

**Paper 5000-2020****Visual Data Insights:  
Image and Ranked Precise Values for Quick, Easy, and Reliable Inference**

LeRoy Bessler, Bessler Consulting and Research

**ABSTRACT**

This paper provides eight widely applicable graphic examples and the embodied design principles for effective visual communication that apply to any graph and any software. Learn how to deliver the best results with Output Delivery System (ODS) Graphics when doing trend or time series plots, horizontal or vertical bar charts, pie charts, and choropleth maps (maps in which geographic unit areas are color-coded to identify their category or data range). As extras, two easy-to-create and easy-to-interpret special cases are included with exhibits and code in a handout, along with a reference list of design guidelines. Even if you are just a requestor or **viewer of other people's data visualization, come to learn what to ask for and what to expect.** This paper assumes no prior knowledge of ODS Graphics.

**INTRODUCTION**

The purpose and scope of this paper are as described in the abstract. I have been trying to get the most communication-effective results from SAS® graphics tools (first SAS/GRAPH®, now ODS Graphics) since 1981, and sharing my graphics experience and knowledge since 1987. Whether you are simply a person who specifies what you would like created as a graph, or a newcomer to ODS Graphics, or an experienced user, I hope that you will find something useful here.

ODS Graphics supports a wide variety of types of graphs, charts, and plots, but the focus of this paper is on those which I deem to be the most frequently used ones for business or organizational reporting or by media. Some of the examples demonstrate design features that are not common, but which add special communication value. They are a subset of the current state of my journey on a quest to get beyond the defaults to what I deem to be the best that SAS graphics can deliver for communication-effective data visualization.

Though I use SAS software for a variety of purposes, visual communication is a use that I especially enjoy. In this era of Data Science, I think of myself as a Data Artist.

The focus of this paper is on principles and examples. A zip file of adequately commented code for the examples can be requested via email. For one example, a reference is provided in this paper for previously published code.

**KEY DESIGN PRINCIPLES DEMONSTRATED HERE**

Published elsewhere as my Principia **Graphika** & my Principia **Color** ("color" is the Latin word for the American English word "color" and the British English word "colour") are longer lists of design principles for communication-effective use of graphics and communication-effective use of color. Below is a subset of them that are particularly relevant here. Feel free to bring to my attention any instances where these standards are not met in this paper.

- Readability is a fundamental design requirement
  - Assure high contrast between text/numbers and color background
- Provide precise numbers, whenever technically feasible
  - Image = quick, easy inference
  - Precise Numbers = reliable, accurate inference

- **Show Them What's Important**
  - Use Ranking
  - Use Subsetting (see Figure 10 and Reference 1 for one example)  
Let Part Stand for the Whole (When Sufficient)
- For Reliable Color Distinguishability
  - Text and Lines must be thick enough
  - Plot Markers and Legend Samples must be big enough
  - Discrete colors, not color gradients, must be used
- Avoid red and green for color-coding in the same image  
Red-Green is the commonest color blindness
- No Vertical Axis Labels or Horizontal Axis Tick Mark Values  
We read left to right—make all text easy to read
- Suppress, whenever inessential, graphic paraphernalia
  - axis values
  - tick marks
  - axis lines
  - axis labels if obvious from axis values (if present) or if explained in title/subtitle
- The suppression recommended above is not some purist matter of minimalist design. It is IMPOSSIBLE to reliably determine precise numbers by comparing point locations (or bar ends) to axis values. Running your eye from a plot point or a bar end to an axis, and mentally interpolating the approximated corresponding point on the axis to estimate a number based on the framing tick mark values is a futile, unreliable, unacceptable way to attempt to get the precise value. For numbers, accuracy and precision are the keys to knowledge and truth and reality.

## **WHY ARE THERE NINETEEN FIGURES FOR EIGHT EXAMPLES?**

Some of the figures are bad examples to be replaced with my recommended alternatives. In one case, the good example comes in both a full-color and a black-and-white version which requires a slightly different design. In another case, two different ways are offered to accomplish the same good outcome. And with Pie Charts, there is not just one good solution, and all of them need to be shown as alternatives to less desirable, or even downright disastrous, design choices that might be made. So, though I might have over-delivered on my promise, I think that I have done no harm.

## TIME SERIES OR TREND PLOTS

### COLOR MULTI-LINE PLOT

There is no XAXISTABLE (or YAXISTABLE) feature in SAS/GRAPH. I did this same thing many years ago with some tricky SAS/GRAPH coding. I am grateful that this solution is available in ODS Graphics.

If there are too many X values and/or they are too many digits, this can become infeasible. However, with a small enough font and a wider display area (e.g., on a web page and a wide screen high-resolution monitor), you can fit a lot. I once did something like this successfully for a seven-line plot that spanned 37 months (latest reporting month, back three years to the same month).

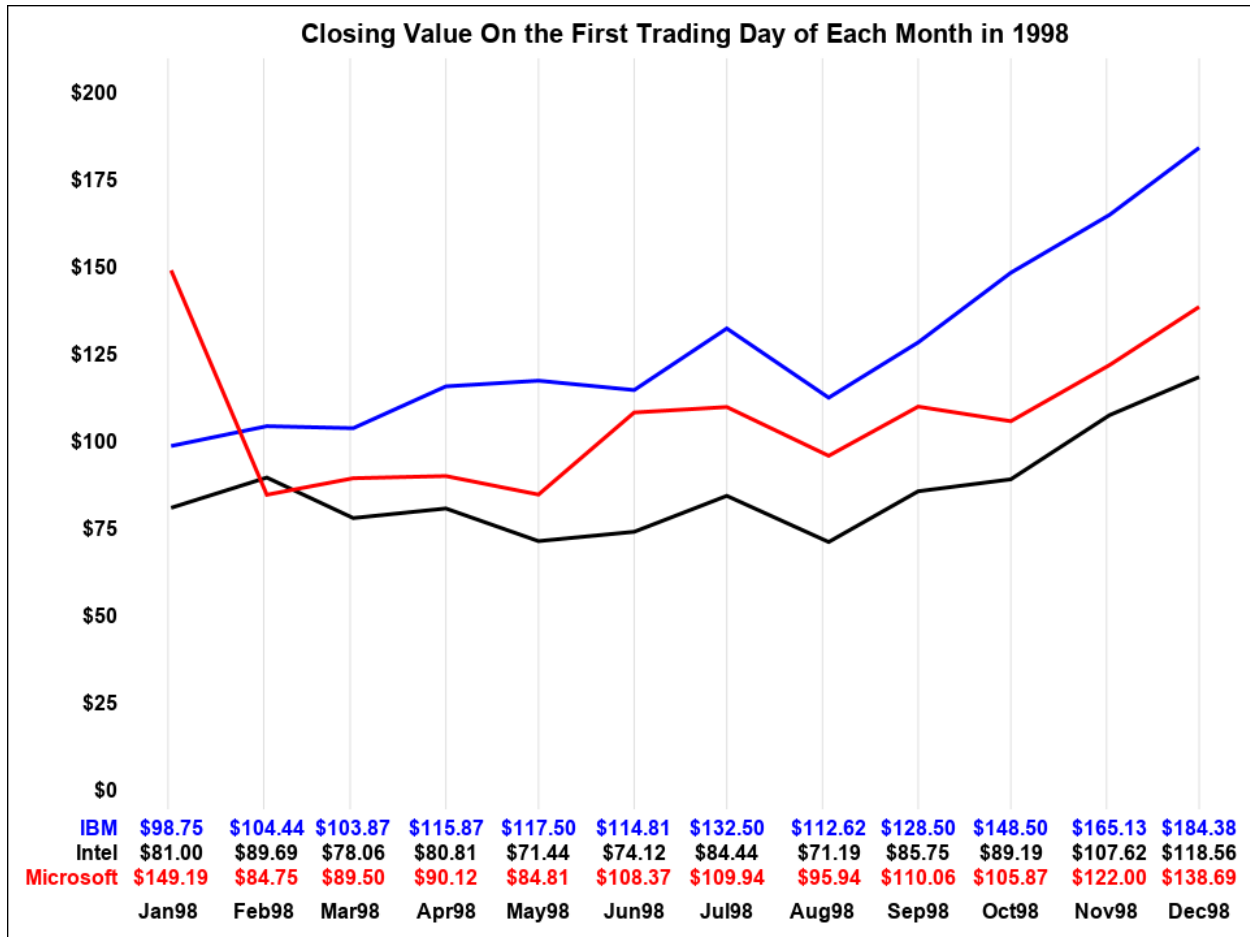
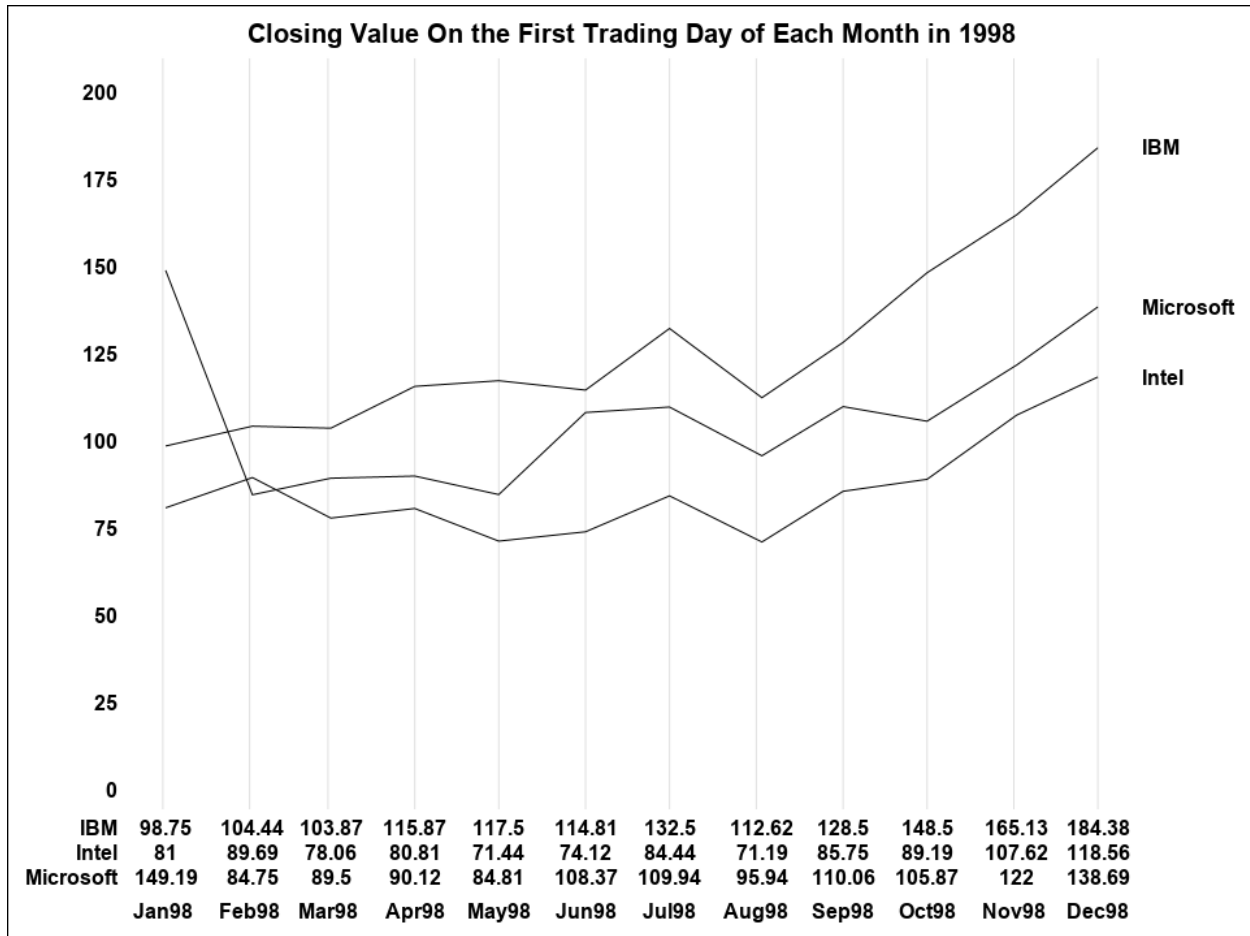


Figure 1. Color-coded table eliminates any need for a legend.

## BLACK AND WHITE MULTI-LINE PLOT



**Figure 2. Using the ODS Graphics CURVELABEL option.**

For black & white output, use the CURVELABEL option so that the viewer's eyes don't need to travel from a line to a legend to discover the line's identity. If curve labels collide, use a plot markers and a legend.

## VERTICAL BAR CHART FOR CATEGORICAL DATA

(A vertical bar chart is also sometimes used for time series data. If the bars are few enough, their ends can be directly labeled with values.)

### SUPPLYING THE TOTAL IN THE TITLE WITH A MACRO VARIABLE

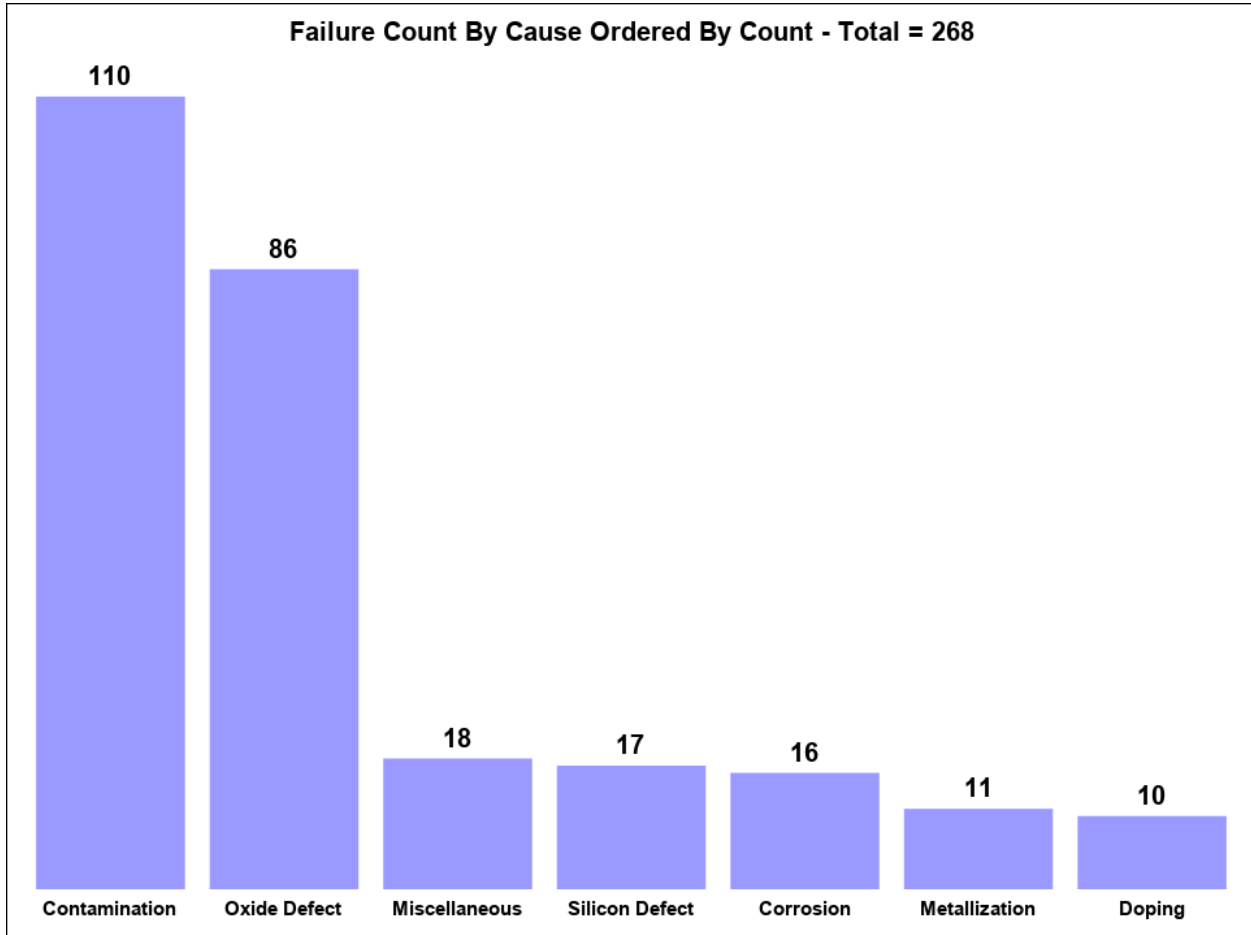
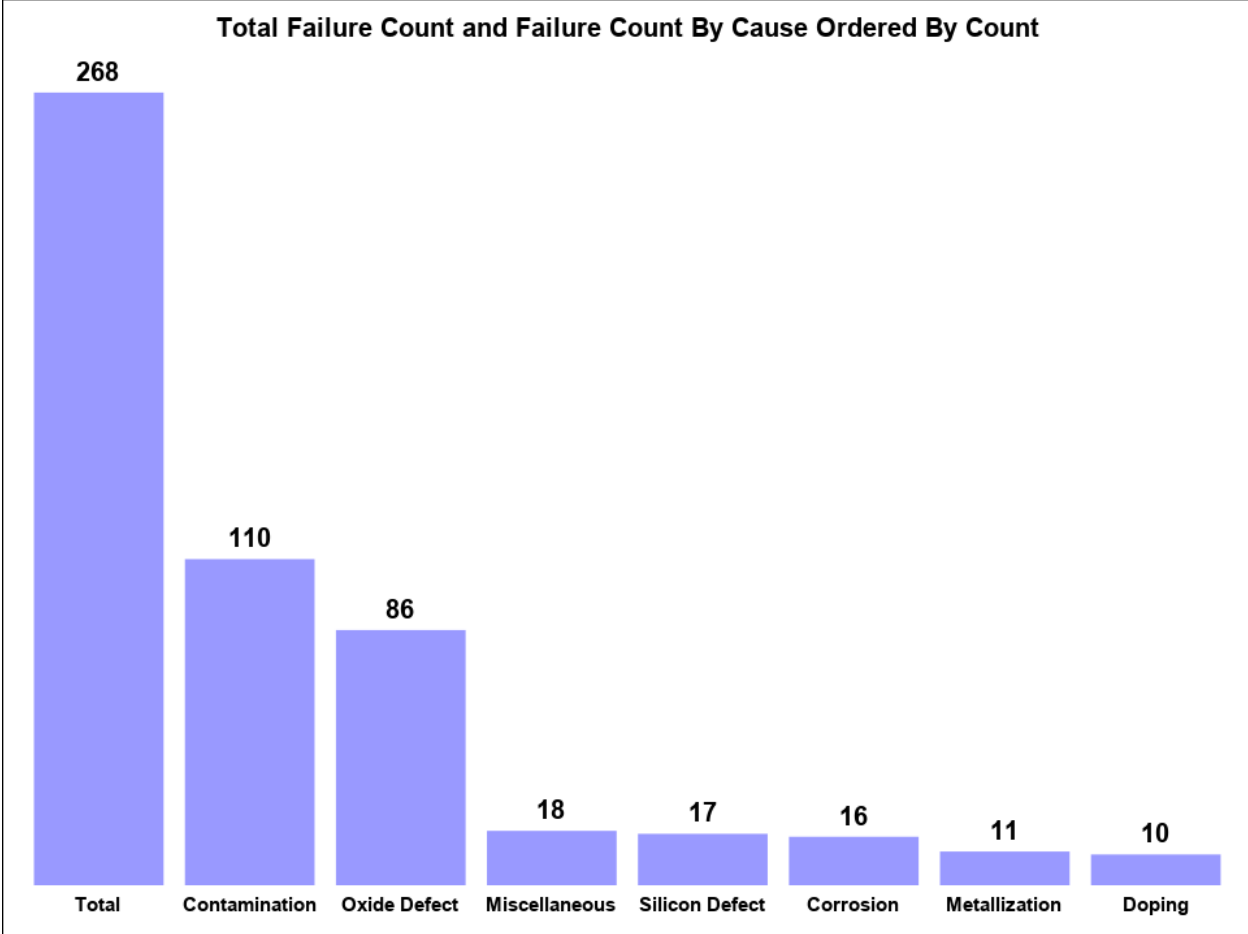


Figure 3. Easily Coded

**SUPPLYING THE TOTAL AS AN EXTRA BAR**



**Figure 4. More Complicated Coding, and Undesirable If the Total Bar Dwarfs the Constituent Bars**  
If the constituent bars all become tiny, their lengths can be less distinguishable.

## IMPROVING THE WATERFALL CHART

VERY popular, even if its advantage over a traditional vertical bar chart that includes a Total bar (like Figure 4) is not apparent to me, but that's just a personal problem, I admit.

### STANDARD WATERFALL CHART

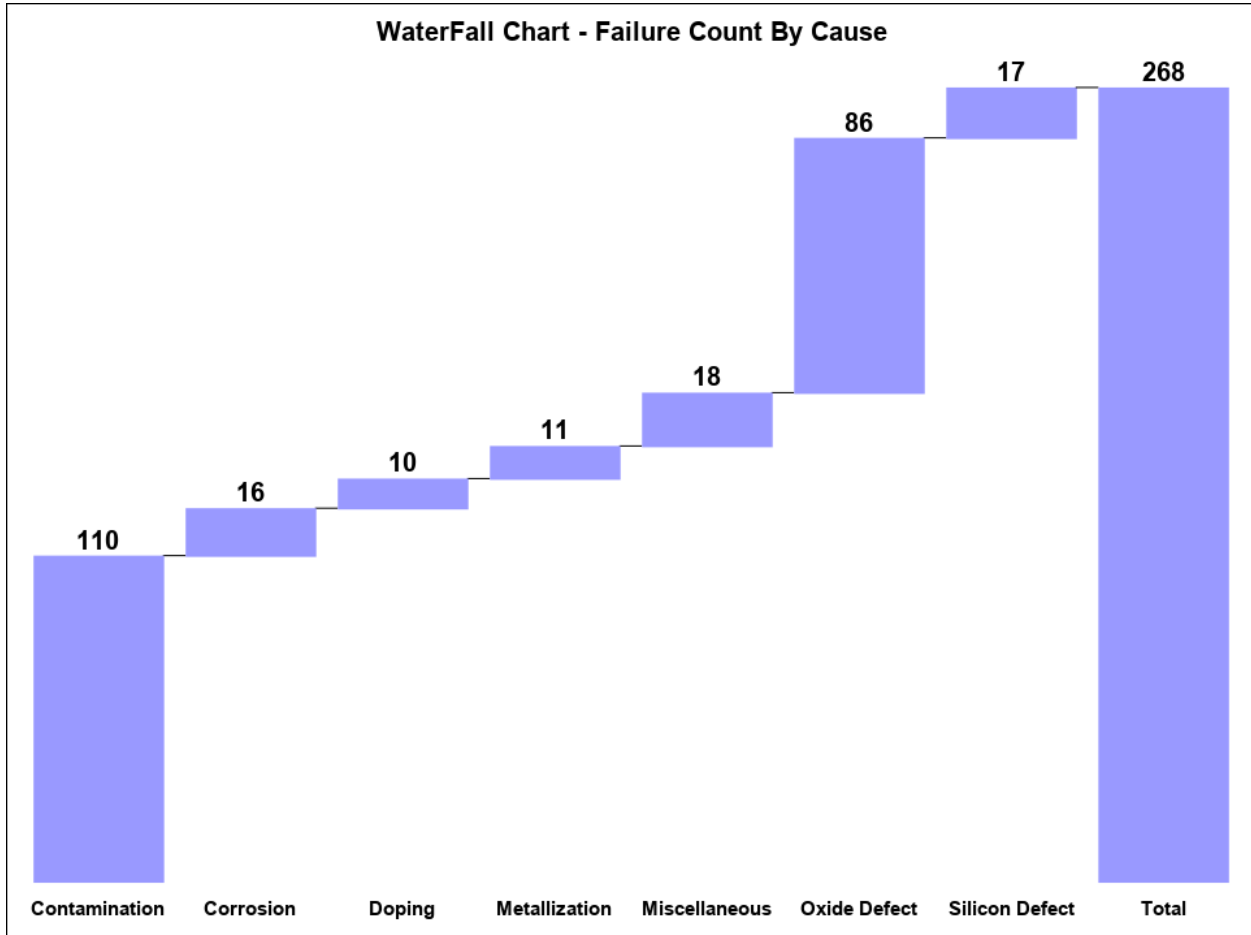
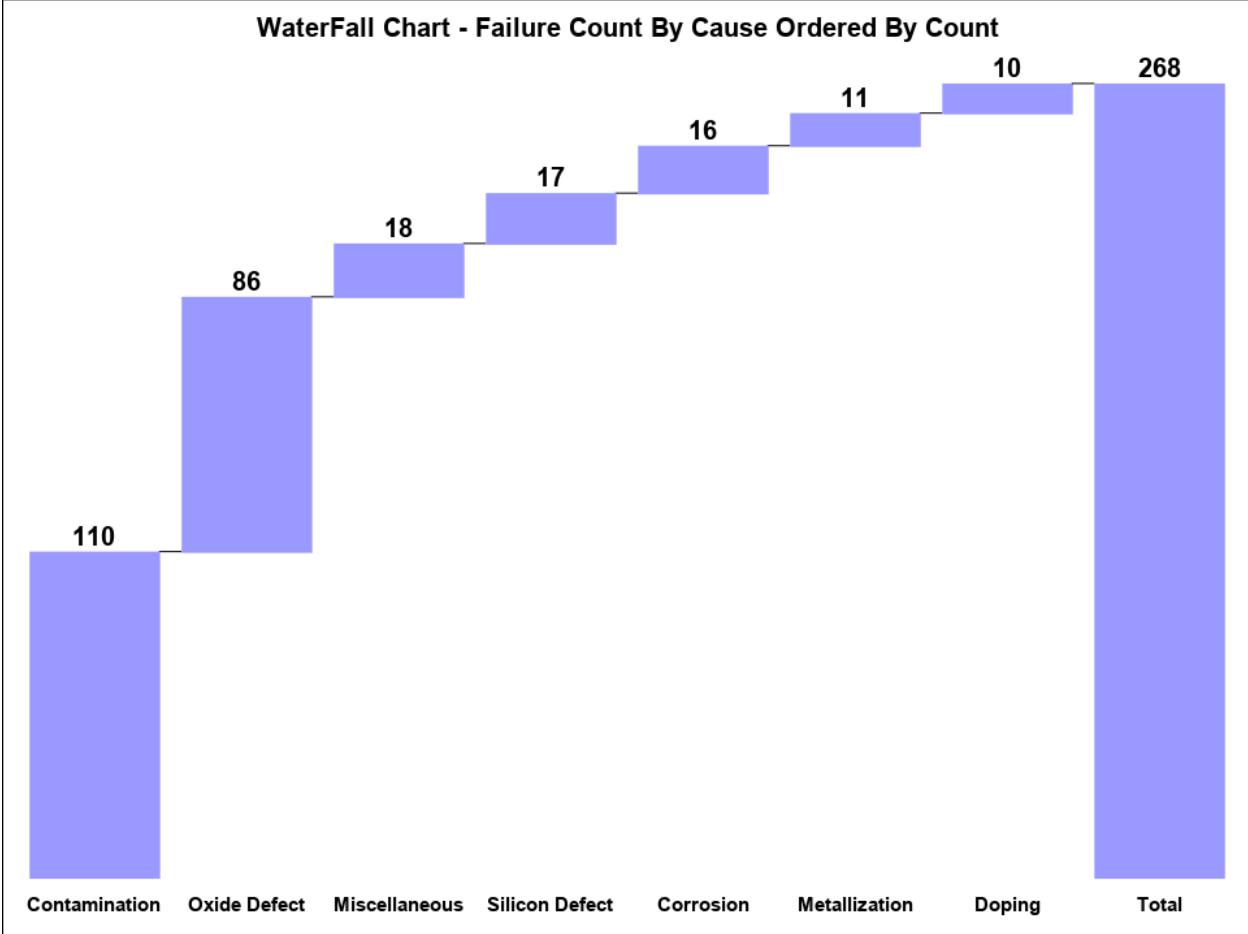


Figure 5. The Constituent Bars Are Ordered By Category Label

**ORDER THEM BY VALUE—THAT’S WHAT MATTERS!**



**Figure 6. More Useful, But Requires Complex Preprocessing. Do a Vertical Bar Chart Instead!**



## STACKED BAR CHARTS

### UNUSABLE STANDARD STACKED BAR CHART

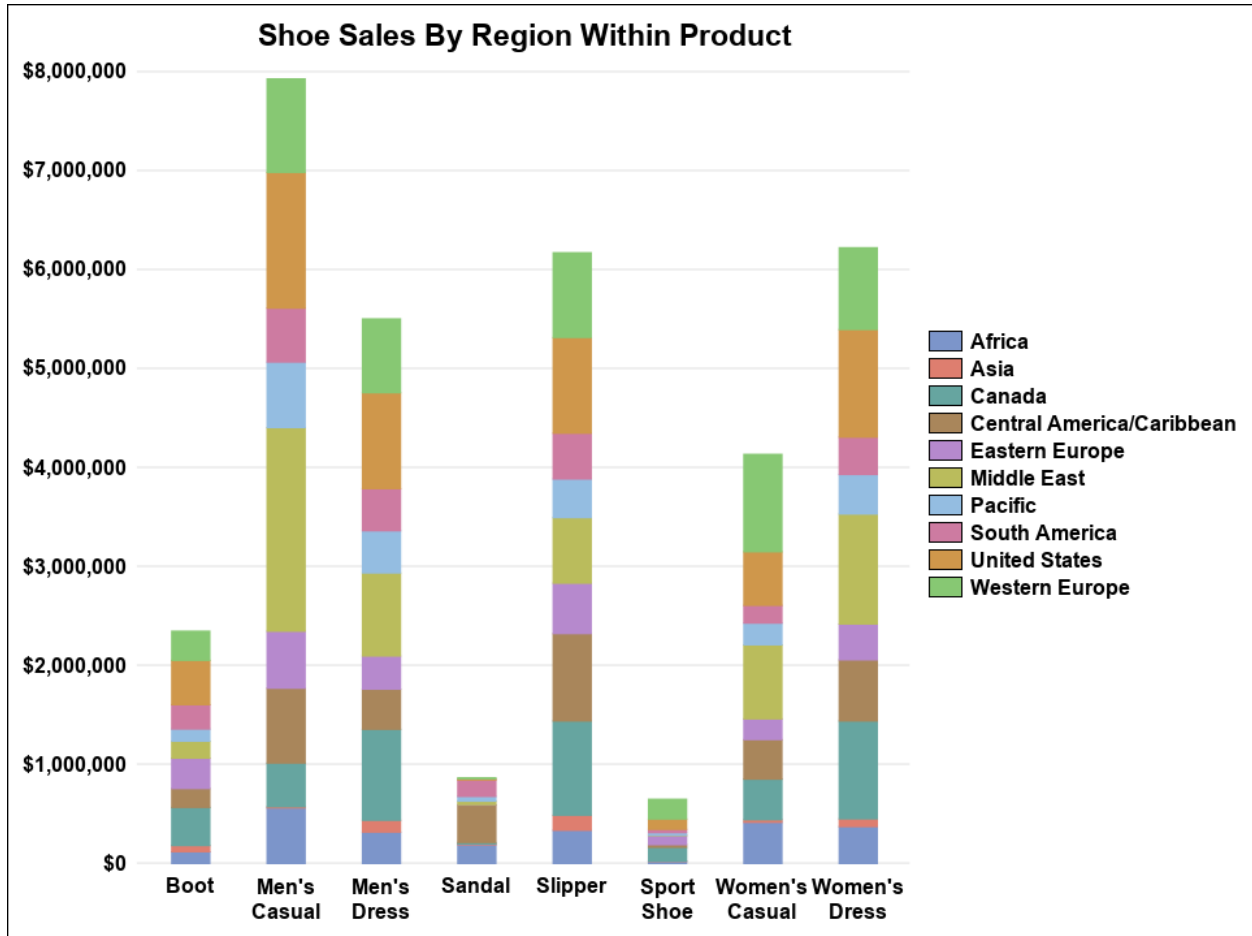


Figure 7. This is NOT practical if you want to know the numbers.

**USABLE STACKED BAR CHART IMBEDS NUMBERS WITH AN XAXISTABLE**

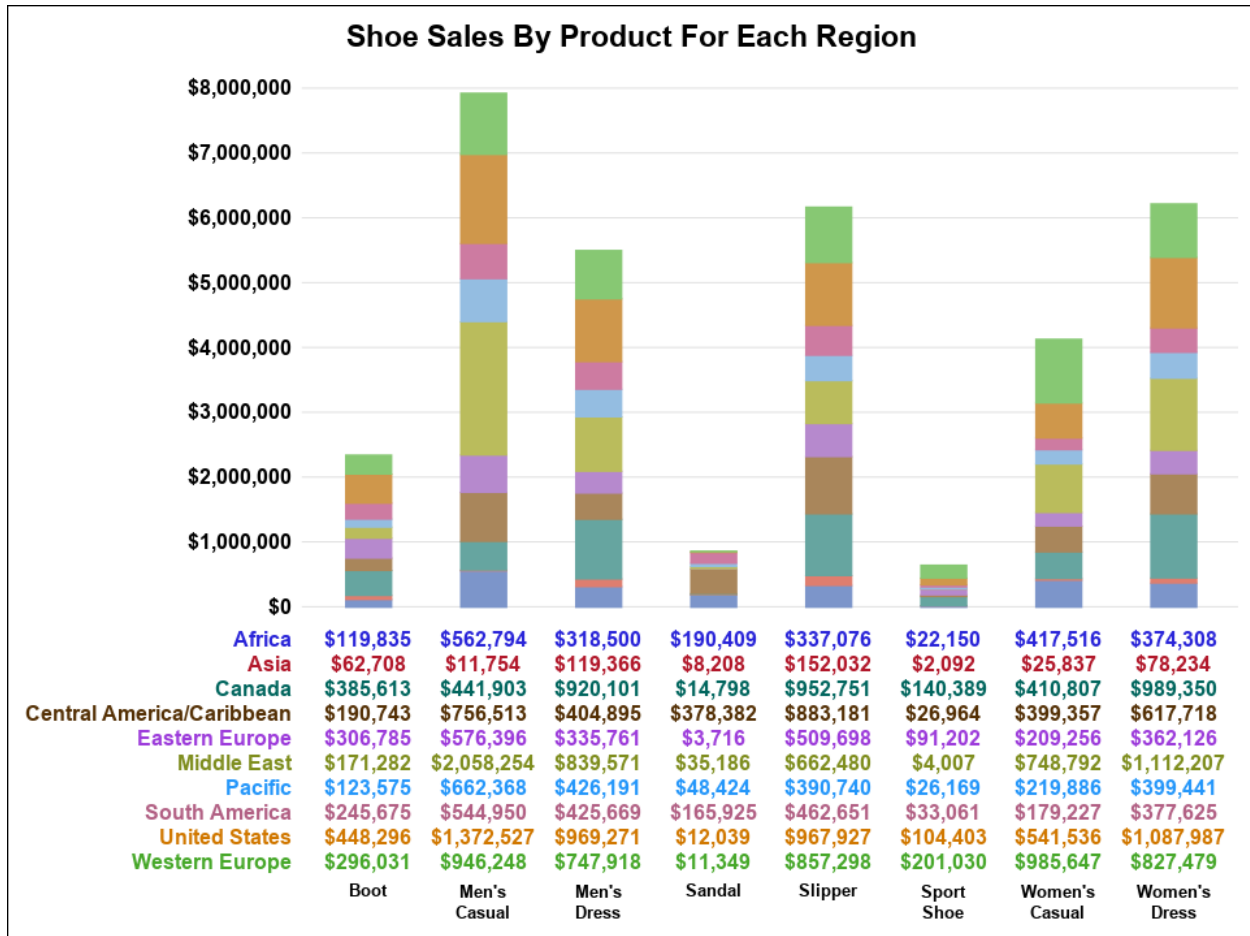


Figure 8. IT REALLY WORKS.

NOTE: Thick text is mandatory if you want your viewers to be able to reliably distinguish colors.

**MAXIMALLY INFORMATIVE HORIZONTAL BAR CHARTS**

**AN EASY IMPLEMENTATION**



Figure 9. The same quantitative information as a Pie Chart, but a different image for comparison

**WHO COULD ASK FOR MORE? HERE YOU HAVE EXPLICIT RANKING, SUBSETTING, LINKS TO OTHER SUBSETS AND THE ENTIRE LIST**

Other subsets: Top 85% of the Grand Total & Sales Greater Than Or Equal to \$1,000,000. They & ALL back to here. The whole package is web-enabled. For the code, see Reference 1.

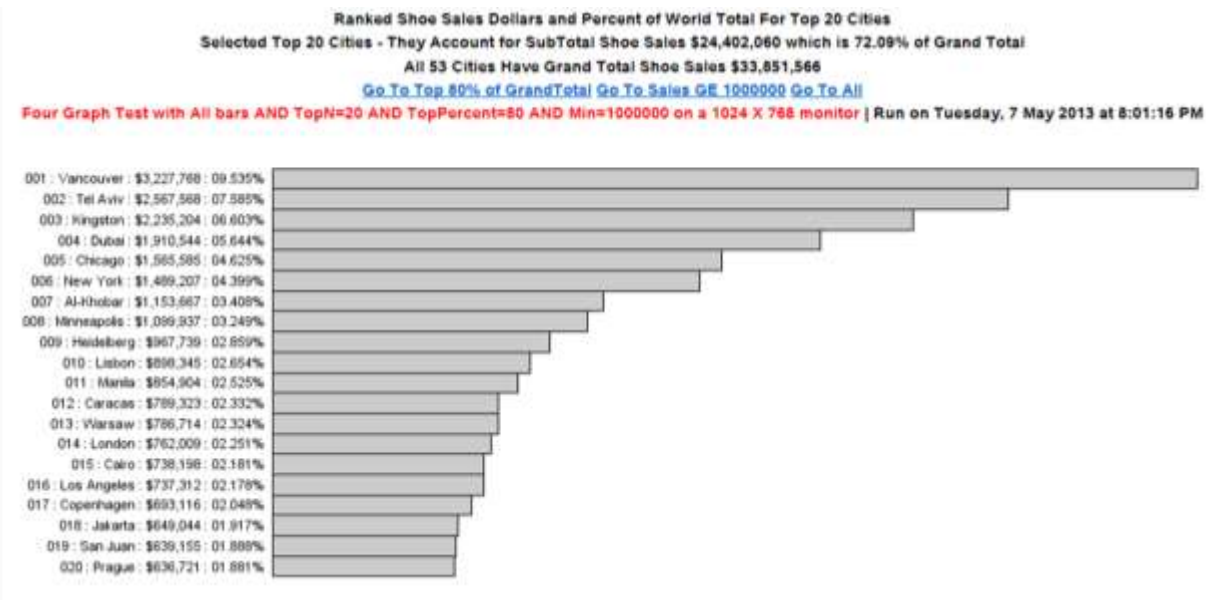


Figure 10. Explicit Rank Numbers and all information adjacent to Category identity  
Non-adjacency can be a visual association challenge for very long bars

**A CLOSEUP OF THE TOP EIGHT LABELS**

001 : Vancouver : \$3,227,768 : 09.535%  
 002 : Tel Aviv : \$2,567,568 : 07.585%  
 003 : Kingston : \$2,235,204 : 06.603%  
 004 : Dubai : \$1,910,544 : 05.644%  
 005 : Chicago : \$1,565,585 : 04.625%  
 006 : New York : \$1,489,207 : 04.399%  
 007 : Al-Khobar : \$1,153,667 : 03.408%  
 008 : Minneapolis : \$1,099,937 : 03.249%

Figure 11. Rank, Category, Value, and Percent of Grand Total as Bar Label – All you need to know

# CLUSTERED HORIZONTAL BAR CHART

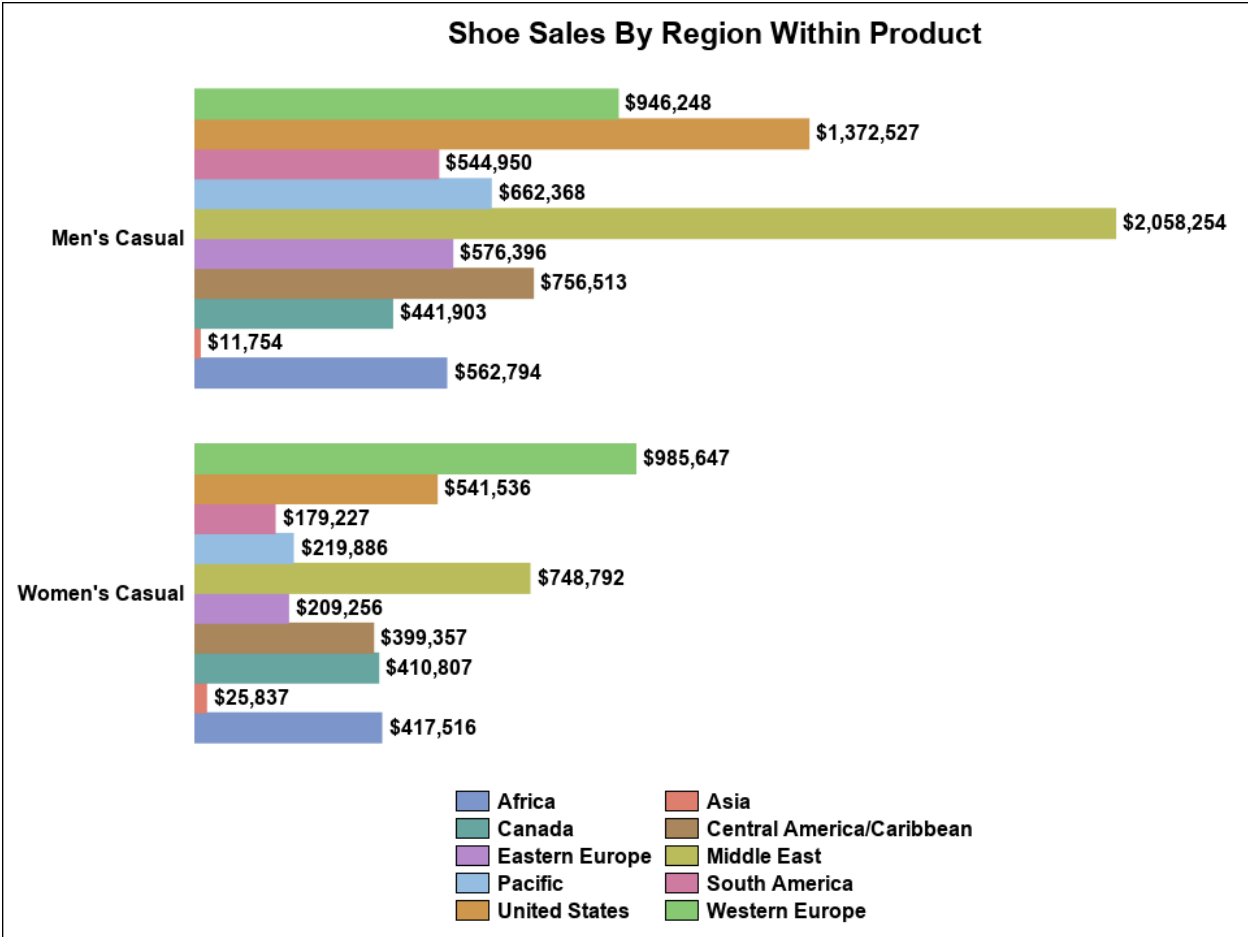


Figure 12. Easier To Digest Than A Stacked Bar Chart, The Viewer Actually Gets Precise Numbers

# PIE CHARTS

Decades ago, bias against pie charts was engendered by academic observation that it can be difficult to accurately visually determine the comparative size of pie slices. The solution, provided below and repeatedly advocated by me, is simple: **DISPLAY the actual values and the percents of whole. Optionally, accelerate/facilitate the comparison by ORDERING the pie slices by size.** Ordering is a disadvantage only when it brings the labels of small slices too close together and collisions occur between the labels. In that case, you could present larger slices between small pairs of slices, or you could move all of the labels into the legend as shown in Figure 17, which delivers ALL of the data, even when one or more slices are vanishingly narrow.

Unnecessary claims of pie chart graphic ineffectiveness were based on the research "discovery" of an Unnecessary Truth.

## MAXIMALLY INFORMATIVE PIE CHART

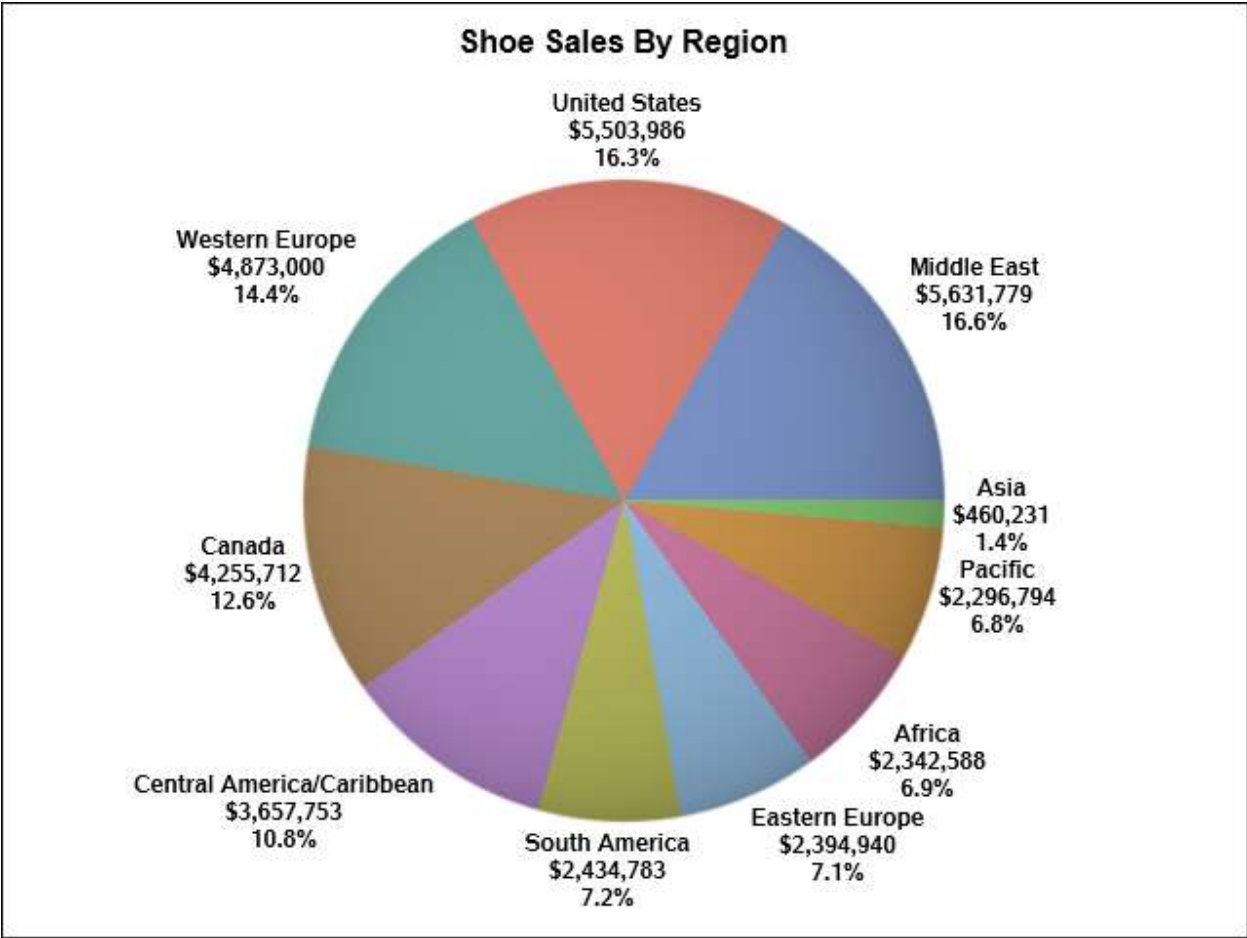


Figure 13. Maximally Informative Pie Chart (Ordered, with Description, Value, & Percent of Total)

**Note:** In ODS Graphics, the default way of displaying pie slices is to center the first slice at zero degrees and to present the remainder in counterclockwise order. Above, the edge of the first slice is at zero.

**SAME BENEFITS USING THE CALLOUT OPTION, BUT MORE COMPACT**

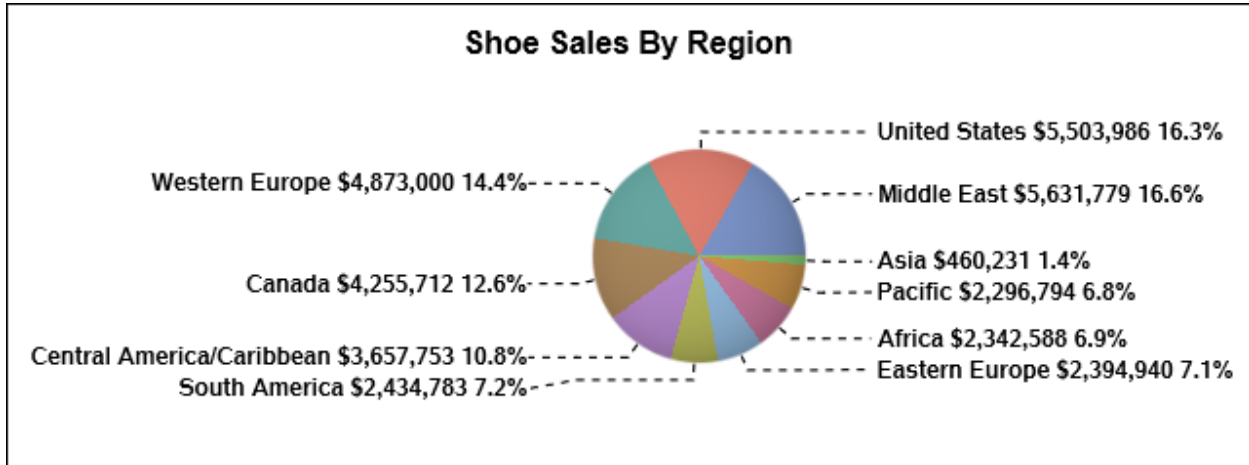


Figure 14. Optimum Height Pie Chart at Microsoft Word Document Page Width (6.5 inches)  
Increasing the height wastes vertical space. Decreasing it shrinks the size of the pie.

**DISADVANTAGE OF USING AN ODS GRAPHICS PIE CHART WITH A LEGEND**

To determine the numbers for a Region your eyes must go to two places.

**Default Slice Layout and Default Legend**

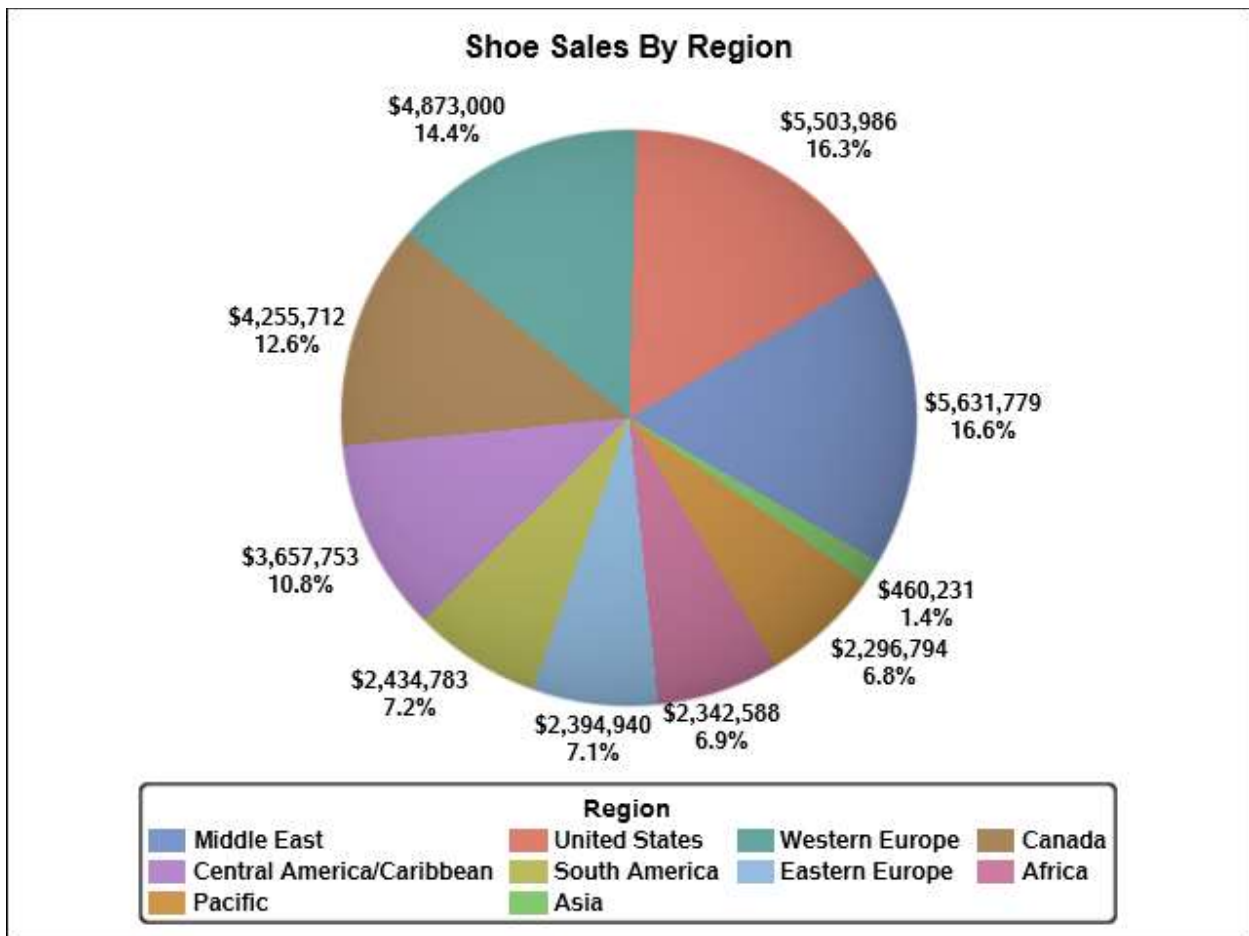
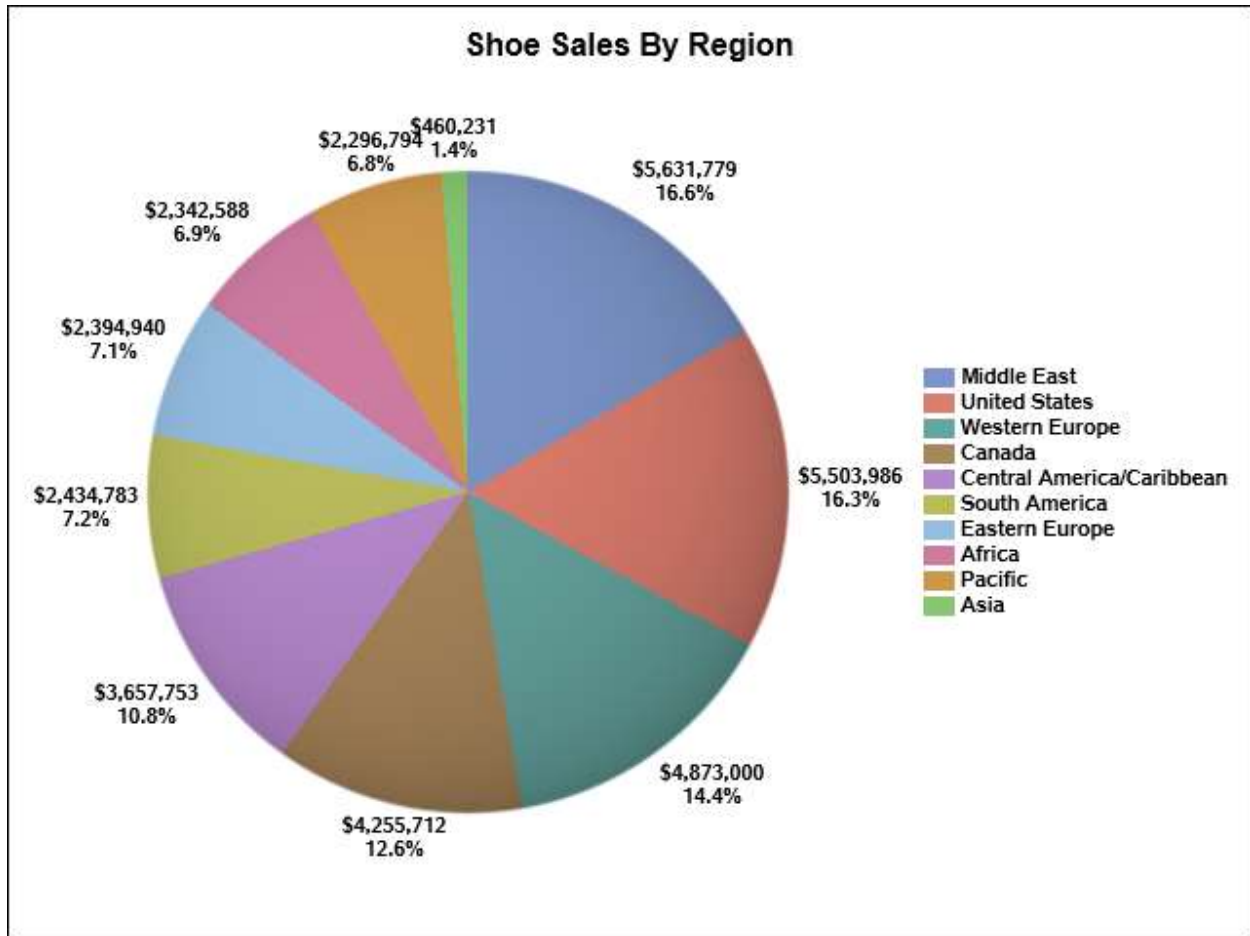


Figure 15. Legend Color Samples Were Optimized to Make Colors More Easily Distinguished

## Better Pie Chart: Slice Order and Legend Entry Order Are “Visually Aligned”



**Figure 16. Edge of first slice is at 90 degrees with descending value slices displayed clockwise**

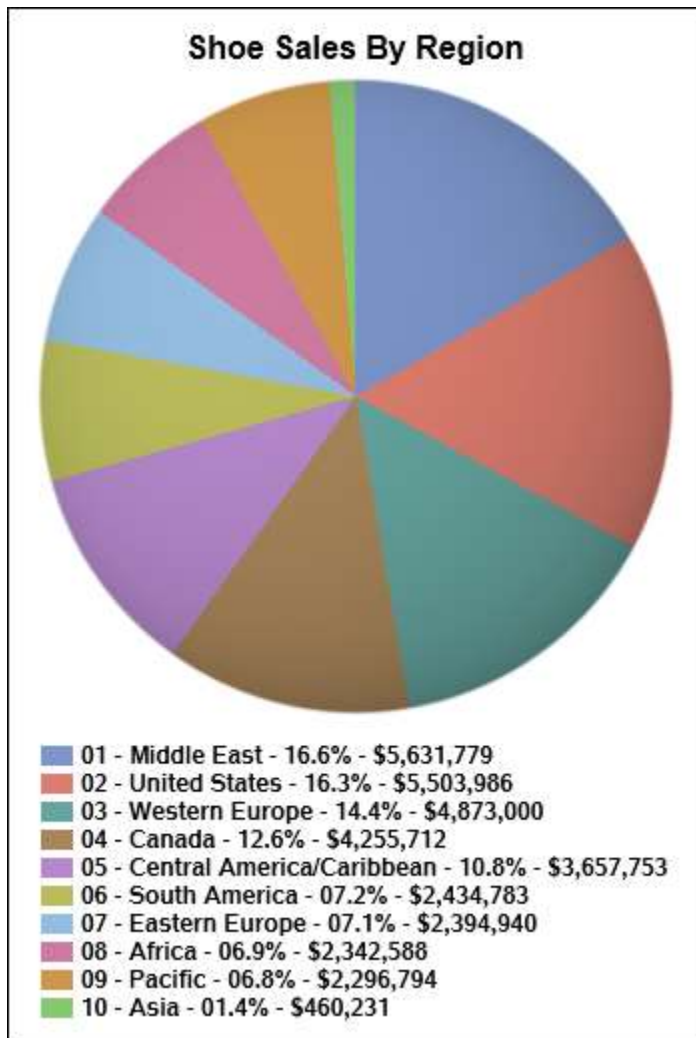
By “visually aligned” I mean that your eye starts at the top of the pie and goes down, and your eye starts at the top of the legend and goes down. At least that is what MY eye does. The default would have the pie slices start with the first one centered at “Three O’Clock” and go counterclockwise. Here the pie slices start with the edge of the first one at “Twelve O’Clock” and go clockwise.

Data Labels for Pacific and Asia are touching, but still readable. Reducing font size for all text (and fill height) from 8pt to 7pt can eliminate that, but would make text slightly more difficult to read.



## Maximally Informative Legend

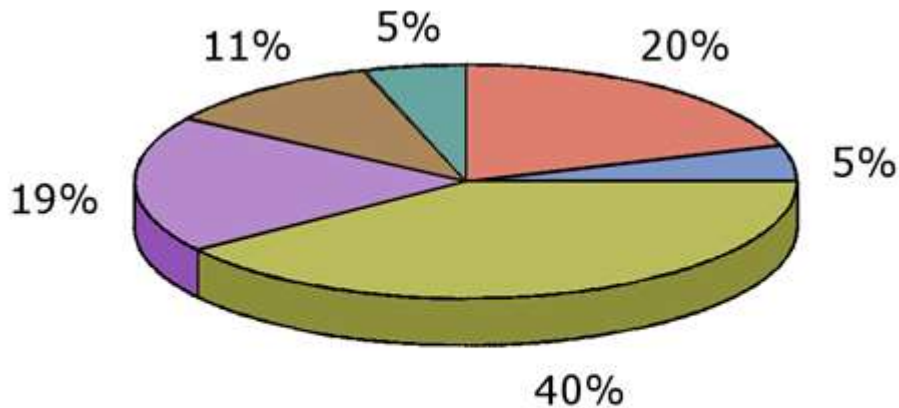
Avoids any possible data label or callout collisions



**Figure 17. Pie Chart with No Possible Data Label Collisions (because it has no data labels)**

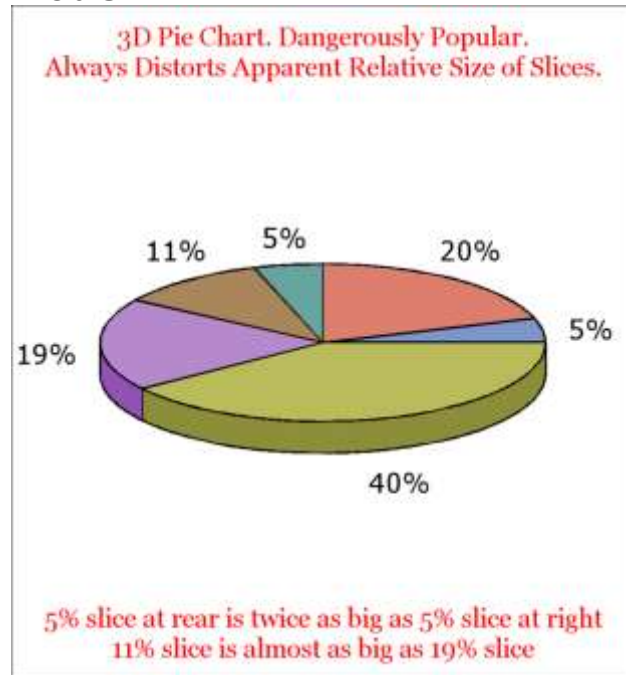
This pie chart is, in effect, a table with an image as an easy and quick visual comparison accessory—an inverse to the situation where data labels are a precise quantitative accessory to an image.

**3D ALWAYS CREATES A MISLEADING PIE CHART**

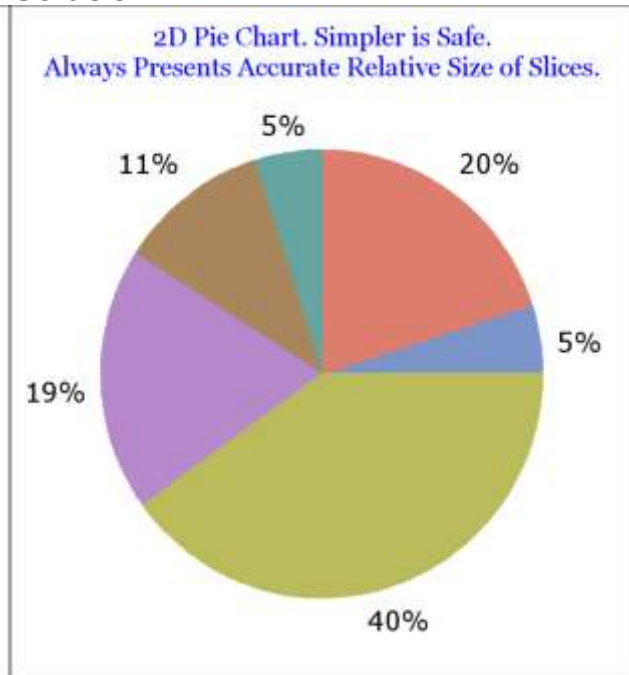


5% slice at rear is twice as big as 5% slice at right  
 11% slice is almost as big as 19% slice

**Problem**



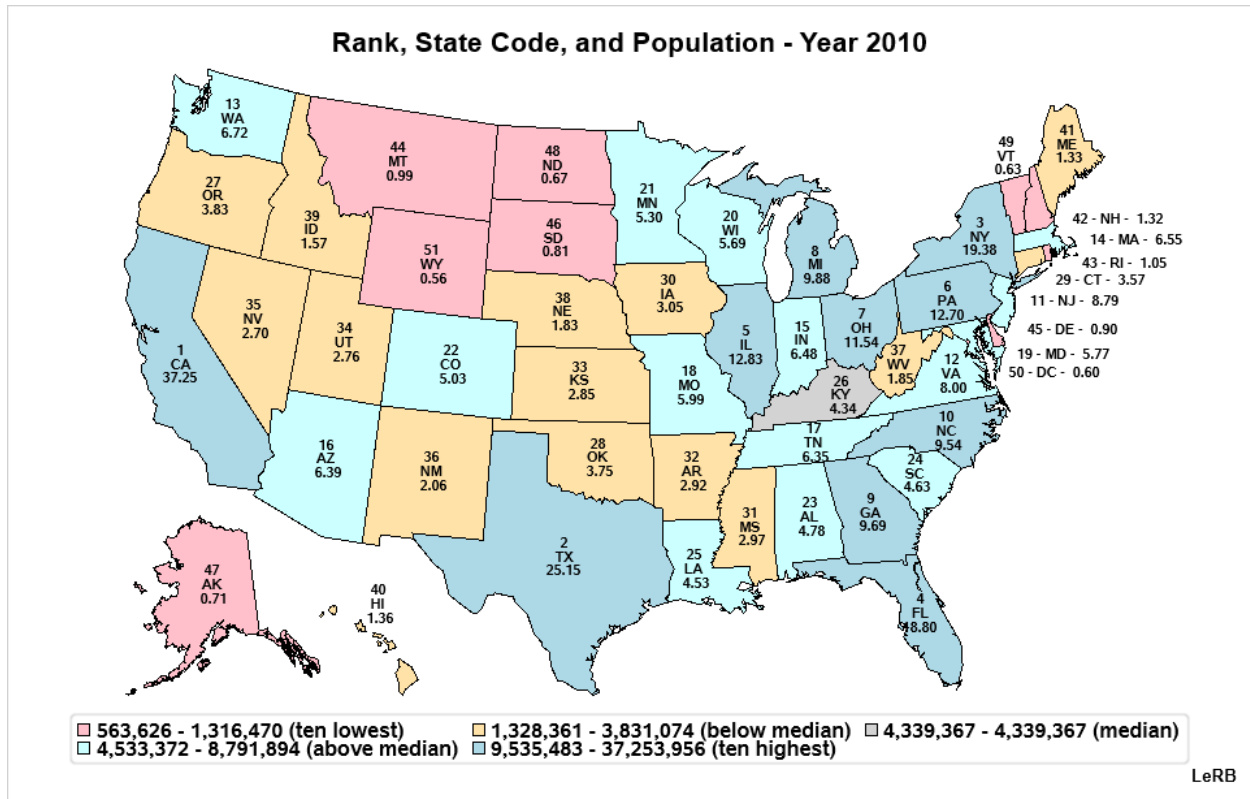
**Solution**



**Figure 18. 3D ALWAYS Creates a Misleading Pie Chart (unless two slices of equal size)**

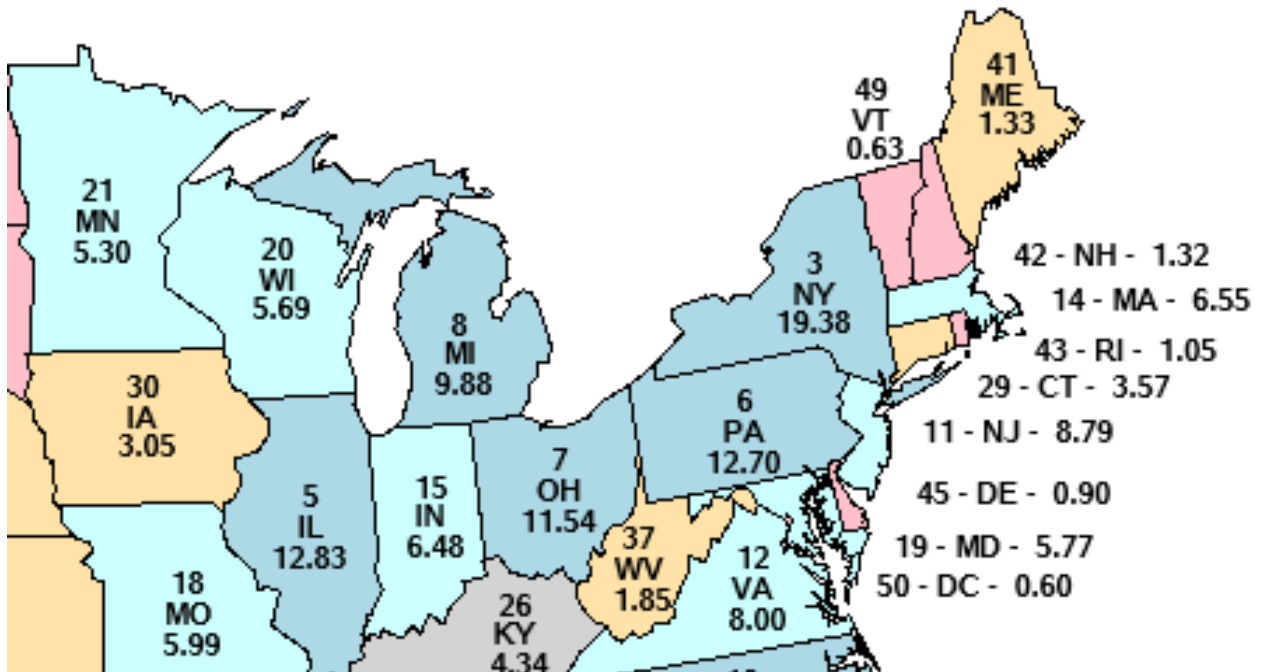
I first demonstrated this underappreciated problem at the 1995 SAS Users Group International Conference, and have never stop bringing it up. ODS Graphics does not support 3D pie charts, but SAS/GRAPH eventually did, after users kept resorting to various folderol to create them. 3D is more flashy, more exciting, maybe more interesting. I have always had a weakness in favor of accurate, reliable, effective visual communication.

## SHOW THEM WHERE IT'S AT WITH A CHOROPLETH MAP



**Figure 19. Maximally Informative Map of State Populations (with DC treated as a state)**

Range boundaries are dynamically determined via pre-processing and are delivered as macro variables to the LEGEND statement. This map design is one of my favorite ways for grouping geographic unit areas into categories. On/In-State Annotation is impossible for some mid- to upper-Atlantic states and for some in New England. One could annotate lines from the labels to those state centers. This close-up shows the annotation more clearly.



## CONCLUSION

I hope you have enjoyed seeing these ways that you can deliver image and ranked precise values using the most popular and widely used types of data visualization, as well as the maximally informative choropleth maps. I have enjoyed developing these examples, refining them, and sharing them. Thanks for your interest. Feel free to request the zip file of commented code via email.

## REFERENCES

1. Bessler, LeRoy. 2013. "Data Visualization Power Tools: Expedite the Easy, Implement the Difficult, or Handle Big Data" in *PharmaSUG 2013 Conference Proceedings*, Chicago, IL: PharmaSUG. Available at:  
<https://www.pharmasug.org/proceedings/2013/DG/PharmaSUG-2013-DG11.pdf>

## ACKNOWLEDGMENTS

I'm grateful to Geeta Kersellius and other volunteers at SAS Global Forum 2020 who made it possible for me to share my ideas, knowledge, and experience with fellow SAS users, to Chris Potter SAS Users Group International Graphics Section Chair who in the early 1990's encouraged me to share my graphic design ideas, to Allen Paller a graphics expert speaker who encouraged me to become the graphics support and good design advocate at Miller Brewing Company, to Thomas S. Cain at Miller who gave me that responsibility, to Steven J. Subichin the Miller colleague who shared his SAS map annotation method with me, to all of the people in SAS Technical Support and SAS software development who have helped me in my SAS graphics quest since 1981, and to Alexandra Riley for her review of this paper.

## CONTACT INFORMATION

Your requests, questions, comments, suggestions, and ideas welcome. Contact me at:

LeRoy Bessler PhD  
Bessler Consulting and Research  
Visual Data Insights™  
Strong Smart Systems™  
Le\_Roy\_Bessler@wi.rr.com

## ABOUT THE AUTHOR

Dr. LeRoy Bessler, SAS consultant and trainer, has presented at conferences in the USA, Canada, and Europe, on communication-effective data visualization (using graphs, plots, charts, tables, web pages, maps, or color), SAS to Excel, tools for SAS server administrators, users, and managers, and Software-Intelligent Application Development for Reliability, Reusability, Maintainability, Extendibility, and Flexibility. His experience includes SAS application development as well as SAS Administration, supporting users, servers, software, and data. An energetic and enthusiastic and frequent practitioner of data visualization, in this era of Data Science LeRoy thinks of himself as a Data Artist. He first wielded the brush of SAS/GRAPH in 1981 when it was only two years old, and now enjoys working with ODS Graphics.

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.  
Visual Data Insights and Strong Smart Systems are trademarks of LeRoy Bessler PhD.