

Is your picture worth a thousand words? Creating Effective Presentations with SAS/GRAPH®

Justina M. Flavin, Pfizer Global Research & Development, La Jolla Laboratories, San Diego, CA
Arthur L. Carpenter, California Occidental Consultants, Oceanside, CA

ABSTRACT

SAS/GRAPH® contains an abundance of procedures and options that can be used to create any number of types of graphical output. And if that is not enough, with the capabilities of the Annotate facility, the sky and your imagination may be the only limits to your graphical creations.

The problem is that all of this power and flexibility may not result in output that conveys the intended message. Graphical presentation of data, when done well can enlighten and inform, but when done poorly can confuse or mislead.

Since effective communication should always be the goal of each graph and chart, this paper will present effective techniques that can be incorporated into SAS/GRAPH programs and will discuss some techniques to be avoided when creating graphical output. Application of these ideas will enable you to design useful and informative presentations that maximize the level of information transfer and ensure that your pictures are truly worth a thousand words.

KEYWORDS

SAS/GRAPH, graphical communication, color, font

INTRODUCTION

In designing graphs, charts, or tables that facilitate effective communication, there are two questions that need to be raised:

- What is the objective of the data display and what information is it to convey?
- What needs to appear on the output to support the objective and enable the viewer to draw the correct conclusion(s)?

In the authors' opinions, the graphical output that is used to answer these questions will possess the characteristics of being simple, consistent, and straightforward. The importance of these characteristics will be discussed as they apply to the three main components in a graph – the text, the symbols and lines, and the colors, fonts and patterns – and the importance of these characteristics when creating multiple graphs. What happens when these underlying principles are violated is presented in the section on creating misleading output.

MANAGEMENT OF TEXT

Every graph should tell a story, so it is necessary to include enough information on the graph to tell that story, but without bogging down the viewer with unnecessary details. Avoid the use of too much text – less is more in this regard.

TITLES AND FOOTNOTES

Text may be added to the graphic output through the use of the TITLE, NOTE, and FOOTNOTE statements, as well as through the use of the ANNOTATE facility. Since up to ten TITLE statements may be specified as well as up to ten FOOTNOTE statements, much text can be added to the graph.

The use of large numbers of TITLE and FOOTNOTE statements is not recommended, since it is the graph and NOT the text that should convey the information. Too much text can distract from the graph in a couple of ways.

- Each title or footnote reduces the area available on the page for the graph itself. Large numbers of title or footnote lines can also cause severe distortion in the vertical aspects of the graph. Each line of text reduces the vertical area available for the graph output, while the horizontal area is unaffected - unless the ANGLE=90 or ANGLE=-90 options are applied - resulting in an elongated graph.

- If more than a few titles or footnotes are felt to be necessary for the observer to fully understand the graphical output, then perhaps the way the data are being displayed needs to be rethought. Text on the display should inform the viewer of something he needs to know about the graph, but it should not be used to explain the output.

After all, if the picture is worth a thousand words then it shouldn't require a thousand words to explain it to the observer.

TEXT SIZE

When multiple title lines are used, these should be displayed in descending level of specificity. So the most general title will appear first, followed by more specific titles. If different font sizes are desired in the title lines, these too should be ordered by descending size.

The graphics options HTITLE and HTEXT can be used to set the default size of the text. The use of these GOPTIONs provide not only a consistent default text size, it may also reduce the number of times that the HEIGHT= option is needed.

ANNOTATE

If text is desired on the graph itself, this can also be added using the ANNOTATE facility. One potential problem with adding text to the graphic area is that there may be overlap of the text and data display itself. Using this option to add text to a data display may add some real value to the interpretation of the data – or may simply add noise and clutter to the output.

AXIS STATEMENT

The AXIS statement allows the user to define and control the appearance of labels, tick mark values, and descriptions for the horizontal and vertical axes. By default, the label for the vertical axis is positioned horizontally at the top of the axis, while the label for the horizontal axis is centered underneath the axis.

Through the use of the ROTATE and ANGLE options, labels, tick mark values, and text descriptions may be rotated so that text is not displayed horizontally, but rather, vertically or on an angle. In general, although angles up to about 55 degrees seem to work well, non-horizontal text is harder to read, so judicious use of these options is advised.

Axis labels should be concise and if applicable, should include the units of the tick mark values. Tick mark values and descriptions should be controlled using the ORDER and VALUE options so that enough are present to define the axis values, but not so many that the axes are cluttered.

If the tick mark values are dates, use a date format that does not lead to ambiguity such as DATE9. The DDMMYY. and MMDDYY. formats are more confusing since dates such as 02/10/2003 can be interpreted as either 10 February 2003 or 02 October 2003. If all dates are from a given year, it may be better to specify this in the axis label so that it is not necessary to repeat the year on each tick mark. Picture formats can be very useful for specialized date/time presentations (Carpenter, 2004).

LEGEND STATEMENT

The LEGEND statement controls the location and appearance of a legend on the output. The LABEL and VALUE options can be used to control the text that appears on the legend. If the LABEL option is not specified, then the variable's label or the variable name is used. The VALUE option can be used to control the text description for each value of the variable that is being plotted. If this option is not used, then the text defaults to the variable's formatted value. In some data displays, a well thought out title along with the text portion of the values may be descriptive enough that a legend label is not necessary.

The LEGEND statement is also used to control the location, size, color, and layout of the legend. Be sure to reduce the legend to its necessary components and then structure it to complement the graph itself.

MANAGEMENT OF SYMBOLS AND LINES

The SYMBOL statement controls the appearance of plot symbols and plot lines in output produced by the GPLOT procedure. If SYMBOL statements are not specified, default definitions are generated by cycling through the default set of colors and plot symbols. Most of the time these defaults will not produce acceptable output and it will be necessary to define SYMBOL statements.

If multiple SYMBOL statements are needed, it is best to use a unique combination of plotting symbols, line types, and colors for each SYMBOL statement. This creates the situation of having a large amount of variation between the different groups of data being displayed on the output.

SYMBOL SELECTION

As much as possible, choose symbol values that are distinctly different in shape and design. For instance, use of "STAR" and ":", or "%" and "&" on the same graph is not recommended, while "CIRCLE" and "DOT" contrast well.

Special Plot Symbols					
Plot Symbol	V=	Plot Symbol	V=	Plot Symbol	V=
.	POINT	⊙	—	*	:
⋯	PAW	☆	=	♂)
Z	Z	♁	'	☾	(
Y	Y	♁	&	♂	/
#	HASH	♁	%	♂	'
△	TRIANGLE	◇	\$	♂	<
◇	DIAMOND	♥	#	♂	>
□	SQUARE	♁	"	♂	+
*	STAR	□		⊕	*
X	X	○	CIRCLE	♀	@
+	PLUS	●	DOT		

Moving away from the default set of symbols by specifying a font with the FONT= option dramatically increases the number of available symbols. The MARKER font for instance contains a number of useful symbols. For users working with one of the Windows operating systems, any of the TrueType® fonts, such as WingDings can also be used. A number of these symbols can be either empty or filled which adds contrast without dramatically increasing the complexity of the graph. Remember the principle of simplicity – it is easy to get carried away with too many silly symbols.

Selected Fonts

SIMPLEX ABCDabcd
 DUPLEX ABCDabcd
TRIPLEX ABCDabcd
 SWISS ABCDabcd
SWISSB ABCDabcd
~~ORIENT~~ ABCDabcd
~~SCISSOR~~ ABCDabcd
BRUSH ABCDabcd

Specialty Fonts (selected values)

S I M P L E X U	A B C D E F G H I J K L M N O P Q R S T U V
C A R T O G	↑ ↓ × ↻ ↺ ↻ ↻ ↻ ↻ ↻ ↻ ↻ ↻ ↻ ↻ ↻ ↻ ↻
M A R K E R	◆ ▲ ▼ → ↑ ↔ ↕ ↘ ↙ ↗ ↚ ↛ ↜ ↝ ↞ ↠ ↡ ↢ ↣ ↤ ↥ ↦ ↧ ↨
M A R K E R E	◀ ▶ ▴ ▽ ↔ ↕ ↔ ↕ ↘ ↙ ↗ ↚ ↛ ↜ ↝ ↞ ↠ ↡ ↢ ↣ ↤ ↥ ↦ ↧ ↨
S P E C I A L	☐ ○ ♥ ◇ ♠ ♣ ♡ ♂ ♀ ☆ ★ ▪ ▲ ✱ ⚡ ⚡ ⚡ ⚡

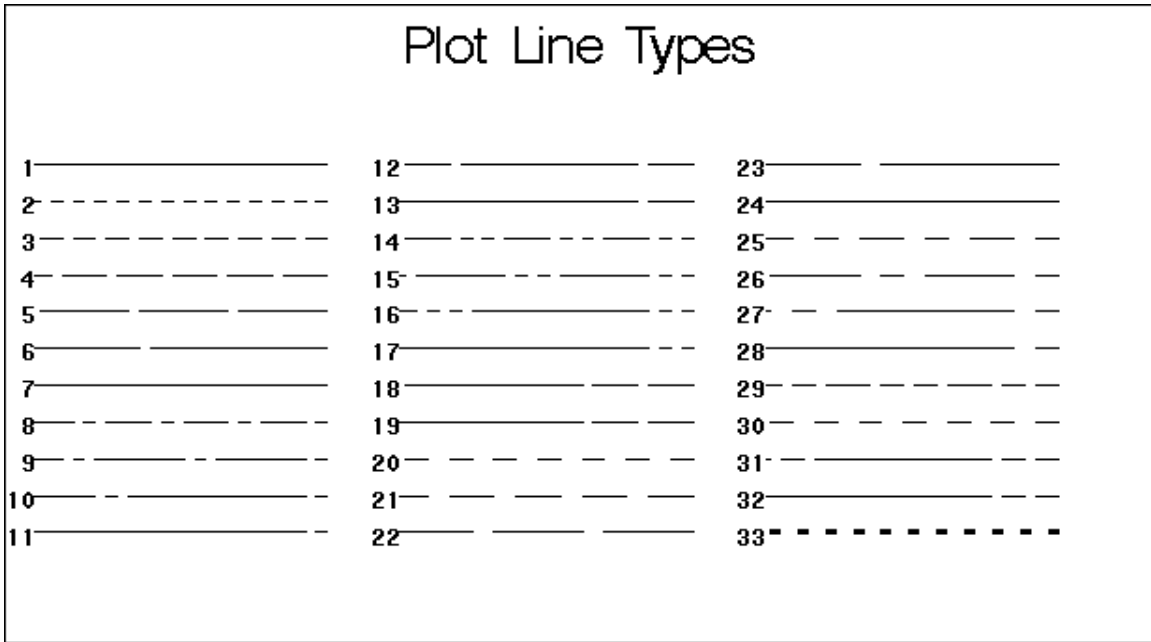
Font values are selected using the above letters (shown in SIMPLEXU)
 e.g. a 'C' in MARKER will show a ▲ and a 'L' in CARTOG will show a ↻

NOTE: Since symbol sizes vary by font these values do not line up vertically

LINE SELECTION

As with symbol selection – choose line styles that are quite different from each other. Although over 40 line types can be specified using the LINE= option, only a few of these are truly different. Most of the line types are a combination of long and short dashes. Under most circumstances the dashes are too long and may look almost solid. The authors have found the following to be the most useful:

- LINE=1 Produces a solid line
- LINE=2 Produces a dashed line - Under some situations (plots with lots of twists and turns) the dashes are too long
- LINE=33 Short dashes (almost dots) - Under some device drivers this line type tends to fade and may need to be widened using the WIDTH= option.



MANAGEMENT OF COLORS, FONTS AND PATTERNS

Given the wide range of choices for colors, fonts, and patterns it becomes necessary to develop some guidelines for their use.

COLOR

The use of color is recommended, but it is important to remember that much output is still produced (or reproduced) in black and white. So even if the original output is to be produced in color, it is still a good idea to create output that can still be meaningful in the absence of color. The use of gray scales is preferred when the graphic is to be rendered in black and white.

While color can be added to just about every aspect of the output – in TITLE, FOOTNOTE, AXES, LEGEND, and SYMBOL statements – caution is advised. Since the viewer's attention should be drawn to the data being graphed or displayed, it is best to reserve most color for use in this area of the graph.

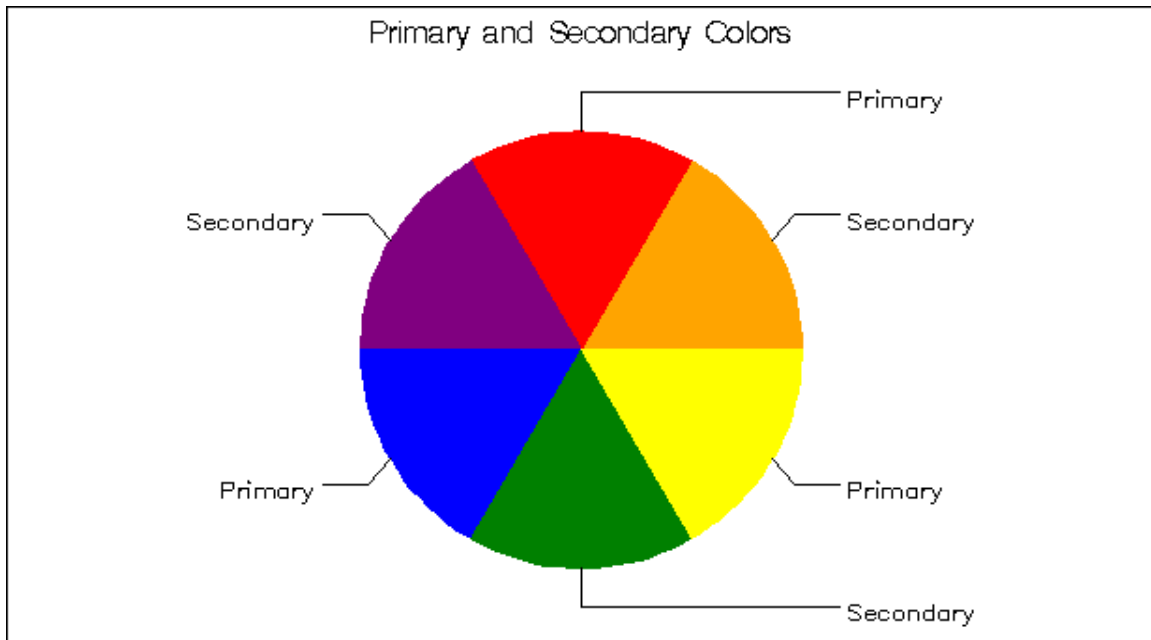
If color is used in the titles and footnotes, all title lines should be of the same color. Likewise, multiple footnotes should all be of the same color. The only exception to this rule might be in the production of a series of graphs. If several title lines that do not change from graph to graph are used, it may be useful to change the color of the title line that does uniquely identify each graph.

In the AXIS statements, different colors may be used for the axis lines, axis labels, and axis tick mark values. Unless the use of different colors for each of these items truly adds additional information which aids in the interpretation of the graph, a good rule of thumb is to stick to one color for all aspects of the axis. Colors can also be used in the label, frame, and values description portion of the legend but this is probably the portion of the graph where it is best to not use color since there may be color in the legend value portion.

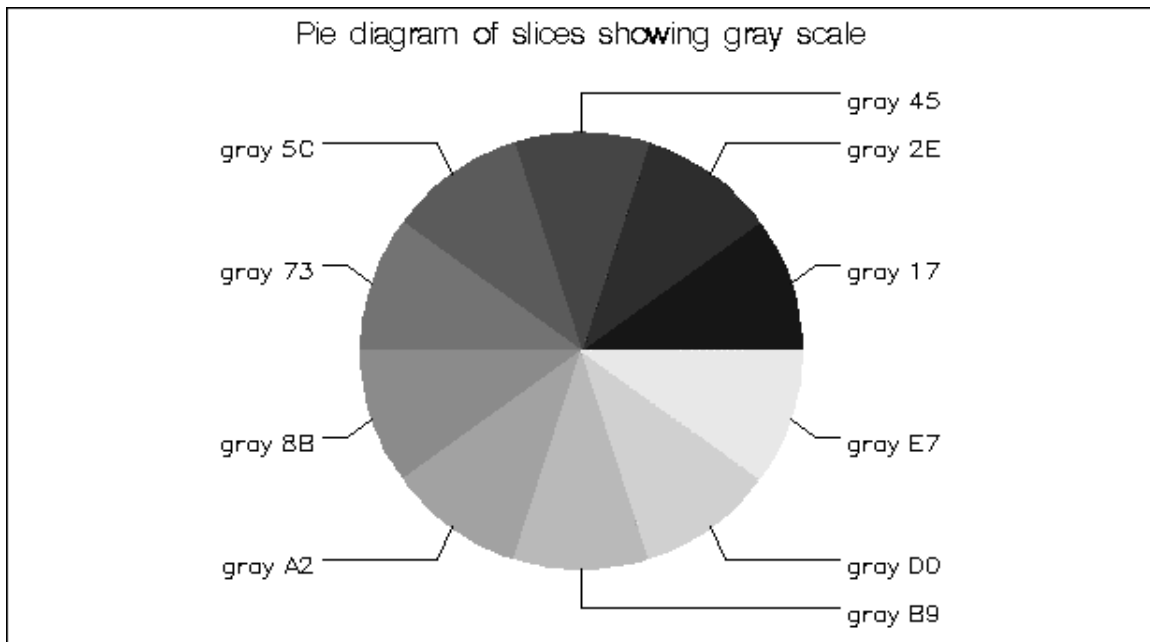
SAS/GRAPH can generate millions of colors, however it should be obvious that there is little advantage to using 'lots' of colors. The human eye can only distinguish amongst a few and subtle color variations can often not be differentiated.

In adding color to the different groupings of data in the graph, it is best to choose a different color for each data grouping and to stick with the primary (red, yellow, and blue) and secondary (orange, green, and purple) colors. Secondary colors are created by combining any two primary colors. A complementary pair is

defined as a primary color and the secondary color that is the combination of the other two primary colors – red and green, yellow and purple, blue and orange. Since complementary pairs provide the highest visual contrast, colors should be chosen to maximize the number of complementary color pairs in the display. An analogous color pair – a primary color paired with one of its secondary colors (red and purple, red and orange, yellow and orange, yellow and green, blue and green, blue and purple) – provides a lower visual contrast than complementary color pairs. If at all possible, assign colors in such a way so that data that are closer together on the display use complementary colors. For data groupings that are further away from each other, the higher contrast is less critical and analogous color pairs are acceptable. It is not advisable to use more than six colors on a graph, since any other color added is nothing more than a tertiary color or hue of a color already being displayed. Both Perry Watts and LeRoy Bessler have published extensively on the topic of the use of colors with graphics.



An alternative to the use of color is that of gray scale. It is possible to generate up to 256 shades of gray, however as was mentioned earlier, the differences are too subtle to be useful. A macro (%PATTERN) found in Carpenter (2004) will generate N maximally contrasting gray scales.



FONTS

Like colors, different fonts may be specified for each aspect of the graph where text is displayed. One, or at most two, fonts should be chosen and all text should utilize this font. Choose fonts that display characters without a lot of embellishments (*sans serif*). Good choices are Swiss or Century. Poorer choices are Old English, Brush, and Script.

Fonts may be either proportional or non-proportional. Most fonts in SAS/GRAPH are proportional, but can be made to be non-proportional by adding the suffix 'U' to the name of the font. In a non-proportional font the lower case letter 'i' will take up the same horizontal space as an upper case 'W'. These fonts (such as SAS Monospace, Courier New, and SIMPLEXU) are useful for generating columns of numbers, but they can waste space when used in titles, labels, or other text.

Always be careful to specify the font to be used. The FTEXT graphics option or the FONT= text option should **always** be used in each TITLE, FOOTNOTE, SYMBOL, and PATTERN statement. When a specific font has not been specified, SAS will be forced to pick one and it may pick what is known as a hardware font. Many of the default hardware fonts tend to be non-proportional and tend to ignore or modify the sizing (H=) options.

PATTERNS

Starting in V7 the default fill pattern became SOLID. In prior versions various combinations of cross-hatching were used. While cross-hatching works well with pen plotters, this type of pattern has a number of problems, some of which are described in the section below on techniques to avoid, and it is appropriate to move away from the use of this type of pattern.

Since most patterns are now either solid or empty, color becomes more crucial to distinguish amongst the patterns. Again consider the use of gray scales if the plot is to be reproduced or copied.

STANDARDIZATION AND CONSISTENCY IN CREATING MORE THAN ONE GRAPH

Often it is necessary to produce more than one graph for inclusion in a report or presentation of the data. Within the report, different types of graphs may be produced for different types of data, but consistency should be maintained in the elements that do not change from graph to graph.

If identical text appears on every graph, these lines should occupy the same position on every page and be of the same size, color, and font on each page. Choosing one font style will also create consistent looking output.

While axes and legends will not be identical on every graph since these are generally data driven, the colors, fonts and text position of these items can be defined using global `AXIS` and `LEGEND` statements to ensure consistency. In particular, for graphs which require a legend, the legend should be positioned in the same place for all graphs.

One good way of ensuring consistency and standardization of graph size and text on the graph is through the use of the `GOPTIONS` statement. The size of the output can be controlled using the `HSIZE/VSIZE` options while the `HPOS/VPOS` options control the size of the text characters appearing on the plots. Default fonts can be specified for text, titles, and by-lines using the `FTEXT`, `FTITLE`, and `FBY` options, default colors using the `CTEXT`, `CTITLE`, and `CBY` options, and default character size using the `HTEXT`, `HTITLE`, and `HBY` options. Any of these default values can be changed by reissuing a `GOPTIONS` statement in the program or changing the option within another statement (such as a `TITLE`, `NOTE`, or `FOOTNOTE` statement).

Creating a series of similar graphs on the same data presents a whole new set of problems. For instance, in creating a series of plots on laboratory data, use of a `BY` line would seem to be an obvious choice to cycle through all lab tests. Generally though, each lab test needs to be further broken down by treatment or dosing group, so that one graph is produced for each lab test-treatment group combination. While the variables for both lab test and treatment group can be used on the `BY` line statement, the graphs generated in this fashion will, for the most part, not have consistent axes across treatment group graphs for the same lab test, since SAS only uses the data in the current graph to determine the range of the axis values, as opposed to all data within that laboratory test. Thus the outcome is usually unacceptable as it makes treatment group comparisons more difficult to observe. A solution to this problem entails the use of macro code to determine appropriate upper and lower values for each lab test, and to then use these values to generate global axes statements that remain the same for each lab test, regardless of the treatment group.

Another common consistency problem can occur when generating plots using the `GPLOT` procedure and using a third variable, such as a treatment or dose group, as a grouping variable. `GPLOT` automatically uses the values of this variable in the `LEGEND`, but very often the designers of the plots will want particular symbols, line types, and/or colors to be specified for the display of specific values of this variable. If the program is subsequently run on a subset of the original data, some values of this discrete variable may be absent in the subset. This can result in an inconsistent representation (symbol, line type, color) of a particular value of the variable in a series of plots. The authors of this paper presented solutions to this problem in Flavin and Carpenter (2004).

Starting in SAS 9, some procedures that support the generation of presentation quality graphics - especially those in SAS/STAT® - will be able to take advantage of graphics styles. These styles are coordinated through ODS and lend a consistency to the resulting graphs. Refer to Rodriguez (2004) for further details.

TECHNIQUES TO AVOID

A great deal has been written on what makes a good graph. The following examples are based on Carpenter and Shipp (1995) and Carpenter (1994). Most of the following examples highlight things that are for the most part just common sense. However, the problems that they represent appear in the literature often enough to warrant inclusion in this paper.

The following lists in decreasing order of our ability to discriminate various graphical elements (Cleveland, 1984):

- 1 Position along a common scale
- 2 Position along nonaligned scales
- 3 Length, direction, and angle
- 4 Area
- 5 Volume and curvature
- 6 Shading and color saturation

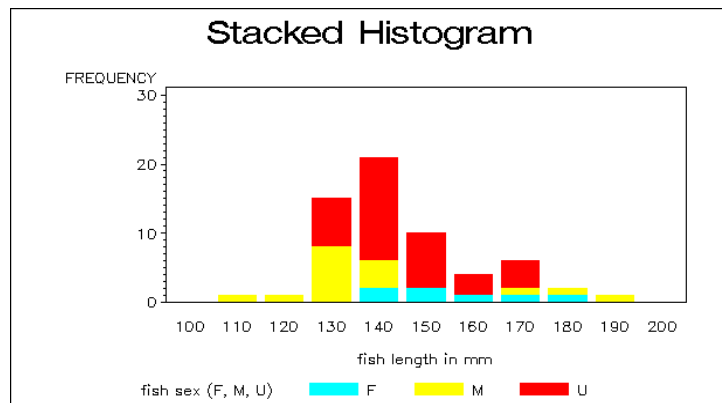
As the numbers indicate, there is a gradation in how well we can interpret graphics which utilize the above elements. The values of two points positioned using a common scale might be easy to compare, while the same values would probably be more difficult to compare if color shading was used to distinguish them.

SAS/GRAPH can easily be used to produce graphics that ignore the failings of our minds.

STACKED HISTOGRAMS

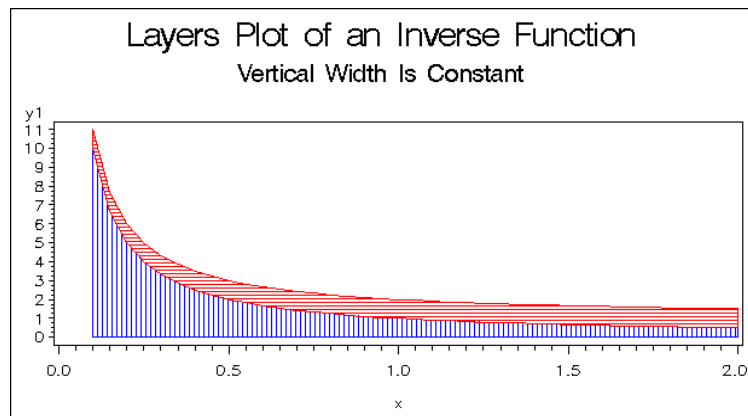
Stacked bar charts are used to show the values of two or more variables (or categories) on each bar. Both PROCs CHART and GCHART allow the use of the SUBGROUP= option to easily create this kind of graphic. The problem with this type of chart is that the stacked values are perceived as being on a common scale, when actually the scales are nonaligned. Not only are they nonaligned, but only the first set of vertical values have a common base.

Because the stacked segments do not have a common base it is difficult to compare values such as the number of Males in the 130mm category and the number of fish with Undetermined gender in the 150mm category.



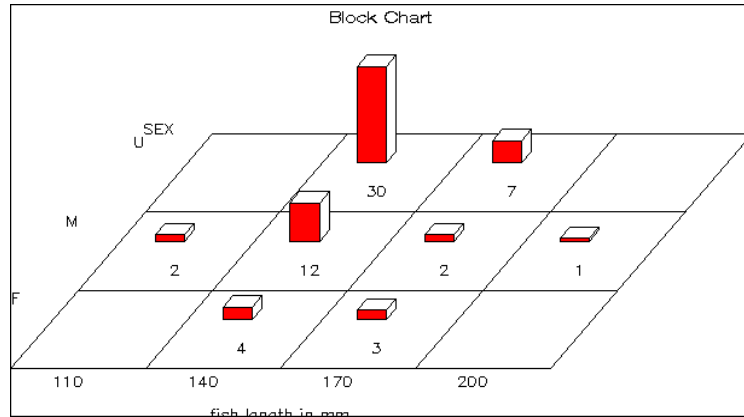
PLOTS IN LAYERS

The PROC GPLOT equivalent to the stacked values discussed in the previous example uses the AREAS= option in the PLOT statement. Unlike the stacked histograms, the AREAS= option is designed to plot two or more variables that always take on monotonically greater than values (i.e. for any value of the horizontal variable $y_1 < y_2 < y_3$ etc.). While this provides an aligned (vertically and horizontally) scale, our eye perceives a non-aligned scale. The plot shows two curves with a constant vertical difference, however the use of the AREAS= option (aggravated by the choice of PATTERNS) makes it very difficult to see the constant difference. Our eye sees the horizontal spacing as well, thereby totally hiding the constant difference.



BLOCK CHARTS

Block charts are created by using the BLOCK statement in PROC GCHART. SAS will create a series of 'pseudo' three dimensional towers with the height of each representing the value of the selected variable. When two classification variables are selected, the towers are arranged like a small city of skyscrapers.

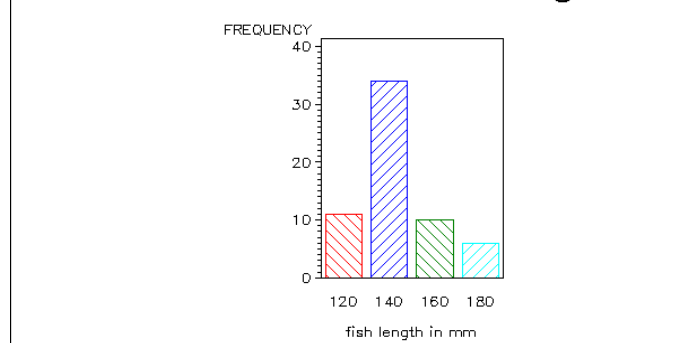


In addition to being a problem of nonaligned scales, our eye wants to see a three-dimensional picture, and will attempt to use perspective where none has been added. Because our eye wants to see the tower in a back row as being farther away, this can cause a tower in the back row to appear taller than one of equal height in the front row. Notice that the parallel grid lines running from the front to the back seem to diverge. This is an optical illusion, an artifact of the eye trying to put the graph into perspective.

OPTICAL ILLUSION AND THE PATTERN STATEMENT

Some combinations of lines and patterns just do not go well together. As we design our graphic and choose our fill patterns, we must guard against occurrences of charts such as the one shown here (this is another good reason to use SOLID and EMPTY as your pattern type).

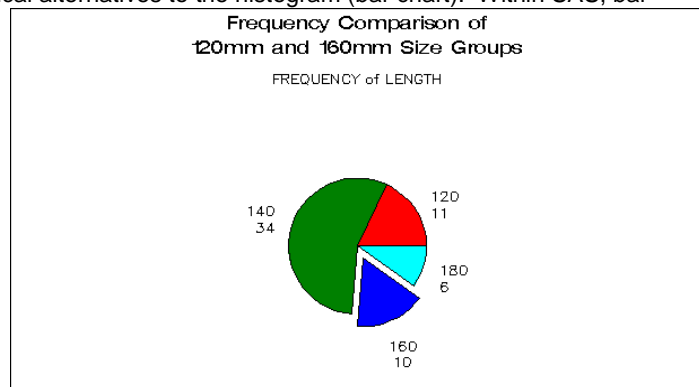
Patterns Seem to Create Leaning Towers



In this chart, although the vertical bars are parallel and equally spaced, they appear to be leaning toward or away from their neighbor.

PIE CHART

The pie chart is one of the classical graphical alternatives to the histogram (bar chart). Within SAS, bar charts and pie charts are created by the same PROCedures. Pie charts are not necessarily bad, but the programmer (and wouldn't it be nice, the user of the graphic) needs to be aware that our minds can more accurately distinguish the length of bars with a common scale than we can the area of a circular sector. Exploded pie charts (EXPLODE= option in the PIE statement) are even worse because they trick the eye into seeing too much area.

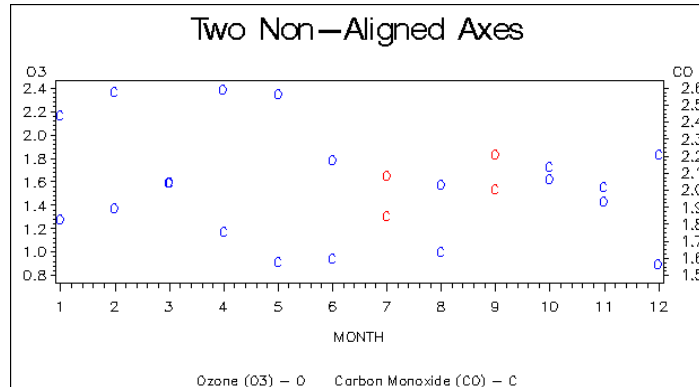


It is for this reason that some readers may perceive the exploded section (160mm) in the lower right corner to be larger than the 120mm section when in fact it is not.

NONALIGNED AXES

The PLOT2 and BUBBLE2 statements in PROC GPLOT allow the user to specify a second vertical axis when overlaying two variables with a common horizontal axis. These two axes may or may not have values in common, however, our eye will tend to want to compare the magnitude of two values by looking at their height above the horizontal axis rather than by seeking out their true values.

The non-aligned axes in this example make it difficult to compare concentrations of ozone and carbon monoxide. In July (month=7) and September (month=9) it first appears that the O₃ concentration is greater since it is higher on the plot. A closer inspection, however, shows that this is not the case. Our eye has been fooled by the non-aligned axes.

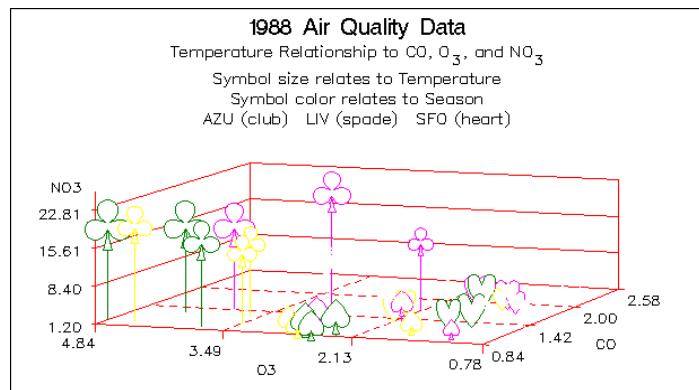


BUBBLE PLOTS ETC.

The BUBBLE and BUBBLE2 statements in PROC GPLOT, and the SCATTER statement with the SIZE= option in PROC G3D produce symbols whose size represents the magnitude of another variable. Unfortunately, the ability to compare areas is not very high on the list of accurate graphical tasks, consequently effective size comparison is difficult.

At least with the BUBBLE and BUBBLE2 statements, comparisons are made among symbols that have the same shape. In a PROC G3D

SCATTER plot, different shapes can be used to add another dimension to the plot, but our eye does not easily compare the volumes of a heart, spade, and star. The problem is compounded by the 'pseudo' three-dimensional effect as was discussed above.



INCLUSION AND EXCLUSION CRITERIA

Avoid the use of 3-D effects and shadowing. While these may appear to make output look more sophisticated, the reality is that they cause more confusion than enlightenment.

In clinical trials data, graphs are often produced on individual subject data or on data pooled by treatment group. Individual subject plots can be useful for identifying subjects with extreme outlier values. If each subject's data are plotted on a separate graph, this can yield a large number of graphs. Conversely, if the data for all subjects are presented on one plot, it may be impossible to distinguish one subject's data from another. Graphs displaying results grouped by treatment group most often present the information in the most useful format.

Data that include extreme outlier values can cause problems in generating graphical output. If these values are included and the user allows the procedure to internally determine axes scales, the output generated may be unacceptable due to the non-outlier data becoming clumped in a small part of the graph. One solution is to exclude extreme outliers from the graph, but to then list these excluded values somewhere on the output. If it is important to include these on the graph, then creating a broken axis using the ANNOTATE facility may be a possible solution.

There are many different procedures available in SAS/GRAPH for creating different types of output. Sometimes there are several different types of displays that could potentially be used, so it is necessary to assess the available choices. For instance a bar chart and a pie chart present two different ways of

displaying the same information. When more than one option exists, it is often necessary to consider the audience, the message to be conveyed, and where the output will be seen or presented.

Another consideration to weigh is whether or not a graphical display truly adds value to the data being presented. As a general rule, a SAS/GRAPH program usually takes more development time to produce acceptable output than the development of a program to generate summary statistics or tables. If resources are an issue and the same conclusions can be easily drawn from tabular output, then perhaps a graphical display is not necessary.

CONCLUSION

Effective communication does not necessarily mean pretty. At the same time an effective graph or chart will by necessity not be ugly. The ideas presented in this paper have hopefully presented some ideas to help the reader differentiate between 'effective' and 'pretty' as well as the aspects of the elements that form an attractive and effective graph or table.

Several of the papers listed in the references provide additional perspective on the general topics of best graphical practices. One of the cited authors, LeRoy Bessler, has written a number of additional papers on the topic.

The standards within companies, groups, and projects vary, however the principles outlined here will almost always apply. There are times when graphs are meant to NOT present the information accurately or in a way that can be completely understood. Assuming that is not the case, remember the story that the graph is meant to tell and then boil the graph down to its essentials. What is the message and what is needed on the graph to convey that message? It is not enough to be able to transfer what is being visualized into SAS code. Always be aware that the choice of graphical elements in constructing the graph will determine how those who look at the graph will interpret what they see.

REFERENCES

Bessler, LeRoy, "*Easy, Elegant, and Effective SAS® Graphs: Inform and Influence with Your Data*", Proceedings of the 28th SAS User Group International Conference, Cary, NC: SAS Institute Inc., 2003.

Bessler, LeRoy, "*Communication-Effective Use of Color for Web Pages, Graphs, Tables, Maps, Text, and Print*", Proceedings of the 29th SAS User Group International Conference, Cary, NC: SAS Institute Inc., 2004.

Carpenter, Arthur L., "*Techniques to Avoid: What Your Momma Did not Tell You About SAS/GRAPH®*", Proceedings of the 19th SAS User Group International Conference, Cary, NC: SAS Institute Inc., 1994.

Carpenter, Arthur L. and Charles Shipp, *Quick Results with SAS/GRAPH® Software*, Cary, NC: SAS Institute Inc., 1995, 249pp.

Carpenter, Arthur L. and Richard O. Smith, "*Controlling Graph Size: Building Thumbnails and GIF Files Using SAS/GRAPH®*", Proceedings of the 29th SAS User Group International Conference, Cary, NC: SAS Institute Inc., 2004.

Carpenter, Arthur L., "*Building and Using User Defined Formats*", Proceedings of the 29th SAS User Group International Conference, Cary, NC: SAS Institute Inc., 2004.

Carpenter, Arthur L., *Carpenter's Complete Guide to the SAS® Macro Language, 2nd Edition*, Cary, NC: SAS Institute Inc., 2004, 476pp.

Cleveland, W. and R. McGill, "Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods", *Journal of the American Statistical Association*. Volume 79. Number 387. pp. 531-554. 1984.

Cleveland, William S., *Visualizing Data*, Hobart Press, Summit, NJ, 1993, 360pp.

Cleveland, William S., *The Elements of Graphing Data*, Hobart Press, Summit, NJ, 1994, 297pp.

Flavin, Justina M. and Arthur L. Carpenter, "*Is the Legend in Your SAS/GRAPH® Output Telling the Right Story?*", Proceedings of the 29th SAS User Group International Conference, Cary, NC: SAS Institute Inc., 2004.

Rodriguez, Robert N., "*An Introduction to ODS for Statistical Graphics in SAS® 9.1*", Proceedings of the 29th SAS User Group International Conference, Cary, NC: SAS Institute Inc., 2004.

Watts, Perry, "*Working with RGB and HLS Color Coding Systems in SAS® Software*", Proceedings of the 28th SAS User Group International Conference, Cary, NC: SAS Institute Inc., 2003.

Watts, Perry, "*New Paletts for SAS 9 Color Utility Macros*", Proceedings of the 29th SAS User Group International Conference, Cary, NC: SAS Institute Inc., 2004.

TRADEMARK INFORMATION

SAS, SAS Certified Professional, and SAS/GRAPH are registered trademarks of SAS Institute Inc. in the USA and other countries.

® indicates USA registration.

CONTACT INFORMATION

Justina M. Flavin
Pfizer Global Research & Development, La Jolla Laboratories
10777 Science Center Drive (B85)
San Diego, CA 92121
(858) 622-7376
justina.flavin@pfizer.com

Justina Flavin is a Manager in the Clinical Operations Strategy Implementation Department at Pfizer Global Research & Development in San Diego. She currently teaches SAS programming classes for the University of California, San Diego Extension – Bioscience Division and has taught seminars at the PharmaSUG and WUSS conferences. Justina served as the conference chair of PharmaSUG '99, was the facilities and seminar course coordinator for WUSS 2000, and serves on the EC of SANDS, the San Diego SAS User Group. Justina has a B.A. in Applied Mathematics from the University of California, San Diego.

Art Carpenter
California Occidental Consultants
PO Box 430
Oceanside, CA 92085-0430
(760) 945-0613
art@caloxy.com
www.caloxy.com

Art Carpenter's publications list includes three books and over four dozen papers and posters presented at SUGI and other SAS user group conferences. Art has been using SAS since 1976 and has served in leadership roles in various local, regional, and national user groups. Art is a SAS Certified Professional™. Through California Occidental Consultants he teaches SAS courses and provides contract SAS programming support nationwide.