, improved public perception of your company, and the regional benefits of increased tourism.

You can prevent the slow decay in GIS functionality by putting in place a good system of change detection. Technological and institutional changes can both affect how well the GIS system continues to perform for the organization long-term.

Photogrammetry: the science of exact measurement of features on aerial photography.

Potential questions to ask vendors when purchasing a GIS system:

* What services do you offer?
* How will this product meet my needs?
* What data formats are supported?
* How do you handle communication and change requests?
* What hardware expertise do you have?
* What does the price include?
* How long until the system is operational?
* What happens if the system crashes?
* What are your quality control procedures?

Depending on your application and your specific needs, you may want to develop your own datasets. Developing your own datasets allows you to have total control over the content and accuracy, not to mention the ownership, of your data.