

Too Many Pixels: A User's Guide to Incomprehensible Web sites, Charts, Graphs and Other Presentations

Peter Parker, U.S. Dept. of Commerce, Washington, DC

peter_parker@ita.doc.gov

Last updated 01/12/1999

Abstract

After almost two decades of working with charts, graphs, web sites, and other presentations, I have seen too many that were over-produced, over-slick, data-impooverished, incomprehensible, but "pretty" works, more concerned with style than substance or ease of use. 3-D charts and other multi-colored, multi-spatial, over-graphical monstrosities should be avoided. Just because hundreds of fonts are available, doesn't mean they should all be used. These colorful "presentations" may appeal to unsophisticated, easily impressed audiences, but usually are of limited use to analytical viewers. A good presentation should be able to summarize information, without distractions, so that users can infer and conclude, while delving deeper into the data if needed. Making useful and user-friendly presentations is an art and can lose clarity if too much science is used. My preferred method of demonstrating proper information broadcasting is to show examples of poorly rendered ones and then transform them into presentable ones suitable for my target audience. In this paper, I critique a web-site, a chart, a map presentation, and finally the familiar SASware ballot®. Much of my methodology draws heavily on the research of Edward Tufte and his reference works.

Introduction

Speak the speech, I pray you, as I pronounced it to you, trippingly on the tongue: but if you mouth it, as many of your players do, I had as lief the town-crier spoke my lines. Nor do not saw the air too much with your hand, thus; but use all gently: for in the very torrent, tempest, and, as I may say, the whirlwind of passion, you must acquire and beget a temperance that may give it smoothness. O, it offends me to the soul, to hear a robustious periwig-pated fellow tear a passion to tatters, to very rags, to split the ears of the groundlings, who, for the most part, are capable of nothing but inexplicable dumb shows and noise: I could have such a fellow whipped for o'erdoing Termagant; it out-herods Herod: pray you, avoid it. (Hamlet, Act III, Scene II.)

The graphical display of information is meant to enhance the transmission of information, not to distract from it, or to make an artistic statement. Rather, graphics should help you understand data. One would be pressed to imagine what a billion dollars looks like. How much more is a billion compared to a million? However, one can create the vivid graphic image in one's mind of a million dollars being a stack of one-dollar bills being equal in height to the Washington Monument, while a stack of billion dollar bills would reach up about one hundred miles.

Recently, I attended a speech given to about three hundred high school seniors. The speaker wanted to relay statistical information about their probable futures and he easily, but undramatically, could have presented the data using basic bar charts. Instead, he handed each student a few slips of numbered paper. When he asked holders of number 24 to rise, most students stood up (about 96%). That was the number of students who would be going to college. When he asked only holders of number 13 to remain standing, only about one-fourth stood. That was the number who would not graduate from college. He asked the holders of number 56 to rise. Only about two stood up. That was the expected number who would not live past this summer. And so forth, for the number who would go to jail for violent crimes and other statistics. Visually, this presentation produced a powerful effect.

We think in pictures. Ten percent means something to visually oriented teen-agers when they can see thirty people dispersed among three hundred standing up. Otherwise, ten percent is barely distinguishable from one or thirty percent. This clever means of presenting information is the bottom line of graphical data, making information comprehensible.

Besides making information more digestible, graphics also enhances bad information, making it more effective in misleading the user. Recently, I read a newspaper article about bicycle riders having problems with impotence from bicycle seats, presumably from pinching a particular artery. A study of runners and bikers had shown that while 1% of runners suffered from this malady, 4% of bikers did as well. Anyone with a statistical background can see that if one assumes that the standard error is plus or minus three percentage points (common in survey data), then there is no statistical difference between the runners and bikers, or even from zero occurrences.

What would be interesting would be comparing bikers to the national average. If bikers suffered this problem at a significantly greater rate than the general public, then one could conclude that there might be some cause and effect. Yet, the statistically illiterate portion of the public may believe that bikers are four times as likely as runners to suffer from impotence. A bar chart likely will show these two data points with two bars (or some undignified graphic), one bar four times as big as the other. Then we would have a health crisis, with stickers on bicycle seats warning that this product may be anti-family.

The preceding examples are cases of good and bad visualizations of information, meant to define the problem. A metaphor from life can be found in visiting any gym with weight lifting equipment. You'll find potential fitness buffs

arching their back on barbell presses, overextending elbows, and bouncing weights while doing curls with dumbbells. They think they're getting a good workout, but instead, they're barely pumping up, and likely will injure themselves.

As in presentations, the neophyte speaker or webmaster may think that they are getting their message out, but they may be only obscuring it instead. In the paper, I'll examine other examples of information display, and criticize their failings, in order to emphasize what one must consider before going public with any type of visual product. Specifically, I'll examine a web-site, a chart, a map presentation, and finally the well-known but truly incomprehensible SASware ballot.

Incomprehensible Websites

Finding an incomprehensible web site is an easy task. Browse the Internet for a few minutes and you'll find many embarrassingly bad sites. Bad practices are constantly harped on by Web page purists, and since this information

is so widely available, I will only mention them. Basically, one should stay away from gaudiness and any unnecessary features that delay a web page from loading. Avoid:

- 1) large or unnecessary pictures (30k maximum)
- 2) large or flashing fonts
- 3) super-saturated links (e.g., a picture of a globe with a link to every country- use a pick list instead)
- 4) music
- 5) animation
- 6) too many clicks to get anywhere

My paper will focus on incomprehensible web pages that are more concerned with impressing the leadership of the institution rather than in the primary purpose of the web-site, which is providing information to clients. For this purpose, I will describe an imaginary web site, rather than risk insulting my colleagues. Here's the recipe for making a truly hideous "institutional" web page.

Visualize a web page for the fictional business-promoting Office of Obfuscation:

	<h2 style="text-align: center;">Welcome to the Office of Obfuscation</h2>	
<p>"Keep up the Great Work!</p> <p>Tyger! Tyger! burning bright In the forests of the night, What immortal hand or eye Could frame thy fearful symmetry?...</p> <p>The Tyger, Blake "</p>		<p>...And we are here as on a darkling plain Swept with confused alarms of struggle and flight, Where ignorant armies clash by night.</p> <p>Dover Beach, Matthew Arnold"</p>
<ul style="list-style-type: none"> ● Compliance Customer Policy Division 	<ul style="list-style-type: none"> ● Policy Compliance Customer Division 	<ul style="list-style-type: none"> ● Customer Fulfillment Division
<ul style="list-style-type: none"> ● Customer Policy Compliance Division 	<ul style="list-style-type: none"> ● Customer Satisfaction Center 	

<ul style="list-style-type: none"> ● Our Mission Statement ● A Word from our Department Head ● A Word from our Division Head ● A Word from our Section Head ● A Word from our Sub-Section Head 	<ul style="list-style-type: none"> ● How to contact us. ● Tell us what you think.
---	---

Please register with us before continuing further.

While this web page prototype may be visually appealing to some users, most people will find it incomprehensible, and, therefore, useless. This page was rendered in the "institutional" mode, making sense to the office that maintains the page, but not to users seeking information. Users don't care about the graphics or superfluous information. Just because it's a federal home page does not justify having a picture of the President. Everyone knows what he looks like. Why a photograph of the other government official? While he may not be widely known, no one cares what he looks like. The graphic of the arrows in a circle is merely another decorative graphic to impede users with slow modems. Of course, each official includes a memorable statement, relevant or not to the information needs of the web user. All levels of the structure want to be visible, and will be certain to include biographies and photographs of their top officials. No one needs or wants to hear from the department, division, section and subsection heads.

An institutional approach would not be complete without reference to the power structure of the office. What is a "Customer Satisfaction Center"? Perhaps if the user is lucky, they can find the information they're seeking there. How does "Compliance Customer Policy Division" differ from "Customer Compliance Policy Division"? "Customer Fulfillment" from "Customer Satisfaction?" Does anyone





care except those who work in those divisions? The only information the user may crave is a brief statement on the functions of the Office of Obfuscation.

How to contact us? If this web page was more information-effective, perhaps the user wouldn't have to contact the office.

Tell us what you think? Do you really want to know?

Please register with us before continuing further. Another click, another obstacle. A user merely wants the data and as quickly as possible. At least, this view should be a user's first thought. However, this registration may be the only well-designed part of this page. A simple registration process could be used to see how many different people are using this business-promoting page and how useful it is, such as generating sales leads. If only three people use this page and all three found insignificant improvements to their business, then this page should be discontinued. However, if several thousands of people use this page every month and generate substantial business because of it, then it may be worthwhile for the office to expend more resources into this page, perhaps providing more data.

A better approach to a simpler and comprehensible web-site may look like this:

<p>We blur the lines! We have all sorts of data! Big and small numbers!</p>	<p>Welcome to the Office of Obfuscation</p>	<p>Data has been updated for November 1998 Trade Statistics!</p>
<ul style="list-style-type: none">  Who we are.  How to contact us.  Useful links. 		<ul style="list-style-type: none">  Data: <ul style="list-style-type: none"> Imports Exports Balance of Trade Report Other Reports

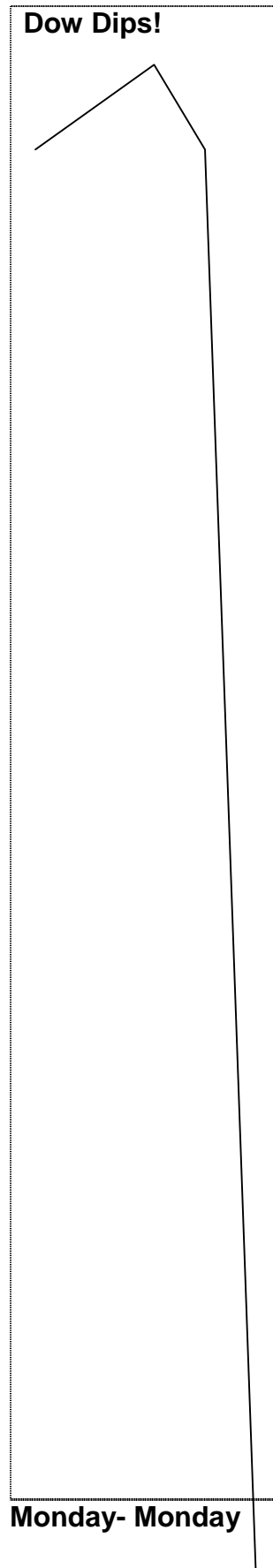
The data is only a click away. Graphics are minimal. One could add a small logo to enhance the look of the page, as long as the graphic is unobtrusive and has something to do with the subject of the web page. A graphic of a cargo ship would be only useful for emphasizing that the page is about foreign trade. Also, a company logo or government seal graphic would remind the user who's the sponsor providing them with information. Even with additional visual information, keep the page simple. The users want to enter quickly, find their information and then get out. If they care about the institutional structure, they can click on [Who we are](#). The page clearly identifies the office, including a brief pithy message "We blur the lines! We have all sorts of data! Big and small numbers!" to clearly relay what the office does. Simplicity works!

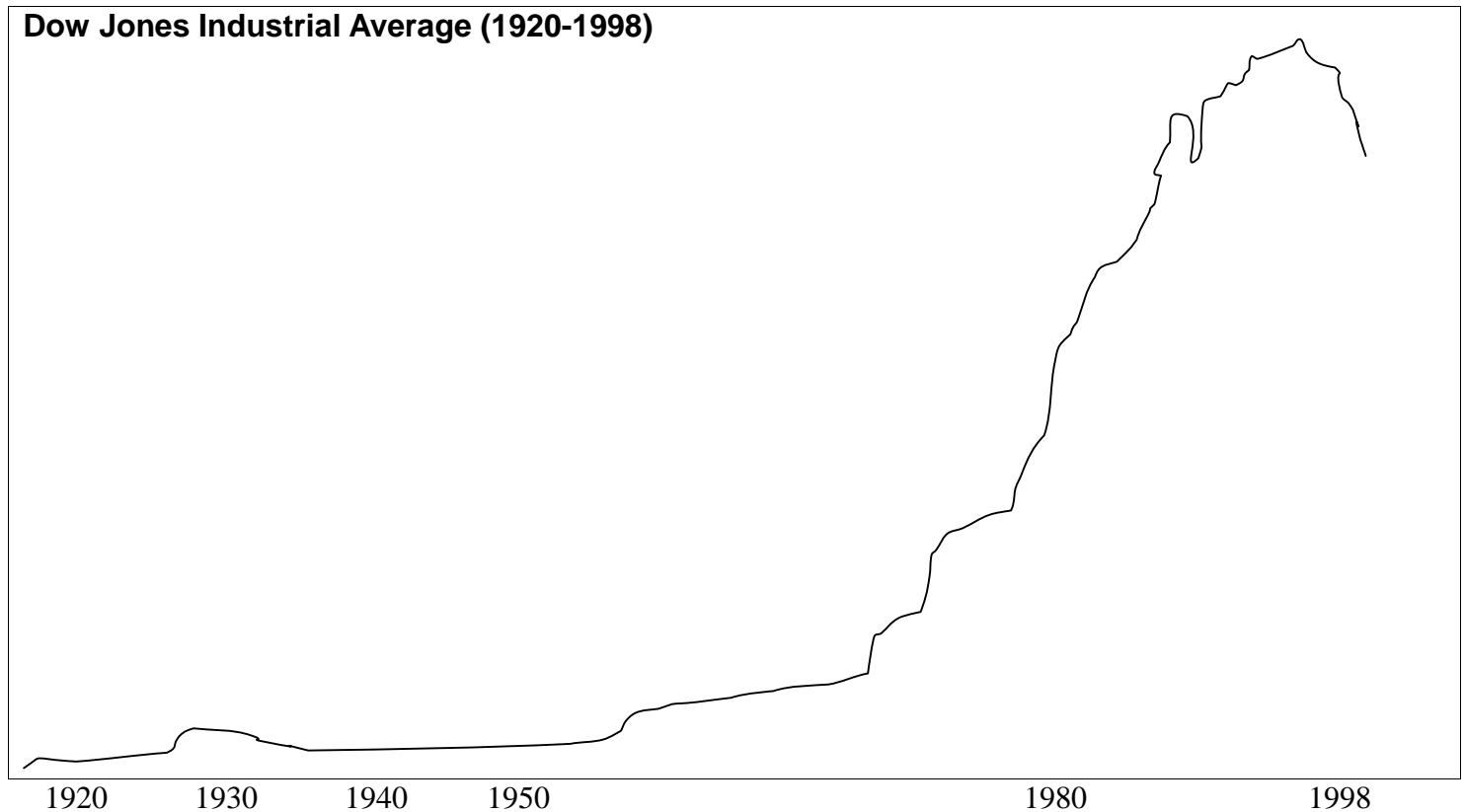
However, there is another approach that may justify graphic-intensive information-feeble home pages. If one views web pages from a marketing perspective, this strategy may make perfect sense. First of all, every company needs a web-site, regardless of content, merely as a credential. By keeping up with popular technology, the company will not be perceived as behind the times. Web pages are also advertisements for businesses. Big and bright catches the eyes. As you go through phone book yellow pages, which listing do you call? The one with the biggest ad, of course. Additionally, a web page with high quality graphics will enhance the professional perception of a company. For selling to the general public, style always trumps content.

Incomprehensible Charts

With SAS/GRAPH® software, or any other graphical software, we can tell any story. During the late summer of 1998, the Dow Jones Industrial Average plummeted over the course of a week. The graphical chart in one newspaper had a tragic story to tell. The market was falling through the floor. Here is that chart approximately reproduced, although the original was more exaggerated, running along the entire left hand margin of the paper, showing an almost straight line drop. Note that there are only six data points being shown (for six days, excluding the weekend). The newspaper dressed it up, adding the day in big letters, the level of the Dow, and the percent change on each data point. The line was in thick red, with extraneous grid lines. The figure here is only to give the reader an idea of what a bad day this must had been for some speculators.

However, another paper decided to show the changes in broader context. It mapped the past eighty years of the Dow Jones Industrial Average, to keep the numbers in historical perspective. Obviously, this graphic was more intellectually honest. Incidentally, a few days later, the market rebounded close to the average of a week before. This "plummet" became a small blip.





This chart is a rendering of the actual graph, only to give reader the perspective of a growing trend with blips, versus a minor plummeting. This chart tells a better story of where the Dow Jones Industrial Average started, where it has been, and how it has progressed. It had been comparatively flat for decades, taking off during the eighties.

Unfortunately, looking from the marketing perspective again, data-deprived charts may be more convincing to the general statistics-ignorant public. A color pie chart with two or three data points may raise the hackles on a statistician's neck, but may be enticing to unsophisticated users, desiring an easily digestible message. Hence, the former plummeting stock market chart will make a clear concise message, that the market is dropping rapidly. I call these charts "Headline charts." Like a newspaper headline, they only give the simplest of information that could be understood with minimal mental effort. And like newspaper headlines, they are woefully inadequate by themselves. They need the rest of the story.

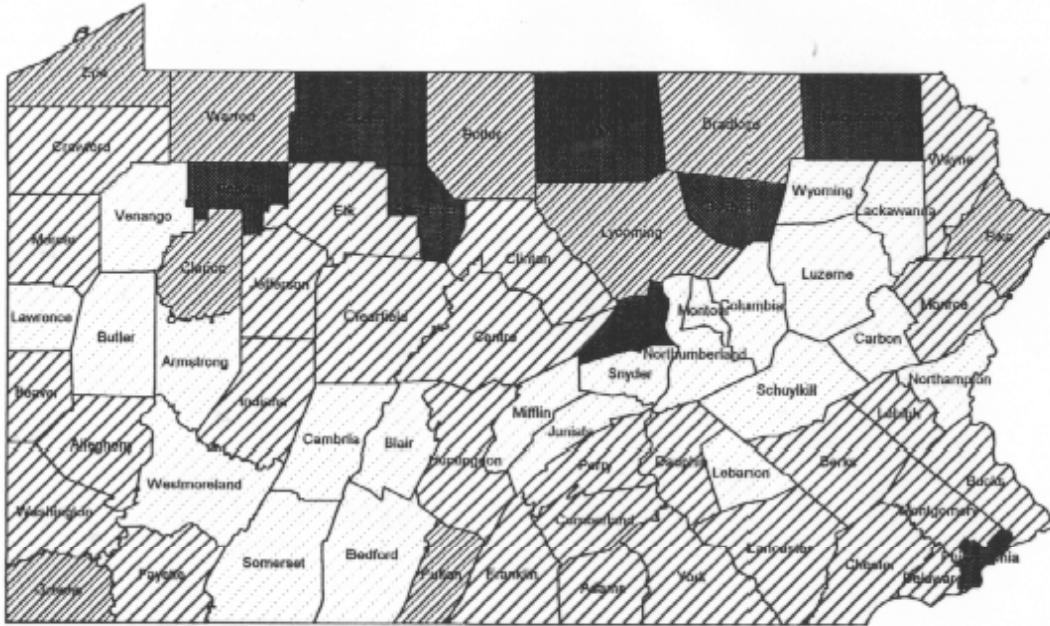
Incomprehensible Map Presentations

Incomprehensible map presentations are easy to pick apart. They tend toward over-inflated graphics to cover up poor analysis. Information tends to become clouded instead of clearer in these over-wrought depictions. It is the merging of sampling data with the spatial relations of

political geography that tends to obfuscate the message. Visually it may be appealing, but informationally it may be misleading. Rather than detail all of the problems associated with this type of presentation, I am showing a "good" example of a bad map to point out what you should avoid.

Suppose you wanted to display information on a study of baldness for men over sixty-five years old for the counties in the state of Elbonia. It can be displayed as follows with different colors for different percentages of the sample with baldness. I've seen colorful charts with pinks, greens, reds and blacks to show different baldness rates. How does pink to green to red offer intuitive clarity? A clearer approach would use black to show the least baldness and white, the worst. As the baldness rates increase, the blacks become grayer until white, reflecting a shinier pate. Even with this intuitive darker to lighter scaling, this map has a lot of problems.

Elbonia- Baldness Rates for Men, 65 years and older, for 1990-1995



Baldness Rate

Black	- 10 %
Dark Gray	- 20 %
Light Gray	- 30 %
White	- 40 %

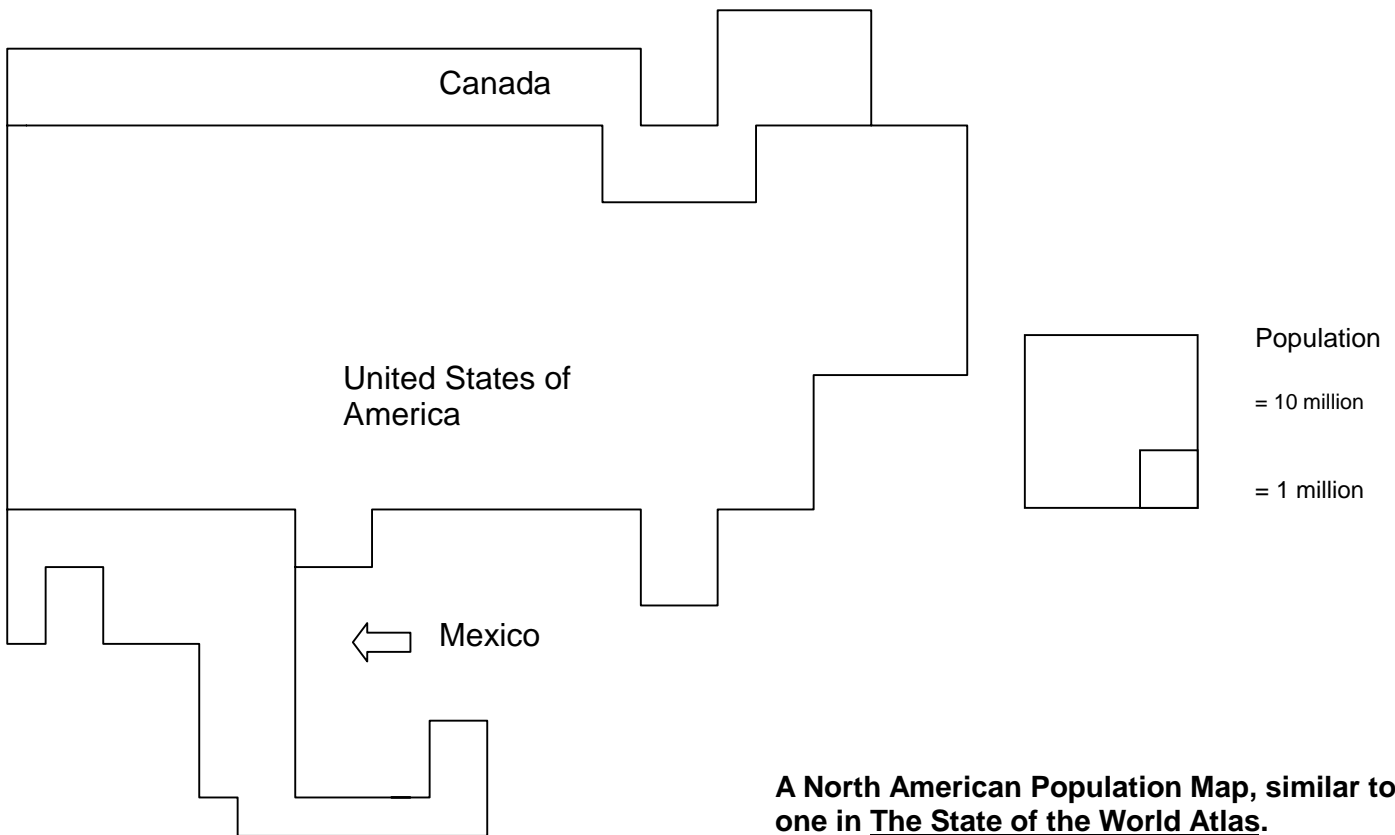
From this map, one may decide that the southern area of Elbonia is conducive to baldness, while the Northern region is safer. However, this map begs for the following questions to be answered:

- 1) What is the "rate?" Are these rates adjusted for population size? What if the light regions have high concentration of elderly men while the darker ones have insignificant one. Perhaps this rate should be based on percent of baldness per 100,000.
- 2) Do the extreme white areas have more urban areas than the black areas. Perhaps the "safer" rural areas would pull down the numbers in some counties. The artificial county barriers may hide the numbers. Perhaps a better approach would be to ignore the county lines and color the chart on a smaller unit, such as cities. One may see a "white" county with some black spots, or vice-versa.

I prefer a totally different approach than overlaying a county map with data to fit within those counties. In the book, *The State of the World Atlas*, the authors presented the world map in several ways, based purely on one parameter, such as population. For instance, a $\frac{1}{4} \times \frac{1}{4}$ inch block could represent a million people. Hence, a stylized set of blocks in the shape of the United States would be many more blocks than the geographically immense but less populated Canada. Japan becomes several times larger with respect to other countries, while

the former Soviet Union would be depicted much smaller. Other block proportional representations of maps could be based on population growth rates, birth rates, death rates, per capita income, or any other economic factor. For the county map of Elbonia, one could redraw the map with blocks, based on population. Each block could represent 50,000 people, and each block could be colored white, gray, or black, based on the baldness rate. The analyst will still have the spatial relation since the county blocks will be relative to each other in location, like in a political geography map, but minus the useless details of irregular borders and empty rural areas that may skew the analysis and produce flawed conclusions.

Since we don't have the data to render the previous map of Elbonia in the block format, I'm presenting a fragment of a world map from *The State of the World Atlas*. While retaining the basic shapes of the US, Canada and Mexico, the authors were able to give a different view of size comparisons between these three countries, based on population size. Canada becomes a sliver of a country to the USA's larger population. The authors also colored the three countries according to their population growth rates, which can not be duplicated in this black on white paper. Not only can one visualize the population sizes of these three countries, but also they can see the US is growing at a smaller rate than the other two. In fact, at the time of this book's publication in 1981, Mexico (at 3%) was growing three times as fast as the US (at less than 1%).



A North American Population Map, similar to one in The State of the World Atlas.

Deciphering the SASware Ballot

Any presentation should be user-friendly. One should be able immediately to pick up a chart or a table, and with only minimum effort, decipher its meaning. Ideal general software such as word processors and spreadsheets should be mastered within fifteen minutes of dabbling. The user manual should only be required for occasional reference. The SASware ballot, beloved wish list of

SAS® software programmer needs and desires, surely would fail any reasonable user-friendly criterion. After examining the SASware ballot of SUGI'23 at the conference last year in Nashville, I left more bewildered than enlightened.

Consider this actual fragment of the SASware ballot, reproduced here:

1998 SASware Ballot Results

I. SAS Language

General

<u>Overall Rank</u>	<u>Part Rank</u>	<u>Number of Votes</u>	<u>Item Number</u>	
4.0	1.0	676	23	provide the ability to create stand-alone executables.....
9.5	3.5	552	22	allow options on the Title....
11.0	5.0	541	1	provide the ability to display.....
.
29.0	14.0	41	512	allow separate justification.....
55.0	20.0	340	9	add an option to display
.....				

This is not a particularly sophisticated survey and the ordinary user easily should understand it. However, I'm puzzled by the results. What is an overall rank? How does it differ from the part rank? Is the item number useful if the request description is included? Extraneous information is "chartjunk," (Tufte's term) cluttering the design, but is not useful. It is not cost-effective with precious space, limited to an 8.5 x 11-inch sheet of paper or a screen on a computer monitor.

Does it matter how many votes the item received? Suppose the number of votes did matter. What was the total number of votes cast? Per section, how many votes can one cast? Without the context, the number of votes has little meaning. In the above chart, 676 votes are much more than 552 votes, but that doesn't relay much information. Suppose one could cast three votes out of ten choices, ranking them by priority, then one could see a ranking based on weighting. Is that what the "Part Rank" and "Overall Rank" fields are trying to tell us? If so, a simple explanation along the border of the table would be useful.

The next to last entry in that fragment shows that item #512 is ranked higher than item #9, but it received only 41 votes, versus 340 votes for the latter. Not only are the ballot results confusing but seemingly illogical. Even with a liberal weighting scheme, it would appear that an item receiving roughly eight times the number of votes would be higher ranked by the SAS software sub-culture.

Obviously, the SASware ballot suffers from too much information that is not usable and from a lack of clarity. Even a simple explanation on the terms and method used would be a move to user-friendliness. However, for purely transmitting usable information, I would remove most of these fields. Instead, I would create a more detailed chart with full explanations on a web page for those who desire to drill down to that level of detail. This ballot, while of immediate interest of SUGI attendees, might better serve the SAS software public on a less static medium, such as a web page. This technology would allow the SAS Institute to add their own updates on the status of this wish list. Here is what an improved SASware ballot on a web page might look like (cleaner and more informative):

1998 SASware Ballot Results

I. SAS Language (Ranked by Priority of SAS Users)

(For more detailed results with statistics, [click here](#))

General

<u>Rank</u>	<u>Item</u>	<u>Status (with date)</u>
1	provide the ability to create Stand-alone executables	8/31/1998- No way!
2	allow options on the Title....	6/04/1998- In testing stage (for more details).
.	.	.
.	.	.
7	allow separate justification	5/15/1998- Will be implemented in release 7.01
8	add an option to display	9/30/1998- On hold
.....		

(note: I used a subdued gray color bar on alternate lines of text to help the eye in reading across the line.)

Summary:

I have hit the high spots but only touched upon information visualization issues. I took the approach of displaying and critiquing examples of flawed presentations to make the point of what one should avoid. This preferred method is learning through bad examples. Or rather, learning through what doesn't work well.

The underlying message throughout this paper is that while presentations can be used to promote one's organization or views, the honest presentation would provide primarily for an effective transfer of information. All other messages should be secondary for the truly impartial information disseminator. As statisticians seek to provide unbiased analysis in all aspects, the applications developer should use this same methodology as well. The message needs to be clean and authentic. Above all, the message needs to be understandable.

This paper should help presentation designers think clearly about their visualization products. Careful thinking will yield clarity of idea presentation- and acceptance.

References:

Allison, Alexander W., et al., Eds. The Norton Anthology of Poetry, Revised. New York: W.W. Norton & Co., 1975.

Fussell, Paul. Bad, or the Dumbing of America, New York: Summit Books, 1991.

Kildron, Michael and Ronald Segal. The State of the World Atlas, New York: Simon and Schuster, 1981.

Shakespeare, William. The Complete Works of Shakespeare, New York: Oxford University Press, 1938.

Strunk, William S., Jr. and E.B. White, The Elements of Style, New York: Macmillan Co., 1959.

Tufte, Edward R. Envisioning Information, Cheshire, Connecticut: Graphics Press, 1995.

Tufte, Edward R. The Visual Display of Quantitative Information, Cheshire, Connecticut: Graphics Press, 1990.

Tufte, Edward R. The Visual Explanations, Cheshire, Connecticut: Graphics Press, 1997.

SAS, SAS/GRAPH, and SASware Ballot are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration

Other brand and product names are registered trademarks or trademarks of their respective companies.