

UC Santa Barbara

GIS Core Curriculum for Technical Programs (1997-1999)

Title

Unit 47: On-Screen Visualization

Permalink

<https://escholarship.org/uc/item/12q7v9xv>

Authors

Unit 47, CCTP
Edsall, Robert M.

Publication Date

1998

Peer reviewed

UNIT 47: ON-SCREEN VISUALIZATION

Written by Robert M. Edsall, Department of Geography, Penn State University

- [Context](#)
 - [Example Application](#)
 - [Learning Objectives](#)
 - [Preparatory Units](#)
 - [Example Application Revisited](#)
 - [Follow-Up Units](#)
 - [Resources](#)
-

Context

Effective interpretation of the results of your analysis -- both by you and by others -- depends to a large extent on the design of the on-screen display. Visualizing information in the form of a map, chart, image, or other form of graphic takes advantage of humans' most perceptive sense -- vision -- to uncover patterns, associations, and processes that would likely be missed if the data were displayed in tabular or some other non-graphic form. However, as Bittenfield (1996) points out, it's this very ability that might lead a viewer of a graphic display to be as easily misled as informed by the contents of the graphic. A slight alteration of the color scheme or classification method of a map, for example, might produce two different graphics that appear to tell two very different stories -- even though they represent the exact same data. Thus it is crucial that GIS users be aware of the power of a graphic image to communicate ideas and of the considerations necessary to design an effective on-screen visual display. This unit will outline some guidelines, developed by cartographers over many decades, which will assist you in the logical design of easily understood GIS display.

Example Application

Guidelines for on-screen visualization should be applied to all situations which require the production of some graphic display of information. In a GIS context, this includes almost every possible application. No one application can incorporate all of the guidelines covered in this unit; the following is chosen because it represents a relatively complex but not

uncommon situation encountered in GIS use.

A GIS specialist has been asked to produce a map showing possible environmental and social impacts of the construction of a new four-mile highway in a rural area. The specialist is faced with a host of choices, completely separate from any geographical analysis of the impacts, concerning how best to visualize the multiple variables of various types on a single display. The data she has been asked to include is the following:

- political boundaries
- results of a local survey concerning the desire for the highways construction
- hydrology
- land use
- wildlife reserves, game lands, and other protected areas
- existing road network
- buildings, including houses, barns, churches, schools, and commercial buildings

This task requires careful selection, simplification, and classification of the data given and creative use of visual variables like shape, texture, location, hue, and lightness to create a cogent, easily understood, and informative graphic. Each choice she makes will impact the overall look of the graphic, altering its visual balance, contrast, and, ultimately, the information communicated by the display.

Learning Objectives

Awareness

The student should be aware, in a general sense, of visualization design concerns and strategies

Competency

The student will:

- design for and construct the appropriate type of display for the data being mapped
- make logical choices in selecting data to include in a display
- manipulate visual variables like color, size, and pattern to your advantage
- understand issues of classification and generalization and how they may effect the message of the display

Mastery

- design effective graphics for multiple variables
 - incorporate logical color schemes for the display of each variable
-

Preparatory Units

Unit 10 - Projecting Data
Unit 11 - Register and conflate data
Unit 23 - Creating maps with CAD

Awareness

[GIS and Visualization](#)

[General Considerations](#)

Competency

[Form of the map](#)

[Selection and Generalization](#)

[Data classification guidelines](#)

[Color use guidelines](#)

[Use of other visual variables](#)

[Legend design](#)

[Use of Text on the Display](#)

Mastery

[Form of the map](#)

[Data classification guidelines](#)

[Color use guidelines](#)

[Example Application Revisited](#)

Follow-Up Units

Unit 48 - Designing products for printing
Unit 51 - Preparing Digital Presentations

Resources

The International Cartographic Association's Commission on Visualization web site includes many current theoretical and practical discussions about visualization as it applies to geography and cartography.

Brewer, C. (1994). "Color Use Guidelines for Mapping and Visualization," in *Visualization in Modern Cartography*, A.M.MacEachren and D.R.F.Taylor, eds. London:Pergamon. pp. 123-148.

A comprehensive guide to the logical use of color in cartographic and visualization applications.

Buttenfield, Barbara P. (1996). "Scientific Visualization for Environmental Modeling: Interactive and Proactive Graphics." *GIS and Environmental Modeling: Progress and Research Issues*, Goodchild, M.F., et al., eds. Fort Collins, CO: GIS World Books. pp. 463-468.

This article and several others in this volume discuss visualization strategies in the context of GIS.

Hearnshaw, H. and Unwin, D. eds. (1994). *Visualization in Geographical Information Systems*, New York: Wiley & Sons. 243 pp.

A set of papers on visualziation and GIS written by professors and other professionals in the field. Good for in-depth consideration of a variety of different issues in GIS.

Imhof, E. (1975). "Positioning Names on Maps," *The American Cartographer*, 2(2): 128-44.

MacEachren, A.M. (1994). *Some Truth with Maps: A Primer on Symbolization and Design*. Washington: AAG. 129 pp.

A very readable primer on cartographic techniques, tips, and pitfalls, with many illustrations.

Monomier,M. (1991). *How to Lie with Maps*. Chicago: U. of Chicago. 176 pp.

A fascinating discussion, with many graphic examples, of the ways in which cartography can be manipulated and manipulative. Monmonier has also written other books and articles on the subject which are also worth investigating.

Robinson, Arthur H., Joel L. Morrison, Phillip C. Muehrcke, A. Jon Kimmerling, and Stephen C. Guphill (1995). *Elements of Cartography*, sixth edition. New York: Wiley & Sons. 674 pp.

A standard introductory cartography texts used in American universities.

Created: May 14, 1997. Last updated: March 11, 1999.

GIS and Visualization

Visualizing large amounts of information interactively is one of the most attractive and useful capabilities of GIS. High-powered computers can alter any element of the display "on the fly," changing not only the look of the graphic image but also its interpretation. This ability to create multiple perspectives -- both literally and figuratively -- enhances a viewer's perceptive abilities to understand the phenomenon being studied like never before.

Visualizing data using current computing technology and interactive GIS has many advantages over doing so using traditional paper maps. The following is a partial list, one that grows with each new version of software and each new advance in hardware.

- GIS is fully interactive. Adding new fields of data, taking them off, changing the color scheme or form of the map, adding text, moving symbols, and a host of other capabilities give a user tremendous flexibility and power.
- GIS displays are zoomable and pannable. Moving around in the display offers a user new perspectives, greater (or less) detail, and new insight.
- Users can take advantage of computationally intensive functions such as "draping" a perspective view over a surface (like a digital elevation model) or creating the impression of three dimensions on a 2D display (the computer screen) using complex rendering and shading algorithms.

Many of the tips and guidelines outlined in this unit were developed by cartographers over many decades for the design of *paper* maps. They are all relevant to the design of interactive displays within GIS. In fact, with the vast array of choices facing a GIS display designer and with the luxury of complete interactivity, awareness of the power of the visual image to persuade a viewer has never been more important.

General Considerations

In the most general sense, any designer of on-screen displays should understand and design around some basic concepts. Any strategy for visualizing should begin with the following guidelines.

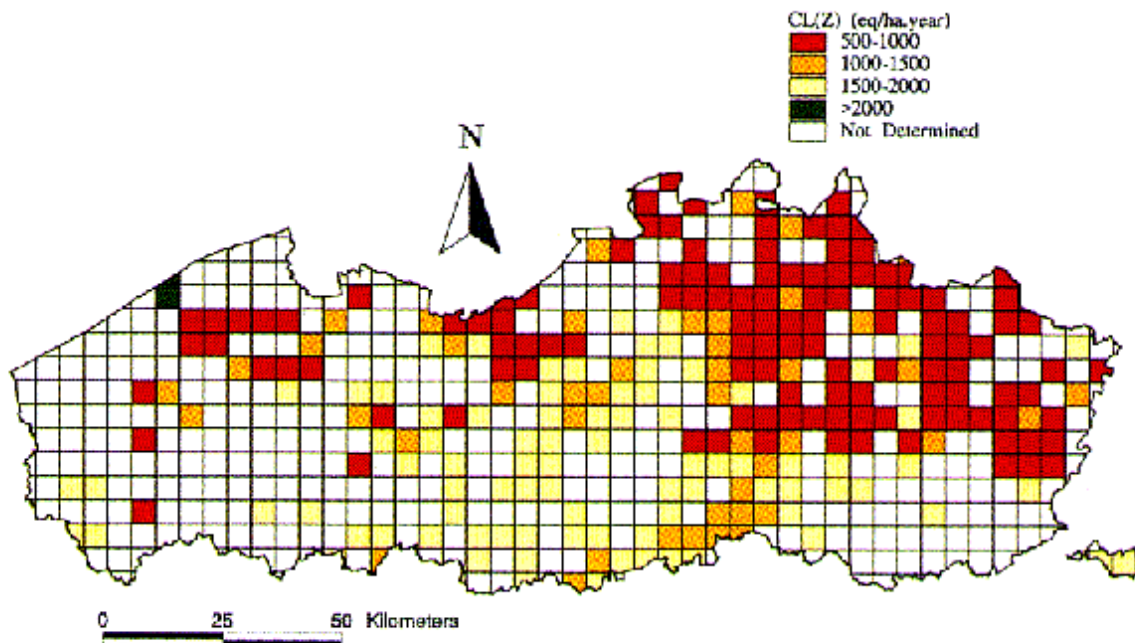
Know your audience. Their level of expertise will determine your design strategy. If this is a display for the general public, for example, a high degree of selection and generalization may be required, whereas expert viewers may desire a large amount of detail and depth.

Get your point across. Does the dominant visual feature on your map relate to the purpose of the map? Certain elements of a graphic object, like its hue, location, size, and contrast, tend to be more easily and more intuitively interpreted by human vision. Make sure that the dominant object on the display supports or at least represents the primary theme of interest of the graphic.

Avoid unnecessary bells and whistles. Including a visually striking digital terrain model on a map of voting results may make the display look better, but it in fact may also serve to distract the viewer to the extent that the main point of the display is lost. More information on a map, thus, is not always better.

Strive for visual balance. Use the space of your display to its fullest by avoiding large empty areas. However, don't crowd the image to the point of being difficult to interpret or even look at. Use a blend of text and graphic elements, and center your main image and its title to emphasize its importance.

Critical loads for acidity



from Craenen, et al., in *GIS: Our Common Language*, ESRI Map Book, Volume 12, p. 42.

The map above achieves these objectives to some extent. The map is visually balanced, with informative but not distracting or confusing text, a useable legend, and straightforward graphics. The message of the map, however, is unclear due to the color scheme used. The eye is drawn to the darkest (red) values, but we notice that this color represents the *lowest* of the attribute values. It is likely that low values of this theme are, in fact, the most important (showing regions of a low threshold), however, if this is in fact the case, the highest values (>2000) should not be represented by an even darker hue (black). The data is ordered, yet the color scheme used does not imply order.

Form of the map

The form of your data dictates the form of the map. Raw count data, for example, should not be represented using a [choropleth](#) map, since this will yield a display that puts visual emphasis on the largest area units of the map. A choropleth map of the number of cars in southern California, for example, would be dominated by the count of cars in massive (but sparsely populated) San Bernadino County, and would mislead a viewer to perceive that there are not very many cars in southern California. A more appropriate map form for this data would be a [proportional symbol map](#), with circles or other icons of sizes in proportion to their attribute value (count of cars) centered in each spatial unit (county).

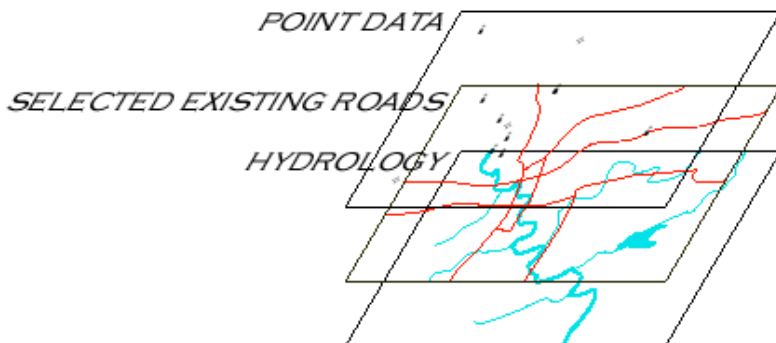
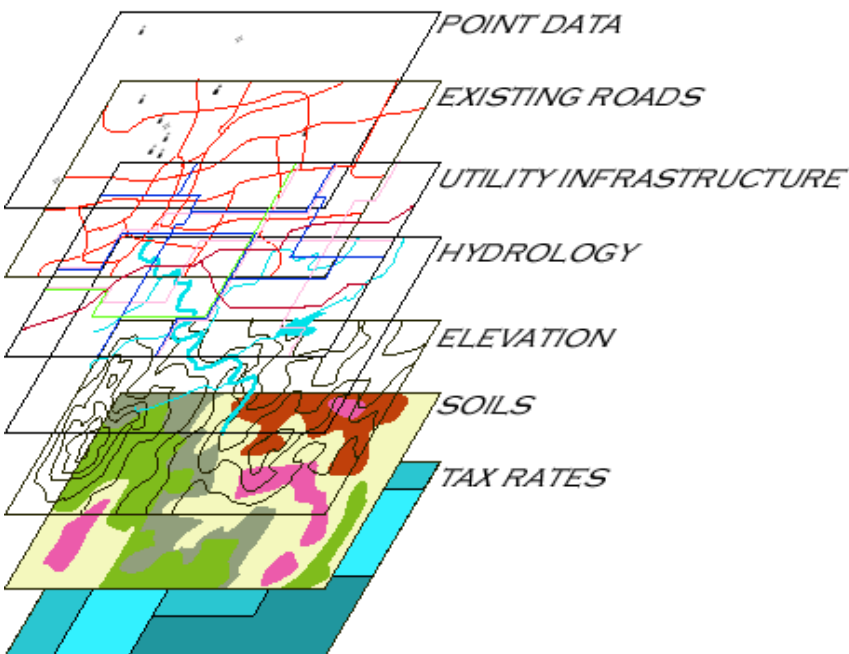
This distinction is illustrated in the [map of voting results from the Czech Republic](#). Raw data, in this case total number of votes cast in each district, is represented using proportional symbols, in this case, circles. Ratio data, in this case percent of eligible voters who voted, is shown using a choropleth map underneath the proportional symbols.

[other map forms](#)

Selection

A vital task in creating effective graphics for visually relating ideas is the filtering of vast fields of data into a units that are simple enough to be understood quickly and intuitively but detailed enough to convey useful information. Often, creators of visual displays are too ambitious with the amount or detail of data that they attempt to present on screen. Make sure that you select and generalize if necessary.

Recall that, in all visualization applications, one of the primary considerations is the audience for whom the display is being created. To this end, a designer should carefully *select* those themes or variables that are most appropriate for the communication of ideas. If, as in many cases, the audience a single GIS specialist who is exploring data privately, searching for connections, trends, and patterns, it is likely that the display of highly detailed information is most important. Even in this case, however, proper visualization techniques as outlined in this unit will make data exploration faster and more accurate. However, for public presentation, only those data that are immediately relevant and quickly understood should be included.



An example of this type of selection is the display of the results of a study of potential development sites for a new housing complex. A group of GIS specialists may have the street and hydrology network, with labels, an elevation model, and other layers of information on hand to make a decision. When exploring this data, they might create a complex multivariate map with symbolization techniques known to them alone. However, when presenting this data to a planning commission or contractor, they should limit their display to a small number of variables -- like surrounding streets, streams, parks, and schools.

Generalization

Such refinement may be taken a step further by generalizing the data both visually and mathematically. Visual generalization might involve, for example, including only those streets that give a viewer reference to the location of the housing complex, or perhaps labeling only the major streets in the vicinity. Other visual generalization techniques involve aggregating objects as scale is decreased (such as transforming 10,000 buildings into an "urban" area then into a dot as the geographical region of the display is increased), or including only those features with an attribute greater than some threshold (such as displaying only those airports with runways over 3000 feet).

Mathematical generalization techniques vary from simple to extraordinarily complex. In most GIS applications, linework generalization is the most important application of these techniques. It is seldom necessary in practical situations to carry through the most detailed linework of coastlines, rivers, or highly irregular political boundaries. As mentioned, such detail may be not only unnecessary but also distracting. In vector GIS, this linework is stored and displayed as an exhaustive series of line segments connected at points. Selectively eliminating, for example, three of every four of these points creates a smaller database and a less complicated display while still maintaining enough visual information for recognition and understanding.

Data Classification

If you are developing a [choropleth map](#) of ordered data, one of the first decisions to be made deals with classification: which values should be associated with each color. In other words, which units should be in the lowest class, which units should be in the highest class, and how the rest of the units should be distributed among the remaining classes.

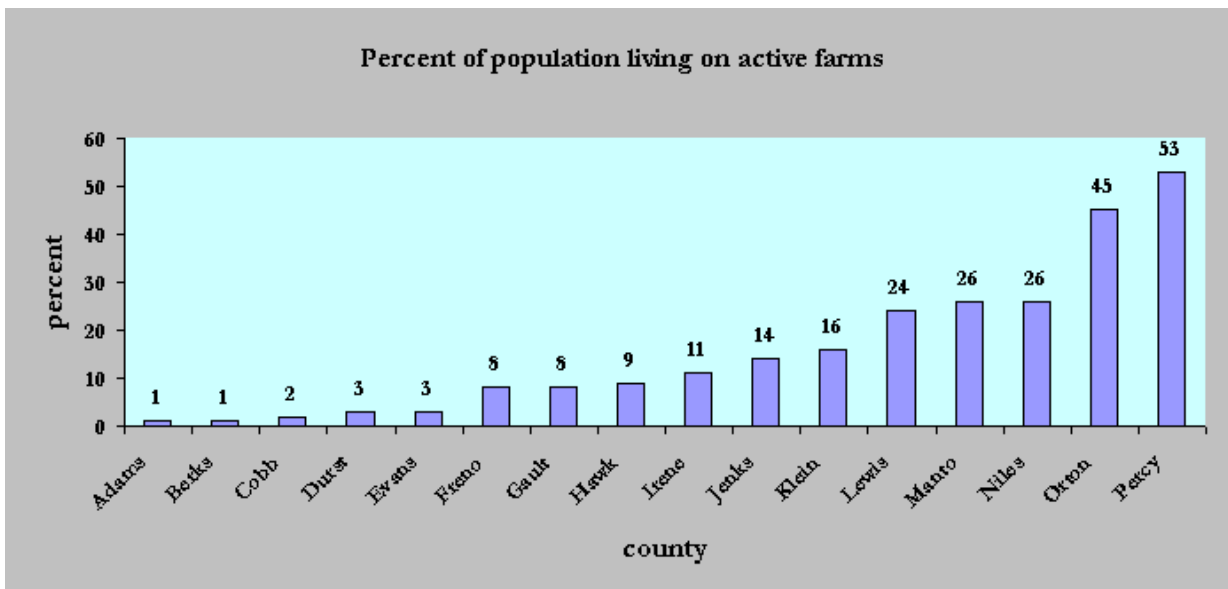
Among the many choices made by analysts and designers, data classification decisions might among the most important, but also the most difficult to understand. A GIS specialist must make choices about not only how many different classes that the data should be categorized into, but what the value ranges of those classes should be. A slight adjustment of the "breaks" in the value ranges of ordered data, for example, might alter the map significantly and reveal trends that were not detected previously (or are not in fact there).

In this section, two of the most common "default" methods of classifying data are presented. These are two of many choices available in recent versions of [ArcView](#), and a designer should be aware of the differences among all of the methods. Each has advantages and disadvantages.

Quantiles. This method classifies data into a certain number of categories with an equal number of units in each category.

Equal Intervals. This method sets the value ranges in each category equal in size. The entire range of data values (*max - min*) is divided equally into however many categories have been chosen.

These methods are illustrated using the following data:



The data above are classified in Quantile and Equal Interval schemes in the table below. The first letter of each county is indicated in the "counties" column, and the corresponding data values (the height of the bar in the bar graph) is listed in the "ranges" column, representing value ranges for each of the four classes.

--	--	--

	Counties	Ranges	Counties	Ranges
Class 1	ABCD	1-3	ABCDEFGHI	1-13
Class 2	EFGH	3-10	JKLMN	14-26
Class 3	IJKL	10-25	-	27-39
Class 4	MNOP	25-52	OP	40-52

Notice right away the possible pitfalls of these methods. With a four-category quantile classification, there are an equal number of counties in each class, but note that Durst and Evans Counties, though they have identical attribute values, are placed in different classes. In addition, Manto and Niles Counties are placed with the extreme Orton and Percy Counties, rather than with Lewis County, to which they are much more similar. This method, in this case, leads to a misleading visualization.

Switching to a four-category equal interval method, the most obvious problem is that only three of the four classes actually contain data points. The ranges of each class (13%) are the same, but because this data is skewed (has a few data points that are very different from the rest), no county's attribute value actually falls into the third class. This reduces visualization effectiveness by effectively eliminating one fill color.

[Look how different](#) these maps -- of identical data -- appear, depending on the classification scheme used.

[More effective methods](#) for visualizing this data exist.

Why would anyone use them, if they are so limited?

The two classification schemes above are the most easily computed and one or the other is usually the default classification in most GIS. These methods *are* adequate to display data that varies linearly, that is, data with no outliers that tend to skew the mean of the data far from the median. A chart of such data like the one above would contain columns that increased in height equally from left to right; the "steps" between the columns would be the same.

All ordered data should be examined to determine whether or not altering the classification method has a significant impact on the display. If so, a method other than the two discussed in this section should be used for the most accurate and informative visual portrayal of the information.

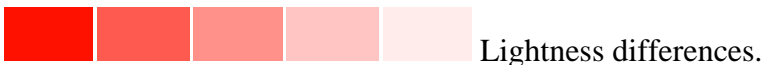
General principles of color use in visualization applications

One of the most powerful ways to stretch the capabilities of on-screen displays is by taking advantage of the extraordinary ability of humans to use color information to aid in visualizing data. Using color logically requires a basic understanding of the three "dimensions" of color -- hue, lightness, and saturation.

Hue. Color can be described by its wavelength; red at one end of the visible spectrum, violet on the other. A change in wavelength of visible light is manifested by a change in hue. "Red," "green," and "violet" are all hues.



Lightness. Color can also be described by its shade, or its degree of lightness. Pink-white, pink, rose, red, and maroon are all composed of the same hue (wavelength) but have different lightness. This dimension of color can also be called "value," or "brightness."

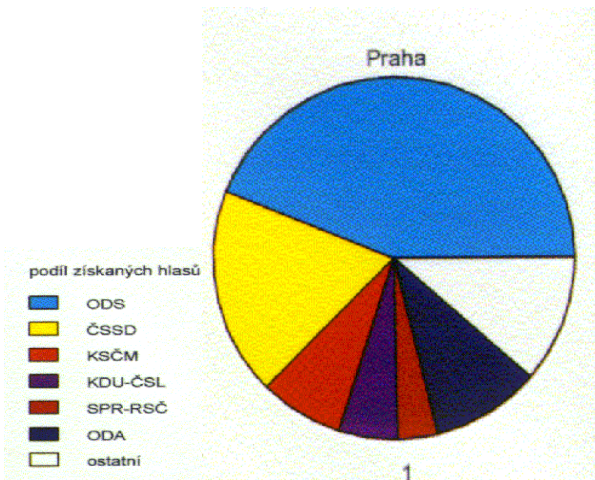


Saturation. A third dimension of color describes how "pure" the color is. If a color is made up of only one hue, it is called "saturated." The color of monochromatic light, like that which comes from a laser beam or neon sign, is completely saturated. However, if you imagine mixing together every glob of acrylic paint on an artist's palette, the result is a *desaturated* color, one made of many hues. Saturation could be called the "grayness" of the color.



As mentioned, color on the whole is one of the most powerful of the so-called "visual variables" in visualization applications. Differences in color can be used in limitless ways to differentiate objects or areas on a display. The following are some general guidelines for the logical use of color in an on-screen display.

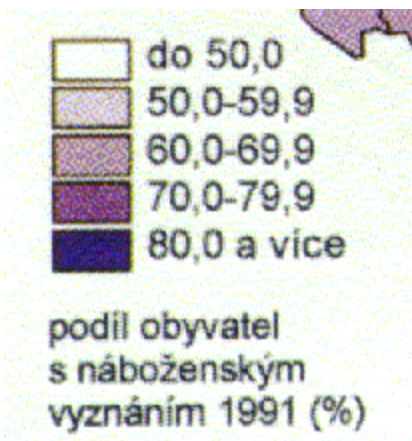
Use of hue in visualization. Hue differences are best used to differentiate between two items that have no inherent order; that is, objects that differ in *kind (nominal)* and not *amount (ordinal)* are best represented by varying hue. Geologic maps use hue to a great extent to differentiate between soil or rock types. Land-use types, political party affiliations (as in the example below), and area code zones all represent nominal differences and should also be differentiated by a nominal visual variable like hue. Theoretically, hue differences imply no inherent order ("green" does not in itself imply "less" or "more" than blue). However, it has been argued -- successfully -- that a "spectral" color scheme, where blue and indigo represent small values (like cold temperatures) and red represents high values (warm temperatures), is an intuitive ordered use of hue.



Use of lightness in visualization. Unlike hue, differences in lightness do imply order. It has been shown that darker colors indicate higher values of a variable. Lightness variation, thus, is one of the most intuitive and powerful ways to represent ordered data, like rainfall totals, voter turnout, or traffic density.

It is generally accepted that the human eye can successfully discern as many as seven different hues and seven different lightnesses on a single display. If your display requires more classes than seven, you should exploit other visual variables, like texture, density, or shape differences.

Tip. In [ArcView](#), select two colors for the outermost classes and then use Ramp Color to create intermediate colors in the intermediate classes.



These two examples come from the well-designed map of [Czech voting results](#) made by ARCDATA PRAGUE.

Use of saturation in visualization. Saturation is the least understood of the three dimensions of color. Saturation differences seem to be the least discernable by the eye, and thus this dimension should be used sparingly. Saturation has been best used to represent binary data --

data that can have one of two values, like public/private ownership of land, or attributes that either have or haven't been field-checked.

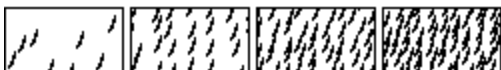
Other important visual variables

Shape. Objects with different shapes are presumed to be different in kind. A star as a symbol for a capital city indicates that it is different, not necessarily larger or smaller, than a city represented by a circle or a dot. A single-dashed line indicates a border of a different kind than a solid line or one that is double-dashed.

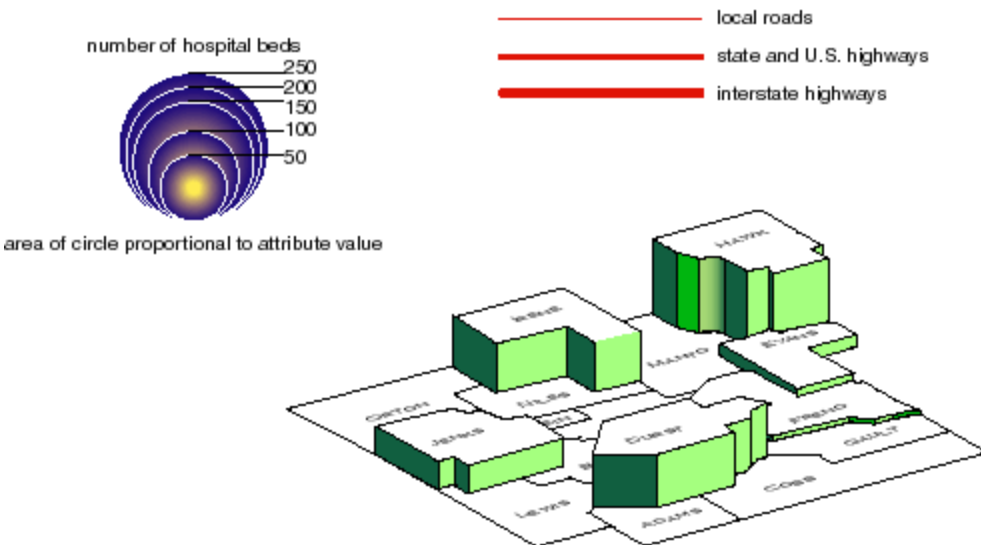
Texture. Like shape, a texture or pattern difference, in general, implies no inherent order. Overlaying a semi-transparent textured pattern over a solid color is an effective way of visualizing multiple area variables like forest cover and ownership. Textures include, but are certainly not limited to, cross-hatching, rows of dots or other shapes, and "brick" patterns. The swamp icon on U.S.G.S. topographical maps is an example of the use of texture to represent an area feature.



Density. The density of a texture can represent ordered data to a limited extent. A semi-transparent overlay of a dot pattern that varies in density can be used to effectively represent such information as levels of an airborne toxin or relative uncertainty of the data "underneath" the pattern.



Size. Not surprisingly, size is a very important visual variable for ordered data. As in [proportional symbol](#) maps, the relative size of icons gives the viewer a qualitative estimate of the relationships of the attributes of the spatial units in the display.



Legend Design

An often-overlooked aspect of any good display design is the usability of the legend. From a graphic design standpoint, items in the legend should appear as similar as possible to the icons to which they refer, in size, shape, color, texture, and all other visual variables which apply. The legend should be compact and well organized; in general, area features are grouped together, as are line and point features. Annotations such as the title, date, creator, affiliation, and source materials should all be included in the legend, which should be self-contained and large enough to fully explain the icons on the display.

Use of Text on the Display

Lettering and typography have a long cartographic history. The guidelines for the use of text on a GIS computer display are no different from those developed for paper maps. Choices of type style, form, size, color, and placement are key in the design of a legible and organized display. Text in visualization applications can be treated as another symbol, one imbued with meaning through language and not through graphical abstraction. As such, choices of size and color should be made according to the guidelines of these visual variables outlined elsewhere in this unit.

Style. Popularly referred to as "font," text style is the geometric design of the letters themselves. Elements of text style include serifs ("finish" strokes like the little horizontal base of the lowercase L), line thickness, relative dimension ratios, and decorative ornaments. A GIS is often capable of producing dozens of fonts, but be aware that overuse of this capability leads to a confusing display. Generally, the fewer the type styles on the display the better. That said, there are some general conventions often used by cartographers. Since style differences are much like [hue differences](#) in that they are best at distinguishing nominal differences, one common convention is to use serified fonts for natural features like rivers, mountain ranges, and islands, and "sans serif" (without serif) fonts for human-built features like cities and political units. Good type style choices for displays are easy to read and do not detract or distract from the overall message of the display. In other words, the viewer should notice the data and not the unusual font of the text.

Form. A type style can take many forms, like "bold," "italic," "light," or "upright." There are few cartographic conventions dealing with form, since the choices are far more limited than those of style. One form convention worth noting, however, is the frequent italic form of labels of water features. Almost universally, rivers, lake, oceans, bays, and other water features are labeled using italics, perhaps to represent flow and fluidity.

This is a serified type style (Garamond).
This is sans serif (Helvetica).

6 point is quite small

12 point is a common size for text.

Water feature labels are often italicized. And blue.

Labels for M O U N T A I N R A N G E S can be stretched.

Size. Be very careful with size. GIS displays have the ability to scale type to any size, from 0.5 point (very small, unreadable, less than 1 mm in height) all the way to 180 point (2 inches tall, the size of a banner headline) or more. Cartographers often overestimate humans' ability to see very small text and our ability to notice small differences in size of text. As a general rule, do not use text smaller than 8 points in a display. In addition, do not use more than three sizes of the same type style and form on the same map, and use size variations of over 25%. That is, use a "small" size of, say, 10 points, a "medium" of 12.5, and a "large" of 16.

Placement. As discussed [elsewhere](#), GIS displays have a significant advantage over paper

maps in that symbolization and design choices can be altered very easily and quickly. This is an important consideration in placing text on the display. As the data displayed changes, text can be moved to accommodate new data while ensuring legibility of the text. There are some general rules, however, that should be observed at all stages of GIS display design. This is a partial list from Robinson (p. 416):

- names should be entirely on land or on water.
- lettering should be oriented to match the orientation structure of the map. On a large-scale map, this means that type should be parallel with the upper and lower edges of the map; in small-scale maps, type should be parallel to the parallels.
- type should not be curved unless it is necessary to do so
- disoriented lettering (that is, not parallel to the map edge or the lines of latitude) should never be set in a straight line but should always have a slight curve.
- names should be letterspaced as little as possible. That is, there should not be wide spaces between the letters in the name. Exceptions to this guideline involve labeling large area features like mountain chains or oceans.



For a detailed list of placement do's and don'ts, see Imhof.

Map forms

Choropleth maps

These maps, where each spatial unit is filled with a uniform color or pattern, are appropriate for data that have been scaled or normalized in some way. For example, density information, expressed as "per unit area," is appropriately represented using a choropleth map. Tax rates, expressed as "per cent," are also a form of ratio data, which is most easily and correctly visualized using choropleth maps. Choropleth maps are also appropriate for indicating differences in kind of an area data, like ownership or use of land or type of forest cover.

The purple fills of the [Czech voting results map](#) represent percent of eligible voters who voted. This representation is (correctly) a choropleth map.

Isopleth maps

Isopleth maps differ from choropleth maps in that the data is not aggregated to a pre-defined unit like a political area or watershed basin. These maps can take two forms:

- Lines of equal attribute value are drawn such that all values on one side are higher than the "isoline" value and all values on the other side are lower, or
- Ranges of similar attribute value are filled with similar colors or patterns.

This type of map is used to represent continuous area data that varies smoothly over space. Temperature, for example, is a phenomenon that should be mapped using isoplething, since temperature exists at every point (is continuous), yet does not change abruptly at any point (like tax rates do as you cross into another political zone). Elevation maps should always be in isopleth form for this reason.

Proportional symbol maps

As the name implies, these maps scale icons (most often circles) according to the data they represent. Proportional symbol maps are not dependent on the size of the spatial unit associated with its attribute. In other words, on a proportional symbol map of Europe, tiny Liechtenstein would have the same visual importance as Spain if their respective attribute values were the same. This would not be the case with a choropleth map.

Scaling proportional symbols. Much research has gone into the optimal scaling for proportional symbols. As a general rule, make sure that the *area*, rather than linear proportions like radius or length of a side, is the scaled parameter. For example, if there are four times as many cars in Los Angeles County than in Orange County, the *area* of the symbol should be four times greater for Los Angeles County. If the symbol choice is a circle, the radius of the Los Angeles symbol should thus be only twice as great (since area scales with the square of the radius).

The total number of voters in each Czech district is represented with proportional symbols on the [Czech voting results map](#).

Dot maps

Used to represent themes that vary smoothly over space but are discrete, dot maps create a visual impression of density by placing a dot or some other symbol in the approximate location of one or more instances of the variable being mapped. Dot maps should be used only for raw data, not for data normalized or expressed as a ratio. Appropriate themes for dot maps include livestock farms, utility poles, and population distribution in a region.

Dot map parameters. When constructing a dot map, two parameters must be considered: the graphical size of each dot and the value associated with each dot. For example, you might stipulate that each dot be 2 pixels in diameter, and each represent 100 persons. In general, many small dots, each representing relatively few instances of the attribute, is more effective than a few large dots, but is more tedious to construct.

Warning: Dot map generation in present GIS is, from a cartographic standpoint, fatally flawed in that dots are arranged within each spatial unit in a random fashion. For example, [a dot map of alligator population](#) aggregated to the entire U.S. would be rendered with just as many dots in Iowa as in Florida. To create a dot map that does not mislead, the map must be saved as a vector image and brought into a drawing package like FreeHand or Illustrator, where each dot is physically moved to a more logical location. Unless you wish to go to this trouble, dot maps should be used only for very detailed spatial information (for example, county- or township-level data of alligator population for the entire U.S.) or with a great deal of caution and awareness.

Many reference maps could be considered dot maps, the location of each dot representing the location of a point of interest, a building, or some other discrete point information.

Data classification II

In general, the classification method chosen should accomplish two things:

1. maximize the between-class differences
2. minimize the within-class differences In some case, quantile and equal interval methods meet these criteria, but often the arbitrary nature of these methods is inferior to those mentioned here.

The most straightforward way of accomplishing the criteria above is to input the class breaks manually after careful study of the data. For ordered data, this is done by sorting the data from lowest value to highest, as done with the "[percent living on active farms](#)" data, and looking for large gaps, or *natural breaks*. It is here that the class breaks should be placed in order to meet the two criteria above. Some advanced GIS packages will do this procedure for you automatically, using a technique called Jenks Optimization, which will iteratively calculate best classification strategy.

One other technique worth noting is *nested means* classification. In this method, the mathematical mean of the attribute values, \bar{m} , is calculated and a class break is placed at \bar{m} . This separates the data into two classes -- those values above and those values below the mean. Data is further classified by calculating the means of the values within these two categories, and inserting class breaks at each of those two points. This leaves four classes, the breaks for which have been determined not arbitrarily but by methods specific to the data set and its characteristics. One more level of means may be calculated, leaving eight classes. This nested means method is mathematically straightforward and can be a good compromise for skewed data.

[Compare natural breaks and nested means](#)

Advanced color guidelines

Color models. The color of an object is generally described by its location in some color space or model. Most often, in computer displays, this model is RGB, or Red, Green, Blue. These three components of the color are each listed in values from 0 to 255 for *16-bit color* displays, for over 1.6 million color combinations. For example, a saturated red would be (255, 0, 0) in RGB space, whereas a desaturated dark green might be (150, 220, 150). Each of these colors has a corresponding coordinate value in HLS (Hue, Lightness, Saturation) space. Hues in such a model are generally listed from 0 to 255, and Lightness and Saturation are some real number between 0 and 100. Printers refer to the CMYK (Cyan, Magenta, Yellow, Black) model, based on the light-absorbing qualities of ink on paper. Combining inks of these colors is the basic idea of *four-color process printing*.

Advanced color techniques. Displaying information with an emphasis on its *divergence from the mean* or from some other arbitrarily chosen attribute value can best be done by using a diverging two-hue color scheme. In this scheme, low values are indicated by a saturated, dark hue (for example, maroon), values around the mean are indicated by a desaturated gray tone, and values much higher than the mean are indicated by a saturated second hue (for example, royal blue). This gives the visual emphasis on the locations of the extremities of the data. Data on relative poverty and wealth would be appropriate to represent using this color scheme, since a viewer would likely be interested in both the poorest and wealthiest regions. [Click](#) to go an example of this color scheme.

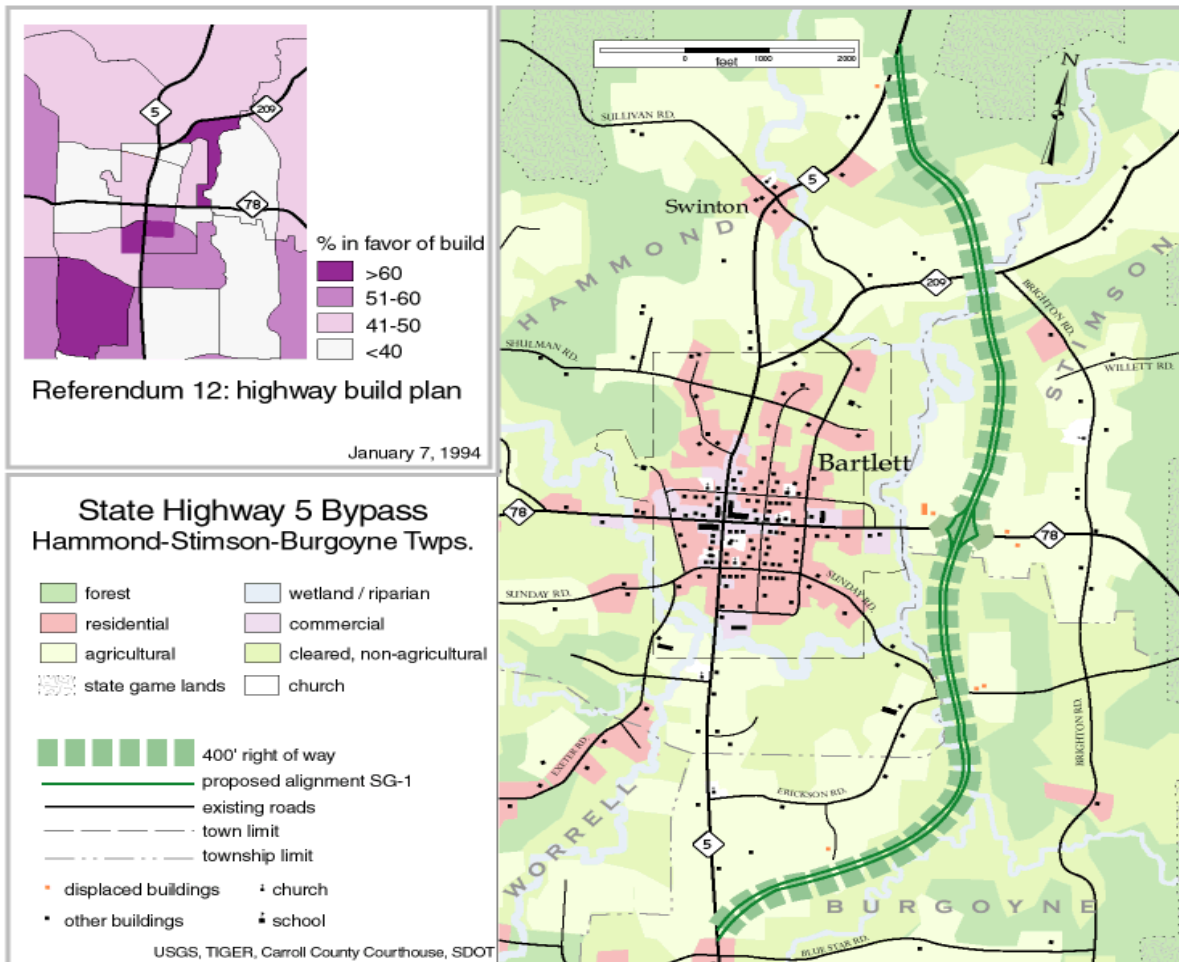
A designer interested in emphasizing the relative amount of two qualities that differ in kind from one another would be wise to use a *hue transition* scheme. A display showing the party affiliation of the state congresses for all fifty states might use a bright red for a heavily Republican state congress, a purple (red-blue) for a 50-50% mix of party affiliation, and a bright blue for a heavily Democratic representation.



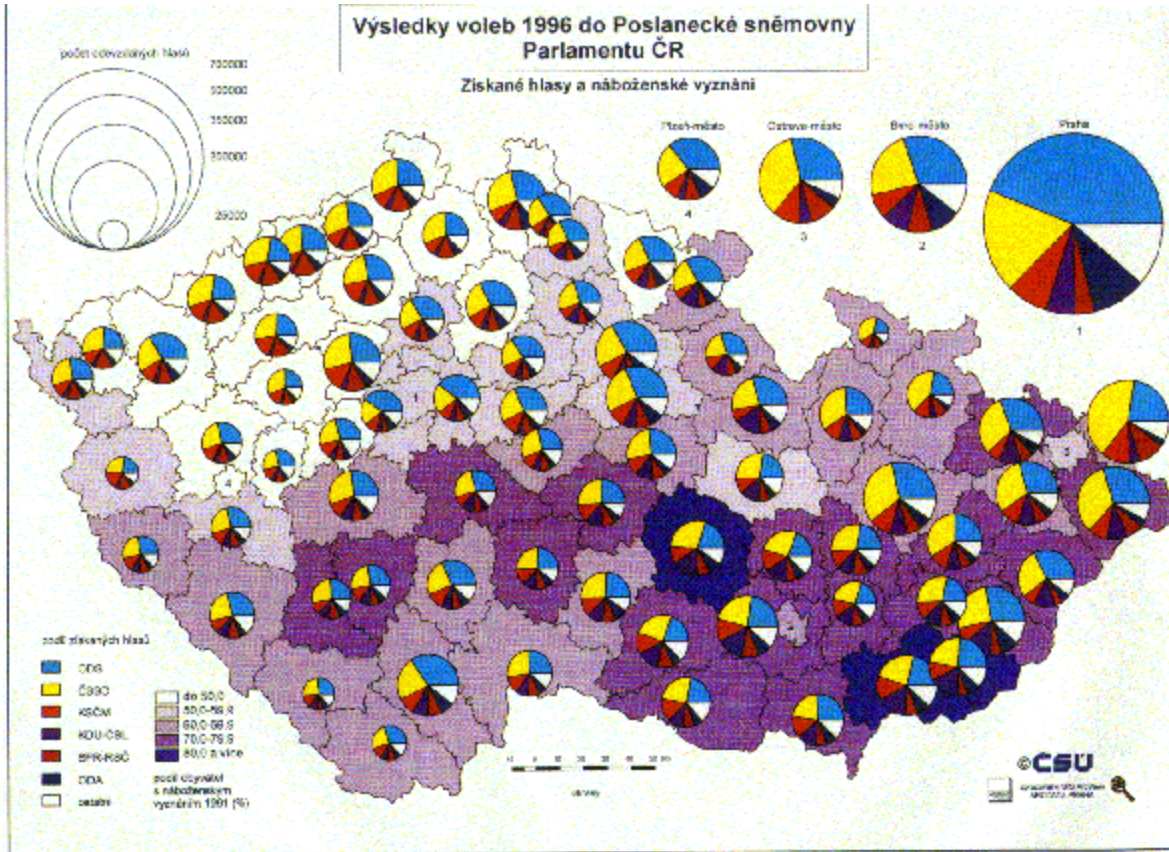
***Opacity.* One further dimension of color which is becoming more widely available in on-screen displays is a color's opacity. Three-dimensional displays of time series data or atmospheric phenomena might take advantage of this characteristic of color by enabling a viewer "see through" a surface.**

Example Application Revisited

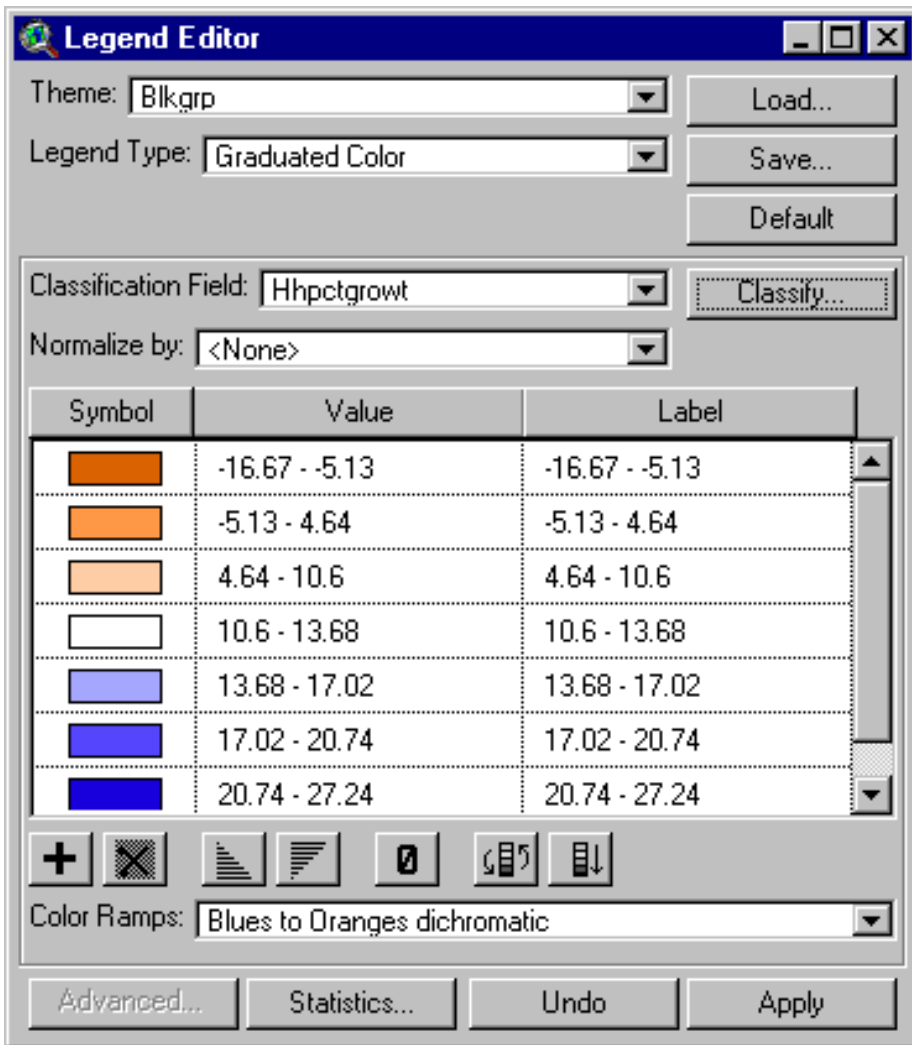
The GIS specialist in our example application can use the concepts above to create a cogent and logical display of the proposed highway. Let's examine the choices she made:



- Using different types of linework, her display included all of the political boundaries in the area. She left them highly detailed, particularly in the immediate vicinity of the planned highway corridor, knowing that townships, counties, and other political divisions are important in policymaking decisions of this kind.
- Land-use, nominal-level data, was represented with area fills of different hues. The hues were of similar lightness - light shades of each hue that would not visually dominate the display but that were easily discernable.
- Wildlife and other protected lands were indicated by the overlay of a pattern fill.
- Road and hydrology data was carefully selected to illustrate only those segments that would be directly affected by the construction of the highway. The visual emphasis of the display is on the detailed area.
- The color of the buildings indicates whether they lie in the right-of-way for the new highway, indicated by a green dashed line. Their shape indicates the building type.
- The location of the highway itself is indicated by a high-contrast solid green line.
- The survey results, which are aggregated to township level, are shown in a separate choropleth map with an ordered lightness difference indicating the respondents' sentiment.



from ARCDATA PRAGUE, in *GIS: Our Common Language*, ESRI Map Book, Volume 12, 1997.



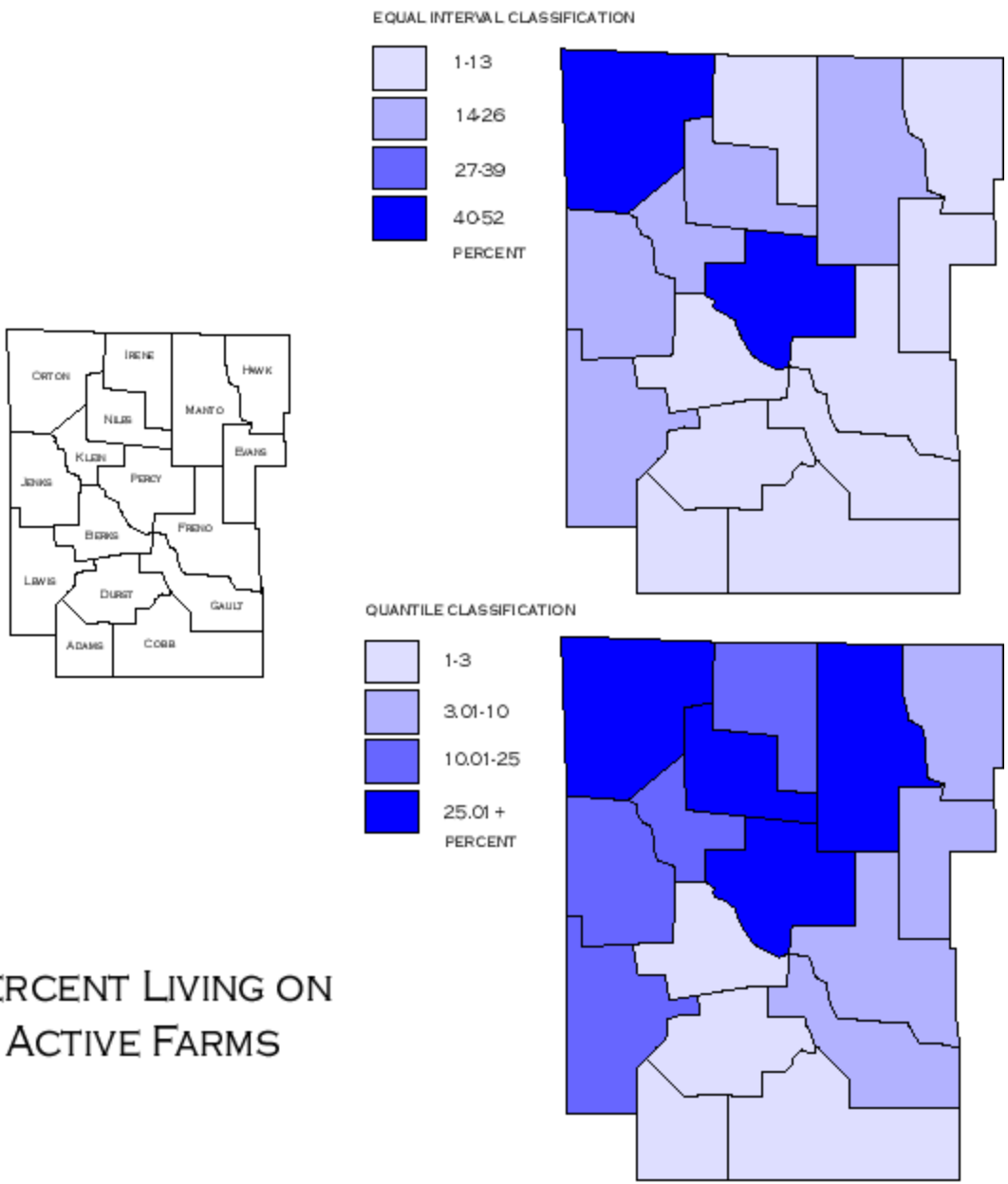
Current versions of ArcView contain very useful tools for visualization in the Legend Editor, including automatic color schemes, map forms, and classification. By double-clicking on the theme which you'd like to adjust in ArcView's Table of Contents in the View window, the Legend Editor appears. In the above example, the settings create a choropleth (Graduated Color) map of the field "Hhpctgrowt." The user has selected the natural breaks classification method (adjusted by clicking the Classify... button) in a seven-class dichromatic scheme. This scheme emphasizes the extremes in the data, with dark colors assigned to the very lowest (in this case, negative) and the very highest attribute values. The field, "Percentage growth in the number of households, 1990-1994," is most useful as an indicator of the fastest (and slowest) growing residential regions, and these extremes in the data should be highlighted. In the "Color Ramps" pull-down menu, different schemes may be selected, like "monochromatic" (value-difference) schemes, spectral schemes, and other commonly used schemes. Their relative appropriateness depends on the data being mapped and the argument being presented.

The symbolization of each category can also be manipulated by double-clicking on the appropriate swatch of color in the Legend Editor. Options such as a change of color, a change

of fill, or other symbolization changes are made available through the Palette Editor window which opens.

These are but a few of the many visualization options available to users of ArcView. Consult the documentation for other ideas.

Comparing Equal Interval and Quantile Classifications



What is different about these two maps? The data are the same, though the interpretation of each map (and thus of the data) might be quite different. The only difference between the maps is the methods used to classify the data: The top map uses the [Equal Interval](#) method, and leaves the impression that the percent of population living on active farms is about the same from the northeast corner of the region through the entire southern one-third.

The bottom map shows this same data classified according to a [Quantile](#) classification method, and this map shows what appears to be a large percentage of active farmers throughout the north.

The classification method you choose can have a large bearing on the final impression produced by the map.

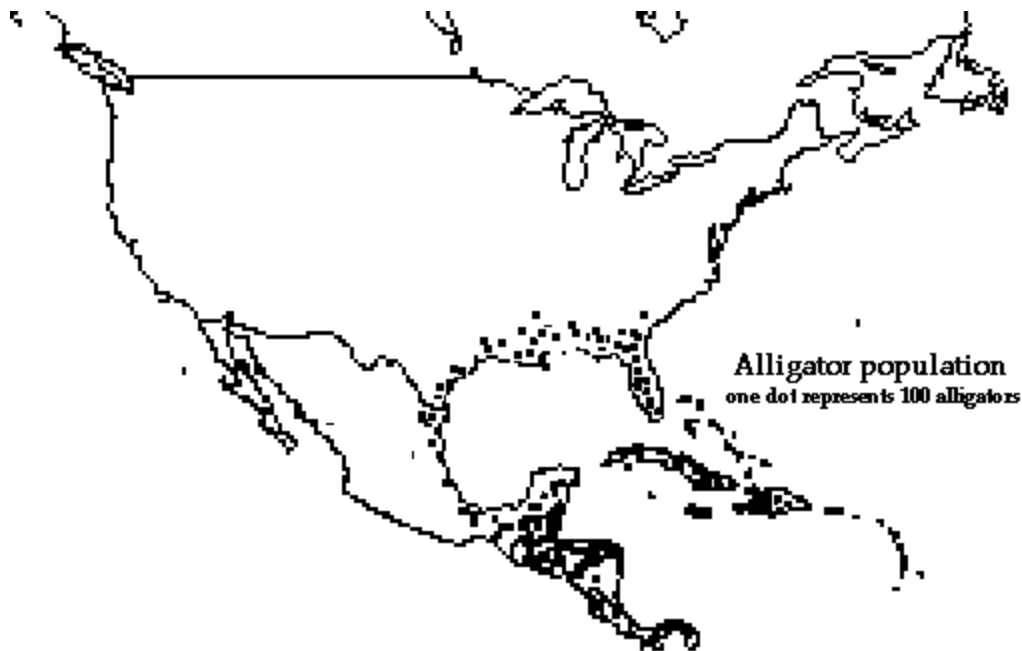
[Take a look](#) at other classification methods.

Dot maps

Dot maps can be very misleading if not created properly. The map below was created in a version of the analysis software Atlas Pro which places dots randomly within the enumeration units (in this case, countries of North America). Note that this method results in the implication that there are at least as many alligators in the midwest and the desert southwest as there are in Florida.

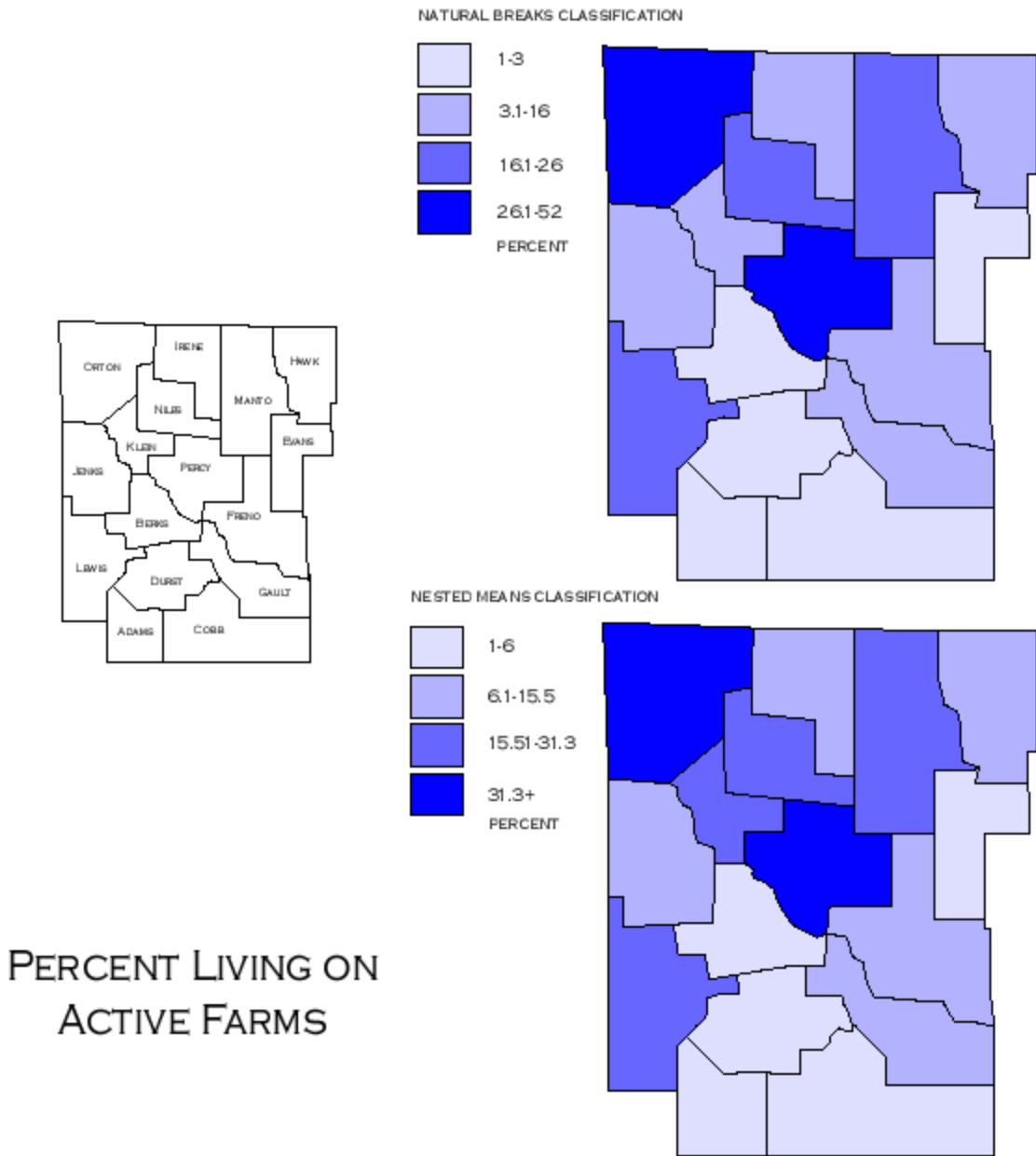


The map below was transferred to a drawing program (FreeHand) and the dots were moved to more logical positions. Dot maps are difficult and cumbersome to create appropriately; unless a good deal is known about the processes underlying the phenomenon (like the fact that alligators prefer swamps to deserts), dot maps should be avoided or used with caution and awareness.



Note: there are numerous other design problems with these maps. The color of the dots, for example, is the same (black) as the color of the country boundaries; the map would be much easier to read had the dots been red.

Comparing Natural Breaks and Nested Means Classifications.



These maps were classified more carefully than those created with [cruder methods](#) of classification. The top map shows a so-called "[natural breaks](#)" method, where the user assigned class breaks where there were large gaps in the data. There is very little difference between this method and the more rigorously mathematical "[nested means](#)" method, shown in the lower map. Both of these schemes almost always produce more accurate representations of the phenomenon in the display than the quantile or equal interval methods.