13 Cartography

Aileen Buckley, Paul Hardy, and Kenneth Field

Contents

A. Buckley (✉)
Esri
Redlands, CA, USA
e-mail: abuckley@esri.com

P. Hardy
Cambridge, UK
e-mail: paul@paulhardy.net

K. Field
Esri
Redlands, CA, USA
e-mail: kfield@esri.com

Abstract

This chapter reviews the geographic data that underpins a map, the process required to portray that information in graphic form, basic principles of map design, and considerations for page layout. While this is not a comprehensive review of all aspects of cartography, several excellent resources that you can refer to later are cited throughout the chapter and in the References. After reading this chapter, you should find that you are better able to understand what a map shows and how it is shown. You should also be more critically aware of how information can be portrayed effectively on maps. This chapter will also help you better appreciate cartography as a crucial part of a geographic information system (GIS).

A review of the development of cartography sets the scene for the current state of the art and science in this field (Sect. 13.1). The principles and practices of cartography are presented to provide a clear rationale for the importance of map design in communicating relevant geographic information. Looking at the various types of maps and their construction allows you to consider the constraints on map design and the consequences for how data is portrayed on maps (Sect. 13.2). A review of the design and use of symbols, color, and type on maps illustrates the many possibil-

W. Kresse, D. Danko (Eds.), *Springer Handbook of Geographic Information*, Springer Handbooks,
https://doi.org/10.1007/978-3-030-53125-6_13

ities and choices available to the mapmaker (Sects. 13.2–13.6). Relief portrayal techniques illustrate how data about a terrain surface can be used to create a three-dimensional representation of the physical environment (Sect. 13.7). The final sections cover the intricacies of map design and page layout (Sects. 13.8 and 13.9). Map production and reproduction are not covered here.

13.1 Cartography in Review

Cartography is often defined as the study and practice of making maps. For this chapter, we concentrate on the practical aspects of mapmaking, which involve a combination of art, craft, and science. Cartography has continually evolved to use new technologies, of which GIS is among the most recent. Manual techniques that were the mainstay of cartographic design and production as recently as the 1980s have now been completely replaced by digital workflows, and new methods that were previously unavailable to or unimagined by cartographers are now pushing the frontiers of cartography.

This technological transformation has had consequences for cartography. Early innovations focused on providing high-quality design required for the full range of maps. GIS software interfaces and parameter defaults tended to lead the innocent user to produce garish and cluttered maps. Today, GIS routinely incorporates the accumulated best practices from cartographic practice and research and even pushes the limits of map design. As technology further evolves from the isolated desktop computer to server, mobile, and cloud environments, cartographic methods and practices have adapted yet again. However, the fundamental tenets of good map design remain largely unchanged; technology simply leads to different modes of implementation, as well as exciting new avenues in the creation of map-based products.

13.1.1 Brief History of Cartography

Cartography is one of the oldest of human graphic portrayal methods, dating back tens of thousands of years to Stone Age rock carvings showing routes to important geographic sites, such as good hunting grounds and foraging sites. Cartography progressed to include Babylonian clay tablet village plans; Polynesian stick charts for navigation (Chap. 14); Egyptian, Greek, and Chinese papyrus and parchment maps; medieval *mappa mundi* (on cloth); and portolan navigational charts on vellum. With the introduction of the printing press in the fifteenth century came paper. Computers have more recently been used to produce not only printed maps, but also ephemeral maps displayed on the screens of computers and mobile devices. Along the way, a wealth of knowledge has been accumulated about the best practices for making maps of different types. This chapter summarizes many of these practices, especially with regard to GIS and mapping.

13.1.2 Cartography and GIS

In its broadest sense, the term *cartography* can be described as the entire mapmaking process—from landscape survey, data collection, and information distillation to graphic compilation of the map and, finally, production and reproduction. In the context of GIS—a technology predicated on spatial data—the term *cartography* relates primarily to the graphic and production aspects of mapping. The early stages of data compilation and analysis are considered largely noncartographic in nature, although visualization of the data leads to initial insights by revealing unknowns [1]. The current view is that geographic analysis is about finding an answer to a geographic question, and cartography is about communicating that answer, generally to a larger audience than the analyst or the mapmaker. In this context, geographic analysis and cartography are complementary, and GIS is essential to both.

The relationship between GIS and cartography has not always been so harmonious. In the early days of GIS, excitement over the nascent technology led software developers to create computer mapping tools without investing time to learn from trained and experienced cartographers, who subsequently recognized and lamented the poor quality of GIS-designed maps. Users accepted the maps created with GIS software as being not only of sufficient quality but superior to traditional maps, simply because they were created with an innovation. The result was a breed of geospatial professionals who were more than capable of developing or using GIS but possessed little cartographic knowledge or training. However, in recent years, GIS has contributed to a rebirth in cartography, and cartography has helped elevate the stature of GIS. GIS tools are now capable of meeting the high-quality standards set by cartographers. In this Information Age, people have a greater appetite for information about the world. Thus map design and production have become relevant and vital not only for geospatial professionals but also for their map-reading audiences. People demand high-quality maps, and GIS developers and users have accepted the challenge and responsibility to ensure that the maps created with GIS technology meet these demands.

A high-quality map is much more than a visual report of GIS data—a well-designed map has elegance and style, clarity and meaning, and the appropriate content to convey the necessary level of precision and accuracy for the intended message. The responsibility of the cartographer is to ensure that the map communicates clearly, without clutter or confusion. A good mapmaker will often adhere to the axiom "Perfection is achieved not when there is nothing more to add, but when there is nothing left to take away." (Antoine de Saint-Exupéry). A good example of a well-made map is the iconic London Underground Tube map by Harry Beck, which removes all aboveground detail other than the River Thames; dispenses with positional accuracy and scale uniformity in order to clearly show connectivity; and demonstrates an innovative design, use of color, and layout. It has thus become a cartographic classic.

Cartographic design is rarely right or wrong in an absolute sense, although some decisions clearly result in incorrect representations of information. Instead, map design can be viewed as more—or less—effective. The challenge for novice cartographers is to learn what constitutes effective cartographic design so that they can ensure that theirs are better maps.

13.2 Types of Maps

Maps take different forms and support a range of uses. Each map type has specific requirements in terms of design and compilation to communicate its message purposefully and support its proper use. It is difficult to categorize all maps neatly by a specific type; nonetheless, cartographers often classify maps into one of three broad categories: reference maps, charts, and thematic maps.

13.2.1 Reference Maps

A reference map (sometimes called a general-purpose map) displays both natural and cultural features in the geographic environment. These features are used to identify the spatial location of objects in an absolute (for example, in latitude and longitude coordinates) or relative sense and in relation to one another. Reference maps focus on the location of a variety of geographic features in an area, with limited focus on feature attributes (aside from type and, sometimes, name). Such a map provides a picture of the geographic character of the mapped area and the spatial configuration of features within that area (Fig. 13.1).

A globe is an example of a reference map that gives a general overview of the geographic nature of the entire earth. Reference maps can also give more detailed views of a portion of the earth's surface. For example, reference maps found in atlases illustrate whole countries or continents, while the maps in a national map sheet series show smaller areas in greater detail.

A topographic map—so called because it portrays the topography or shape and features on the surface of the

Fig. 13.1 Portion of a reference map showing physical features of North America (courtesy of Tom Patterson, http://www.shadedrelief.com/north-america/)

Fig. 13.2 Portion of a 1 : 24 000-scale topographic map of San Francisco, California (courtesy of United States Geological Survey (USGS))

earth—is another example of a reference map (Fig. 13.2). Not only does it depict landforms, but it also includes other natural features, such as coastlines, rivers and creeks, vegetated areas, and lakes, along with cultural features, such as buildings, roads, cemeteries, and other human-made objects. Topographic and other types of reference maps also show features that do not exist physically in the real world yet provide important context for describing the character of a region. One example is political or administrative boundaries, which may, in part, follow features that do exist in the landscape, such as a coastline or road, as with the recreation area outlined in red in Fig. 13.2. (Note that in some regions, the terms *reference map* and *topographic map* are used interchangeably.)

Reference maps of small areas have a high level of precision, making them appropriate and useful for accurate distance and direction finding, position finding, and navigation. To support these types of uses, some maps that started out as topographic maps may be redesigned to create special-purpose maps. For example, road maps are designed to facilitate navigation over road networks, and hiking maps help readers follow trails on foot. Because of their ability to support a wide range of important map functions, reference maps are used in such diverse areas as engineering, resource management, and city planning.

Reference maps also include very detailed maps of small portions of the earth's surface, such as a construction site or subdivision of city lots. These maps sometimes resemble engineering plans and provide the dimensions and direction of boundary lines and precise areal calculations of land parcels. These types of maps are called plans or plat maps—a plat being a plot of land (Fig. 13.3).

Plans or plats are often the most detailed maps available for an area because they result from data acquired by surveying. A plan whose function is based on delimiting property and landownership boundaries is called a cadastral plan. It provides the basis for legal documents used in land records. Plans and plats are important because other maps of the same area are often derived from them. For example, cities will use these maps to create tax maps, school boundary maps, transit maps, water utility maps, and other products.

13.2.2 Charts

A chart is a special type of reference map whose specific function is to aid navigation and allow readers to determine their position and plot routes or courses. Aeronautical charts display water features; landscape features, including terrain; obstructions to flight, such as towers, smokestacks, and powerlines; and flight-related features, such as airports and airspaces. Nautical charts (Fig. 13.4) show water features and characteristics, including depth, rocks, reefs, and substrate; and aids to marine navigation, such as buoys, beacons, and lighthouses. Charts are often overprinted with navigational tools to assist users in manually plotting a course; for example, a compass rose (shown in the land area in Fig. 13.4) provides the directions of geographic (or true) north, marked by the star, and magnetic north marked by the arrow.

13.2.3 Thematic Maps

Maps that focus on a specific subject or theme are called thematic maps. Unlike reference maps, which show many types of features but emphasize none, thematic maps focus on a single type or characteristic of a feature. While reference maps focus on the variety and relative locations of different features, thematic maps emphasize the geographic distribution of the features or phenomena that relate to the

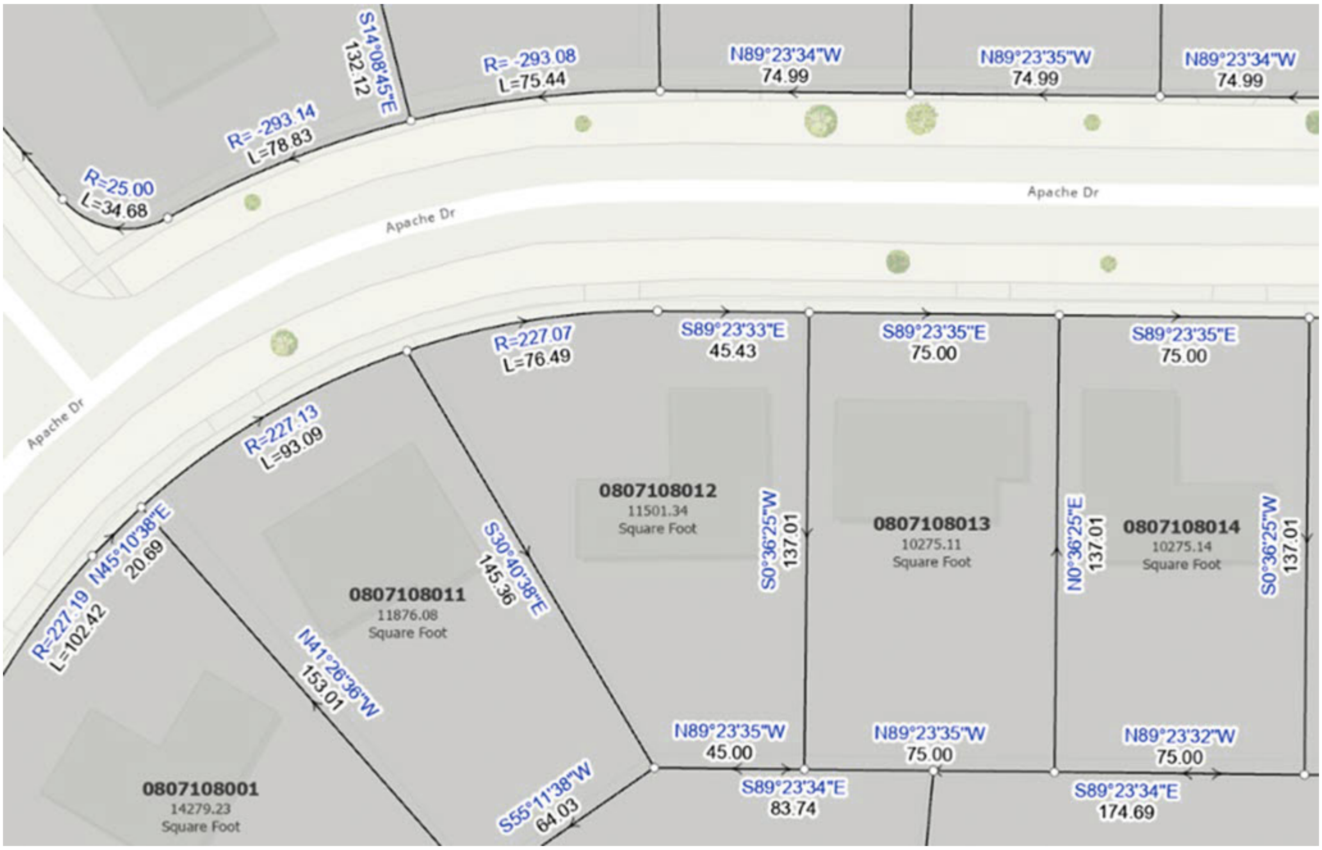


Fig. 13.3 Portion of a tax parcel map in ArcGIS (courtesy of Esri)

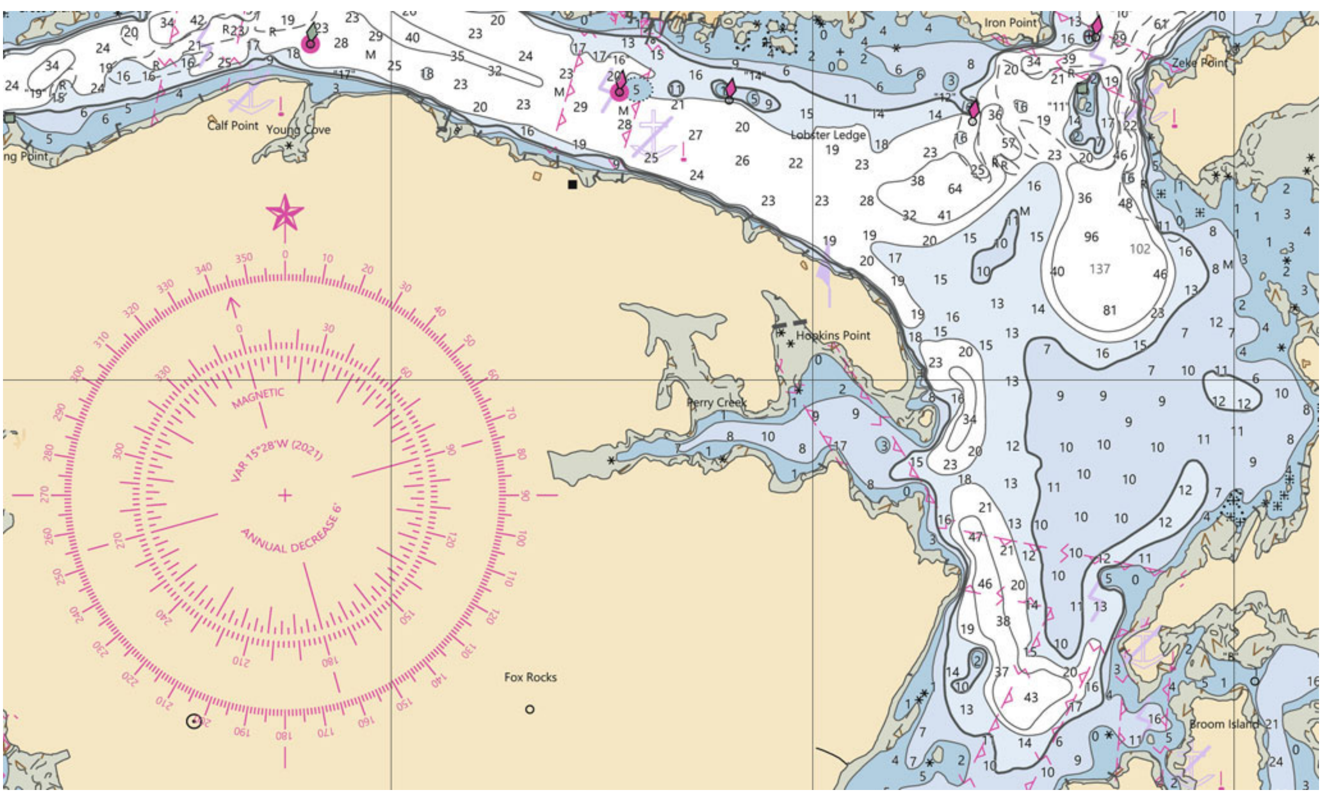


Fig. 13.4 Portion of a nautical chart for the Fox Island Thorofare, Maine, USA (courtesy of National Oceanic and Atmospheric Administration (NOAA), https://www.charts.noaa.gov/OnLineViewer/13308.shtml)

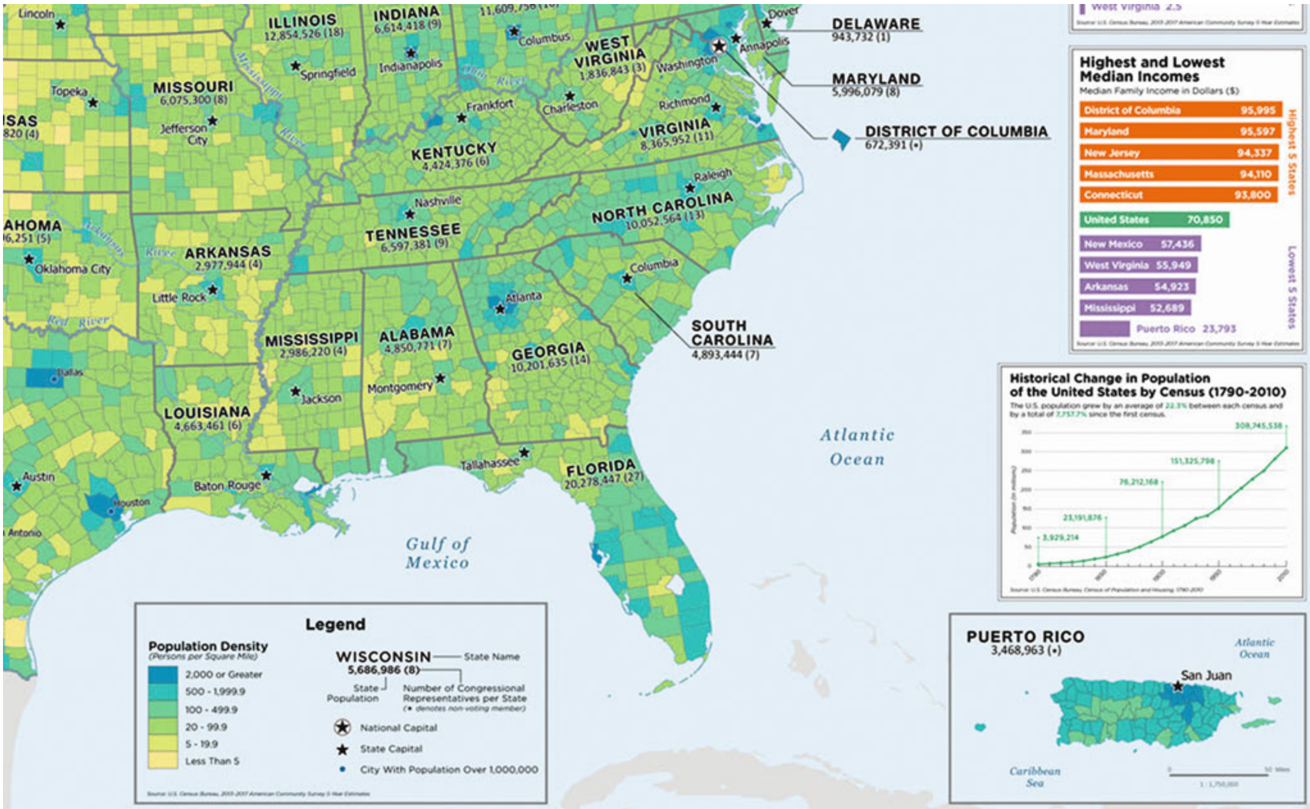


Fig. 13.5 Portion of a thematic map showing population density of the United States (courtesy of US Census Bureau, https://www.census.gov/programs-surveys/sis/2020census/2020-resources/2020-maps/understanding-us-pop-hs.html)

theme of the map, many of which do not physically exist on earth. An example is a map that shows population density (Fig. 13.5). You cannot see this phenomenon in the physical environment, although you may see evidence of it (for example, more houses or apartment buildings).

Thematic maps can depict qualitative information that varies in type but not quantity—such as land use, zoning, and soil classes—or quantitative information based on magnitude—such as population density, annual rainfall, and stream flow.

13.2.4 Qualitative Thematic Maps

Qualitative thematic maps show variation in the kinds of things that are located in the mapped area. This type of map could be used to display the different kinds of trees in a park, the types of crimes that have occurred in a city, or variations in land use. For these maps, it is important that the mapmaker shows the location of features and the variation between feature types, if there is more than one type. Basic reference information, such as political boundaries or city locations, is included to provide geographic context for the reader, but the theme is visually prominent as the most important message of the map. An example is the qualitative map in Fig. 13.6, showing California cities categorized by type with different symbols for the state capital, the county seats, and all other cities. Although the cities themselves provide geographic context, the only other reference information on the map is the county boundaries.

13.2.5 Quantitative Thematic Maps

Quantitative thematic maps show the quantitative attributes (which represent counts, amounts, magnitude, or intensity) of features. Although quantitative thematic maps generally show a single numerical attribute or variable, some are designed to show two (bivariate maps) or more (multivariate maps) variables (Sect. 13.4.2).

Cartographers have developed a variety of standard methods for mapping quantitative data. The appropriate use of these methods depends on the type of information being mapped and the intended use of the map. Choropleth maps, for example, differentiate quantitative attributes of areas through variations in color lightness. These maps can be used when the numerical value within the area is assumed to be homogenous (for example, population density) and when abrupt breaks can be expected to occur between areas (for example, at county boundaries). Figure 13.7 shows a choropleth map represent-

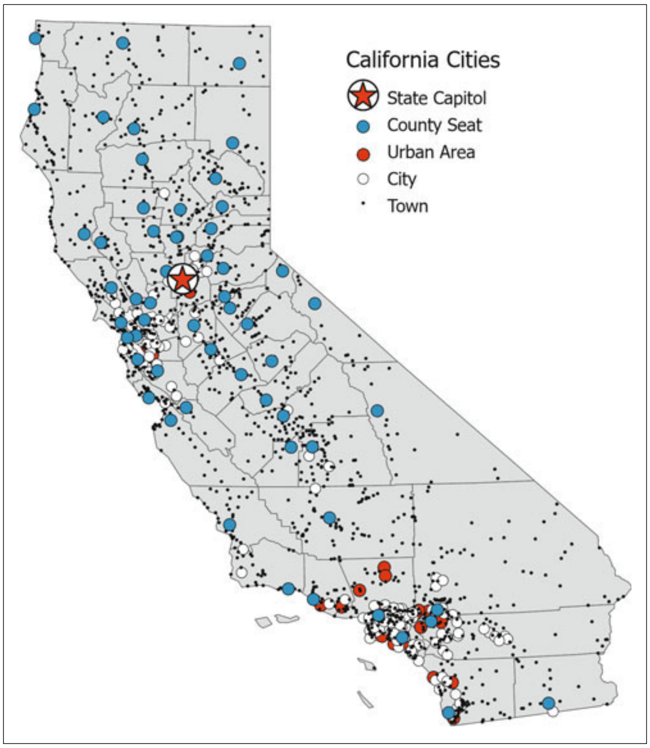


Fig. 13.6 A qualitative map of city types in California (data courtesy of Esri)

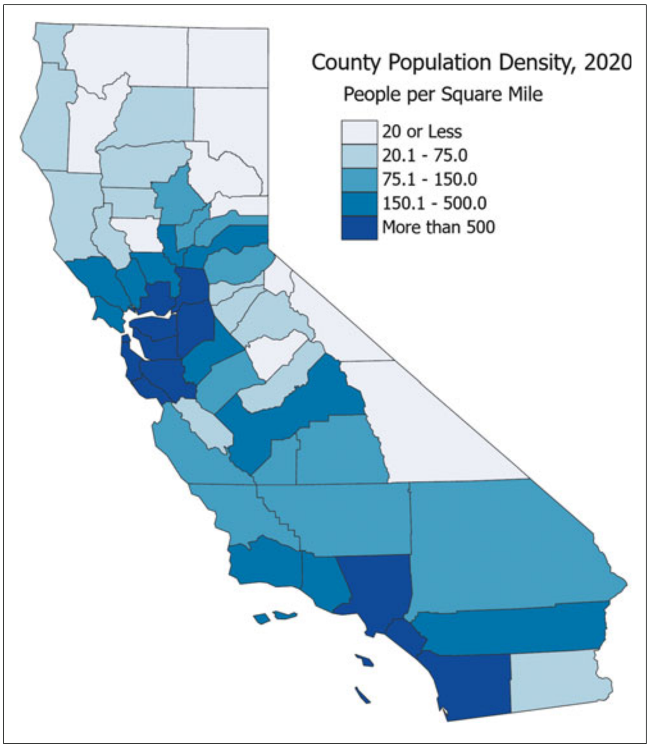


Fig. 13.7 A choropleth map of the 2020 population density for Californian counties (data courtesy of Esri)

ing the population density of California counties. The map is designed to show that some counties have a higher population density (illustrated by darker colors) than other counties (shown with lighter colors). The map also shows areas that display similar characteristics—for example, counties with the same color lightness have the same population density.

An important thing to note about choropleth maps and other maps that assume that the values within the mapped areas are homogenous (for example, Figs. 13.9 and 13.14) is that count (the number of features) and amount (measurable but not countable) data are not appropriate to display with these mapping methods. For example, it would be incorrect to map the raw number of people in a county because all those people could be concentrated within a single area of the county rather than being distributed equally across the county. Notice, for example, that the cities in Fig. 13.8 for the largest county in the state—San Bernardino County, in the southeast—are concentrated in the southwest portion of the county and that there are no cities of significant population size elsewhere in the county. Therefore, when making choropleth maps, it is important to remember to normalize count or amount data by adjusting values so they can be measured on a common scale. Converting population counts to a density (e.g., number of people/area) or a proportion (e.g., number of people in a class/total number of people) are good examples of normalizing data.

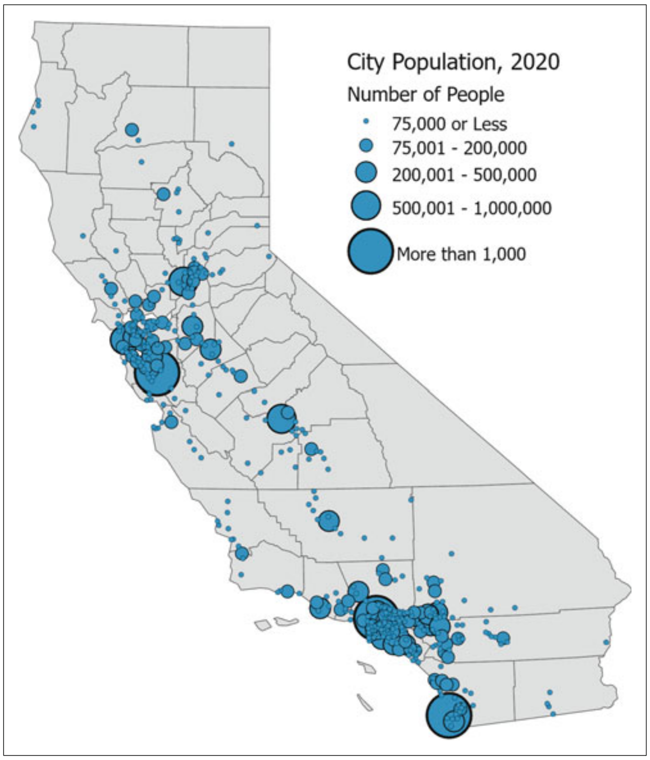


Fig. 13.8 A graduated symbol map of the 2020 number of people in California cities (data courtesy of Esri)

13

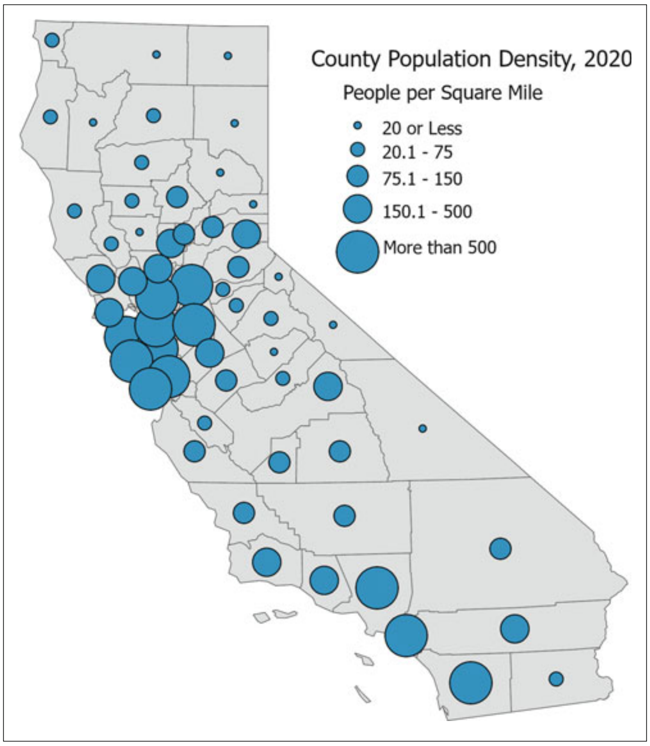


Fig. 13.9 A graduated symbol map of the 2020 population density for California counties (data courtesy of Esri)

A better way to show counts and amounts is provided by graduated or proportional symbol maps, in which the size of the symbol varies relative to the magnitude of the attribute value. Graduated symbol maps divide the range of values into classes, while proportional symbol maps scale the size of symbols in direct relation to the values. These maps can be used to show the values of features at specific locations, such as the number of people in the cities (Fig. 13.8). Strictly speaking, when graduated or proportional symbol maps are used to show values within an area (Fig. 13.9), the data should be normalized for the same reasons that apply to choropleth maps. However, in practice, cartographers will also use these maps to show counts or amounts, as with the proportional symbols used to show number of people in Fig. 13.23.

An isoline is a line connecting points of equal numeric value. For example, the brown lines, called contours, on the USGS topographic map in Fig. 13.2 are isolines of equal elevation, and the gray lines, called bathymetric contours or isobaths, on the nautical chart in Fig. 13.4 are isolines of equal depth. Sometimes, mapmakers color between these lines to show areas with equal values. Isoline maps are conventionally used for portraying data that is sampled at points but that represents a continuous surface, such as elevation, atmospheric pressure, or probability of precipitation (as in Fig. 13.10).

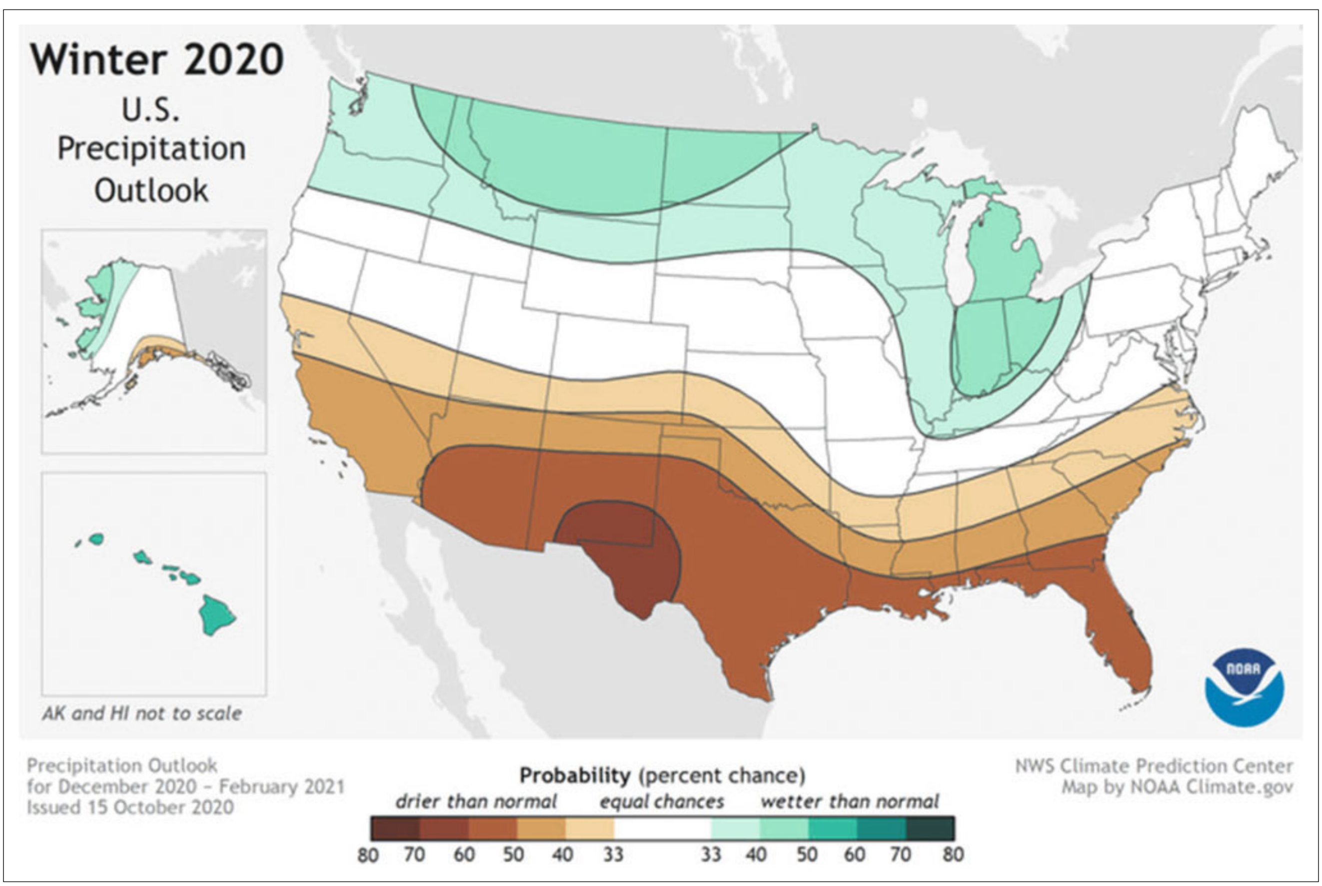


Fig. 13.10 The areas between the isolines are colored to show the precipitation outlook for the United States (courtesy of NOAA, https://www.noaa.gov/media-release/us-winter-outlook-cooler-north-warmer-south-with-ongoing-la-nina)

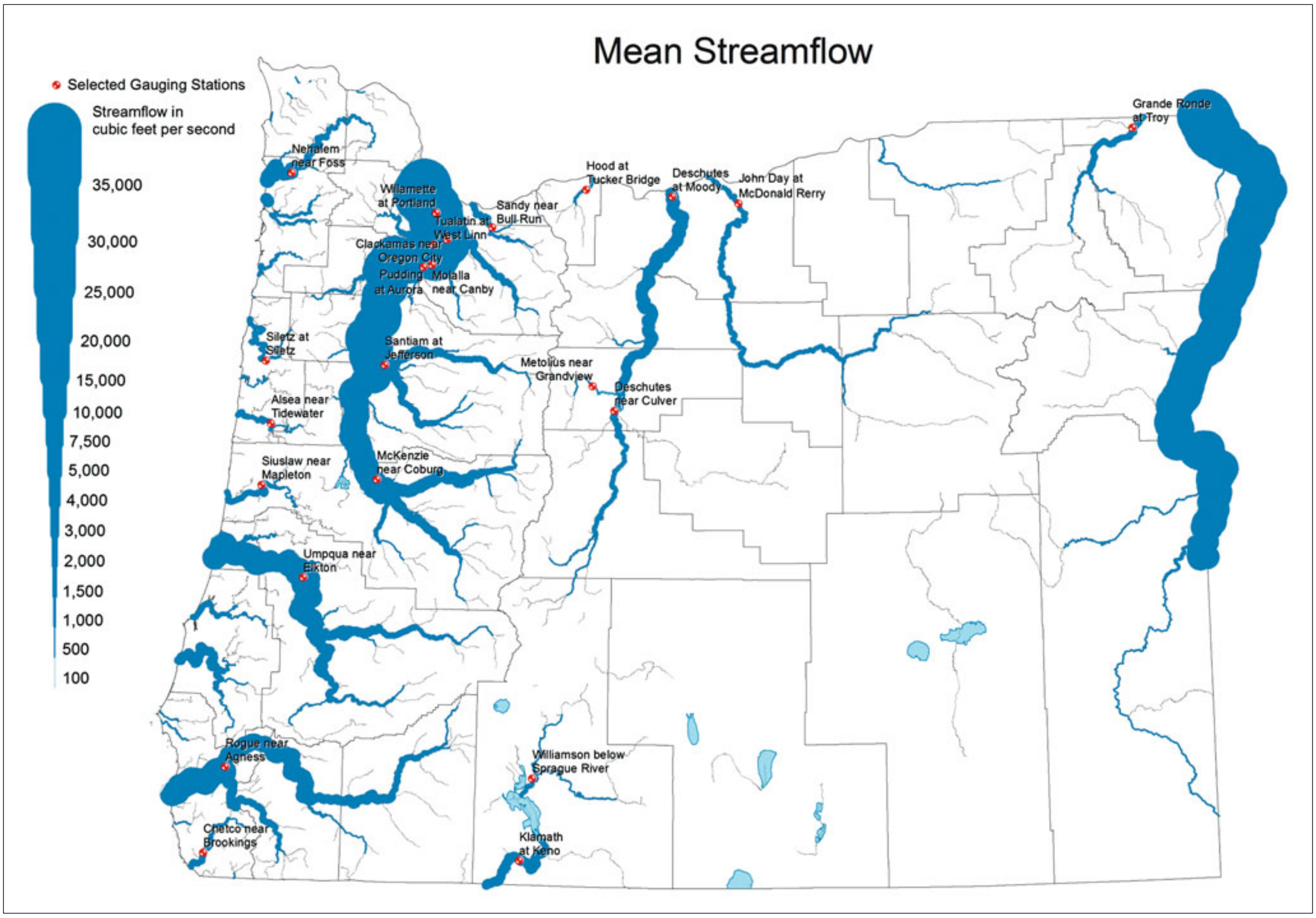
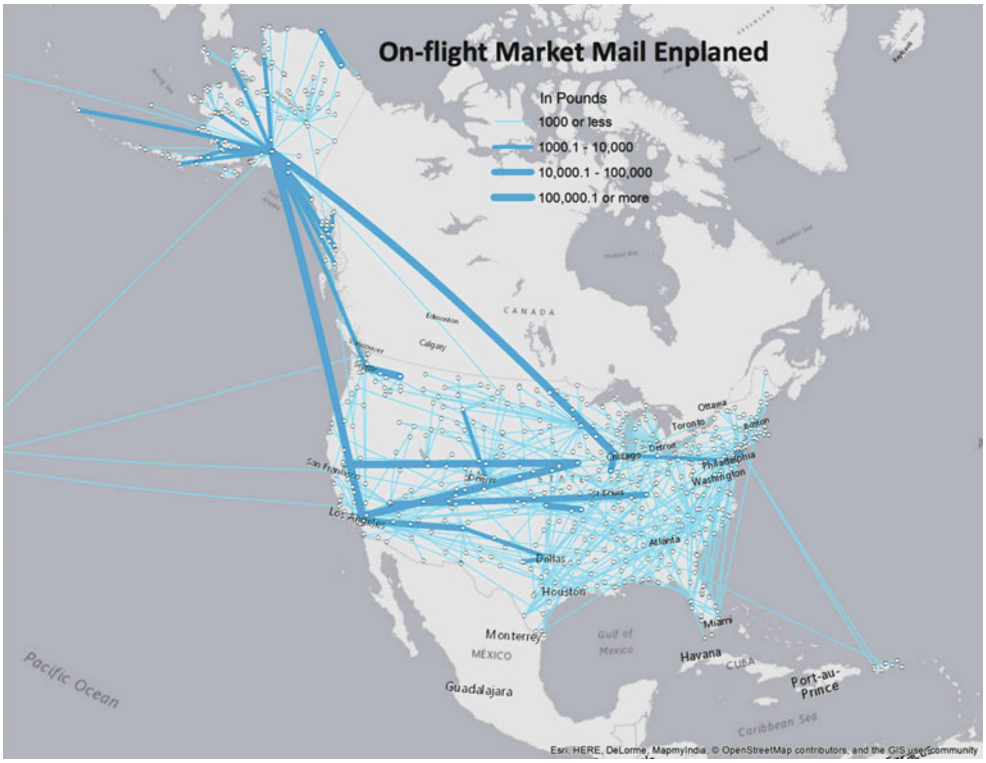


Fig. 13.11 A network flow map showing average annual streamflow for Oregon rivers (data courtesy of USGS and Esri)

Fig. 13.12 A radial flow map showing the amount of mail flown between US airports (data courtesy of Federal Aviation Administration (FAA) and Esri)



13

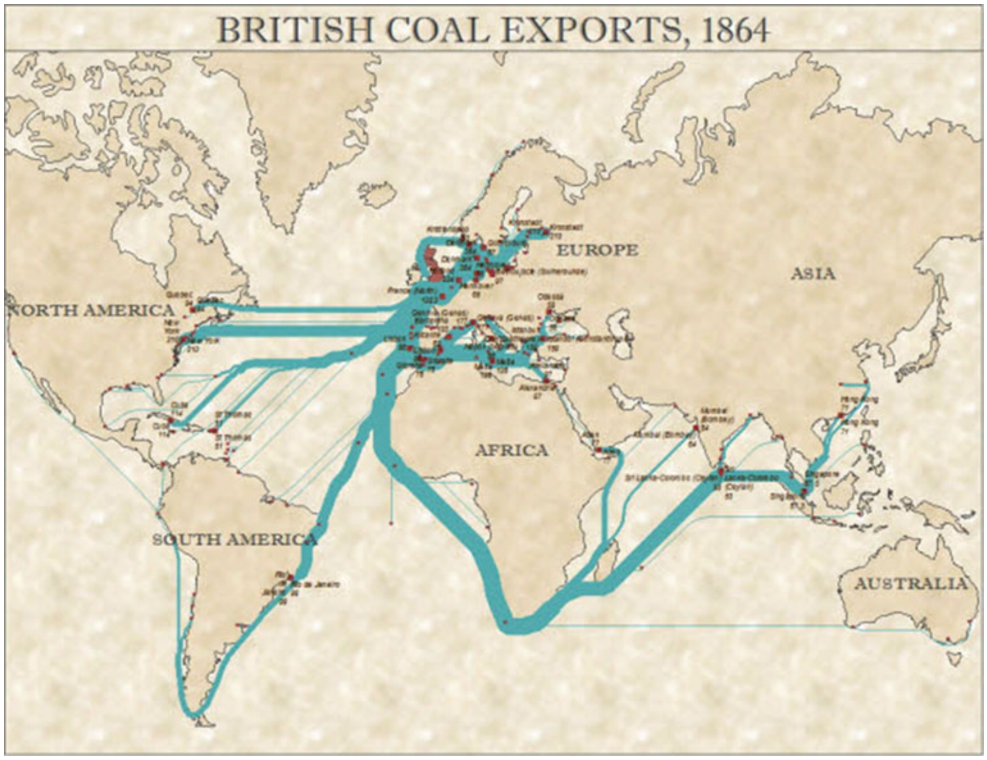


Fig. 13.13 A distributive flow map showing coal exports from Britain in 1864 (data courtesy of Esri)

Another common quantitative map type is a flow map. On these maps, the path or direction of movement is indicated with lines or lines with arrowheads called flow lines. There are three main types of flow maps: network, radial, and distributive. Network flow maps are used to show interconnectivity between places and are usually based on physical networks, such as transportation or hydrographic networks (Fig. 13.11). Radial flow maps have spoke-like patterns because the features and places are mapped in nodal form, often with the shortest route between origins and destinations (Fig. 13.12). Distributive flow maps typically show the distribution of commodities or some other type of phenomena that are delivered from one or a few origins to multiple destinations (Fig. 13.13). The flow lines on these maps may approximate the general supply routes, but they are generalized, unlike the more geographically correct routes on network flow maps (Fig. 13.13).

While there are other methods to map quantitative data, those presented here are the most predominant ones, if we also include dot density maps. Dot density maps—also sometimes called dot maps—show the density of a phenomenon within areas with point symbols (commonly, dots—hence the name of this mapping method). A unique characteristic of dot density maps is that the dots within the areas are shifted to the places where the phenomenon is most likely to occur and excluded from areas where it would not. The density of dots gives the impression of a variation in magnitude (density) of the phenomenon.

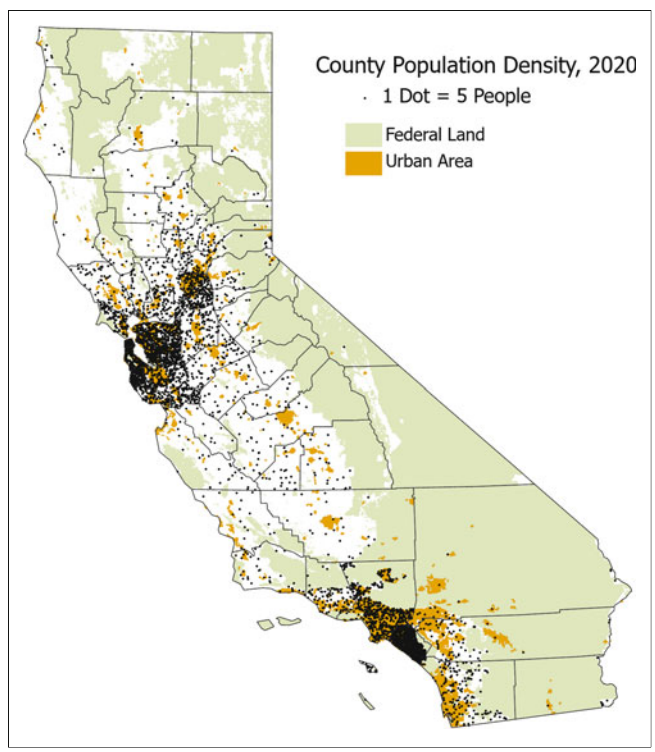


Fig. 13.14 A dot density map of the 2015 population density of California counties, excluding federal lands and including urban areas (data courtesy of Esri)

Dot density mapping is very different from filling a polygon with a pattern of randomly-placed dots, which is actually a form of choropleth mapping because the dots are the fill for the areal symbol. Another important distinction between a dot-filled choropleth map and a dot density map is that the dots on the choropleth map are distributed across the entire area (recall that one of the assumptions for choropleth maps is that the phenomenon being mapped is homogenous across the area), whereas on dot density maps, the dots are excluded from areas where the phenomenon would not exist and concentrated in areas where it would exist. On a dot density map, the dots cannot exist at the same locations as the features because each dot represents more than one feature. The legend on these maps should always indicate the value that each dot represents.

An example of a dot density map is shown in Fig. 13.14. For this map, dots are excluded from the federal lands, shown in green, and included in the urban areas, shown in orange. Notice that the population density in San Bernardino County is represented appropriately, with the dots concentrated in the southwest portion of the county. Dots in the other areas of the county may be located in small areas of public land or pockets of public land within the federal lands.

13.3 Cartographic Compilation

It is sometimes easy to forget that maps are abstract representations of the geographic environment, and people often imbue them with a stronger sense of reality than they should be given. It helps to remember that there are certain properties of maps that result from the cartographic compilation process. That process starts with projecting a scaled down representation of the earth onto a, most often, flat surface. Then cartographers reduce complexity and increase clarity during compilation of the map through selection, generalization, classification, and symbolization of the features on the map. Thus, maps share the following characteristics.

- On all maps, there is a systematic reduction from ground distance to map distance—this is called map scale.
- Except for globes, all maps are made using a map projection, which is a mathematically defined transformation of locations on the spherical (or elliptical) earth to the flat surface used to display the map (for example, a sheet of paper or a computer screen).
- Maps are generalized representations of the environment. Features important to the map are selected and, if appropriate, classified; then they are represented in a generalized form.
- Maps are symbolized representations of the environment. The generalized features are shown graphically, using map symbols and labels.

Because all of these properties are important to consider when a map is being made or used, they are discussed in further detail below.

13.3.1 Map Scale

Map or cartographic scale is the relationship between distances on the map and their corresponding ground distances. Map scale can be expressed as a representative fraction (RF), for example, $1/250\,000$ or $1:250\,000$. In this example, 1 unit of measurement on the map equals 250 000 units of measurement on the ground—for example, 1 inch to 250 000 inch or 1 cm to 250 000 cm. Scale as a ratio is often casually abbreviated to just the denominator—for example, a 250 000 or 250 K map. Map scale can also be expressed as a verbal phrase, such as *1 inch to a mile*, in which 1 inch on the map relates to one mile on the earth. Map scale can also be expressed as a scale bar, which looks like a small ruler on the map or in its margin.

Maps can be classified as either large- or small-scale maps. These terms come from the numeric value of the RF—$1/x$ or $1:x$. For example, the numeric value 1/24 000 is much larger than $1/100\,000\,000$. Thus, the smaller the denominator of the RF, the larger the map scale. A topographic map at a scale of 1 : 24 000 is a large-scale map (Fig. 13.2) in which small areas are shown in great detail, and a world map at a scale of 1 : 100 000 000 is a small-scale map (Fig. 13.13). Note that most of the maps in this chapter are not shown at their original map scale because they have been resized for this publication. This is an important thing to remember about map scale, as the only truly useful indicator for map scale on resized maps is a scale bar (provided it has been resized exactly the same way the map has).

Web maps provide varying levels of detail because additional detail is revealed when users zoom in on an area—these types of web maps are examples of multiscale maps. Each time the user zooms in on the map, the map scale becomes larger. Unfortunately, the map scale is often not indicated on these maps. If this is the case, it is sometimes possible to determine the map scale if the zoom level is known (Table 13.1).

One thing that map scale does for map readers is indicate the appropriate usage of the map. To make accurate distance, direction, and area measurements, use maps at a scale of about 1 : 250 000 or larger. The change in scale across the extent of these maps is negligible, so the map can be trusted as a geometrically correct representation of the small portion of the earth it portrays. If a generalized representation of a large area—such as a state, a country, a continent, or the entire earth—is required, choose a small-scale map. For these maps, remember that scale changes continuously across the map, so if an RF is provided, it only gives the scale at a particular point or along a given line or lines, but not for the entire map. As a result, these maps are not appropriate for calculating measurements of distance or area.

Table 13.1 Map scales associated with the zoom levels for web maps that use the web Mercator projection and tiling scheme (see https://developers.arcgis.com/documentation/mapping-apis-and-services/reference/zoom-levels-and-scale/)

Zoom level	Map scale	Usage suggestion
Level ID:0	591 657 527.59	Global
Level ID:1	295 828 763.80	
Level ID:2	147 914 381.90	Subcontinent
Level ID:3	73 957 190.95	
Level ID:4	36 978 595.47	
Level ID:5	18 489 297.74	Large country
Level ID:6	9 244 648.87	
Level ID:7	4 622 324.43	Small country/US state
Level ID:8	2 311 162.22	
Level ID:9	1 155 581.11	Large metropolitan area
Level ID:10	577 790.55	
Level ID:11	288 895.28	City
Level ID:12	144 447.64	Town
Level ID:13	72 223.82	Village
Level ID:14	36 111.91	
Level ID:15	18 055.95	Small road
Level ID:16	9027.98	Street
Level ID:17	4513.99	
Level ID:18	2256.99	Address
Level ID:19	1128.50	Street intersection
Level ID:20	564.25	
Level ID:21	282.12	
Level ID:22	141.06	
Level ID:23	70.53	

13.3.2 Map Projections

A map projection is a geometric transformation of the earth's surface onto the flat surface used to display the map, most commonly a sheet of paper or a computer screen. This transformation results in distortion of the earth's spherical nature on the flat medium. The map projection process influences the basic geometric properties of the earth's representation, including direction, distance, area, and shape. All maps aim to minimize these distortions or preserve a particular geometric property at the expense of others. The problem is that all the properties cannot be preserved at the same time; therefore, it is important that mapmakers and readers understand how projections maintain or maximize their critical geometric properties and how the other properties are affected.

Many GIS users can and do ignore the map projection issue, particularly if their maps cover a small area not near the poles, because distortion will be minimal on these large-scale maps when the majority of map projections are used. In other cases, GIS users rely on GIS data that is already stored in a projection suitable for the part of the world they are mapping, or they use a projection that has been dictated to them by an authoritative entity or through convention. However, nearly all mapmakers and users will need to know how to make educated projection decisions at some time, particularly for maps of large areas (for example, many countries, the continents, and the world). The challenge is that there are an infinite number of map projections.

So how does the informed GIS user distinguish one projection from another or choose the appropriate one? One way is to organize the wide variety of projections into a limited number of map projection families based on the distortion of the geometric properties discussed above. Therefore, cartographers have developed this set of map projection families:

- Direction (azimuthal)
- Local shape (conformal)
- Area (equal area)
- Distance (equidistant)
- Shortest route (gnomonic).

Some map projections retain more than one of the geometric properties. For example, an equidistant azimuthal map projection preserves both direction and distance. However, no map projection can simultaneously preserve shape and size. To learn more about map projections and their distortion properties, see standard cartography textbooks, such as [2], Snyder's classic book [3], or web resources, such as those listed at http://en.wikipedia.org/wiki/List_of_map_projections.

Despite the infinite number of projections to choose from, in practice there are a relatively small number of projections that are regularly used because they are appropriate for common combinations of extent, map scale, and intended map use. A resource that summarizes the commonly used map projections is the USGS map projection poster (https://www.usgs.gov/media/files/map-projections-poster).

Although there is a wealth of information on map projections, all too often map projections are used incorrectly. A good example is the Mercator projection (Fig. 13.15), which is unique in that any line drawn on a map using this projection is a line of constant direction or bearing. This property makes the projection good for navigational maps and charts, since navigators can draw a line from their current position to their destination and simply set the compass bearing in that direction. If they keep following the bearing (not accounting for air or water currents and the like), they will eventually arrive at their destination. In practice, the process is more complex than this, and navigators will first find the line of the shortest route (which can be done using a gnomonic projection), then convert that to a series of shorter straight lines on a Mercator projection map, which will be used to set and reset their course.

Although constant direction is an important and desirable property for maps that support navigation, it is countered by great distortion in shape and size (Fig. 13.15), which makes the Mercator projection inappropriate for almost any other use. Nonetheless, for decades, wall maps in children's classrooms were made with this projection. These maps would be a better teaching aid if they were made with a more ap-

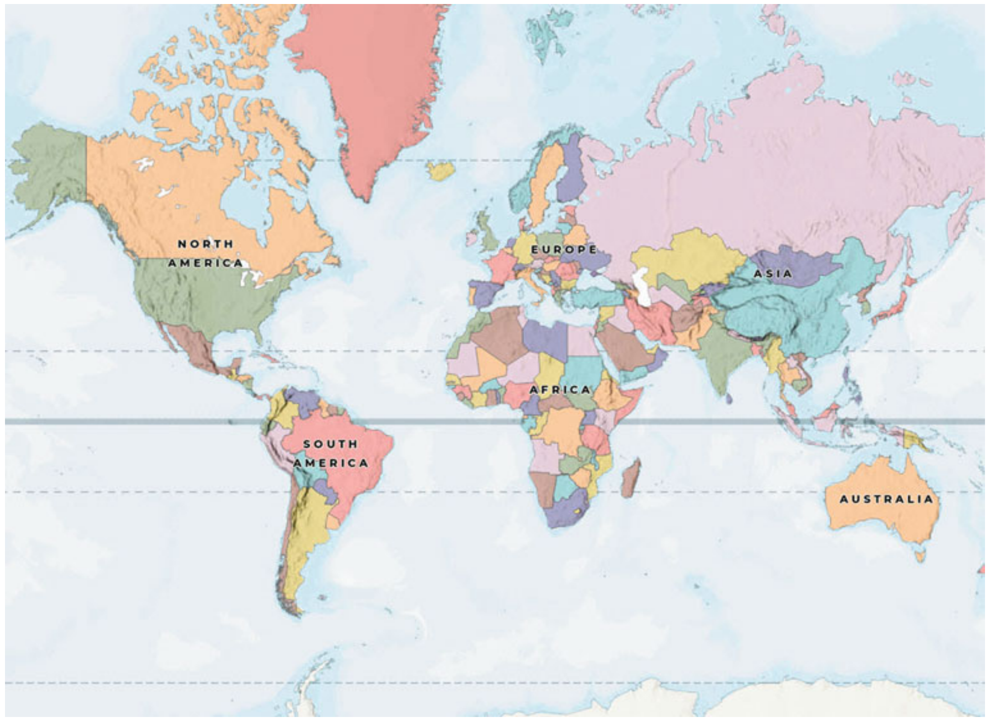


Fig. 13.15 World countries shown on a map with the Mercator projection (data courtesy of Esri)

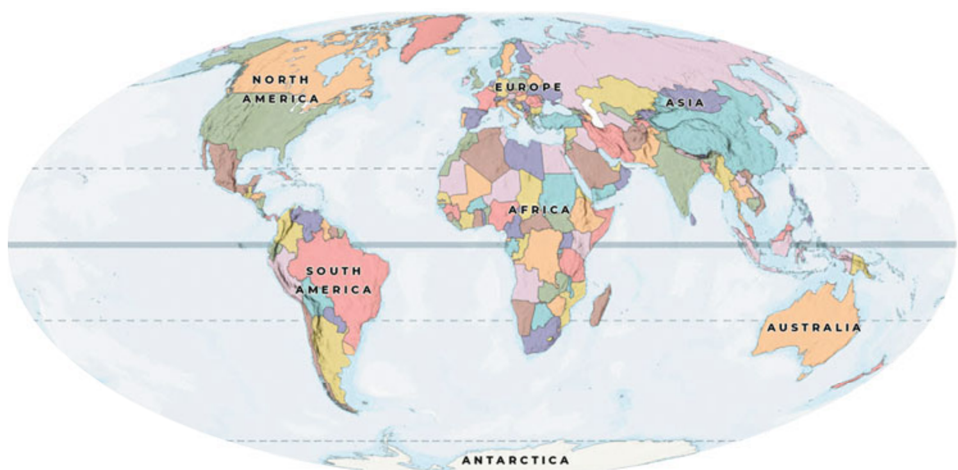


Fig. 13.16 World countries shown on a map with the Mollweide projection (data courtesy of Esri)

propriate projection, for example, the Mollweide projection (Fig. 13.16), which is an equal area map projection used for the world.

A modified web Mercator projection has become the de facto standard for web maps. (See [4] for an in-depth examination of the web Mercator projection.) While this is fine for viewing small areas (zoom levels 11 through 23, as shown in Table 13.1) that are not near the poles, it is inappropriate for most smaller scale maps of continents and the world. Nonetheless, the web Mercator projection has found widespread use, and examples of world maps made with it abound. As a result, generations of map users, including users of modern web maps, may have a distorted mental view of the relative size and shape of the world's land masses.

13.3.3 Selection

The first step in abstracting information about the world into something that can be represented on a map requires selection—deciding what type of and how much information to portray. This selection of information reduces the complexity of the world to make the map more intelligible. Selection not only increases clarity but also helps the map reader understand what is important on the map.

A map that will be used for a range of purposes, such as a topographic map, must include a lot of information of various types, although only some of it may be relevant for particular map uses. For these maps, selection is a more arduous task. For thematic maps, only the locational information that is required to provide geographic context and is relevant to the theme should be included, because superfluous information can result in distraction from or misinterpretation of the map's message. For thematic maps, the challenge is to avoid adding too much superfluous content.

13.3.4 Generalization

Once the features important to the map have been selected, they are then represented in a generalized form. Cartographic generalization reduces the amount of information on a map through a change in the geometric representation of features. Common generalization operations for vector data are shown in Fig. 13.17. A generalization operation may be applied to a single feature or a single type of feature, but the results of that operation may affect other features or categories of features. For example, smoothing a line symbol that depicts a river may result in the generalized line overlapping streamside buildings or roads or being misaligned with bridges and contour lines. As a consequence, an additional generalization operation—for example, displacement—may be used. Most generalization tasks, therefore, require that the results be considered contextually (relative to other features) to ensure graphic clarity and allow inherent geographic relationships and patterns to be recognized—this is known as contextual generalization. It is easy to understand why multiple generalization operations are often used. The most important consideration for a cartographer is that, whether applied to a single feature or a collection of features, the outcome of any generalization operation or set of operations should always preserve the characteristic appearance and nature of the local geography.

Raster data can also be generalized. For example, the satellite image in Fig. 13.18 can be classified to show land cover (Fig. 13.19). A filter can be used to replace the values of isolated cell with those of the majority of their contiguous neighboring cells. A boundary smoothing operation can be used to simplify the boundaries between classes. Pockets of cells that do not meet a size threshold can be grouped into larger regions. These are but a few examples of raster generalization operations, the goal of which is to produce a more generalized map with larger groups of cells that have similar values (Fig. 13.20).

13.3.5 Classification

Selected and generalized data can be further simplified through classification. Classification is the grouping, ordering, or scaling of features into categories (for qualitative data) or classes (for quantitative data). This is usually done based on attributes of the features, although sometimes it can be done by some other criteria, such as using the location of features to create clusters.

For quantitative data, there are several commonly used classification methods (or classing methods). Classification methods are procedures used to assign class intervals or ranges to numerical distributions. The selection of class intervals and the number of classes both have a significant impact on the appearance of the quantitative information on a map. Although class intervals can be set manually, most GIS systems provide a set of common classification methods that users can choose from, three of which are described below.

With quantile intervals, the same number of features is assigned to each class. If the number of classes is four, the classes are quartiles, five classes are quintiles, and so on. Data can also be classified by equal range intervals, commonly called equal intervals, in which the range of data values is divided by the desired number of classes. Another classification method that is commonly used in GIS is to specify classes based on breaks between groups in the distribution of data values. With natural-break intervals, the goal is to minimize the variation within classes and maximize the variation between classes. This can be done using a method described by Jenks [6]—thus, this is often called the Jenks natural-breaks classification method or Jenks optimization method. For a complete description and illustrations of these methods and others, see [2].

13.4 Symbols

The final step in the cartographic abstraction process is symbolization, in which features and their attributes are represented by graphic elements, called symbols. Symbols do not always take on the appearance of the geographic feature they represent, for example, a simple circle does not look like the city that it represents on a map (see Fig. 13.6). Nonetheless, the map reader should be able to interpret the symbol's meaning accurately and relate it to its real-world counterpart appropriately.

13.4.1 Properties of Symbols

Cartographers commonly consider three properties of data that guide their design and use of the symbols used to represent the data dimensionality, level of measurement, and visual variables [8]. Dimensionality is an expression of the spatial extent—especially width, height, or length—of a feature. For cartographic purposes, dimensionality relates to the following types of symbols:

- Point symbols (zero dimensions)
- Line symbols (one dimension)
- Area symbols (two dimensions)
- Surface representations (2.5 dimensions; Sect. 13.7)
- Volume representations (three dimensions).

The dimensionality of the feature in the real world may be the same as the dimensionality of the feature's symbol on the

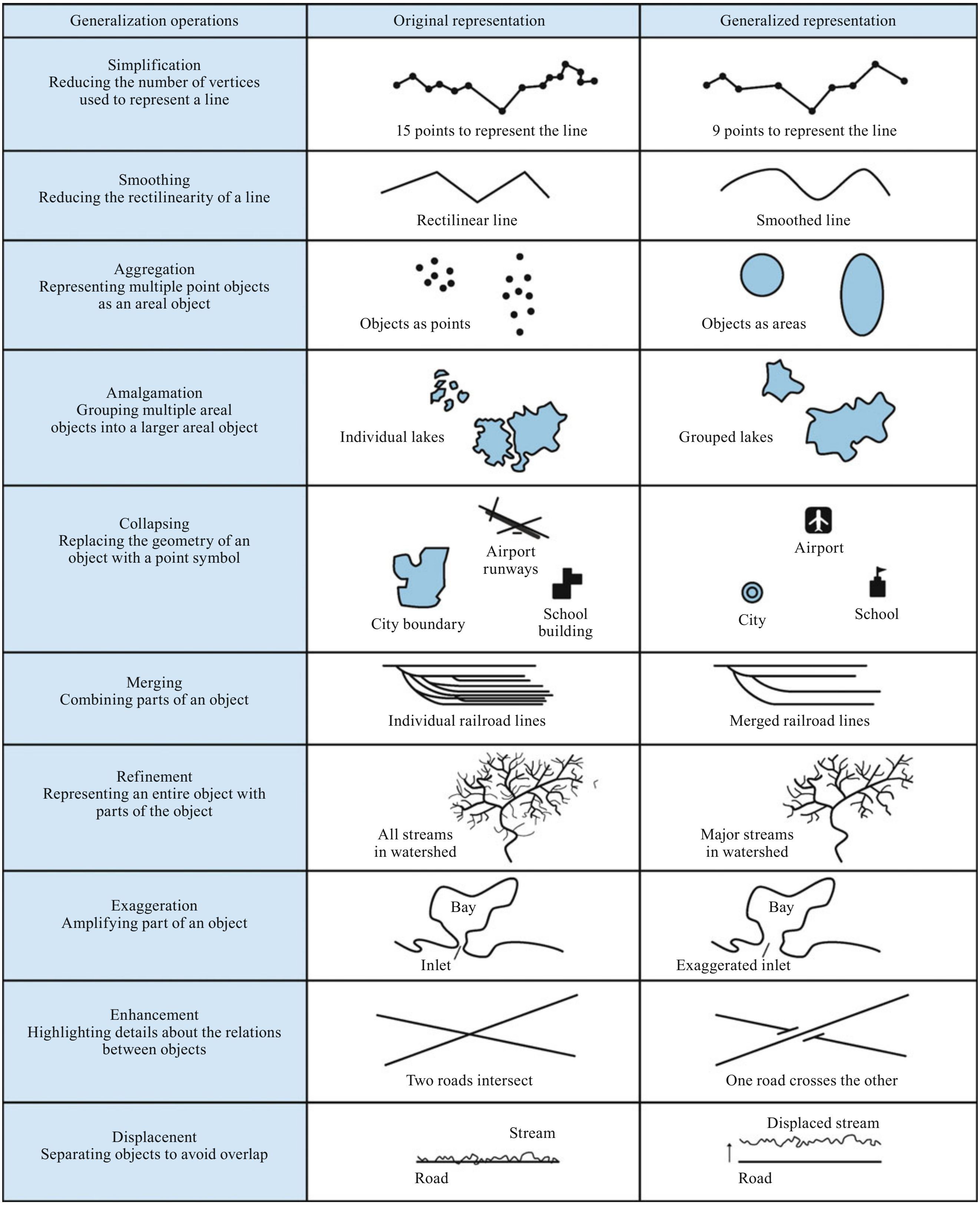


Fig. 13.17 Common generalization operations for vector data (adapted from [5])

Fig. 13.18 A small portion of 30-meter Landsat 8 imagery (bands 6,5,2) focused on the area around Plaster Rock, New Brunswick, Canada (data courtesy of USGS and Esri)

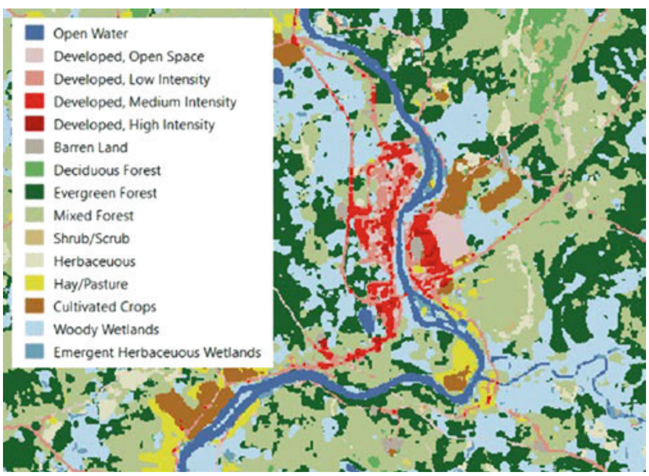


Fig. 13.19 A portion of the land cover map after classifying the much larger satellite image (data courtesy of USGS and Esri)

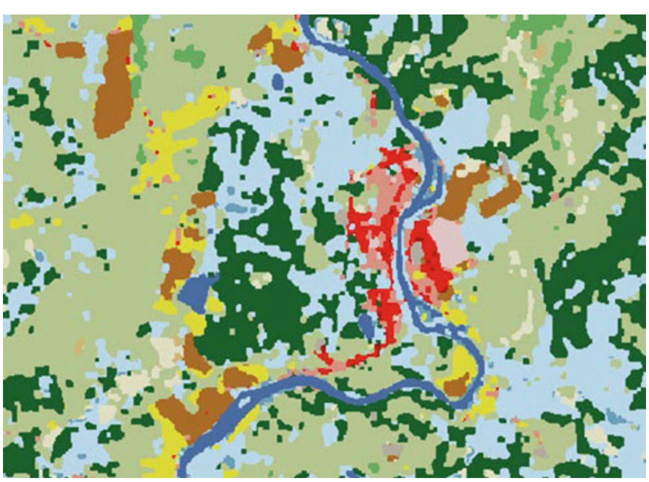


Fig. 13.20 The land cover map after raster generalization operations have been applied (data courtesy of USGS and Esri)

map. For example, an object that exists as a linear feature in the geographic environment, such as a powerline, may be mapped using a line symbol, such as a black line; a vegetated area may be mapped using a green polygon; and a surveyed location may be symbolized with a brown cross. The correspondence between the dimensionality of a feature and its symbolized representation is not always so straightforward. This is due to the other factors that relate to the properties of maps, including map scale, map projections, and generalization, but foremost among them is map scale. As map scale changes, the spatial dimension of a mapped feature may also change. For example, on a large-scale map, a city may be represented as an area with two dimensions. However, the same city might be better represented as a point feature with zero dimensions on a small-scale map (as with the state capitols in Fig. 13.5), because the space available to show the city is restricted.

Another property of symbols relates to the level of measurement of the data they are being used to represent. For cartographic purposes, it is often sufficient to classify the data as either qualitative or quantitative. Thematic maps showing these types of data were discussed in Sects. 13.2.4 and 13.2.5.

The third property of symbols relates to how their appearance can be altered or varied. In cartography, the graphic variations of map symbols are called visual variables, a concept introduced by Bertin in 1967 [7]. Visual variables provide mapmakers with a range of possibilities to consider when designing symbols and assigning them to features on maps. Visual variables can be categorized as those that differentiate and those that order. Differentiating visual variables are used to show qualitative differences among features, while ordering visual variables show quantitative differences. Cartographers refer to these as the qualitative and quantitative visual variables, respectively.

The qualitative visual variables include shape, orientation, and color hue (Fig. 13.21).

- The shape used for point symbols can range from simple geometric forms (circles, squares, or triangles) to miniature representations that mimic the feature (called mimetic symbols), such as a pictorial sketch of a landmark building, as shown in Fig. 13.21. Shape, in reference to line symbols, relates to the marks that make up the line. For example, different combinations of solid or dashed lines and point symbols are used to portray different kinds of linear features. As with line symbols, shape for area symbols relates to the marks that make up the symbol's fill. The marks in line and area symbols also range from simple geometric shapes to detailed mimetic symbols.
- Color hue refers to the dominant wavelength of visible light that forms our color association (see Sect. 13.5.1). What we commonly recognize as the colors red, blue, and yellow are more accurately referred to as hues. Strictly

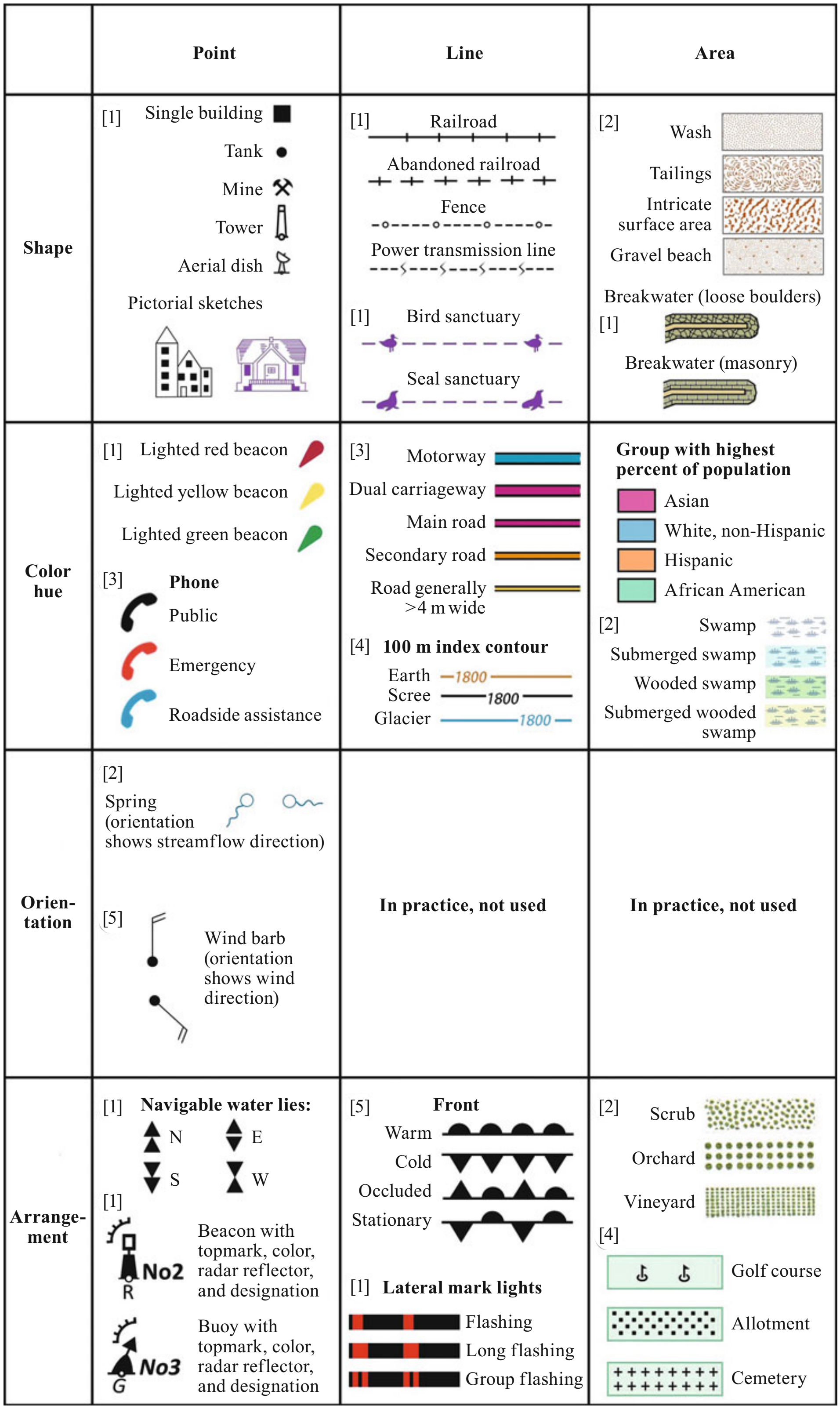


Fig. 13.21 Qualitative visual variables (adapted from [2] NOAA nautical charts, [3] USGS 1 : 24 000-scale topographic maps, [4] Ordnance Survey 1 : 25 000-scale Explorer maps, [5] swisstopo national maps, and [6] NOAA weather maps)

speaking, variation in hue should not impart a message of magnitude, as red is not more than blue, which is not more than yellow. However, some colors have an inherent lightness (for example, yellow) or darkness (for example, purple), so there may be some internal ordering by the human visual system. Additionally, some hues have an inherited meaning through common usage. For example, on weather maps, red is often associated with warmer temperatures and blue with cooler temperatures. These considerations should be kept in mind when choosing hues for qualitative map symbols or reading color maps.

- Orientation relates to the angle of the symbol (or a portion of the symbol) for point features, as with the spring symbols in Fig. 13.21. In common practice, orientation is not used for line or area symbols since the orientation of the corresponding features is already determined by their geography [7]. Although orientation can be measured in angles from 0–360°, this numeric value does not impart a magnitude message because north is not more than east, which is not more than south.
- Arrangement relates to the positioning of marks within a symbol. An example is a random arrangement of green circles to denote an area with scrub or a regular array of similar symbols to show orchards and vineyards, as illustrated in Fig. 13.21.

The quantitative visual variables include size, color lightness, color saturation, and pattern texture (Fig. 13.22).

- The size of a symbol can be varied, which both the point and line examples in Fig. 13.22 illustrate. The visual impression of point and line symbols varying in size is one of a measurable or ordered difference in the magnitude of the attribute being symbolized. In common practice, size is not used for area symbols, since the area of a feature is already dictated by its geography [7].
- Color lightness, also called color value, refers to the darkness or lightness of a color (see Sect. 13.5.1). The human visual system will naturally interpret a darker color as *more*, so darker colors are used to represent features with greater quantities or higher quantitative data values than features shown with lighter symbols. In Fig. 13.22, the color lightness of the point symbols for each of the race and ethnicity categories (distinguished by color hue) represents the percentage of population.
- Color saturation, also referred to as color chroma, is the purity of a hue (see Sect. 13.5.1). Saturation is varied by mixing a hue of high purity with a gray of the same lightness—the addition of gray results in lower saturation. The terms *brilliant* or *intense* can be used to describe fully saturated colors, and *muddy* or *dull* describe colors with low saturation. In Fig. 13.22, color saturation is used to show percent slope for each of the aspect categories, which are differentiated by hue. In practice, color saturation and color lightness are often used together to create sets of colors ranging from light to dark that can be used to symbolize quantitative data (Sect. 13.5.6).
- As with the shape and arrangement of qualitative visual variables for area symbols, pattern texture relates to the marks that make up the symbol. Denser symbols with less space between the marks will appear to be darker. Thus, pattern texture acts in the same way as the color value quantitative visual variable so that denser patterns denote higher values, as illustrated in Fig. 13.22.

13.4.2 Bivariate and Multivariate Symbols

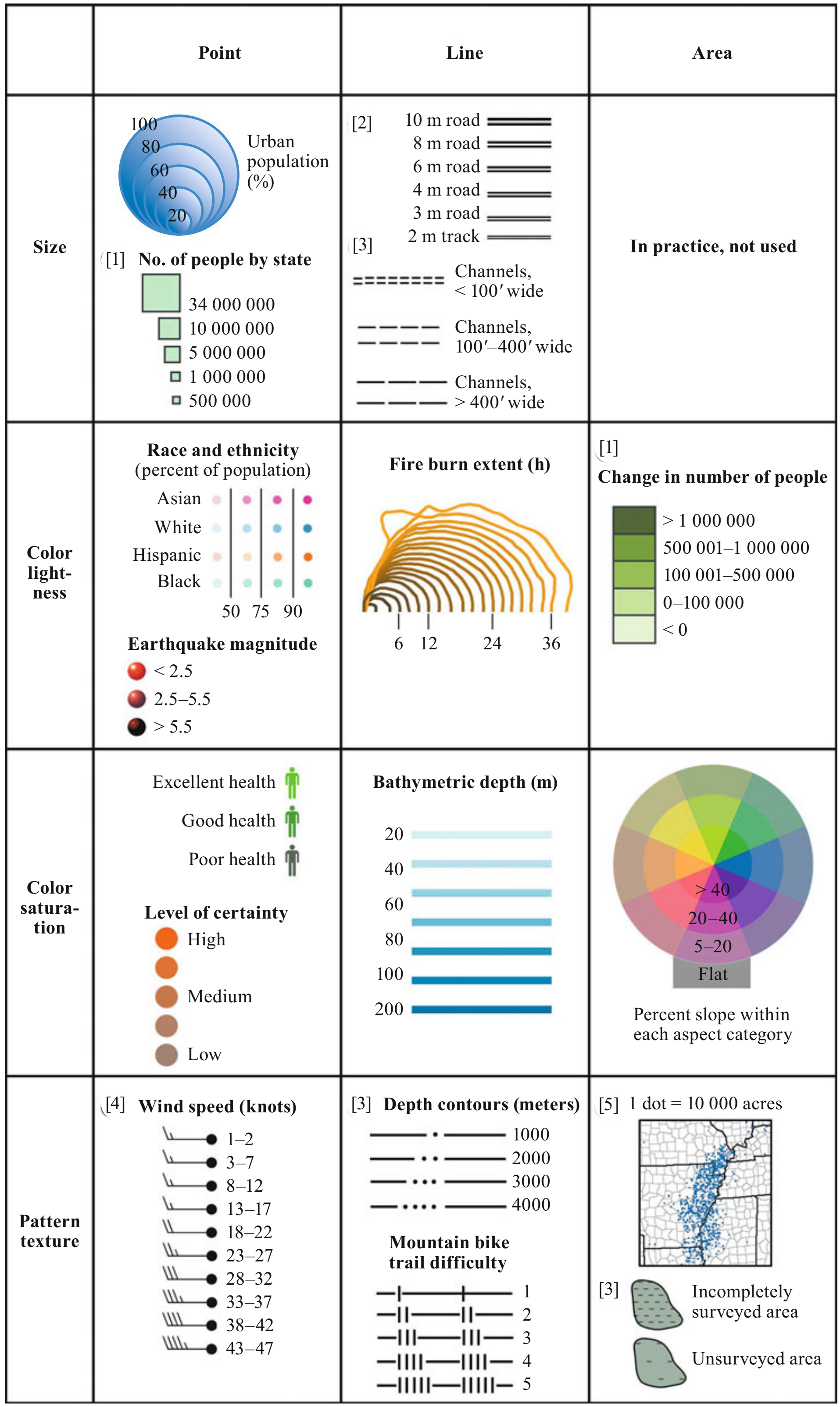
Mapmakers sometimes combine two or more visual variables to emphasize a single attribute or to show multiple attributes. To emphasize a single quantitative attribute, mapmakers can use a primary visual variable, such as size, augmented by a secondary visual variable, such as color value (lightness), in which darkness serves to exaggerate the message that is conveyed by a larger size. Cartographers might also combine two visual variables to display two different attributes. Figure 13.23 provides an example in which color hue is used to differentiate the year and symbol size is used to express the number of people. Mapping two variables is achieved with a bivariate symbol, as with the circle symbols in Fig. 13.23; mapping more variables requires a multivariate symbol, such as the beacon and buoy symbols in Fig. 13.21. An example of a bivariate choropleth map is shown in Fig. 13.24. On this map, color hue differentiates the two categories of attributes (cancer mortality for males and females), and color value portrays the quantitative value of the attributes in the categories. Purple hues denote areas where the values for both males and females are similar.

Of course, the more complex the symbol, the harder it is for the map user to read the map; the longer it takes to understand the map; the greater the chance the map will be misinterpreted; and the harder it is to design the symbol (because the guidelines for each visual variable must be correctly applied). Therefore, the need for bivariate and, especially, multivariate mapping should be well justified.

13.4.3 Design Guidelines for Symbols

The range of possibilities when designing cartographic symbols is virtually limitless. While this can seem daunting to the budding mapmaker, seasoned cartographers base their decisions on the conceptual foundation for symbol design described in Sect. 13.4.1—that is, the triad of dimensionality, levels of measurement, and visual variables. The examples in Figs. 13.21 and 13.22 illustrate how these three concepts come together. Notice that the levels of measurement are reflected in the two sets of examples—qualitative visual variables in Fig. 13.21 and quantitative visual variables in Fig. 13.22. Also notice that the dimensions are represented in the columns of the two figures, and the visual variables are represented in the rows. (Symbols for 3-D features are not discussed in this chapter.) For the most part, when the three concepts are considered carefully and in concert, cartographic symbol design is controlled, confined, and

Fig. 13.22 Quantitative visual variables (adapted from US Census Bureau atlas maps [2], swisstopo national maps [3], NOAA nautical charts [4], NOAA weather maps [5], and US Department of Agriculture (USDA) agricultural atlas maps [6])



often conventional, although there is still room for creativity. Many symbol design guidelines have been provided earlier in Sect. 13.4; others are provided in the discussions of color (Sect. 13.5) and type (Sect. 13.6).

Sometimes map symbols can be interpreted without the use of a legend (Sect. 13.9.1 gives more details on legends). This can be achieved by using symbols that are familiar because of cartographic conventions (such as a blue line for

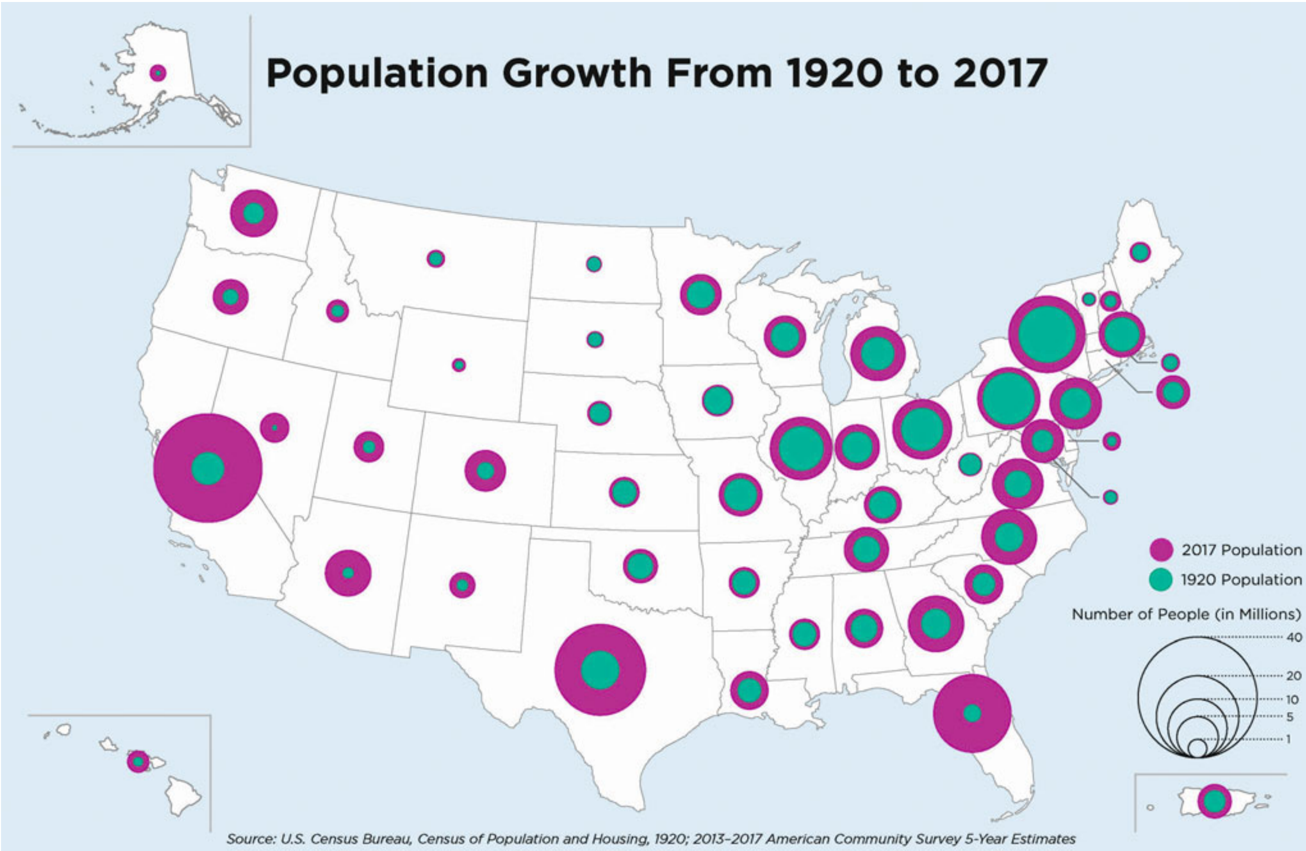


Fig. 13.23 A bivariate symbol map for US counties showing population growth from 1920–2017 (courtesy of US Census Bureau, https://www.census.gov/programs-surveys/sis/2020census/2020-resources/2020-maps/understanding-us-pop-hs.html)

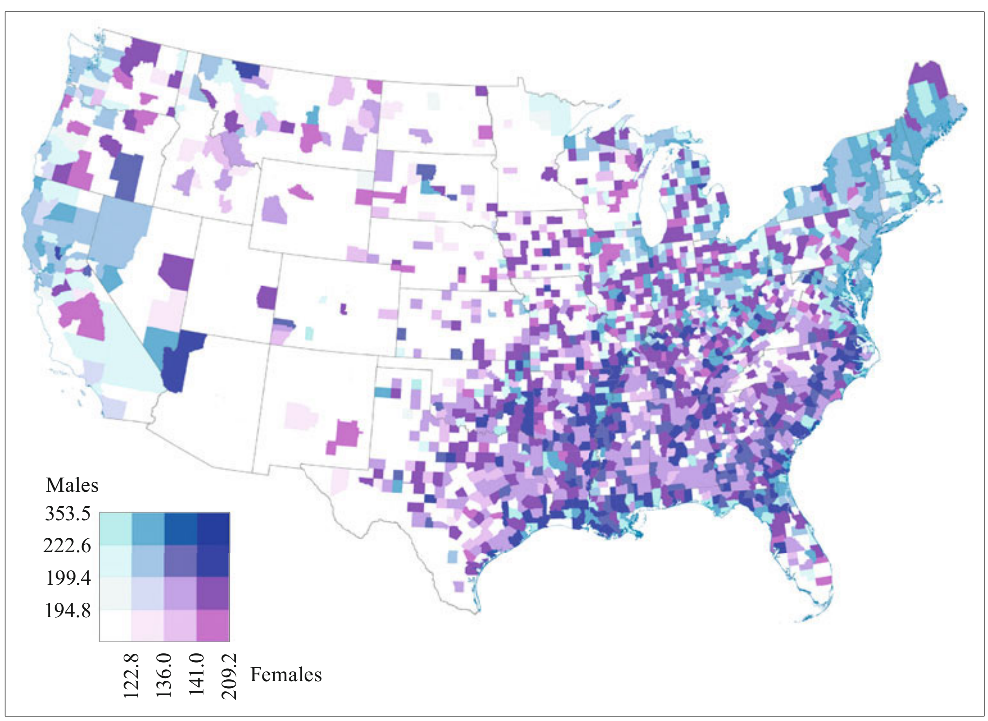


Fig. 13.24 A bivariate choropleth map for US counties showing 1970–1994 cancer mortality rates for both males and females (data courtesy of National Cancer Institute (NCI) and Centers for Disease Control and Prevention (CDC))

a river), using unambiguous symbols (for example, using a tent symbol for a campsite), or labeling features directly on the map (as with a trail or railroad). If any opportunity for misinterpretation exists, a legend should accompany the map, and the symbols in the legend should appear exactly as they do on the map. A telltale sign of a novice mapmaker is a legend with symbols or colors that are different from those on the map (for example, when a point symbol in the legend is larger or smaller than the corresponding symbol on the map or when transparency that is used on the map is not reflected in the colors in the legend).

13.5 Color

Color not only makes maps more interesting to look at, but it can also be a great aid to mapmakers, since it allows more types of features to be distinguished on the map and better order to be imposed within the map (Sect. 13.8). One of the major challenges for mapmakers is to use color in an effective manner so that it aids in communication rather than complicating and confusing the message. Today's computer systems with 24-bit color depth can display 16.7 million colors (32-bit color depth with an alpha channel supports almost 4.3 trillion color combinations), making the challenge of color selection even more difficult. Mapmakers, therefore, benefit from basic knowledge about the properties of color, familiarity with common color models, and guidelines on how to use color on maps, some of which were described in Sect. 13.4.1.

13.5.1 Properties of Color

Although one person's conception of blue may be slightly different than another person's, most people with normal color vision do not differ substantially in their color perception. Therefore, maps can usually be safely designed based on the assumption that the map user has normal color perception. There are exceptions when designing maps for specific conditions of use (for example, to be read in low light) or by specific audiences (such as readers with color vision deficiency).

Several systems based on observers with normal color vision have been developed to describe color properties and define colors. These systems have to be more detailed than simply assigning a specific color to the label *blue*, for example, since the range of possible blues renders the description meaningless. Furthermore, in English, there are only 11 basic color terms: black, white, gray, red, green, yellow, blue, brown, purple, orange, and pink (see Fig. 13.25, top). Consequently, color descriptions must be more precise. This is possible because we can describe color based on the physical stimulation of color and the perception associated with viewing colors. Being able to define color with precision is important in order to control color throughout the mapping process, from data input to graphic display to map output. For print maps, for example, the process involves selecting a color (often from a color chart) and ensuring that the color's settings are used throughout the entire design process. Because a particular color has been specified, the mapmaker can be confident it will look correct when printed on paper or displayed on a computer screen.

Color specifications are often based on the three basic elements of color (Fig. 13.25):

- Hue: commonly referred to as color, such as red, green, or blue.
- Lightness: the lightness or darkness of a hue; also called value. Lighter colors are sometimes referred to as tints—the mixture of a color with white; darker colors are called shades—the mixture of a color with black.
- Saturation: the extent to which a color departs from being neutral, such as gray, or in simpler terms, its *colorfulness*.

These three elements of color must be considered relative to the medium used to display the map and the conditions of map use. For example, color lightness can be varied to create lighter and darker versions of a hue (see Fig. 13.25,

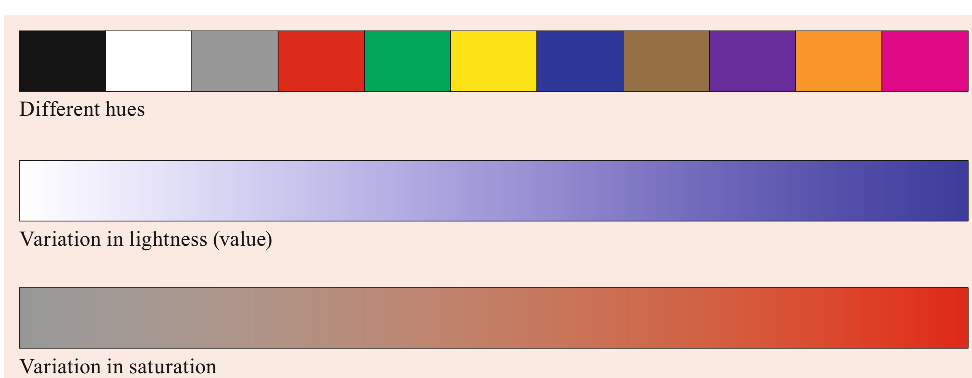


Fig. 13.25 The elements of color

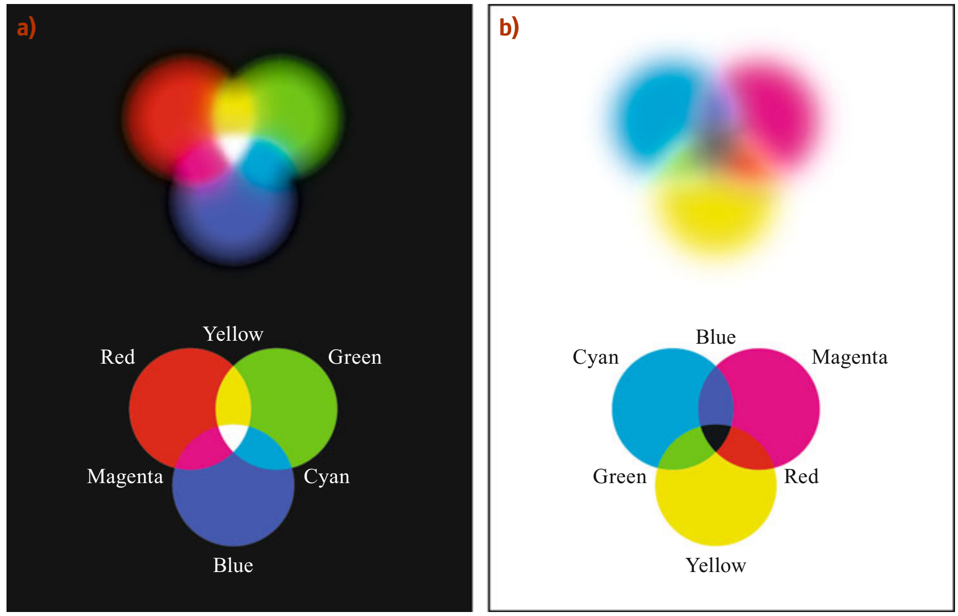


Fig. 13.26 The additive, RGB (**a**) and subtractive, CMY (**b**) color mixing systems

middle). Color saturation can change the colorfulness of a hue (Fig. 13.25, bottom). However, *lightness*, *darkness*, and *colorfulness* are relative concepts. If a map printed on paper is seen under a bright light, then the amount of reflected light is increased, and the lightness of all hues on the map will also be perceived to increase (although the relative differences between values of the hues will remain the same). This principle does not apply to a map displayed on a computer screen. In this case, the increased light of the hue is added to the light emitted by the monitor, resulting in a decrease in the relative differences in brightness in the screen image. This is one reason why problems arise when designing color for a print product on a computer and why it is best to work with color specifications rather than apparent color (an observer's perception of color, which varies depending on the light in the viewing environment).

13.5.2 Color Mixing

For cartography and other modes of graphic communication, color mixing is often based on two primary systems that relate to the technology used for production and presentation of the graphics (Fig. 13.26). The difference between the two systems lies in whether the hues are mixed as light sources or as ink pigments. For light sources, the RGB system is used. This is so named because the primary hues are red, green, and blue (RGB). For mixing ink pigments, the CMY system—based on the primary colors of cyan, magenta, and yellow (CMY)—is used.

The process of mixing colors using the RGB system is called additive color mixing. Red, green, and blue are referred to as the additive primaries in this color system. If three lights have red, green, and blue filters attached, they create beams of red, green, and blue light. Where all three beams overlap, the reflected light is perceived as white. Where only two beams overlap, the reflected light produces a mix of color. Blue and green mixed together form cyan, blue and red produce magenta, and red and green make yellow (Fig. 13.26a, bottom). By altering the intensity of each of the beams of light, different colors can be created, as in the interior space in Fig. 13.26a, top. Additive color mixing is used to display colors on computer screens in which extremely small phosphor dots arranged on a liquid crystal display (LCD) monitor emit red, green, and blue light in varying intensities. From a distance, the varying intensities reveal a sum of light and, consequentially, a color. The background color on a monitor is black, which represents no color emission.

Printing on white paper is achieved by using subtractive color mixing of three transparent inks—cyan, magenta, and yellow. These colors are referred to as the subtractive primaries or process colors. Cyan transmits green and blue light, magenta transmits red and blue, and yellow transmits red and green (Fig. 13.26b, bottom). For example, if yellow ink is printed on top of cyan ink on white paper, the yellow ink absorbs blue light but transmits red and green—the cyan ink then absorbs red light, so only the green light reaches the paper and is reflected back to the observer. Subtractive color mixing is used for digital printing with toners (like in laser printers) or liquid ink (for larger printers), and for offset printing in which the image is transferred, or offset, from a plate to a rubber blanket, and the image on the blanket is rolled onto a sheet of paper.

Because ink tones are not completely pure (in practice, they do not absorb or transmit 100% of the theoretical wavelengths of light), mixing the three subtractive primaries does not produce pure black. Hence, a black ink is usually used as a fourth printing color. Because black is normally printed first and the other colors are *keyed*, or registered, to it, it is assigned the acronym *K*. Thus, you will also see this referred to as the CMYK color mixing system.

13.5.3 Common Color Models

Colors can be defined using different color models, which are systems for specifying colors numerically according to their individual components. A color model is visualized as a three-dimensional space that contains combinations of the model's primary colors to produce all possible colors. Although a variety of color models has been developed, the RGB, CMY, and HSV models are primarily used in GIS.

The RGB color model specifies color based on the relative intensities of the red, green, and blue additive primaries on a computer monitor. The numerical values range from 0 to 255 for each of the three colors yielding 256^3 (16 777 216) possible colors. The primary colors can be represented on three axes of a cube, so that the cube becomes the color space in which any point can be given a numeric value for red, green, and blue that precisely specifies the color (Fig. 13.27). In this color model, black is defined as 0,0,0 and white as 255,255,255. Gray is scaled along the diagonal line between the two opposing corners between black and white. The secondary colors (cyan, magenta, and yellow) are situated at the remaining corners of the cube.

In the CMY color model, the CMY primary colors are analogous to the RGB primaries. The CMY color space can also be visualized as a cube with its primary colors at three corners, and its secondary colors (red, green, and blue) at the opposite corners (Fig. 13.28). Black and white are in opposite positions from their locations in the RGB color model,

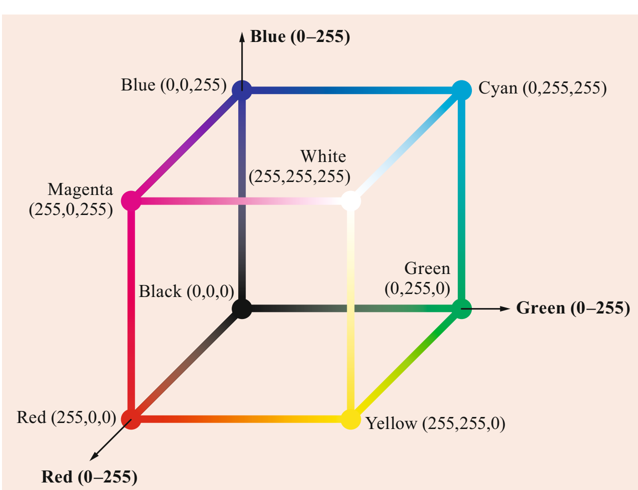


Fig. 13.27 The RGB color model

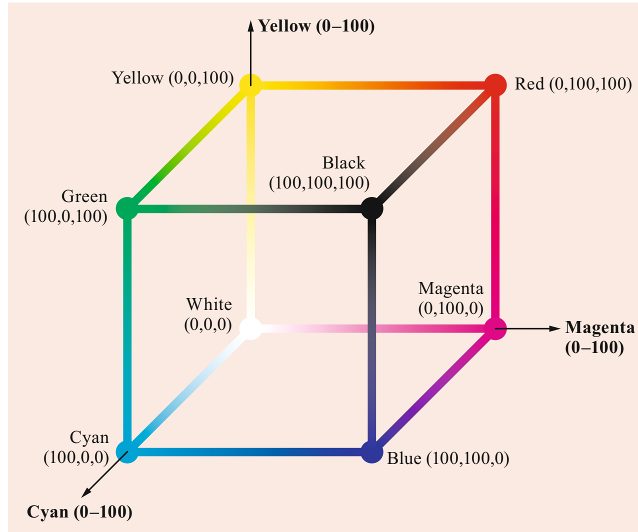


Fig. 13.28 The CMY color model

and gray is scaled between these two corners. Numeric values for the primary colors in this color model are expressed in percentages and range from 0 to 100 (although you will sometimes see the values ranging from 0 to 1 instead).

Another common color model used in GIS is the hue, saturation, and value (HSV) color model. The HSV color model can be visualized as a cone in a cylindrical coordinate system (Fig. 13.29). In this color model, the hues are arranged in the circle that represents the base of the cone inverted in Fig. 13.29 (for visualization purposes) so that the hue element of color can be expressed as an angle varying counterclockwise from red at zero. Saturation ranges from 0% (gray) at the middle to 100% at the face, or the curved surface of the cone's exterior, where the colors are fully saturated. Value ranges from 0% (black) at the apex of the cone to 100% at the base. White is located at the center of the base as the point with 100% value, 0% saturation, and no hue. Gray tones fall along the vertical axis from this point to the apex of the cone, where black is defined as 0% value, 0% saturation, and no hue.

The HSV color model is actually a rearrangement of the geometry of the RGB and CMY color models. It is more intuitive to many mapmakers because it allows them to work directly with the hue, saturation, and value (lightness) visual variables, as opposed to proportions or percentages of color determined by the technology used for production.

13.5.4 Design Guidelines for Color

When designing cartographic symbols, some relationship should exist among the color that is chosen, the data it is portraying, and the meaning to be communicated. For example, bare earth might best be symbolized with a brown hue because that is how dirt appears in the environment

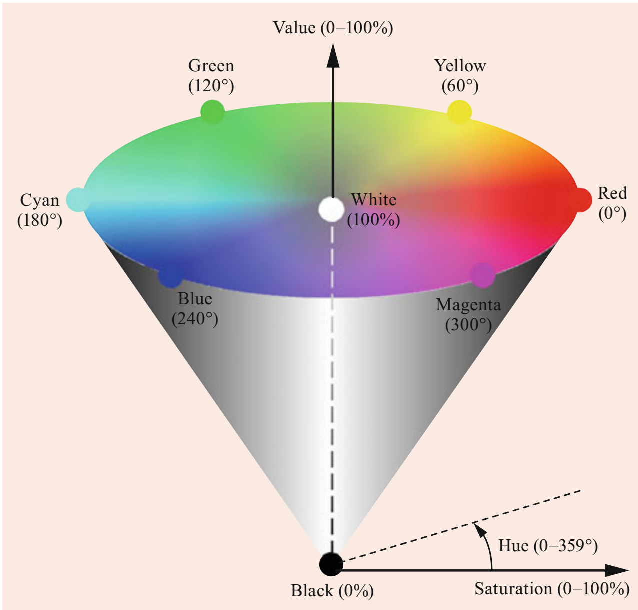


Fig. 13.29 The HSV color model

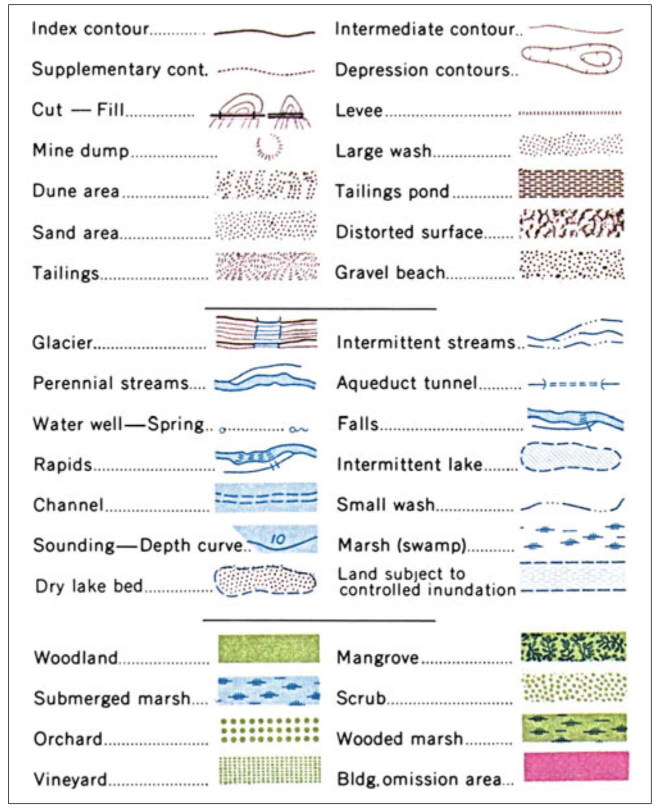


Fig. 13.30 Some of the symbols for the 1 : 24 000-scale USGS topographic map series (courtesy of USGS)

(Fig. 13.30). For some maps, the decisions about color and many other properties of symbols are dictated by specifications that override any creative decision-making on the mapmaker's part. For example, symbols on the 1 : 24 000-scale USGS topographic map series have been meticulously specified and the mapmaker is obligated to adhere to the relegated standards (Fig. 13.30).

Over time, a history of using certain colors for specific features has resulted in cartographic color conventions. Examples include blue for water features, green for forested areas, brown for contours, and black for political boundaries. It is important to keep in mind that color and other symbol conventions vary by locale and map type. For example, in one country's topographic map series, major roads may be shown with a line symbol that has black on the outside edges and a color, such as red, in the middle (this is an example of a cased line symbol), as with the first four symbols in Fig. 13.45. In another country, the interior of the symbol may be shown in a different color, for example, blue and pink in Great Britain (see the road examples in Fig. 13.21). It is also important to note that color conventions can conflict, as when two different symbols are used to show two different attributes of a feature. For example, an area may be in a certain jurisdiction that should be colored, for instance gray, and it may also be vegetated. If jurisdiction is the more important attribute for the map's purpose, its symbol will take precedence over the green fill that would conventionally used for vegetated areas. Nonetheless, color conventions allow mapmakers some decisiveness in their use of color.

When color conventions do not exist or maps are not designed to meet a specification, mapmakers can fall back on color guidelines developed through practice or research in cartography. A few are listed here as examples.

- When possible, colors that mimic the visual appearance of the features being mapped should be used, as with the green symbol for woodland in Fig. 13.30.
- To show associated features, color hues that are related in some fashion should be used. For example, different green colors could be used to show variety in the types of forested areas.
- The background of a map is often shown with a light, neutral or near-neutral color to ensure that it recedes visually; the foreground detail on a map is often symbolized using darker or more saturated colors.
- Symbols can be augmented with other visual variables, such as the submerged marsh and wooded marsh symbols in Fig. 13.30 in which the blue fill indicates submerged areas and the green fill indicates wooded areas.
- The emotive use of color should be avoided (for example, using highly saturated reds to symbolize disease).
- To draw attention to potentially ignored, small, or prominent features, use more saturated or darker colors than the surrounding areas.

Once the color use guidelines are known, they can be broken with a mindfulness toward achieving a specific goal or affective design (the look and feel of the map). For example, the use of darker colors for the background and lighter colors for the thematic information is a departure from the norm, so it can create a visual impression that is unique and, therefore, intriguing to map readers. The same holds for maps in black and white, since most maps today are made in color.

Color is extremely important in map design; therefore, understanding color definitions, perceptual qualities, and cognitive associations is helpful to mapmakers. For more on color in cartography, see [2] and [9].

In this section, much of the discussion focused on the use of color for reference maps. In the next section, the focus is on color use for thematic maps.

13.5.5 Color for Qualitative Thematic Maps

When considering color for thematic maps, cartographers often choose among cartographic color schemes, which are sets of colors to symbolize qualitative data categories or quantitative data classes on a map. Qualitative color schemes with varied hues should be used to represent different types of features (Fig. 13.31). For example, the land cover map in Fig. 13.19 has categories for water, developed land, forests, pastures, crops, and more. Each category should be visually distinct, without giving prominence to any one category or suggesting order in the categories. The simplest way to achieve this is to ensure that the different hues maintain a similar contrast with the background color of the map by controlling the value and saturation of each color. In general, this is not a problem for maps on which a single color or a limited number of colors are used for the background or maps with a limited set of classes or categories. However, this can be problematic on other types of maps, such as a thematic map overlaid on an aerial or satellite image in which the background has a lot of variation.

13.5.6 Color for Quantitative Thematic Maps

Color lightness, color saturation, or their combination should be used to represent order or numerical differences in quantitative attributes. In general, larger values or magnitudes should be shown with darker and/or brighter colors. Good examples of quantitative (also called sequential) color schemes are shown in Fig. 13.32.

The order of the quantitative classes shown on a map is also an important consideration in the selection of colors for a color scheme. On maps in which the areas between isolines are colored (Fig. 13.10), each class will always be portrayed next to its adjacent class (if the class is at the extreme of the range) or classes (if it is in the middle of the range). In contrast, on maps like choropleth maps (Fig. 13.5), any class has the potential to appear next to any other class or the same class. This can lead to the problem of simultaneous contrast in which two colors, side by side, influence each other and change the way we perceive each color. The effect is even more noticeable when complementary colors, such as red and green, yellow and purple, or orange and blue, are involved.

The number of classes used to portray the range of quantitative data is also important to consider, because it will impact how visually distinct each class is from the others. The smaller the number of classes, the larger the visual distinction, and vice versa. This gives rise to the conventional cartographic wisdom that choropleth maps should have no more than eight classes, although for some hues, like yellow, the number of classes should be even more limited. When too many classes are shown with a single color hue, map readers will find it difficult to determine to which class a feature on the map belongs, even when they study the legend. With isoline maps, the number of classes can be greater because of the forced adjacency of classes and their

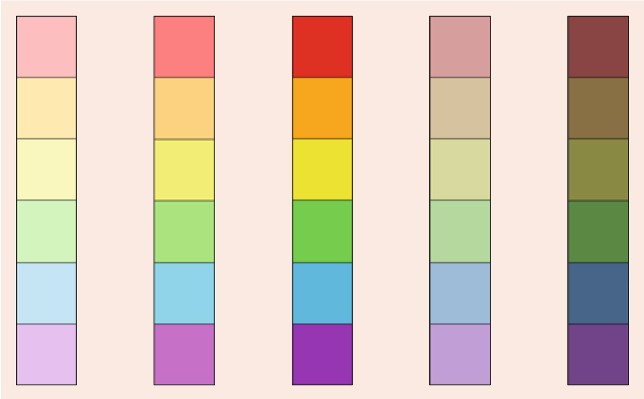


Fig. 13.31 Qualitative color schemes

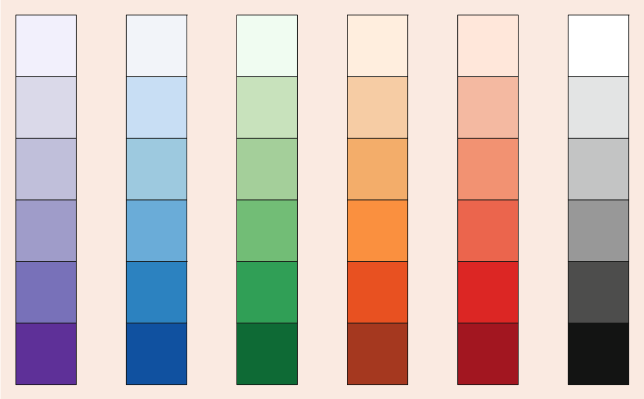


Fig. 13.32 Quantitative (sequential) color schemes

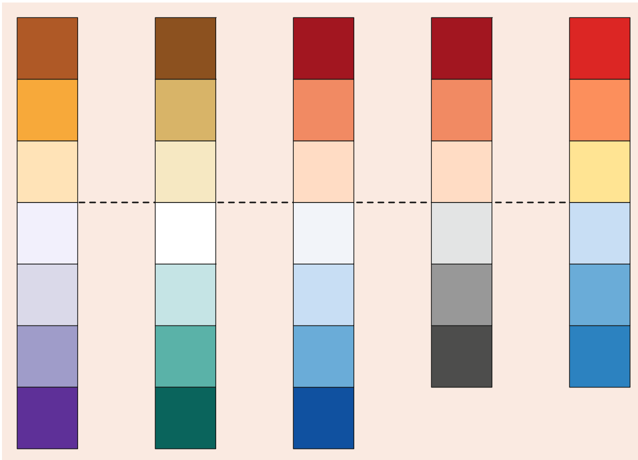


Fig. 13.33 Diverging color schemes

colors on the map as with the 11 classes in the legend in Fig. 13.10.

Figure 13.10 is shown with a diverging color scheme (sometimes called a bipolar color scheme) which is used for data that varies from a midpoint or critical value. A diverging color scheme portrays the midpoint or critical value with a neutral or near-neutral color, and values above and below that are shown with progressions of two different hues varying in lightness and/or saturation (Fig. 13.33). The midpoint or critical value can be either the single class in the middle or the break between the two different color progressions.

Choosing an appropriate color scheme is a challenge for many mapmakers, even seasoned professionals. A great resource is ColorBrewer 2.0 [10], an interactive web-based tool that presents color scheme options based on the mapmaker's requirements, such as the nature of data being mapped, number of classes, color and complexity of the background, usability for color blind readers, print friendliness, and ability to be photocopied. Like the examples in Figs. 13.31–13.33, ColorBrewer shows colors for symbolizing areal features; however, the same colors can also be applied to the design of point and line symbols.

13.6 Type

The great majority of maps make use of cartographic type to convey important or additional information about mapped features. Typical uses of type on a map include labeling places with their names ("Fort Scott" in Fig. 13.2) or numbers (101 the highway shield in Fig. 13.2), labeling features with values (as with the elevation values shown on the contours in Fig. 13.2), describing the environment ("Fort Point National Historical Site" in Fig. 13.2), locating a feature without using symbology ("Baker Beach" in Fig. 13.2), or indicating features of vague extent ("South Bay" in Fig. 13.2). Type is also used in titles, legends, text blocks, and other elements that accompany the map (see Fig. 13.5, for example). As with color, mapmakers benefit from basic knowledge about the properties of type and have, over time, developed guidelines on how to design and place type on maps.

13.6.1 Properties of Type

As in any profession, type designers have a specialized vocabulary to describe type and its properties (Fig. 13.34). Familiarization with this terminology makes it easier to communicate about typefaces and their properties and to recognize the underlying structure of various designs and the differences among them.

A typeface is a group of characters—letters, numbers, punctuation marks, mathematical symbols, and alternate or other characters—that share a common design or style. Helvetica, Arial, and Palatino are examples of typefaces. The different options available within a typeface make up a font family or type family—a set of characters that have the same basic qualities in their design, although their sizes, styles, and weights can vary. For cartographic purposes, we can divide typefaces into two main categories: serif and sans serif. A serif is a short line or stroke appended to or extending from the open ends of a character (Fig. 13.35). Sans serif (which literally means without line) fonts do not have these appendages.

A font is a set of characters within a typeface. A font is distinguished by its form (upright or italic), weight, width, ornamentation, and designer or foundry.

While the fonts within a family will vary based on their form, weight, and other characteristics, their essentially similar design is what ties them together. Thus, they are a family of fonts or a type family. Examples of font families for the Arial typeface include Arial, **Arial Black**, Arial Narrow, **Arial Rounded MT Bold**.

Many type families are available, at a minimum, in styles that include weight (such as ultralight, light, demibold, and bold or black) and form (upright, also known as roman, and italic, which is a uniquely designed version of a typeface that slants from left to right). Other families are much larger and contain options for additional variations, such as the width of the characters (for example, narrow, which is also sometimes called condensed, and expanded) and ornamentation (serif or sans serif). Size is another type property. This is generally expressed as a number of points (there are 72 points in an inch) or count of pixels (which is a digital expression) representing the height of the characters.

All these properties of type can be varied and give a font its unique appearance. Figure 13.34 illustrates a small set of

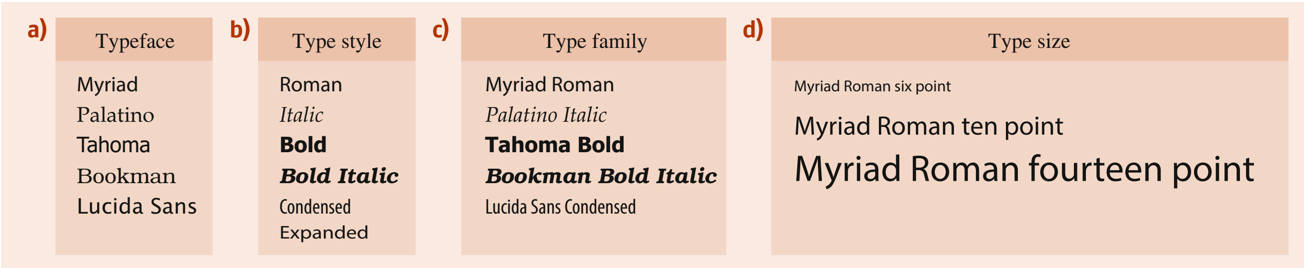


Fig. 13.34 Variations in type

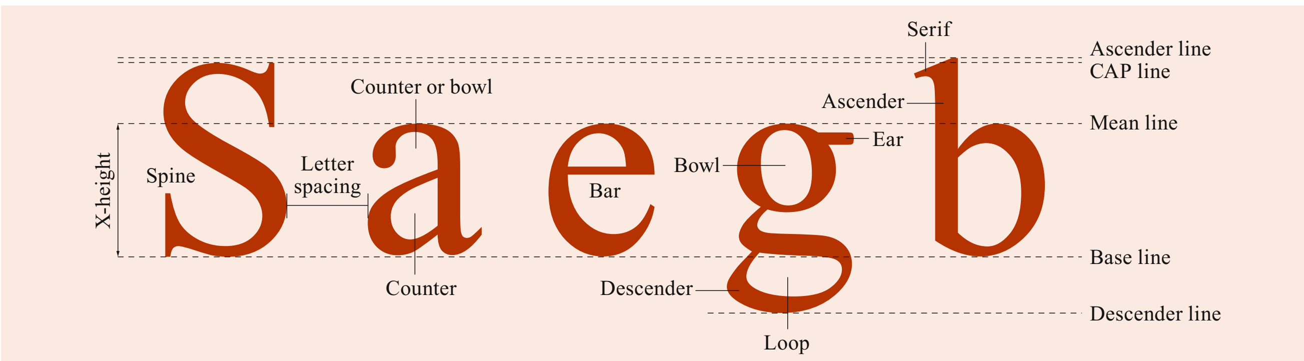


Fig. 13.35 The anatomy of a character (adapted from [11])

typefaces; some possible variations in style; and font or type families, which can then be varied in size.

A font can also include uppercase, mixed case (also known as title case), or lowercase characters. Some fonts are designed so that all characters are in one of these cases, but for most fonts, the user can set the capitalization of the characters. The user can also underline the characters.

There are thousands of typefaces, with new ones being developed all the time. So how can one typeface be distinguished from another? One important step is to train your eye to notice the details that set one font apart from another, as described above. Another is to examine the anatomy of the characters, as illustrated in Fig. 13.35.

For mapmakers, an important consideration of a character's anatomy is whether it has serifs, and whether to use a serif or sans serif font is one of the first decisions a cartographer will often make when choosing a font. Serif fonts are considered decorative and appear somewhat old-fashioned, but they have also been credited with increasing readability, especially for longer blocks of text, because they help the eye travel more quickly and easily across a line of type. For mapping purposes, labels are generally short and small, making a good case for the use of a sans serif font, which often has a more modern appearance. Some sans serif fonts are more legible than serif fonts at any size, especially if the sans serif fonts have open counters, a large x-height, and wide letterspacing (Fig. 13.35). These types of fonts are especially good for web maps, which are often displayed at a lower resolution on-screen than maps that are printed (especially professionally) on paper.

13.6.2 Design Guidelines for Type

A selection of general guidelines for the design of type used on a map is offered here. See Brewer's book [9] for a more complete explanation and illustrations.

- The number of font families used on a map should generally be no more than three.
- Short labels, as for contours, should be shown with a sans serif font.
- When there is a scarcity of space, a condensed font can be used to fit the label into the space available.
- A serif font should be used for large blocks of text.
- Mixed case should be used for most labels on a map.
- Larger, bold, uppercase, and/or underlined fonts can be used to label larger, more important, or prominent features.
- An italic font should be used to label water bodies and other features in the natural landscape; an upright (roman) font should be used for cultural features.
- Color can be used to distinguish the labels for different types of features. For example, a black font can be used to

label cultural features, brown to label terrain features, and blue to label hydrographic features.

Guidelines such as these offer mapmakers a good foundation from which to start making type design decisions. Departures from these norms can lead to interesting design alternatives that may improve a map. However, the legibility of labels is paramount (see Sect. 13.8.1 for more on legibility). Type overlap (or overlap, for short) occurs when text is overlaid on a symbol or other text. This not only affects legibility but also looks sloppy and should be avoided if at all possible.

Fig. 13.36 Label placement guidelines vary with map scale: Large scale (**a**) and small scale (**b**) examples are shown for Puerto Rico; the Columbia River, which is the border between Washington state to the north and Oregon to the south in the United States; and El Salvador (data courtesy of Esri)

13.6.3 Placement Guidelines for Type

Designing type on maps requires consideration not only of the appearance of the type but also its placement. Placing type *by hand* (positioning each label individually, even on a computer) is a slow and painstaking task. Most GIS systems have the capability to generate labels from attribute values and place the labels relative to their associated features. Labeling engines allow the user to specify rules for both type appearance and placement. Advanced labeling engines allow mapmakers to define a greater number of parameters, particularly for placement, as well as assign label and feature weights. These weights are used to specify the relative importance of labels and features to resolve conflicts when there is a situation involving overlap.

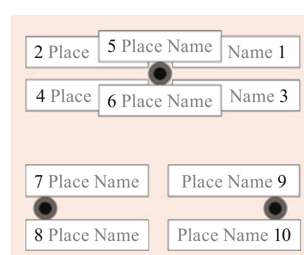
There is a tradeoff between simple but fast labeling engines, which may result in labels in dense areas being placed on top of one another or on underlying features, and intelligent labeling engines, which iterate solutions to find the best places to position labels for minimum conflict (that is, overlap). Advanced labeling engines can even change properties of the labels to fit the geographic context (for example, expanding the letter and word spacing to fill a large area or condensing the spacing or using an abbreviation to fit a label into a tight space). Advanced label placement requires contextual analysis and conflict resolution, which can be computationally intensive tasks.

When placing labels manually or setting up rules for automated label placement, cartographers often rely on best practices that they have developed practically and tested scientifically over time. Their first source of reference for label placement is often Imhof's article [12]. Imhof starts by advising that placement guidelines are always dictated by map scale. For example, on a large-scale map, the label for an areal feature may fit within the area, but on a small-scale map, the label should be placed as though it relates to a point or linear feature (Fig. 13.36).

Imhof's treatise on label placement is quite extensive, and it includes labeling the following: point, line, and area features (an example for point features is shown in Fig. 13.37); points along linear features, shorelines, or coastlines; mountain features; and features on polar maps. Imhof also offers guidance on lettering direction, label spacing, label overlap, type combined with other map contents, and the total impression of lettering. Instead of having to remember all these details, mapmakers can take comfort in knowing that many of Imhof's guidelines have been encoded in GIS labeling engines.

Although mapmakers attempt to avoid placing type over other symbols or labels on the map so that the type is not difficult to read, this is sometimes impossible, especially when type is placed over features to label them (for example, contour labels over contour lines). When it is not possible to

Fig. 13.37 Preferred label positions for point features (adapted from [12])



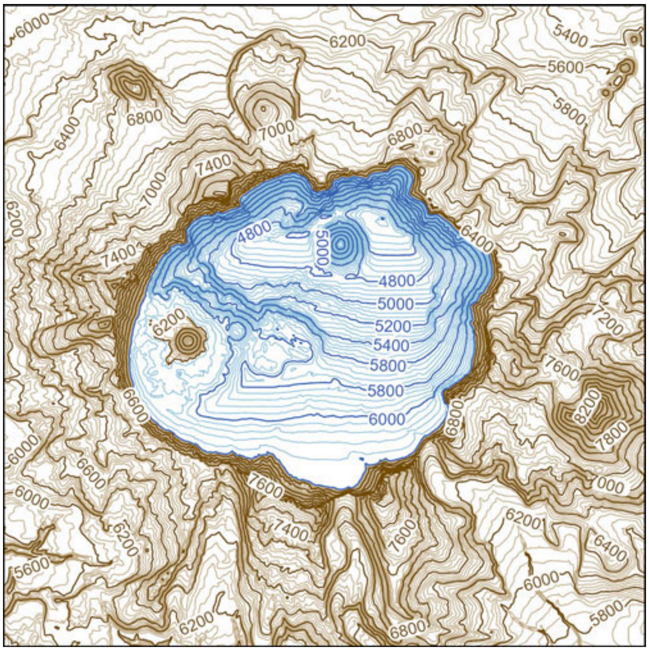


Fig. 13.38 Contours and bathymetric contours and their labels for Crater Lake National Park in southern Oregon, USA (data courtesy of Esri)

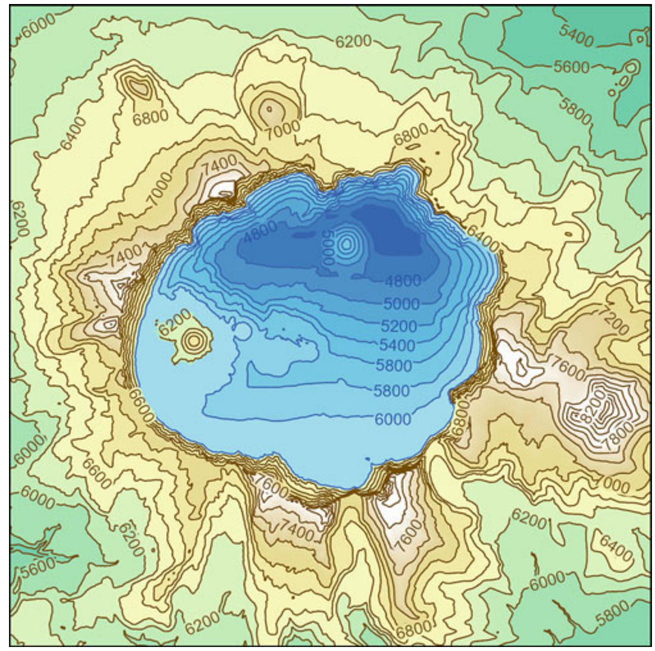


Fig. 13.39 Layer tints for the terrain and bathymetry (data courtesy of Esri)

avoid such overlap, cartographers rely on techniques to mitigate the impact of overprinting. A common technique is to use a halo—a buffer around the label filled with the color of the background, such as the contour labels with white halos in Fig. 13.38. Another technique is to knock out, or erase, only the conflicting areas of the underlying symbols or labels, as with the contour and bathymetric contour labels in Fig. 13.39. This is often a better solution when the background is multicolored.

13.7 Relief Portrayal

Relief portrayal is used on maps to show the topography (elevation) and form of the land surfaces, as well as the bathymetry (depth) of water bodies (Figs. 13.38 and 13.39). Because GIS has made it easy to display relief on maps; however, if the relief is not directly relevant to the subject of the map, as is often the case for thematic maps, then it should not be portrayed, as this could confound the message of the map and compromise the user's ability to clearly see the most important features. A definitive work on the theory and methods of relief portrayal in cartography was written by Imhof [13]. Many of his methods have been encoded in GIS software, resulting in the ability to portray the land's form in many interesting and useful ways, often with relative ease. GIS can be used to portray absolute relief by showing precise elevation (distance above sea level) and depth (distance below sea level) information and relative relief to give a general impression of relative heights and depths. Both absolute and relative relief can be shown on maps, and often they are shown together, as described below.

13.7.1 Absolute Relief Portrayal

One of the easiest ways to show relief is to label the elevation of surveyed locations on land with spot heights, sometimes improperly called benchmarks (see the small black labeled circles in Fig. 13.46), and labeling depth under water with soundings (see the gray labels in Fig. 13.4). Maps with these types of labels are often augmented with contours (isolines of equal elevation) or, for water areas, bathymetric contours (also called isobaths or depth curves). For terrain, the contours are traditionally shown as brown lines at a specified contour interval (the difference in value between contours), as illustrated in Fig. 13.38. Every fourth or fifth line (depending on the contour interval) is labeled and shown with a thicker line (these are called index contours), while the remaining intermediate contours are thinner and sometimes lighter in color (Fig. 13.30).

Despite the simplicity of the concept of contours, many map readers have difficulty interpreting relief based on this mapping method. Mapmakers will, therefore, sometimes color between the lines, creating what cartographers call a layer tint (Fig. 13.39)—this is called a hypsometric tint for land elevation and a bathymetric tint for water depth. A layer tint aids in interpretation if the colors are chosen appropriately. The goal is to select colors that mimic the land cover

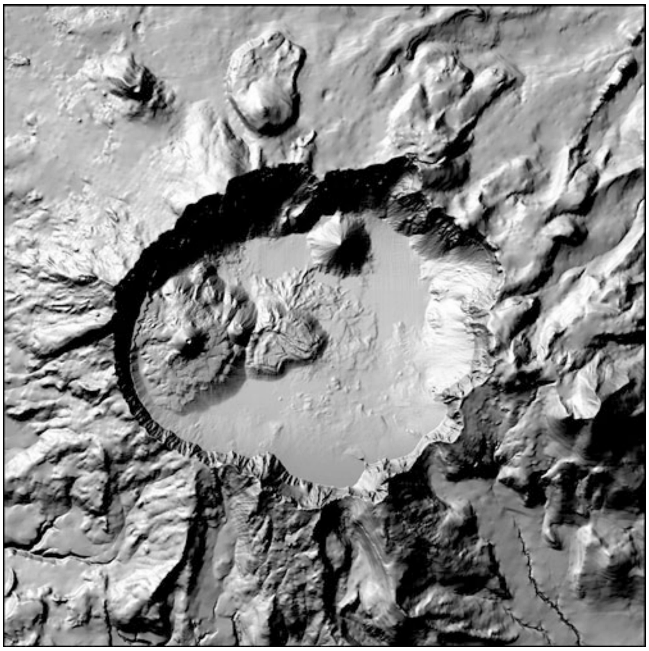


Fig. 13.40 A default hillshade (data courtesy of Esri)

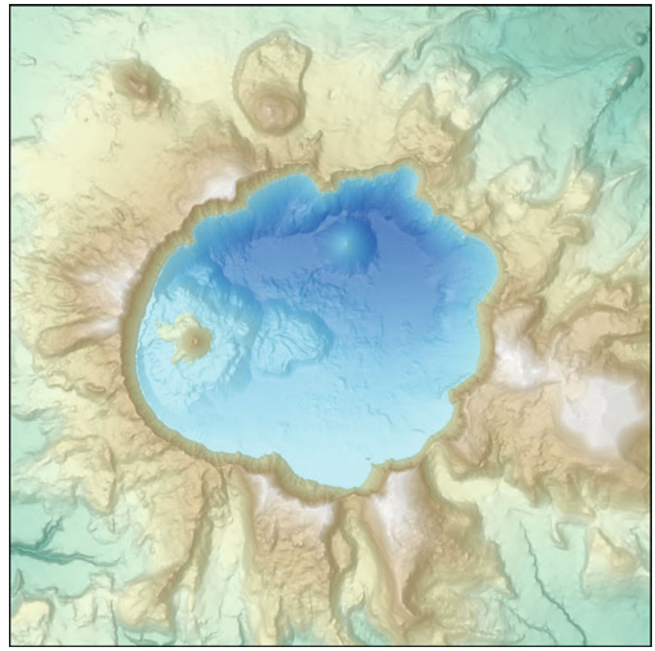


Fig. 13.41 A multidirectional hillshade overlaid with a hypsometric tint (data courtesy of Esri)

or give an impression of the depth of the water. For example, in the Crater Lake area shown in Fig. 13.39, lowlands and valleys are shown in green; higher, scrub-covered elevations are yellow-brown; treeless mountaintops are brown; and the snowcapped peaks are white. The bathymetry of the lake is shown in blue, with deeper depths having darker shades.

13.7.2 Relative Relief Portrayal

Elevation data is often available in a raster format called a digital elevation model (DEM) in which each pixel or cell has an elevation or depth value. DEMs can be used to generate a variety of relief representations, of which the most common is a hillshade. Hillshading (also called shaded relief and relief shading) simulates the light and shadows that would be cast on the surface from a source of illumination, such as the sun. In Fig. 13.40, the illumination is from a northwest source that, by default, is at 45° above the horizon and 315° from north.

A relief portayal method that is currently popular is called multidirectional hillshading, which involves illuminating the surface from light sources in multiple directions, although the strongest light still comes from the northwest, as in Fig. 13.41. Hillshades are often overlaid with a layer tint, which can also be generated from a DEM. While the colors are again selected to simulate the land cover or the depth of the water, the visual effect is one of continuous change rather than abrupt breaks at contours (Fig. 13.41). This eliminates the need for contours and labels, freeing the background to support overlaid features and their labels. However, the result is a relative representation of relief, rather than an absolute one.

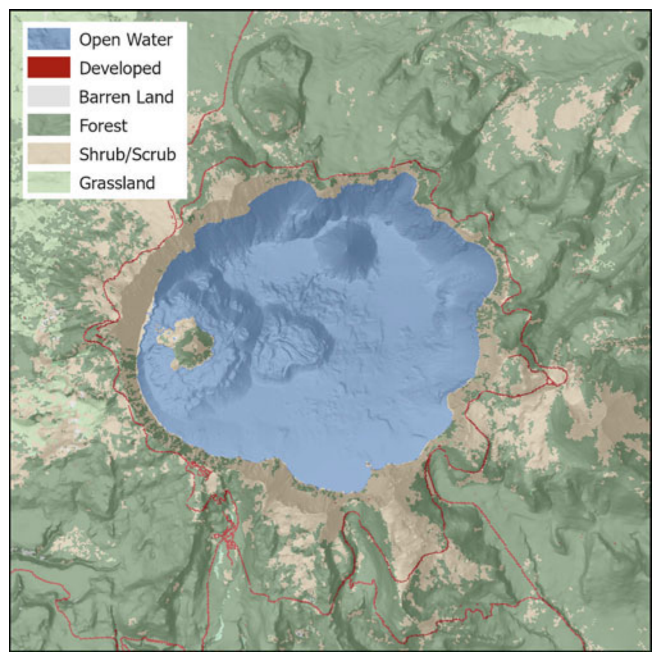


Fig. 13.42 A multidirectional hillshade overlaid with National Land Cover Data (NLCD) (data courtesy of Multi-Resolution Land Characteristics Consortium and Esri)

For terrain surfaces, a hillshade can be a subtle graphic (using a range of light gray tints), so other layers, aside from hypsometric tints, can also be displayed over the surface. It is often best to overlay only those themes that have some rela-

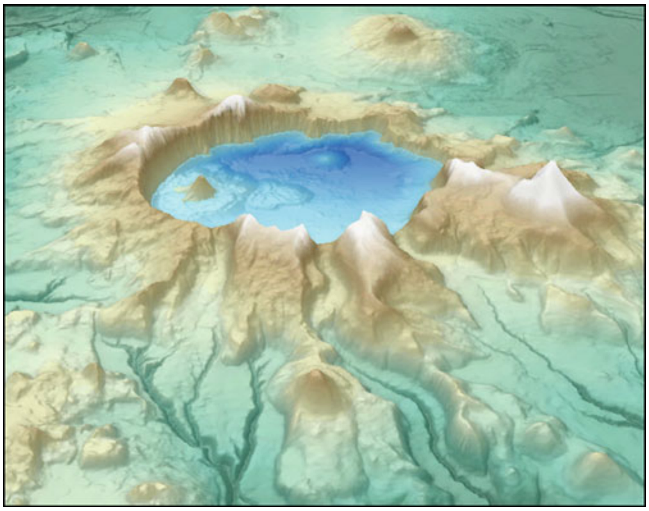


Fig. 13.43 An oblique view of the terrain with 3x exaggeration (data courtesy of Esri)

tion to terrain, such as geology or land cover (Fig. 13.42), but you will often find other non terrain-related layers displayed over hillshaded surfaces. It is also common to see an aerial or satellite image draped over the surface and sometimes even a scanned image of a printed map.

Another method of relief portrayal is to switch from a planimetric view looking down on the landscape from directly above to an oblique view looking across the surface from an angle. Sometimes, vertical exaggeration is added to the DEM values to make the view more dramatic. The terrain in Fig. 13.43 is exaggerated by a factor of 3, such that the elevation values in the DEM are multiplied by 3. For comparison purposes, the bathymetric data has not been exaggerated. It is important to remember when looking at these types of portrayals that the variations in relief on the ground are not as great as they appear on the map.

Three-dimensional cultural features, such as buildings, and natural features, such as trees, as well as other symbols and text can also be superimposed on the oblique view. For maps of this type, it is useful to provide the user with the interactivity to walk or fly through the view to inspect details and see obscured parts of the map.

13.8 Map Design

We have seen that symbolizing and labeling mapped features are major tasks in the mapmaking process. Another important step is to compose the map as a complete, clear, and compelling graphic communication product. When designing the map (or, at times, set of maps) and laying out the contents on the page or screen, cartographers employ many principles of basic graphic design. Knowing these design principles helps mapmakers to produce better maps and map readers to gain a better understanding of how maps are presented.

13.8.1 Legibility

For all symbols and type on the map, mapmakers must carefully consider legibility, which is the ability to be seen and recognized. The legibility of a symbol or label is related to its size and its contrast with the background and other symbols and labels (Sect. 13.8.2). Symbol and type size guidelines are offered in Tables 13.2 and 13.3. Because type characters tend to be more complex than simpler map symbols, the sizes for the type are slightly larger than for symbols at any viewing distance. For complex map symbols, it is safer to use the type size guidelines than the symbol size guidelines to ensure legibility.

A useful online resource to explore the design, arrangement, and legibility of symbols is the Map Symbol Brewer by Schnabel [14].

13.8.2 Visual Contrast

Visual contrast relates to the visual distinction between a symbol or label and what it overlays, such as the map background. Difficulties arise when the overlaid objects are too similar in appearance (for example, red text on an orange background), or when too many things are overlaid such that there is no space to see between objects to the background.

Table 13.2 Symbol size guidelines (adapted from [2])

Viewing distance (ft; ft = 0.3048 m)	Computer screen (pt)	Print maps (pt)	Computer screen (in.)	Print maps (in.)	Computer screen (mm)	Print maps (mm)
1.5	6	4	0.08	0.06	2.0	1.5
2	8	6	0.10	0.08	2.7	2.0
3	11	8	0.16	0.12	4.0	2.9
4	15	11	0.21	0.15	5.3	3.9
5	19	14	0.26	0.19	6.6	4.9
10	38	28	0.52	0.38	13.3	9.8
20	75	55	1.06	0.77	26.6	19.5
30	113	83	1.57	1.15	39.9	29.3
50	188	138	2.62	1.92	66.5	48.8
100	377	276	5.24	3.84	133.0	97.5

Table 13.3 Type size guidelines (adapted from [2])

Viewing distance (ft; ft = 0.3048 m)	Computer screen (pt)	Print maps (pt)	Computer screen (in.)	Print maps (in.)	Computer screen (mm)	Print maps (mm)
1.5	8	6	0.11	0.08	2.8	2.1
2	11	8	0.15	0.11	3.8	2.8
3	16	12	0.22	0.17	5.6	4.3
4	21	16	0.30	0.22	7.5	5.7
5	27	20	0.37	0.28	9.4	7.1
10	53	40	0.74	0.56	18.8	14.2
20	107	80	1.48	1.12	37.6	28.4
30	160	121	2.22	1.68	56.4	42.6
50	266	201	3.70	2.79	94.0	70.9
100	533	502	7.40	5.59	188.0	141.9

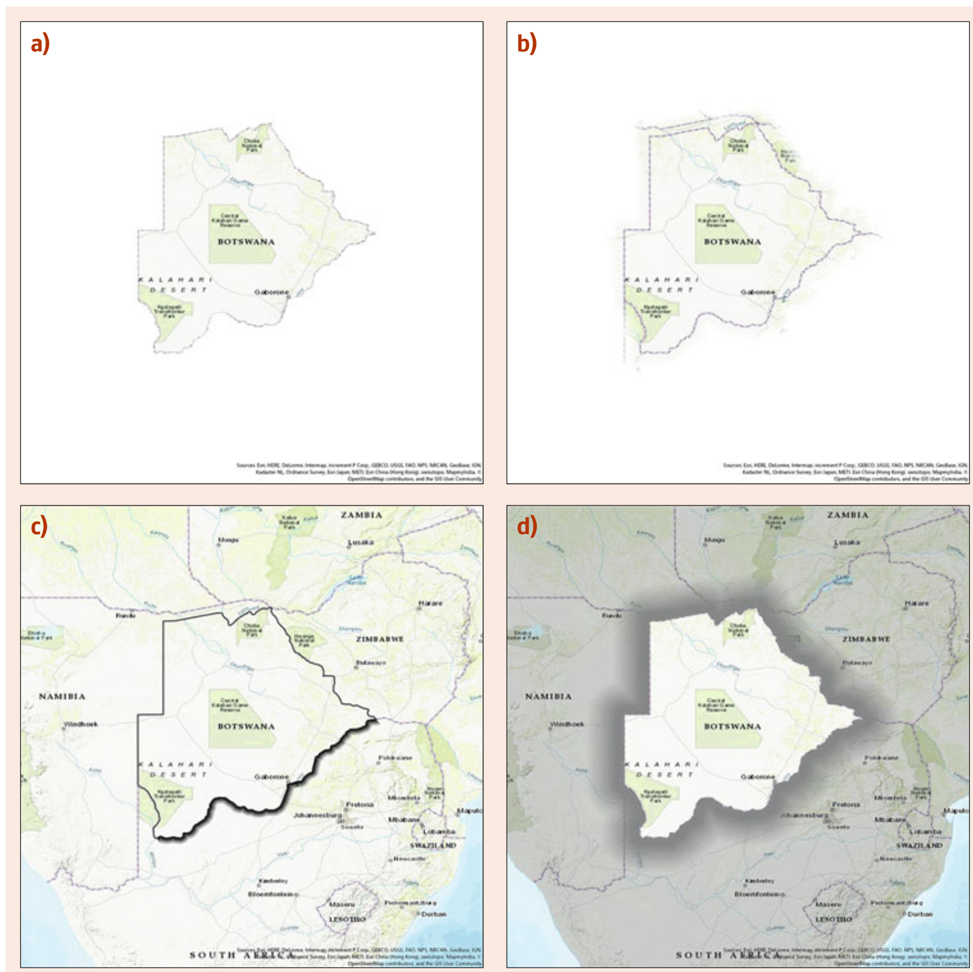


Fig. 13.44 Methods that can be used to promote figure-ground organization include a closed form (**a**), a vignette (**b**), a drop shadow (**c**), and illumination (**d**) (data courtesy of Esri)

Thus, maintaining visual contrast becomes more difficult as objects are added to the map. The most important design decision to consider in promoting visual contrast is choosing the colors of an object and its background. Map scale, function, and specifications (as illustrated in the 1 : 24 000 scale topographic map symbols in Fig. 13.30) will also impact the mapmaker's ability to maintain good visual contrast throughout the map.

The methods used to promote visual contrast also relate to the conditions of map use. For example, if the map is to be used in low-light conditions, the design will necessarily change. This is apparent on mobile maps (such as those displayed on smart devices) that automatically change from dark objects on a light background in daylight conditions to light objects on a dark background in low-light environments.

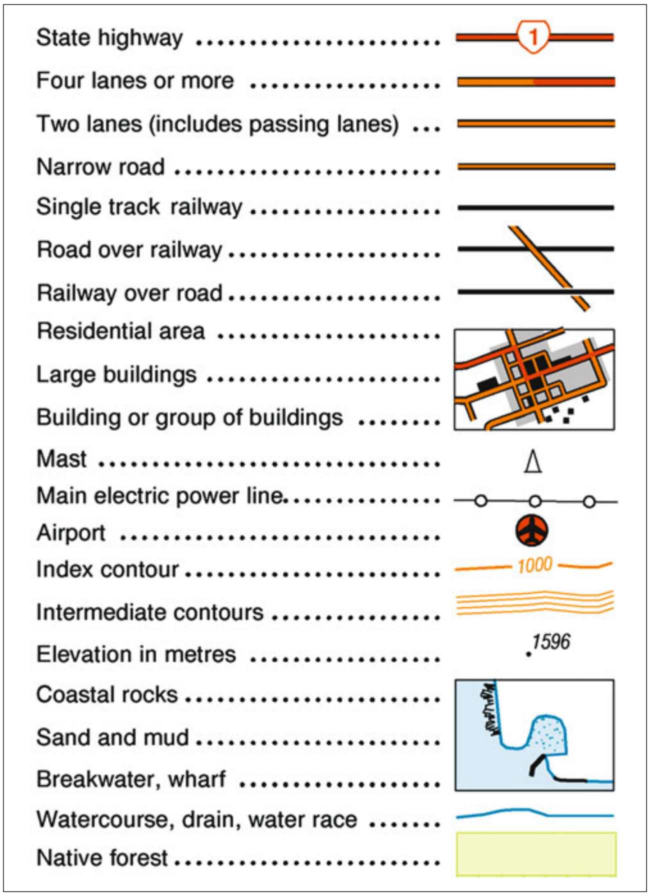


Fig. 13.45 Legend for the map in Fig. 13.46 (courtesy of Land Information New Zealand (LINZ))

13.8.3 Figure-Ground Organization

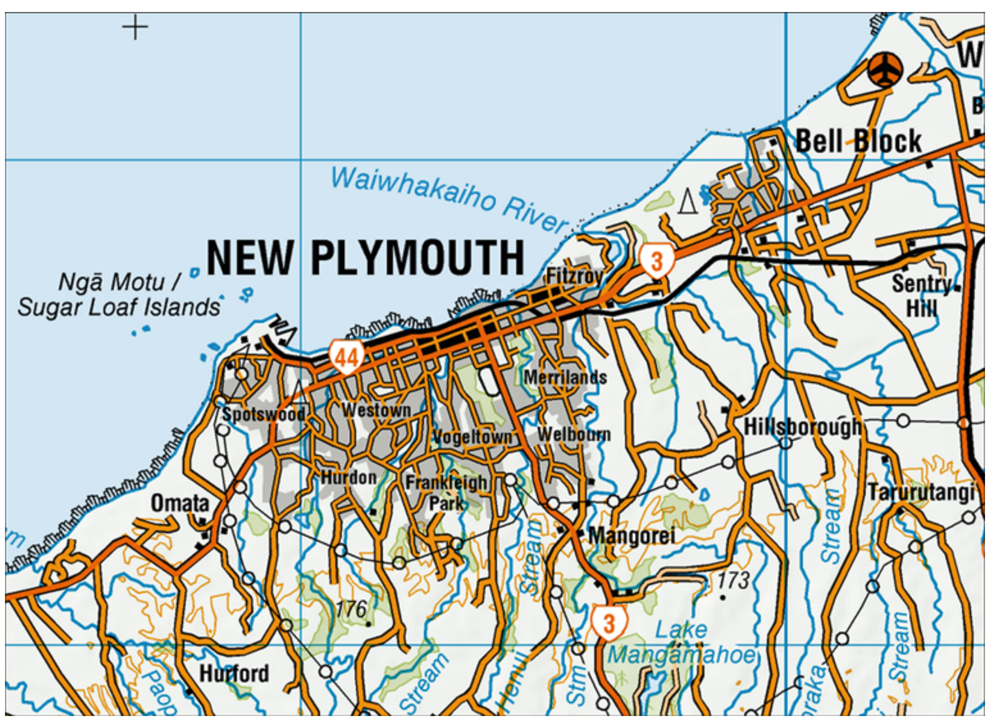
A graphic design concept that is similar to visual contrast is figure-ground organization, which is the visual separation of the map into the mapped area (the figure) that stands out against an amorphous background. Examples of cartographic methods to promote figure-ground organization are illustrated in Fig. 13.44. These include use of a closed form (a polygon that represents the entire mapped area), a vignette (a buffer around the perimeter of the figure that fades into the background), a drop shadow (a thin vignette along the perimeter of the figure offset in only one direction—generally, the southeast—giving the impression that the figure is raised above the ground), and illumination (a vignette that fades into a background that is darkened thus "illuminating" the figure). Other methods also exist, and many of the methods can be combined.

13.8.4 Visual Hierarchy

Visual hierarchy relates to the internal, graphic structure of the content on the map. It is used to communicate the relative importance of mapped features through a visual impression that categories of features are prioritized or ordered. Correctly applied, visual hierarchy reflects an appropriate intellectual order by graphically emphasizing the most important map details and de-emphasizing less important features.

The categories involved in the hierarchy and their order are dictated by the map's use. For reference maps and charts,

Fig. 13.46 A topographic map with good visual hierarchy (courtesy of LINZ)



categories of map features and their labels should be organized according to a hierarchy that relates to how the features are organized structurally in the physical landscape. For example, using the legend in Fig. 13.45, it is apparent in the 1 : 250 000-scale topographic map in Fig. 13.46 that terrain, shown with contours, is at the lowest level, followed by land cover (vegetation and built-up areas); hydrography (rivers, streams, and water bodies); transportation (roads and railroads); power lines; and, at the highest visual and conceptual level, place names and other labels. For thematic maps, all reference data may recede to the background, while the thematic information takes visual prominence on the top layer (see, for example, Fig. 13.11).

13.9 Layout

Map products often require design and compilation of more than the main, and often, single map, which is primarily what has been discussed to this point in this chapter. Other content contributes to the map's message by helping to explain what is on the map or providing information that cannot be portrayed on the map. What makes up this ancillary content is often referred to as map elements (or sometimes map marginalia, although it will not always appear in the margins). When the map and its elements are organized and displayed together, the concept of layout comes into play.

Layout is the arrangement of type, graphics, and space on the page (or in the case of web and digital maps, the computer screen). Layout can be used to create a story, provide a voice, promote engagement, and impart a style, all relative to the map's intended message and use. For expediency, in this chapter, we use the term *page* with the understanding that the discussion also applies to a computer screen.

Layout starts with the page—a blank space with visible and invisible properties. Other elements are added to that space. Additional graphic design concepts arise when considering not only the map and its internal content but also the external content (the map elements) shown with the map. Three of these concepts—proportion, balance, and harmony—are discussed later in this section.

13.9.1 Map Elements

Common map elements and some of their general design guidelines are presented below.

- The title and subtitle—an explanation of what is shown on the map, including the subject or theme, location, date, and other pertinent information (number 10 in Fig. 13.47).
- The legend—an explanation of the symbols and labels found on the map, usually shown in graphic form (see, for example, Fig. 13.5), although sometimes indicated as a word phrase (such as "One dot equals 5 people per square mile" in Fig. 13.14). The legend should only include symbols shown on the map or set of maps, and all symbols in the legend should appear exactly as they do on the map (the same size, shape, orientation, level of transparency, etc.).
- The grid or graticule—an array of north-south, east-west lines based on latitude and longitude (the graticule is shown with the blue lines in Fig. 13.46); a grid coordinate system, such as universal transverse Mercator (UTM); or a grid cell coordinate system, such as the Military Grid Refence System (MGRS). (See [2] for a complete explanation and illustrations.)
- The map scale indicator—an indication of the map scale in one or more of the formats described in Sect. 13.3.1 (numbers 13 and 14 in Fig. 13.47). For small-scale maps, scale will often vary across the map, so an indication of map scale will only be true at a point or along a line or lines—as such, map scale indicators are often not used on these types of maps, and the graticule can be used instead. Map scale indicators are not pertinent to oblique views of an area.
- The orientation indicator—an indication of the orientation of the map, often shown as a north arrow, although a compass rose is commonly used on charts (the letter c in Fig. 13.47). An orientation indicator is essential if the map is not in normal orientation (with north at the top of the page). An orientation indicator should not be used on maps in which the direction to north varies (as shown by the light gray lines indicating the graticule on the map in Fig. 13.1). In these cases, the graticule can be used instead.
- A locator map or location diagram—a small map or diagram (sometimes shown as an array of rectangles or trapezoids) that indicates the extent of area shown on the main map and sometimes the extent of adjacent maps.
- An inset map—a large-scale map that shows an enlargement of an area with dense symbology and/or labeling on the main map that would be difficult to see at the main map scale (number 18 in Fig. 13.47). Multiple inset maps may be used if there are multiple congested areas on the main map.
- Projection and datum information—text indicating the map projection and, especially for maps that are used to determine distance and direction, the datum, which is the reference system against which position and elevation and depth measurements are made (number 3 in Fig. 13.47).
- Credits and publication notes—text that indicates the author, source(s) of data (number 17 in Fig. 13.47), date of edition or publication (number 6 in Fig. 13.47), publication (number 4 in Fig. 13.47), and other pertinent information to authenticate the map.
- Neatline—an indicator of the map extent. The neatline may be simple (see the right and bottom sides of the map in Fig. 13.47), or decorative. It can also be anno-

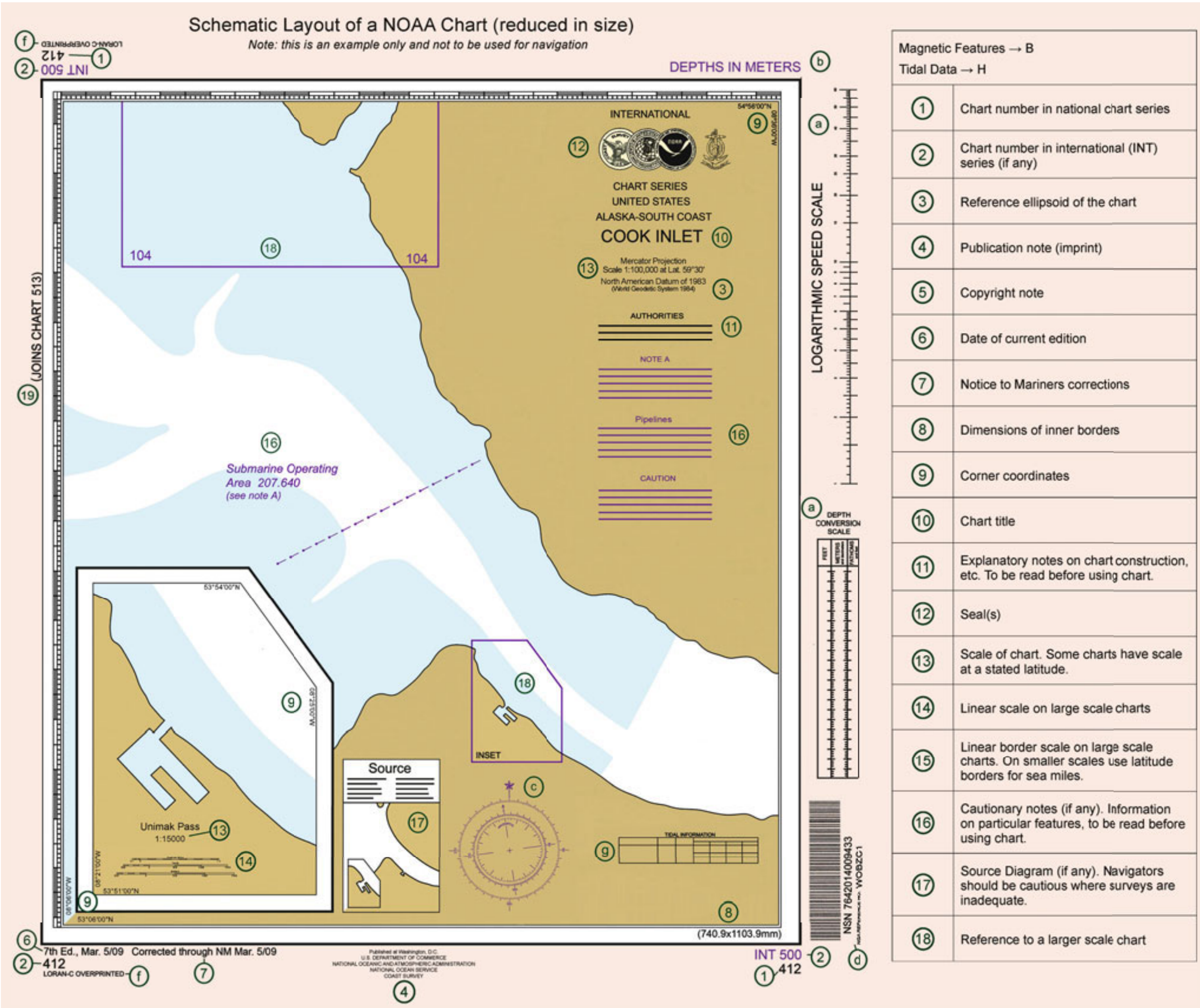


Fig. 13.47 The page layout for a NOAA nautical chart showing many of the common map elements (courtesy of and adapted from NOAA, https://nauticalcharts.noaa.gov/publications/docs/us-chart-1/ChartNo1.pdf)

tated and graduated to show the grid coordinate system or systems used on the map; this is common on charts and topographic maps. On most nautical charts, longitude is graduated along the top or bottom of the chart, and latitude is graduated along a side (see the top and left sides of the map in Fig. 13.47). The subdivisions are in degrees, minutes, half-minutes, and tenths of minutes—and, on older charts, in seconds. It is important to remember that only the vertical, latitude scale (but not the horizontal longitude scale) on a Mercator chart can be used to measure distance because a minute of latitude equals 1 nautical mile anywhere in the world, while the distance between minutes of longitude decreases between the equator and the poles.

- Text blocks, graphs, photographs, and other ancillary elements—additional content that helps explain the map and its intended message (for example, numbers 11 and 16 in Fig. 13.47).

Not all maps need all these elements, but some maps—such as topographic maps and charts—use many of them. Additional map elements may also be required, depending on the map type and intended use, as shown in Fig. 13.47.

With increased use of the web to share maps, many or all of the map elements may be compiled, along with the map (or several maps), in what is called a story map. A story map is a sequence of maps that narrate a story. It is often augmented with text, photos, and/or video, and the web application (or app) provides functionality, such as panning and zooming, pop-ups, a magnifying glass, a swipe tool, and/or a time slider—all of which help readers better understand the story.

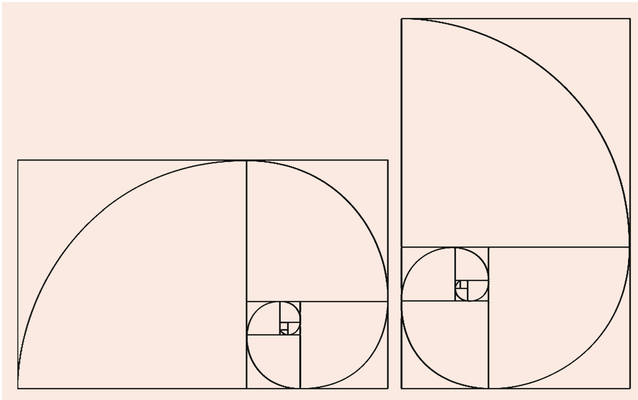


Fig. 13.48 The geometry of the golden ratio (3 : 5)

13.9.2 Proportion

Proportion is the relationship between one part or element of the map and another or the whole map. Proportion relates to both the map elements (for example, the legend) and parts of the elements (such as the legend boxes). Good proportion is achieved when a correct or desirable relationship exists between the elements with respect to size, shape, scale, color, quantity, degree, visual weight, or setting. Design concepts that relate to proportion include symmetry (one side being a reflection of the other), balance (equivalence in the visual weight of the elements), and harmony (elements appearing to belong together in size and distribution).

Symmetry is often not something that cartographers can control on their maps or, sometimes, even their layouts. The geography of the features in the mapped area may or may not be symmetrical (notice how Crater Lake in Fig. 13.38 is more symmetrical than California in Fig. 13.6). There is nothing to be done about this because that is simply the geography of the area being mapped. Also, notice that the map elements in Fig. 13.47 are not symmetrical, partly because there is a need for only one title, one legend, one scale bar, one compass rose, and so on. Sometimes symmetry can be achieved using groups of elements rather than individual elements. Balance and harmony are easier to achieve in a layout and are discussed further in Sects. 13.9.3 and 13.9.4.

When making decisions about where to place map elements in a layout, cartographers often rely on the principle of the golden ratio. This is a mathematical ratio found in both human design and nature that can, when used properly, help create aesthetically pleasing compositions in map design. The golden ratio (which is known by many alternate terms, such as the golden section, golden mean, divine proportion, or the Greek letter *Phi*) is approximately equal to 1.618, such that if one side of a rectangle is 1 unit, the other is 1.618 units. This ratio (which can also be roughly expressed as 3 : 5) is considered to produce images that are pleasing to the eye. Figure 13.48 illustrates this design concept.

The golden ratio relates to both the space on the page and the space used for the elements placed within those areas. For example, a legend designed using a 3 : 5 ratio will be more visually pleasing than one based on a 1 : 1 ratio. Similarly, if the legend itself contains rectangles to show areal feature symbols, the rectangles can also be designed using a 3 : 5 ratio to promote an aesthetically pleasing design. Notice how the design of the page and many of the elements in Fig. 13.47 employ the principle of the golden ratio.

13.9.3 Balance

Balance relates to the arrangement of the map and its elements on the page. The objective is to create visual equilibrium rather than the appearance of being off-balance in any direction. To achieve pleasing and effective visual balance, the page contents should be positioned around the visual center of the page, which is a point just above the geometric center of the page. This is the point on which the eye first focuses, and it serves as the fulcrum, or balancing point, for the page.

The impression of visual balance is controlled by the size, visual weight, and location of the symbols and labels on the map and the map elements on the page. For example, one large map element may tip the layout out of balance—to counter this, several smaller map elements can be grouped together and balanced against the larger element.

Poor visual balance leads the reader to see individual components as competing for space or to see an abundance of either congested or empty space. However, it is not always desirable to fill all the empty space in a layout. White space helps to give the impression of breathing room and create a sense of freedom for the eye to choose what to focus on. On the other hand, large unfilled spaces can throw the layout out of balance and give an impression of lack of attention to basic design. This can be mitigated in a number of ways, such as using a faded background (Fig. 13.44d) or showing the graticule outside the mapped area (Fig. 13.46).

13.9.4 Harmony

When a layout has harmony, the individual elements on the page present a meaningful whole. Harmony is achieved through the arrangement of elements to create a pleasing image in which the map and its elements complement each other and work together visually. As each element is added, its effect on the layout must be assessed, and use of the remaining space must be reevaluated. The layout may need to be reorganized several times to achieve harmony.

Harmony can be enhanced through the alignment and distribution of elements in the layout. Alignment is the arrangement—generally in an orthogonal direction relative to the page—in a straight line or in an appropriate relative position. Note, for example, that the elements at the right side of the map in Fig. 13.5 are aligned with the right side of the page, and the elements along the bottom are aligned with the bottom of the page. Proper alignment creates an ordered appearance by ensuring that the elements have an obvious and inherent relationship with each other. Proper alignment of elements eliminates the apparent messiness that can occur when elements are placed seemingly randomly. In most GIS systems, elements can be easily aligned using alignment tools or guides (usually dotted lines that appear in the layout view). Intelligent guides can be snapped to when elements are repositioned and provide an indication of when the elements are properly aligned.

Distribution is the spacing of elements so that they are equidistant. Distribution can be employed with features within elements, such as the point symbols in the legend in Fig. 13.5, or with a set of related elements, such as the elements along the right side of the page in Fig. 13.5. Most GIS systems have tools to distribute selected elements in the map layout so that they are spaced equally in either the horizontal or vertical direction.

Conclusion

This chapter reviewed the process of transforming geographic data into a map. This involves at least a basic understanding of the following: map projections and map scale; how to select, generalize, and portray that geographic information in graphic form; how to use symbols, color, and type on maps; how relief is shown on maps; and basic principles of map design and page layout. This chapter also described how cartography is a crucial part of GIS. From this chapter, you will hopefully have gained an understanding of how map use is integrally related to map compilation and design, and how effective map design cannot be achieved without keeping map use in mind.

References

1. MacEachren, A.M.: Visualization in modern cartography: Setting the agenda. In: MacEachren, A.M., Taylor, D.R.M. (eds.) Visualization in Modern Cartography, Pergamon Press, Oxford (1994), p. 1–12
2. Kimerling, A.J., Buckley, A.R., Muehrcke, P.C., Muehrcke, J.O.: Map Use: Reading, Analysis, Interpretation, 8th edn. Esri Press, Redlands (2016)
3. Snyder, J.P.: Map Projections: A Working Manual. United States Geological Survey, Washington (1987)
4. Battersby, S.E., Finn, M.P., Usery, E.L., Yamamoto, K.H.: Implications of web Mercator and its use in online mapping. Cartographica **49**(2), 85–101 (2014)
5. Dent, B.D., Torguson, J.S., Hodler, T.W.: Cartography: Thematic Map Design, 6th edn. McGraw-Hill, Boston (2009)
6. Jenks, G.F.: The data model concept in statistical mapping. Int. Yearb. Cartogr. **7**, 186–190 (1967)
7. Robinson, A.H., Morrison, J.L., Muehrcke, P.C., Kimerling, A.J., Guptill, S.C.: Elements of Cartography, 6th edn. Wiley, Hoboken (1995)
8. Bertin, J.: Semiology of Graphics. Esri Press, Redlands (2011)
9. Brewer, C.A.: Designing Better Maps: A Guide for GIS Users, 2nd edn. Esri Press, Redlands (2015)
10. Brewer, C.A., Harrower, M.: ColorBrewer 2.0 (2013). http://colorbrewer2.org/
11. Slocum, T.A., McMaster, R.B., Kessler, F.C., Howard, H.H.: Thematic Cartography and Geovizualisation, 3rd edn. Prentice Hall, Upper Saddle River (2009)
12. Imhof, E.: Positioning names on maps. Am. Cartogr. **2**(2), 128–144 (1975)
13. Imhof, E.: Cartographic Relief Presentation. Esri Press, Redlands (2007)
14. Schnabel, O.: Map Symbol Brewer (2007). https://www.oschnabel.carto.net/mapsymbolbrewer/

Further Reading

15. Allan, A.L.: Maths for Map Makers. Whittles, Dunbeath (2004)
16. Axis Maps: Cartography Guide (2020). https://www.axismaps.com/guide
17. Brewer, C.A.: Designed Maps: A Sourcebook for GIS Users. Esri Press, Redlands (2008)
18. Brown, A., Feringa, W.: Colour Basics for GIS Users. Longman, Harlow (2002)
19. Cauvin, C., Escobar, F., Serradj, A.: Thematic Cartography. Wiley, Hoboken (2010)
20. Dent, B.D., Torguson, J.S., Hodler, T.W.: Cartography: Thematic Map Design, 6th edn. McGraw-Hill, Boston (2009)
21. Kraak, M.-J., Ormeling, F.: Cartography: Visualization of Geospatial Data, 2nd edn. Pearson, Harlow (2003)
22. Krygier, J., Wood, D.: Making Maps: A Visual Guide to Map Design for GIS. Guilford, New York (2005)
23. MacEachren, A.M.: How Maps Work: Representation, Visualization, and Design. Guilford, New York (2004)
24. MacEachren, A.M.: Some Truth with Maps: A Primer on Symbolization and Design. Association of American Geographers, Washington, DC (1994)
25. Monmonier, M.: How to Lie with Maps, 2nd edn. Univ. of Chicago Press, Chicago (1996)
26. Snyder, J.P.: Flattening the Earth: Two Thousand Years of Map Projections. Univ. of Chicago Press, Chicago (1997)
27. Spence, M., Darkes, G.: Cartography: An Introduction. British Cartographic Society, London (2009)
28. Tufte, E.R.: The Visual Display of Quantitative Information. Graphics, Cheshire (1983)
29. Tufte, E.R.: Envisioning Information. Graphics, Cheshire (1990)
30. Tyner, J.A.: Principles of Map Design. Guilford, New York (2010)

Aileen Buckley Dr Aileen Buckley, a research cartographer at Esri since 2003, focuses on developing and sharing best practices for mapping. She publishes and presents widely on many aspects of mapping and GIS. She is a coauthor of Map Use: Reading, Analysis, Interpretation and a co-editor of the Atlas of Oregon. She holds a PhD in Geography from Oregon State University, and she was formerly on the faculty at the University of Oregon.

Paul Hardy Paul Hardy has an MA in Computer Science from Cambridge University, England. He spent 28 years at Laser-Scan in Cambridge, England, as Chief Programmer, Product Manager, and Principal Consultant, developing software for computer graphics and digital mapping. From 2003 to 2006, he was Product Manager for Cartography at Esri, implementing advanced cartographic capabilities into ArcGIS. From 2007 to 2016 he was Business Development Consultant for Esri Europe, advising national mapping and cadastral agencies. He has been since retired, apart from occasional consultancy.

Kenneth Field Dr Kenneth Field is a cartographer at Esri. He was formerly Principal Lecturer and GIS Course Director at Kingston University London, England. He holds a PhD in GIS from Leicester University and has presented and published widely. He is a Fellow of the British Cartographic Society and Royal Geographic Society, and he is a Chartered Geographer (GIS).