Module 4: Data Visualization

Module 4 Learning Objectives

- Learning the common pitfalls of visualization and examples of poorly designed maps
- Learning the major components/processes of cartographic communication
- Learning the basic principles of map design and cartography communication
- Learning the major modern geo-visualization techniques used for effective cartography communication

1. Common pitfalls of geospatial data visualization

1.1 Definition of geo-visualization

Geo-visualization, short term for “Geographic Visualization” can be defined as a set of tools and techniques to support geospatial data analysis through the use of interactive visualization. Like the related fields of scientific visualization and information visualization, geo-visualization emphasizes information transmission. Geo-visualization communicates geospatial information in ways that, combined with human understanding, allow data exploration and decision-making processes.

To summarize, geo-visualization is an interesting and useful field of research for different reasons:

1) can reduce the time to search information, and support decision-making;

2) can enhance the recognition of patterns, relations, trends and critical points etc.;

3) can give a global vision of a situation, a phenomenon, etc.;

4) enables the use of human visual memory and the capability of perceptual processing of data;

5) permits a better interaction between user and the information system;

6) and can possibly lead to the discovery of new bunches of knowledge.

A core argument for geo-visualization is that visual thinking using maps is integral to the scientific process and hypothesis generation, and the role of maps grew beyond communicating the end results of an analysis or documentation process. As such, geo-visualization interacts with a number of disciplines including cartography, visual analytics, information visualization, scientific visualization, statistics, computer science, art-and-design, and cognitive science; borrowing from and contributing to each.

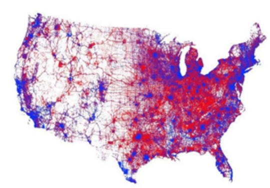
1.2 Common pitfalls of geo-visualization

Indeed, graphics are to date the most efficient means of communicating information as well as effecting data analysis and display in GIS environments. Maps are the primary tools of geo-visualization to present, use, interpret, and understand geospatial data. The development of computer software and Geographical Information Systems (GIS) has made it possible for anybody who can operate a computer to get a map output from the GPS coordinates.

So any person, (not necessarily a surveyor or a cartographer) who wants a map, takes out a GPS, goes into the field, gets coordinates, downloads them into a GIS system and gets a map. Maps are a great way of displaying and analyzing statistical information but they need to be properly designed. This can be a tricky business until you know what you are doing and desktop mapping and GIS systems rarely provide much help. Most software packages will allow us to produce really bad and misleading maps. Poorly designed maps restrict the communication of information, and may convey false ideas about the facts contained in the data as displayed. This is particularly dangerous if cartographic output is presented to persons (e.g. executives) who are only remotely involved in a particular project but are involved nonetheless in making final decisions. The non-cartography trained persons may easily produce maps that may look good, but are not consistent with any established cartography standards or conventions. Such maps may not transmit the intended meanings to map users at all. Sometimes they can only be misleading. Figure 1.1 shows examples of misleading maps.

<table>
<tr>
<td>

(a)
</td>
<td>



(b) (c)
</td>
</tr>
<tr>
<td>Figure 1.1 (a)Even the familiar image of the globe inaccurately distorts the size of countries on the equator, making them seem smaller. Source: Pixabay.</td>
<td>(b)The choropleth map tweeted by Donald Trump (l) did not show population density.
(c) Using the same data, Kenneth Field presented an example of how different the map would look if it showed where people live and where they voted on a more granular and accurate level.</td>
</tr>
</table>

Based on (Ildefonso, 2021), some common cartographic mistakes are listed below:

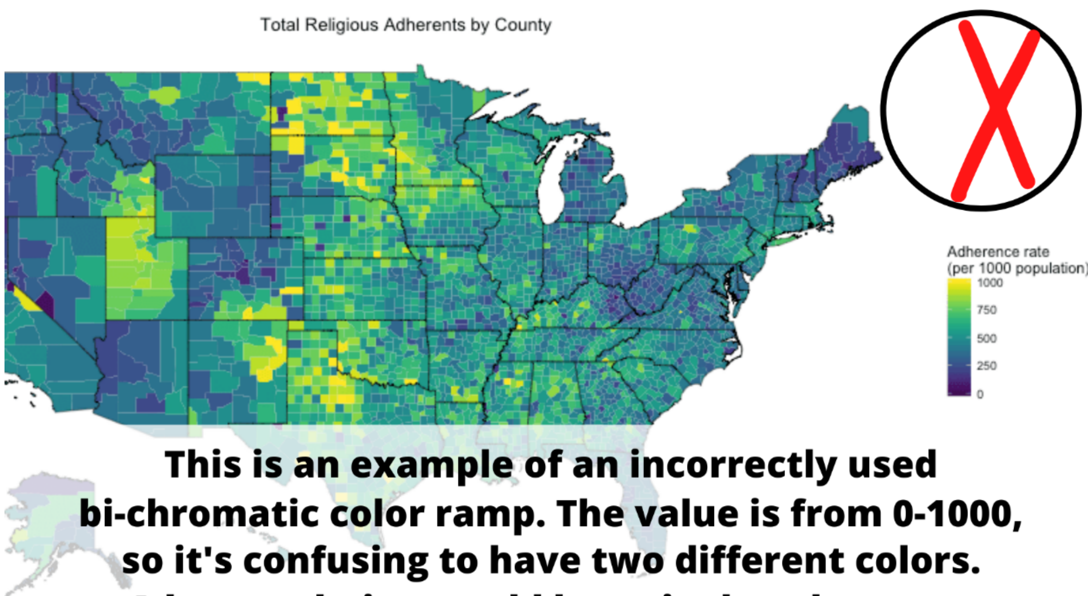
1. Showing too much data. This is a mistake made by even veteran GIS practitioners. Mapmaking is a delicate balance between trying to represent a complex reality and producing a readable map. To help address this problem, find ways to contextualize your map, such as integrating your map into a Story Map or embedding it on a website where you can provide more information.

2. Adding unnecessary mapping elements. GIS Beginner are often so excited when it’s finally time to jazz up their map that they go a little overboard and include fancy scale bars and north arrows that are either unnecessary or that clash with the style of the map.

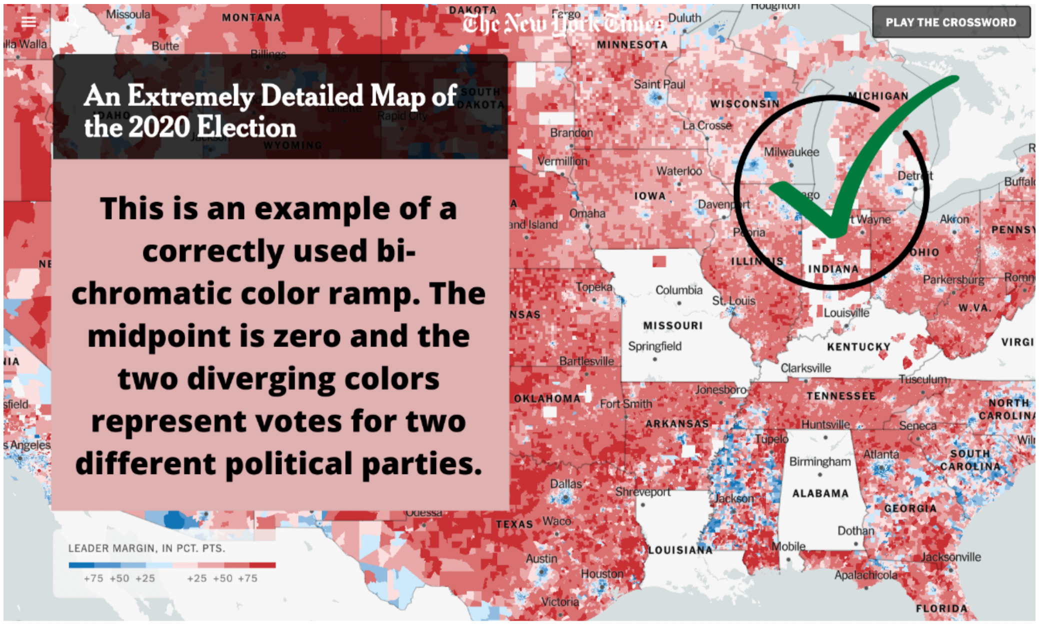
3. Using a multicolored categorization for quantitative data. Multicolored categorizations are best for qualitative data, such as a map that shows different types of vegetation. Additionally, the color choices shouldn’t be random but reflective of the category—woodland (brown), forest (green), tundra (white or light brown), etc.

4. Using choropleth maps to display absolute numbers. Choropleth maps represent values by shading patterns. They should be used to display normalized data, which is data that has been compared to the whole population (e.g. rates, percentages, proportions, per capita values, medians, or averages).

5. Choosing the wrong color ramp. A bi-chromatic color ramp (the one with two distinct colors) should be used with data that goes above and below a midpoint (e.g. above and below sea level). A single-color ramp should be used to display a continuum without a midpoint (e.g. percent of population fully vaccinated).

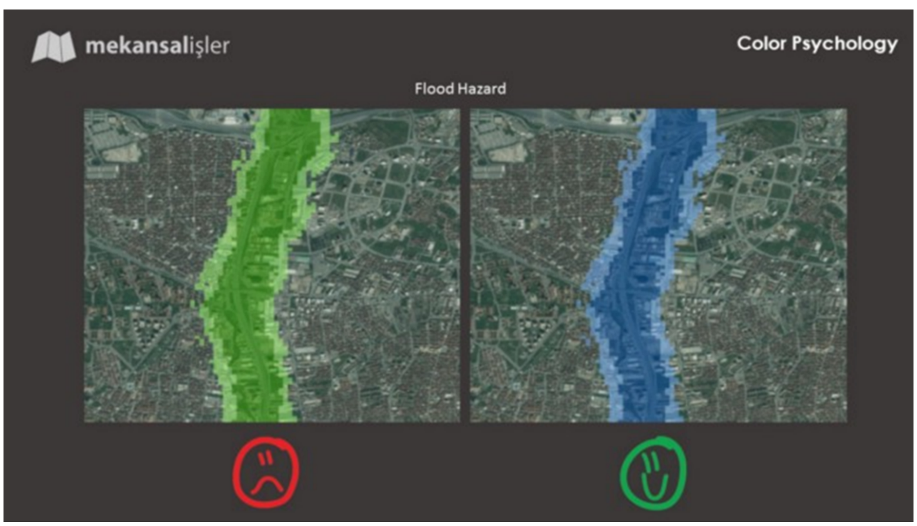


This is an example of an incorrectly used bi-chromatic color ramp. The value is from 0-1000, so it's confusing to have two different colors. A better choice would be a single-color ramp.

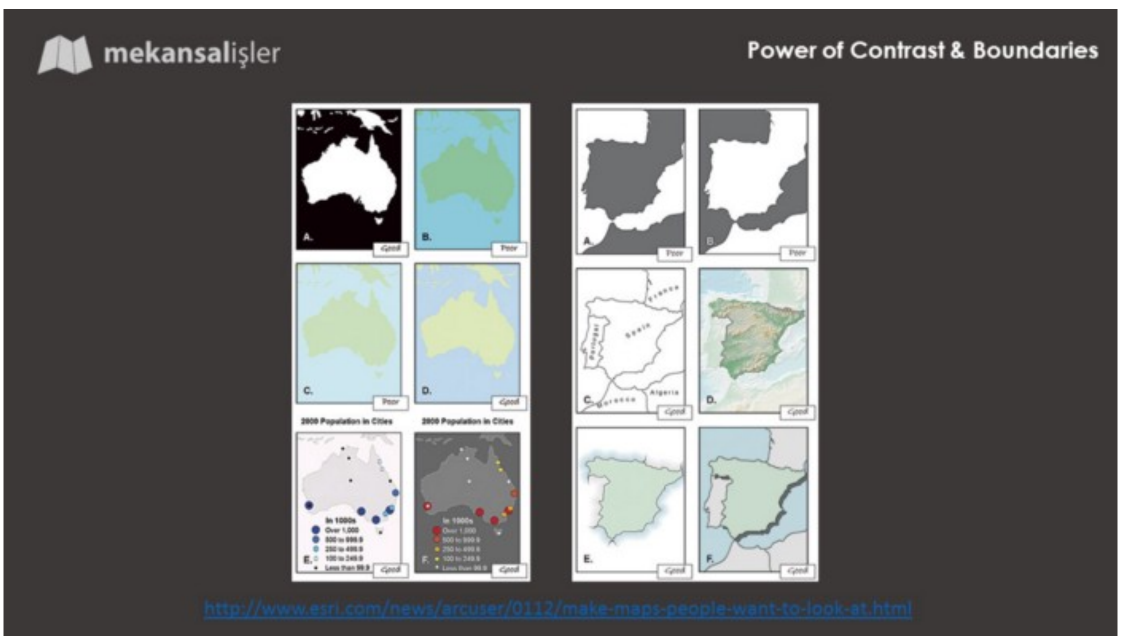


Based on Celik (2015), some pitfalls of spatial data visualization are listed below:

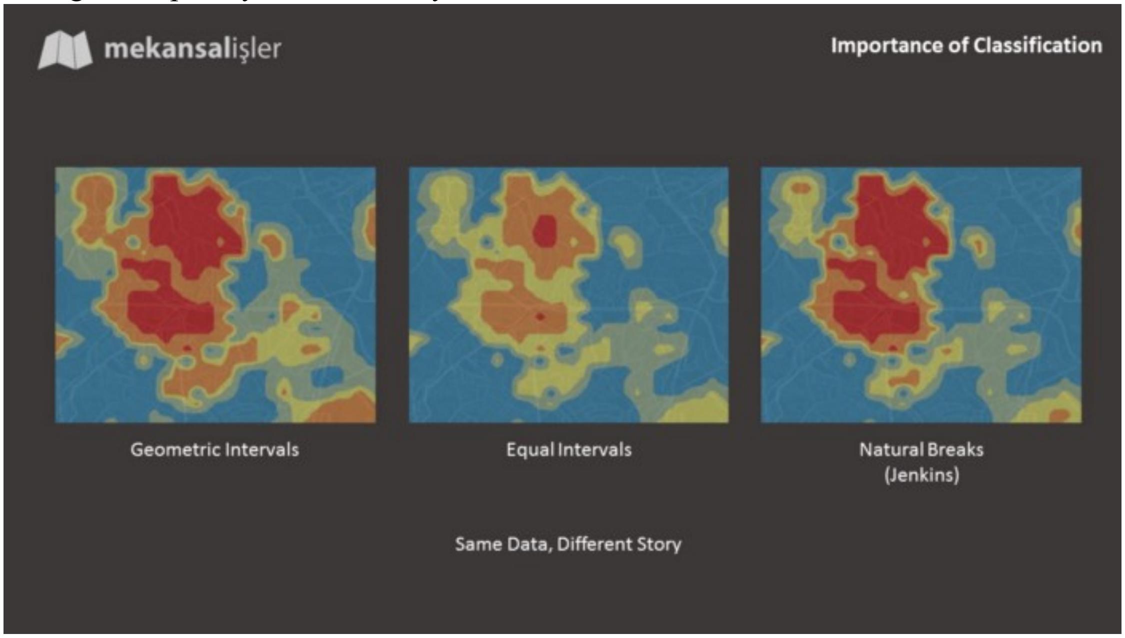
- Color Psychology: Every color has multiple meanings & feelings. Therefore, you should be careful when you use a color for your data visualization. People expect to see rivers in blue, parks in green and danger in red.



- Power of Contrast & Boundaries: What you see is usually defined by what you see around it. Your choice of colors should be as contrast as possible and you need to seperate your data points with right visualizations. Some good & poor examples are provided by ESRI. As you see slide above, some maps are harder and some are easier to understand. What makes a visualization is easier to understand is usually the good use of contrast colors & boundaries.



- Importance of Classification: Below figure shows 3 different maps that are visualized from same data using same number of classes. However, all three of them tells a different story. First map is classed with Geometric Interval algorithm, that optimizes both the number of data points in a class and variance within the class. Second map is classed with Equal Intervals algorithm, that basically divides the data with equal intervals.Third map is classed with Natural Breaks algorithm and it keeps the variance within the class minimum while increasing the variance between classes. The mission of analyst is to use the correct classification algorithm to pass the message of data. One mistake, and you are telling a completely different story.



- Volume != Area: This is a common mistake that we do everyday. When you look at this map, you feel like company ABC has more market share than the company XYZ. This is due to the landsize of the states. Administrative boundaries are usually meaningless for these kind of datasets.

Monmonnier (2018) produced a very exciting book entitled “How to lie with map?” In this book, he showed many examples of how reality can be distorted or disguised. Figure 1.2 give an example of a funny busline. In Figure 1.2(a), one can see the “real” map whereas in Figure 1.2(b) lays a commercial announcement. In this announcement, it is easy to see that the busline look quicker: it is a fraudulent map. However in Figure 1.3(b), the caricature of a road network. (Figure 1.3(a)) is more easily understandable than a map nearer to the reality showing too many details hiding the more important issues.

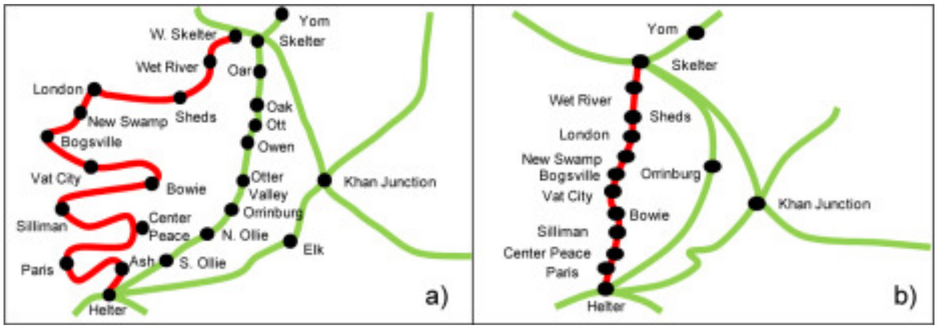


Figure 1.2 A funny busline from Monmonier: a) a “real” map; b) a fraudulent map. Redesigned from Monmonnier (2018)

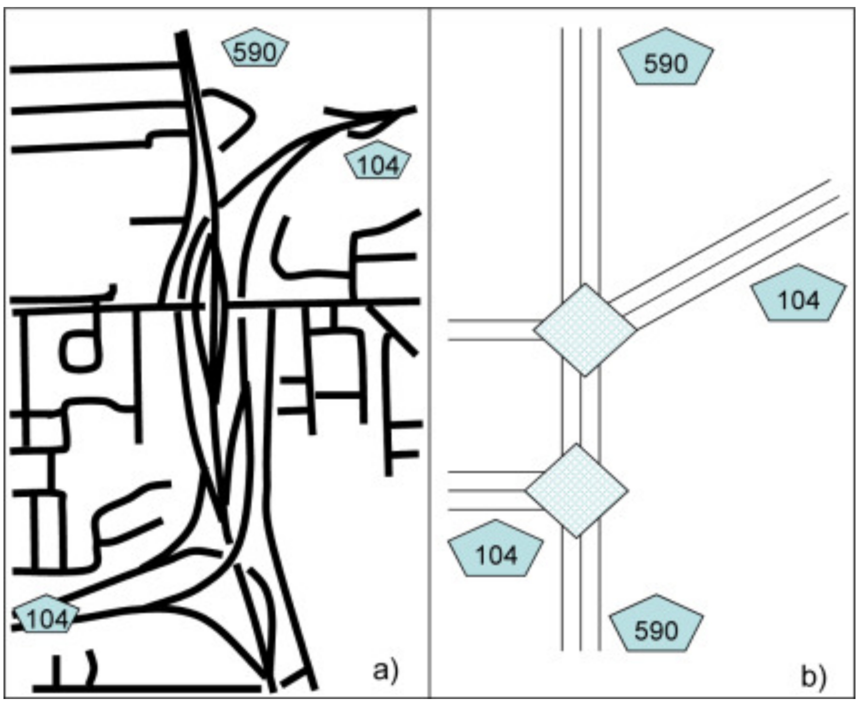


Figure 1.3 A road network and its caricature: a) a “real” map, not so easy to understand; b) a caricature more easily understandable. Redesigned from Monmonnier (2018)

For more on this topic, check out these helpful resources on standard cartographic practices:

1. Book: How to Lie with Maps by Mark Monmonier
2. Book: Ways of Representing the World by Daniel Dorling and David Fairbairn
3. Article: Which Color Scale to Use When Visualizing Data

2. Basic principles of map design and cartography

2.1 Major Components/Processes of Cartographic Communication

2.1.1 Map projections

A map projection is a way to flatten a globe's surface into a plane in order to make a map. This requires a systematic transformation of the latitudes and longitudes of locations from the surface of the globe into locations on a plane. The creation of a map projection involves three steps in which information is lost in each step: 1) selection of a model for the shape of the earth or round body (choosing between a sphere or ellipsoid); 2) transform geographic coordinates (longitude and latitude) to plane coordinates (eastings and northings); 3) reduce the scale (in manual cartography this step came second, in digital cartography it comes last).

If a surface can be transformed onto another surface without stretching, tearing, or shrinking, then the surface is said to be an applicable surface. The sphere or ellipsoid are not applicable with a plane surface so any projection that attempts to project them on a flat sheet will have to distort the image (similar to the impossibility of making a flat sheet from an orange peel). A surface that can be unfolded or unrolled into a flat plane or sheet without stretching, tearing or shrinking is called a 'developable surface'. The cylinder, cone and of course the plane are all developable surfaces since they can be unfolded into a flat sheet without distorting the projected image (although the original projection of the earth's surface on the cylinder or cone would be distorted).

There really is no one best map projection. Each map projection distorts shape, distance, direction, scale, or area and cannot preserve all map properties at the same time. You can't represent Earth's surface in two dimensions without distortion.

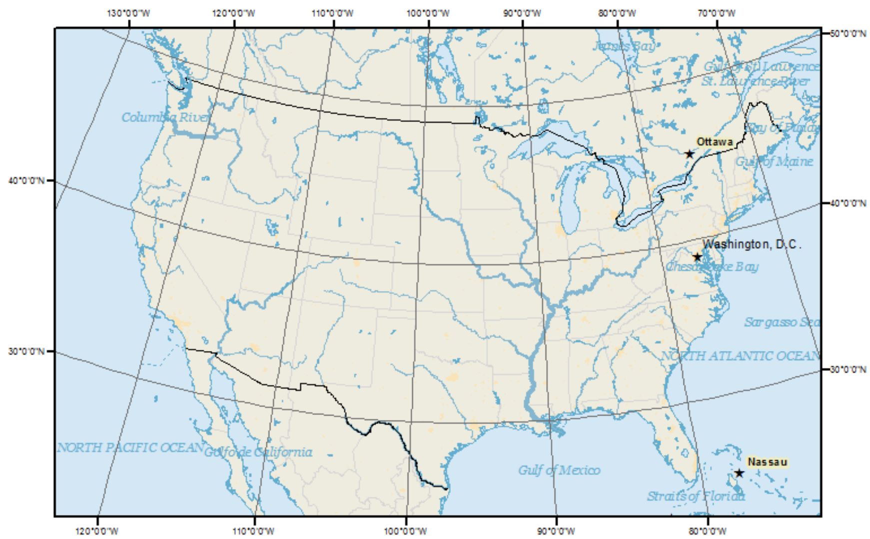
There are several different types of projections that aim to accomplish different goals while sacrificing data in other areas through distortion. They are listed as below:

- Area preserving projection – equal area or equivalent projection
- Shape preserving – conformal, orthomorphic
- Direction preserving – conformal, orthomorphic, azimuthal (only from a the central point)
- Distance preserving – equidistant (shows the true distance between one or two points and every other point)

Azimuthal projections touch the earth to a plane at one tangent point; angles from that tangent point are preserved, and distances from that point are computed by a function independent of the angle. Azimuthal equidistant projection is used by amateur radio operators to know the direction to point their antennas toward a point and see the distance to it. Distance from the tangent point on the map is equal to surface distance on the earth. Conformal map projections preserve local angles and shapes. Often, meridians and parallels intersect at right angles. A map projection cannot preserve angles and shapes at the same time. Mercator projection wraps a cylinder around the earth; the distance from the equator on the map is being geographical latitude, on a scale where the earth's radius is 1. Equal-area projections preserve area.

There are thousands of map projections that are in existence today! Two of the most common map projections used in North America are the Lambert conformal conic and the Transverse Mercator.

Lambert Conformal Conic



North America: Lambert Conformal Conic

The Lambert Conformal Conic is derived from a cone intersecting the ellipsoid along two standard parallels. When you "unroll" the cone on a flat surface, this becomes the mathematically developed surface.

The most distortion occurs in the north-south directions. In general, distortion increases away from the standard parallels. For example, this map projection severely expands South America.

Universal Transverse Mercator

North America: Mercator

Universal Transverse Mercator (UTM) coordinate system is a standard set of map projections with a central meridian for each six-degree wide UTM zone. Even though Google maps used the Mercator projection because it preserves shape decently, and north is always up.

But Mercator map projections are really bad at preserving area. For most of us, the projection is common enough that it looks fine. In reality, Africa is huge on the globe. But Greenland appears to be as large as Africa, even though in reality it is only 1/14th the size. The Mercator puzzle game illustrates this point.

Spatial referencing systems (latitude and longitude) are used to locate a feature on the Earth's spheroid surface. The location of any point on Earth can be defined using latitudes and longitudes. These points are expressed in angular units such as degrees, minutes, and seconds.

Most maps in a GIS are in two-dimensional form. To make use of these maps, you need reference systems that use a pair of coordinates.

However, when you transfer a spherical shape to a flat surface, you approximate the true shape of the Earth. Depending on the map projection you choose, some projections may cause distance between features on a map to be preserved while distortion is introduced to shape. In some cases, the area may be preserved while the direction is distorted.

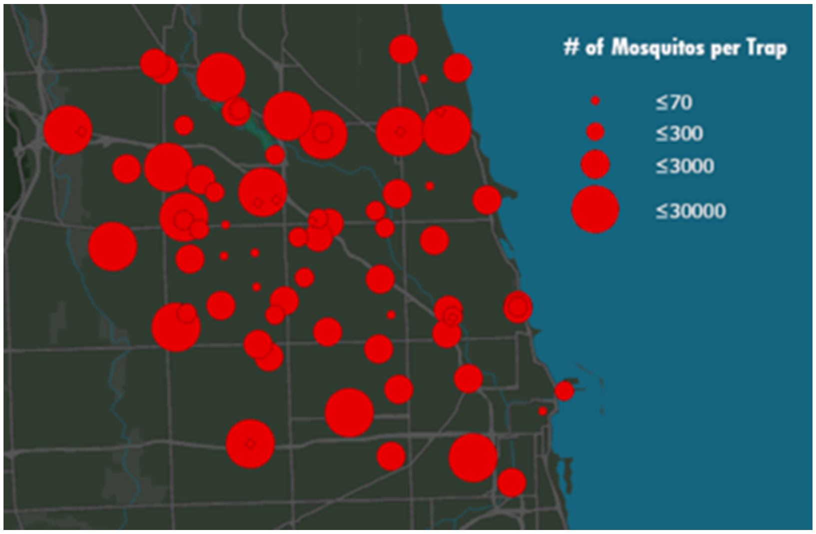
Cartographers choose map projections that best represent the purpose, size, and shape of the area of interest on the map.

2.1.2 Symbolization

Symbolization is an important skill in cartography, or map making. It is the process of choosing an appropriate representation for specific features on a map. We can symbolize point features as dots, squares, triangles, flags, or other shapes and we can symbolize line features using solid, dashed, or other patterns. The symbols we choose should help describe additional information about the features on the map.

In general, we associate large size with greater numerical values and intense color with strong events. For example, when symbolizing earthquakes on a map, using dots that are all of the same color highlights the locations of the individual earthquakes. To emphasize the difference in distribution between high magnitude and low magnitude earthquakes, we can symbolize them using dots of varying sizes, with the largest dots representing the highest magnitude earthquakes. We can use color in a similar manner, symbolizing the higher magnitude earthquakes with intense shades, such as dark red and lower magnitude earthquakes with lighter, pastel shades, such as pink.

Graduated symbols are used to show a quantitative difference between mapped features by varying the size of symbols. Data is classified into ranges that are each then assigned a symbol size to represent the range. For instance, if your classification scheme has four classes, four different symbol sizes are assigned. The color of the symbols stays the same.



This map uses graduated symbols to show which mosquito trap locations in Chicago have the highest counts of mosquitos tested for West Nile Virus.

Symbol size is an effective way to represent differences in magnitude of a phenomenon, because larger symbols are naturally associated with a greater amount of something. Using graduated symbols gives you a good degree of control over the size of each symbol, because they are not related directly to data values as they are with proportional symbols. This means you can design a set of symbols that have sufficient variation in the size that represents each class of data to make them distinguishable from one another.

Chart symbology

A chart is a type of statistical graphic that represents data. Charts can be used as multivariate symbology to show quantitative differences between attributes, with each part of the chart representing an attribute value that contributes to the overall whole set of values. Chart symbology can be used with point, line, or polygon features. For example, you can use pie chart symbols to represent the ethnicity of a district in a city. Each section of the pie chart comprises one of the ethnicities. Each chart symbol can then be sized proportionally by the district's total population.

Census tracts in New Orleans use pie chart symbols to show ethnicity percentage. The symbols are sized by total population.

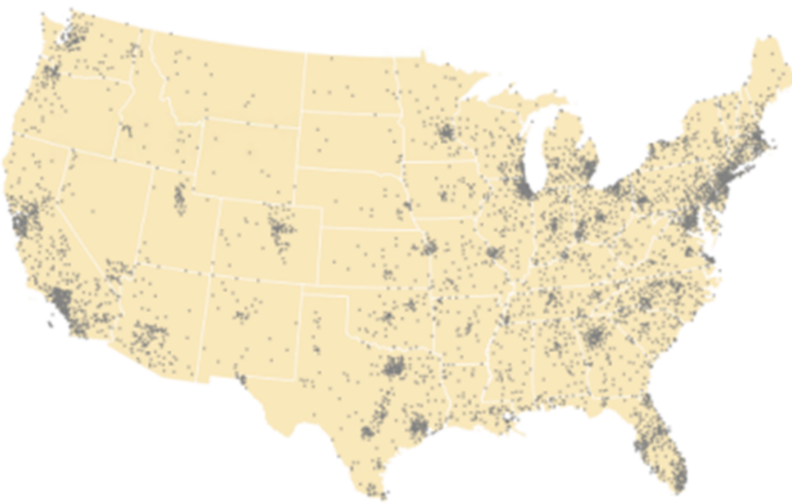
Proportional symbology is used to show relative differences in quantities among features. Proportional symbology is similar to graduated symbols symbology in that both draw symbols sized relative to the magnitude of a feature attribute. But where graduated symbols distribute features into distinct classes, proportional symbols represent quantitative values as a series of unclassed symbols, sized according to each specific value.

Proportional symbols can be defined for point, line, or polygon feature layers. When applied to point or line symbols, the feature's size is modified directly. When applied to polygon features, a proportionally sized point symbol draws at the center of the polygon. For reference, you can specify a uniform background symbol for the polygons that draw below the points.

This map uses proportional symbols to show the precipitation in Mexico.

Dot density symbology is one way to represent quantities within polygons on a map. With dot density symbology, the data you symbolize is not classified. Instead, quantitative values for one or more fields are represented as a collection of point symbols (typically solid circles or dots) within each polygon. Each dot represents a constant number of things, or people, or other quantifiable phenomena. The dots are equally sized, even when multiple fields are symbolized together within a layer.

To mimic a natural data distribution, a good approach is to symbolize a fine-grained polygon dataset, such as United States counties, with dot density symbology using a completely transparent polygon symbol. Then, overlay this layer over a coarser-grained categorization polygon layer, such as the states in the United States, symbolized with single-symbol symbology. This gives a better indication of where people actually live than if the dot density symbology were applied to the coarser states polygons instead. An example is shown below

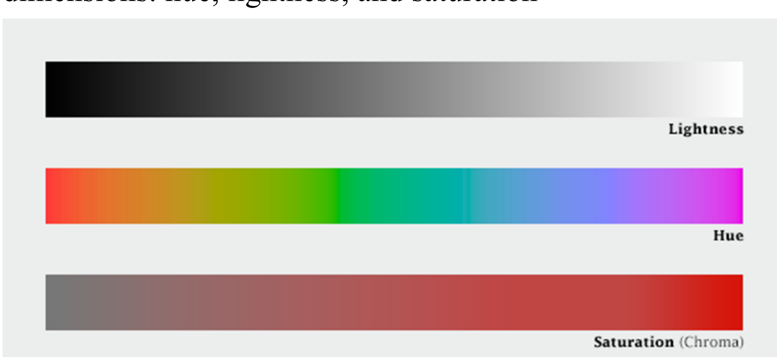


2012 population per U.S. county drawn with dot density symbology in gray, shown above U.S. states, shown in light orange

2.1.3 Colors

The use of color in maps and data visualizations has a long tradition. Color, along with position, size, shape, value, orientation, and texture is one of the primary means to encode data graphically.

When you hear the word "color," words such as blue, red, and green likely spring to mind. Though these are colors in the colloquial sense, these are better described as color hues. When using color as a visual variable, each color is specified not just by color hue but by three dimensions: hue, lightness, and saturation



Credit: NASA Earth Observatory(link is external)

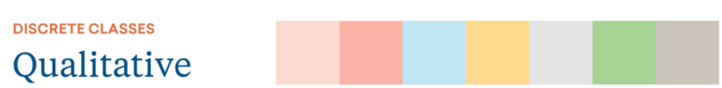
Hue, in its most basic meaning, is the classification of the rainbow color. More specifically, it is the dominant wavelength of color depicted by the name such as red, blue, or green. Hue is especially important in the area of cartography where it is often used to distinguish between different elements on a map. Certain conventions exist for the use of some hues in cartographic design, such as blue for water or green for vegetation; such conventions have been used for many centuries. Hue is especially useful in displaying nominal, or qualitative data, where different hues represent different, unrelated phenomena. Hue is especially useful in displaying nominal, or qualitative data, where different hues represent different, unrelated phenomena.

Lightness is another dimension of color; it describes how perceptually close a color appears to a pure white object. Lightness is also commonly called value, though cartographers sometimes avoid that term, as value is also used to describe data values—using the same word for both items can cause confusion. Lightness is perhaps the most important of the three perceptual dimensions when it comes to data representation and is used to show ordered differences. Lightness works well for visually encoding the order and/or magnitude of thematic data values—typically, lighter colors signify lower data values (i.e., less signifies less), and darker, more visually-prominent features signify higher data values.

The third dimension of color is saturation. Saturation is also sometimes called chroma. In map design, saturation is generally less important than hue and value, but it still can play an important role. Highly saturated colors are particularly useful for calling attention to small but important map elements that would otherwise be lost.

The basic principle is that variations in hue visualize nominal/categorical differences while variations in lightness visualize ordinal differences. But the strict application of this rule varies from one case to another: qualitative schemes may apply plenty of variations in lightness, especially when there is a large number of categories to display and sequential scales can benefit from hue variations when they are first and foremost ordered by lightness.

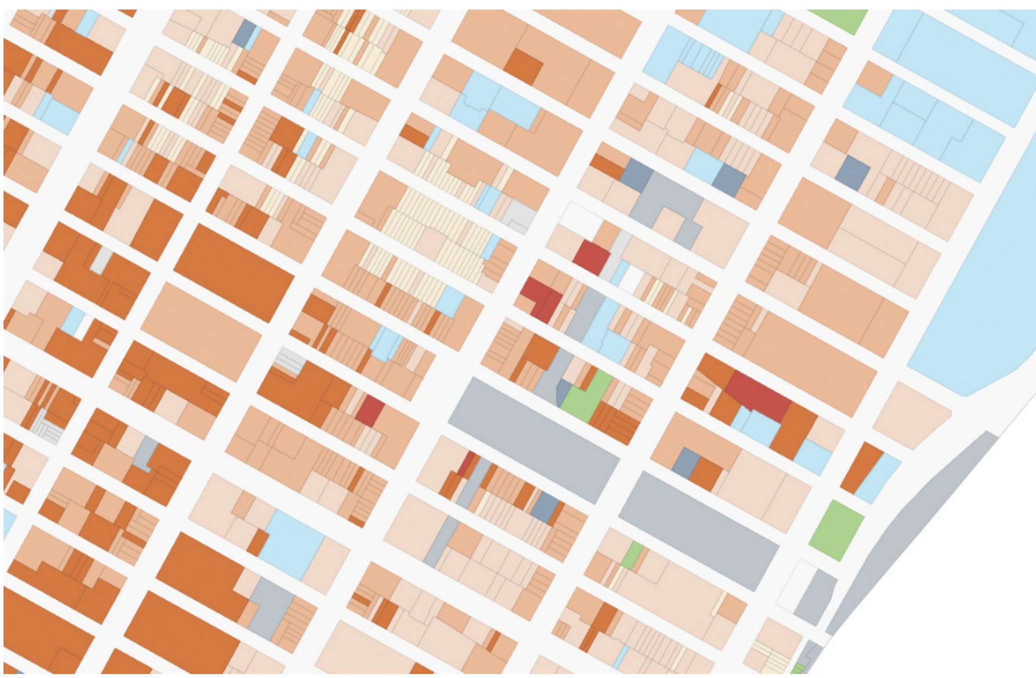
There are three main types of color schemes applied to maps: Qualitative, Sequential and Diverging.



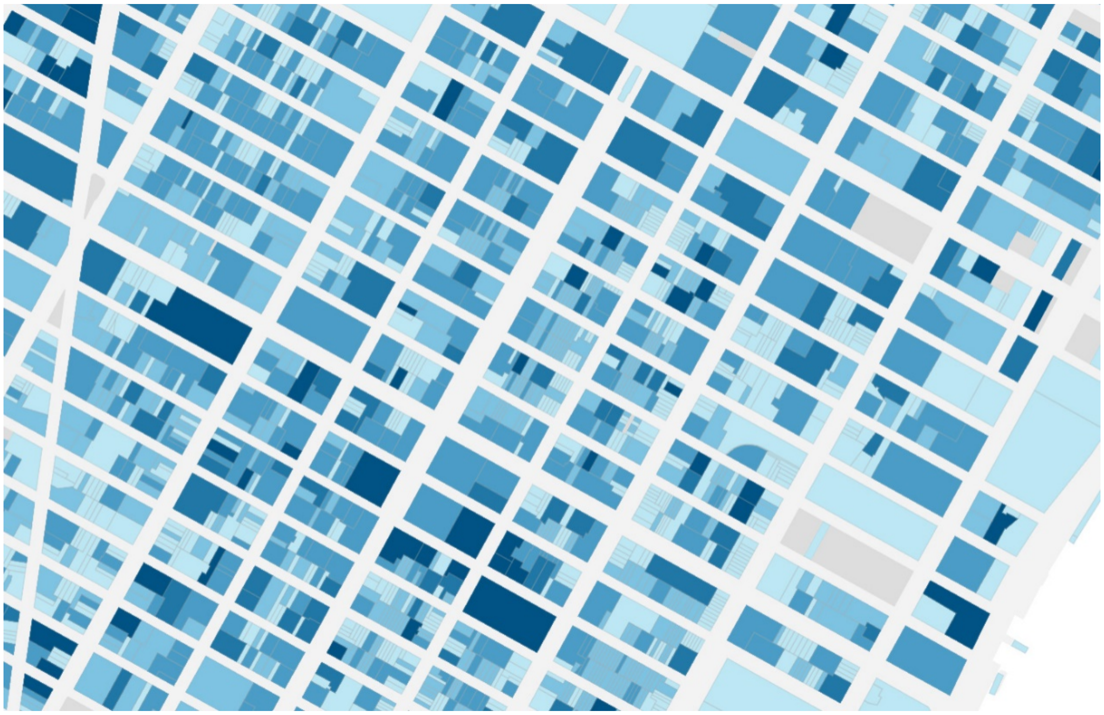
Qualitative schemes are applied to discrete unordered classes of nominal data such as race or ethnicity. They are not appropriate for mapping ordered numerical data. The distinction between classes becomes visible through variations in hue, ideally with no or slight lightness differences between colors. If a class needs to be highlighted it is possible to use a darker or more saturated color to visualize it. Qualitative schemes may also consist of paired hues with lighter and darker shades of the same color, applied to related categories (ex: related land use categories such as single and multifamily residential buildings).

Sequential schemes are applied to ordered, often numerical data such as floor area ratio per lot or population density per square mile. Changes in color lightness correspond to the progression from low to high: light colors are used for lower values and the dark colors are used for higher values. Sequential schemes can derive from both single and multi-hue combinations. The higher the number of data classes – the more difficult the distinction between each step.

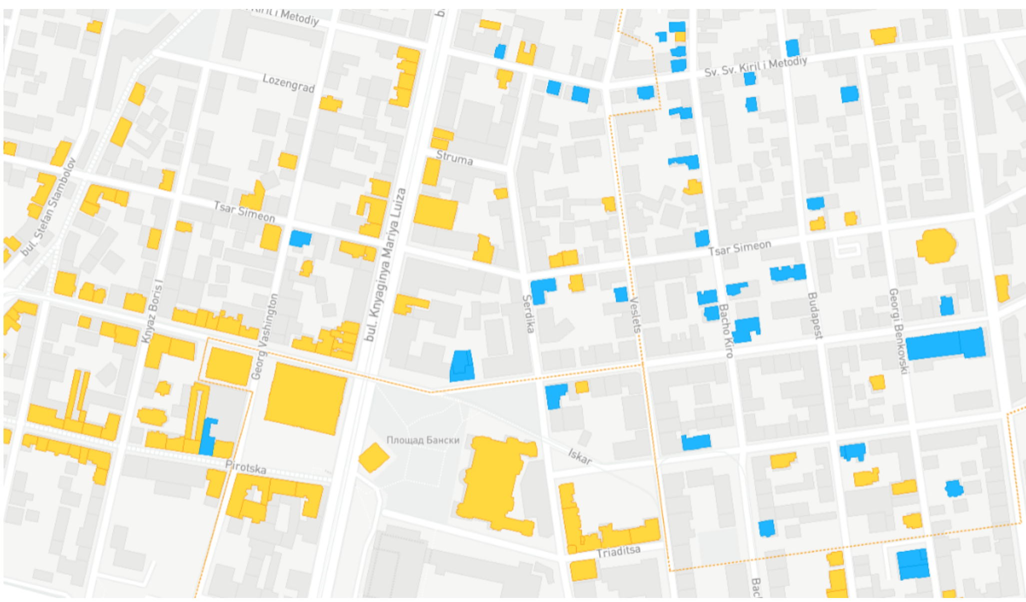
Diverging schemes are often described as a combination of two sequential schemes with a critical break point in the middle. The two sequences "diverge" from a shared light color that stresses important mid-ranges in the data. The two extremes are visualized by contrasting dark hues while changes in lightness are used to display intermediate values. Diverging schemes are usually symmetrical but specific data distribution may require shifting the break point towards either one of the extremes. Common examples of data suitable for diverging color scales are temperature variations and stock exchange dynamics.



Using a qualitative color scheme with both hue and lightness variations to map twelve categories of land use in New York City. Map by Morphocode. (from https://morphocode.com/the-use-of-color-in-maps/)



Using a sequential color scheme to map floor area ratios in New York City. Map by Morphocode. (from https://morphocode.com/the-use-of-color-in-maps/)



An example of a binary color scheme. The map shows buildings that are currently part of the cultural heritage in Sofia in yellow and buildings that are no longer listed as such in blue. Map by Morphocode. (from https://morphocode.com/the-use-of-color-in-maps/)

2.1.4 Map Generalization

Map generalization is the name of the process that simplifies the representation of geographical data to produce a map at a certain scale with a defined and readable legend. To be readable at a smaller scale, some objects are removed; others are enlarged, aggregated and displaced one to another, and all objects are simplified. During the process, the information is globally simplified but stays readable and understandable.

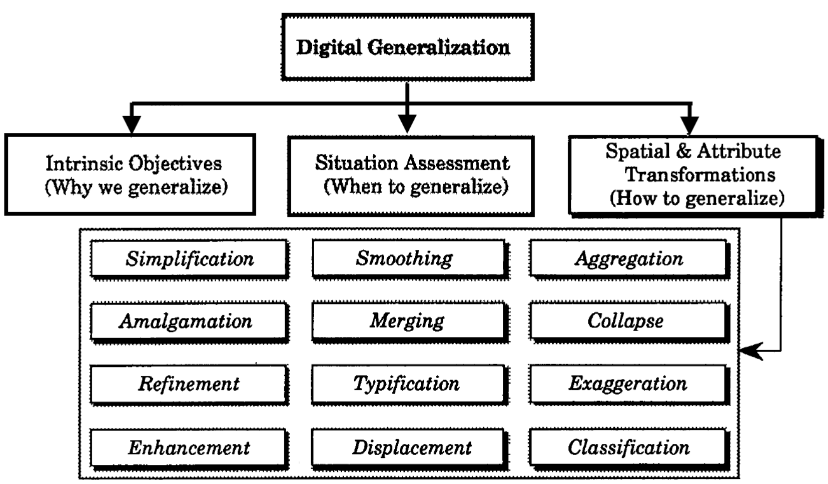
The smaller the scale, the less information is given per square kilometer. Conversely, the larger the scale, the more detailed is the area mapped for the same map size. For a given size of map sheet, nearly the same quantity of information is given for different scales, either privileging the density of field information (for larger scale) or the spatial extension (for smaller scale).

It is impossible to represent every detail of the world on a map; therefore, every map has been generalized to some extent. Generalization of maps has become necessary due to automatic production of maps on the web, and the increased amount of detailed GIS data available.

Generalization is closely related to the concept of visual hierarchy because visual detail helps to emphasize the most important map elements while less detailed features tend to attract less attention. After generalizing map data, the importance of what is remaining on the map must outweigh the insignificance of items that were generalized.

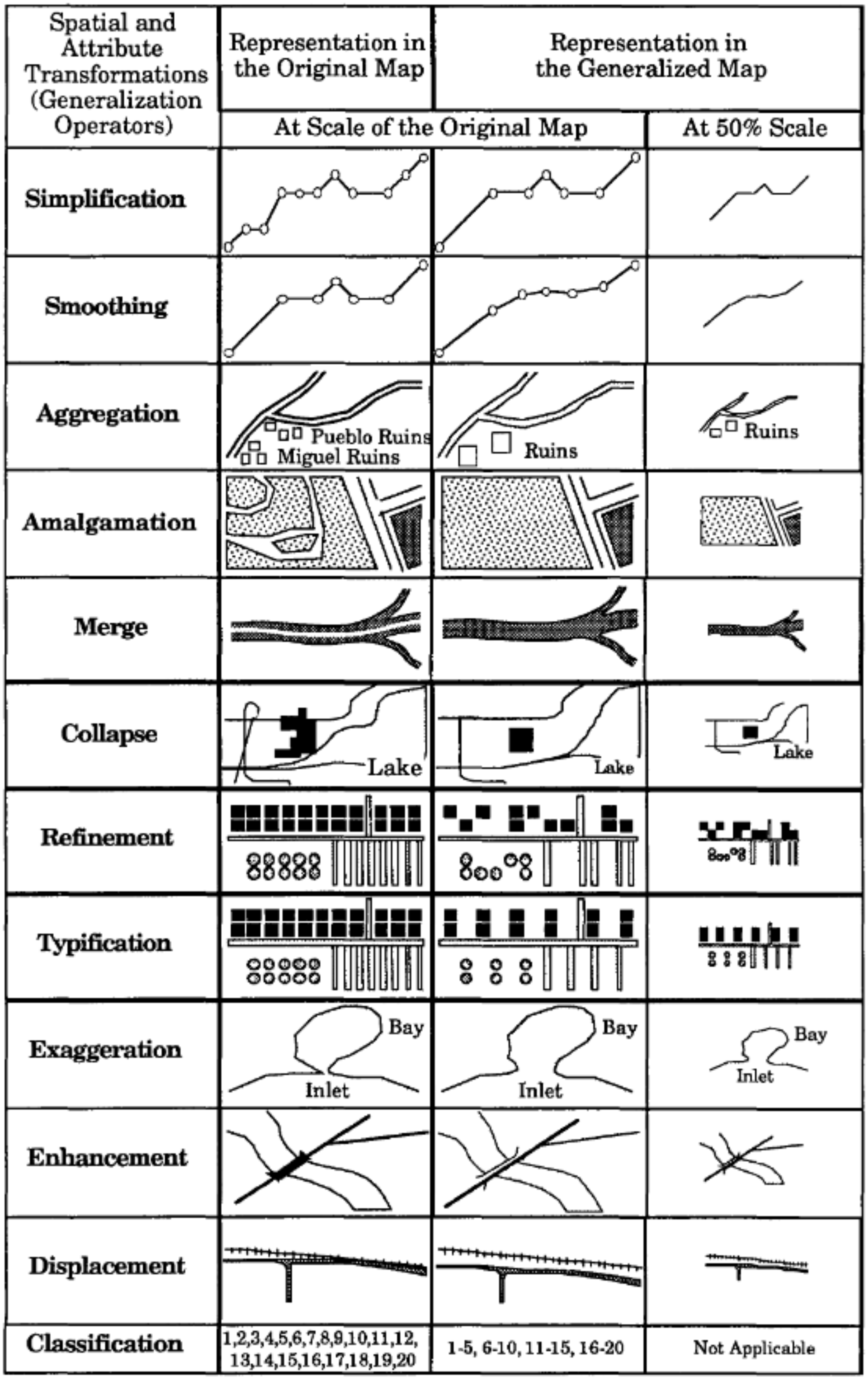
Generalization is important in GIS applications other than cartography because increased data detail requires more storage space, data entry time, and processing time. This would encourage GIS users to use less detailed data, but issues are also created in generalizing (especially aggregation) data, such as the modifiable areal unit problem. When thousands or even millions of records or pixels are being used, generalized data can make a big difference in geoprocessing run time if fewer digits or characters are in the attributes of the data.

Shea and McMaster (1989) list the following twelve categories of operators for cartographic generalization:



(from Shea, S. K., and McMaster, R. B.: Cartographic generalization in a digital environment: When and how to generalize, Proceedings of AutoCarto, Vol. 9. (1989))

Examples for the twelve operators are shown in the following image:


Spatial and Attribute Transformations (Generalization Operators)
Representation in the Original Map
Representation in the Generalized Map
At Scale of the Original Map
At 50% Scale
Simplification
Smoothing
Aggregation
Pueblo Ruins
Miguel Ruins
Ruins
Ruins
Amalgamation
Merge
Collapse
Lake
Lake
Lake
Refinement
Typification
Exaggeration
Bay
Inlet
Bay
Inlet
Bay
Inlet
Enhancement
Displacement
Classification
1,2,3,4,5,6,7,8,9,10,11,12, 13,14,15,16,17,18,19,20
1-5, 6-10, 11-15, 16-20
Not Applicable

(from Shea, S. K., and McMaster, R. B.: Cartographic generalization in a digital environment: When and how to generalize, Proceedings of AutoCarto, Vol. 9. (1989))

2.1.5 Map Scales

Map scale refers to the relationship (or ratio) between distance on a map and the corresponding distance on the ground. For example, on a 1:100000 scale map, 1cm on the map equals 1km on the ground.

Map scale is often confused or interpreted incorrectly, perhaps because the smaller the map scale, the larger the reference number and vice versa. For example, a 1:100000 scale map is considered a larger scale than a 1:250000 scale map.

A map scale is the relationship between a distance on a map and the corresponding distance on the earth. Map scale may be expressed as an equivalence, usually by different units (e.g., 1inch = 1mile or 1:63,360); or graphically, as a bar scale. Large scale maps make each feature look larger, and show a smaller geographic area. Small scale maps make each feature look smaller, and show a larger geographic area.

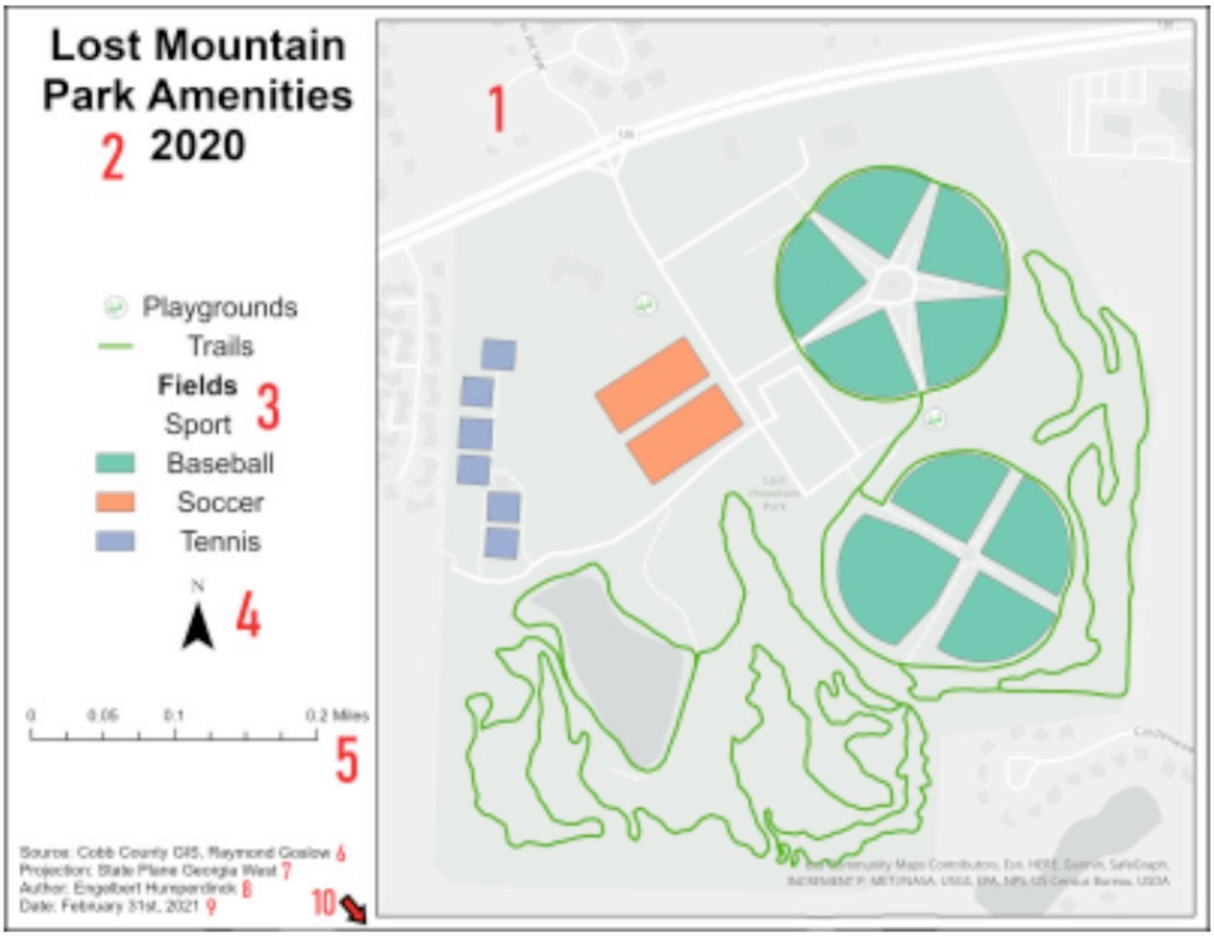
The map on the left is a large scale map of New York City where every map feature is drawn 1/63,360 the actual size on the earth. The map on the right is a smaller scale map of New York where every map feature is 1/12,672,000 the actual size on the earth. (from https://www.caliper.com/glossary/what-is-a-map-scale.htm)

2.2 Map Layout

Map Layout is the assembling of the various elements of a map into a single whole, including the map itself, its legend, title, scale bars, and other elements. Map symbolism can rarely stand alone to sufficiently depict all the necessary information that a map is trying to tell; additional explanation and context is usually needed. Their primary purpose is to give a place identity, orientation, subject matter, symbolization, etc. This term usually refers to the combination of the map image with auxiliary elements; the assembling of the geographic symbols within the map image is called Map composition.

<u>The 10 Required Map Elements</u> (from Patterson 2021)

Every map that you create MUST contain all 10 basic map elements. These elements are the bare minimum information that your maps must include. The map below shows each one with an explanation.



This map contains all 10 basic map elements.

1. **Map Body** - The map body shows the area under discussion and all the data being included. You might call it the "actual map", even though the map is not complete without the other 9 elements. It should take up most of the space on the page layout.
2. **Title** - The title tells what the map is all about. It should be in a prominent area and in a large font.
3. **Legend** - The legend explains what all the symbols on the map mean. It should be easy to find and easy to read.
4. **North Arrow** - The north arrow helps the map reader get oriented.
5. **Scale** - The scale informs the reader of the size relationship between the map and the real world. It can be shown in a bar or described in words or numbers.

6. **Source** - All the sources you used for map data need to be credited here.
7. **Projection** - No flat map can perfectly represent a 3D globe - the projection describes how that transformation is done. It can also be a called a spatial reference.
8. **Author** - You matter! Give yourself (and anyone who helped you) credit for making the map.
9. **Date** - The date you finished making the map. In a classroom setting, this can instead be the date the assignment is due, even if you made it earlier.
10. **Neatline** - This is the thin line containing all the other map elements so none of them are "loose".

2.3 Overall Design of Maps

Cartographic design is one part of a larger process in which maps play a central role. This cartographic process begins with a real or imagined environment or setting. As map makers gather data on the subject they are mapping (usually through technology and/or remote sensing), they begin to recognize and detect patterns that can be used to classify and arrange the data for map creation (i.e., they think about the data and its patterns as well as how to best visualize them on a map). After this, the cartographer compiles the data and experiments with the many different methods of map design and production (including generalization, symbolization, and other production methods) in an attempt to encode and portray the data on a map that will allow the map user to decode and interpret the map in the way that matches the intended purpose of the map maker. Next, the user of the map reads and analyzes the map by recognizing and interpreting the symbols and patterns that are found on the map. This leads the user to take action and draw conclusions based on the information that they find on the map. In this way, maps help shape how we view the world based on the spatial perspectives and viewpoints that they help create in our mind

<u>Basic principles of map design</u> (from Buckley 2021)

Cartographers apply many design principles when compiling their maps and constructing page layouts. Five of the main design principles are legibility, visual contrast, figure-ground organization, hierarchical organization, and balance. Together these principles form a system for seeing and understanding the relative importance of the content in the map and on the page. Without these, map-based communication will fail.

Visual contrast and legibility provide the basis for seeing the contents on the map. Figure-ground organization, hierarchical organization, and balance lead the map reader through the contents to determine the importance of things and ultimately find patterns.

Blow introduces you to these five principles and explains their importance in cartography. It's worth noting that these principles are not applied in isolation but instead are complementary. Collectively, they help cartographers create maps that successfully communicate geographic information.

1) Visual Contrast

Visual contrast relates how map features and page elements contrast with each other and their background. To understand this principle at work, consider your inability to see well in a dark environment. Your eyes are not receiving much reflected light, so there is little visual contrast between the objects in your field of view and you cannot easily distinguish objects from one another or from their surroundings. Increase illumination, and you are now able to distinguish features from the background. However, the features will still need to be large enough to be seen and understood so that your mind can decipher what your eyes are detecting.

The concept of visual contrast also applies in cartography (Figure 2.1). A well-designed map with a high degree of visual contrast can result in a crisp, clean, sharp-looking map. The higher the contrast between features, the more some features will stand out (usually features that are darker or brighter). Conversely, a map that has low visual contrast can be used to promote a more subtle impression. Features that have less contrast appear to belong together.

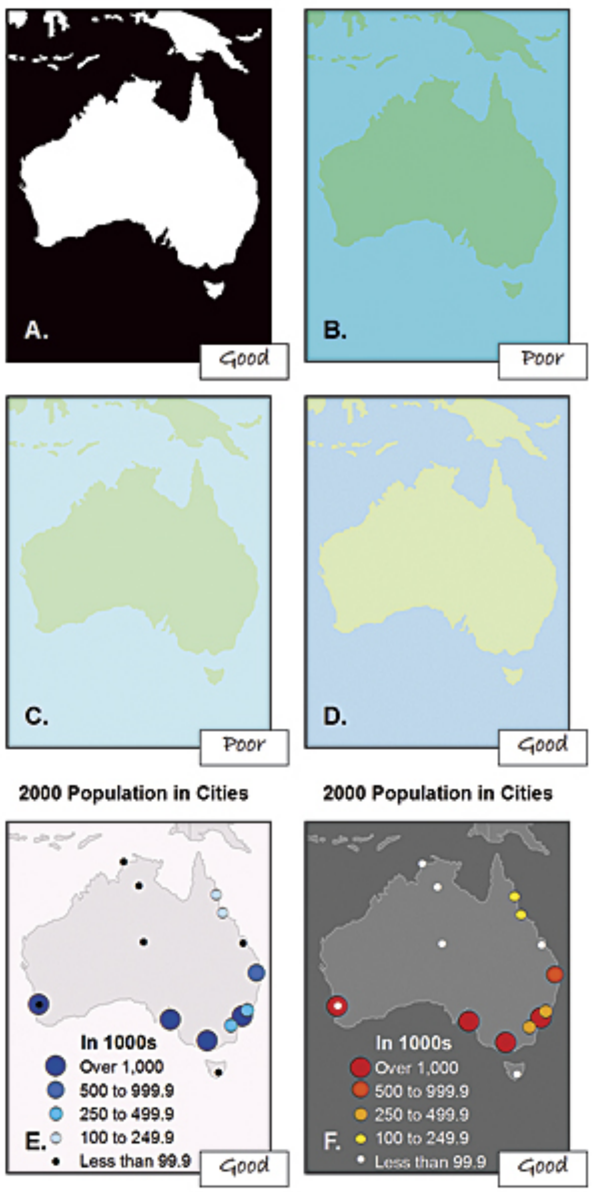


Figure 2.1 Although black and white (A) provide the best visual contrast, this is not always the best color combination for maps. When using colors of similar high (B) or low (C) saturation (brightness), the hues (blue and green, in this case) must be distinguishable. If they are not, varying the saturation or value (lightness or darkness) of a color (as with the water in D) can create the contrast that is missing. Operational overlays should contrast with the basemap (E and F).

2) Legibility

Legibility is the ability to be seen and understood. Many people strive to make their map contents and page elements easily seen, but it is also important that they can be understood. Legibility depends on good decision making when selecting symbols. Choosing symbols that are familiar and are appropriate sizes results in symbols that are effortlessly seen and easily understood (Figure 2.2). Geometric symbols are easier to read at smaller sizes. More complex symbols require more space to be legible.

Visual contrast and legibility can also be used to promote the other design principles: figure-ground organization, hierarchical organization, and balance.

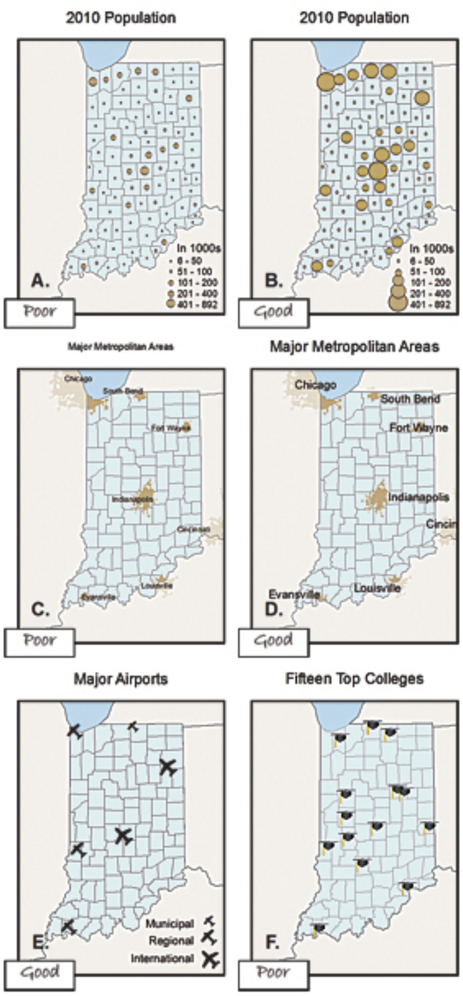


Figure 2.2 Symbols (A) and text (C) that are too small are illegible. Appropriately sized symbols (B) and text (D) can be easily distinguished and read. Using familiar geometric icons, such as an

airplane for airports (E), helps readers immediately understand the meaning of the symbol. More complex symbols, such as a mortarboard for universities (F), need to be larger to be legible.

3) Figure-Ground Organization

Figure-ground organization is the spontaneous separation of the figure in the foreground from an amorphous background. Cartographers use this design principle to help map readers focus on a specific area of the map. There are many ways to promote figure-ground organization, such as adding detail to the map or using a whitewash, a drop shadow, or feathering.

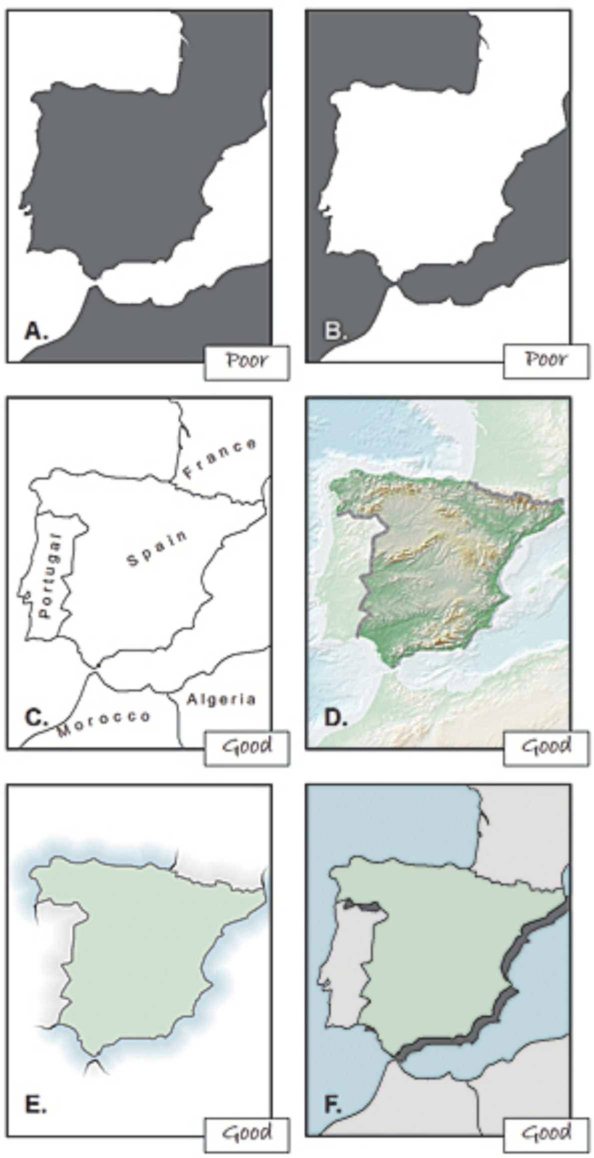


Figure 2.3 It is sometimes hard to tell what is the figure and what is the ground (A and B). Simply adding detail to the map (C) can help map readers distinguish the figure from the ground. Using a whitewash (D), feathering (E), or a drop shadow (F) can also help.

4) Hierarchical Organization

As noted in *Elements of Cartography*, Sixth Edition, one of the major objectives in mapmaking is to "separate meaningful characteristics and to portray likenesses, differences, and interrelationships." The internal graphic structuring of the map (and, more generally, the page layout) is fundamental to helping people read your map. You can think of a hierarchy as the visual separation of your map into layers of information. Some types of features will be seen as more important than other kinds of features, and some features will seem more important than other features of the same type. Some page elements (e.g., the map) will seem more important than others (e.g., the title or legend).

This visual layering of information within the map and on the page helps readers focus on what is important and lets them identify patterns. The hierarchical organization of reference maps (those that show the location of a variety of physical and cultural features, such as terrain, roads, boundaries, and settlements) works differently than for thematic maps (those that concentrate on the distribution of a single attribute or the relationship among several attributes). For reference maps, many features should be no more important than one another and so—visually—they should lie on essentially the same visual plane. In reference maps, hierarchy is usually more subtle and the map reader brings elements to the forefront by focusing attention on them. For thematic maps, the theme is more important than the base that provides geographic context.

Figure 2.4 When the symbols and labels are on the same visual plane (A), it is difficult for the map reader to distinguish among them and determine which are more important. For a general reference map (B), using different sizes for the text and symbols (e.g., city points and labels), different line styles (e.g., administrative boundaries), and different line widths (e.g., rivers) are some of the ways you can add hierarchy to the map. When mapping thematic data (C), the base information (e.g., county boundaries and county seats) should be kept to a minimum so that the theme (e.g., soils) is at the highest visual level in the hierarchy.

5) Balance

Balance involves the organization of the map and other elements on the page. A well-balanced map page results in an impression of equilibrium and harmony. You can also use balance in different ways to promote edginess or tension or create an impression that is more organic. Balance results from two primary factors: visual weight and visual direction. If you imagine that the center of your map page is balancing on a fulcrum, the factors that will tip the map in a particular direction include the relative location, shape, size, and subject matter of the elements on the page.

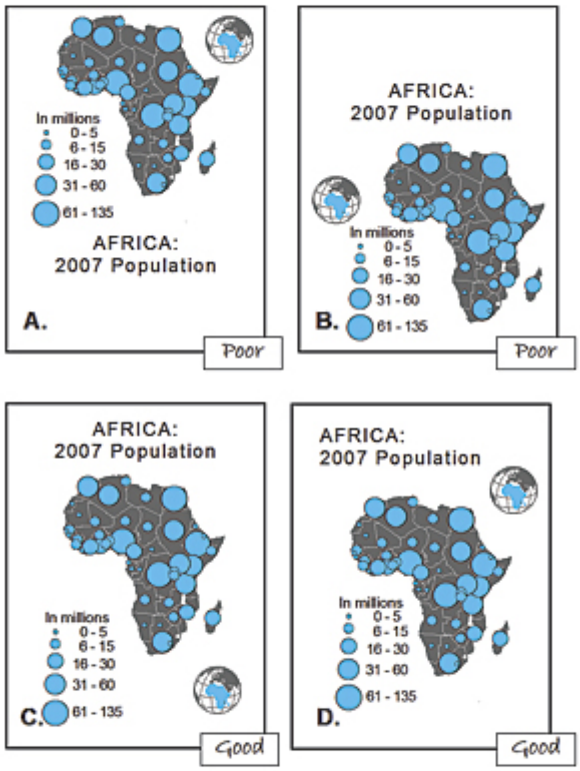


Figure 2.5 Positioning heavier elements together can make the page look top-heavy (A) or bottom heavy (B). Centering the map slightly above center (C) ensures that it is in the most prominent position on the page. The position of elements can also cause the eye to move in a desired direction. In D, the title is the first thing read, followed by the locator map, then the map of Africa, and finally the legend.

Together these five design principles have a significant impact on your map. How they are used will either draw the attention of map readers or potentially repel them. Giving careful thought to the design of your maps using these principles will help you to assure that your maps are ones people will want to look at!

Resources

These cartography textbooks provide more in-depth discussions of the design principles described in this article and how they are applied in cartography.

Dent, Borden D., Jeffrey S. Torguson, and Thomas H. Hodler. 2009. *Cartography: Thematic Map Design*, Sixth Edition, 207–222. Boston, MA: WCB-McGraw Hill.

Robinson, Arthur H., Joel L. Morrison, Phillip C. Muehrcke, A. Jon Kimerling, and Stephen C. Guptill. 1995. *Elements of Cartography*, Fifth Edition, 324–330. New York City, NY: John Wiley & Sons, Inc.

Slocum, Terry, Robert B. McMaster, Fritz C. Kessler, and Hugh H. Howard. 2009. *Thematic Cartography and Geographic Visualization*, Third Edition, 212–221. Upper Saddle River, NJ: Pearson Prentice Hall.

3. Modern geo-visualization techniques (from Nöllenburg 2007 and Buckley et. al. 2004)

Geo-visualization integrates approaches from visualization in scientific computing (ViSC), cartography, image analysis, information visualization, exploratory data analysis (EDA), and geographic information systems (GISystems) to provide theory, methods and tools for visual exploration, analysis, synthesis, and presentation of geospatial data. It is clear that geo-visualization research is a multidisciplinary task. Since it is the human who uses visualizations to explore data and construct knowledge, effective geo-visualization techniques must above all take the user needs into account.

3.1 Driving Forces of Geo-visualization

So what is the reason for the increasing interest in geo-visualization over the last 15 years? There are three driving forces for geovisualization.

The first is the rapid advances that have been made in graphics and display technology. The availability of both low-cost 3D graphics hardware in personal computers and the development of highly immersive 3D virtual environments resulted in investigating the potential that these technologies have for visualizing geospatial data. However, this emphasis on realism contrasts with the history of cartography that points to centuries of successful abstraction making the world easier to understand according to MacEachren. Indeed, maps filter out unnecessary details of the environment in order to highlight interesting information. For example, a road map based on satellite images would be extremely hard to use. The challenge is to study the relative advantages and disadvantages of realism and abstraction in geo-visualization and then, depending on the problem context, potentially integrate both abstract and realistic displays in a single geo-visualization environment.

The second driving force for geo-visualization is the need to analyze and explore a dramatically increasing amount of geospatial data that are routinely collected these days by a multitude of scientific and governmental institutions, private companies, and individuals. This is due to an increasing availability and decreasing cost of technology to acquire, store, and process these data. A majority of the data, MacEachren and Kraak estimated up to 80 percent, contain geospatial references, e.g., coordinates of environmental measurements, census data, positions of vehicles, ships, planes, and parcels, addresses of customers, etc. These data, often characterized by a high dimensionality, are a vast source of potentially valuable information for research and decision making, e.g., in studying disease incidence patterns, traffic flows, credit card fraud, or climate change. The large volume of many data sets poses challenging problems for their exploration. With increasing data volume and complexity humans quickly reach the limit of their capacities in analyzing raw numeric and textual data. The goal of geo-visualization is to combine the strengths of human vision, creativity, and general knowledge with the storage capacity and the computational power of modern computers in order to explore large geospatial data sets. One way of doing this is by presenting a multitude of graphic representations of the data to the user, which allow him or her to interact with the data and change the views in order to gain insight and to draw conclusions.

Finally, the third driving force for geo-visualization is the rise of the Internet and its development into the prominent medium to disseminate geospatial data and maps. On the one hand, the Internet facilitates collaboration of expert users at different places, which is one of the ICA Commission's research challenges, and, on the other hand, it enables geo-visualization applications to address the public. Reaching the public is an important aspect both for governmental agencies and for business companies who provide and sell services based on geospatial information.

3.2 Visualization Methods and Techniques

A number of techniques and methods adapted from cartography and scientific visualization are studied and applied in geo-visualization. Below introduce the main methods and techniques.

3.2.1 2D Cartographic Visualization

The most common visualization method for geospatial data is a cartographic display of some form, i.e., a map where the area under consideration is depicted, and onto which the data of interest are plotted at their corresponding coordinates. *Space* is used to depict space by mapping latitude and longitude to the coordinate axes of the map drawing area. This might seem to be the most natural way of using this graphic variable. However, there are good reasons of linking for example population to space resulting in a *cartogram* where area on the map is not proportional to a certain geographic area but in this case to the number of people living in that very area. Still, cartograms usually try to preserve the users' mental map by keeping similar shapes and by preserving the adjacencies between the depicted areas. An example of a world population cartogram is shown in Figure 3.1. Especially when the focus of a map is on social, economic, or political issues, cartograms help to draw the users' attention to population as the map's theme while avoiding to emphasize large but sparsely inhabited regions.

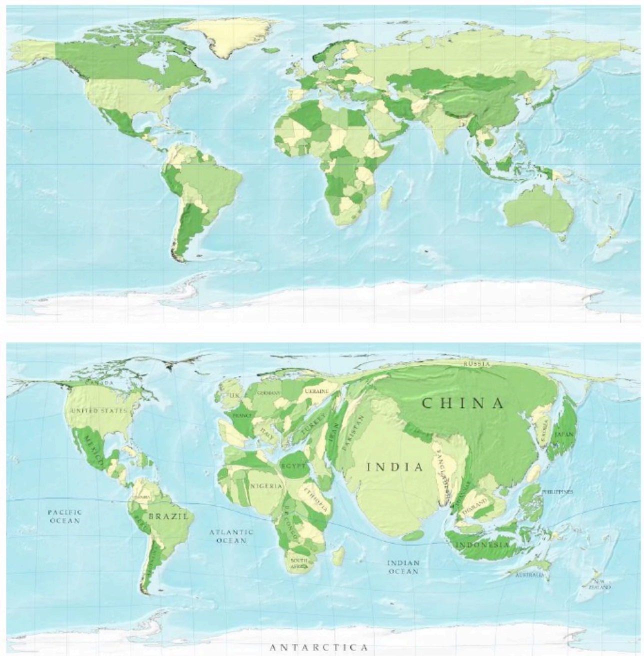


Figure 3.1 Geographical world map (top) and world population cartogram (bottom). In the cartogram the size of a country is proportional to the size of its population. *Image courtesy of M. Newman.*

For both cartograms and geographic maps the interesting aspect is of course how to depict abstract attributes of the data or at least a subset of them. Among the most popular methods to represent categorical but also numerical data are *choropleth* maps. A *choropleth* map uses the graphic variables describing properties of color or texture to show properties of non-overlapping areas, such as provinces, districts, or other units of territory division. A number of categories is mapped to distinct colors or textures which are used to fill the areas accordingly. Examples are land cover/use with categories like forest, crop land, housing, etc. or election results per district, e.g., displaying the percentage of votes for a certain party or the voter turnout as in Figure 3.2. For unordered data well-distinguishable colors are needed while for ordered data it is important to find a lightness or hue scale that represents the original range of numbers efficiently, i.e., that the user can estimate values and differences from the colors. Alternatively, a continuous range of attribute values is mapped to a continuous color range without assigning values to a fixed number of classes. While choropleth maps help to show general trends in the data there is certainly a loss of information because the user cannot map a certain color to its exact numerical value. Furthermore, a choropleth map can only express one or two attributes of the data (by using a two-dimensional color scheme or by combining color and texture). Andrienko and Andrienko described a selection of methods to represent single and multiple attributes in a map. Depending on the type of the attributes (logical, numeric, or nominal), they used bar and pie diagrams common in statistic visualization. Similarly, glyph-based techniques from visual data mining can also be combined with map displays. Using their geospatial reference, glyphs or statistical diagrams are placed on the map and thus both spatial and multidimensional abstract attributes are represented on a single map. However, if the number of symbols or attributes exceeds a certain limit the symbols become hard to compare and other non-map based techniques from visual data mining should be applied in addition to the display of a map or cartogram, see Section 3.2.3.

Other approaches for displaying high-dimensional data reduce the dimensionality of the data, e.g., by applying statistical techniques like principal component analysis or by calculating compound indices representing, for example, the socio-economic development of a region. The disadvantage, especially for explorative visualization, is that through the loss of information potential patterns of some attributes might get lost.

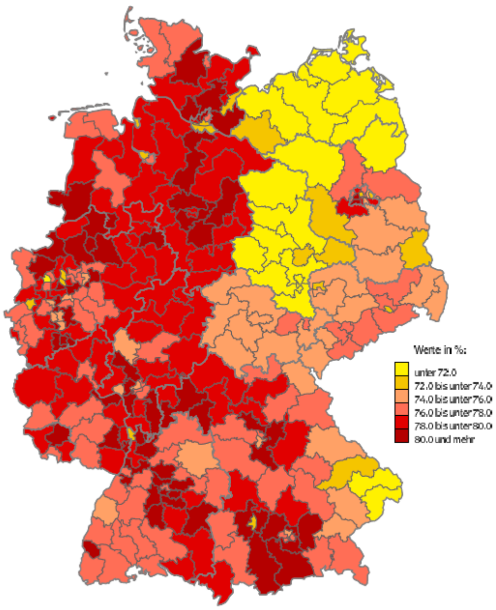


Figure 3.2 A choropleth map showing turnout of voters in the 2005 federal elections in Germany. Image courtesy of Statistisches Bundesamt, Wiesbaden.

3.2.2 3D Cartographic Visualization

In contrast to traditional paper maps and two-dimensional visualization methods, geovisualization can go one step further and use the potential of increasingly experiential representation technologies. 3D visualization includes the full range from regular 3D graphics hardware in desktop computers to immersive 3D displays, CAVEs (Cave Automatic Virtual Environments), and Power Walls providing stereoscopic views. Since humans live in a three-dimensional environmentour perception and cognition is adapted to processing 3D visual stimuli. But there is still little known about when 3D visualization is appropriate and how it can effectively enhance visual thinking.

Cartography has a long and successful tradition using abstraction to depict a wide range of data on maps. In contrast, the focus of computer graphics technology is on producing increasingly realistic images and virtual environments. Virtual reality techniques are widespread, for example, in architecture and landscape planning where realism is very important. Depending on the geo-visualization task, realism can be a distraction and insight is more likely when using abstract visual symbolism. But as MacEachren et al. pointed out, there had been only few efforts exploring abstract visualizations of geospatial data in 3D.

In terms of the 3D representation of the data, MacEachren et al. distinguished between using the three dimensions of the representation to display the three dimensions of physical space, using one or two dimensions for non-spatial data, e.g., income or time, and using all three dimensions for abstract data.

Today, the most widespread use of 3D is at the level of visual representation while the display is a 2D screen. It is important to be aware of the implications that the projection of a 3D representation onto a 2D plane has. Depth cues such as perspective and occlusion also cause problems because distances are harder to estimate, and occlusion hides objects depending on the viewpoint. Ware and Plumlee observed that due to occlusion humans cannot perceive much information in the depth direction while the x- and y-directions, orthogonal to the line of sight, can convey complex patterns. A set of interactive navigational controls are necessary to move within the 3D representation, e.g., zooming or flying. As Wood et al. pointed out the effectiveness of the virtual environment metaphor relies to some extent on navigational realism. While moving (e.g., walking or flying) slowly through the visual space maintains a sense of orientation, faster modes of movement such as teleporting lose the context and the user has to reorientate.

3.2.3 Visual Data Mining Tools

Visual data mining, also denoted as exploratory data analysis (EDA) in statistics, is a human-centered task that aims at visually analyzing data and gaining new insights. This contrasts computational data mining techniques which use algorithms to detect patterns in the data. Effective visual data mining tools need to display multivariate data in a way that the human viewer can easily perceive patterns and relationships in the data. Visual data mining in general is not tailored specifically for geospatial data. Since geospatial data usually have many abstract attributes these general techniques can be applied for displaying non-spatial attributes of the data. Visualization techniques for multivariate data were broadly classified as geometric, glyph- or icon-based, pixel-oriented, and hierarchical by Schroeder and Keim and Kriegel. In a geo-visualization context, geometric and glyph-based techniques are most common. Graph-drawing techniques that depict relationships between individual data items are also covered in this section.

Geometric Techniques Two geometric techniques commonly used in geo-visualization are scatter plots and parallel coordinate plots. Scatter plots in their basic two-dimensional form depict objects as points in a coordinate system where the axes correspond to two selected attributes, see Figure 3.3(a) for an example. Elements in the data set with similar values in these attributes form visual clusters in the scatter plot. The idea of a scatter plot can be extended to three dimensions but then phenomena as occlusion and different perception of depth may occur. The extension to more than three dimensions is often implemented by drawing a scatter-plot matrix containing one scatter plot for each pair of attributes. This, however, makes the identification of multidimensional patterns difficult because many plots in the matrix need to be linked mentally.

Parallel coordinate plots (PCP) are a means of displaying high-dimensional data in a single plot. In a PCP, one dimension is used to place multiple parallel axes, each of which represents one attribute of the data. Each element of the data set is then characterized by the values of its attributes which are connected along the axes and thus build a geometric profile of that element as depicted in Figure 3.3(b). Since all elements are plotted in this way, the user can identify similar objects by comparing the geometric shape of their profiles. However, depending on the number of profiles, overplotting occurs and may result in poor legibility. Keim and Kriegel estimated that about 1,000 items could be displayed at the same time. Moreover it becomes difficult to compare profiles based on an increasing number of attribute axes. Another important aspect of PCPs is the order of the attributes plotted along the parallel axes since this order has a strong influence on the shapes of the profiles. Hence, a user should be able to rearrange the attributes manually or based on sorting algorithms.

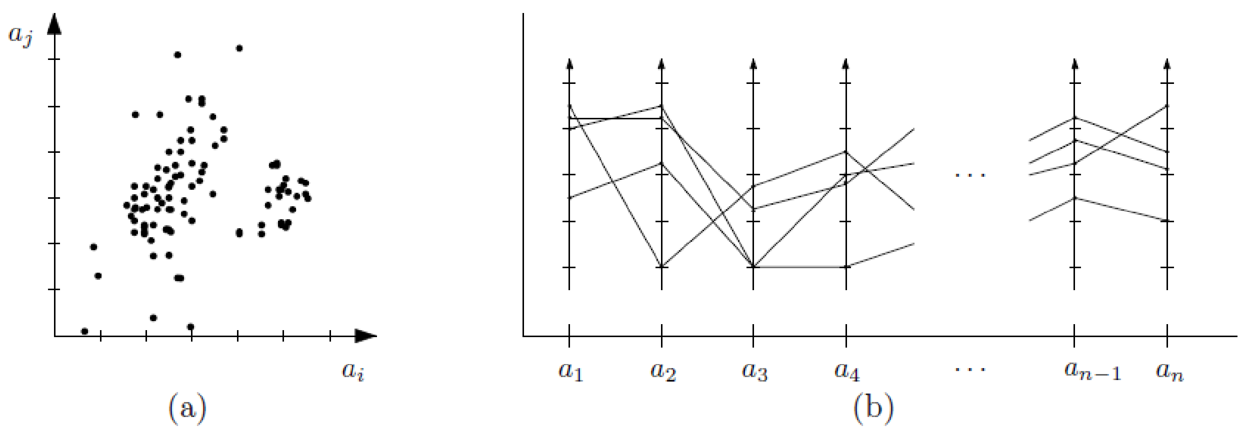
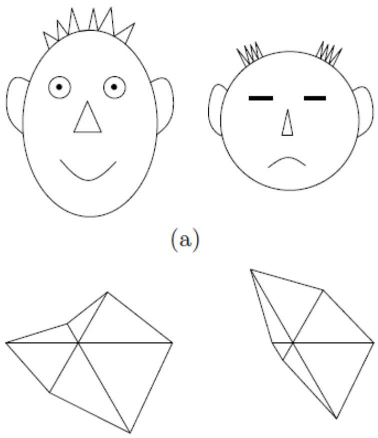


Figure 3.3 Example of a 2D scatter plot in subfigure (a) and a parallel coordinate plot in subfigure (b).

Glyph-Based Techniques Glyph-based or icon-based techniques use a mapping of multiple attribute values to a set of different visual features of a glyph which in turn represents one data object. Two examples of such techniques are Chernoff faces and star plots. In a Chernoff face, different variables of the data are related to facial features of an iconic face, such as size and shape of mouth, eyes, ears, etc., see Figure 3.4(a). The motivation of using faces to depict multidimensional data is that human mind is used to recognize and compare faces. However, different features, e.g., shape of the eyes and area of the face, are hard to compare and Chernoff faces are, in contrast to human faces, not perceived pre-attentively such that there is no advantage over other types of glyphs.

Star plots depict the value of a set of attributes by the length of rays emanating from the center of the glyph. The endpoints of the rays are connected to create a closed shape as depicted in Figure 3.4(b). While the maximum number of facial features in Chernoff faces is reached quickly, star plots can display data with higher dimension by increasing the number of rays. Again, as for parallel coordinate plots, the order of the attributes influences the shape of the star plots.



(a)

(b)

Figure 3.4 Two examples of glyph-based techniques: subfigure (a) shows two Chernoff faces and subfigure (b) shows two star plots on six attributes.

A nice property of glyph-based techniques is that they can be easily combined with map displays by placing each glyph according to its geospatial coordinates on the map. However, with an increasing number of symbols or attributes glyph-based techniques are of limited use due to the difficulty of visually recognizing patterns and distinguishing features on a display with too many glyphs or glyphs with too many features.

Graph-Drawing Techniques Geospatial data often contain links between related elements, e.g., routes, trade connections, etc. Exploring such data sets includes the search for patterns in the link structure between items. Data containing relationships between elements are mathematically modeled as a graph consisting of a set of nodes, the data elements, and a set of (weighted) edges, the links between elements. The research area of graph drawing provides a multitude of algorithms for visualizing such graphs. In general, for graph drawing the emphasis is on finding a layout, i.e., positions of nodes and edges of a given graph that satisfies certain aesthetic criteria, e.g., few edge crossings. In geo-visualization, there are usually certain constraints on such a layout since nodes already have a spatial location. In that case, finding a legible layout for the edges is of interest, for example in schematizing road networks. In other cases, such as drawing metro maps, the network topology is more important and node positions are only required to satisfy certain relative positions (e.g., left, right, above, below) in order to preserve the user's mental map. Finally, some data is best analyzed by putting no restrictions to node positions and using a general algorithm to find a graph layout in which link patterns can be identified visually. Such methods are applied in visual social network analysis. In the latter cases, where node positions are modified, a map display of the true geography in combination with a graph layout focusing on the link topology is helpful for identifying both spatial and link-based patterns. An example by Rodgers visualizing trade volume between regions of the world as a graph is shown in Figure 3.5. The stronger a trade relationship between two regions the more they attract each other in the graph layout.

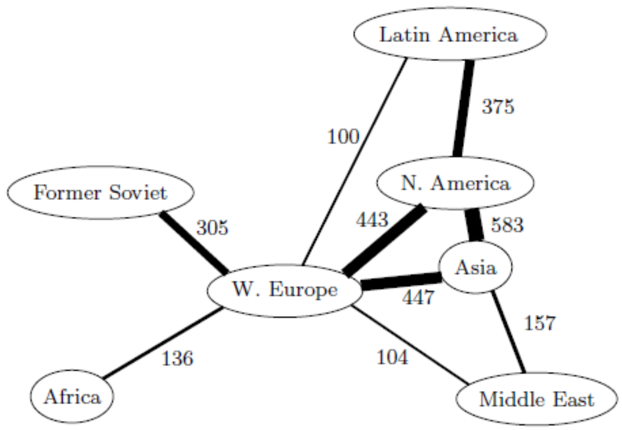


Figure 3.5 A graph showing trade relationships. Edges are weighted by trade volume and drawn shorter and thicker with increasing weight. Only edges with at least 100 billion dollars trade volume are shown.

3.2.4 Animation

The methods described so far are primarily static displays of geospatial data. They can all be printed and studied on paper. In geo-visualization, however, dynamic and interactive displays are core concepts today. Since the 1930s cartographers are experimenting with map movies and animated maps. Leaving interaction aside, animated maps are using time to add another visual dimension to the display. It is intuitive to relate time to time, just as space depicts space in most maps. In this case, the time period of the data is mapped to the animation time. Each scene or frame of the animation shows the state of the data at one moment accordingly. Thus, the temporal change of the attributes becomes visible. It may be necessary to smooth the animation using interpolation to generate intermediate frames. Scenes can also be reordered from chronological order to an order based on attribute values. This may be helpful for studying at what points of time events with similar properties took place. For example, earthquake data can be ordered by the number of human fatalities in such a way that the beginning of the animation shows the least and the end the most catastrophic earthquakes. Animation can also be used to display spatial features, e.g., animations of flights over the terrain. In other cases, the temporal dimension is used to display quantitative attributes by mapping their values to the blinking frequency of symbols or to highlight classes in a choropleth map by blinking. The presence of the temporal dimension in dynamic visualizations also introduces the potential to use acoustic variables. Although visualization is mostly concerned with visual aspects, using sound to complement dynamic graphics expands the possibilities of visualization. Krygier summarized the role of sound in geo-visualization from attracting attention over narrative voice and representing quantitative attributes (e.g., using the variable pitch) to sound maps for visually impaired.

Bertin's notion of visual variables has been extended to dynamic animated displays. Six dynamic variables are suggested: (1) temporal position, i.e., when something is displayed, (2) duration, i.e., how long something is displayed, (3) order, i.e., the temporal sequence of events, (4) rate of change, e.g., the magnitude of change per time unit, (5) frequency, i.e., the speed of the animation, and (6) synchronization, e.g., the temporal correspondence of two events.

Animation and its set of additional visual variables represents a powerful tool for map designers. However, animation should not be used carelessly and it is always worth asking why do I need to animate these data? From a user's perspective, things that change on a map attract more attention than the static background and moving objects attract more attention than objects that appear and disappear. The fact that animated displays change over time has an important disadvantage, especially if there are no interactive controls: there is always the risk that the user will miss important information. While a static display can be analyzed at an individual speed, an animation has to be followed at the predefined pace. Therefore, it is hard to compare data displayed at different points of time as the human brain usually forgets most details of previous frames of an animation. A study by Rensink et al. showed how difficult perceiving changes in an animation can be.

Harrower gave some guidelines for designing effective animated maps. Because dynamic variables attract more attention than static variables the information conveyed by static variables should be kept simple, e.g., by using a choropleth map with only few data classes (high, medium, and low) and a rather low level of detail. Details are often more effectively displayed using static visualization techniques. Temporal exaggeration, i.e., displaying durations not to scale, is often necessary since otherwise a short event, e.g., an earthquake, in an animation spanning several decades will be missed by the viewer. Directing the user's attention to critical events can be done for example by initially flashing new symbols on the map. In general, animated maps are better suited to depicting geographic patterns, e.g., growth or shrinkage of an area, rather than specific rates of change according to Harrower. Finally, he observed that people are less confident with animated maps than with static maps, due to less experience and training.

Many of the above problems can be avoided by giving the control of the animation to the user. In interactive animations, where the user can control the displayed level of detail and the speed of the animation, information is less likely to be missed and users feel more confident with the animation. Still, the study of dynamic displays with regard to their geospatial expressiveness is identified as one of the challenges and further usability studies are required.

3.2.5 Spatio-Temporal Visualization

Spatio-temporal data are very common in the earth sciences and related disciplines. The goal of many studies is to reveal, analyze, and understand patterns of temporal change of phenomena, such as global warming, population development, or spread of diseases. The previous section has presented animation as a means of displaying temporal data. Animation works well in displaying patterns if they are based on the same temporal sequence as the animation itself, e.g., showing trends like urban growth over time. However, Andrienko et al. criticized that for less evident patterns it is necessary to compare the data at different points in time which involves memorizing a large number of states in an animated display, even if interactive controls allow to pause, jump, and step through specific points in time. Thus it might be more effective to statically display selected moments in time simultaneously using small multiples. Then, an analyst can directly compare attribute properties of different points in time at his or her own speed. However, the number of simultaneous images on the screen is limited and long time series have to be evaluated piecewise. Andrienko et al. argue that, for all these reasons, spatio-temporal data exploration must be supported by a variety of techniques, possibly in combination with an animated display.

Andrienko et al. classified spatio-temporal data according to the type of temporal changes: (1) existential changes, i.e., appearance and disappearance of features, (2) changes of spatial properties, i.e., change of location, shape, size, etc., and (3) changes of thematic properties, i.e., qualitative and quantitative changes of attributes. Following this classification, they presented corresponding visualization techniques. All techniques involved a map display to visualize the spatial attributes of the data.

Data of existential changes usually consist of events or observations at specific moments or time periods during which a certain property holds, e.g., road congestion data. Hence a map showing these data always considers a selected time interval. If data items are represented by glyphs, one way to display the time associated with them is by using textual labels. Another possibility is using a color scheme to represent the age of the data. A 3D representation of space and time is a third and common method. In such a space-time cube, the third dimension corresponds to time while two dimensions represent geographical space. The reference map is usually displayed in the coordinate plane corresponding to time 0 and data items are positioned above the map depending on their spatial locations and their times of appearance. An example of a space-time cube is shown in Figure 3.6. It shows the trajectory of Napoleon's troops during the French campaign against Russia in 1812.

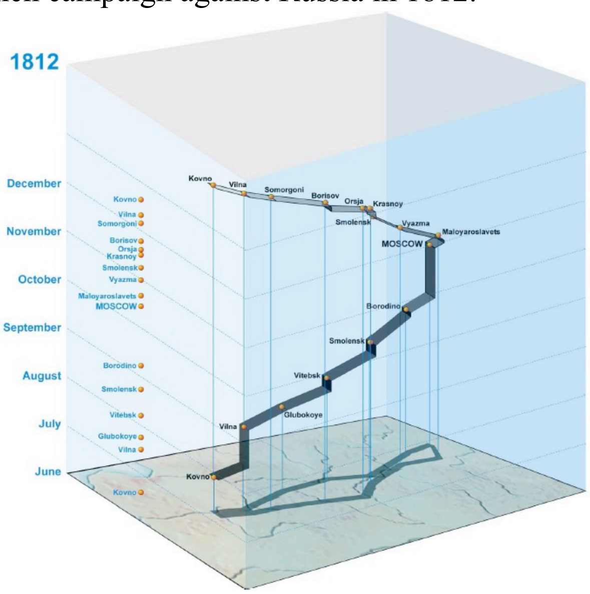


Figure 3.6 A space-time cube visualization of Napoleon's march in Russia. Image courtesy of M.-J. Kraak.

For data that contain moving objects, comparing object trajectories is of interest. Static 2D maps are able to show the trajectories of a small number of objects but in this simple form it is not possible to evaluate aspects like speed or whether two objects met at a crossing or just visited at different points in time. Andrienko et al. suggested animating object movements, either as a sequence of snapshots in time in which at each moment objects are shown at their current positions or using the movement history and showing the trajectories up to the current point in time. Movement history can optionally be limited to a specified time interval. It was found that the snapshot technique was suited for a single object while several objects were better observed displaying also the movement history. MacEachren suggested using the space-time cube to display trajectories which avoids the disadvantages of 2D trajectories mentioned above as it shows when and not just if an object visited a point.

There are several methods of displaying thematic attributes on a map. A very effective and common method is the choropleth map. Animating a choropleth map is able to give a good overview of the values in a selected attribute. However, it is difficult to estimate trends in a particular area on the map or to compare trends between different areas. Change maps, adapted from conventional cartography, use the choropleth map to show the differences of an attribute between two selected points in time. Mapping increase and decrease in attribute value to shades of two different colors allows to evaluate regional changes for two moments. Such a map is restricted to two points in time and the map can be misleading because information on the actual attribute values is lost, e.g., concerning crime data two areas can have very different burglary rates but still be colored the same if the rates both decrease by the same value. Andrienko and Andrienko combined time-series graphs with maps to avoid these disadvantages. A time-series graph is a two-dimensional plot of the temporal variation of attribute values, where time is represented on the x-axis and attribute values on the y-axis. Plotting all data in the same time graph gives an overview of the dynamics of the whole data set. To assess local behaviors, Andrienko and Andrienko plotted the time-series data individually for each area on the map and used the closed shape of the plot as a symbol superimposed on each area similarly to the glyph-based techniques, see Figure 3.7 for an example. This technique allows to evaluate changes and actual values of an attribute for the whole time period under consideration. The user can explore both spatial patterns and patterns in the attribute space in the same view.

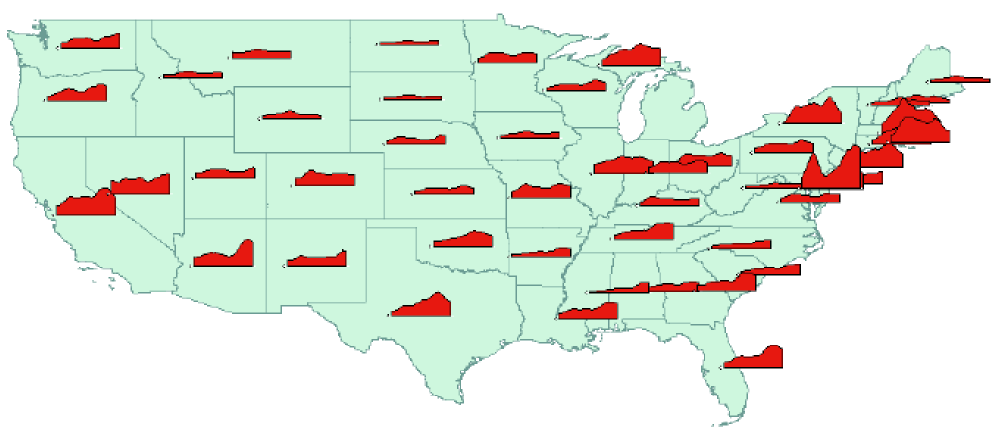


Figure 3.7 Cartographic representation of the spatial distribution of the burglary rates in the USA. Image courtesy of G. Andrienko.

Shanbhag et al. presented three techniques that modify choropleth maps in order to display temporal attribute data. They did not color each district area in the map uniformly but partitioned it into several regions, each representing one point of time in the data. Their first technique builds on a cyclical, clock-like metaphor for perceiving time and partitions the area polygon into wedges. The second technique draws from the metaphor of annual rings of a tree trunk and assigns time points to `rings' of the polygon. Finally, they suggested time slices for a linear perception of time, i.e., polygons were partitioned into vertical slices. Using any of the three techniques, temporal trends in each area could be detected by observing the variation (e.g., brightness) of the different regions of the district area. However, for an effective visualization the number of simultaneously displayed time points must be limited with respect to the size of the polygons in order to avoid clutter.

For detecting periodic temporal patterns, Hewagamage et al. suggested using spirals to depict time in a 3D representation similar to MacEachren's space-time cube. They used this technique to display events, i.e., data with existential changes like the sight of a bird at a specific time and place. Depending on the semantics of the data, events often show some periodic appearance patterns, e.g., bird migration depends on the season and observed birds may rest at a certain place every year. Hewagamage et al. took a linear time line and coiled it such that one loop of the resulting three-dimensional spiral corresponded to a user-specified time interval, e.g., a year or month. At each location of interest such a spiral was positioned, and the events at that position were placed as small icons along the spiral. Thus, points in time whose temporal distance was a multiple of the selected time period were vertically aligned on the spirals. Since parts of the spirals were occluded the display needed to have interactive controls for zooming and panning, as well as for changing the period of the spirals.

3.2.6 Interactive User Interfaces

Interaction is paramount in geo-visualization, especially for visual exploration. The communication aspect of geo-visualization is also shifting towards higher levels of interaction. Dykes introduced guided discovery as a communication task. For example, consider a student who is learning by interactively (re-)discovering known relationships in a data set. Dykes saw interaction as the key defining characteristic of visualization early in the 21st century and according to MacEachren interaction is a key factor distinguishing geo-visualization from traditional cartography. He continued that geo-visualization is an active process in which an individual engages in sorting, highlighting, filtering, and otherwise transforming data in a search for patterns and relationships. In this sense, the present section focuses on common interaction principles and interactive user interfaces in current geo-visualization systems.

Map use in traditional cartography, too, has been benefiting from interaction albeit to a far lesser degree than current geo-visualization. For example, drawing with a pencil or using colored pins to mark spots on a paper map is a way of interactively changing the map. For centuries using these and similar techniques were the only way of exploring geospatial data. However, such manual interaction techniques are time consuming and often give only very limited insight into the data. Generating individual maps on demand had to be done manually by cartographers and at a prohibitive cost for most explorative purposes. This has changed drastically in computer-based geo-visualization.

Buja et al. introduced a taxonomy for general interactive multivariate data visualization. In the following, interactive visualization techniques are grouped using this taxonomy. Their two main classes are focusing individual views and linking multiple views.

Focusing By focusing Buja et al. mean any interactive modification that selects what to see in a single display and how it is seen. They compared focusing to operating a camera: choosing a perspective, deciding about magnification and detail, etc. The individual visualization methods introduced in the previous sections all profit from one or the other form of interaction--as static displays most of them are of very limited use. Two- and three-dimensional maps usually come along with a set of navigational controls to move within the map space either by scrolling, shifting, or rotating a map or by walking or flying through a virtual 3D environment. Further controls allow users to zoom in or out of the map and thus to increase or decrease the level of detail and the amount of generalization of the underlying geographic features (rivers, cities, lakes, etc.). Once a perspective on the map is chosen, its appearance can be modified further, depending on the type of map. The display of geographic background features such as rivers, mountains, roads, etc. can be switched on or off. Depending on the user's task such surface details may be informative or distracting. In Figure 3.8, a screenshot of the satellite image viewer Google Earth is shown. With the navigational controls in the bottom of the figure the user can move around and change the scale. The check boxes on the left can be used to hide or show information like roads or borders.

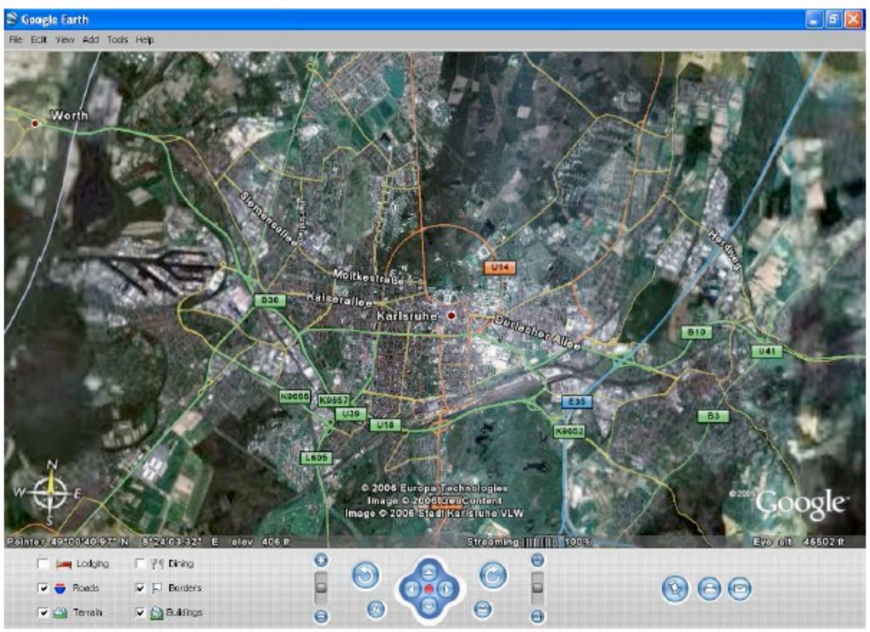


Figure 3.8 Screenshot of Google Earth.

The next mode of interaction concerns the way that the actual data items are displayed. For the general case of multivariate data the user must be able to select the attributes of interest and the type of their visualization, such as choropleth map or glyph-based map. Deselected attributes as well as the actual numeric values of selected attributes should be accessible, e.g., as tool tip information when moving the mouse over a symbol or area on the map. Dynamic isolation was used by Andrienko and Andrienko to highlight a subset of data items in maps that use different symbols to depict different classes of data items, e.g., to select only coal mines in a map of natural resources. Spatial patterns may be cluttered in a map of all resources while they become evident once dynamically isolated.

Modification of the color scheme in a choropleth map is another important aspect. Assigning colors to classes or color ranges to numerical ranges is a very simple way of interacting with choropleth maps. While traditional paper maps use well-established color schemes that are suitable for the majority of map users these schemes may still not be optimal for an individual analyst and hence should be adaptable. In an unclassed map, i.e., numbers are mapped to color shades directly, Andrienko and Andrienko used dynamic visual comparison to display attribute values with respect to a number N in the value range. To this end they applied a diverging color scheme which maps N to a neutral color, e.g., white, values lower than N to shades of red, and values higher than N to shades of green. The greater the difference the darker the color. This is a way to visually split the areas on a choropleth map. The reference value N can be selected by clicking on an object in the map to select its value or by using a slider unit representing the value and color range. Statistic values like mean, median, or quantiles are typically used as reference value. In classed maps the value range of an attribute is partitioned into a set of intervals and a distinct color is assigned to each interval. Here dynamic classification was applied, i.e., interactive modification of the underlying classes by shifting interval boundaries or changing the number of classes.

Finally, the mapping of a numeric value range to color shades in a choropleth map or to height in bar maps can be modified. This mapping is usually a linear function and hence outliers in the value distribution may cause a poor color resolution for the bulk of values. To reduce this effect Andrienko and Andrienko used dynamic focusing to map the full scale of colors to a subrange of values with the outliers removed. Values to the left or right of the subrange are simply colored by the minimum or maximum color value. Using non-linear functions to map values to colors, as it is common for statistical graphics, is undesirable for choropleth maps as visual analysis greatly depends on the immediate expression. Similar colors would no longer necessarily reflect numeric closeness of values. For scatter plots, however, non-linear transformations (e.g., logarithmic scale) are acceptable if the goal is to find a functional relationship between two attributes.

As mentioned in Section 3.2.3, visual tools like the parallel coordinate plot (PCP) or glyph-based techniques depend on an order of the attributes. Ordering the axes in a PCP or star plot and assigning facial features to attributes in Chernoff faces has an important effect on the geometric shape of PCP profiles and glyphs. Giving users the opportunity to change these properties allows them to see the same data from multiple perspectives. Concerning outliers and visual resolution, similar arguments and solutions as for mapping a value range to a color range in choropleth maps apply for the scale of PCP axes.

For animated maps there is a need to give control of the animation to the user. Even simple animations have VCR-like interfaces to control the animation speed, to pause the animation, or to select individual scenes. Harrower described these abilities to navigate in time as equivalent to navigating in space, i.e., zooming and panning on static maps. Tools to navigate in time vary from linear time lines with sliders, which the user can move to select points or intervals in time, to more sophisticated devices like the time wheel. The time wheel supports a cyclic view of time and consists of three concentric circles, the innermost representing hours, the next one days, and the outer circle represents months. In each circle, the respective time units can be marked and thus it enables to select recurring periods in time, e.g., the hours between 7 and 9 p.m. on the first five days of January, April, and May.

Linking and Brushing The full potential of interaction in geo-visualization lies in linking multiple views of the same data on the screen which is the defining criterion of the second class of Buja et al.'s taxonomy of interactive multivariate visualization. Linking basically means simultaneous highlighting of data items in multiple views. It is usually combined with brushing, i.e., selecting display objects by pointing on them or encircling them on the screen. Brushing was originally used for scatter-plot matrices, where points highlighted in one scatter plot are simultaneously highlighted in the other plots of the matrix to evaluate for example whether a relationship in two attributes also holds for other pairs of attributes. Monmonier extended this idea as geographic brushing and links a map with scatter plots or other visual data mining tools. An example of linking and brushing is shown in Figure 3.9, where an outlier point in the scatter plot of per-capita income and percentage of poor is selected. The map shows the geographic location of the corresponding county New York, which has the highest average income in the USA. The PCP in the bottom of the screenshot allows to evaluate the remaining statistical attributes of New York.

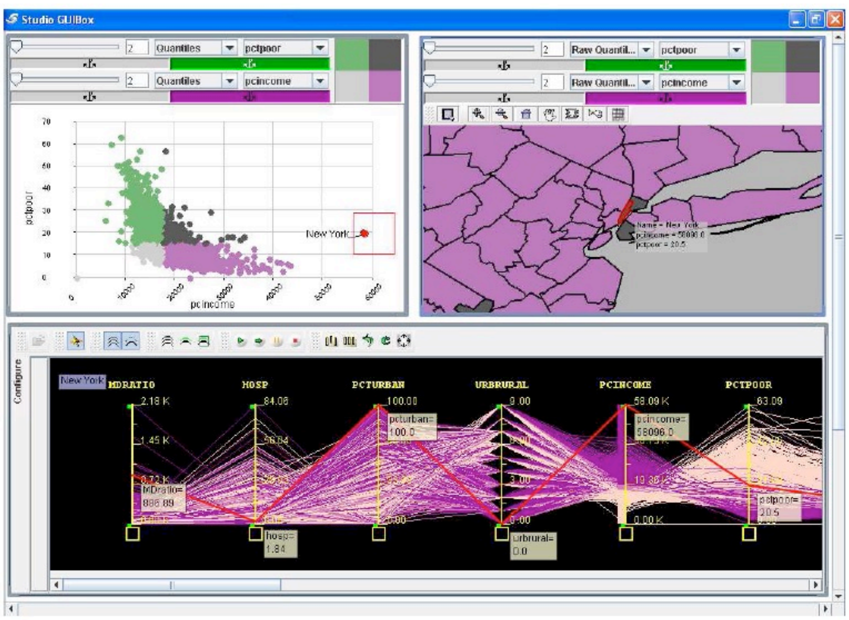


Fig. 3.9 Screenshot of GeoVISTA Studio showing linking and brushing between a scatter plot, a map view, and a PCP. The red item marked in the scatter plot is simultaneously highlighted in the other views.

Different views should be linked in a geo-visualization system. Highlighting a point cluster in a scatter plot or a cluster of PCP profiles thus shows the spatial pattern of the corresponding objects in the associated map view. If there is, for example, a set of objects that are visually similar in a PCP the analyst might ask whether these object are located in the same region on the map or whether they are spread over the whole country but only in rural areas etc. Consequently, outliers that deviate from a general pattern can be located and examined more thoroughly. Similarly, one can select spatial object clusters on the map and subsequently analyze their behavior in attribute space using the remaining views. If the map is a choropleth map then linking and brushing can be done automatically using the same color scheme for the representations of objects in all views. Then, the analyst can mentally connect multiple views because the classes of the choropleth map are marked identically in all views. Changing the class assignment in the choropleth map immediately updates the linked displays. Conversely a scatter plot can be used to define new class boundaries for the map. A very simple example is the cross map. It divides the value ranges of both attributes in a scatter plot into a lower and upper range thus defining four classes, each of which is represented with a distinct color on the map. Changing the class assignment is as simple as moving the class-break point, i.e., the center of the cross separating the four classes in the scatter plot.

The possibilities of interacting with multiple linked views through highlighting and brushing are numerous. The number, type, and arrangement of the views depend on the specific geo-visualization task, the individual user, and the available space on the screen. In any case, it is linking and brushing that make the use of multiple views more than simply the sum of its parts. The interactive principles introduced in this section all concern a core aspect of geo-visualization: stimulate visual thinking by presenting the data in different ways and from a large number of perspectives. This is a key aspect of avoiding both Type I (seeing-wrong) and Type II (not-seeing) errors as false patterns are unlikely to be visible from many perspectives and patterns hard to see in a single view are more likely to be discovered in other views.

3.2.7 Combining Visual and Computational Exploration

The visualization and interaction techniques described so far seem most successful for data exploration tasks in small and medium-sized data sets, i.e., up to a few hundred items and a few tens of attributes. Geospatial data sets in practice, however, are continuously growing in size and are often characterized by a high number of dimensions as a report by the US National Research Council observed. In such large data volumes with high-dimensional attributes human vision cannot be successful in isolating patterns. Visualizing large data sets result in maps and other displays that are cluttered with overlapping items and small symbols such that properties of data items are hardly visible. Zooming is no remedy to these problems: it does help to avoid information overload and magnifies small symbols but at the cost of losing the overall view of the data which is just as important.

On the other hand, computational methods have been developed in areas like machine learning and data mining that can analyze large data volumes and automatically extract knowledge. But data mining methods have limited pattern interpretation abilities. They are susceptible to missing patterns with unusual, non-linear shapes. Interpreting potential patterns is also extremely difficult for computational methods which do not have the domain knowledge of a human expert.

With their strengths and weaknesses, computational methods and visual approaches complement each other. The integration of both approaches to combine their advantages promises further advances in the exploration of geospatial data. MacEachren and Kraak reported integrating advantages of computational and visual approaches for knowledge discovery as one of the four primary themes on the ICA research agenda. Data exploration with tools that integrate both approaches is an iterative process. Results of the initial computational analysis are displayed graphically for further examination by an analyst. Using visual tools the patterns detected automatically need to be explored and interpreted. Questioning the results of a single run is very important. The user must be able to verify patterns and their stability by interactively changing the parameters of the data mining methods.

Examples of this integrative approach comprise the use of self-organizing maps (SOM) as a form of neural networks by Guo et al. and Koua and Kraak as well as applying k-means clustering by Andrienko and Andrienko. Both methods are used for detecting clusters in a data set. A cluster denotes a subset of data items that are similar to each other and different from items in other clusters. This implies the need for a similarity measure on the attribute space, e.g., based on the Euclidean distance. In an explorative environment, where the goal is to discover unknown relationships, patterns have no predefined shape. Therefore, it is important to apply computational methods that do not impose a-priori hypotheses about the shape of patterns and instead let the data speak for themselves. Kohonen's SOM is an example of such a method. The basic idea is to project high-dimensional data to a given grid of nodes on a two-dimensional surface such that (potentially non-linear) similarities are preserved and transformed into spatial proximity: data items within the same node are most similar, and the similarity between an item in one node and items in other nodes decreases with increasing grid distances to these nodes.

A SOM for geospatial data by itself is missing some important information. First, the geographic locations of clustered items cannot be extracted from the SOM view and second, the attribute values representative for a node cannot be displayed. Guo et al. solved this problem by linking a view of the SOM with a map and a PCP view which provided the missing information. A screenshot of their tool is shown in Figure 3.10. The map depicts 156 counties in the US states Kentucky, Pennsylvania, and West Virginia. They used an appropriate (user-adjustable) 2D color scheme to color SOM nodes such that nearby nodes have similar colors. This can be seen in the upper right window in Figure 3.10. The hexagons in this windows colored with shades of grey depict the distance in attribute space between adjacent SOM nodes. In the choropleth map view, each county is colored according to the SOM node it belongs to. This enables the user to compare proximity in attribute space with proximity in geographic space in a single view. A PCP is used to display the summarized attribute values of all SOM nodes. Summarizing data items in the same node avoids overplotting in the PCP since the number of SOM nodes (e.g., 10 £ 10) is usually much smaller than the number of data items. The profile of each node is again colored identically and the line thickness is adapted to the number of items in that node. Brushing and linking of the three views are supported and the user can select either counties on the map, profiles in the PCP, or nodes in the SOM. The corresponding objects in the other displays are immediately highlighted.

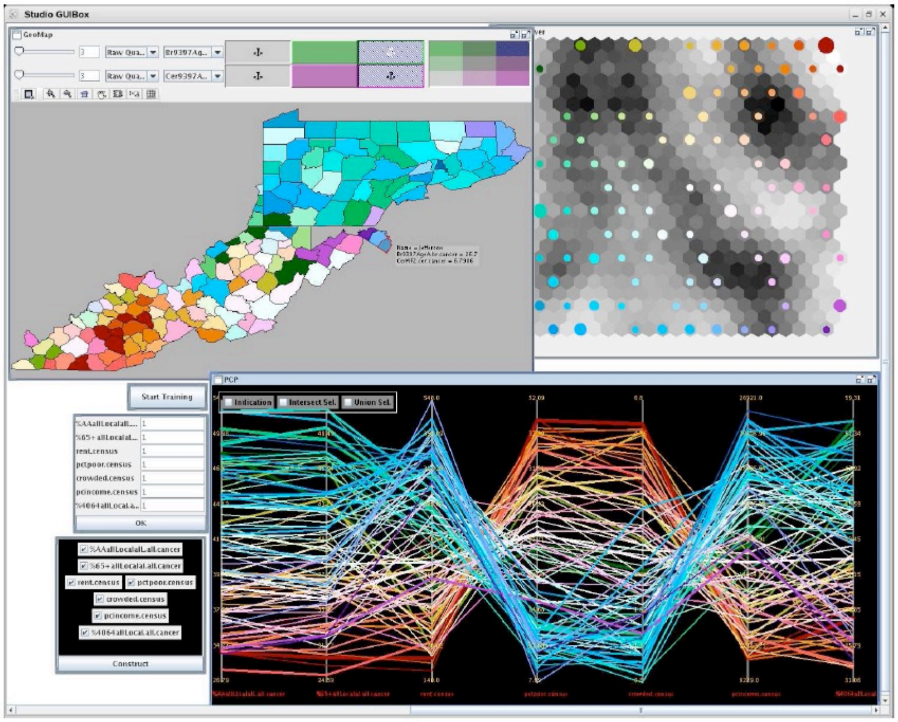


Fig. 3.10 Visualization of a cancer data set with a choropleth map, a parallel co-ordinate plot, and a SOM view. Screenshot of an application built with GeoVISTA Studio.

The example of Guo et al. shows that the integration of a computational data-mining technique into a geo-visualization system allows the exploration of large geospatial data sets. The visual information load is reduced by automatically clustering the data and only displaying summary information while details are still available on demand. Users can explore the data and generate hypotheses by interacting with the system and by bringing in their expertise.

3.3 Geovisualization Tools

In the previous section, a wide range of methods and techniques used in geo-visualization have been presented. However, the challenge for designers of geo-visualization tools is to put these methods together effectively in order to help users solving their respective tasks. It has become clear that both users and tasks vary considerably. Hence, there cannot be a universal tool that fits all users and tasks. Instead, tools need to be flexible and easily configurable by users in order to be applicable to more than just a tightly defined problem. Users should be able to select freely from the methods discussed in the previous section and to link multiple views. In the following, five examples of geo-visualization systems are briefly described to show how multiple views are combined and linked in practice.

ArcGIS or ArcGIS Pro. ArcGIS or ArcGIS Pro is a commercial tool for visualizing and analyzing geographic data. It is one of the world's leading geographic information systems. ArcGIS or ArcGIS Pro offers a range of methods for creating customized thematic maps and analyzing spatial data. It is primarily used in administrative and industrial settings, e.g., for emergency planning, site planning, or marketing.

XGobi. XGobi, initially released in 1996, is a general data visualization system supporting linked scatter-plot matrices and parallel coordinate plots. XGobi is a freely available tool, and it can also be linked to ArcGIS.

Cartographic Data Visualizer. The Cartographic Data Visualizer (CDV) integrates the mapping and abstract data visualization components into a single application. Its latest release is from 2000, and it is freely available on the Internet. CDV offers dynamic choropleth maps, population cartograms, and statistical graphics like PCPs. Graphic symbols serve as interface elements to access detailed information or to select subsets of items. The views are linked such that highlighting items in one view is passed on to the other views.

CommonGIS. CommonGIS is another integrated geo-visualization environment developed at Fraunhofer AIS which is in practical use in academic, administrative, and commercial settings. It combines a multitude of geo-visualization techniques from (multiple) dynamic choropleth maps, optionally combined with bar plots and pie plots, over animated maps, time-series diagrams, and space-time cubes to multivariate visualizations like PCPs or scatter plots, see Figure 3.11. The user interface of CommonGIS supports interaction through focusing, brushing, linking of multiple views, and dynamic range selection for attributes. It is also possible to complement the visual data analysis by computational data-mining techniques. One focus in the development of CommonGIS was that it could be used even by users with no expertise in cartography and geosciences. The tool is commercially distributed, but it is free of charge for academic users.

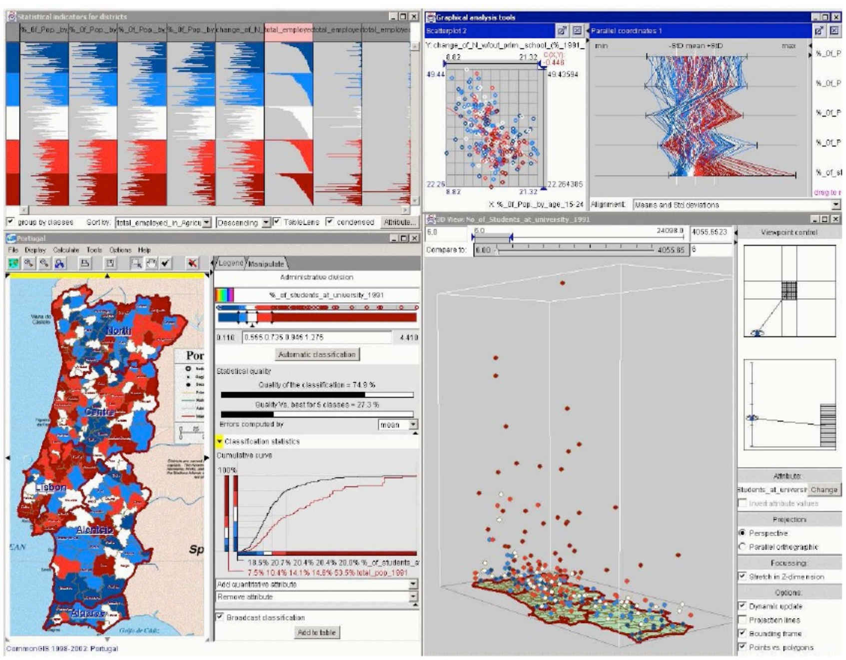


Figure 3.11 Screenshot of CommonGIS. Image courtesy of G. Andrienko.

GeoVISTA Studio. The approach taken by GeoVISTA Studio is different. It is an open source software development environment for geo-visualization rather than a static application. It is a component-oriented system with the goal to integrate a wide range of both computational and visual analysis activities and ultimately to improve geoscience research. Creating custom applications is done via a visual programming user interface that allows to connect different geo-visualization components, provided as Java Beans, according to the desired data and control flow. Visual programming is a key aspect according to Takatsuka and Gahegan because it allows geoscientists with little computational background to rapidly create prototypes when they are searching for useful insight. GeoVISTA Studio comes with a range of standard components, such as choropleth maps, 3D renderers, PCPs, scatter plots, color maps, spreadsheet views, and computational tools as k-means clustering and self-organizing maps. Additionally, Java Beans created by third-party developers can easily be plugged in. Once a specific application or applet has been designed, its deployment over the Internet, for example to students for educational purposes, is supported. Sample applications can be downloaded from the project web site. One example is the SOM-based tool by Guo et al. described in Section 3.2.7 and shown in Figure 3.10 which is built using GeoVISTA Studio.

3.4 Future Trends

How will geo-visualization develop over the next years? In the future, the map are still seen as the primary tool to present, use, interpret, and understand geospatial data. However, it has become apparent in the past that the map can and will take a variety of forms, some of which quite astonishing. The map has evolved from its traditional role as a presentational device to an interactive and highly dynamic interface to access and explore geospatial data.

Another common feature of current geo-visualization tools is that they consist of multiple linked displays, each depicting an alternative representation of the data under examination to stimulate visual thinking. This will certainly remain a central aspect in geo-visualization.

Concerning the tasks and users of geo-visualization tools a transition from explorative, individual use by experts towards the whole range of tasks from exploration to presentation and heterogeneous groups of users is taking place. While individual expert tools will continue to develop, there will be an increasing number of applications designed for the public and disseminated over the web. It is clear that the usability requirements differ significantly between research tools for specialists and public applications for the mass of non-expert users. Human-centered aspects need to play a key role in the design of future tools in order to make them fit the needs of their audience. The goal of universal usability, i.e., creating applications for wide ranges of users from children to senior citizens in different languages and respecting visually disabled users, remains a big challenge.

Advances in hardware technology will also have a strong influence on geo-visualization. On the one hand, large 2D and 3D displays and virtual environments will be used for geo-visualization tasks, especially for collaborative use by a group of experts and decision makers. Usability studies for these new technologies will have to investigate their advantages and disadvantages over traditional visualization methods as well as possible cognitive and social impacts. In particular, the right balance between abstraction and realism needs to be determined for 3D displays. On the other hand, portable devices like PDAs or mobile phones will provide location-based services, e.g., route finding in foreign places which build on some sort of map display. This means additional challenges for application developers in terms of efficient memory and bandwidth use as well as in terms of visualization design for multi-platform, small-size displays.

A last point that is raised by Gahegan is the need for interoperability of geo-visualization components. Following the efforts of the Open Geospatial Consortium to define open standards for geospatial information systems, Gahegan promotes a similar initiative for geo-visualization. Visualization components should define standard interfaces in order to be reused and integrated into a variety of geo-visualization tools. A lack of interoperability means that for developing a new tool too much effort is spent on re-implementing basic functionality and too little effort can be invested in developing new methods.

The coming years will show what directions geo-visualization research will take and how it will influence both research in the earth sciences and related disciplines and our everyday handling of geospatial information.

References:

- Buckley, A. R., Gahegan, M., & Clarke, K. (2004). Geographic visualization. In A research agenda for geographic information science (pp. 313-333). CRC Press.
- Buckley A. 2021. Make Maps People Want to Look At. https://www.esri.com/news/arcuser/0112/make-maps-people-want-to-look-at.html
- Celik A. 2015. Colors, Methods and Mistakes in Spatial Data Visualization. https://medium.com/@acelik/colors-methods-and-mistakes-in-spatial-data-visualization-60d02e7f09fd
- Ildefonso O. 2021. Top Mapping Mistakes. https://digitalfellows.commons.gc.cuny.edu/2021/05/12/top-mapping-mistakes/
- Monmonier, M. 2018. How to lie with maps. University of Chicago Press.
- Nöllenburg, M. (2007). Geographic visualization. In Human-centered visualization environments (pp. 257-294). Springer, Berlin, Heidelberg.
- Patterson, M. 2021. Info The 10 Required Map Elements. https://storymaps.arcgis.com/stories/e050502512ea404f8dc2489695ffbda4