Infovis as Seen by the World Out There

2007 in Review

Stephen Few

Perceptual Edge
&
University of California, Berkeley

Those of us who have now sat through several days of presentations have been immersed in the details of many research studies. Our brains are throbbing from trying to cram so much in so quickly. The last thing you need now is for me to take you on yet one more microscopic tour. Instead, I'm going to invite you all to relax, to clear your heads, and to back way up to a point far enough out in space to see the big picture of what has happened in infovis during the last year or so.

Not only will I invite you to zoom out for a broader view, I'll also ask you to completely shift perspective from the familiar view of an infovis insider to that of the greater world out there—the perspective of an outsider. There are people in the world out there who stand to benefit from what we do, and the view that they have of infovis is usually different from our own. It's important that we understand it if we wish to make our efforts count in the world.

Although I'm an insider, my involvement in the academic infovis research community is relatively new, and my interest in information visualization didn't emerge from within academia. My involvement began naturally as a businessperson in search of solutions to persistent business problems. I was searching for better ways to help people make sense of business data and then communicate what they uncovered to decision makers—not primarily to statisticians, scientists, or engineers, but to everyday business people whose jobs depended on their ability to make smart decisions based on data, especially quantitative data.

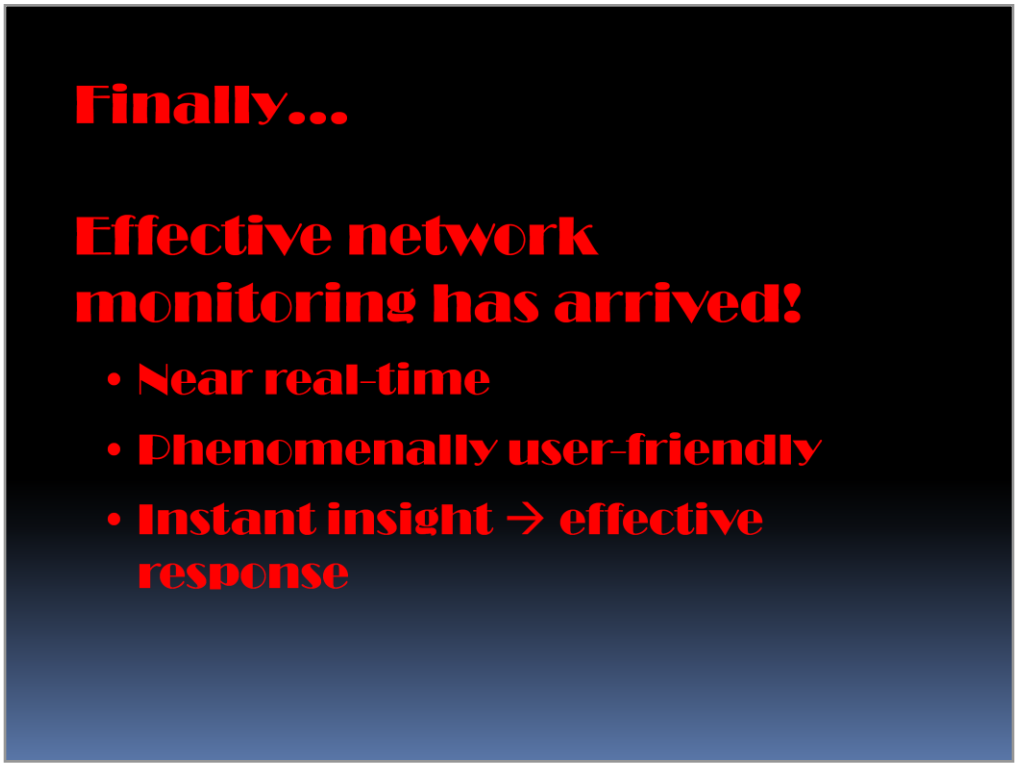
I loved Ola Rosling's analogy during yesterday's panel about social visualization. In regards to our relationship to the world out there, he said to imagine photographers responding to the proliferation of cameras throughout the world with the attitude:

> *This is a disaster. What will become of photography now that the huddled masses have them. These novices don't know what they're doing. They wouldn't know a good photograph if it bit them. If they're going to play in our space, they'll have to follow our rules or get the hell out. Photography will be ruined and our profession will cease to exist unless we can stop this travesty.*

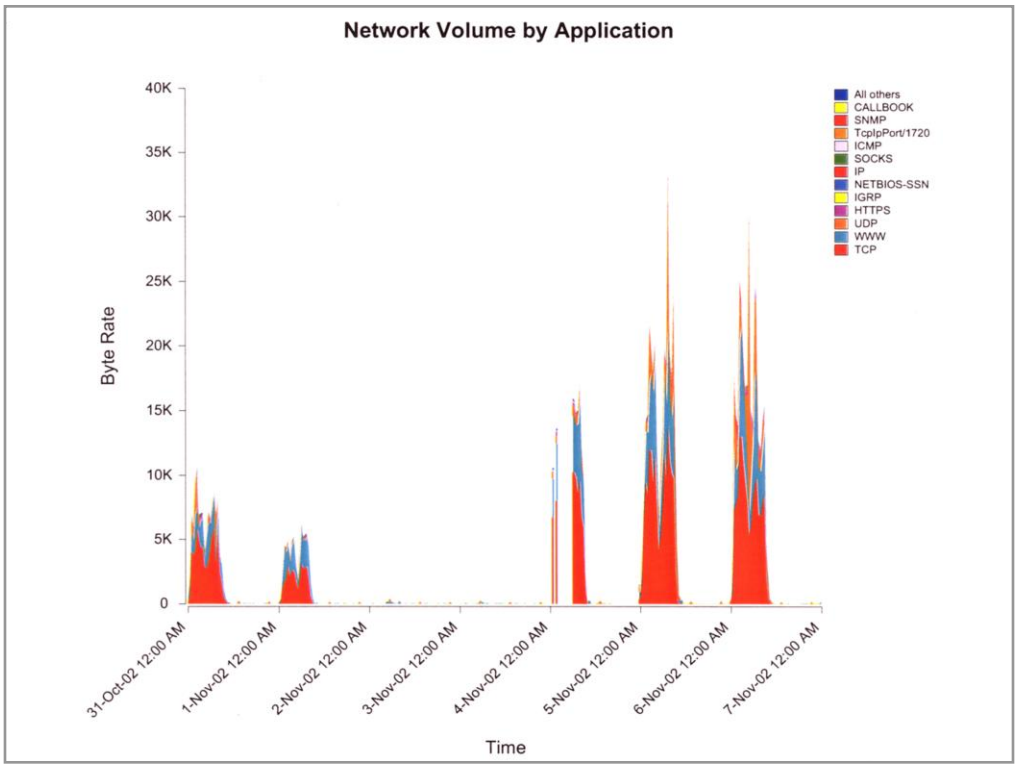
When Ola said this the room responded with a nervous laugh of embarrassment. His words tapped into a thought and a feeling that many of us have had. Of course we ought to invite the world to play in our sandbox. Together, we can build some beautiful sandcastles.

I want to extend Ola's analogy, however, to the world of people with whom I work. Few of them prior to meeting me even knew that most of the visualizations that we have created exist. All, however, have been exposed to visualization in a primitive form, however. To stay true to Ola's analogy, these folks try to photograph the world using ancient brownie cameras with dirty lenses and expired film and they were never trained to use. They just point and shoot, trusting brownie to do a heck of a job. When I show them my Canon Digital SLR camera and the pictures it can take with proper training, they can't wait to get their hands on it. They not only want better visualization tools, but they want guidance in how to use them effectively.

Most of the 25 years that I've worked in information technology have been spent focusing on what was originally called "decision support," then morphed into "data warehousing," and is today called "business intelligence," although there is increasing pressure to replace this name with "business performance management."



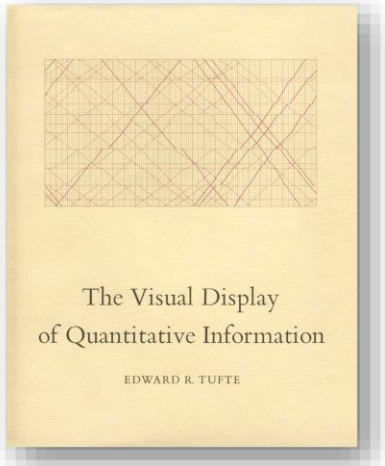
Several years ago, I managed a large team of business intelligence professionals at one of the world's largest software companies. One day I walked into a meeting, took my seat at the conference table along with several others, and waited for yet one more boring presentation to begin. Eventually, a young fellow eagerly stood at the front of the room and opened the meeting with a slide that was something like this. Bursting with excitement, the speaker announced that everyone in the room would now receive a daily report that would tell us how the computer network was being utilized so we would no longer be in the dark about what was going on, and then the graph on the next slide appeared.

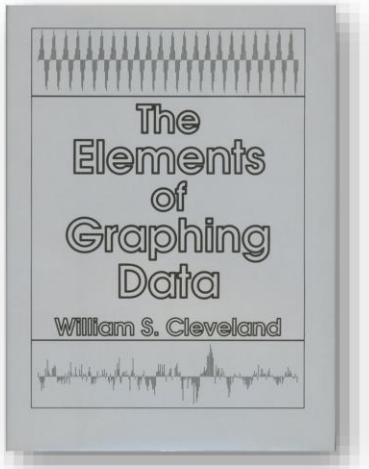


I sat there with my mouth hanging open in disbelief, but what shocked me most was the fact that everyone around me seemed to be nodding with enthusiasm, and none more enthusiastic than the Chief Information Officer. If I had asked for anyone in that room to tell me what the graph said, no one could have told me anything intelligible, but they were enthusiastic nonetheless. It occurred to me that perhaps no one was willing to admit to being the only person in the room that didn't get it.

It was on this day that I realized: "We have a problem."

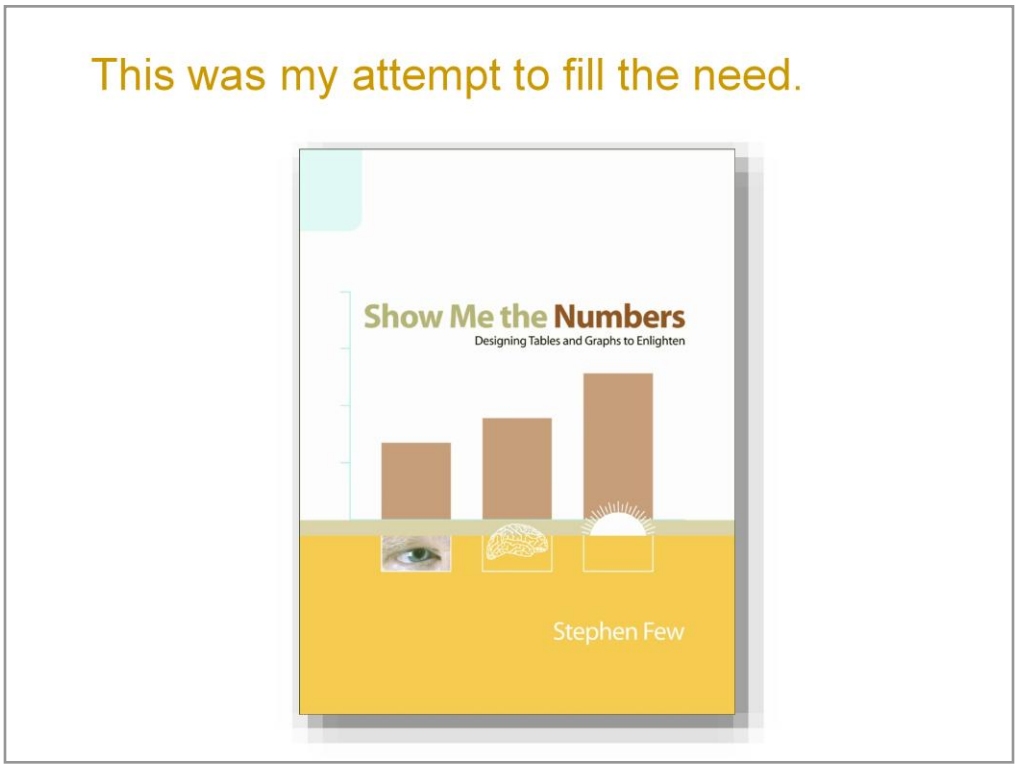
The world needs accessible and practical training resources.





These don't teach practical skills for non-statisticians.

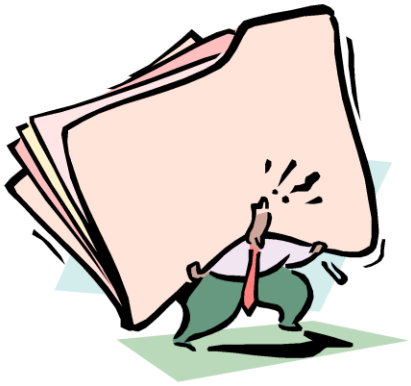
Even among my team of business intelligence professionals, people whose fulltime jobs involved analyzing and reporting information, no one had the skills that were needed to communicate quantitative information effectively. To remedy this problem, I began to search for resources that they could use to learn these skills. I was already familiar with the work of Edward Tufte, but considered his books too theoretical and disconnected from the practical needs of business. I found several other sources, some of them quite valuable, such as books by William Cleveland, but none that taught the simple, practical skills that business people needed in the language that business people would find familiar and relevant. After awhile, I realized that the resource I was looking for did not yet exist, so I decided that I might as well be the one to write it.



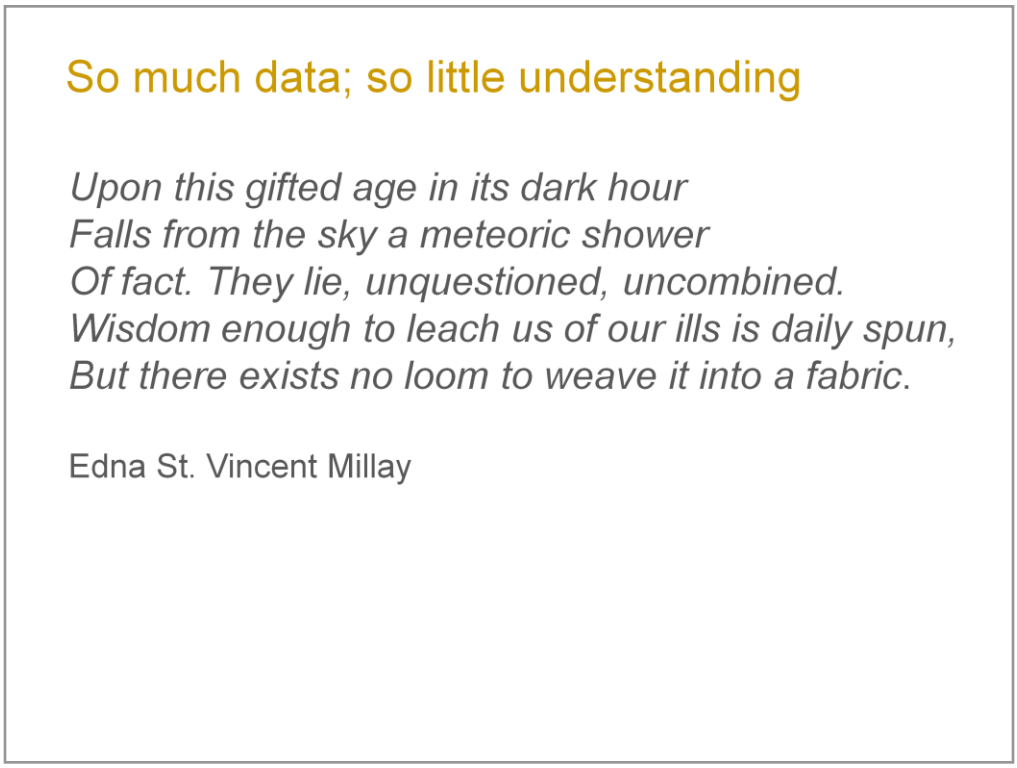
I quit my job, proceeded to fill in the gaps in my knowledge, including extensive reading of infovis research, and eventually started writing, which resulted in the book *Show Me the Numbers*.

The world is buried in data.

"Just show me the numbers!"

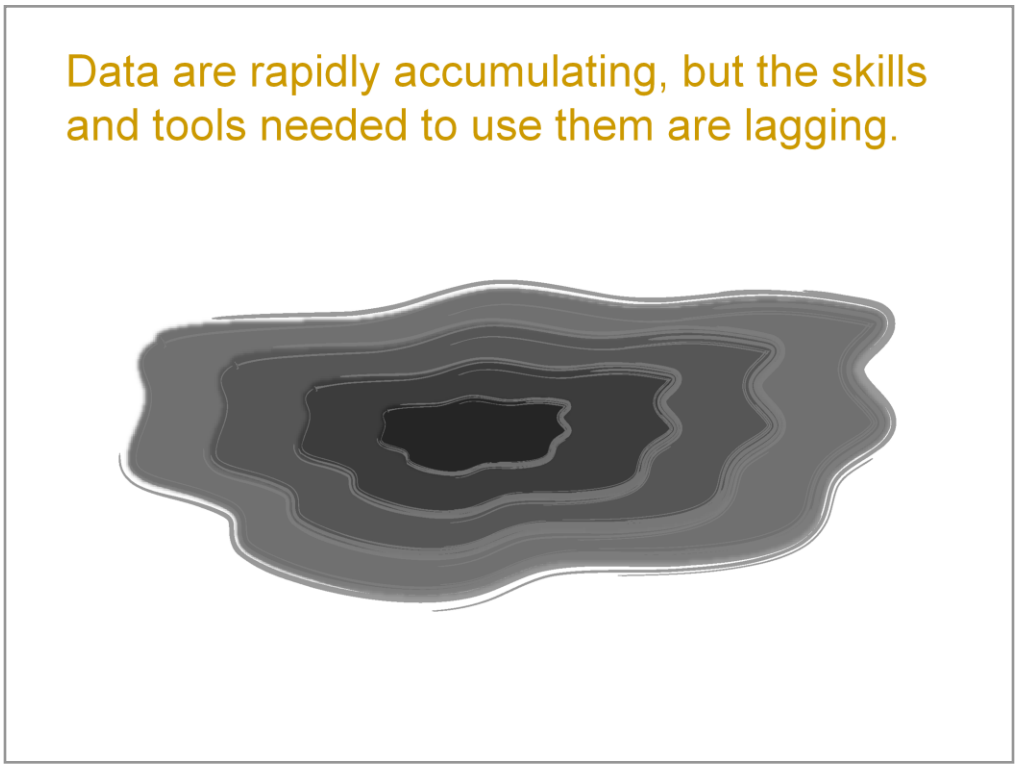


I didn't choose this title arbitrarily. I actually heard this expression at times on the lips of business people, especially sales executives, who ran around in frustration saying "I don't need your reams of reports—just show me the numbers that I need to see to know what's going on."



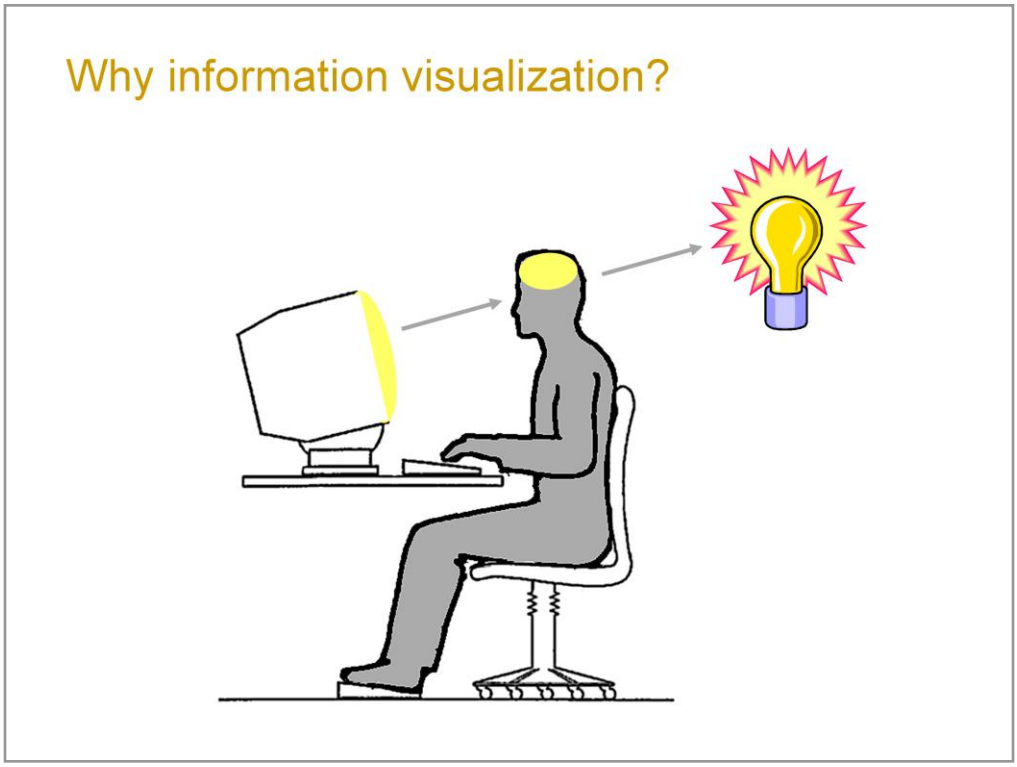
This poem by Edna St. Vincent Millay eloquently describes our situation today. The amount of information that is available has increased dramatically in the last few years, but the ability to make use of it has increased little, if any.

> *Our networks are awash in data. A little of it is information. A smidgen of this shows up as knowledge. Combined with ideas, some of that is actually useful. Mix in experience, context, compassion, discipline, humor, tolerance, and humility, and perhaps knowledge becomes wisdom.*
>
> *Turning Numbers into Knowledge*, Jonathan G. Koomey, 2001, Analytics Press: Oakland, CA page 5, quoting Clifford Stoll.



Technology has done a tremendous job of enabling the world to capture, store, and distribute massive amounts of data. Far too little effort has been made, however, to help people make good use of that data—to uncover the important meanings that reside therein and to communicate those meanings to others in ways that are clear, accurate, and useful.

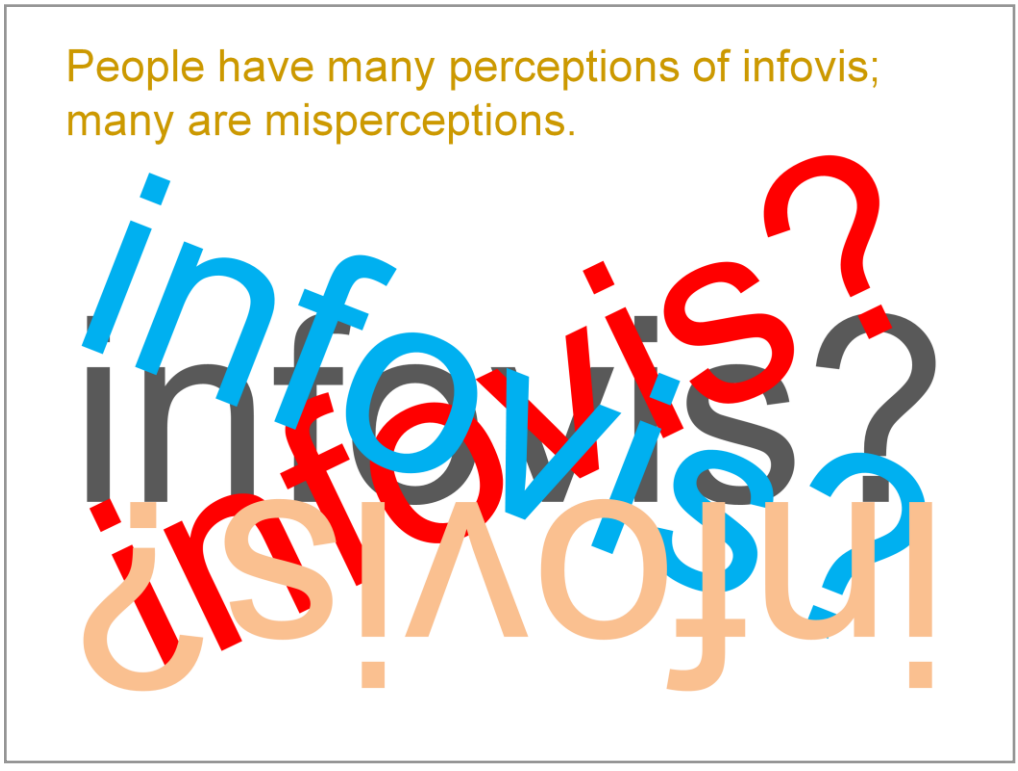
Most people who are responsible for analyzing data have never been trained to do this. Knowing how to use Excel or some other software that can be used to analyze data is not the same as knowing how to make sense of data.



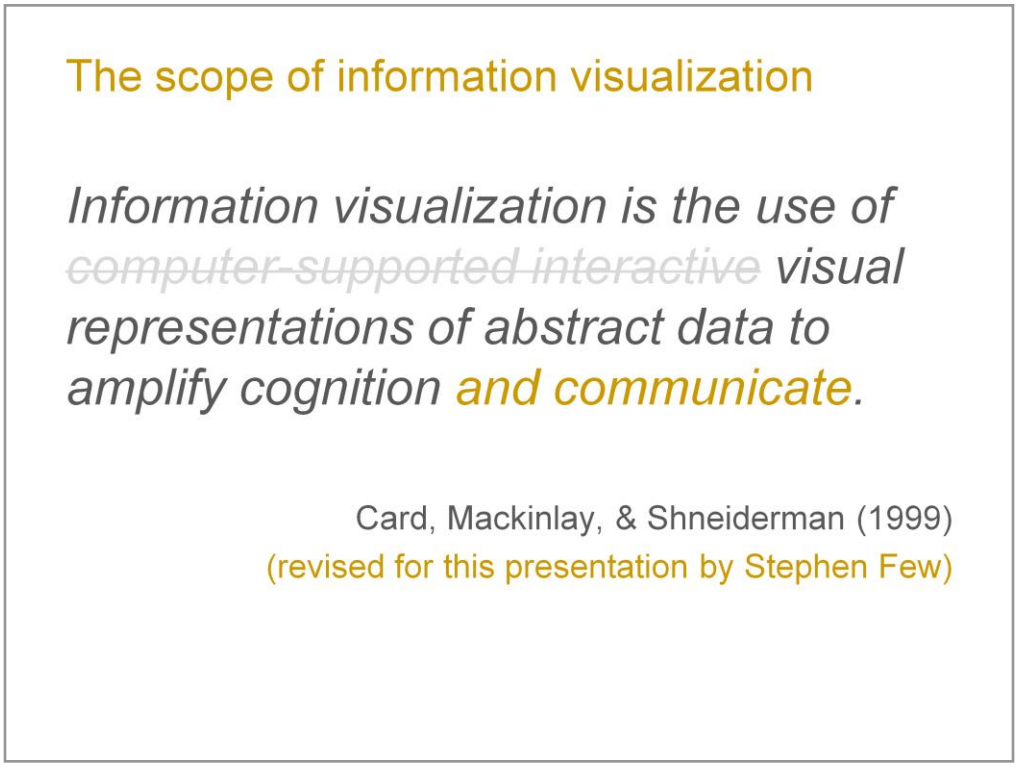
I became fascinated by information visualization—the use of our eyes in close collaboration with our brains to make sense of and communicate information—not because I'm a visually oriented person (I in fact am much more verbal than visual) and not because I find the science of this field more interesting than other areas or the algorithms more elegant. I began to explore and eventually specialize in infovis because it offers solutions to real problems in the world that concern me, solutions that stand apart in their ability to enlighten.

Through information visualization I have found a way to give something useful to the world. This is what drives me; this is what makes me excited and proud to be a part of this community. This is also, however, what brings me so much pain and makes me so frustrated and even angry when I see infovis done poorly, see it distracted by pursuits that simply don't matter, or see it presented to the world in ways that misrepresent what it is and undermine its value.

I began by saying that I want to help you see infovis this afternoon from a different perspective. How we as insiders see and understand InfoVis is quite different from how it is seen by most people in the world out there. Most people get only brief glimpses of what we do, and those glimpses rarely tell the story clearly.



Think about what people outside of our community have seen of infovis this year through marketing, blogs, and other sources. This picture is rife with misperception that makes it difficult to catch glimpses of the truth. I'd like to take you on a tour of what the world has seen of InfoVis in the last year or so: both the promising highlights and the discouraging failures that have shaped the world's perception of our beloved and important work. The world needs what we do, but its needs remain largely untouched and definitely unsatisfied.



Let me pause for a moment to say that I am intentionally using the definition of "information visualization" a bit more broadly today than the formal definition. I like the definition that Card, Mackinlay, and Shneiderman gave us in their book *Readings in Information Visualization*. I use it often in my work. Today, however, I'm expanding the scope of visualization to fit what I usually refer to as "data visualization" to differentiate it from the strict definition of infovis. Today I'm including in the scope static visual representations of abstract data, which may appear on a computer screen, but could also appear on paper, and I'm adding to the process of amplifying cognition the goal the process of communication. Although there are some differences in the design and use of visualizations that are meant for sense-making versus those that are meant for communicating information to others, the visual representations themselves follow the same basic rules. And besides, the process of sense-making and the process of communicating what we discover in the data are intimately connected and complementary.

As we look back at the year 2007, there is no doubt that we will long remember it as the year of one particular great leap in infovis technology. Everyone in the infovis certainly community knows that one technological breakthrough of this year has caused all else to pale in comparison. Rather than keeping you waiting for the obvious, I'll jump immediately to the unparalleled infovis innovation of 2007.

I am obviously referring to the ambient orb. No longer will anyone need to remain in the dark about what's going on, for you will constantly be kept informed by this simple but elegant glowing globe of many colors. When I first ran across an ad for the ambient orb about a month ago, I read about what the New York Times has hailed "the newest concept in how to monitor everyday data." I sat in awe with nothing to say but "Why didn't I think of that?" Quoting from the advertising copy itself:

The Ambient Orb may look like a crystal ball on acid, but it's really more of a giant mood ring—plugged straight into the fluctuations of the stock market or anything else you care to track. The orb can be wirelessly configured to track any individual stock, any market index or your personal portfolio. "People want information, but they don't want to invest a lot of time in getting it," says Ambient president David Rose. "This makes getting information a 'glanceable' thing."

Not easily outdone, however, and unwilling to spend $150 plus shipping to get one of my own, I decided to improve on the orb with an invention of my own. I call it the “Illuminating Lava Lamp.” Unlike the orb, it isn’t limited to a single message at any one time; it is able to present multivariate information. Just think of every possible shape that these globules of wax can form, a rich palette of information about an infinite number of variables, simultaneously displayed and simultaneously perceived through the enormous power of visual perception. Training isn’t required, because the shapes coincide perfectly with Bertin’s visual vocabulary to communicate intuitive messages, much like the hieroglyphics of ancient Egypt.

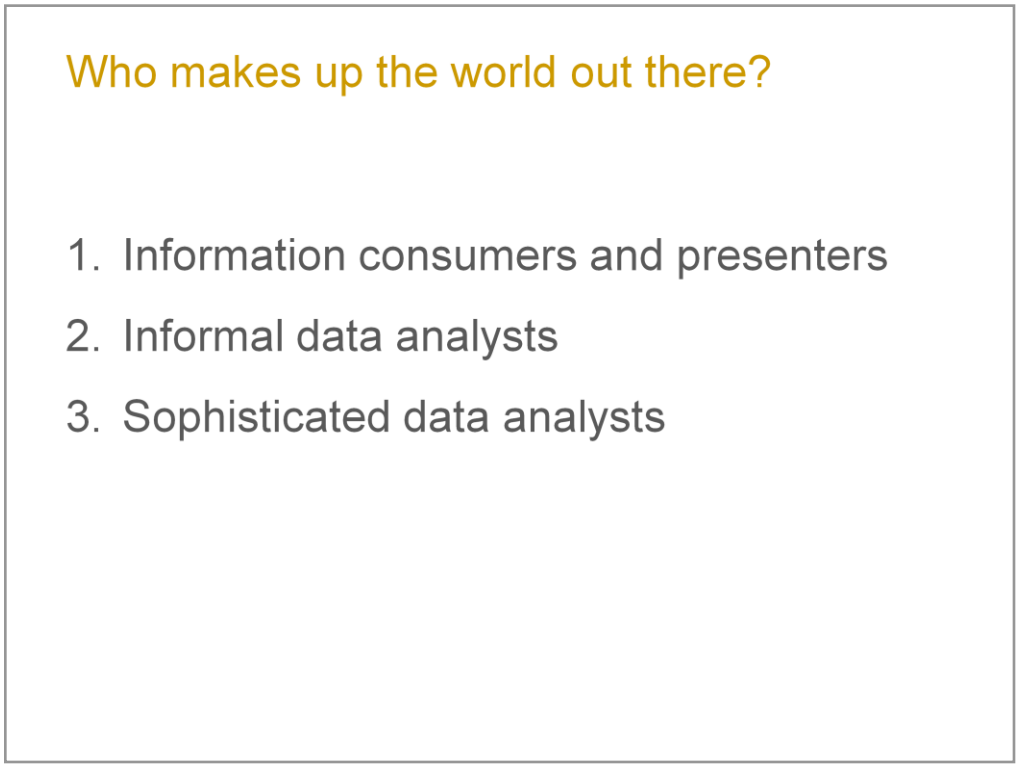
For serious practitioners of infovis, I simply call it the “Illuminating Lava Lamp,” but to make it appealing to the world at large, I extend its name to the “Temporal Hyperbolic Reality Illuminating Lava Lamp”—THRILL. It isn’t yet available commercially, but it will soon be available as a prototype in several stylish colors from www.info-obsure.com.

Before looking at what's happened specifically in the world of infovis, let's first take a moment to look at a few of the significant events in the world at large that have affected the way people view infovis.

The war in Iraq continues to poignantly remind us that we invite a world of trouble and suffering when we make decisions that are not based on sound evidence and reason. This sensitizes world citizens to the need for better information; the ability to extract what's meaningful and true from the onslaught of data that bombards us.

The eventual victory of good science over powerful corporate interests has finally awakened the world to the inconvenient truth about global climate change. This awakening has given us reason to believe that good science and honest research about things that matter can win the minds of the public, even against powerful corporate and political opposition. It also sends a stern warning that by ignoring or resisting good science for too long, the cost of delays can be devastatingly expensive, potentially beyond the point of redemption.

The book *Super Crunchers* by Ian Ayres argues that the world, especially the world of business, is turning more and more to the statistical analyses of large datasets to inform decision making as an alternative or at minimum a complement to traditional reliance on the intuition of experts. Ayres makes a case for the superiority of statistical analysis—evidence based decision making. What he doesn't do directly is something that we must: We must answer the question, "Does infovis have a role to play in the world of super crunchers, and if so, what should we be doing to support the effort?"



Who makes up this world out there, those who are not a part of the infovis research community but stand to gain from what we have to offer? Everyone who needs to make sense of information and to communicate what they find to others in an effort to make better decisions. This is not a neatly defined, homogeneous audience. It is made up of people of all types in all walks of life with varying needs and perspectives. One way to separate these folks into meaningful groups is to categorize potential infovis users based on the frequency and complexity of their information needs. One possible classification consists of the following three groups:

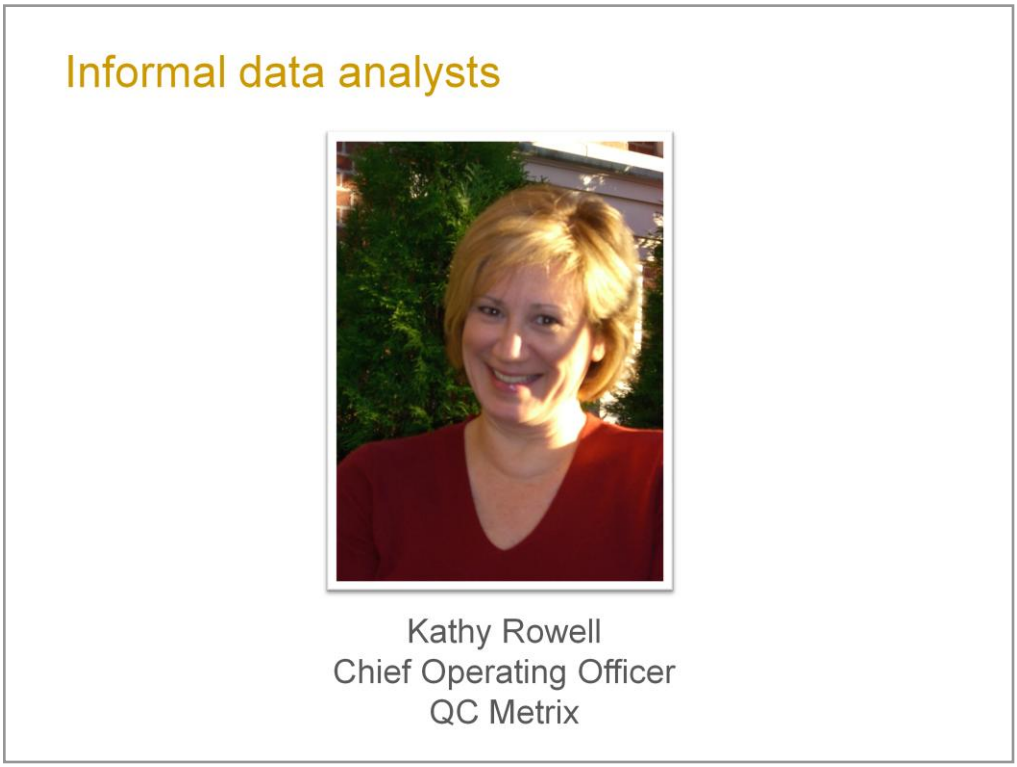
1. Information consumers and presenters (around 80%)
2. Informal data analysts (around 19% and growing)
3. Sophisticated data analysts (less than 1% and in need of growth)

I believe that the first group is by far the largest today, consistently perhaps of 80% of infovis users. The second group is much smaller, perhaps around 19%, but is beginning to grow. The third is tiny by comparison, less than 1% of people who use infovis in the world of business, but desperately in need of growth.

Information consumers and presenters

Christian Rohrer
Director of User Experience Research
eBay

Christian Rohrer, the Director of Customer Intelligence for Products at eBay, resides firmly in the first category of information consumers and presenters. Christian leads three teams at eBay that are concerned with new product opportunities, user experience research, and product health. These global teams of researchers focus on delivering insights to inspire, inform, and assess user experience and product designs, utilizing methods such as ethnography, usability, data mining, and survey research. As their manager, Christian consumes the products of their research, often presented in visual form, and then must report those results in simple terms to senior management, again often relying on graphical means to tell the stories, usually in PowerPoint.



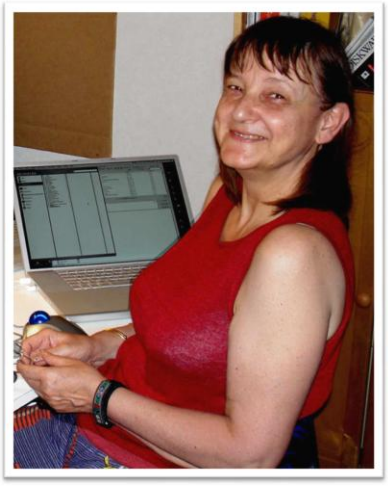
Kathy Rowell, the Chief Operating Officer of QC Metrix, represents the second group. QC Metrix collects, analyzes, and reports back on medical information, especially surgical outcomes, to about 200 client hospitals. Even though she is an executive, she gets her hands into the analysis of these data personally because she cares deeply about it. Through their work they have been able to detect several post-surgical problems that affect patients that were not on the hospitals' radar, resulting in dramatic improvements in post-surgical care, such as the reduction of infections. Kathy is not a sophisticated, full-time statistician, but with the right infovis tools, she is able to make important discoveries that are used to improve the lives of thousands of people.

Sophisticated data analysts

Marcel Baumgartner
Business Analyst
Nestle Corporation, Switzerland

Marcel Baumgartner of Nestle Corporation in Switzerland is a good example of someone who belongs to the sophisticated analyst group. I became acquainted with Marcel when he invited me to teach two of my courses for the Swiss Statistical Society, which I did last June. Marcel's work focuses on the analysis of data related to Supply Chain Management, for example, indicators of forecast accuracy and customer service levels. To better understand how these indicators change over time, and to what degree they depend on each other, he needs to visualize the data, which he does using S-Plus and Tableau. Another aspect of his work is related to forecasting demand. Marcel believes that to judge the plausibility of statistical forecasts, visualization is a very efficient means to make a good judgment (e.g. by showing historical and forecasted together).

Another sophisticated data analyst



Anne Chappuis
President & Associate Professor
Decision Graphics & University of Rouen-Normandie

Another person who fits into this category as a sophisticated real-world user of infovis, is also an infovis insider. Anne Chappuis began doing visual data analysis in the 1970s. A native of France, she refers to herself as a disciple of Jacques Bertin and has made extensive use of his work. I became acquainted with Anne only recently when she sent me an email to say that she appreciated my work and considered me someone who applied Bertin's work effectively. She is currently working in India where she started an NGO, Visual Information Systems for Action (VIStA) (www.vista-info.net) . She wrote in an email to me, "We believe that the villagers, even if illiterate, are clever and capable of analysing their situation and planning for their future if they are provided with information, and specially with visual information." Over the years she has relied heavily of visual analysis tools to support many organizations in France and India.

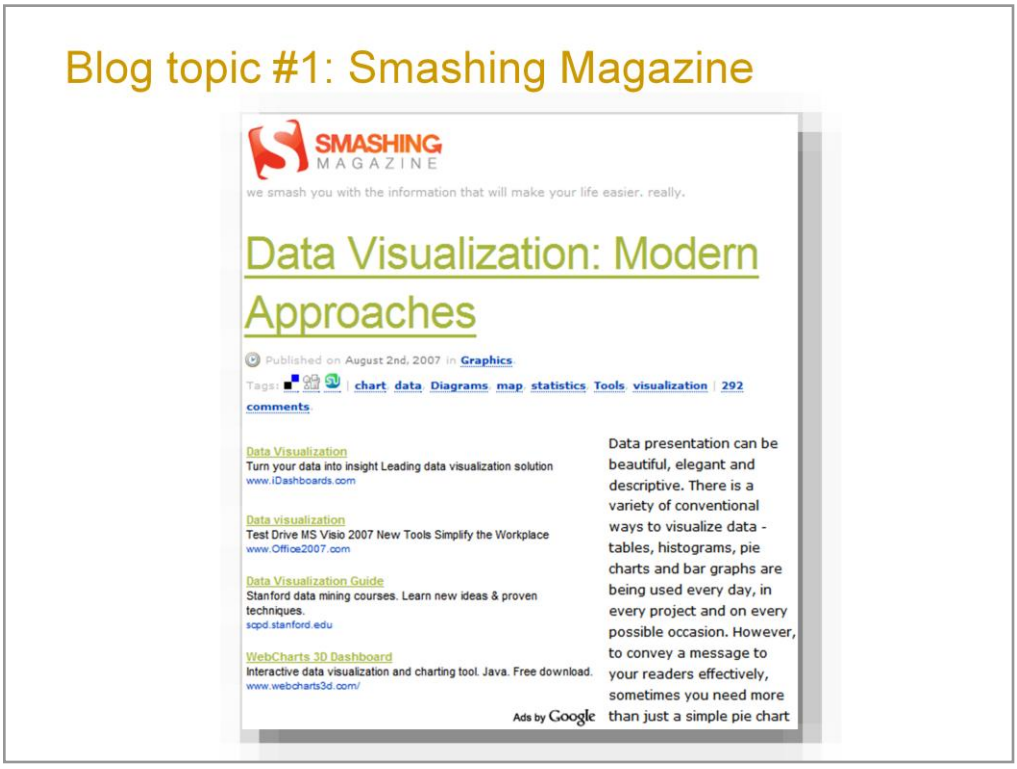
People outside of the infovis community don't learn about our work by reading our research articles. I hope this isn't a surprise to you. Their exposure to what we do comes mostly from four sources:

1. Blogs
2. Product marketing
3. Books
4. Presentations by the *Visualati*

As you know, for good or ill, blogs have increasingly become a source that people turn to for information. I use a series automated Google searches that run every day to find anything that has been posted about infovis on the Web, including in blogs. Blogs occasionally spread information about some aspect of infovis far beyond the walls of our community. Unfortunately, they don't always present that information accurately and they certainly latch onto bad examples of visualization just as often as good, perhaps more often.

During 2007 and extending back a ways into 2006, three infovis-related topics have created significant buzz in the blogosphere.

Blog topic #1: Smashing Magazine



The first topic the caught the attention of bloggers and was exposed by them to a broad audience was an article by Smashing Magazine (www.smashingmagazine.com) entitled “Data Visualization: Modern Approaches”. Here’s what Vitaly Friedman, the publisher of Smashing Magazine, said about the article.

> *The showcase of modern data visualization methods has been created just because we were extremely impressed by Infographics.com and methods presented there. So we’ve searched for further methods - for three-four days in a row. The search was worth it. Take a look what we’ve come out with in the end. It’s extremely fascinating to observe that even University professors of data visualization have found something useful in our research. This is what makes our site keep going on; this is what helps us to maintain the quality of our work.*
>
> (From an interview with Vitaly Friedman of Smashing Magazine in Fadtastic, September 3, 2007)

Unfortunately, Smashing Magazine has no expertise in data visualization and as a result, exercises no discrimination in what they present as viable solutions. The exposure that they provide is a mixed blessing. In this more recent article about graphs, they state that "Server-based solutions, implemented with Flash, JavaScript or pure CSS, offer a more flexible alternative." But almost all of the solutions that they feature include horribly designed graphs like these.

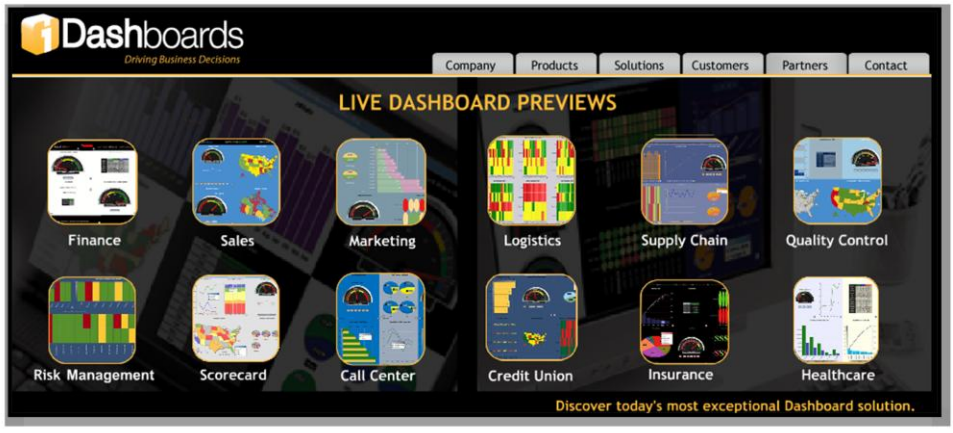
Their readers respond almost unanimously with accolades, with the exception of one voice of reason:

> *I am really troubled to see more and more articles like this. Most of these tools create pretty flashy charts, that's for sure, but they don't teach anyone how to visualize data properly. Why would you ever need a donut chart or a 3D stacked cylinder chart? What does that add to the data that a simple 2D bar chart can't show you? There is so much of what Edward Tufte refers to as "non-data ink" on these charts that it just smothers the "data ink" completely.*
>
> *We need to teach analysts the techniques to use simple and straight-forward charts to let the data shine through. All these different primary colors and 3D stuff have nothing to do with the data. Check out the work of Edward Tufte and Stephen Few for some good visualization principles…*
>
> *Rian*

The ads illustrate this lack of discrimination.

Smashing Magazine's lack of discrimination should concern us, because it has an affect on how the world out there perceives visualization. I think that the Google ads that appear along with the article aptly illustrate the problem.

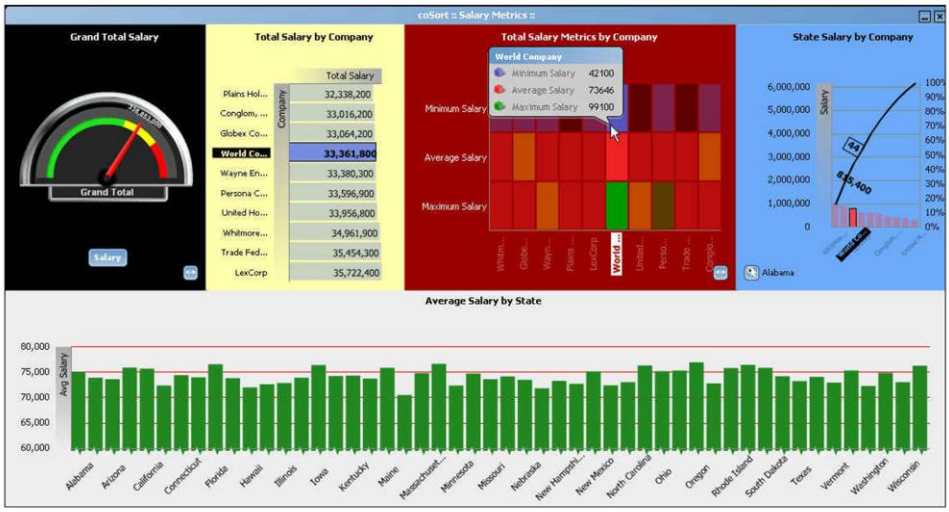
Ad #1: iDashboards



Business Intelligence vendors are gaga for widgets.

The first ad is for iDashboards, one of the prominent dashboard software vendors in the business intelligence market. I'll talk a bit about dashboards, one of my areas of specialty, a little later. For now, let me simply point out that I have never seen a dashboard produced with the iDashboards products that could be used effectively. The dashboard market is dominated by the notion that the more dazzling and colorful the dashboard appears and the more tricks the display widgets can do, the better the display. Few of the vendors who are selling dashboard products have bothered to ask the question, "How should a dashboard display be designed to help people monitor what's going on in the business efficiently and accurately?"

An iDashboards example



Here's an example, which unfortunately is all too typical. This was created by the person who developed the iDashboards software and wrote a book about dashboards. If you put yourself in the position of a Human Resources Director who had to rely on this to monitor compensation across several companies, you would spend hours to get very little insight into what's going on.

Ad #2: Microsoft Office 2007

Microsoft Office suggests that we should focus on making our graphs pretty.

The second ad features Microsoft Office 2007. The video that this ad links to presents Microsoft Office's information visualization vision featuring the new charting functionality of Excel. I'll talk about this more in a little while when I show you some of the marketing messages that have influenced people's perceptions of infovis. For now, let it suffice to say that the new charting engine renders better graphics, but it did nothing to improve the charts' ability to communicate data effectively.

Ad #3: Stanford statistics courses

Academia can appear theoretical and intimidating.

The third ad is for a data mining course offered by the Statistics department at Stanford. I'm all in favor of statistics courses and believe that businesses especially would benefits from a much deeper and broadly distributed understanding of statistics. This ad, however, reinforces one of the notions that many business people have about infovis: that it is complicated and only relevant to statisticians.

Ad #4: WebCharts

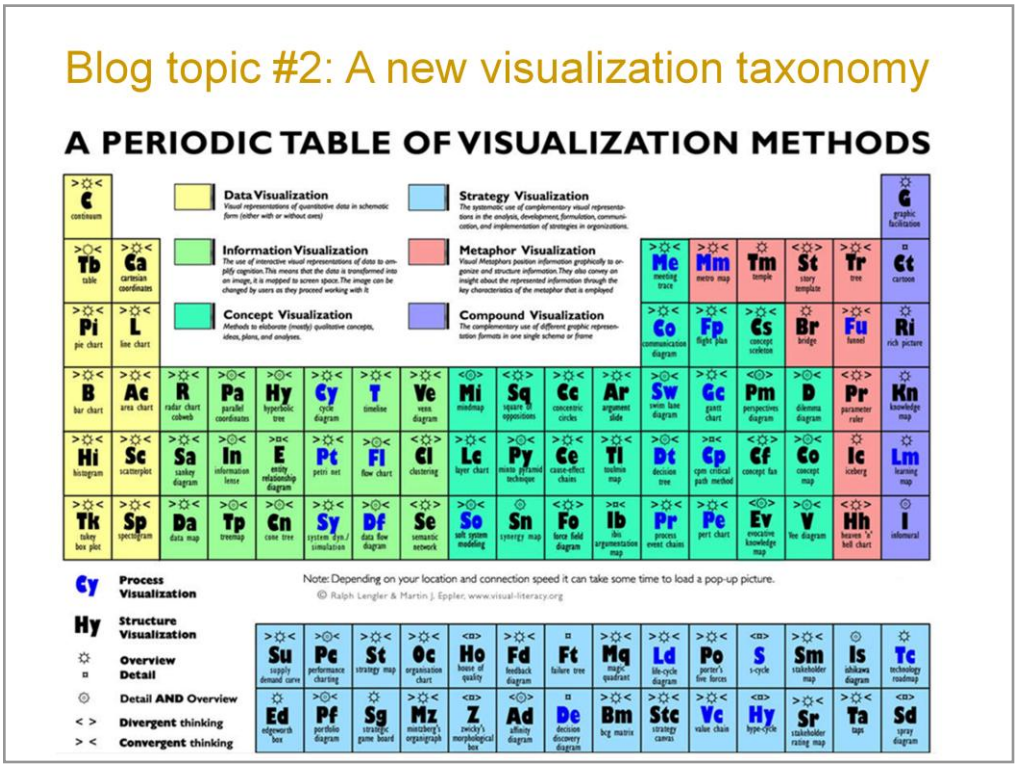
Sheer eye candy with empty calories

The fourth and final ad features one of many web-based graphing products, which almost unanimously focus on dysfunctional visual effects with no clue whatsoever about the purpose of graphs and how they should be designed to present data effectively. They look like they were developed by bored engineers who tried to outdo one another with nifty tricks and cute visual features that make the graphs look like perceptual puzzles.

The view of infovis presented by these ads

Either cute widgets and visual fluff, or the realm of highly-trained statisticians.

Each of these ads promote one of two perceptions of infovis that are common in the world out there, especially among business people: Infovis is either about making data look visually dazzling or its is about highly sophisticated tools that are only relevant to and approachable by the statistical elite.



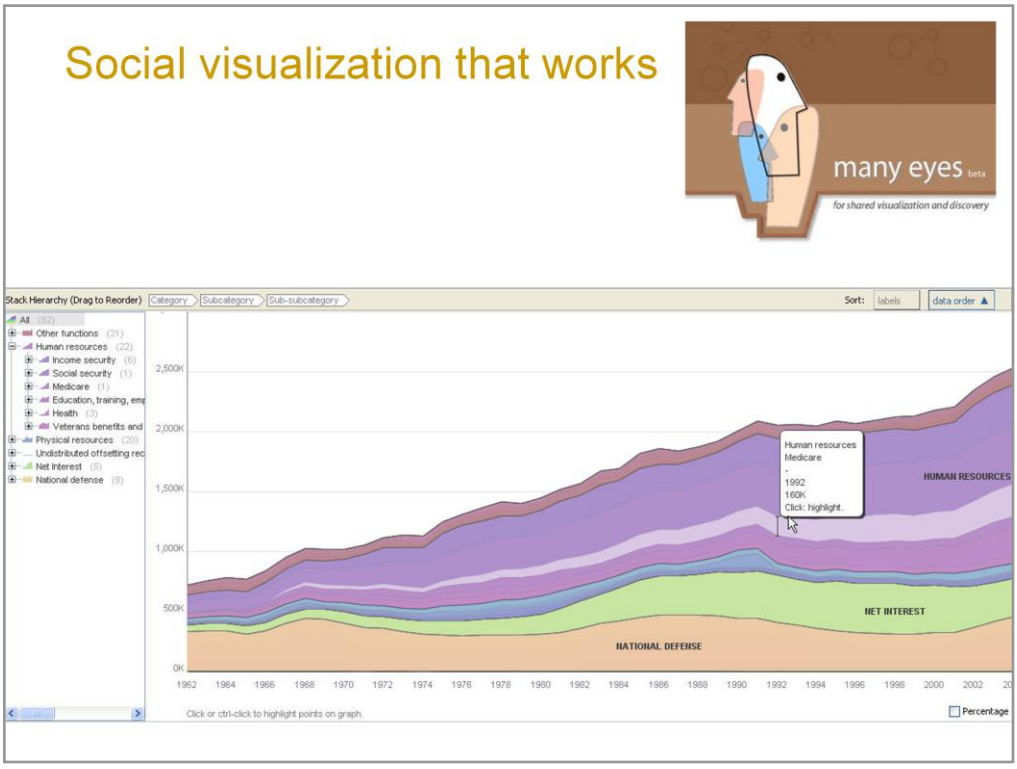
Another infovis related topic that made a splash in the blogosphere beginning in January of this year, with ripples continuing today, is the taxonomy of visualization methods called “A Periodic Table of Visualization Methods” created by two fellows who work at the Institute of Corporate Communications at the University of Lugano, Switzerland. Perhaps even more than the Smashing Magazine article, this has exposed a broad audience to visualization, and for that I suppose we can be grateful, but the effects of this exposure have a downside.

I’m constantly haranguing those business intelligence vendors that dabble in infovis without understanding it for their clueless designs, but here’s a case where the academic community, which I hail as the source of what’s good and true, designed a presentation of visualization methods that breaks several rules of visual communication.

The original Periodic Table of the Elements, unlike this table that was modeled after it, is quite effective in design. It works because its organization reflects the meaningful attribute of each element’s atomic number (the number of protons in the elements atomic nucleus). The two-character abbreviations that it uses to label each element work because they are the standard abbreviations that are used by scientists. This new periodic table of visualization methods, however, exhibits no such organizing principle. What is the point of doing an entire research project to force a list of visualization methods into a unrelated paradigm? Apart from breaking the visualization methods into general categories (information visualization, concept visualization, etc.), this table exhibits no useful organization.

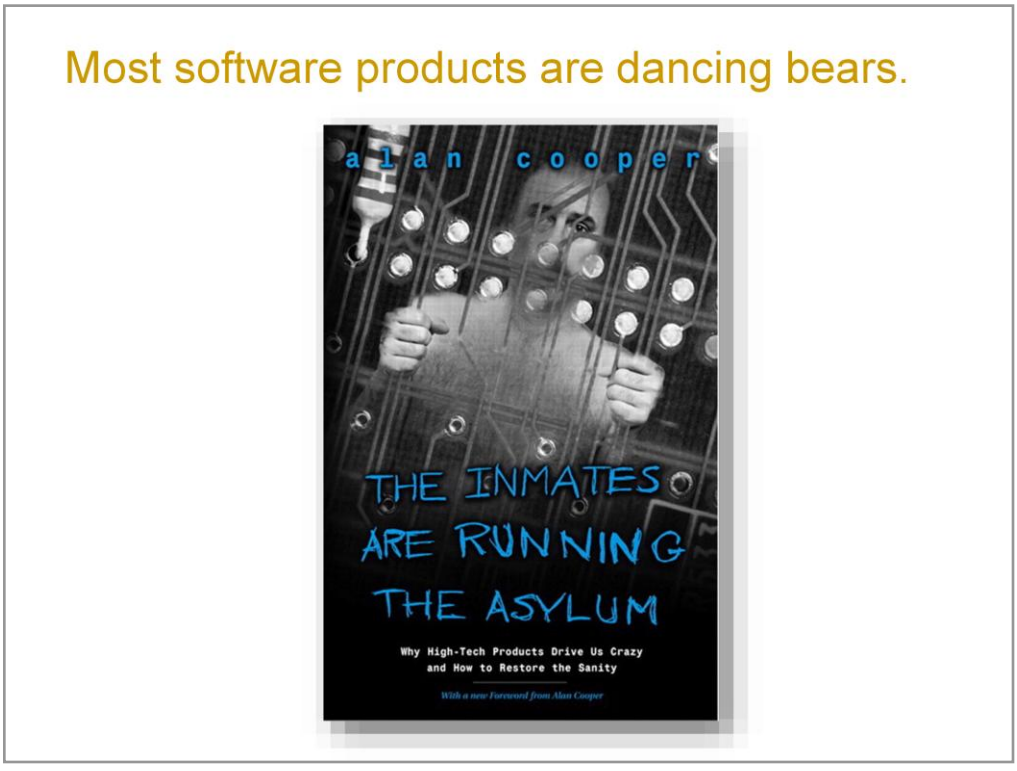
Also disappointing is the fact that this taxonomy of visualization methods makes no use of visual attributes to present them, other than colors to organize them into questionable categories. One attribute of visualization that it sought to identify was whether each visualization method presented detail or overview information. I only got as far as looking at the second entry—Tb for Table—which was designated as a means to present overview information only, before giving up on this particular distinction as useless. One of the beneficial characteristics of tables is that they can be used to display overview (summary) information, detail information, or a combination of both together. I also found the distinction between “divergent thinking” (”adding complexity”) and “convergent thinking” (”reducing complexity”) confusing. I believe that good visualizations neither add nor reduce complexity, but accurately and as simply as possible represent the level of complexity that exists in the data.

The final infovis topic that has made its mark in the blogosphere this year is social visualization, also known as collaborative visualization. The first publicly-available website for collaborative visualization that I became aware of was Swivel, which has probably garnered the most attention in the blogosphere. I responded with a mixed review. I loved the concept of making it easy for people to share data with one another visually and to make sense of it collaboratively, but the graphical capabilities of Swivel were not designed to effectively enable this. The other aspect of the site that bothered me was the way it encouraged people to playfully feature relationships in data that were absurd, such as this graph that asserts a negative relationship between wine consumption and violent crime. I believe that in their good intentions to make data and its analysis a more democratic process, they forgot that it is important to make use of expertise and best practices to guide people in doing it meaningfully. I like the fellows behind the site. They are smart, thoughtful, and well-intentioned. They welcomed my critique and expressed their intentions to improve the visualization capabilities of the site. I was encouraged, until they announced their next great visualization feature: the ability to "bling your graph."



Not long after learning about Swivel, which caused me to become interested in a better way to do web-based collaborative visualization, a new site was announced that answered my prayers called Many Eyes. Here was a site that not only understood information visualization, it seemed to understand collaboration also. This site was able to succeed where Swivel seemed to fall short, perhaps because of one significant difference: it was developed by who were already experts in infovis and diligent researchers of collaboration—Martin Wattenberg, Fernanda Viegas, and their colleagues at IBM Research. As a site that was designed to bring people together around data, Many Eyes has been discovered by a broad audience of folks, many of whom had little prior knowledge of visualization. The exposure they have received and the perceptions of infovis that it has forged in this case are positive and true.

Let's move now to the ways that the marketing of commercial software products has helped to shape the world's notion of infovis in the last year or so.



Before we look at some of the product marketing that has helped to shape the world's perception of infovis this year, let me admit that I share designer Alan Cooper's opinion that most software products are dancing bears—that is, they are unnecessarily difficult to use and far less effective than they could be. Cooper argues in his provocative and excellent book *The Inmates are Running the Asylum* that most software works poorly, which results from the way that most software companies approach the development process. This approach is intimately tied to the fact that these companies are managed from an engineering perspective rather than a design perspective. Cooper writes:

> *High-tech companies—in an effort to improve their products—are merely adding complicating and unwanted features to them. Because the broken process cannot solve the problem of bad products, but can only add new functions, that is what vendors do…The high-tech industry has inadvertently put programmers and engineers in charge, so their hard-to-use engineering culture dominates…Programmers aren't evil. They work hard to make their software easy to use. Unfortunately, their frame of reference is themselves, so they only make it easy to use for other software engineers, not for normal human beings…While we let our products frustrate, cost, confuse, irritate, and kill us, we are not taking advantage of the real promise of software-based products: to be the most human and powerful and pleasurable creations ever imagined…All it requires is the judicious partnering of interaction design with programming.*
>
> (*The Inmates Are Running the Asylum*, Alan Cooper, 1999, SAMS Publishing: Indianapolis, Indiana, page 8, 15, and 17)
>
> T*he sad thing about dancing bearware is that most people are quite satisfied with the lumbering beast. Only when they see some real dancing do they begin to suspect that there is a world beyond ursine shuffling. So few software-based products have exhibited any real dancing ability that most people are honestly unaware that things could be better – a lot better.*
>
> (*Ibid*, page 59)

I believe that the biggest and most familiar software company in the world—Microsoft—exemplifies what Cooper laments.

Despite the great work in infovis by Microsoft Research, based on Microsoft's products, we don't usually think of Microsoft as an infovis company. In a primitive sense, however, the most prolific infovis product in the world is Microsoft Excel. Despite the primitive nature of Excel's visual analysis and charting functionality, it is used more than any other product to make sense of data and, in combination with PowerPoint, to present data to others. Almost everyone who takes my table and graph design course wants to know, more than anything else, how to apply the data presentation principles that I teach to Excel.

For this reason, when I heard from my friend Marti Hearst of U.C. Berkeley about two years ago that Microsoft was in the process of completely redeveloping the charting engine that resides in Office products, including Excel, I got really excited about Microsoft's opportunity to improve the lives of millions by improving the effectiveness of its charts. Marti tracked down the name of someone involved in charts at Microsoft, and I immediately wrote him to offer some free advice about what could be done to make the product better. My request was cordially received—the fellow was familiar with my work—so I submitted a set of requirements. I was thanked for my efforts and assured that I would likely be quite pleased with Excel's new charting capabilities. I then waited, with great expectation for my first glimpse of the product.

A few months later I saw in the official Microsoft Office 2007 blog my first examples of the new charting functionality, and my hopes were crushed. What I saw was a product that ignored almost every suggestion that I made, but instead could produce charts like this pie chart that are so pretty you want to bite right into them. Rather than making any of the improvements that I knew were desperately needed by millions of Excel users, Microsoft sent the message to the world that what's important about graphical communication is that it look sharp, with the ability to manipulate the visual appearance of charts in every way imaginable, not that it is able to communicate data effectively.

It isn't that the folks responsible for the charting functionality of the office products didn't work hard with the best of intentions—I'm sure they did—but they were in my opinion given the wrong set of marching orders and no time to do anything else.

Databars: one of Excel's near misses

Cause of Death in Year 2000	White	Black	American Indian or Alaska Native	Asian or Pacific Islander	Hispanic or Latino	White, not Hispanic or Latino	Total
Diseases of heart	253.4	324.8	178.2	146.0	196.0	255.5	**1,353.9**
Malignant neoplasms	197.2	248.5	127.8	121.9	134.9	200.6	**1,030.9**
Cerebrovascular diseases	58.8	81.9	45.0	52.9	46.4	59.0	**344.0**
Unintentional injuries	35.1	37.7	51.3	17.9	30.1	35.3	**207.4**
Chronic lower respiratory diseases	46.0	31.6	32.8	18.6	21.1	47.2	**197.3**
Diabetes mellitus	22.8	49.5	41.5	16.4	36.9	21.8	**188.9**
Influenza and pneumonia	23.5	25.6	22.3	19.7	20.6	23.5	**135.2**
Chronic liver disease and cirrhosis	9.6	9.4	24.3	3.5	16.5	9.0	**72.3**
Suicide	11.3	5.5	9.8	5.5	5.9	12.0	**50.0**
Homicide	3.6	20.5	6.8	3.0	7.5	2.8	**44.2**
Human immunodeficiency virus (HIV) disease	2.8	23.3	2.2	0.6	6.7	2.2	**37.8**
Fake data #1	0.0	0.0	0.0	0.0	0.0	0.0	**0.0**
Fake data #2	-10.0	-12.0	-11.0	-2.0	-5.0	-8.0	**-48.0**

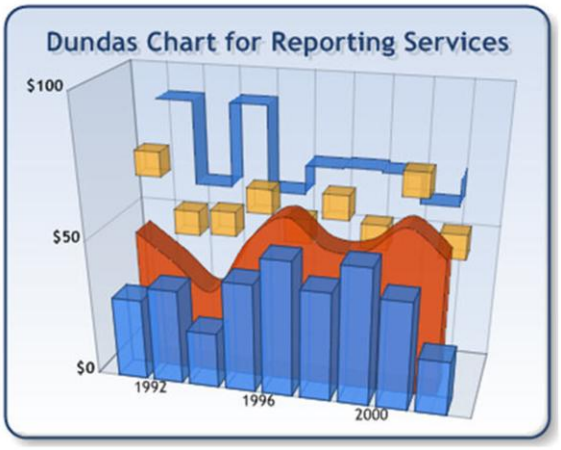
Good ideas cease to be good when implemented poorly.

One potentially useful feature that made the list was the addition of databars, which are bars like those in a bar chart, except these appear in the cells of a spreadsheet to encode the values entered into those cells. As you can see, databars can serve as a primitive approximation of a table lens display, which is really useful, but two aspects of the implementation undermine their effectiveness.

1. Notice that the right end of each databar fades away. The most important part of the bar, the part the marks the value, is the hardest to see. And this is not an ill-chosen formatting option that can be turned by the user; it is a feature that can't be changed.
2. An even worse problem, however, is that the bars' lengths do not correspond to the values they represent. For instance, they are never shorter than around 10% of the columns width, even when the value is zero.

What's sad is the fact that the folks at Microsoft who are responsible for this work that they are aware of the product's failures and want to make the product excellent, which I know from direct conversations with them, but they work under constraints that sometimes make it difficult to do what they'd like. Regarding databars, these folks are currently responding the problems that I and others have pointed out and solicited feedback in an effort to get it right. Not only that, but the person who leads the team that is responsible for the charting functionality of Microsoft's office products is here at the conference, doing his best to plug into the insights that this community has to offer. He deserves a great deal of credit for that, because very few vendors in the business intelligence software space participate in this conference, and many aren't even aware that the infovis research community exists.

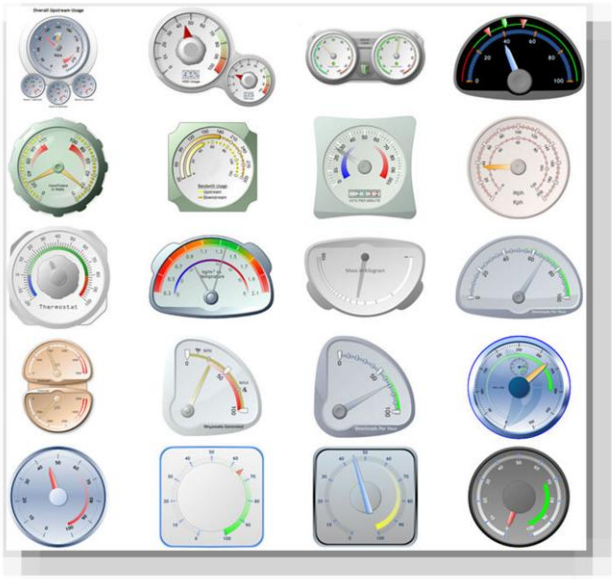
Dundas to the rescue!?



Visualization software should be designed by visualization experts.

Speaking of the business intelligence market, this year Microsoft began making an explicit effort to establish its presence in this space, demonstrated by its first annual business intelligence conference. In an effort to boost its position in this space, Microsoft recognized its need for a good web-based charting product. As is often the case, Microsoft decided to solve the problem by acquiring existing products: in this case those from a company named Dundas. Unfortunately, Dundas' products are not very good. In the last few years I've seen several software companies try to break into the visualization space by acquiring products that look cool on the surface, if you know nothing about infovis, but simply don't work very well. If you throw a bunch of engineers in a room, lock the doors and feed them only Coke and Snickers bars for a year, without any interaction with real customers or knowledge of infovis, this is the result.

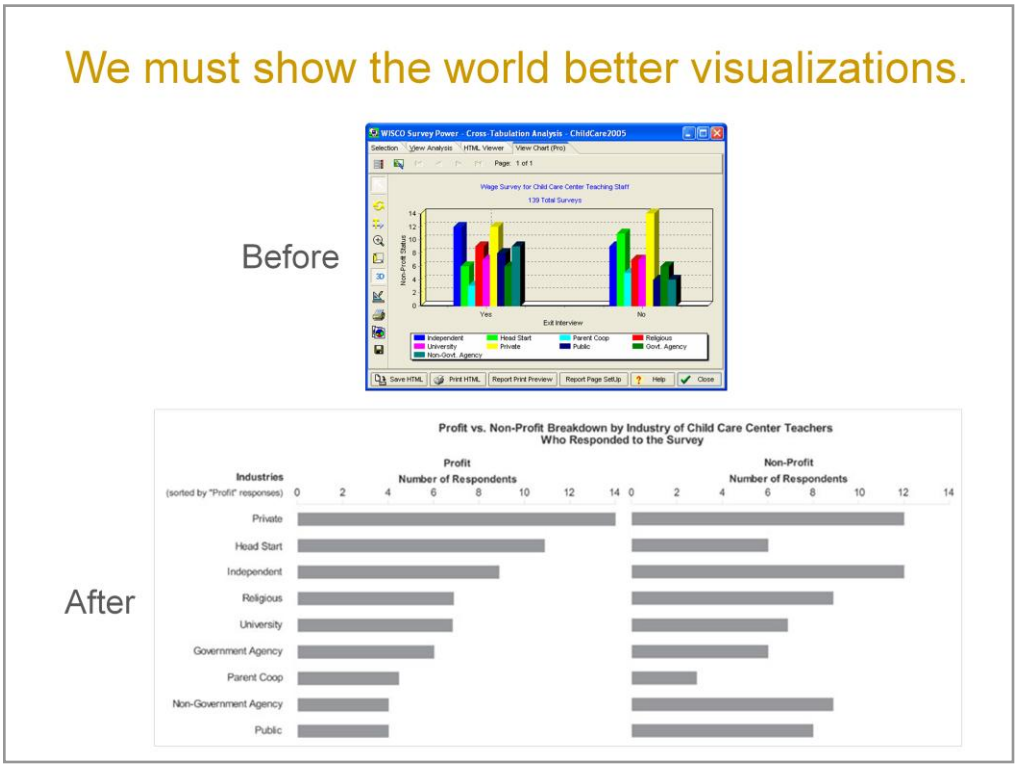
Dundas to the rescue!?



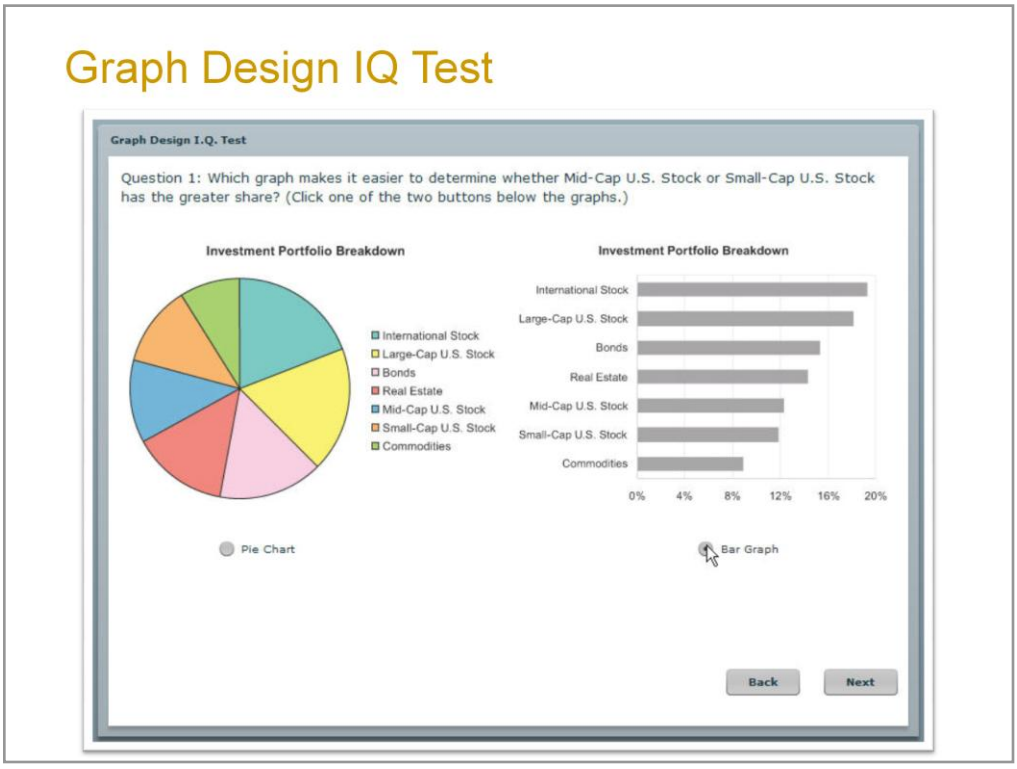
An infinite supply of silly gauges is not what we need.

Widgets galore, all beautifully rendered, but equally absurd as a means to communicate real information. I once had one of these guys, not from Dundas in this case, call me over to his booth on a conference exhibit floor and proudly show me how, when the values change on one of his gauges, the needle bounces back and forth like one of those old mechanical gauges, only gradually settling down to point to the new value.

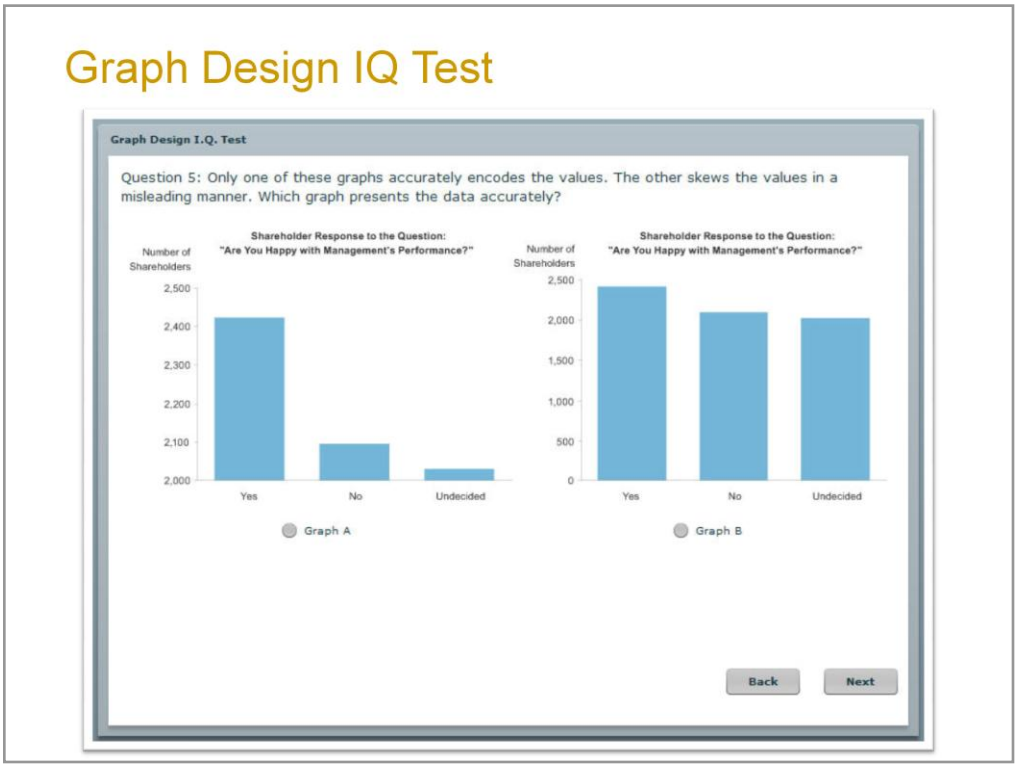
I was once contacted by David Cunningham, the President and CEO of Dundas, who asked if I would be willing to work with Dundas to create a website that would serve as an information visualization educational resource. In an attempt to be avoid being rude, I told him that such a site, however useful, probably should not be associated with a particular software vendor, so as to remain completely product agnostic. In fact, however, I was horrified at the thought of teaming up with a company that understood infovis so poorly for such a purpose.



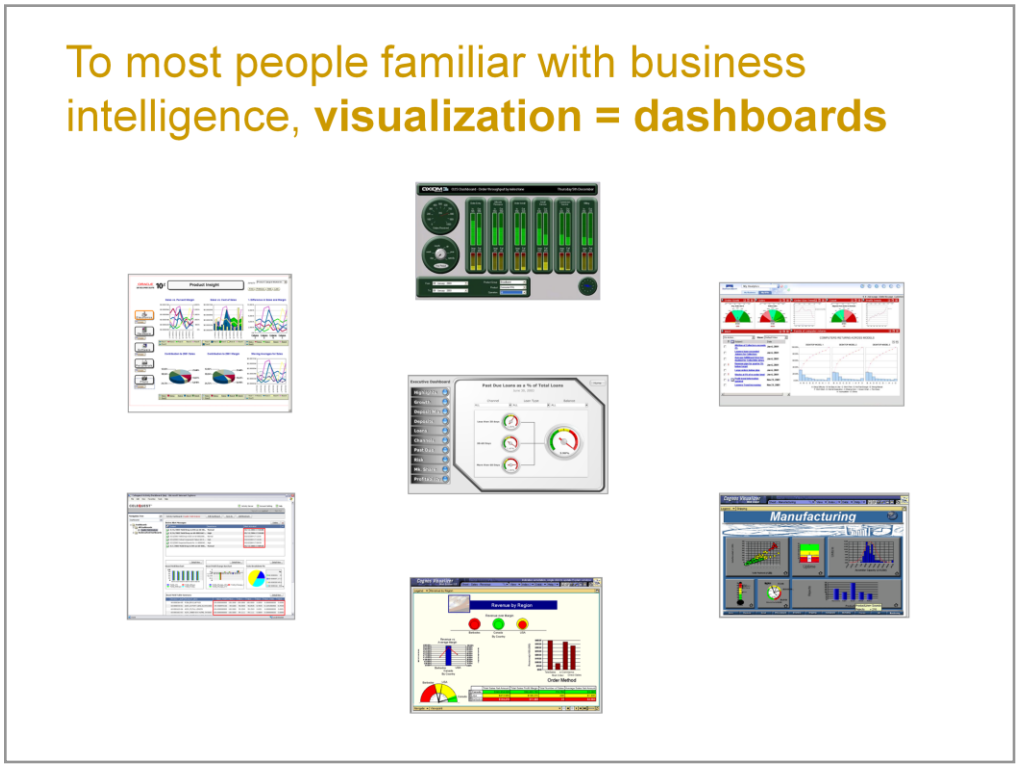
The world beyond our walls is exposed to horrible examples of visualization every day at an increasing rate. We need to show the world what visualization really looks like when it is done right. One of the ways that I try to do this is through examples, by taking a graph that is typical of what vendors promote on their websites, such as this one that you can find in the Examples section of my website (www.perceptualedge.com), critiquing its effectiveness to explain how it fails, and then by offering a better design, along with an explanation for why it succeeds where the other fails. I've found consistently in my work that, when people are shown effective alternatives to the bad visualizations that are common and familiar, they easily recognize the difference. We need to combat the bad visualizations that dominate the market by exposing people to visualizations that really work.



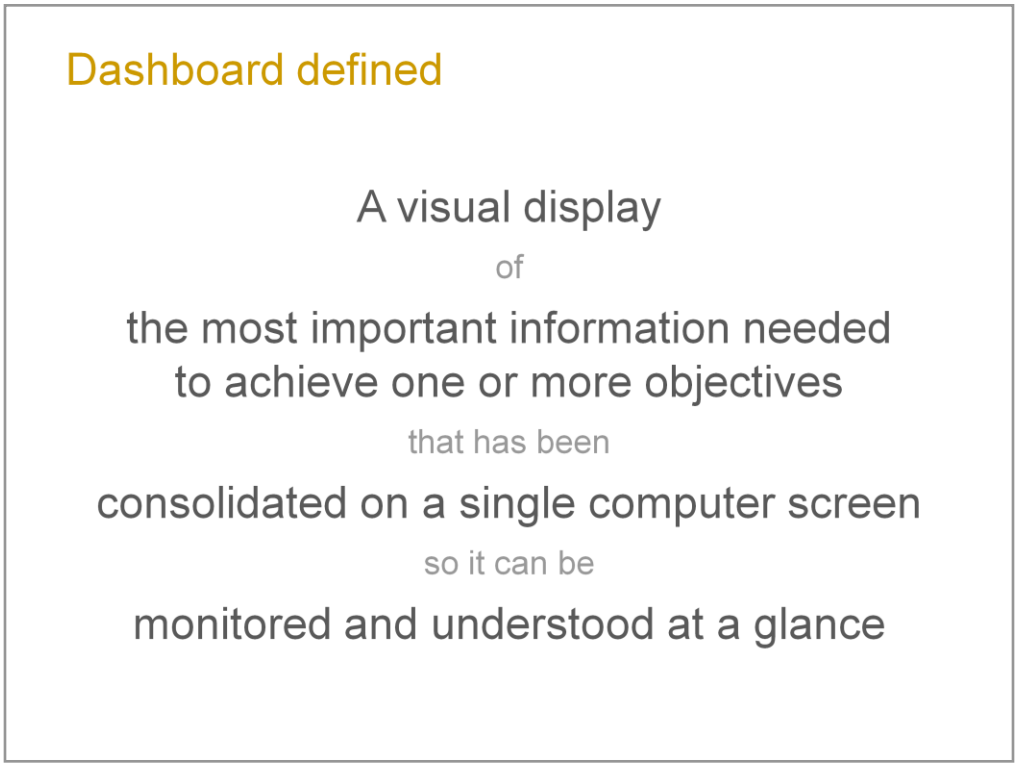
A few months ago it occurred to me that I could use this approach of bad and good examples in an entertaining way to spread the word by creating what I called the “Graph Design IQ Test.” It’s simple Flash program that walks people through a series of questions that ask them to choose the better of two visualizations. Thousands of people have taken this test and passed it on to their colleagues.



Everyone, even without any training in graph design, scores highly when they take this test, because the difference between what works well and what doesn't is generally easy to see when a bad visualization and an effective alternative are shown side by side. That's the whole point: I want them to see how easy it is to design a graph effectively when they are shown alternatives, so they'll be encouraged to seek out resources for learning these skills. When they do happen to get a question wrong, I use the opportunity to educate them a little, explaining why their choice doesn't work, of course, after chiding them a bit with humor.



To many people in the business world today, especially those involved in business intelligence, visualization means one thing: dashboards. I doubt that there is a large company in the U.S. that doesn't have at least one dashboard. They have taken information delivery by storm.

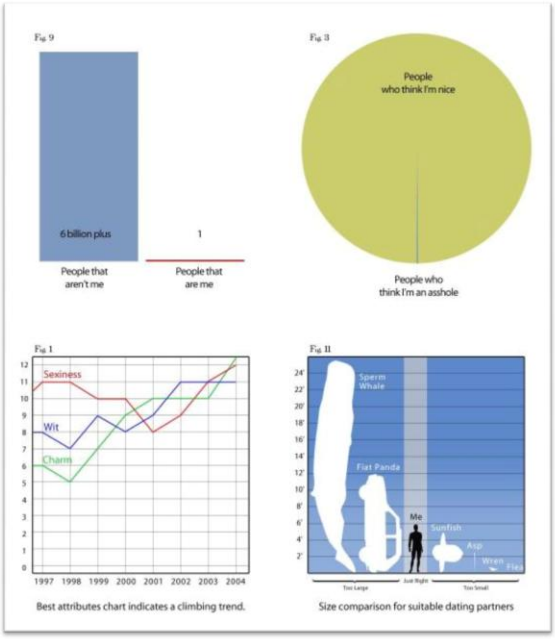


Because dashboards are probably not familiar to everyone here, let me take a moment to define the term.

> *A dashboard is a visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at a glance.*

This definition first appeared in the March 20, 2004 issue of *Intelligent Enterprise* magazine in an article written by Stephen Few entitled “Dashboard Confusion.”

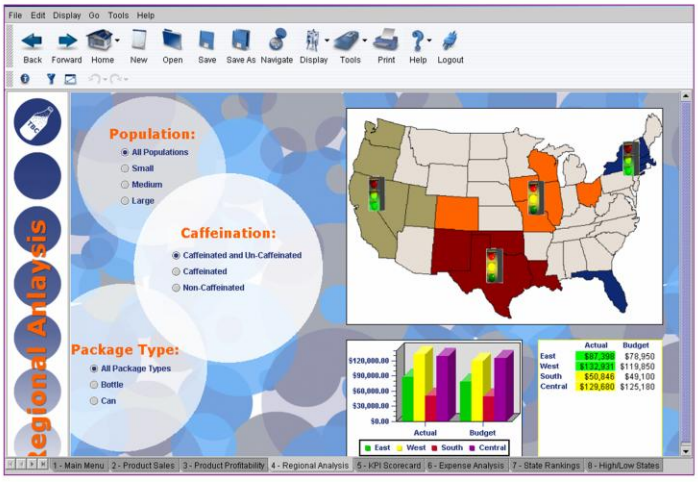
There's a dashboard for every occasion.



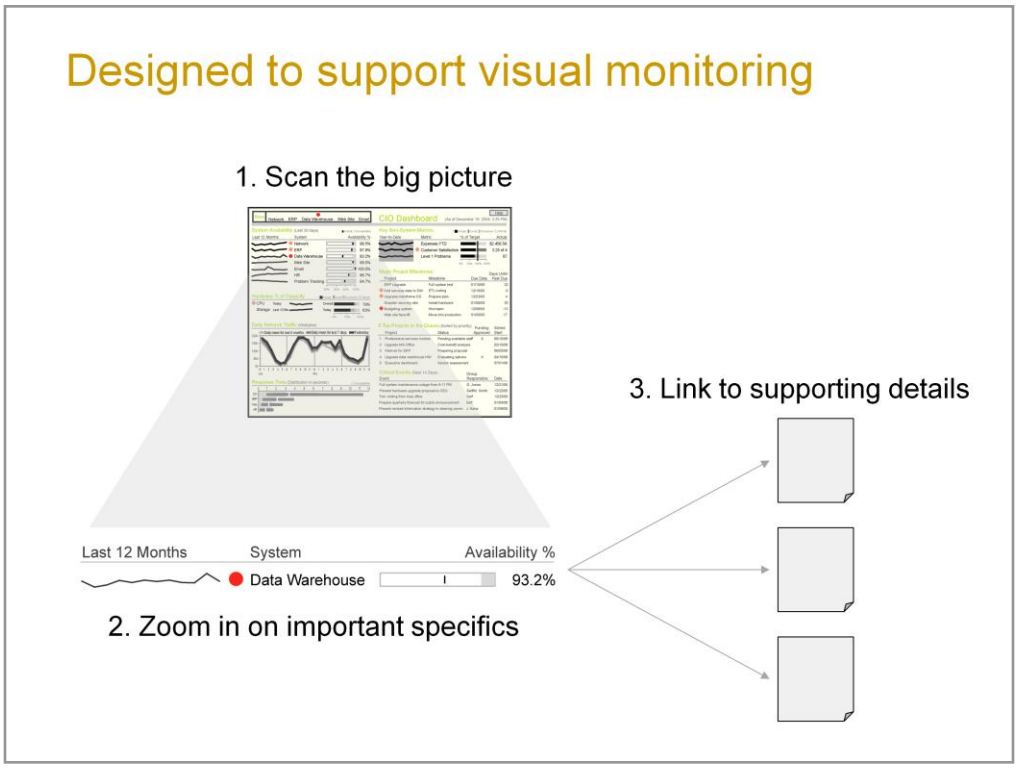
Although most dashboards are used by businesses to track performance, this particular variation appeared in the personals section of Craig's List, which was used by a fellow to market his dating potential.

Most dashboards fail.

- They say too little
- What they do say, they say poorly

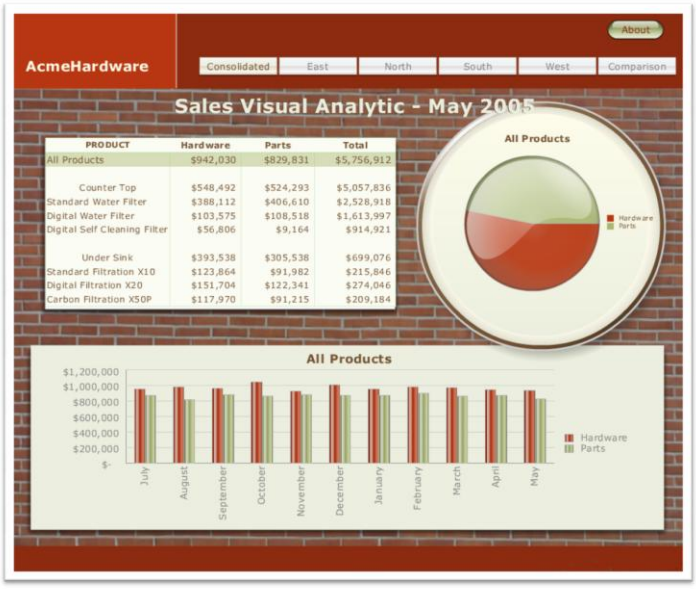


The fact that this is what many people in the business world think of as data visualization should concern us, because most dashboards fail: they say too little and what they say they say poorly.



For a screen full of information to effectively support the task of monitoring, it must be designed in a particular way. I believe that it must by designed in a way that allows the person viewing it to quickly get the big picture of what's going on, to then identify particular items that most need her attention, to get enough information about those items to determine if action is required, and then to easily access any additional information that might be needed to determine how to respond.

Dashboards often obscure what could be shown clearly.

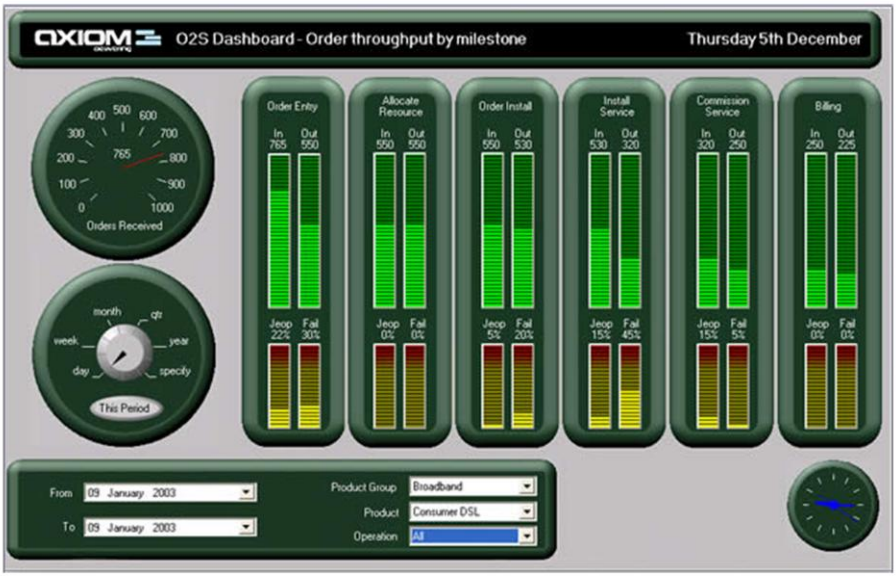


Most people associate dashboards with video-game-like displays filled with gauges and meters. Visualizations that jump off the screen with bright colors and dancing widgets might be entertaining at first, but quickly become just plain annoying when they fail to deliver real value.

Notice the pie chart on this dashboard. How many slices does it contain? Without concentrating, it looks like it has three slices, because visual effects of light, shadow, and reflection have obscured the chart. In the real world, if light is reflecting off of a shiny object, forcing us to squint in an effort to see it clearly, we find it annoying. Why would we ever want to reproduce this annoying effect on a dashboard? Yet countless hours are spent by software developers to fill dashboards with features like this, and almost no time is taken to step back and ask, what is needed to display information on a dashboard in the clearest, most meaningful way possible?

(Source: Business Objects)

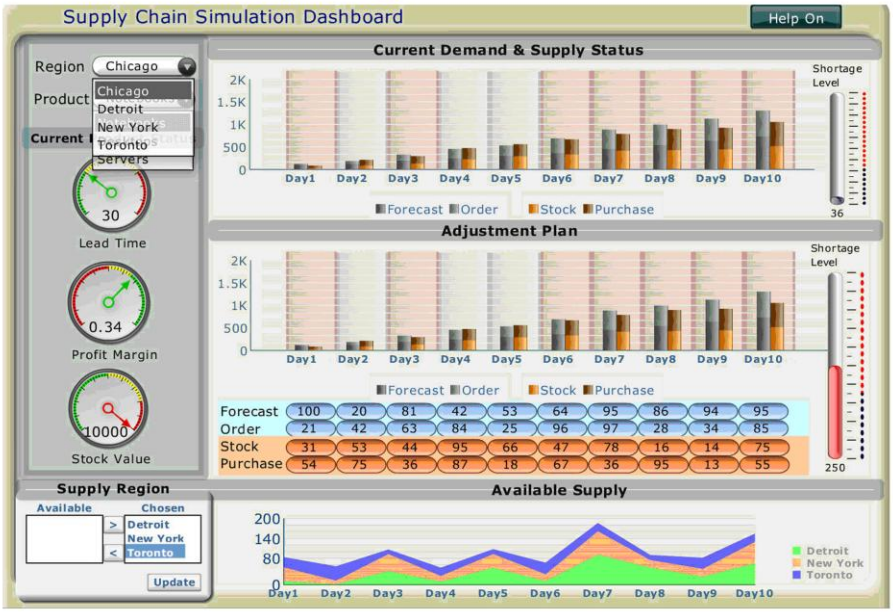
Tremendous effort is dedicated to making dashboards look like something they're not.



The makers of the dashboard above did an exceptional job of making it look like an electronic control panel. If the purpose were to train people in the use of equipment that actually looks like this by simulating it, then this would be great, but that isn't the purpose of a dashboard. The graphics dedicated to this end are pure decoration, visual content that the viewer must process to get to the data.

(Source: Website of Axiom Systems.)

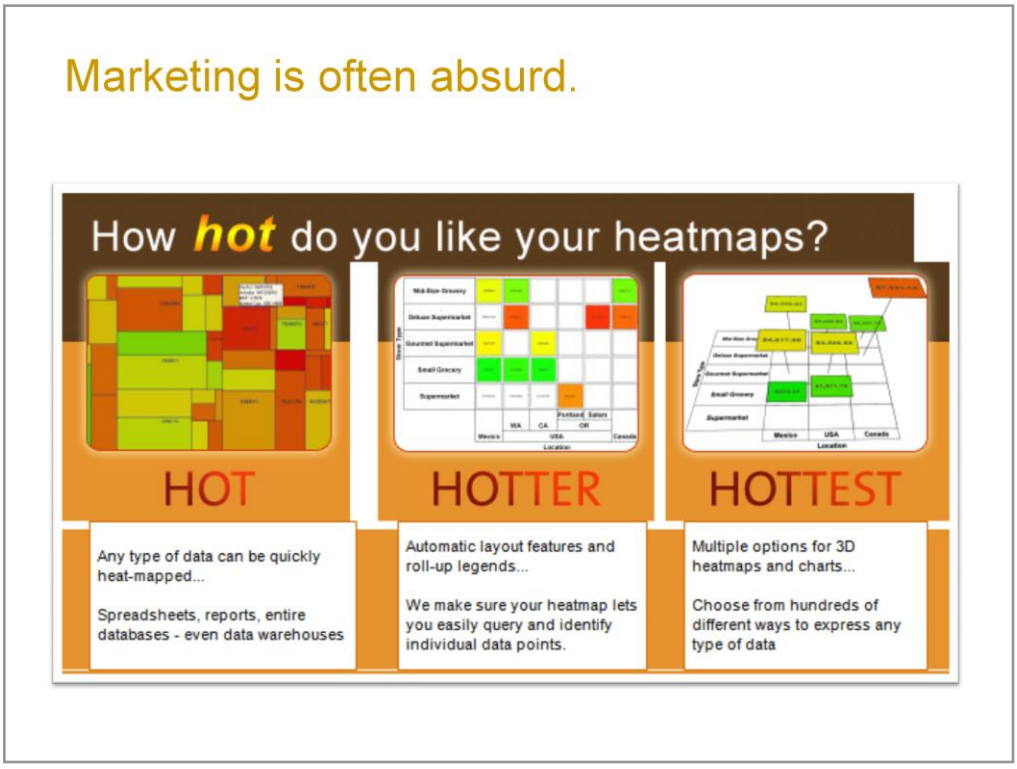
Awards are being given to dashboards that communicate poorly.



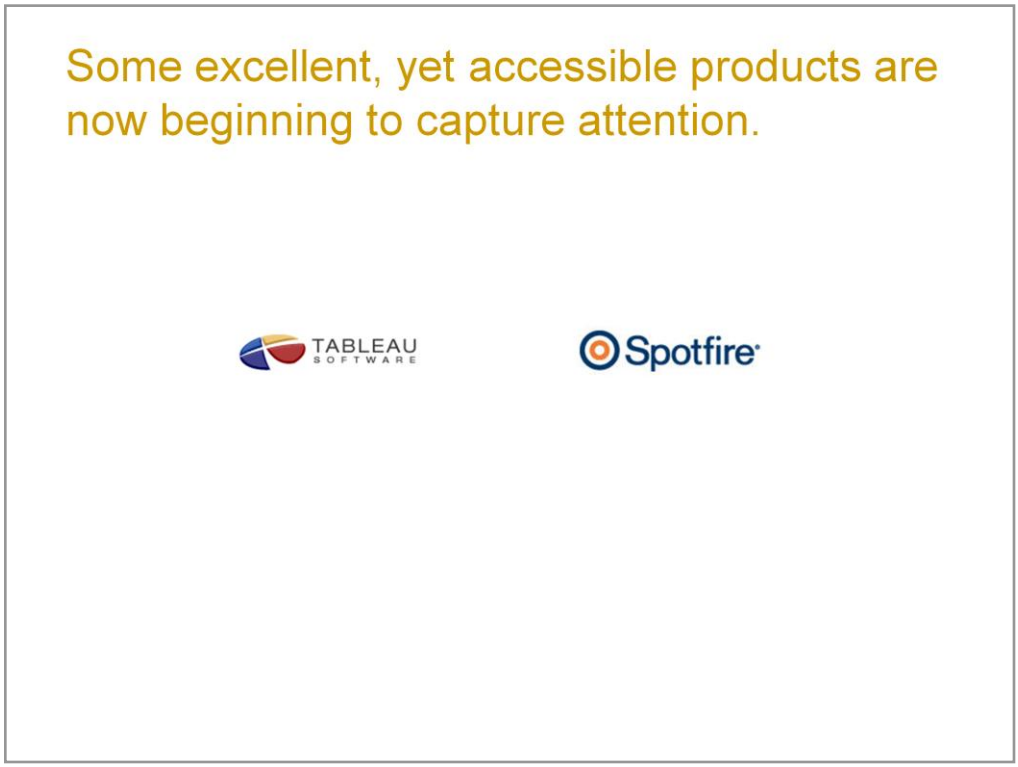
The software company Business Objects awarded $10,000 to the person who created this dashboard as the top prize in its dashboard design competition. Anyone who takes my one-day dashboard design course could do better.

(Note: For a critique of this dashboard's design, go to www.perceptualedge.com/blog/?p=63.)

Visual analysis software has recently made started to garner the attention of business people as a representative of visualization. All business intelligence vendors are now trying to compete by offering products that supposedly support visual analytics, but relatively few understand data analysis or visualization.



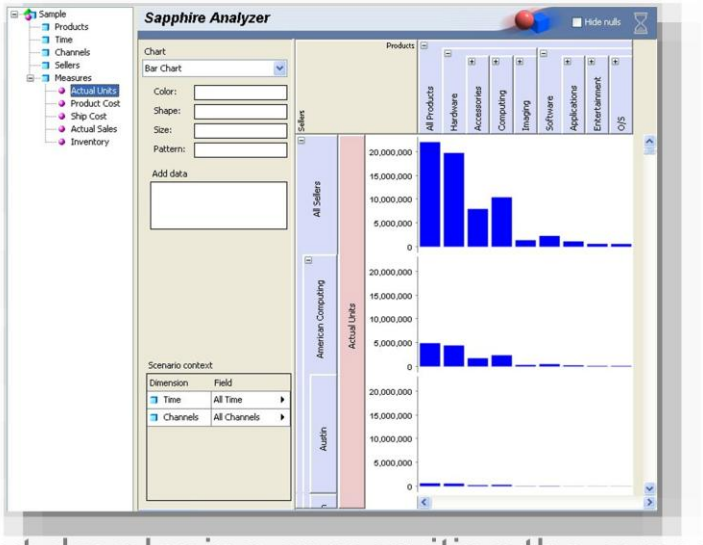
Some vendors that ought know better are adding to the confusion. This ad from AVS (Advanced Visual Systems) illustrates how distorted a view of infovis marketing often provides.



Despite how much bad exposure infovis has received through product marketing, fortunately there are a few good infovis products on the market that are expanding their marketing presence, which will help to correct the misperceptions that exist. Out of around 20 visual analysis products that I place in the good category, the two that have been most successful in marketing to businesses other than only to statisticians are Tableau and Spotfire. Both have excellent products that I'm always happy to encourage clients to consider.

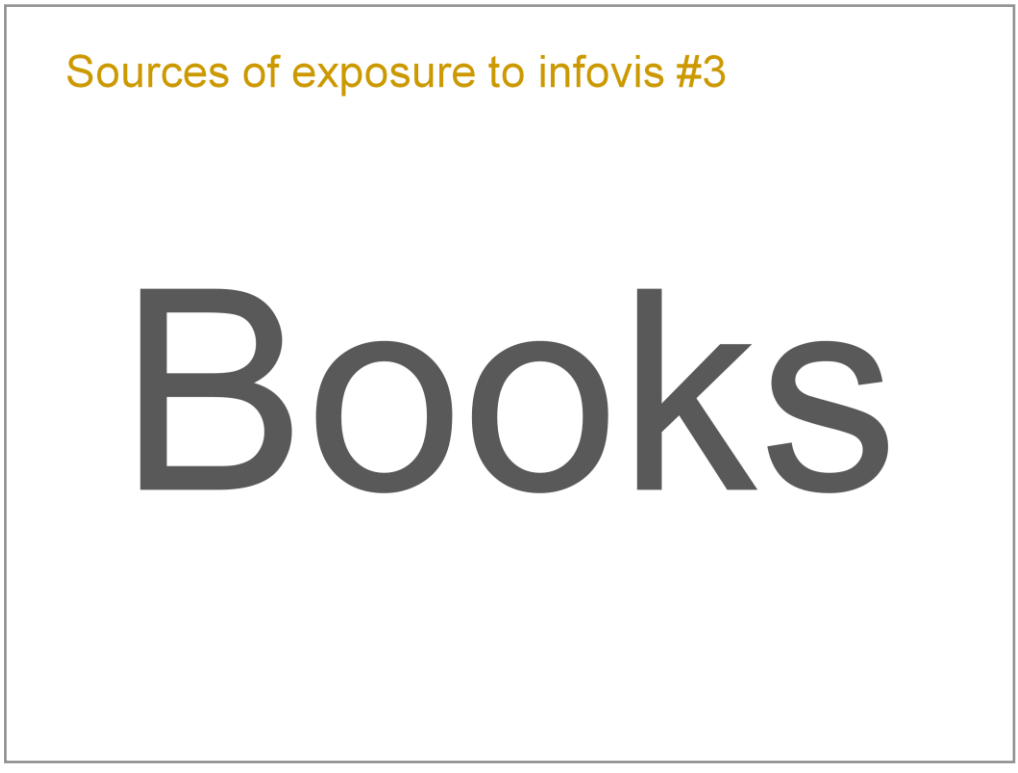
Because of good products like these that have managed to break into the business intelligence market, and persistent badgering by me and a few others, the business intelligence industry is increasingly recognizing visualization as central to what it does. Just last week I was talking to Ted Cuzzillo, who writes for the premier business intelligence research and education organization—TDWI, which stands for The Data Warehousing Institute. Ted is here covering the conference for TDWI. Toward the end of our conversation Ted asked, and I'm paraphrasing, "Would you agree that data visualization is where the business intelligence industry will need to turn to succeed in the future?" I was only too happy and exceedingly pleased to hear these words and heartily agree.

Traditional business intelligence vendors are trying to emulate these successes.



Without developing or recruiting the expertise

This is a prototype that was developed at Business Objects, which goes by two names: Sapphire and Visual Data Analyzer. It is clearly Business Objects' attempt to develop a product to compete with Tableau, but the fact that the work was done with no more than a superficial understanding of information visualization is obvious. In an attempt to outclass Tableau, this product added a feature that in my opinion represents a bad practice, because of the confusion that it would likely produce. You can see in the screen capture above that the bars in each of graphs represents values at three levels of a hierarchy: (1) all products, (2) hardware vs. software products, and (3) lower level product categories such as accessories, computing, and imaging. Representing various levels of a hierarchy in a single bar chart encourages people to compare the magnitudes of unlike values, which will lead to confusion unless the different levels are clearly differentiated in some way in the visualization. Combining values from different levels in a hierarchy also results in a scaling problem in that values from higher levels in a hierarchy will be higher—potentially much higher—than values in lower levels.

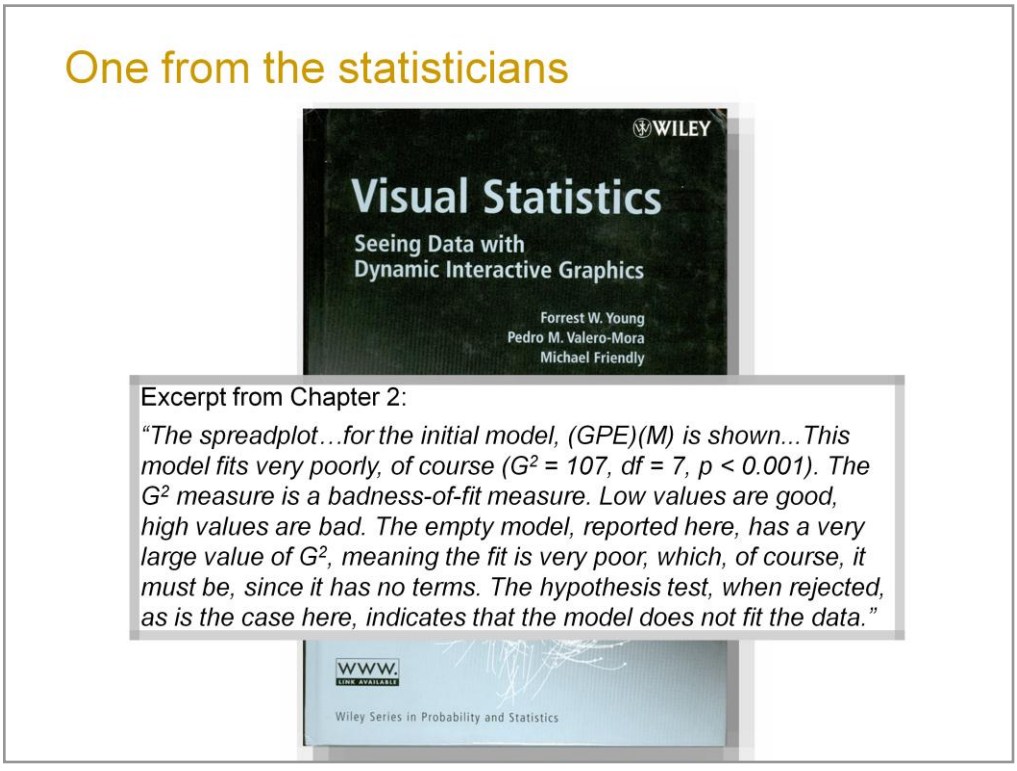


This third source of exposure to infovis that helps to shape the world's perceptions is books. People who wish to learn about infovis sometimes search for books to help them along. In the last year or so, a couple of new books that I'm aware of have been published that might have shown up on the radar.

The first is Edward Tufte's long-awaited fourth book. I owe a great deal to the work of Edward Tufte. Attending his one-day seminar many years ago inspired me to shift my focus in the field of business intelligence to the visual display of information. In July, 2006, his long awaited fourth book, *Beautiful Evidence*, finally became available, which I ordered immediately and eagerly devoured. It contains much wisdom and flashes of brilliance, as expected, but it also seems to stray from the principles Tufte advocates so forcefully in this book and his earlier work.

In the introduction, Tufte states: *"Beautiful Evidence* is about *how seeing turns into showing*, how empirical observations turn into explanations and evidence. The book identifies excellent and effective methods for showing evidence, suggests new designs, and provides analytical tools for assessing the credibility of evidence presentations." (Edward Rolf Tufte, Beautiful Evidence, 2006, Graphics Press LLC, Cheshire, Connecticut, page 9) While holding the book in my hands for the first time, even before opening it, I experienced a foreshadowing of what I would find within. The dust cover is graced with four photographs taken by Tufte of one of his beloved dogs in motion leaping above and eventually splashing into a body of water. The images might be beautiful, befitting the first word of the book's title, but I failed to see how they related to displays of evidence.

I found the contents of *Beautiful Evidence* beautiful throughout, but was frequently distracted by a sense that Tufte focused more on the beauty of its contents than on relevance to the topic or a clear presentation of his case. Reading it felt like spending an evening with Tufte, sipping brandy in his library and conversing while he wandered among the bookshelves pulling and showing examples here and there as they caught his attention. As such, I enjoyed it, but as a treatise on finely crafted presentations of evidence, it lacked the logical unfolding of argument and validation through carefully chosen examples that I expected. Rather than a book that was designed to examine a specific topic to meet the real needs of a particular audience, it seemed more like a series of essays ranging across a variety of topics that were selected mostly because they have occupied his attention in recent years. One striking example was the inclusion of the last two chapters of the nine, "Sculptural Pedestals: Meaning, Practice, Depedestalization" and "Landscape Sculptures", which featured his work as a sculptor—beautiful, but not on topic without stretching the imagination.



The other book that I know of is *Visual Statistics*, which came out in 2006, but I didn't discover it until this year. I was encouraged when I began to read the introduction that this might be a book I could recommend to this audience. The authors' message rang true to my experience and seemed to share my goals:

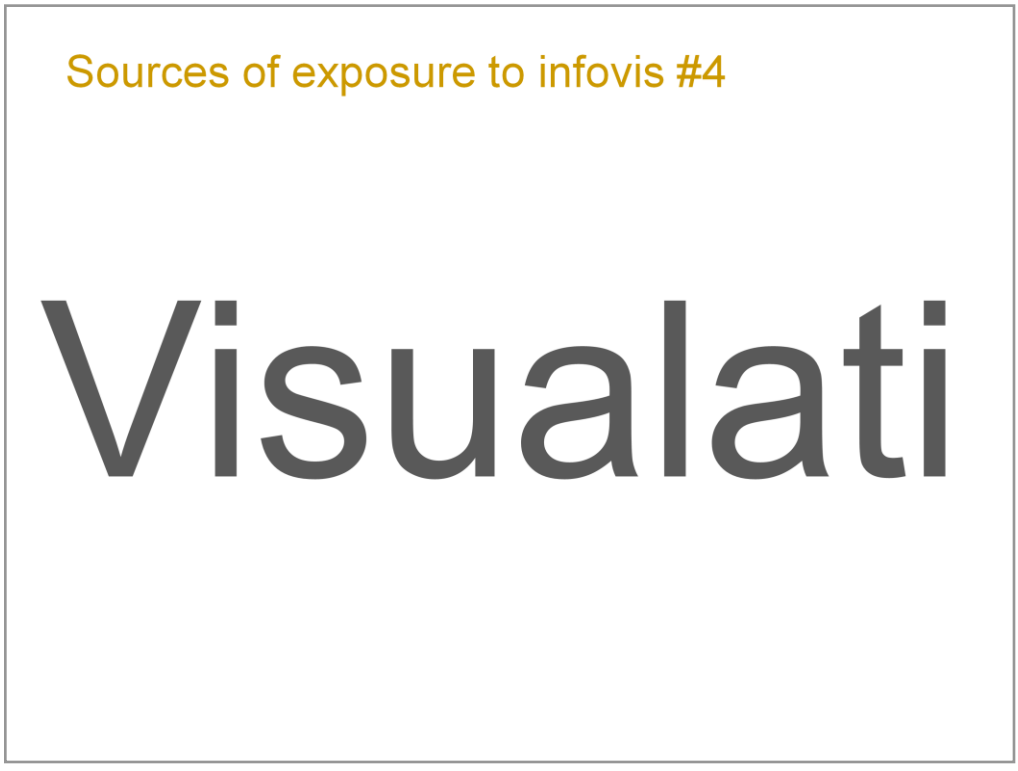
> *Statistical data analysis provides the most powerful tools for understanding data, but the systems currently available for statistical analysis are based on a 40-year-old computing model, and have become much too complex. What we need is a simpler way of using these powerful analysis tools.*
>
> *Visual statistics is a simpler way. Its dynamic interactive graphics are in fact an interface to these time-proven statistical analysis tools, an interface that presents the results of the hidden tools in a way that helps ensure that our intuitive visual understanding is commensurate with the mathematical statistics under the surface. Thus, visual statistics eases and strengthens the way we understand data and, therefore, eases and strengthens our scientific understanding of the world around us.*

As I ventured further, however, my hopes began to evaporate as the language of statistics began to dominate the landscape, making the journey hard going. Near the end if the introduction, the authors described the readers they were hoping to reach:

> *It is our aim to communicate the intrigue of statistical detective work and the satisfaction and excitement of statistical discovery, by emphasizing visual intuition without resorting to mathematical callesthenics* [sic]*…Seldom is there mention of populations, samples, hypothesis tests, and probability levels…This book is written for readers without strong mathematical or statistical background, those who are afraid of mathematics or who judge their mathematical skills to be inadequate; those who have had negative experiences with statistics or mathematics, and those who have not recently exercised their match or stats skills.*

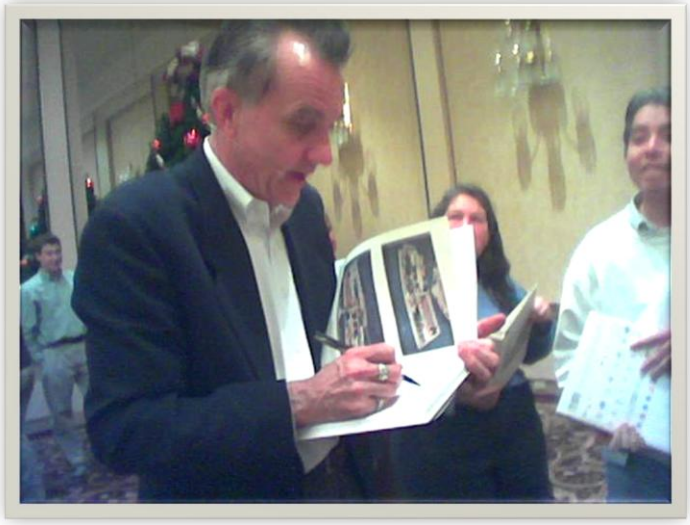
What I discovered in reading the book, however, is that, despite how useful it might be as a primer in visual analysis for statisticians, it is steeped in the concepts and language of statistics, and lacks the explanations that would be needed by non-statisticians to make use of the material. I have no doubt that the authors attempted to reach out to non-statisticians. I suspect, however, that they are too immersed in an academic statistical mindset to recognize when they are using terms and discussing concepts that are unfamiliar to the uninitiated. Terms such as Box-Cox transformation, Euclidean space, kernel density curve, and Pearson's chi square are par for the course. Early in chapter 2, which provides some actual data sets and analytical challenges that are used throughout the book, the reader is already faced with material like the quote above from Chapter 2.At this point, as someone whose statistical knowledge can fit comfortably in a thimble, my eyes began to glaze over.

Please don't misunderstand me. I am not saying that this is not a good book. I suspect that this is a very important book for statisticians, because it introduces them to the power of visual analysis, which most statisticians under-appreciate. This just isn't a book for non-statisticians.



The final source of exposure to infovis that is seen by the world out there, I call the Visualati, which I'm using to describe our fields version of celebrities—not those who are seen almost exclusively within our community, but those who are seen more broadly.

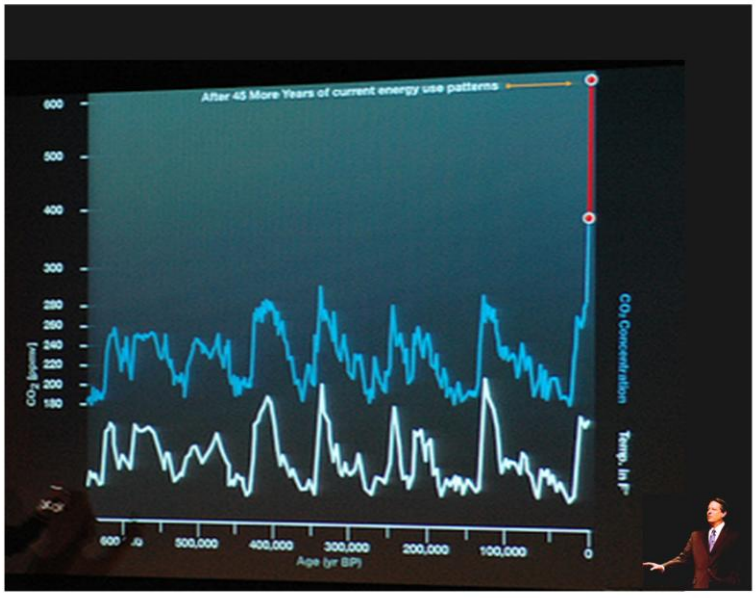
Tufte's popularity continues



But his relevance is dated

The first member of the Visualati has been a visualization celebrity for many years. Edward Tufte's seminars continue to draw large crowds and his books continue to sell well. Tufte has traditionally held screen-based displays in low regard. He correctly points out that screens cannot produce the high resolution of printed displays on paper. In his seminars, until recently he would hold up a copy of one of his books and point to illustrations on its pages, directing his audience turn to that page in their own copy to follow along, rather than use a projected image. Not all displays, however, require the high resolution of the printed page. And something you can't do with the printed page is interact with the data, which is critical to data exploration and analysis. Tufte's has locked himself out of much of the fine work that has been done in our field because of his uncompromising prejudices, which has cause his relevance to our work to decline. I've heard that in his seminars, however, he is now saying that he is turning his attention to computer-based, three dimensional displays, so it will be interesting to see if his contribution to our work will make a comeback over the next few years

Another inconvenient truth

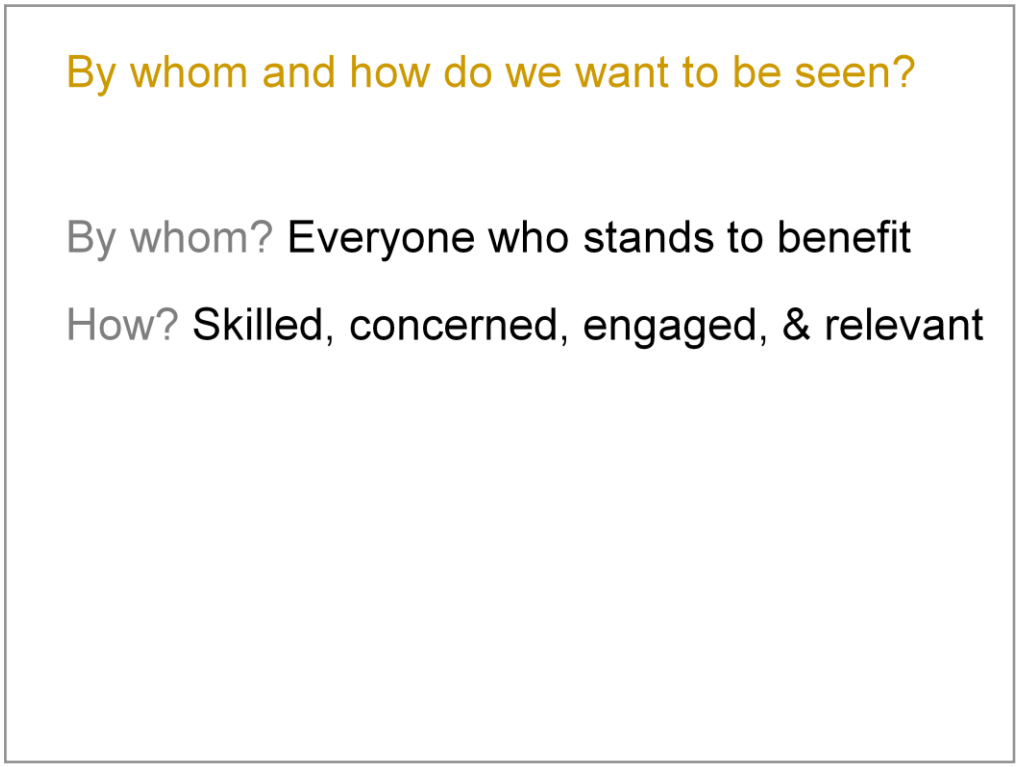


Al Gore invented infovis.

The next member of the Visualiti has only recently made his mark in our field, and you've probably never before thought of him as playing a role in infovis, but when Al Gore stepped in front if that huge chart in An Inconvenient Truth and rode the crane up to the top so he could point to that high projected value of carbon emissions of the near future, he became a data visualization celebrity. Despite his short tenure among us, be it known that Al Gore invented infovis.

I think the real star of information visualization in the last couple of years is Hans Rosling. When Rosling took the stage at the TED conference for the first time in 2006, he managed to get people up on the edges of their seats to watch—believe it or not—a series of graphs, resulting in a standing ovation. For most of the people there, data presentation had never been so compelling.

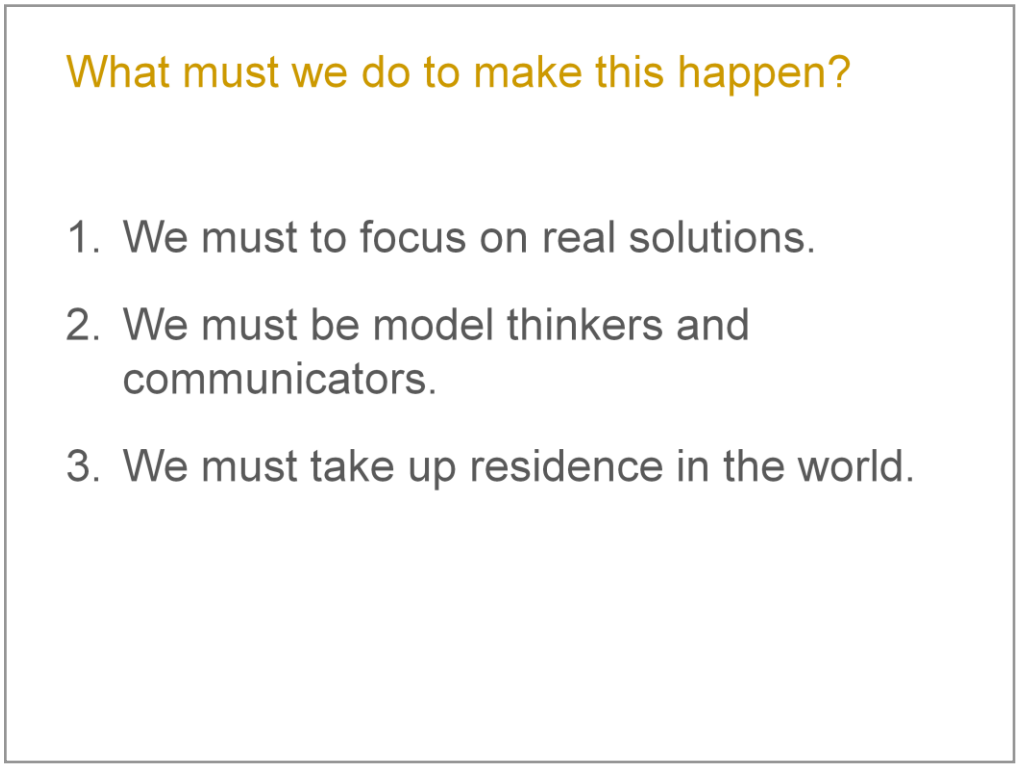
Rosling has used relatively simple visualization techniques, featuring animated plots, to tell statistical stories that are compelling, not only because they are told with great charisma, but because they reveal important truths about the world, such as the changing relationship between wealth and child mortality. I applaud his success, in part because it is success that we can share in, for it illustrates to the world at large what infovis can do when it is done well and it is used for worthwhile purposes.



As we look forward to the year 2008 and beyond, we should all ask the question by whom and how do we want our work to be seen. I'll tell you my answer, which in my arrogance I believe is a good answer that we should all share.

By whom do we want to be seen? By everyone who can use information visualization to better discover and understand what they must know to do what they do, especially those who work to make the world a better place, those who use information to build up what's good in the world and to tear down what stands it way.

How do we want to be seen? I want us to be seen as people who take pride in what we do, who possess both the expertise and the commitment to do excellent research and develop exceptional products and techniques, who are familiar with the problems that our work can solve and are focused on applying the work to those problems. In other words, I want us to be seen as skilled, concerned, engaged, and relevant.



And finally, what must we do to make this happen?

1. First, how the world out there perceives our work need not constrain our efforts, but we ought to be doing work that really matters to the world, work that solves as many real problems as possible. There are plenty to tackle.
2. If we want people to trust our efforts to help them think and communicate more effectively, we ourselves must be model thinkers and communicators. When we fail to do good work, or perhaps do good work but fail to communicate what we've done to people beyond our own tight community, we fail to exemplify the very things that our work is about.
3. We must stop thinking about those outside of the infovis community as "the world out there," contrary to the title of this presentation, and work harder to see ourselves as part of the world, committed to do our part to make it better. Rather than building a bridge to reach those places that need what we do, we must move to those neighborhoods and join the PTA.

And finally, what must we do to make this happen?

1. First, how the world out there perceives our work need not constrain our efforts, but we ought to be doing work that really matters to the world, work that solves as many real problems as possible. There are plenty to tackle.
2. If we want people to trust our efforts to help them think and communicate more effectively, we ourselves must be model thinkers and communicators. When we fail to do good work, or perhaps do good work but fail to communicate what we've done to people beyond our own tight community, we fail to exemplify the very things that our work is about.
3. We must stop thinking about those outside of the infovis community as "the world out there," contrary to the title of this presentation, and work harder to see ourselves as part of the world, committed to do our part to make it better. Rather than building a bridge to reach those places that need what we do, we must move to those neighborhoods and join the PTA.

I know that nothing I've said today is profound, but I hope that many of you have found it useful to see information visualization through the eyes of people who don't do what we do, what desperately need what we do. I wish you all a fulfilling and productive year until we meet again in 2008.

For more information about Stephen Few and Perceptual Edge, including a full library of articles, be sure to visit www.PerceptualEdge.com.