Paper 139-27

Inform and Influence with Image and Data: Communication-effective Web Design for ODS, SAS®, and SAS/GRAPH®

LeRoy Bessler, Bessler Consulting & Research, Fox Point, Milwaukee, Wisconsin, USA

Abstract

SAS/GRAPH has long been rich, and becomes progressively richer, in features. With these features comes the need to apply the technology in the most communication-effective manner. ODS (Output Delivery System) comes with the need to assure that the tabular data presentation options of grids and backgrounds that are now possible are indeed used for communication, not simply decoration. Color can be an obstacle to communication, or an enhancement. The same is true of fonts and 3D. You may not be aware of issues of color combinations, color blindness, or browser-safe colors for the web. Or realize there are certain fonts expressly designed for readability on the web. Or have noticed that 3D graphs can deceive or obscure. Or have considered that a scrollable multiple-image web page can lead to viewer distraction, and is a nuisance when you want to print only one image. This tutorial on The Power to Show explains how to deliver information with graphs, tables, and maps, for web or hardcopy, designed to reveal and persuade, not to impress. You will see how to exploit the power of simplicity. The content is operating system independent, and suitable for all skill levels. The design principles and ideas are software independent.

Introduction

Some TV ads make you remember their action and images, but not what was being promoted. The time and attention of the "audience" for your charts are precious, as are your time and effort to create those charts. Software and hardware are power tools, but defaults produce lots of sub-optimal results quickly. With the web, the whole world is watching.

Charts can be designed to inform and to influence, so that they deliver the messages and meaning in your data with images that are easily and quickly interpreted, and have memorable *significance*. Charts can be constructed to deliver overview impact, for rapid decision-making, *and* precise detail, for reliable decision-making.

The principles and ideas in this paper, developed during 21 years of SAS and SAS/GRAPH use, apply to presentation of enterprise information, not necessarily to analytical or statistical graphs. Some guidelines may not apply to your situation. "Take what you want, and leave the rest."

Due to a six-page publication limit, this paper cannot deal with ODS in detail. For more about ODS matters, please see SUGI 27 papers cited in the section "For More Information". Also, the author will present other illustrations and a demo, as much as presentation time permits. Very little code can be printed in this paper. Other code is available via email.

Communicate with Simplicity and Focus

Dilbert (Scott Adams' cartoon character) said, "I made a study of Internet use in the workplace. The time spent waiting for web pages to load has wiped out all the productivity gains of The Information Age."

1. **Keep your image simple and focused. Design it to accelerate understanding and inference.** Simple loads faster. Never underestimate the power of simplicity. See the Figures for recommended designs.

2. Special Effects are for Movies, 3D is for 3 Variables. **Good design and interesting data can stand on their own.** Omit drop-shadows, clip art, etc. 3D pie charts *always distort* relative size of shares of the whole, defeating the visual communication purpose of pie charts (see Figure 2). 3D bar charts needlessly complicate a simple image, usually making it more difficult to interpret (see Figure 4). 3D maps are either communication problem prone (PROC GMAP's PRISM and BLOCK maps have the solids for high response areas hiding those for low response areas) or impractical (the PROC GMAP SURFACE map).

3. **Remove axis clutter.** (See all the custom charts.) Turn off axis lines. Turn off tick marks. Label (invisible) tick marks sparingly. If not turning off axis labels, supply your own, but not to state the self-evident (e.g., that dates are dates). These paraphernalia are historic holdovers from the bygone era of graph paper, pen, and ink. Communication is most effective with direct annotation of plot points and bar ends.

4. **Use Sparse Annotation.** (See Figures 8 and 9.) Annotate y-values for the critical points of a plot line, and label only corresponding (invisible) x-axis tick marks. *Critical points* are start, end, maximum, minimum, and points where the rate of growth or decline persistently changes. To provide *all* the values, provide flyover text or a linkable table.

5. **Use area fill wisely, if at all.** Omit area fill on line charts. When fill is really needed, use solid colors or grays, not cross-hatching. On bar charts use light gray, or some other light hue. (See Figures 4 and 5.) Dark or intense area fill creates a distracting, vision-dominating mass.

6. **Rank data in bar charts, pie charts, and annotated maps.** (See Figures 2, 4, 5, and 6.) On maps, supply rank as an annotation.

7. **To depict shares of the whole when numerous, use a Customized Ranked Horizontal Bar Chart.** When the shares of the whole are too numerous or too small, there is not room to display pie slice name, value, and percent of whole. (See Figures 3 and 5.) In this case, the pie slices are large enough that area fill colors can be distinguished, so that a customized pie legend is another solution. (See Figure 6.) But, in some cases, the slices are too small, and the hbar chart is the only solution.

8. **When possible, use the powerful Pac-Man Pie Chart to summarize shares of the whole into two.** (See Figure 1.) The idea of a two-part pie chart may seem trivial, if not silly. However, when the share of interest to your message is either tiny or huge, the image is very "impactful" and, therefore, memorable. If needed, supply details for "Other" with a table below the pie chart, or—on the web—via link to a ranked table of data for *all* the slices. Do not blunt the message by splitting the big slice into little ones that may be as small as, or smaller than, the slice whose smallness you are emphasizing.

9. **For line or bar charts, usually start the response axis at zero.** Minimize meaningless fluctuations/differences point-to-point, and avoid magnified "growth" or "decline". Prevent needless questions, and undeserved elation or alarm. See Figures 8 and 9. Note the desirable near-flatness of the upper chart in Figure 9—variation over twenty-two years was less than ten percent. But, when fine structure of a chart is *really* important, devote all the available space to the actual data range.

Communicate with Color

1. For those who can't see a color difference, there is no difference. **The commonest form of color blindness cannot distinguish red and green**, which is a frequently used color combination in the USA. Prof. Jay Neitz of the Eye Institute of the Medical College of Wisconsin: 8 to 10 percent of American males have some form of color blindness (but, due to genetic differences, only about 0.5 percent of females). **For alternatives to the very popular "Traffic Lighting", see Figure 10.**

2. **Color does not improve a bad design. Use color to communicate, not to decorate.** See Figures 6 and 12 for communication-justified and communication-facilitating uses of color. If you have no need to distinguish response levels/categories, use black & white. If you have a few levels or categories, gray shades may suffice. If you have many levels or categories, color is necessary. Note: **It is impossible to reliably distinguish more than five shades of a single hue.**

3. **Color contrast between text and background is essential to communication.** ODS opened the door to trying to enhance so-called "boring" tables with color. Besides the bane of Traffic Lighting, there are problems using yellow with white, or black (or other dark) text on dark or intense background colors. Evaluate the text-background combinations in Figure 10. Also, adequate contrast for an online display does not guarantee the same for hardcopy, which is not brightly backlit.

4. **What are "browser-safe" colors?**

Unlikely as it may seem, many, if not most, web users have monitors or video cards limited to 256 colors. Even when the hardware has a much higher capability, it may be set to display only 256 colors. (To check or change your own setting on Windows, click Start -> Settings -> Control Panel -> Display -> Settings -> Colors.) You may wish to design for the lowest common denominator. Here is why and how.

To deal with audience equipment diversity, web browsers determine the currently set limits of the display, and, if needed, remap colors. Video display tubes produce their colors as combinations of Red, Green, and Blue. All web browsers agree on a universal common subset of 216 "browser-safe" colors. They are RGB colors (Red-Green-Blue combinations), with names, in SAS language, of the form CXrrggbb. Browser-safe RGB colors restrict rr, gg, and bb to six values 00, 33, 66, 99, CC, and FF. (6 X 6 X 6 = 216.) When a web browser detects a color outside this set on a web page to be shown with a 256-color display, it remaps the color to a browser-safe one.

All SAS "predefined colors" (see below) have RGB equivalents, but only seven are browser-safe. (SAS GREEN, contrary to the RGB value still in the manual, changed in Version 6.12, and is no longer browser-safe—even though Green is one of the three RGB primary colors.)

See Figure 11 for samples of browser-safe colors. The basic colors are Red (CXFF0000), Yellow (CXFFFF00), Green (CX00FF00), Cyan (CX00FFFF), Blue (CX0000FF), Magenta (CXFF00FF), Black (CX000000), and White (CXFFFFFF). The upper chart shows the only way for RGB colors to vary lightness with constant hue. It is easy to vary lightness with constant hue when using the HLS color system. If your target is hardcopy, HLS colors are an excellent choice, also providing easy tunability of transition in hue and saturation.

5. **Beware of "Predefined Color Names".** There are 292 "SAS Predefined Color Names", listed in Table 7.2 in the Version 6 and Version 8 SAS/GRAPH manual. They have names such as "PINK", or "LIPK" for "Light Pink". However, many of the names are misleading. If you display or print PINK and LIPK, you will see that Light Pink is darker than Pink. There are other contradictions like this. Also, many of the colors are too dark to be useful. To make wise color choices, create samples. Here is the simplest way to do so:

```
proc gslide;
note f=SWISSB c=black 'C=PINK'
c=PINK h=5 move=(+0,-1) '03'X;
note f=SWISSB c=black 'C=LIPK'
c=LIPK h=5 move=(+0,-1) '03'X; run; quit;
```

There are also about 150 newer color names in the SAS Color Registry. You can find them, and their RGB codes, using this click sequence in your SAS session: Solutions -> Accessories -> Registry Editor -> Colornames -> HTML. With these, too, assume nothing. Make yourself a sample chart. Only seven are browser-safe.

Communicate with Text

0. **Pulitzer's First Rule:** "Make it brief so they will read it."

1. **Make your graph title your headline.** You are using the graphic image to persuade and/or reveal. Don't be reluctant to tell the viewer what you know it means or think it means. (See Figure 9.)

2. **Use Sparse Text to make the graph talk.** Be sure every letter or number in the graph *must* be there. Superfluity detracts from emphasis.

3. **Use high contrast** (best exemplified by Black with White or Yellow). On colored backgrounds, colored text or line must be sufficiently thick. Never use Yellow on White, Black on Dark Blue, etc. (See Figure 10.)

4. **For emphasis,** usually use **Bold**, *Italic*, Underline, or ALL CAPS, not color. If using ALL CAPS for emphasis, use it SPARINGLY**.** ALL CAPS is hard and slow to read. To convince yourself, prepare and read a long paragraph in ALL CAPS, and compare that with Mixed Case.

5. **Check your SAS log** for a note that the undesirable SIMULATE font has been substituted for your requested font. (As of SAS Version 8.2, you still are not always notified.) For a sample of this font (or any other SAS/GRAPH font, specified below with **name=**), use:

```
proc gfont name=SIMULATE nobuild
     SHOWROMAN HEX H=2; run; quit;
```

A common error is a misspelled font name, in which case the SAS log seems to always notify you of substitution. If you use f=NONE, or fail to specify f= where applicable, in your SAS/GRAPH program, the device driver uses its default font. If, when getting a default font, you specify a height other than one cell (i.e., not h=1), then SIMULATE will be used, but you will not be notified. Also, some drivers (e.g., EMF) do not have a default font—SIMULATE is used without notification. *See axis and tick mark labels for the default graphs in Figures 4 and 7.*

GOPTIONS SIMFONT= assigns the default font. The software as shipped has it set to SIMULATE. Make SIMFONT what you prefer.

6. **Set SAS/GRAPH defaults with FTEXT, HTEXT, CTEXT.** Some features of some SAS/GRAPH charts are not controllable with f=, h=, c=. For them, these GOPTIONS parameters are your only recourse.

7. **You can use Windows TrueType fonts in Version 8.** Good fonts include Matthew Carter's creations designed for readability on the screen and the web, Verdana (sans serif—useful for small letters and numbers on your graph) and Georgia (serif—useful for titles and prominent footnotes). See how these fonts are used in the Figures. Rockwell, thickly stroked along the entire contour of its characters, was used in Figure 10 to make numbers more readable on color backgrounds.

Avoid ODS Gridlock, and Make Eye Travel Faster and Easier

If you have a very wide table, and need to find a column's data cell far to the right of a row label, a grid is not necessarily an aid. On the web, let a click on the row label take the viewer back to the same table, but with that entire row highlighted with a background color that does not impair readability. Omit the grid, and use color to communicate.

Tables, Web Publishing, Web Linking, Etc.

For these topics, the author refers you to the SUGI papers mentioned in the next section, and also invites you to request information about his Subsetted Ranking Report (useful for Enterprise Performance Reporting, Balanced Scorecard, etc.) The work cited below offers counterexamples to typical ODS-enabled tables, and provides all the supporting PROC TEMPLATE code. It presents and evaluates three methods, including one developed by the authors, for implementing communication-effective graphic and tabular data presentation on the web.

For More Information

For construction details of web data presentation applications using SAS software, at SUGI 27 Dr. Francesca Pierri and I are jointly presenting "Your Graphs and Tables at Their Best on the Web with ODS" and "%TREND: A Macro to Produce Maximally Informative Trend Charts with SAS/GRAPH, SAS, and ODS for the Web or Hardcopy".

Author Contact Information

Your requests, suggestions, comments, and questions are welcome.

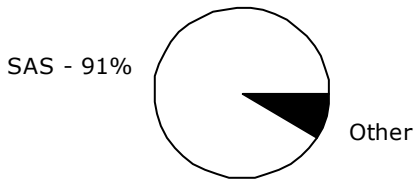
LeRoy Bessler, Ph.D.
Bessler Consulting & Research
PO Box 96
Milwaukee, WI 53201-0096, USA
bessler@execpc.com
1-414-351-6748

Zum Sehen geboren, zum Schauen bestellt
Born to See, Meant to Look
- Goethe

Ancoro imparo
I am still learning
- Michelangelo, at age 87

Fig. 1: The Simplest Pie Chart - The Power of Two

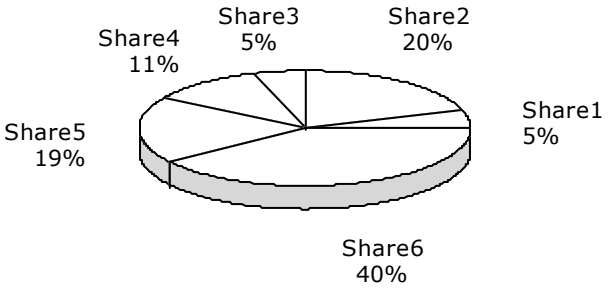
Mainframe Data Analysis Software Market Shares



Information Source: Computer Intelligence, 1993

Fig. 2: Compare 2D and 3D Pie Charts

3D Distorts Apparent Relative Share Size



Descending 2D Chart
Shows Them What's Important

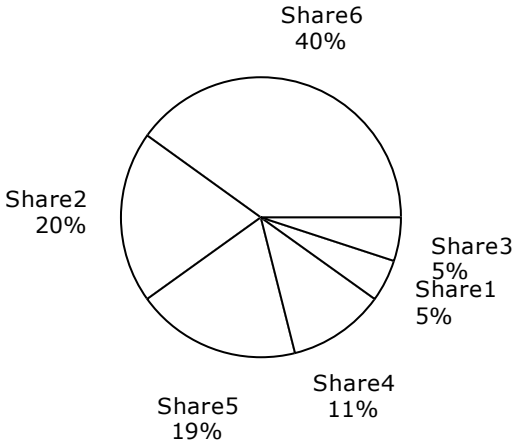


Fig. 3: Problem with Some Pie Charts

Information for Adjacent Slices Can Overlap

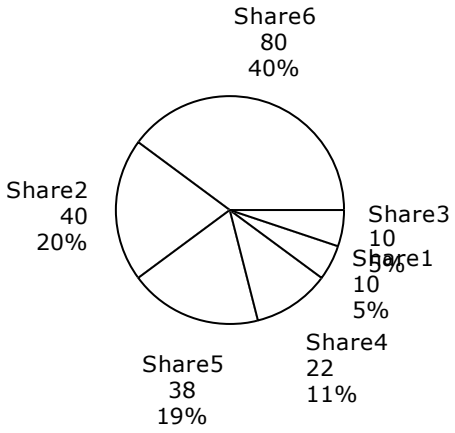
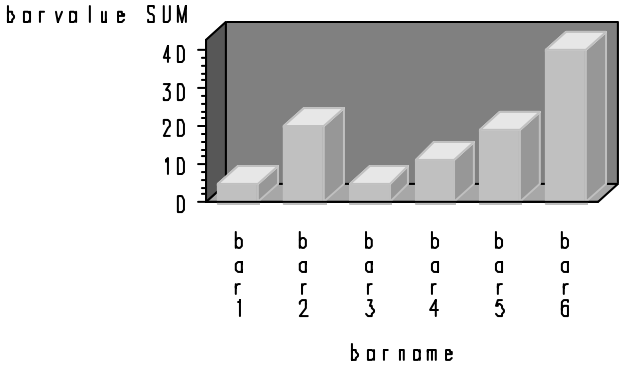


Fig. 4: Compare 2D and 3D Bar Charts

Default 3D Chart
Complex and Harder To Interpret



Optimized Descending 2D Chart
Shows Them What's Important

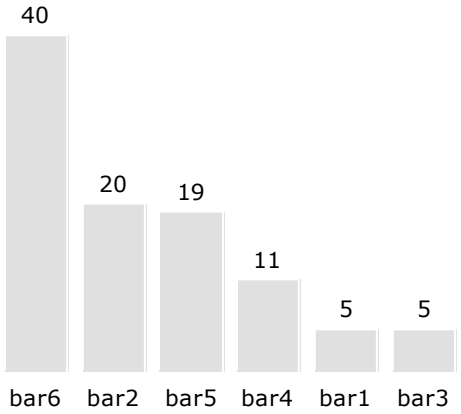


Fig. 5: One Solution When Pie Chart Fails

Descending Horizontal Bar Chart
With Percents Appended to Bar Labels

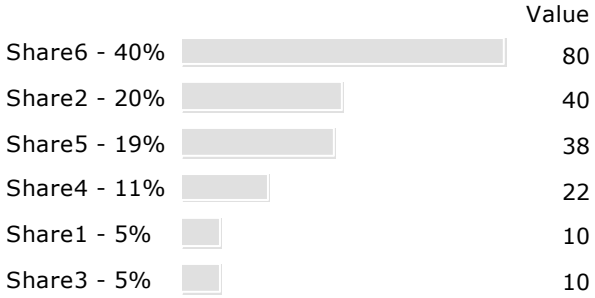


Fig. 6: Pie Chart Fix Using Legend Trick

Imbed Percent & Value in Slice Description

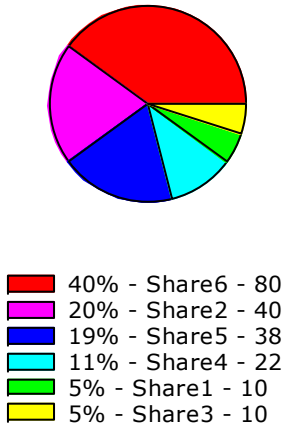


Fig. 7: Default Trend Line

All Possible 3-Point/2-Segment Trend Changes

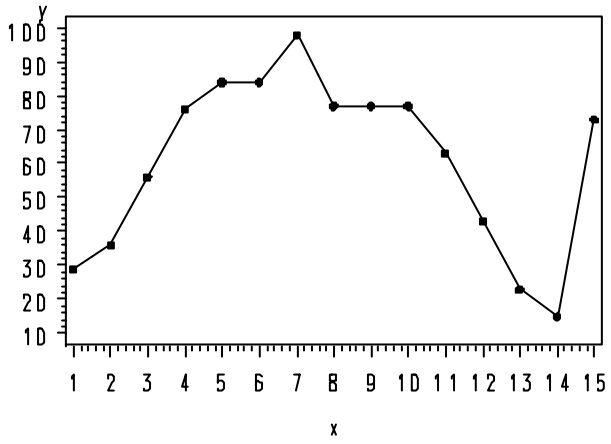
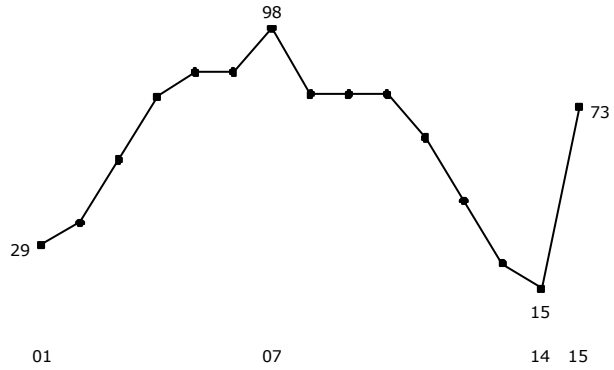


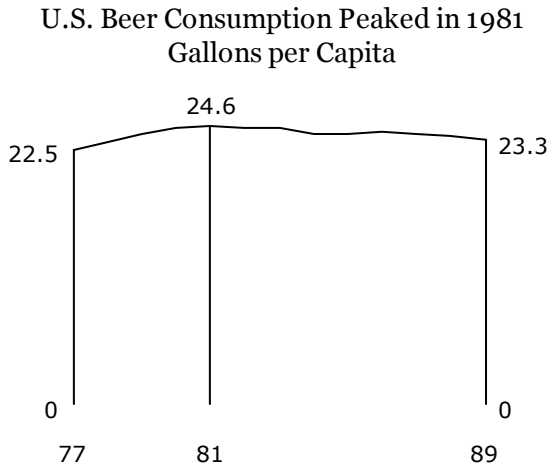
Fig. 8: Sparse Annotation Using a Macro*

Focus is on start, end, maximum, minimum



*TREND Macro also does full annotation
POINTLABEL on SYMBOL statement unusable

Fig. 9: Focus with Sparse Annotation



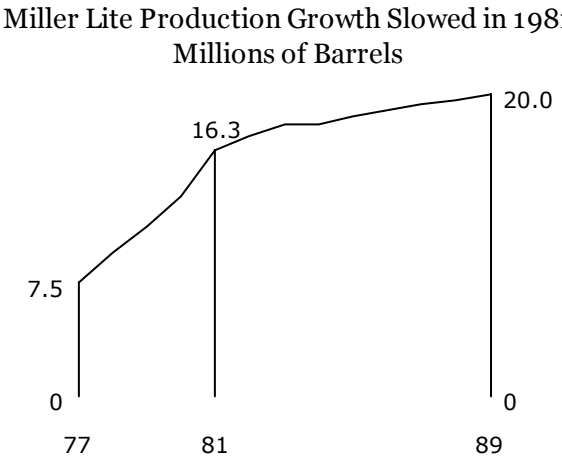
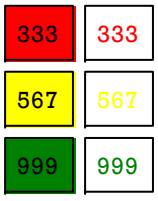
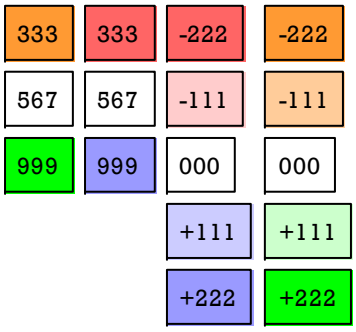


Fig. 10: What Color Should My Data Be?

ODS or Widget Traffic Lighting



Instead, Author Recommends "Flag Lighting" Alternatives*



Common color blindness: Red/Green indistinguishable
For signed data, you can use 2N + 1 colors/categories, still using three hues, but more shades of non-WHITE
*Not every country's flag. What's a better description?

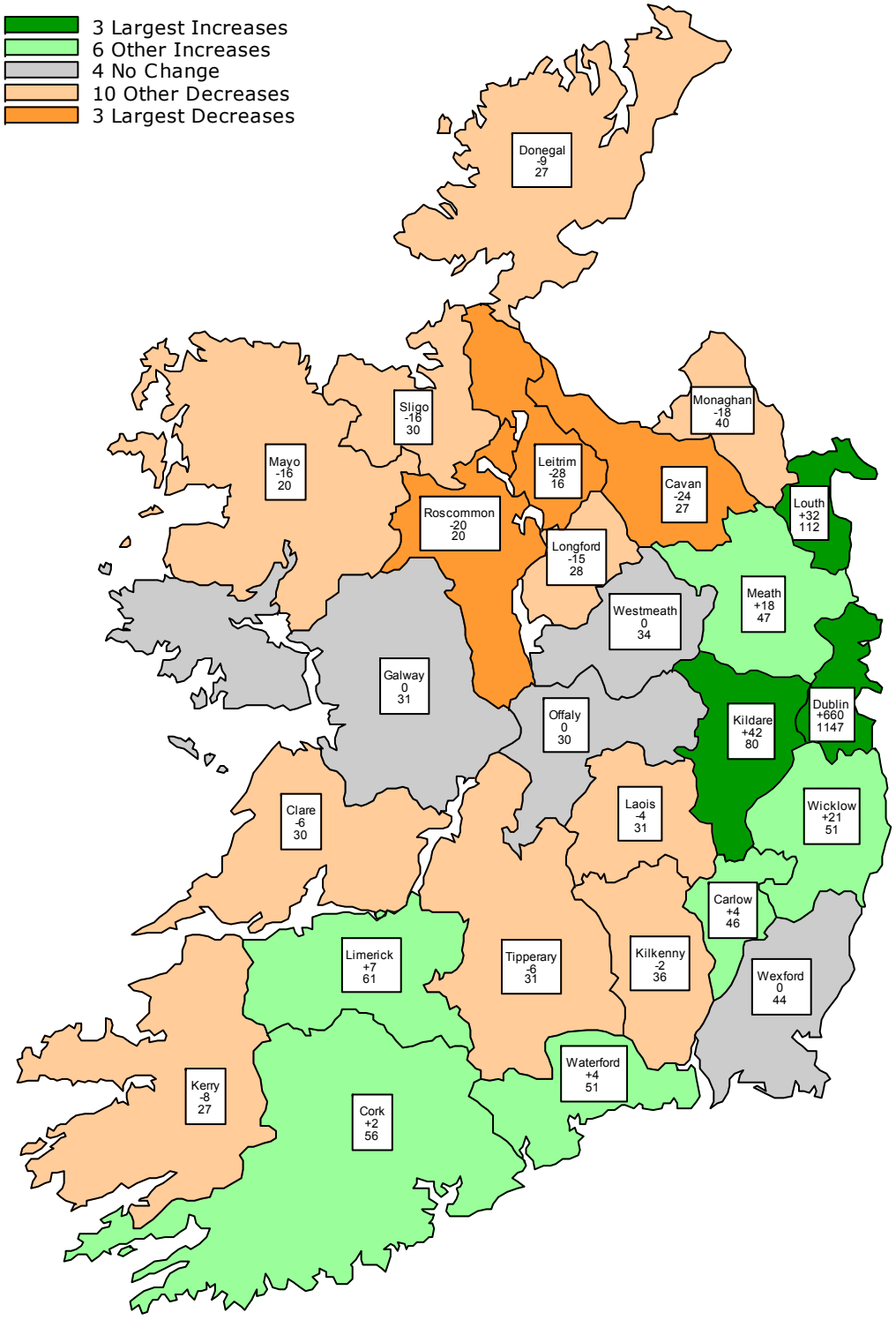
Fig. 11: Samples of Browser-Safe Colors, with Their RGB Codes

Shades of the RGB Primary & Secondary Colors, Grays, & White (CXFFFFFF)
No more than five shades of the same hue are easily distinguished

Other Browser-Safe Colors

Figure 12: Using Color to Communicate

Population Density Change (Per Sq. Km.) in Counties of Republic of Ireland, 1901-1996
A Highlighted "Spatial Table" of County Name, Density Change, & 1996 Density



Other annotation possibilities include numeric rank of density change, or percent of 1996 total population. A macro not only does annotation, but also develops legend text and counts dynamically based on the data. Required macro input parameters include the color palette and the number of areas to be highlighted.