Choosing Colors for Data Visualization

Maureen Stone
January 17, 2006

The problem of choosing colors for data visualization is expressed by this quote from information visualization guru Edward Tufte: "…avoiding catastrophe becomes the first principle in bringing color to information: Above all, do no harm." (*Envisioning Information*, Edward Tufte, Graphics Press, 1990)

Color used well can enhance and clarify a presentation. Color used poorly will obscure, muddle and confuse. While there is a strong aesthetic component to color, using color well in information display is essentially about function: what information are trying to convey, and how (or whether) color can enhance it.

The most important use of color in information presentation is to distinguish one element from another, a function Edward Tufte calls "to label." In Figure 1, for example, the different colors in the scatter plot label different products. But the use of color as a label goes beyond the distinctive data colors, for color in this discussion includes black, white and shades of gray.
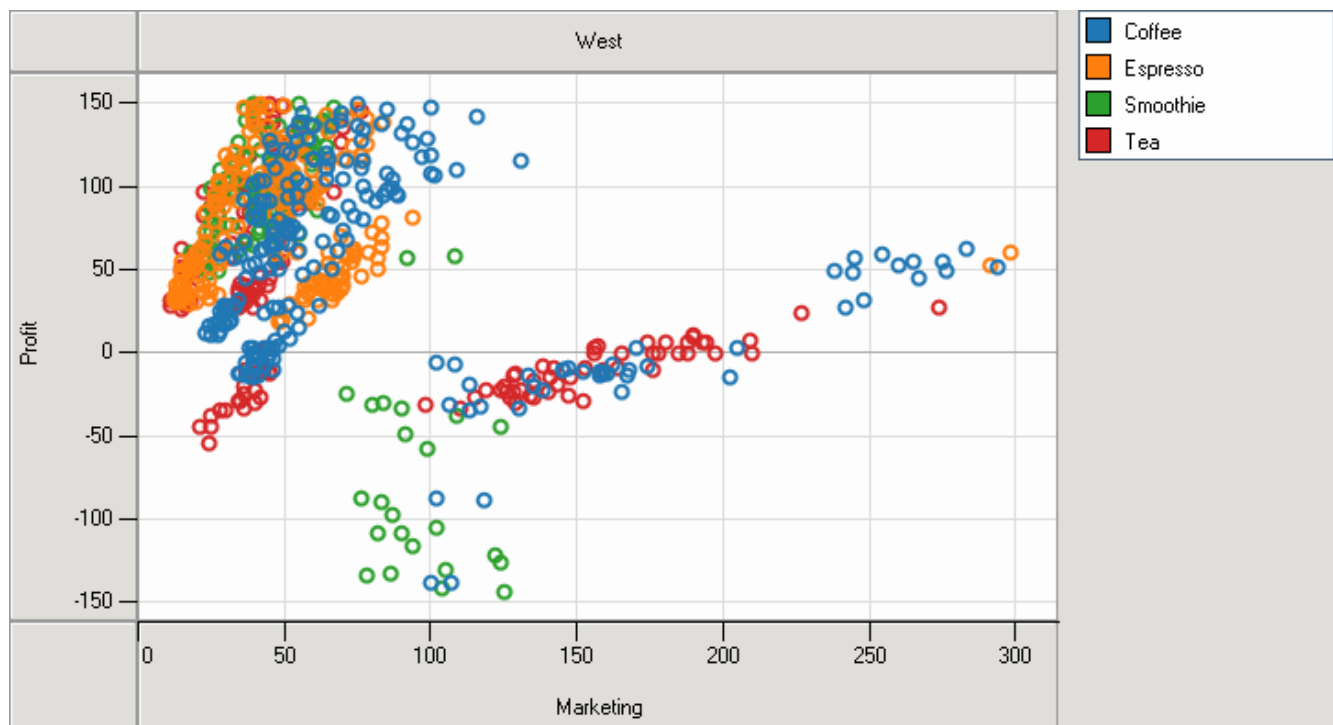

Figure 1. Scatter plot showing marketing vs. profit for four different products.

In Figure 1, color is also used to define the background, draw the grid and axes, and label the chart. All visible parts of a presentation must be some color, all of which must work together to be effective.

An effective design presents information in an organized manner, making it easy for the viewer to understand the roles and the relationships between the elements. A good organizing principle is to define categories of information, grouped by function and ordered by

importance. An effective use of color will group related items and command attention in proportion to importance. And, all will be legible.

For example, all of the contextual information (grid, axis, labels, borders) in Figure 1 are shades of gray, while the data is brightly colored, which makes the data the focus of attention. The data colors are chosen so that all seem equally important, and all are easily visible on the white background. The light gray grid and axes lines are legible, but similar enough to the background white that they don't visually interfere with the overlaid data points. The gray borders are dark enough to frame the plotting area, but light enough that the black text can be easily read.

Given an organized set of data elements, and a clear understanding of their function and importance, how does one go about choosing an effective set of colors? To answer this question, we will first look at some basic principles of color design, then the problem of legibility, and conclude with some guidelines for picking colors based on these principles.

Principles of Color Design

Contrast and analogy are the principles that define color design. Contrasting colors are different, analogous colors are similar. Contrast draws attention, analogy groups. In Figure 2, the contrasting red squares stand out from the field of analogous blue-green ones, and capture your attention first.
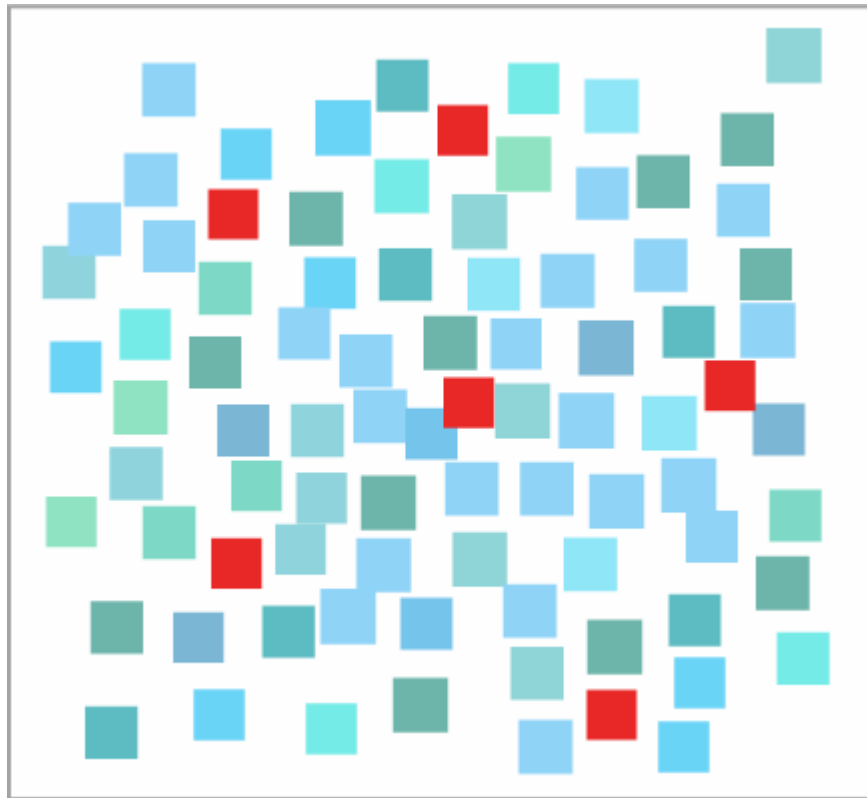

Figure 2. Contrast and analogy. The red squares contrast with the analogous blue-green ones.

In color design, color is specified by three dimensions: hue, value, and chroma. Hue is the color's name, such as red, green or orange. Value is the perceived lightness or darkness of the color. Chroma describes its colorfulness. High chroma colors are vivid or saturated, low chroma colors are grayish or muted.

The hue dimension is circular, typically drawn as a hue circle (Figure 3). There are many different hue circles, reflecting different media and mixing models, but all present the colors in the same order. In any hue circle, analogous hues are close together, most simply variations of the same color name (such as red, or red-orange). Contrasting hues are on the opposite side of the hue circle, though an even smaller separation can be sufficiently distinct.
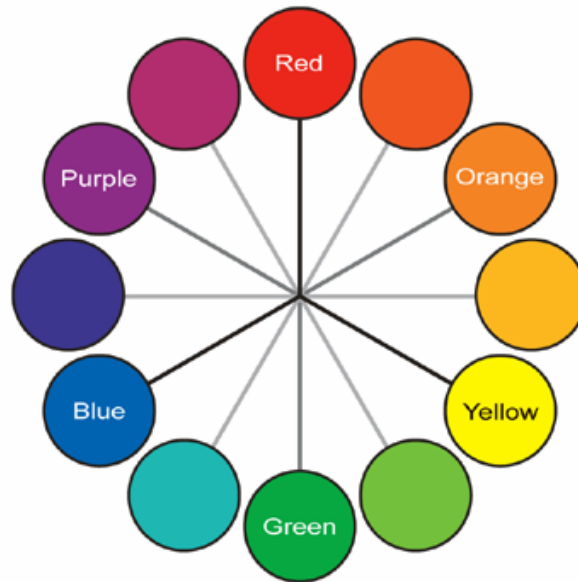

Figure 3. Example of a hue circle (A Field Guide to Digital Color, Maureen Stone, A K Peters, Ltd.)

The value dimension is visually the most important, for contrast in value defines legibility as well as having a powerful effect on attention. It is easy to see the lightness variation in shades of gray, or in shades of a single color. It is more difficult to compare the value of two very different colors. The value scale is typically defined such that black = 0, and white = 100. Figure 4 shows different color gradations that demonstrate (approximately) the same value scale.
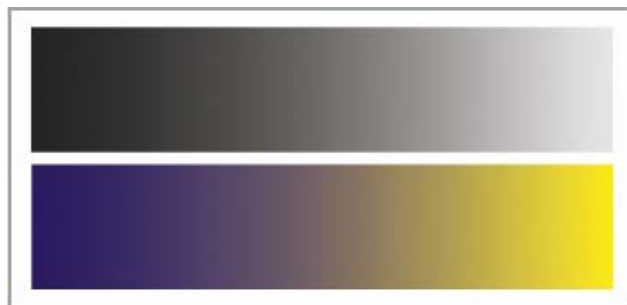

Figure 4. Both color gradations illustrate (approximately) the same value scale (A Field Guide to Digital Color, Maureen Stone, A K Peters, Ltd.)

Chroma indicates how bright, saturated, vivid or colorful a color is. Formally, for any given color, reducing chroma to zero produces a gray of the same value. The maximum chroma for the color will vary with the color and the medium (display vs. print, for example). Fortunately, it is rarely important to be so formal. On a display, the high chroma colors are vivid and bright, and unfortunately, the easiest to select in many applications. Using colors that are darker and grayer, or more pastel (closer to white), has many benefits. The result looks less garish and more sophisticated, is easier to duplicate in print, and allows the use of saturated colors for highlighting. Figure 5 shows examples, organized as in the graphic arts into tints, tones and shades.
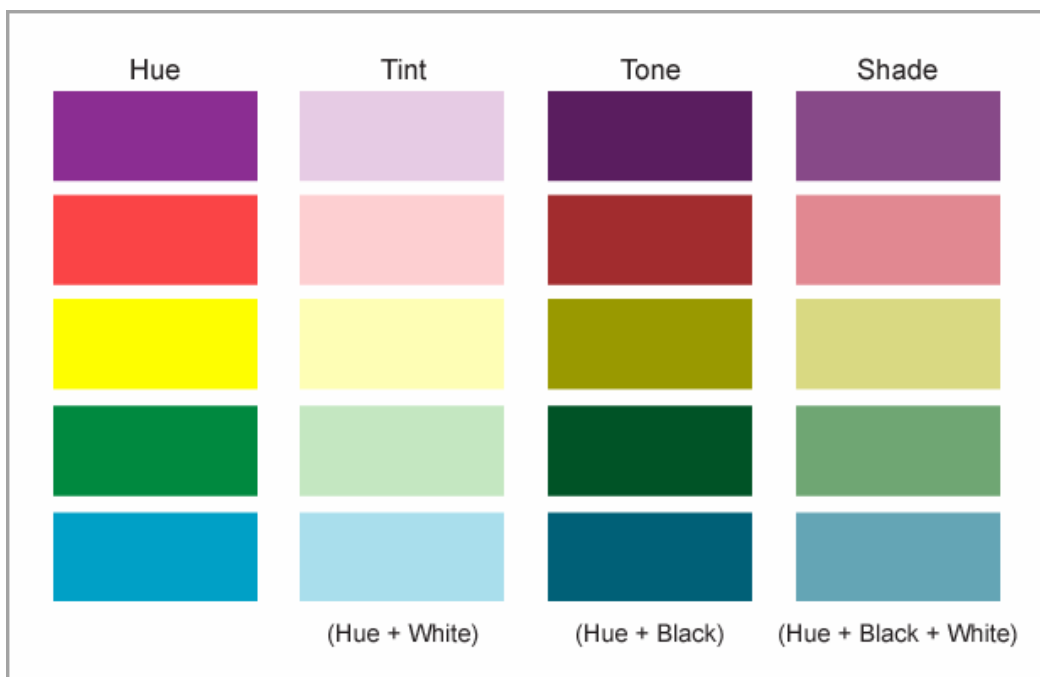

Figure 5. Tints, shades and tones of five different colors. (A Field Guide to Digital Color, Maureen Stone, A K Peters, Ltd.)

As an artistic exercise, contrast and analogy can be applied in any of the three color dimensions. Different dimensions, however, have different application to information display. Making related items the same color (analogous hue) is a powerful way to label and group. However, hue contrast is easy to overuse to the point of visual clutter. A better approach is to use a few high chroma colors as color contrast in a presentation consisting primarily of grays and muted colors. Contrast in value is critical for legibility, and analogous values can also be used to define layers of attention, as I will discuss in the next section.

Legibility

The word "legible" has its roots in the Latin word legere, meaning "to read." Its modern definition includes to read, decipher, discover and to be understood. In information presentation, legibility should be the minimum requirement for any visual element. If it can't be read or deciphered, why include it?

The single factor that determines legibility is the difference in value between the symbol (text, line, etc.) and its background. Differences in hue and chroma do not contribute at all. This difference in value, formally specified as luminance contrast, is the perceptual stimulus that

the human visual system uses to see edges. The higher the luminance contrast, the easier it is to see the edge between one shape and another. If the contrast is too low, it can be difficult to distinguish between similar shapes, or even discern the shape at all.

In text of a medium gray value is displayed over a background that varies from light to dark. Where the contrast is high, the text is legible. Note that this varies with the size of the text, with higher contrast needed for smaller text. Coloring the text green provides a large hue and chroma contrast with the gray background. However, the text still varies in legibility, depending on the value contrast.
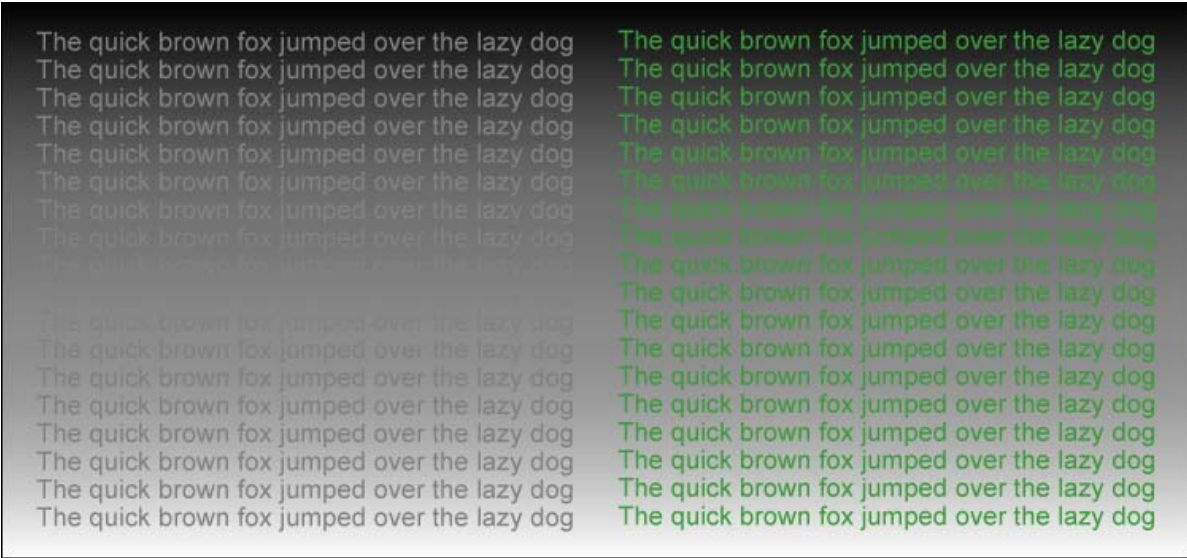

Figure 6. (Left) Medium gray text on a background that varies from black to white. The text is only legible where there is sufficient luminance contrast. (Right) Coloring the text green creates hue and chroma contrast, but this does not improve legibility.

Variation in luminance can also used to separate overlaid values into layers, where low contrast layers can sit behind high contrast ones without causing visual clutter (like the grid lines in Figure 1. Figure 7 demonstrates this principle, and the fact that contrast in hue does not create this separation.
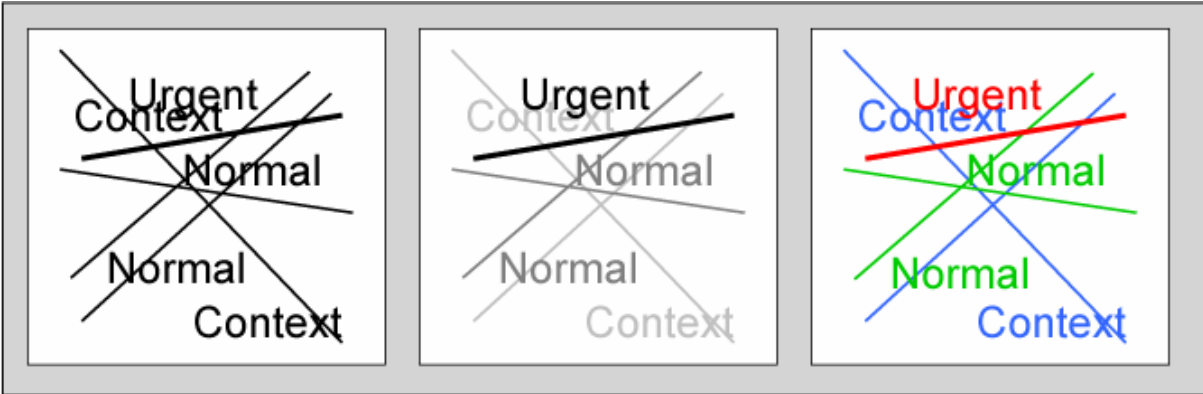

Figure 7. (Left) The overlaid black text is difficult to read; (center) Creating three different levels of contrast separates the information; (right) Using three different hues does not. (Larry Arend, NASA)

The importance of value contrast is why a primary rule in many forms of design is "get it right in black and white." In an ideal design, all important information would be legible even if chroma were reduced to zero everywhere in the presentation, leaving only shades of gray.

Guidelines for Selecting Colors

The principles specified in the previous sections can be summarized as: assign color according to function; use contrast to highlight, analogy to group; and control value contrast for legibility. How do we move from principles to palettes?

In most design situations, the best results are achieved by limiting hue to a palette of two or three colors, and using hue and chroma variations within these hues to create distinguishably different colors. Such a palette is both aesthetically pleasing and functional. It minimizes an over-dependence on hue variation (which can cause visual clutter), and replaces it with careful control of value and chroma. Figure 8 shows several such palettes, and their location on the hue wheel.
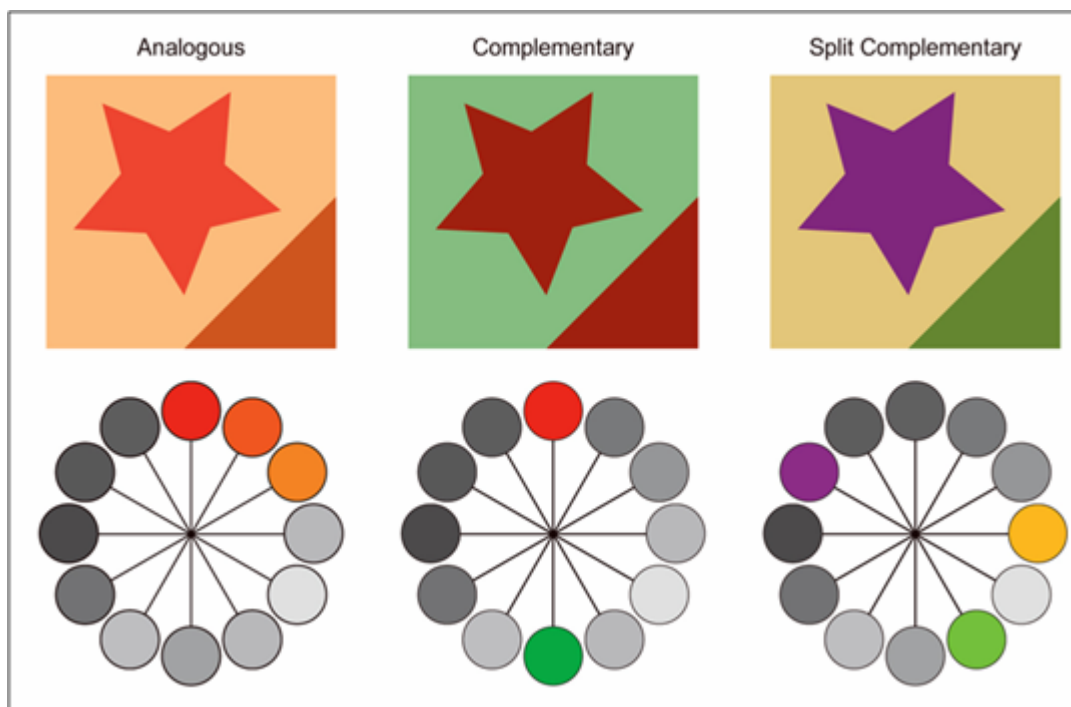

Figure 8. Several different palettes and their relationship to the hue wheel. (*A Field Guide to Digital Color*, Maureen Stone, A. K. Peters, Ltd.)

There are many, many such color palettes; how do we choose a good one? Even for a professional graphic artist, one of the easiest ways to find a good color palette is to use one that someone else has already designed. Most graphic arts books on color are filled with examples of color palettes.

One good source of colors for data display is Cynthia Brewer's "ColorBrewer." While designed for thematic maps, her colors are also useful for charts, graphs, and other data presentation displays.

You can also create your own palettes based on the principles above. Unfortunately, most color selection tools, even those whose labels suggest they model graphic design principles (such as HSV or HLS), provide poor control over value. Using them effectively requires either a discerning eye or a clear understanding of the relationship between RGB and luminance (or both).

Let us conclude with some examples of different palettes applied to the same data to see how different choices affect the emphasis and message.

Figure 9 is a line graph, showing sales trends for four different quarters. Each quarter is labeled a distinctive color of similar value. These colors are approximately equally spaced around the hue circle to create the largest possible hue difference.
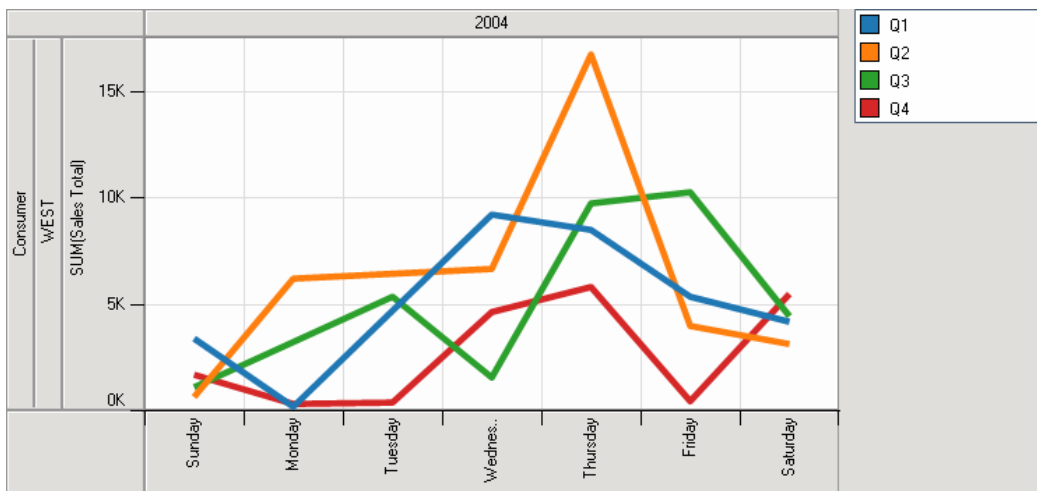

Figure 9. Line graph showing sales on a daily basis, separated by quarter. Each quarter is a distinct hue.

It is not necessary, however, to make the hues so different. Figure 10 shows two examples that span only a segment of the hue circle, yet still create distinctly different colors. Which to use can be chosen to coordinate with the corporate logo, or simply by aesthetic preference.
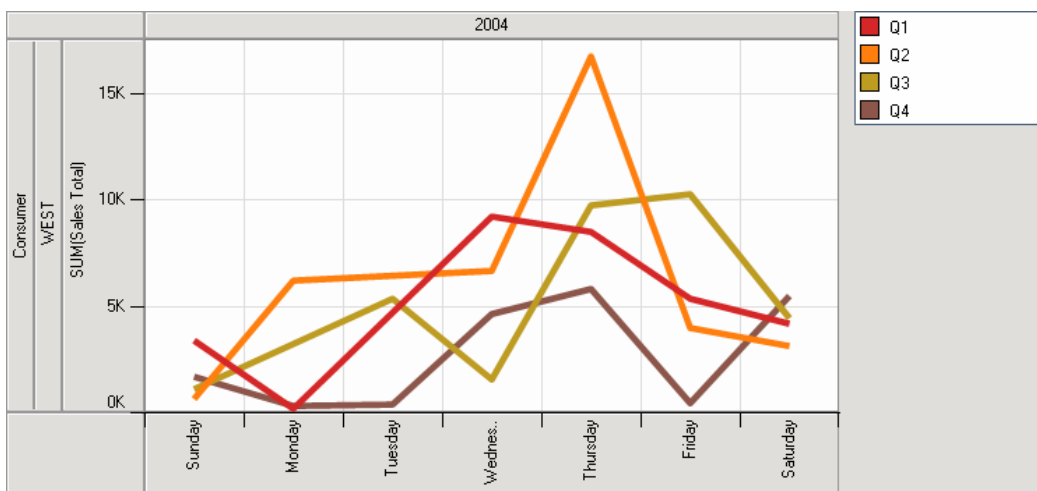

Figure 10. Other colorings of Figure 9, where the hues are more similar but still distinct. There is no functional difference between these colorings and the one below in Figure 11.

Another approach is to use a common hue that varies in chroma and value. Figure 11 uses a palette of different shades of blue, ordered from light to dark. Such a progression suggests an ordered relationship, in this example, time.
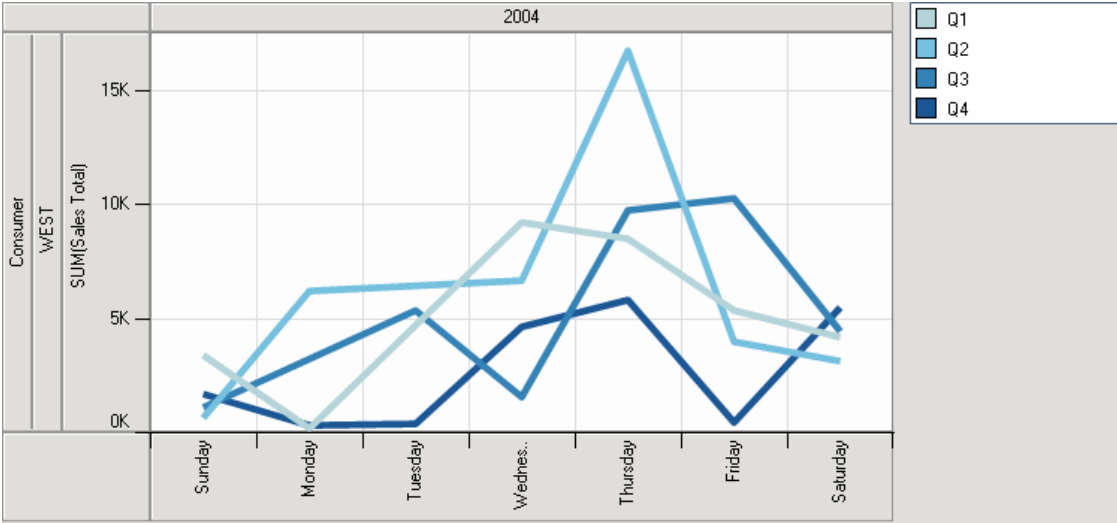

Figure 11. An all blue coloring of the data in Figure 9, ordered from light to dark. This coloring reinforces the sequential order of the quarters.

The same approach can be used to render the image entirely in shades of gray, as shown in Figure 12.
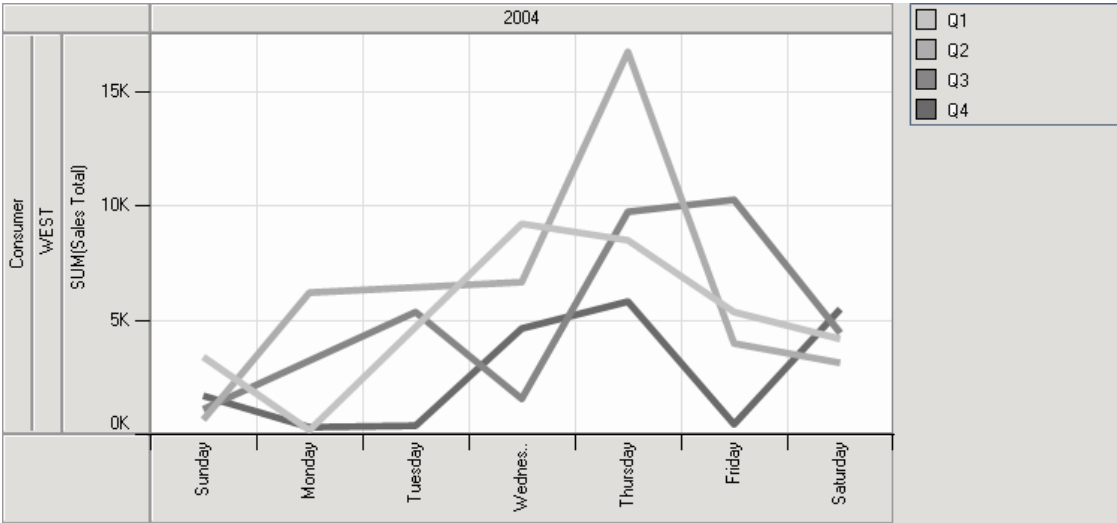

Figure 12. An all gray rendering of Figure 9.

This rendering also provides the perfect background for highlighting with a bright color, as shown in Figure 13. In this example, the focus is clearly the 4th quarter, instead of all four quarters equally.
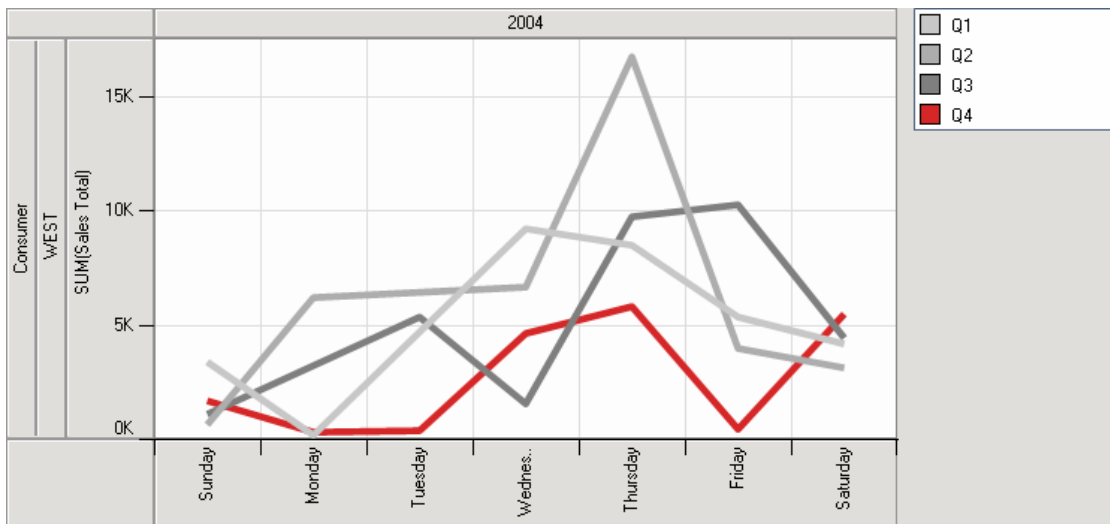
Figure 13. A bright color on gray provides the maximum contrast and emphasis.

In all of these examples, the background is white, and the contextual information (grid lines, labels and their backgrounds) are shades of gray. As a general rule, making the background white and its supporting information shades of gray provides the most effective foundation for your color palette. While dark backgrounds were historically recommended for projected slides or computer displays, this recommendation is no longer valid.

Acknowledgements

All of the data plotting examples (Figures 1, 9-13) were created with the award winning Tableau data analysis software. Figures from my book, A Field Guide to Digital Color, were reproduced with permission from A. K. Peters.

Note—A few words on background color

Most color palettes (including those in the ColorBrewer) are designed to be printed on white paper. When used in digital form, therefore, they should be displayed on a white background. More broadly, a white background gives the appearance of color on paper, which makes it easier to choose colors that both display and print well.

Using white as a background color has perceptual advantages as well. The human visual system is designed to adapt its color perception relative to the local definition of white. It also adjusts its focus slightly as a function of color. Providing a white background gives a stable definition of white, and a stable "surface" to focus on.

The only functional reason to use a dark background is situations where the image will be viewed in the dark. Light text on a dark background minimizes the problem of light from the background "washing out" the text, and is more comfortable to view with dark-adapted vision. Slide projectors, early color displays, and the first digital projectors were dim enough to require a dark room, which created the practice of dark backgrounds for slides and computer generated displays. Modern displays and digital projectors are sufficiently bright that they can be viewed in normal room lighting.

About the Author

Maureen Stone is an independent consultant working in the areas of digital color, information presentation, interaction and systems. Before founding StoneSoup Consulting, she spent 20 years at the Xerox Palo Alto Research Center where she attained the position of Principal Scientist. Maureen's book, *A Field Guide to Digital Color*, was published by A.K. Peters in 2003. She has more than 30 published papers plus 12 patents on topics including digital color, user interface technology and computer graphics.

This article originally appeared on the Business Intelligence Network (www.b-eye-network.com) as one of Stephen Few's guest articles. A library of Stephen Few's articles, as well as other guest articles, is available at www.perceptualedge.com.