This chapter explains how geographers define and use scale in the construction of thematic maps. The choice of map scale is guided by the nature of the problem under investigation, but maps are further constrained by the amount of information that can be presented in a limited space. Generalizing a map involves simplifying the information to be presented while simultaneously enhancing the main ideas the map is designed to show. Scale reductions nearly always prompt the need for an effective generalization of map information. Technology has recently given geographers much more information to map and more precise instruments for making maps, and these developments demand even more sophisticated approaches to the art and science of map generalization.

All fields of study that make empirical observations must consider scale. When the scope of any study is set, decisions are made about what is to be studied in detail and what is to be left in the background. Scale is not a constant in most cases: it is stretched here or compressed there to emphasize what is important at the expense of what is peripheral. Whether space or time is stretched or compressed, the language used to express scale relationships is much the same. Geographers address scale in these conventional ways, but they incorporate scale in a much more specific manner in the construction of maps.

Although it is convenient to think of a map as a miniaturized picture of some portion of the earth's surface (some maps are actual photographs overprinted with additional features and labels), a geographer's map is much more than a pictorial display. It is a thesis, an implicit argument about what is important about some place. A map portrays relationships between various features that
seem to be related in a causal manner. The visibility of these features, suggested in phrases such as “bird’s-eye view” or “as seen from space,” actually has little to do with mapping. In fact, many maps show things such as language use, income, or population mobility that cannot be directly seen at all, but that are mapped in much the same way as more readily visible features. Such thematic maps are constructed for the purpose of showing some set of information in its relevant geographical context (see Chapter 4).

Like a thesis, a map originates in its creator’s mind. A good map is constructed (as is a good argument), by admitting information selectively and then presenting it in sufficient detail (but no more) that its conclusions can be evaluated. Every decision made in the mapping process is informed by the mapmaker’s knowledge of scale—the scale of the things depicted as well as the scale of the map for effectively illustrating the argument.

SCALE AND MAPS

Map scale is an intuitive concept that everyone uses without having to think about it in a formal way. It is simply the ratio between the size of some feature on a map and its corresponding size on the Earth. Scale is most often expressed as a definition of corresponding lengths, such as “one inch on the map represents one mile on the Earth’s surface.” A convention in writing such statements is to express them as fractions, in this case, 1/63,360; or, in words, one inch on the map is equivalent to 63,360 inches on the Earth’s surface. Because the fraction 1/63,360 is a smaller number than, for example, the fraction 1/31,680, we say that a 1/63,360 map has a smaller scale than a 1/31,680 map.

One of the first things a geography student learns, then, is that although scale is an intuitive concept, the terminology of large and small scale is counter-intuitive because the relevant size comparison is in the denominator of the fraction, not in the numerator. In modern practice, scale is usually written 1:63,360 or 1:31,680, but the meanings of large and small scale accord with historic practice. One mile is represented by one inch on the 1:63,360 map, whereas one-half mile is represented by one inch on the 1:31,680 map. Thus, any object on a relatively smaller-scale map will appear smaller than the same object on a relatively larger-scale map, and in this sense the usage of small scale and large scale agrees with our intuition.

Scales written in the form 1:63,360 (or 1/63,360) are called representative fractions. An advantage of this conventional definition of scale is that the representative fraction is unit-free; whether mapmakers and readers measure in inches, centimeters, or any other units, the representative fraction holds true. The reason for measuring in linear units, such as inches or centimeters, rather
than in areal or volumetric units, or even a mixture of all three, is based on underlying mathematical properties. Although linear scale is never consistent over a whole map, there is always at least one point or at least one line where the linear scale is equal in all directions. Areal scales are generally reserved for the class of map projections on which area is consistent over the whole map. Volumetric measurements are rarely made from maps.

Every representative fraction is a ratio, equivalent to what mathematicians call a rational number—a number that can be expressed as the ratio of two integers (whole numbers). And since every rational number can be mapped as a point on a line, any arithmetic operation performed on a map scale (reducing scale or enlarging scale) results in another rational number. Thus, if the scale of a map is reduced from 1:31,680 to 1:63,360, every straight-line segment on the initial map will be half as long as it was (31,680/63,360 = 1/2). But reducing the size of a line segment by half reduces the size of an area by 75 percent. If the definition of scale were in areal rather than linear units, the reduction in scale in this example would be (31,680/63,360)² = 1/4. Any segment of the reduced map's area will be only 25 percent as large as it was, in other words, a 75 percent reduction. Obviously, therefore, doubling the linear scale (63,360/31,680 = 2) yields a four-fold increase (2²) in the map's size. And just as area is the second power of length, volume is the third power. Doubling the representative fraction of a globe would result in a globe eight times the size of the original (although its surface, because it is an area rather than a volume, would be only four times as large as it was).

It follows that reducing map scale by fixed amounts results in an even greater reduction in the amount of area available for depicting the information the map was designed to portray. Nearly all mapping must confront this dilemma, and in practice a variety of scales are used in geographic research.

For example, a convenient map size of 5 by 7 inches that might be found in a textbook can show the following areas at different scales:

- a house at a scale of 1:100;
- a city block at a scale of 1:1000;
- an urban neighborhood at a scale of 1:10,000;
- a small city at a scale of 1:100,000;
- a large metropolitan area at a scale of 1:1,000,000;
- several states at a scale of 1:10,000,000;
- most of a hemisphere at a scale of 1:100,000,000; or
- the entire world with plenty of room to spare at a scale of 1:1,000,000,000.

These examples, ranging from largest (1:10⁸) to smallest (1:10⁹) scale, span eight orders of magnitude and, as a practical matter, cover the spectrum of scales at which geographers are likely to use maps.
It might seem that this sizeable range of scales results from the problem of portraying two-dimensional areas at a variety of sizes and is therefore a natural outcome of the fact that geography deals with areal relationships at many scales. But the problem is neither unique to geography nor does it result solely from a concern with areas. For example, geologists use time scales that have even a broader range. Whereas the earth is about 4.7 billion years old \((4.7 \times 10^9)\), a sand spit along a coastline might have a life span of only a few decades, while earthquake waves travel in units of time measured in minutes \((10^{-5})\) by comparison. Scientists concerned with nuclear processes measure time in units ranging from \(10^{-12}\) to \(10^{-16}\). Thus many sciences use a range of both temporal and spatial scales in their investigations.

Viewed comparatively, geography focuses on a small range of scales. If all the kinds of scientific studies done at scales ranging from the subatomic to the intergalactic are included, geography occupies a narrow middle ground, because its subject matter is focused on place-to-place relationships involving the human use of the Earth. Geographers construct global models of climate, but they are also interested in how humans transform their immediate surroundings. The range in orders of magnitude of geographic scales thus becomes clear. The scales of maps geographers make are those that are relevant for understanding people-environment relationships. Many problems involving the human use of the Earth require simultaneous consideration of many scales, hence we speak of a local-global continuum, referring to the many orders of scale magnitude that separate a local focus from a global one (see Chapter 12).

Geographers are also concerned with time scales. Geologic time, beginning with the origin of the Earth 4.7 billion years ago, can be represented as an outward spiral, with more space on the time map devoted to the better-known later periods than to the distant past (fig. 13.1). Geographers’ time scales are strongly skewed toward the present. A geography of the western plains and mountains of the U.S. might mention the age of the Earth \((4.7 \times 10^9)\), give some details of the origin of the Rocky Mountains and the subsequent development of the Great Plains surface \((6.5 \times 10^7)\), and go on to describe the ice ages \((10^6)\).

But the resolution of the geographer’s focus would then sharpen to treat in some detail the end of glaciation, the arrival of humans, the emergence of the grasslands (about \(1.1 \times 10^4\) years before present), and the eastward diffusion of maize-based agriculture from the Southwest \((2 \times 10^3)\). Greatest attention would focus on the period since permanent Euro-American settlement began in the region \((1.4 \times 10^2)\). Recent trends in geography are defined in terms of changes observable in the most recent decennial census \((10^1\) years ago). The geographer’s time scale is more compact than a geologist’s, broader than a historian’s, and much broader than that of an economist who focuses on contemporary trends. Geology and archaeology, as well as history and economics, are thus temporally cognate fields for geography.
Fig. 13.1. Time and Earth History. SOURCE: Press and Siever 1974:79.

SCALE AND MAP READING

Although the range of scales common in geography is an obvious consequence of its subject matter, there remains an additional reason for the practical restriction of geographical map scales to the middle range of values. All information required to construct a map of the entire world—showing features of the Earth’s surface such as relief, land use, population, and rivers—could be stored electronically in a space many orders of magnitude smaller than the map that would be needed to show that information. Furthermore, a cartographer might construct such a map, translate it into electronic form, store it in a computer, or produce a microphotograph of it. Neither case constitutes an exception to the rule that geographers study at the mid-range of scales for the simple reason that geographers’ conceptions of maps are intimately tied to questions
of whether they can be seen and read. It is in the practice of map reading that the true domain of geographical expertise often is found.

In an age of electronic image processing, some elaboration is required to justify why a field like geography, which is now so thoroughly suffused with computer-based technology, still adheres to the apparently quaint notion of map reading. An analogy might help: geographers read maps for the same reasons cardiologists read electrocardiograms. For both, patterns that appear to the untrained eye as specks of ink on paper or bits of light on a video screen are meaningful. There is no substitute for this direct, visual transfer of patterns from display form into a trained viewer's mind. Although anyone with minimum familiarity with road maps uses them as aids to getting around the countryside, a geographer looking at a road map designed to be read by the general public will see things others do not. The geographer sees patterns and their explanations and implications: a uniform, rectangular grid of roads on the map suggests the historic impact of the geometric land survey, a topography that is most likely flat and free of topographic barriers, and a settlement pattern that is pretty much the same from one area to another; where the rectangular road grid stops there is likely to be a decline in the quality of the land.

Geographers see things that are not explicitly depicted on maps because they know that the patterns they infer go together as a bundle of attributes defining a particular kind of landscape. To geographers, then, any map of an area instantly calls to mind other maps of the same or similar areas. Mental correlations are made. They, in turn, become hypotheses that might be tested either by direct observation or via comparisons with other map patterns. Geographers do not merely correlate patterns. Underlying the kinds of visual comparisons that can be made among maps is a knowledge of physical and social processes appropriate to the scale of a map.

An example of how scale guides geographic research is found in John R. Borchert's (1950) classic study of the North American grassland. Borchert sought to explain the wedge-shaped intrusion of grasses known as the Prairie Peninsula into the humid, forested Great Lakes region of Illinois and western Indiana. Because the Prairie Peninsula's outline does not correspond with isohyets (lines of equal precipitation) on maps of annual total precipitation, some researchers sought explanations of a more local nature that emphasized the role played by forest-destroying fires set by early human occupants of the region. But without some process operating at a regional scale, there was no reason to presume that this particular wedge-shaped area would have been especially affected by fire. Borchert devised a regional-scale interpretation by demonstrating that the Prairie Peninsula, although no drier on an annual basis than the areas bordering it to the north and south, had the same shape as did a zone that was especially apt to be influenced by the eastward flow of warm, dry air from the High Plains at the foot of the Rocky Mountains (fig. 13.2). Droughts in the Prairie Peninsula were correlated with droughts in the dry, western
Fig. 13.2. John Borchert's Explanation of the Shape of the Prairie Peninsula (left) Was Based on a Model of Drought Frequency (right) on a Regional Scale. Source: Borchert 1950:2, 26.

plains. And, since drought conditions favor fire, Borchert's explanation was able to accommodate the evidence concerning fire as a proximate cause of the west's prairie grassland.

Regional-scale atmospheric processes that produce such patterns over North America are components of a worldwide system. But neither a global-scale model nor a microscale model of climate over the Prairie Peninsula itself would have isolated and revealed the relationship Borchert discovered. Climatological phenomena of interest to geographers occur at many spatial and temporal scales, ranging from brief local storms through millennial trends in global temperature and precipitation. Generalizations formulated about trends at one scale are not necessarily appropriate at others, a well-known fact that must be considered whether a subject is being charted through time or mapped over areas (see chapter 12).

**SCALE AND SPATIAL DEPENDENCE**

Although relationships can be identified visually from comparing medium-scale maps such as those Borchert constructed, others are more complex and require simultaneous consideration of processes operating at several scales. Robert Haining's (1978) research on High Plains agriculture illustrates how scale can be made an explicit component of a study. Haining constructed a model that incorporated regional trends and local variations in observed patterns of wheat and corn yields in western Kansas and Nebraska, an area that
exhibits strong year-to-year fluctuations in precipitation and crop yields. At a regional scale, Haining found that yields declined from east to west and from north to south, corresponding to drier and warmer environments, respectively. But these trends, already well known, did not account for all of the systematic variation in the data.

Haining incorporated two other scales at which dependencies between adjacent counties and variations within counties could be detected. First, counties with contiguous borders might be expected to exhibit similarities because of "identical, exogenous, uncorrelated variables" (1978:495), namely weather episodes such as droughts that could well occur on a twenty-five- to thirty-mile scale of variability, which is roughly the distance between centers of counties in the area. Second, some droughts are more extensive and relatively isolated, affecting only a few contiguous counties. A third component, intracounty variation, would reflect the effects of local severe storms that produce wind damage or hail. These geographical scales have direct temporal analogues, such as variations that are diurnal, seasonal, annual, decadal, and so on, to long-term cycles of changes.

Haining's model isolated the relative impacts of regional-, intercounty-, and local-scale effects on the total variability of crop yields by explicitly incorporating spatial autocorrelation—a statistical phenomenon well known to geographers. Spatial autocorrelation designates the frequently observed condition of dependence between adjacent locations; it is a pervasive feature of patterns over the Earth's surface. In practice, the existence of spatial autocorrelation means that if A and B are close together, what happens at A is related to what happens at B, and vice-versa. While recognition that adjacent places are likely to be similar is a fundamental truism of all geographic study, the logical complement of this proposition—that widely separated places are likely to be dissimilar—is neither obvious nor is it always true. The question, instead, is How far apart do places have to be in order to be considered independent? and the answer nearly always involves considerations of scale in the process being studied.

Anthropologists have been intrigued with this question for a century. Known as Galton's problem, from critical remarks made by Sir Francis Galton on a study that purported to show differences among several hundred ethnic units distributed around the world, the question is formulated in terms of the number of independent groups that can be identified in a continuously varying, worldwide array. If two groups have a specific behavior in common, does this constitute one instance of group behavior or two? If the groups were widely separated, the answer no doubt would be two. But if they are close together, it is likely that the two groups share a common influence. Such questions lead to discussions of how many truly different cultures there are on Earth. Behavioral traits can be classified as common or rare by asking How many cultures engage in behavior x? A geographer, however, would be unlikely to formulate the
problem in exactly this way because the central question revolves around the definition of group, which has to have some notion of territorial limit in its specification. The problem becomes one of scale and regions (areas of relative homogeneity with respect to one or more factors). The title of a 1981 book by journalist Joel Garreau, *The Nine Nations of North America*, reflects the common perception that regions, like the anthropologists’ groups, are precisely definable and hence countable. But there is no definitive way to delimit a region any more than there is a definitive way to delimit a human group.

Regions—of landforms, climates, settlement patterns, voting behavior, or any other subject—are discernible to the trained eye on geographical maps at scales ranging from $1:10^3$ to $1:10^8$. Furthermore, the number of possible regions found on a map is not a function of the map’s scale. Given two maps at different scales, there is no reason to suppose that the smaller-scale map, which covers more of the Earth, will have more discernible regions just because it spans more territory. New details appear the more closely one looks, and hence diversity is not necessarily decreased when map scale is increased (or decreased). One can map cultures of the world, cultures of the United States, cultures of Missouri, cultures of the Ozarks, and so on, simply by continued refinement of the definition of what is being mapped.

**SCALE AND MAP COMPLEXITY**

In principle, then, complexity need not be a function of scale. If all place-to-place variations were clinal, that is, if they exhibited only continuous gradations in a given direction, then scale would be an irrelevant property of spatial distribution; a closer look at the data (in other words, more closely spaced observations plotted on a larger scale map) would produce no new insights. In such a case, measurements at equally spaced points along the cline would yield a function describing the rate of change, and a trend could be interpolated to make a larger-scale map or extrapolated to create one at a smaller scale. A variable that can be measured at one scale as well as at another is exemplified by the amount of radiation that the Earth receives from the sun (disregarding the effects of the Earth’s atmosphere). The potential incoming solar radiation at any location on any day is a function only of the latitude of the location and of the day of the year. Gradations in both time and space are continuous around every spatial or temporal point. But geographers find relatively few variables that can be described so simply, and at the opposite extreme lie features of the Earth’s surface that appear different no matter what scale is employed. The classic example of such a pattern is the case of water-land boundaries. An irregular coastline, such as that of the state of Maine, gets more irregular the more closely it is viewed. At a sufficiently large scale, the coast line becomes
a broad zone that has a geography of its own, one that varies periodically with tide cycles. Yet, for many purposes, the coast of Maine can be drawn as a line on a map, and there is no problem of understanding what the line is meant to portray.

There are two issues involved. One problem is that the coastliffe of Maine actually gets longer the more accurately it is measured. There is no single number that can describe the length of a physical feature unless that number is qualified by a statement about the scale and method of measurement. All physical objects are measured by imposing some human standard of measurement. Highly irregular linear features, such as coastlines or meandering rivers, actually come closer to being areal features because of the amount of area that a line representing the feature occupies on a map. Linear features that have space-filling characteristics more like areas and two-dimensional surfaces so irregular that they are more like solids are known as fractals (see chapters 4 and 10). The other scale-related problem lies in the domain of conventional cartography. If a map is constructed from an aerial photograph, the cartographer generally will trace the outline of the feature being mapped with as much accuracy as feasible, but the resulting line will not be exactly the same as the feature. The coast is not a line, and even if it is treated as one, its width varies in both space and time. Some variability is short-term and periodic (waves, tides), but even the average shape of a coast changes, at least on the temporal scales relevant to geologists. So the generalization of a line, like other generalizations, is unlikely to be true to the extent that within it one can find all the properties of the original and no others.

Granting the problem of earth-to-map generalization, what about the apparently simpler issue of map-to-map generalization? Since reducing map scale often requires simplification (information removal) or generalization, (selective information removal), a desirable property of such a procedure would be the ability to ungeneralize the generalized map. If simplification merely muted details, they could be recovered by applying the inverse transformation (enhancement) to the processed map. Waldo Tobler (1969) has considered a class of transformations that can be characterized as having inverses. The height of some point on a mapped surface can be redefined as the weighted sum of its own elevation plus a discounted-for-distance function of its neighbors’ elevations. Applying such a transformation to elevations, whether the heights represent terrain, population density, or any other numerically defined surface, has the effect of smoothing it by decreasing peaks and filling in valleys. The procedure is a two-dimensional analogue of time-series smoothing (such as numbers representing daily prices in a stock market) whereby short-term random shocks are filtered out by redefining each day’s observation as a weighted function of itself and the adjacent days. The problem posed by Tobler is whether a profile that has been smoothed can be restored to its original form. It can, provided starting points (boundaries) are specified and provided certain
conditions are met in the smoothing function. But the process must be deterministic, that is, it must be totally predictable and not subject to random fluctuations; otherwise, no unique reconstruction of the original is possible. These conditions might be met by the standards of a machine; they are less likely to be approximated when humans do the smoothing.

Although Tobler used a mathematical process to reversibly generalize, a physical example illustrates the difference between deterministic and human treatment of scale change. Suppose the coastline of Maine that appears on a 1:250,000 scale map is photographically reduced to a scale of 1:1,000,000. The photographic reduction is deterministic, and it can be undone by enlarging the new map four times \((250,000/1,000,000)^{-1}\), back to where it was. But since the reduced-scale coastline has all the detail it had before—now shown by a line only one-fourth as wide—a human cartographer would most likely smooth the line by eliminating some irregularities and, for the sake of visibility, draw the coast as a line almost as wide as it was before. A wiggly line may get straightened a bit when it is redrawn, but it may also acquire new wiggles that weren’t there originally. The result is that the coastline has been generalized as well as simplified, and it can no longer be restored to what it was, given the inevitable random fluctuations introduced by redrawing it.

THE SCALE REDUCTION PROBLEM

Although geographers are apt to describe map symbols as representing point, line, or area features, these designations are merely conceptual. A zero-dimensional point (dot) representing a small town would be invisible unless it had a diameter; a one-dimensional line, signifying a road, would be invisible unless it had a width. Because symbols or letters on a map have to possess area in order to be seen, and because the conventional description of map scale is linear (a representative fraction), the scale reduction problem refers to the fact that area reductions are the second power of linear scale reductions, and the area available to show the same amount of information therefore decreases rapidly with decreasing map scale.

The purpose of either simplification or generalization is to reduce the information shown on the map because of the necessity to reduce scale; too much information on a small-scale map results in excessive complexity, whereas too little information on a large-scale map produces a map that is too empty. The dotted lines on fig. 13.3 show how information reduction accompanies scale reduction provided that information is kept constant relative to the amount of area on a map. If the scale of the map is to be reduced from \(S_1\) to \(S_2\), then simplification—reducing the amount of information the map depicts—is one way of coping with the problem; but it results in a large decrease in map infor-
Fig. 13.3. Map Simplification and Map Generalization.

Information, from \( I_1 \) to \( I_2 \). Alternatively, if the definition of information is changed, such as by aggregating small, irregular units into larger, more regularly shaped units, information need not be decreased as much (only from \( I_1 \) to \( I_2' \)). Map generalization thus evades the scale reduction/information reduction problem by redefining map categories and symbols to retain only essential information.

**SIMPLIFICATION VERSUS GENERALIZATION**

Simplification refers to removing information from a map so that when scale is reduced the map is not too cluttered. Simplification may involve smoothing irregular features and removing extraneous information not necessary to the map's message. When the information to be mapped is stored in a computer, the computer may select only a part of the data for subsequent use. Whether the data are thinned by a person or by a computer, however, removal is undertaken for the purpose of reducing the amount of information to be mapped.

Simplification and generalization are undertaken with the intended message of the map in mind; the message, of course, is determined by the nature of the problem. The map of the Prairie Peninsula described earlier offers an example. Vegetation patterns are highly discontinuous, but that characteristic was not
what Borchert wished to illustrate. His purpose was to show the general outline of grassland vegetation, especially its wedge-shaped eastward penetration.

Some insight into the nature of generalization is gained by comparing Borchert's map, printed at a scale of approximately 1:32,000,000, with one redrawn from the highly detailed wall-sized natural vegetation map of the United States published by A. William Küchler (1964) at a scale of 1:3,168,000, reduced here to match the Borchert map (fig. 13.4). The linear map scales differ by one order of magnitude and their area scales therefore by two orders. Comparison of the two maps at the same scale shows how much detail Borchert's map does not show. Although Borchert did not have the more detailed map (published fourteen years later) to work from, one may nonetheless see in Küchler's map the relative accuracy of Borchert's generalization of the eastern edge of the prairie. The fact that grassy vegetation grows on uplands rather than on river bottoms is obvious on the Küchler map but totally absent on Borchert's. The role of fire in creating grassy expanses on the flat uplands while leaving the valleys unscathed is an additional message of the Küchler map.

Why would two geographers map the same thing at areal scales such that one is more than one hundred times larger than the other? And how can we avoid concluding that Küchler's larger-scale map is preferable? The answer to both questions comes from understanding the purposes the maps served when they were constructed. The scale of Borchert's vegetation map was appropriate for comparison with the mesoscale climatic patterns he was studying. He was not invoking climate as an explanation of vegetation differences between up-
lands and valley bottoms, but rather as an explanation of the shape of the Prairie Peninsula. His highly generalized map suited his purpose. Küchler, on the other hand, had as his primary aim the depiction of vegetation patterns for their own sake; his 1964 map depicted 116 different vegetation types and used colors as well as pattern textures to distinguish among types. One would consult the Küchler map for the answers to many questions, but the Borchert map for the answers to only a few.

The difference between simplification and generalization thus becomes clear. The small-scale Borchert map is not merely a simplification of the details of a larger-scale map such as Küchler’s. It is, instead, a highly selective generalization in which some details were suppressed for the sake of conveying a single message. To know how to generalize a map, then, is to understand the relationships one is trying to map. Random removal of information will not accomplish the task, and even orderly removal, such as deleting every nth piece of information from the map data, does not constitute an informed generalization; information has to be reduced, when going from large scale to small, at a rate proportional to the second power of the scale change. But all of the essential information can be retained—and even enhanced—by an appropriate choice of symbols, lines, and patterns.

**SCALE AND GEOGRAPHICAL RESEARCH**

It was not many years ago that only small-scale maps could be drawn with confidence for some areas of the world, especially areas that were sparsely populated, remote, and rugged. Ground-based mapping had achieved coverage of many countries, but extensive blank areas remained. A land-use map of the world that might have appeared in a standard atlas was accurate to the scale of 1:10⁴ for, say, Great Britain, but only 1:10⁷ for Siberia. Such a map would have been printed at a single scale, but its accuracy would have varied from place to place because details were lacking in some areas. Today, given the availability of satellite images, there need be few such place-to-place differences. Most surface areas are equally accessible to this technology, although for political reasons not all of the detail reaches the general public.

Accompanying the advent of space photography has been a shift in emphasis from speaking in terms of scale to speaking of image resolution. Resolution refers not to the image but rather to the device that records the image (such as a scanner) or to the medium on which it is recorded (such as film in a camera); it describes how finely features may be differentiated, given the construction of the device. Resolution requires attention to the largest-scale elements that such a device can record. The term *pixel* (from picture element) has been coined to define any two-dimensional area that is the smallest nondivisible
element of an image. In other words, a high-resolution device will yield information in small pixels and is appropriate for constructing accurate, large-scale maps. The quest for greater accuracy on maps has naturally followed developments in high-resolution image-recording devices, including cameras, scanners, and digitizers.

At this point the art and the science of geography begin to diverge a bit. Map generalization is at least as much an art as it is a science. Many factors of judgment are involved in deciding how to reduce scale while preserving essentials. The art tends to be forgotten, however, when the only concern is greater accuracy, that is, higher resolution. The search for greater accuracy nearly always leads to an exponential increase in the amount of information. A map at a scale of 1:10⁴ requires one hundred times as much information as a map at 1:10⁶ in order to keep the accuracy per unit area constant. This comparison sounds staggering, until it is realized that the resulting map will also be 100 times as large. But there is little chance that atlases will become tomes 100 times larger than they now are just for the sake of showing all the new information that has been collected. Although the search for greater accuracy demands higher resolution and produces more information, that information must necessarily be generalized down to nearly the same scale as before simply to keep the product a manageable size.

From this observation it is natural to conclude that collecting more information, ever more accurately, produces diminishing returns for the effort expended. This is true in geography and in every other science that collects data: increasing the size of the sample adds new insights, but at a diminishing rate. Although increasing the size of a sample produces fewer new conclusions per observation as the sample gets larger, confidence in those conclusions continues to grow. And so it is with gathering high-resolution data to make more accurate maps: geographers and cartographers continue to generalize the information down to an appropriate scale, but their confidence about what the map reveals becomes stronger the larger the amount of information upon which it is based.

The information explosion that has accompanied the availability of high-resolution satellite-based images of the Earth has made small-scale maps much more accurate and has made large-scale maps of many areas possible. But it has not led to a general increase in the scales at which geographers study things. Physical processes do not change their natures just because scientists are able to monitor them more accurately or more frequently. What the new technology has provided instead is a clearer idea of the interrelatedness of the whole global-atmospheric system. The information gained from large-scale resolution, combined with the ability for frequent updates, has fostered global (small-scale) models that could only have been dreamed of a generation ago. Computer technology has advanced in response to the information-processing requirements that these models pose.
Rapid machine processing of millions of data points also has made it possible to use existing information to create new map products. An example is the shaded relief map of topography in the United States recently published by the U.S. Geological Survey (see fig. 6.9). What does the land surface of the United States look like when viewed from space? An actual photograph from space would seem to be the answer, but the terrain itself would be masked by clouds in some places and obscured by vegetation in others, and the sun’s angle of illumination would vary continuously across the image. Instead, semiautomated processing of elevation data from 1:250,000-scale topographic maps allowed the map’s authors to encode more than two billion elevations at a resolution of approximately one observation every 200 feet on the ground. The map information was then both simplified and generalized. First, elevation data were thinned to a more manageable file of twelve million observations appropriate to the scale of the final product. Then the data were generalized in several intriguing ways. One was to simulate the effect of low sun-angle illumination by letting the computer produce shaded relief to portray more effectively the ridges and valleys. The sun never shines at the same angle over the entire United States (when it is late afternoon in California it is dark in New England), nor does it ever strike at a horizontal angle that is particularly useful for representing surface form. But the consistency and choice of angle selected by the map’s authors provide a meaningful visual image uninhibited by the complexities of reality.

Their map is a good illustration of how simplification and generalization can be combined to create a map that can be read and appreciated by specialists as well as the general public. The map reader gains an instant impression of the topography of the United States. The map is not a picture, even though it was deliberately created to look like one. It is a thematic map that, like all others, is a generalization—a selective portrayal of relevant information, with some details suppressed and others enhanced, that was created for a specific purpose.

**SCALE AND VOLUME: THE DATA DIMENSION**

So far, discussion has concentrated on scale as it relates to two-dimensional space. Had the depiction of landforms been a three-dimensional model, say in plaster or plastic, the model builder would have had to choose a vertical scale as well as a horizontal one and would undoubtedly have chosen a vertical scale larger than the horizontal. The highest point in the contiguous United States is less than three miles above sea level, whereas the east-west extent of the nation is more than three thousand miles. Thus, a model of the United States one foot wide would have a vertical variation about the thickness of a piece of paper. In
such situations, geographers typically exaggerate the vertical scale in order to enhance surface contours.

Even when mapping data on two-dimensional maps rather than on three-dimensional models, vertical scaling is a vital consideration. The problem of categorizing data for mapping purposes is essentially one of scaling those data. Mapmakers decide whether to divide data into few classes or many, and whether to divide the data range evenly or to use some other means of clumping values into groups. Scaling the data dimension requires special attention to the range of values being depicted. Any categorization that does not divide the range evenly emphasizes some values at the expense of others. The Pike-Thelin map provides an example of scaling the data dimension. The authors converted their elevation data to logarithms, the effect being to create greater contrast in low-relief topography so that valleys in the Great Plains, for example, are more visible.

## SCALE IN GEOGRAPHY

Scale is an important consideration in nearly all geographical studies. Careful selection of appropriate scales in space and in time is required to formulate and answer meaningful geographical questions. Because trends discernible at one scale are often invisible at another, geographers commonly use a range of scales to ensure that the conclusions drawn from a study are well matched to the economic, social, and physical processes known to underlie observed patterns. A map can effectively portray a geographer’s thesis when an appropriate scale is chosen. A geographer’s skills in constructing a map that shows only the relevant information resemble the crafting of a well-argued point in a debate, and when the map argument succeeds, a geographical thesis is substantiated.

## REFERENCES


