CHAPTER 31

CARTOGRAPHY AND VISUALIZATION

Robert B. McMaster and Ian Muehlenhaus

31.1. Introduction to basic cartography

Once you have spent considerable time collecting, verifying, and editing your spatial data, your next step is to craft a powerful map that will clearly and succinctly communicate the patterns and idiosyncrasies of the dataset. This chapter will review the processes you must go through in order to do this.

Traditionally cartography has been defined as both the “art and science” of creating maps. However, this binary is an artificial one. It presupposes that art and science are independent and separable components in mapmaking. In reality, cartography is best thought of as a “craft” combining knowledge from both graphic design theory and mathematics. In order to become a stellar cartographer, one needs to hone, expand, and synthesize their skills in both of these realms – they are not separable. Thus, before we continue, we recommend you consider yourself an artisan – cartography is something that you will only master via continual education and practice. This chapter will present you with a basic overview of the rudimentary information necessary to design an intelligible map and to jumpstart, or refresh, your knowledge. Throughout the text we provide references to more nuanced readings concerning the topics covered here.

Cartography is thousands of years old. In fact, maps predate written language as a form of communication. For millennia, cartography has been used to enhance our understanding of the Earth – its shape, distances, directions, and areas. In fact, flattening the spherical Earth into a multitude of two dimensional shapes that preserve one or more of the aforementioned attributes was first done by the Greeks. Though this chapter does not attempt to provide an historical overview of the cartographic discipline, it should be noted that today’s marvelous technologies that allow us to accurately depict locations within centimeters are the culmination of thousands of years of cartographic development. We are but the tip of the iceberg, and as with any knowledge, the tip is dependent upon its base to stay above the water.
31.2. Map Purpose

The first thing one must do as a cartographer is identify the purpose of your map. Before one begins symbolizing any data, before even laying the map out on the page (or more likely computer screen), take a moment and jot down who you expect will be reading the map – the intended audience – and what the purpose of the map is. Though this may seem rudimentary, it is not. In fact, having this written down when you begin designing the map will save you much consternation later on, when we begin generalizing and symbolizing. The intended audience and map purpose must guide the map creation in order for you to communicate effectively.

Map types can be broken down into two broad categories – general- and special-purpose maps. General-purpose maps are those that provide a geographical base and have a functional rationale. For example, a city map would be considered “general purpose.” You can use the city map to find directions from your house to a local rummage sale, or you can utilize it to determine distances between your house and the nearest fire department. You cannot use the map to gauge the median incomes of different neighborhoods nor the odds that your backyard is sitting atop a radon hotspot. The purpose of such maps is to provide people reference data of some sort. A map of the world may show the location of all countries, and perhaps even different countries’ flags, which is great for referencing. However, it will not allow you to compare different countries’ stockpiles of uranium. Some of the most detailed general purpose maps are topographic maps, which depict the basic location of features on the surface of the earth, including transportation, hydrography, cultural features, and elevation. Topographic maps are like “geographic dictionaries.” They emphasize geographic position and location, as well as planimetric accuracy.

However, not all maps are primarily concerned with position, location, and the accuracy of base map data. Indeed, special-purpose maps are interested in showing the distribution of a theme, such as characteristics of a population or land use. Beginning in the nineteenth century, cartographers, geographers, and statisticians began developing different methods of visualizing thematic data over base maps, including the choropleth, graduated symbol, dot, isarithmic, and daysmetric techniques (among several others). This was a revolutionary idea – to spatially illustrate data that had previously only existed in chart form. However, the usefulness of these techniques quickly came apparent, particularly for state governments attempting to make sense of the increasing amounts of data they were collecting about their populations and territories.1
As opposed to “dictionaries,” special-purpose maps might best be considered “geographical essays.” Like a good snapshot, special-purpose visualizations can quickly tell a story without the need for words. By taking tabular data and illustrating it over a map via location, an audience is able to interpret the data more quickly and readily than if they were to look at numbers. However, a special-purpose map’s readability and communicative power as an essay is only as good as its articulation; this is where cartographic design becomes extremely important. The rest of this chapter will review the five key realms of thematic (special-purpose) map design that you should pay particular attention to when designing maps – scale, generalization, symbolization, data classification, and map design. Notice that these five realms combine both design and math (art and science). In order to become a proficient and successful thematic cartographer, you must practice the craft of melding these four realms together seamlessly.

31.3. Cartographic Scale

The reason we project the Earth onto a flat piece of paper is to look at a particular part of the Earth’s surface in greater detail – something that is impossible to do with a globe due to size constraints. Thus, the first thing any cartographer must do is decide at what scale they want to present the data. Scale is simply the mathematical relationship between the map distance and the commensurate earth distance. This relationship is frequently represented via representative fraction (e.g., 1:500,000 where one unit of distance on a map equals 500,000 of those units in the real world). Map scales are often broken down into large-scale (e.g., ~1:10,000) or small-scale (e.g., ~1:1,000,000). Large scaled maps will contain less mapped area; since the scale is large relative to the size of the Earth, such a map will be zoomed in on the Earth’s surface. This is for the best if you hope to show a lot of detail on the ground (e.g., create a topographic map). However, this is not useful if you hope to show spatial patterns and themes across a wider area of land. In this case, you will want to use a small scale map (small because its representation is tiny compared to the real proportions of the Earth). Small scale maps show more area of the Earth allowing for a larger realm of comparison, but they do so at an expense – the level of detail small scale maps provide may be unacceptable for the purposes of location and distance evaluation or any other general-purpose uses.

To determine the scale at which you create your map, you must take into account three things. First, recall the purpose of the map – which you have written down somewhere hopefully.
Second, think about the intended audience – do they know an area well enough to be able to interpret an extremely large-scale map without any broader reference? Will the scale you choose include all areas affecting the theme of your map and of interest to your audience? (The last thing a cartographer wants to do is incense their audience by zooming in so far that they crop some culturally sensitive or relevant areas off the map.) Once you have determined the scale at which you would like to create your map, you must determine whether you have the appropriate data for mapping at that scale. For example, if you want to make a population map of the ethnic Hungarian diaspora in Central Europe, you may want to zoom in on Hungary and the Romanian Carpathians. Perhaps you envision mapping the percentage of ethnic Hungarians residing in the Transylvanian part of Romania by province. Yet, when you go to look at your data, all you have are the percentage of Hungarians residing in each country (e.g., Romania equals 1.4 million). Thus, you must create a map at the scale of the data – a map that shows both Hungary and Romania in their entirety. Such a representation would be completely ridiculous for the purpose of the map; so perhaps you are better off mapping the Hungarian minority in Slovakia and Serbia as well, zooming out even further to show ethnic-Hungarian populations throughout Central Europe.

31.3.1. The Scalar Dependence of Data Geometries

Geographic phenomena take the form of four different geometric shapes – points, lines, polygons, and volumes. Spatial features in the real world can be represented using these geometries. However, determining which of these geometric features to use will change depending on the scale of your map. Geometry is scale dependent! For example, a country house shown on a large scale map of a French village (e.g., 1:3000) should be drawn as a polygon – a building footprint. Yet, the same house on a map showing a section the French province (e.g., 1:50,000) would be represented by a point. One problem with GISs today is that they do not change the geometry of geographic data automatically depending on the scale of the map. Thus, point data remains point data no matter how far one zooms in on the village. Vice versa, point data remains point data no matter how far one zooms out, even to the global scale. In fact, if the scale makes showing an individual house difficult, and if the house in question is not integral to the purpose of the map, the house should be generalized off the map entirely.

31.4. Database & Cartographic Generalization
The scale of your map will dictate to what extent your map will be generalized; the geometry of your data will determine what types of generalization can be used. The key thing to note here is that every map is generalized – since a map is merely a simplified representation and model of reality, some aspects of reality (typically many) must be trimmed or cut from the representation. Generalization is the simplification, de-emphasis, exaggeration, and enhancement of certain spatial data to make the message of your map clear. The level of generalization necessary is always affected by a map’s scale. Unfortunately, many cartographers do not practice extreme prejudice when generalizing and attempt to cram too much data and detail on a map. Thus, the first thing any cartographer should do before they begin to generalize is again ask – what is the purpose of the map and who is the map audience? If the purpose of the thematic map is merely to show broader spatial patterns, then the accuracy of the base map itself may not be of primary concern, but data accuracy will be important. If the purpose is to show exactly where certain thematic variables occurred, but not necessarily quantify the intensity of these variables, then the data may be simplified but the base map must be accurate enough to show relatively precise location. As for the audience, if a map audience is comprised of mature map readers who are familiar with the data being mapped, perhaps less generalization will be more useful for interpretation. However, if a map audience is comprised of the general public or a younger audience that is less familiar with the data you are mapping, generalization is of paramount importance to declutter and emphasize the thematic data.

Map generalization can be broken down into two types – database generalization and cartographic generalization. The first method of generalization involves the filtering of cartographic information inside the database itself, and does not include any operation to improve graphical clarity. This process is never visual but occurs through the application of internal database operations. Perhaps, for instance, one has a database that was created at a scale of 1:25,000. However, for the purposes of the mapping project, this is too detailed. Thus, one might use database generalization to regenerate and filter the same data for a smaller scale, say 1:50,000. This operation would be conducted within the database itself.

The second method of generalization is extremely visual. Cartographic generalization is the process of modifying map features to make them aesthetically pleasing and clear at a reduced scale of representation. McMaster and Shea (1992) have defined cartographic generalization as
“the process of deriving, from a data source, a symbolically or digitally-encoded cartographic data set through the application of spatial and attribute transformations.” Central to cartographic generalization are the map purpose and intended audience. Obviously, if you are designing a map for airplane pilots, too much generalization might be a problem. On the other hand, if you are designing a map for school children, including all of the geographic features found on an aeronautical chart would make the map overly complex and take away from the message you are trying to communicate.

There is no such thing as “perfect” or even “optimal” generalization. However, there are six fundamental philosophical objectives that should inform and drive all types of generalization:

1. Reducing the complexity of the map/database;
2. Maintaining the spatial accuracy of map data;
3. Maintaining the attribute accuracy of map data;
4. Maintaining the aesthetic quality of the map;
5. Maintaining a logical hierarchy of map elements; and
6. Consistently applying the rules of generalization.

There are 12 fundamental transformations that can be used to generalize the visualization of spatial data on a map (please see Figure 31.1, which illustrates these generalization operations and shows the result of each type of generalization at a 50 percent scale reduction).

Place Figure 31.1 about here

*Simplification* involves the elimination of unnecessary information. *Line smoothing* shifts and eliminates nodes along the path to improve aesthetic quality. This can be done manually or via mathematical splining and averaging methods (McMaster, 1989). *Merging* is crucial for conglomerating two parallel lines at a reduced scale into one (e.g., two banks of a river into one line representing the entire river). *Exaggeration* and *enhancement* are both operators that actually accentuate certain features in order to preserve their significance under scale reduction. This may be important if a bay or other detailed feature such as the shape of a building’s courtyard becomes accidentally simplified or lost due to scale change. Related to exaggeration is the
critical operation of displacement, where features are purposefully shifted apart to prevent coalescence under scale reduction. The aforementioned operationalizations are primarily for line features, but others exist for point and areal features as well. *Aggregation* fuses multiple point features together (e.g., individual trees) to create an areal feature (e.g., forest). Along the same lines, *amalgamation* fuses several smaller polygons into one larger polygon. Whereas the *collapse* operator does the opposite, reducing areas into point features. *Refinement* and *typification* are operations that create a smaller-scale symbolic representation of a set of features.

Figure 31.1. The processes of generalization.
In light of the complexity of taking into account all six of the above keys to successful
generalization, it comes as no surprise that in recent decades numerous attempts have been made
to automate generalization within GISs themselves. This is a tall order for a computer program,
particularly on the aesthetic front, and there is still much progress to be made before these
operations will be efficient and effective in GISs.

31.5. The Measurement Level of Spatial Data

To determine how best to symbolize our spatial data – both via geometric symbolization (point,
line, or polygon) and graphic manipulation (manipulation of the visual variables, to be discussed
in the next section) – we must first analyze the measurement level of the data. Cartographic data
can be categorized into the four established measurement levels of nominal, ordinal, interval,
and ratio.

At their most basic, data can be measured at the nominal level. Nominal data is merely the
classification of data as something – labeling. Data measured at the nominal level are completely
qualitative, meaning they cannot be compared to one another mathematically. Examples of
nominal level data include land-use data (e.g., recreational, housing, commercial, industrial),
race (White, African-American, Asian), or ground coverage (forest, grassland, human-made).
Like comparing apples and oranges, nominal-level data are unique in that all classes lie at the
same level and cannot be compared to one another based on their own inherent attributes. This
fact is crucial to remember when symbolizing this type of data, as one must not accidentally
impart value or rank on nominal attributes – we will discuss this further in the next section.

The next level of measurement, ordinal, involves a ranking (or “order”). Oridinal-level data
allow for differentiation between higher and lower levels. There is a presumed order in the data,
but the level of difference between these rankings is not measurable. Examples of ordinal level
data would include a map showing the location and rating of hotels – from one-star to five-star. It
is impossible to quantify the difference between a one-star hotel and a five-star hotel, but most
people would inherently choose a five-star hotel over the alternative (finances permitting, of
course).

Interval- and ratio-level data are similar in nature and separable from other levels of
measurement due to the fact that they are metric; i.e., interval and ratio levels involve true
numbers that can be used for the purposes of statistical measurement and comparison. The difference between these two levels of measurement is dependent upon the meaning of “zero.” Interval-level data do not have a true, absolute zero-point. Ratio-level data do. Interval-level data are generally found when dealing with raw measurements socially constructed by humans. Ratio-level data include any type of derived data (e.g., any time there is a percentage, there is an absolute zero), as well as data that cannot be negative. These helpful guides do not make determining the difference between interval- and ratio-level data any easier, though. Take, for example, elevation. This concept would seem to have an absolute zero. Moreover, it would seem to be completely natural – not a component of human contrivance. We can define elevation as being associated with intervals, not absolutes. Where is the zero? Sea-level. What is sea-level? Why did we decide that sea-level would be zero. It is an arbitrary zero. Other examples of interval data are temperature (excluding Kelvin) and time (unless you calculate your workday based on the Big Bang). As already mentioned, ratio-level data includes any derived numbers (which are “rates”), as well as any other type of data with an absolute zero. Age is ratio data. You cannot be negative years old. Precipitation is also ratio data; though, we arbitrarily choose how to measure rainfall (inches or centimeters), you cannot have negative rainfall.

These four levels of measurement can be broken down into two broader data types – qualitative and quantitative data. Nominal-level data are qualitative; they cannot be compared to one another. The three other types of data-levels are quantitative; some type of statistical comparison is possible. Before you begin symbolizing your data, it is crucial that you analyze the level of measurement of the thematic data you are mapping. As will be illustrated in the next several sections, if you fail to do so there is a good chance that you will use an incorrect method of symbolization and confuse your map reader. Thus, please add the level of measurement to the list where you wrote down the map purpose and intended audience. It will come in handy.

31.6. Fundamental Visual Variables

Unlike natural phenomena, which can occur volumetrically, paper and computer screens constrain our ability to represent spatial data in three-dimensional form. (Though, as will be discussed, virtual and Web 2.0 technologies are changing our ability to “visualize” and interact with the third and fourth dimensions.) Thus, cartographers have traditionally identified three
types of cartographic symbols – point, line, and area symbols. (Volumes are generally represented with skewed area symbols appearing three dimensional to the human eye.)

The use of visual variables allows us to “design” and manipulate the geometric properties of points, lines, and polygons to represent certain data variables. For well over 100 years, cartographers have been adapting and manipulating different visual variables to create thematic maps. Only in the past few decades, however, did cartographers begin to systematically categorize the different visual variable techniques that can be used, and associate these techniques with different types of data-levels. Bertin is recognized as the first person to point out the prime visual variables with his influential book *The Semiology of Graphics* (1983). Many others have added to Bertin’s original visual variables and modified his original definitions to better suit cartographic endeavors. The visual variables reviewed here and in Figure 32.2 are contrived from the work of DiBiase et al. (1992) and Muehrcke (1983).

![Figure 32.2. The visual variables](image)

The key thing to note is that depending on what level of measurement your data represents, certain visual variables should never be used! Using our analogy of maps as communication,
geometric map symbols are best thought of as the words comprising the graphic message; visual variables act as the grammatical framework structuring the words. For example:

“The grammatical does fails not map of make if just visual to as established follow a use one the and order communicator rules will make one if does structure use variables English as sense correctly not.”

“Just as English does not make sense if the communicator does not follow the established rules of grammatical structure, a map will not make sense if one fails to use and order visual variables correctly.”

The first indented quote makes no sense to anyone. Yet, it contains the exact same wording as the quotation following it. The same kind of confusion confronts map readers when a cartographer uses cognitively inappropriate visual variables and symbolization for the data they are attempting to represent!

Though we often think of literacy as only concerning spoken and written languages, from birth humans also learn graphical literacy (Dondis 1973). Though we can all read graphics, we are not natural communicators. Few of us are lucky enough to have formal training in graphic design. Thus, we need training and practice to effectively harness the grammar – or rules of visual variables and hierarchy – to effectively communicate spatial data. MacEachren (1995) and Slocum et al. (2008) both provide great overviews of how and when to effectively utilize different visual variables. The key thing to note is that the type of visual variable you use, regardless of the geometry of your symbol, must correlate with the data’s level of measurement, otherwise confusion is likely to result.
Figure 31.3. Geographic data type, measurement scale, cartographic symbol, and the qualitative and quantitative visual variables.

31.8. Choosing a Representation for Your Thematic Map

Once you have determined the geometric properties you will use to map your spatial data – point, line, or polygon – your data’s level of measurement, and you understand what type of visual variables work with this data, you must choose how to represent these data on your map (Figure 31.3). How does a cartographer choose what type of thematic map to create?

Cartographers have been researching the effectiveness of different map types for well over 50 years. Different physiological tests have resulted in a variety of results, some more conclusive than others. Perhaps one of the simplest and most effective models, however, came from the cognitive field. MacEachren and DiBiase (1991) argue that spatial phenomena can be arrayed along two axes: continuous-discrete and abrupt-smooth. Depending on the spatial nature of the phenomena for which we have collected data, some methods of symbolization and visual variable manipulation are more appropriate than others. Not only are they more appropriate, they will better communicate the data. The reason for this is that the representation of the data on the map will closely mimic the distribution of the measured phenomena in the real world.

An example of the typology of mappable phenomena is provided in Figure 31.4. Each of the boxes provides an indication of the typical data type found along the continuum as well as a specific symbolization type (below the box). For instance, the number of government employees is both abrupt and discrete and would require a graduated symbol. Government employees are abrupt, because the number of employees is contingent upon strictly demarcated political
boundaries. This data is discrete because the samples occur in specific places, there are no
government employees dotting the landscape; they are typically found in government buildings.
Alternatively the average farm size is both continuous and smooth and would normally utilize
the isopleths method. (The different methods of mapping will be discussed in upcoming
sections.) As Figure 31.4 demonstrates, because farms dot the landscape and sizes can change
fluidly, the nature of this data is polar opposite from that of government employees.

Figure 31.4

Again, a GIS user creating a map must clearly understand the data and the type of
phenomena it represents. More importantly, she must realize what visual variables and
cartographic symbols are appropriate. Misuse of symbolization results in significant errors in
map reading and interpretation, easily derailing the utility of the data one has spent time and
money collecting. Thus, it is imperative that we review the common types of thematic
symbolization found in the above figure so that you have a better understanding of when and
where it is appropriate to use each. The forthcoming thematic mapping techniques – dot,
graduated, choropleth, and isarithmic – have been developed over the past 100 years and are
proven to effectively transmit information clearly and concisely.5

31.8.1. Dot & Graduated Symbol Methods
Two common point techniques used to map discrete data include the dot and graduated symbol methods of thematic representation (Figure 31.5). Dot mapping provides an excellent visual technique for viewing the clustering, dispersion, linearity, and general pattern of a distribution and is often applied to both population and agricultural data. The cartographer must determine both the size of the dot itself, and the unit value, or how many items equals a single dot (e.g., 100 acres of wheat harvested per dot). By taking into consideration the location of environmental or social constraints on wheat production (e.g., wheat is not grown in cities, nor where there is enough precipitation to grow a higher value crop such as corn), one can disperse the dots in relevant areas to more accurately show the distribution of wheat production.

Figure 31.5. The graduated symbol method is excellent for data that are large in number, but close in space. For instance, when mapping the total shipping tonnage brought into ports on the eastern seaboard, the graduated symbol technique is ideal due to the density and size of the data points. Proportional symbols are unique in that their size is directly proportional to their data values. The proportions used to create the sizes can differ depending on the circumstance. There are three choices for proportional symbol creation. Mathematical proportioning creates a direct mathematical ratio between a data point’s value and the size of the symbol compared to other data symbols. Perceptual proportioning skews the sizes of symbols to take into account humans’
incapability of accurately estimating the size of areas. Finally, range-graded proportional circles limit the number of proportional circles used on a map. A dataset’s values are broken down into different ranges and a unique circle size is used to represent all data values falling within an individual range. This way, a map can be made less complex, as perhaps you have 150 data points, but only five differently sized circles on the map, with each circle representing a range of data.

31.8.2. Choropleth & Isarithmic Methods

For continuous data, two mapping techniques are readily available in most GISs – choropleth and isarithmic mapping (Figure 31.6). The choropleth method involves applying value or color intensity to enumeration units (census tracts, counties, states, nations) based on some statistical value. The higher an enumeration unit’s data value, the darker or more saturated the color value. Fundamental to every choropleth method are the concepts of data standardization and classification.

All choropleth data must be standardized. We repeat: a choropleth map may never – ever – be used to map count data. If one maps raw data using the choropleth method, the visualization will suffer from an inherent areal bias. Not all enumeration units are the same size; thus, some enumeration units will naturally have more count data than others simply due to their areal extent. For instance, Texas and California have greater populations than Rhode Island or Connecticut. This should not be a surprise – Texas and California have huge areas compared to the other two states. If you standardize the data by area, however, Connecticut and Rhode Island are far more populated when it comes to the number of people per square kilometer. If you are interested in comparing the raw number of people living in states, you should use proportional symbols.
The isarithmic mapping technique involves lines of constant data value, such as elevation or temperature. These maps are constructed through the “interpolation” of point distributions, with each point having a unique data value. By weaving a pattern of constant data values around these points, a trained map reader is able to visualize a surface with peaks and dips in data values. Many types of data can be depicted using isarithmic maps, including population densities, rainfall amounts, and average incomes.

Isarithmic mapping can be further broken down into two types of representation depending on the nature of the data a cartographer is working with – *isometric* and *isoplethic maps*. *Isometric data* are those actually recorded at positions on the Earth’s surface, for example precipitation (collected and measured in individual containers), barometric pressure (collected by different barometers), temperature (collected at weather stations), and depth to bedrock (collected in different places). Thus each data point represents one individual measurement that one’s lines of constant value can weave around. The *isoplethic technique* is used to map data collected via enumeration unit (e.g., at the county or state level). As with choropleth data, isoplethic data are always standardized by area or some other attribute of the enumeration unit (e.g., number of persons). Thus, isoplethic data would include such things as population density,
per capita income, and average gasoline price. Unlike isometric data points, whose locations are known, isoplethic data requires the construction of a point around which to weave lines of constant value. Thus, when mapping isoplethic data, a cartographer must first create a centroid in every enumeration unit. For example, if a cartographer working for a city government wants to map car thefts by neighborhood, she will first collect the number of car thefts recorded for each neighborhood, standardize this by the number of cars owned in each neighborhood, then find the geometric centroid of every neighborhood, and weave lines of constant value around these data points. If this process sounds complicated, it really is not. Yet, it is more involved in many ways than creating an automated choropleth map, which is one reason cartographers often mistakenly use choropleth maps to map continuous-smooth data when an isoplethic map would better represent the data.

31.9. The Necessity and Rules of Data Classification

Due to the fact that humans cannot differentiate among many different color values, and that cartographers often have multiple enumeration units that need shading, choropleth maps should use classified data. (Please note, however, that range-graded proportional symbols and other types of cartographic visualizations are also equally predicated on the classification of data.) This section will explain what classified data is and review a handful of classification techniques that are useful for illustrating the nature of your data and communicating the message you desire.

Data classification is the mathematical organization of different data points to a limited number of classes (or groups). For example, one may want to create a grayscale map of European wind power based on the number of Kilowatt hours each country produces per person. With well over 30 countries to shade using the choropleth method, a unique shade of gray for each country would make the value and order of different countries’ production hours unintelligible. Thus, the cartographer would “classify” or group like valued countries together and represent their data ranges with one shade. Countries with higher production per person would be shaded extremely dark. Countries with lower production would be shaded a very light gray. Data classification is the process of mathematically determining what countries will get what shade of gray.
Different classifications of exactly the same data can result in significantly different visualizations and interpretations. This makes it imperative that careful consideration be given to each of the following methods of classification on a case-by-case mapping basis. When classifying data, the cartographer must keep in mind several considerations:

1. The purpose and intended audience of the map. Does the classification method chosen enhance the map’s purpose or accidentally cover up the information that is trying to be conveyed? Are the number of classes selected (e.g., seven) suitable for the cognitive level of the audience (e.g., five year-olds) or will the map be too difficult to interpret?

2. The range of data. Does the classification system encompass the full range of data, including both the minimum and maximum values?

3. No overlapping values or vacant classes. Each data value must fall in one class only. Vice-versa, class created must have at least one data value fall within it.

4. The number of classes must be great enough to avoid sacrificing the accuracy of the data, but not be so numerous as to impute a greater degree of accuracy than is warranted by the nature of the collected observations. For example, if you are mapping the 13 provinces and territories of Canada, you should not have 10 classes (i.e., more than one class per two enumeration units) nor only two classes (i.e., dividing the thirteen units into only two groups).8

5. Divide the dataset into reasonably equal groups of observations.

6. Have a logical mathematical relationship if possible (Robinson et al. 1995).

31.9.1. Different Classification Methods

There are many methods and techniques for placing data values down into different classes. We will review some of the most common and rudimentary ones available in nearly all GISs. At the top of the illustration in Figure 31.7 is a set of 20 numbers arrayed along a value line from 10 to 85. The data values range from 11 to 84 – quite a wide range for only 20 different values!
Figure 31.7.

It is good practice to look at the dispersion of your data values along the histogram before doing anything else. Notice how the two highest values are outliers at 82 and 84; there are no data points falling in the 70s at all. The next thing cartographer’s must do is figure out how many classes they want to create. Let us suppose that the audience is an atypical one: adults with training in map comprehension! Using Sturgis’s Rule (see Endnote 8) we have decided that five classifications are ideal for our dataset.

Beneath this histogram are three different methods for classifying data. Each class has a number underneath it representing how many data values fall into each category. Many GISs will allow you to view a histogram and select from a variety of classification methods. The three most common classification methods are equal interval, quantiles, and natural breaks.

An equal interval classification assumes equal distance, or “range,” between the class breaks. In this instance, the equal distance has been set at 15 units, which would yield class breaks of:
Alternatively, a quantiles classification puts an equal number of observations in each class. Thus, if there are 100 observations in a dataset, and the user desires five classes (or in this case “quintiles”), 20 observations will be placed in each category, regardless of whether the class breaks apart observations that are very close to one another or, conversely, forces some that are far apart together. In the example provided in the figure above, a quintiles classification is calculated (20 data values / 5 classes = 4 observations per class), yielding class breaks of:

- Class 1: 10-25
- Class 2: 25-41
- Class 3: 41-52
- Class 4: 53-66
- Class 5: 66-85

Notice that the range of the fifth class has been extended quite a bit, and our outliers in the 80s will now be shaded the same color as several of the values in the 60s.

Finally, a natural breaks classification may be calculated. Here, the user selects the maximum “breaks” along the number line, or where the significant gaps appear in the dataset. The idea is to minimize the internal variation of the data set, while maximizing the variation among the classes. Although the breaks are typically determined through graphical methods – number lines, histograms, and frequency curves – George Jenks created an “optimal classification method” that used an algorithmic approach to determine these ideal breaks (Slocum 2008). For the sample dataset, a possible series of breaks (found by inspecting the number line) might be found at 20, 38, 52, and 75, yielding classes of:

- Class 1: 10-20
- Class 2: 20-38
- Class 3: 38-52
- Class 4: 52-75
- Class 5: 75-85
Many other classification methods exist, including those using nested means, standard deviations, and area-under-the-curve calculations. The references listed throughout this chapter provide greater details on most of the techniques used by geographic information specialists and cartographers. The key thing to take away is that the user of GIS must be aware of both the classification methods provided by the GIS application and the limitations of those methods. The worst thing that a cartographer can do is simply select a classification method based on convenience or what looks “right,” as this rarely leads to an effective or meaningful representation of the data.

31.10. Tying It All Together: Graphic Design, Hierarchy & Figure-Ground

We have reviewed what goes into preparing and shaping the spatial data one is mapping. We have reviewed scale, generalization, visual variables, symbolization, and data classification. What remains to be done is to present this information in a clear and easily received manner that will help map readers avoid any misunderstanding. This section will give a brief overview on map design, visual hierarchy, and figure-ground (Figure 31.8).

Again, before beginning to design the layout of a map, the purpose and audience should be the first thing that pops into the cartographer’s mind. What is the map attempting to communicate? What visual style will best reflect the rhetoric of the data or your message? It is imperative that the cartographer present the mapped elements in a logical manner. Not necessarily logical to the cartographer but to the eventual map viewers.

A map should never be presented as a uniform graphical plane, but one that visually emphasizes certain features and deemphasizes others. This is particularly true when it comes to allowing a map reader to reference what it is they are supposed to be looking at and what is merely tangential information. Information and symbolization of primary importance should be promoted and supplemental information of tertiary worth subdued graphically. The concept of figure-ground pertains to the visual differentiation between the prominent features on a map (figures) and the background components (grounds). Borden Dent (1996) systematizes the approach to the creation of effective figure-ground relationships on maps by arguing that one must establish a hierarchy among the features that are to be mapped. This hierarchy in turn can be matched with an established set of cognitive visual levels (see Figure 31.8).
The differentiation of map levels follows an orderly progression. The actual thematic symbolization (e.g., graduate circles, choropleth values, dots, etc.) should be positioned at the highest visual level. These should be emphasized so that they become the figure. Certain map elements that are instrumental in helping map readers decipher the information should also be prominently displayed at this level (e.g., the title, legend, and cartographic labels). The ground should follow at a lower level, with the base map (e.g., political boundaries, cultural and physical features, and land-water boundaries) deemphasized but still available should someone need to reference the location or relationship among several data points. All other cartographic information should even be further subdued so that it does not take away from the spatial information on the map; such information may include map credits, a north arrow, graticule, logos, or similar items.

![Figure 31.9 Three possible levels of a graphic hierarchy.](image)

Though GISs have become incredibly helpful in almost all facets of cartographic map symbolization, their increasing ubiquity has also resulted in a plethora of poorly designed and illegible maps. GISs will not automate map design. Before a map goes to press, or to the web, one should always double-check with Dent’s hierarchy to make certain that the following information visually jumps out in the following order: (1) thematic data; (2) crucial map
elements; (3) base map data; and (4) all other data. When maps muddle this order and confuse figure-ground, the purpose and message of the map becomes difficult to discern.

31.11. Technological Development & Online Cartography

Technology has always played a role in the evolution of cartography, particularly cartographic methods (Monmonier 1985; 2002). Recent advances in computer hardware and software, coupled with high speed Internet access and the advent of Web 2.0, have changed cartography more rapidly than any other technological developments since the dawn of thematic cartography (MacEachren 1996; Goodchild 2007). These technological innovations have had a great impact on the purpose of cartography, shifting it from merely being a tool of spatial communication to a tool of spatial exploration. A new term to describe this expanded capability is geographic visualization – or geovisualization.9 What is this new visualization? According to MacEachren and Monmonier (1992):

“In the context of scientific visualization, ‘to visualize’ refers specifically to using visual tools (usually computer graphics) to help scientists explore data and develop insights. Emphasis is on purposeful exploration, search for pattern, and the development of questions and hypotheses. Advances in scientific visualization are changing the role of maps and other graphics as tools of scientific investigation.”

DiBiase (1990) argues that technological development has ushered in a whole new realm of possibilities with mapping. He breaks mapping down into two broad uses: visual thinking and visual communication. Maps have always been about communication of spatial data; we have drummed this fact sufficiently already; so let us examine what is meant by “visual thinking.”

Up until recently, maps were static. One could not interact with them. Obviously this is no longer the case. Due to the fact that maps are becoming user-interactive and are also linked to spatial databases via network technologies, maps can now be used for data mining. DiBiase (1990) hypothesized, and research has begun to support, that this new use for maps may be even more powerful than their original role in communication. Maps help humans develop knowledge visually. Moreover, maps are not merely created for public presentation; one can create maps using GIS for private data exploration. DiBiase (1990) and MacEachren (1995) both view this as a major paradigm shift. Cartographers are increasingly adjusting their research to focus on dynamic cartography and geovisualization technologies (Figure 31.9).
One of the more interesting advancements in the area of visualization is the potential for four-dimensional cartography. Unlike the “static” maps discussed so far, dynamic maps include the temporal dimension, which is normally visualized through animations. In such a dynamic world, the visual variables based off of Bertin’s original work are insufficient. DiBiase et al. (1992) proposes that dynamic mapping offers at least three new modes of cartographic expression:

1. Animation: the illusion of motion created from a sequence of still images.
2. Sonification: the representation of data with sound.
3. Interaction: the empowerment of the viewer to modify a display.

They also propose a revised set of dynamic variables that are useful for understanding animation. These include:
1. Duration. Duration is the number of units of time that a scene is displayed. This scene duration may be used as a design variable – longer scenes allow for a more thorough study of a distribution.

2. Rate of Change. Rate of change is a proportion, \( m/d \), where \( m \) is the magnitude of change in position and attributes of entities between scenes and \( d \) is the duration of each scene.

3. Order. Order is the sequence in which the scenes are presented.

Ironically, chronological order is not always the ideal method for exploring a spatial distribution with animation. These three dynamic variables can be used for a variety of visualizations, including emphasizing location, emphasizing an attribute, and visualizing change. Modern GISs and online mapping applications are increasingly incorporating these dynamic capabilities into their software.

### 31.13. Conclusion

The user of a GIS must first be aware of the purpose of the map they are creating and keep in mind the nature of the map’s intended audience. Secondly, a GIS user must be aware of the cartographic capabilities of the system they are using. Rather than take a cookie cutter approach to making a map, they should experiment and explore with the symbology by using different visual variables and exploring different thematic approaches to representing the data. Only through experimentation and tinkering with the arsenal of cartographic techniques and methods available will a GIS user create an intelligible map for her audience. Far too frequently, GIS users lack the background in cartography to adequately and usefully present the data they have spent considerable time and energy analyzing. The worst option is to default to customized approaches packaged within a GIS. Only through taking into consideration all of the cartographic steps and concepts outlined herein, and addressed in more detail in the recommended readings scattered throughout the previous pages, can one be assured of creating a useful and enlightening thematic map.


---

1. A fascinating history of the development of thematic cartography may be found in Robinson’s (1982) *History of Thematic Cartography*. Also of note is MacEachren’s (1979) “The Evolution of Thematic Cartography: a Research Methodology and Historical Review” found in the *Canadian Cartographer* 16(1): 17-33.

2. Many people get confused by large and small scale when they look at representative fractions, often presuming that the fraction with a bigger number must be a larger scale. A simple way to keep this straight is to ask yourself which number is bigger – one-hundred-thousandth or one-millionth? Since the numbers represent fractions, it becomes obvious that 1:10,000 is much larger than 1:1,000,000.

3. Of note to those using GIS, beware of the scale you at which you decide to create your thematic map to make sure it parallels the spatial properties of your data! An increasing problem within modern GISs is the mixing of disparate scales (as database layers) without accounting for differential generalization, detail, and accuracy between and among these layers. The initial scale from which your data were obtained should always be included with the metadata to avoid confusion. *You should never create a larger scaled map than the scale at which your data was collected and compiled, no matter how tempting it may be!*

4. Additionally, with the advent of computer cartography, a fifth type of geographic data, temporal, has been studied; though, the temporal characteristics of geographic data are often ignored on paper maps. There are notable exceptions to this rule; however, including Menard’s now famous map of Napoleon’s march to Russia and back.

5. Many other types of thematic maps also exist, but these four are the most commonly used ones. Flow maps, for example, can be used to create qualitative or quantitative representations of exchanges between two or more places – something that is difficult to do with the techniques reviewed here. However, flow maps are often not created, because they can be extremely time consuming to make in modern GIS systems. For a review of other symbolization and thematic map types, including methods for portraying multivariate data, see the introductory books by Dent (1996), Robinson et al. (1995), or Slocum et al. (2008).

6. A typical method of perceptual scaling is the Flannery method. This option is available in most GISs. Flannery surmised that humans tend to underestimate the size of larger areas, and came up with a formula for creating perceptually accurate areas. Though this formula was the standard method of proportional scaling for several decades, over the past ten years the Flannery method has been largely debunked, because it fails to take into account the impact of neighboring symbols on one’s perception of size.
The dasymetric technique is often desirable over both choropleth and isarithmic methods, but is rarely used in modern GISs due to software limitations and cartographer time constraints. Similar to what can be done with a dot map, this method masks off areas where data points cannot possibly occur and derives choroplethic values based on these logical areas, not default enumeration units.

One useful technique for determining the number of data classes is the Sturges Rule. This rule states that the number of data classes (or $x$) should fall between $2^n < x < 2^{n+1}$. Thus, for the 50 states of the US, the two values of $n$ would be: $2^5 = 32 < 50 < 2^6 = 64$. For 50 observations, the value $x$ falls between $2^5$ and $2^6$, and either 5 or 6 classes would be needed.

For a comprehensive survey of some recent research activities in this area, see the special issue of the journal *Computers and Geosciences* (Vol. 23, Number 4) edited by MacEachren and Kraak (1997).