



The Analytics Age

Stephen Few

Presented at the 2009 Innovators' Summit by SAS Institute

www.perceptualedge.com

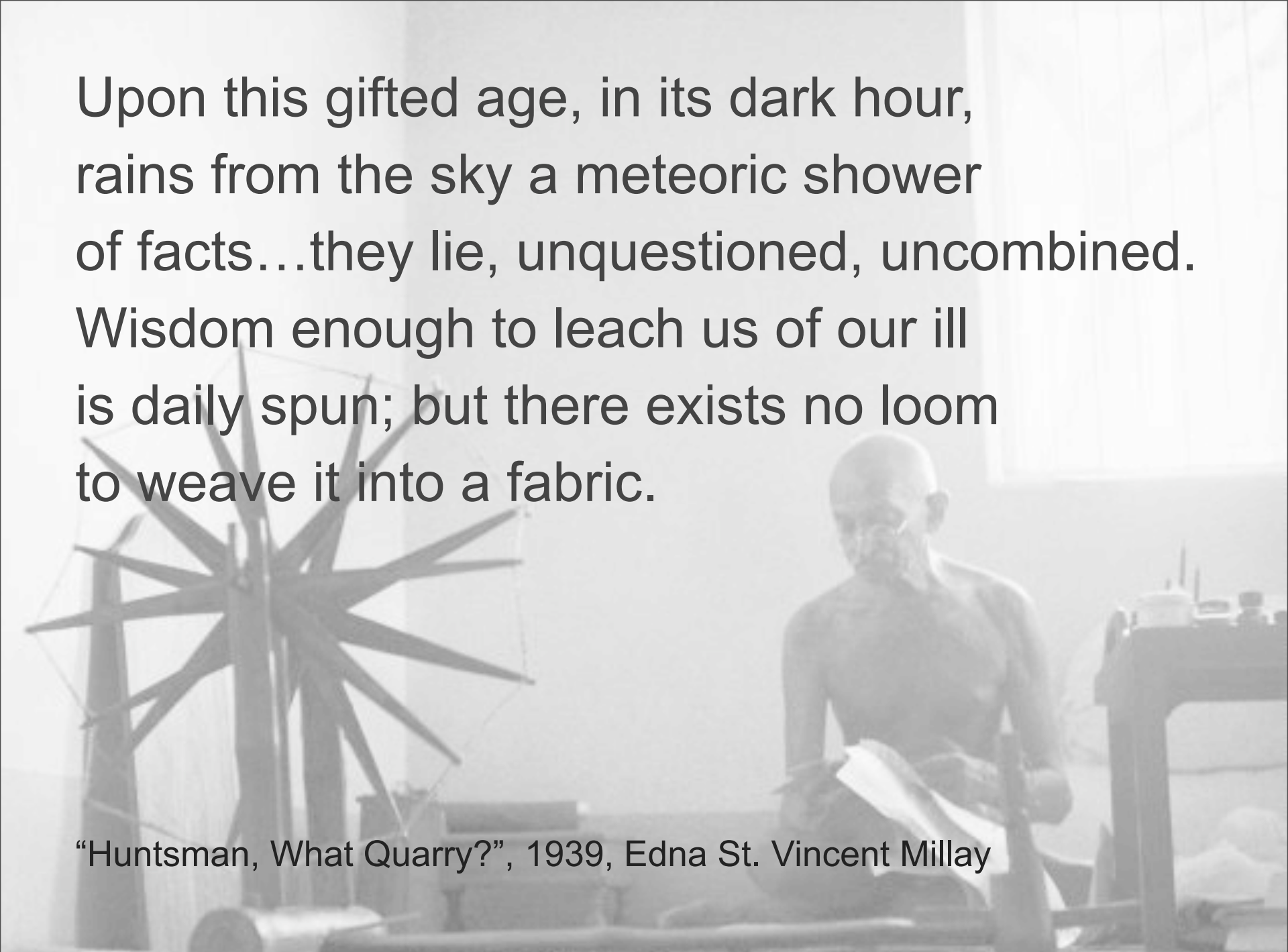
Copyright © Stephen Few 2009



Statistical Porn

Your ticket to paradise.

I was put in the mood for this conference just a few nights ago by an encounter at my local neighborhood wine bar. I noticed that a young fellow sitting near me was studying a math text, so I asked about his studies. He told me that he's in a political science doctoral program at U. C. Berkeley and that just returned from Ann Arbor where he spent the entire summer studying statistics. When I mentioned that I was leaving the next day for Chicago to speak at a SAS conference, his eyes lit up and he asked "Do you know their product called JMP?" I told him that I was in fact going to a JMP conference, he became visibly excited and exclaimed: "That product JMP is statistical porn." He meant this as a great compliment--something along the lines of "JMP gives me free and easy access to all my wildest statistical fantasies."

A black and white photograph of a man in a workshop, likely a weaver, with a large loom in the background. The man is shirtless and appears to be working with a piece of fabric. The loom is a large, complex structure with many vertical threads. The scene is dimly lit, with light coming from a window on the right.

Upon this gifted age, in its dark hour,
rains from the sky a meteoric shower
of facts...they lie, unquestioned, uncombined.
Wisdom enough to leach us of our ill
is daily spun; but there exists no loom
to weave it into a fabric.

“Huntsman, What Quarry?”, 1939, Edna St. Vincent Millay

This poem by Edna St. Vincent Millay is amazingly prescient in its description of the “information age” as we experience it today.

The amount of information that is available to businesses 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)

Most of people who are responsible for analyzing data have never been trained to do this.



The world is now overwhelmed with data, not because there's too much, but because they don't know what to do with it; they don't know how to make sense of it. The resulting pain has now reached a critical threshold; the world is desperate for answers.

I believe that we're now experiencing the dawning of the analytics age. The pieces that are needed to make this happen have come together: readily available information, sound analytical methods, an emerging understanding of the human brain--how people think--which makes it possible to build better analytical tools that take advantage of our strengths and augment our weaknesses, including data visualization, and growing appreciation in society for analytically-informed, evidence-based decision making over management by whim.

IT DASHBOARD beta

- HOME
- INVESTMENTS
- DATA FEEDS
- ANALYSIS
- FAQ



Department of Health and Human Services

Major Investments: 65
Spending on Major Investments: \$2.3 B (FY 2009)

Overall Rating

Agency

Investment Spotlight

Welcome to IT Dashboard



- View Upcoming Features
- Learn More

Blog

Friday, July 17, 2009
Evidence-based decisions
 Vivek Kundra, Federal CIO

Today, the Department of Veterans Affairs (VA), under the leadership of Secretary Shinseki and VA CIO Roger Baker, announced that it will temporarily halt 45 IT projects which are either behind schedule or over budget and work to determine whether these programs should be continued.

[Read This Post](#)

Agency Updates

Monday, July 13, 2009
 Nine agencies have now finished rating 100% of their major investments.

375

Investments Evaluated by Agency CIOs

[View detailed chart by agency](#)

Note: All descriptions, dates, and costs are as reported by agencies. Major investments (Investments Evaluated) represent only a portion of the agency's entire IT portfolio reported in Exhibit 53.

I'm excited that the Obama Administration seems to be working hard to improve its use of technology to support better analytics, but I'm frustrated that many of its efforts so far, such as this Federal IT Dashboard, are making the same silly mistakes that businesses typically make when they don't understand analytics.

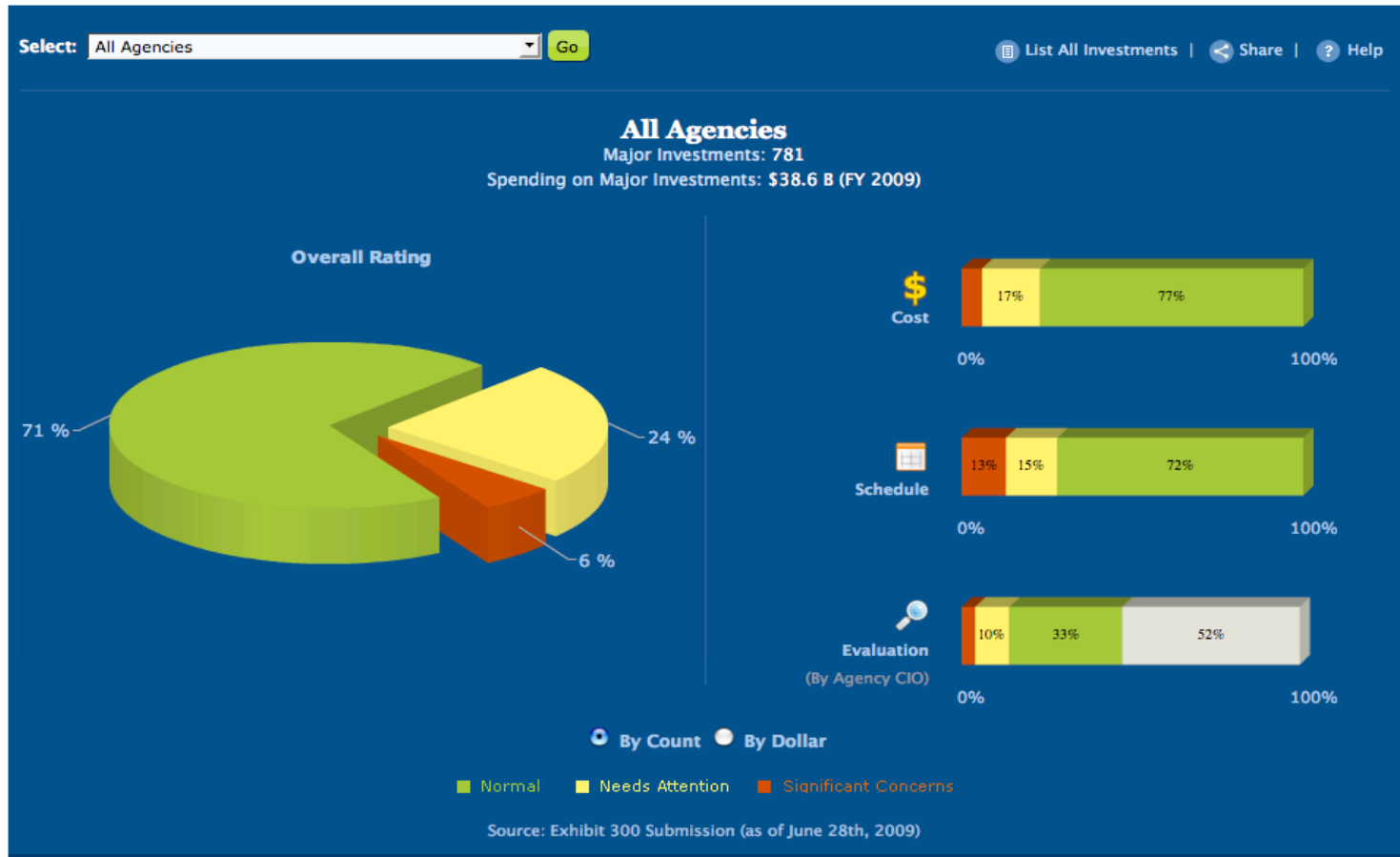
IT DASHBOARD beta

HOME | INVESTMENTS | DATA FEEDS | ANALYSIS | FAQ

Home > Performance Dashboard

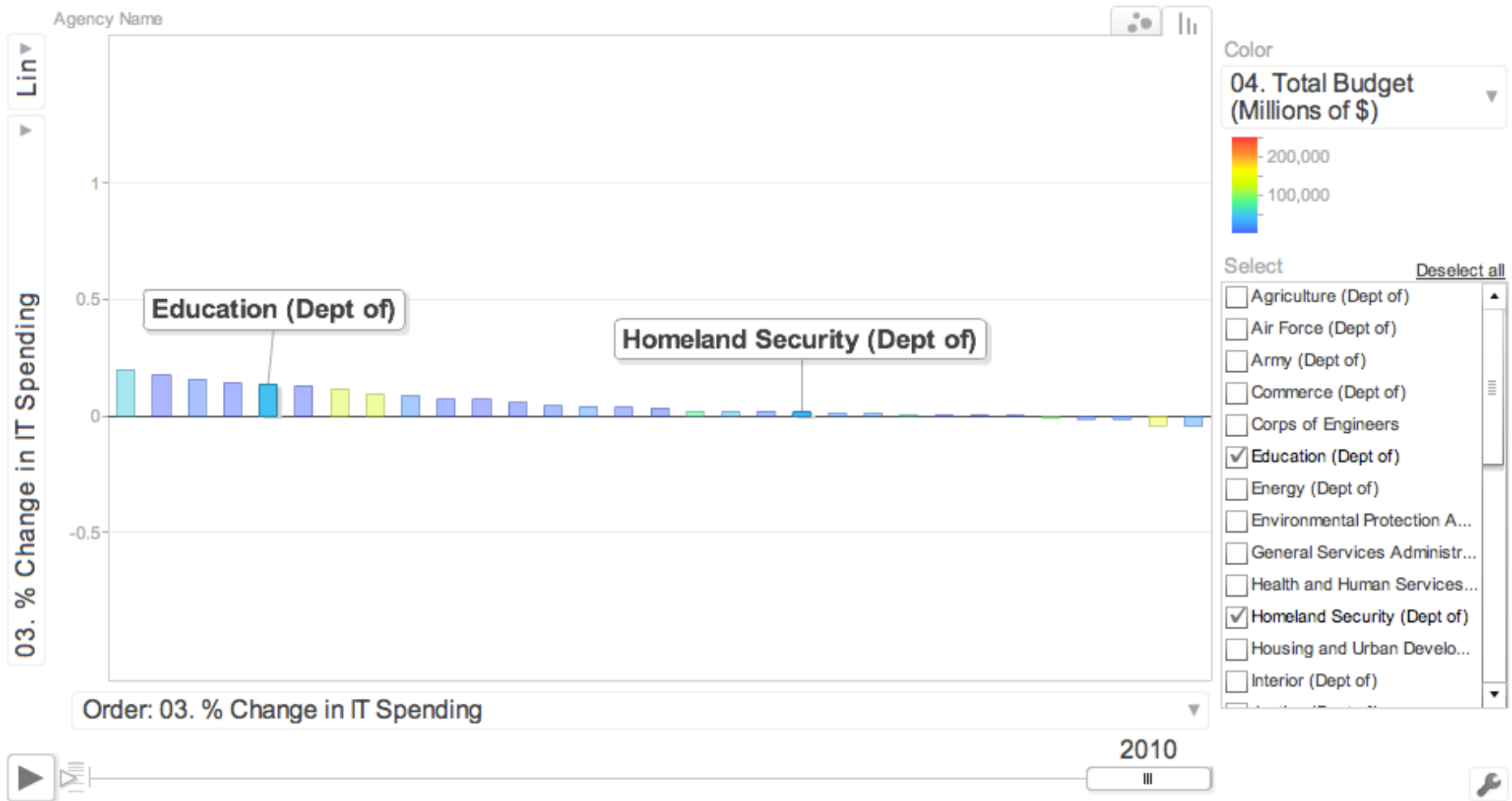
PERFORMANCE DASHBOARD

The Performance Dashboard tracks information technology (IT) investments self-reported by federal agencies and departments as "major". Major investments (Agency's Exhibit 300s) represent only a portion of the agency's entire IT portfolio (Agency's Exhibit 53). You may browse the entire IT investment portfolio using the IT Dashboard's [treemap visualization](#).



Note: All descriptions, dates, and costs are as reported by agencies.

This application was supposed designed for analytics, but the people who put it together clearly don't understand how people make sense of data.



By emulating some of the flashy displays that they've seen, they've provided the means to watch values such as these IT spending figures dance around the screen to show change through time in a way that the human brain can't follow.



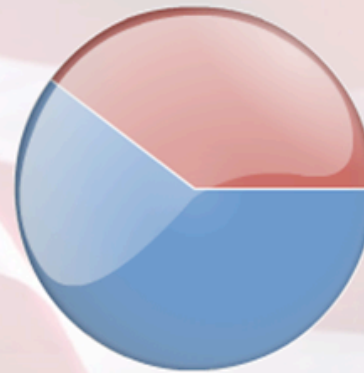
US

US by State

Select a Population Group

- All
- Both
- Men
- Women
- White
- Black or African American
- Hispanic or Latino ethnicity
- Detailed characteristics (to...
- Detailed characteristics (m...
- Detailed characteristics (wo...

All
Both



■ non-institutional population ■ Civilian labor force

Employment



Unemployment



For the last few years, in their desperation for answers, businesspeople have been lured into buying silly products like Crystal Xcelsius that catch their attention with eye-catching visuals but offer no real analytical functionality or substance. Most of the products that claim to support information visualization today were built by organizations that don't know anything about it. You probably won't be surprised to learn that this particular product was built by a team of developers who came directly from the video game industry.

Notice the cute gauges on the bottom. What do they take up all that space to tell us? Two numbers, without any context for making sense of them or even quantitative scales to make the positions of the pointers meaningful. The pie chart sure is shiny, isn't it. Besides its silly design with three-dimensionality and reflected light, the pie chart isn't even being used properly. The red slice--civilian labor force--is actually a subset of the non-institutional population, not a separate segment of the larger population.

People like you who are statisticians and understand analytics need to raise your voices in warning and show the world solutions that actually work. This is not the time to don the cloak of mystery and pretend that statistics is the esoteric realm of an enlightened cult; it's the time to make statistics relevant to the world with open arms.

Statisticians
are now sexy.



Statisticians are about to get the respect they deserve. No longer will you be kept in some back room huddled over your computers. The role of statistician is becoming the sexiest role in the job market. Statisticians will become the priests of the modern world--those who talk with God and can divine the mysteries of data.



Hal Varian

U.C. Berkeley and Google

I keep saying the sexy job in the next ten years will be statisticians. People think I'm joking, but who would've guessed that computer engineers would've been the sexy job of the 1990s? The ability to take data—to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it—that's going to be a hugely important skill in the next decades, not only at the professional level but even at the educational level for elementary school kids, for high school kids, for college kids. Because now we really do have essentially free and ubiquitous data. So the complimentary scarce factor is the ability to understand that data and extract value from it.

I think statisticians are part of it, but it's just a part. You also want to be able to visualize the data, communicate the data, and utilize it effectively. But I do think those skills—of being able to access, understand, and communicate the insights you get from data analysis—are going to be extremely important. Managers need to be able to access and understand the data themselves.

You always have this problem of being surrounded by “yes men” and people who want to predigest everything for you. In the old organization, you had to have this whole army of people digesting information to be able to feed it to the decision maker at the top. But that's not the way it works anymore: the information can be available across the ranks, to everyone in the organization. And what you need to ensure is that people have access to the data they need to make their day-to-day decisions. And this can be done much more easily than it could be done in the past. And it really empowers the knowledge workers to work more effectively.

“Hal Varian on how the Web challenges managers”, www.McKinseyQuarterly.com



Greet the dawn of analytics by...

...demanding effective tools

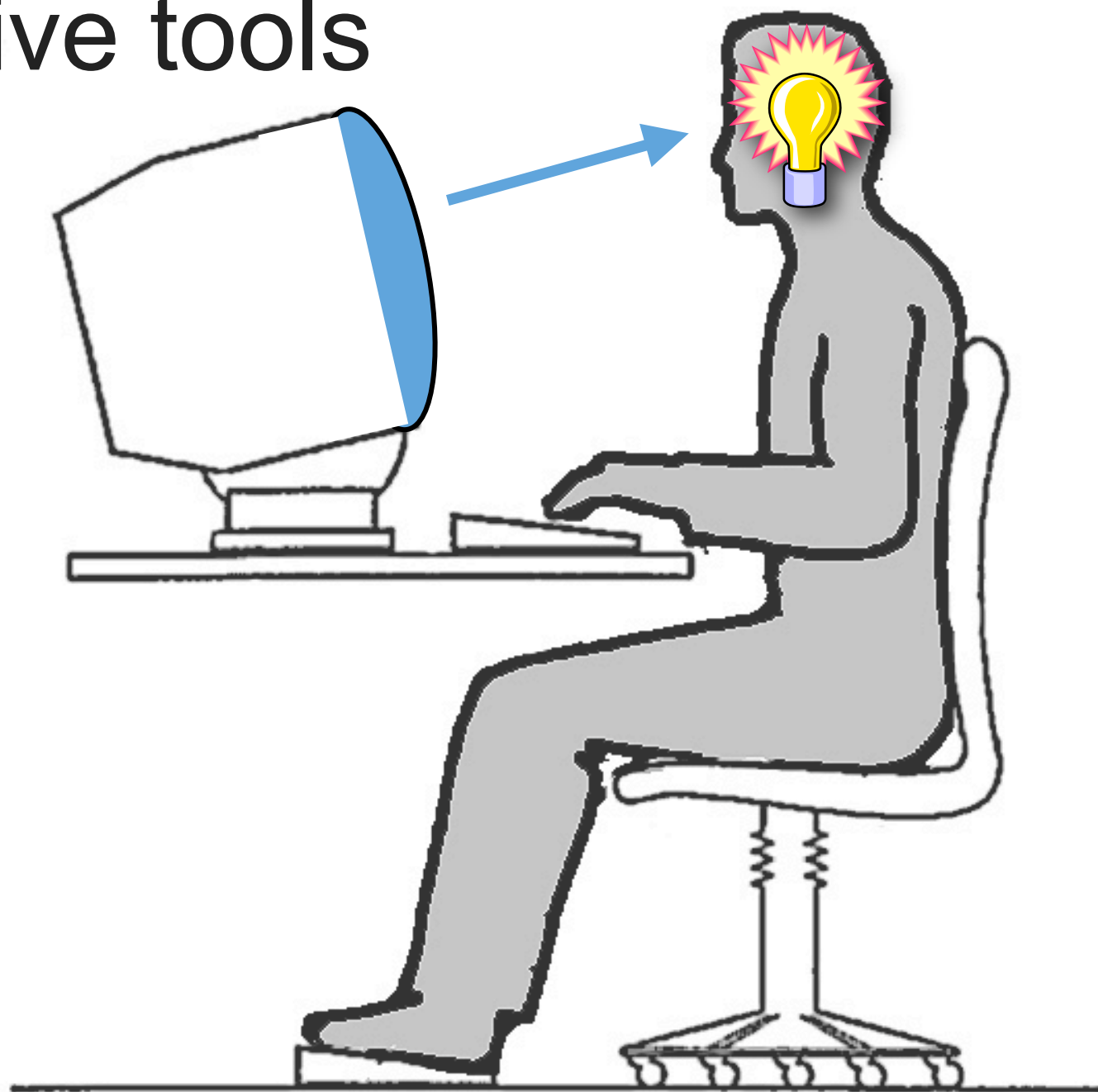
...sharing the work

...learning to communicate

As statisticians, you are in a unique position to usher in the dawn in a way that allows the light of understanding to light the world or you can restrict the sun's rays to light far less--the choice is yours. To allow the light of understanding to spread its rays liberally, you can do the following:

- Demand effective data analysis tools. Many silly tools are being marketed to a desperate world today that don't know the first thing about data analysis. For instance, as the potential of data visualization has become recognized, software vendors with no clue about data visualization are selling tools that are embarrassingly bad. The days of organizations getting by with nothing but Excel must be put behind us.
- Share the data analysis workload with others in the organization, including non-statisticians, supporting their efforts to do what they can handle so that you can focus on the complicated stuff.
- Learn how to communicate your analytical findings in a way that people can understand. You haven't done your job if you can't tell the stories in clear and compelling ways to those who need the information.

Effective tools

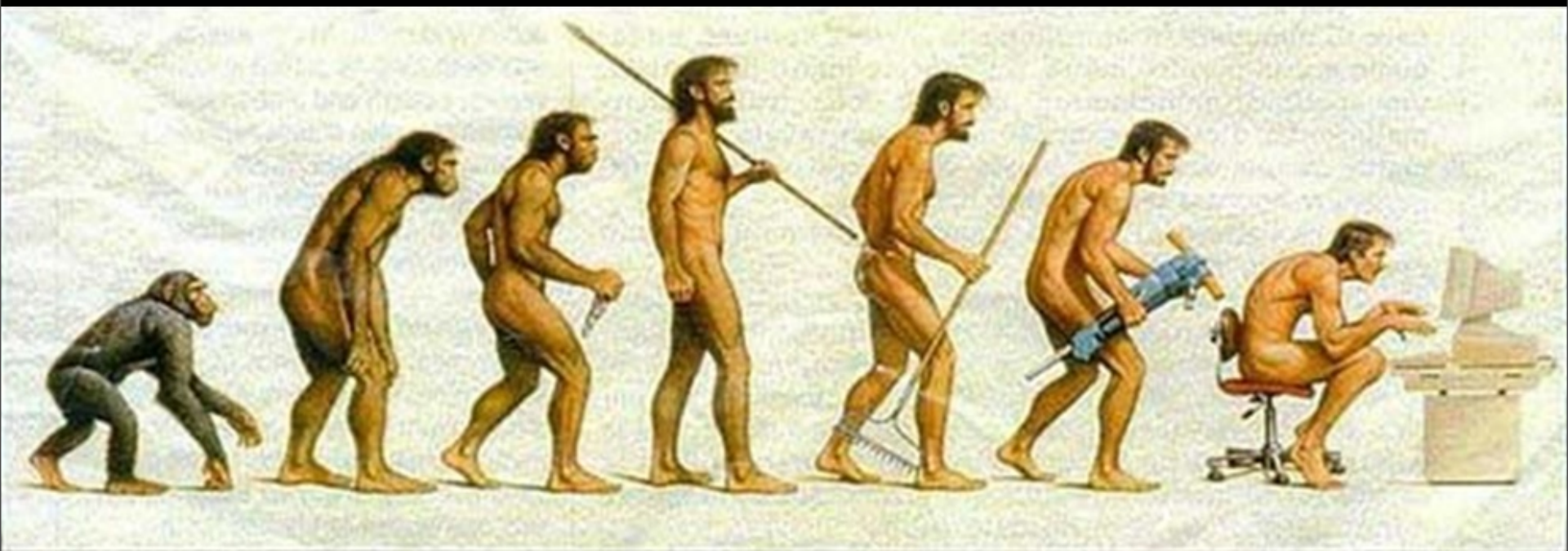


Technology has an important role to play, but computers can't make sense of data; only we can. It is up to us to bring our powerful senses and minds to the task, but no matter how good we are, we'll struggle without good tools that can present information to our eyes in ways that they can easily and clearly perceive, and can augment our minds to work around their limitations. JMP is one such tool. No matter how good our tools are, however, we must never forget that they're only tools. We are the analysts, and it's up to us to develop the skills that will make analytics successful.



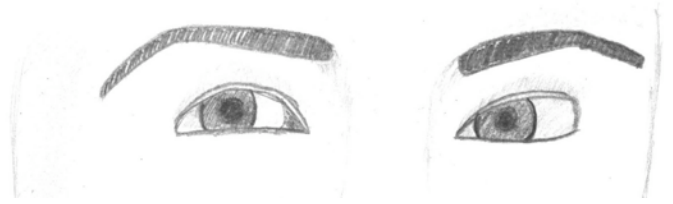
Long before we had tools such as hammers and saws, we knew how to build houses. We are builders by nature. Tools simply support what we already know how to do by extending our reach, increasing our strength, and speeding the process.

Tools



Tools only help us when we use them properly. When misused, our reliance on them can make us lazy, weak, and stupid.

The right balance



The process of data exploration and analysis is one that, to a great extent, ought to involve using our eyes. There is an intimate connection between seeing clearly and thinking clearly. Information visualization makes possible an ideal balance between unconscious perceptual and conscious cognitive processes. With the proper tools, we can shift much of the analytical process from conscious processes in the brain to pre-attentive processes of visual perception, letting our eyes do what they do extremely well.

Why the eyes?



I became fascinated by data visualization—the use of our eyes in close collaboration with our brains to make sense of information—not because I’m a visually-oriented person (I’m actually much more verbal than visual) but because it offers solutions to real analytical problems that concern us, solutions that stand apart in their ability to enlighten.

“I see!”

There is an intimate connection between seeing and thinking. For this reason, when we talking about understanding, we tend to use terms that are related to vision. When we suddenly come to understand something we say “I see!” It’s hard to think of terms that describe understanding that don’t refer to sight.



70%



John Medina

Human perception is amazing. I cherish all five of the senses that connect us to the world, that allow us to experience beauty and an inexhaustible and diverse wealth of sensation. But of all the senses, one stands out dramatically as the primary and most powerful channel of input from the world around us, and that is vision. Approximately 70% of the body's sense receptors reside in the eye.

Edward Tufte, says: "Clear and precise seeing becomes as one with clear and precise thinking." (*Visual Explanations*, Edward R. Tufte, Graphics Press: Cheshire, CT.1997 page 53)

In his wonderful book *Brain Rules*, John Medina points out that "Vision trumps all other senses." He illustrates this with an account of an experiment involving expert wine tasters. Wine experts wax prolific with detailed and imaginative descriptions of the qualities they can discern when they smell and sip a glass of wine: "Aggressive complexity with just a subtle hint of shyness." Distinct vocabularies exist for white and red wines, which should never overlap. A group of researchers in Europe decided to test the objectivity of these perceptions and did so using test subjects from the heart of the wine world, the University of Bordeaux. They wanted to know, what if we dropped tasteless, odorless red dye into white wines and then asked these experts to taste and describe them, would their perceptions be altered by what they saw? Can you guess what they found? Every one of the wine experts used red wine vocabulary to describe what they tasted. What they saw overrode what they smelled and tasted.

Fortunately, visual perception is not only dominant, it is powerful and can be relied on to do much of the work when analyzing data.



Analytical Zen

Effective data analysis is a *way of seeing*. We can all learn to look at information in ways that open our minds to greater discovery and understanding. Unfortunately, what usually passes as data analysis results in an assortment of disconnected facts that are rarely woven together into a meaningful fabric. This is a travesty, because simple data visualizations and interactions can be used to immerse us in the flow of analysis.

It's appropriate to relate the philosophy of Zen to data sense-making. The Zen journey involves peeling away the veils of illusion that cloud our view, layer by layer, until we can see clearly. When applied to data sense-making, it involves coaxing our data from the shadows into the light of understanding, and from there into the realm of wisdom by using our newfound awareness to make the world a better place.

The path to enlightenment



Simple

Contextual

Focused

Flexible

Multifaceted

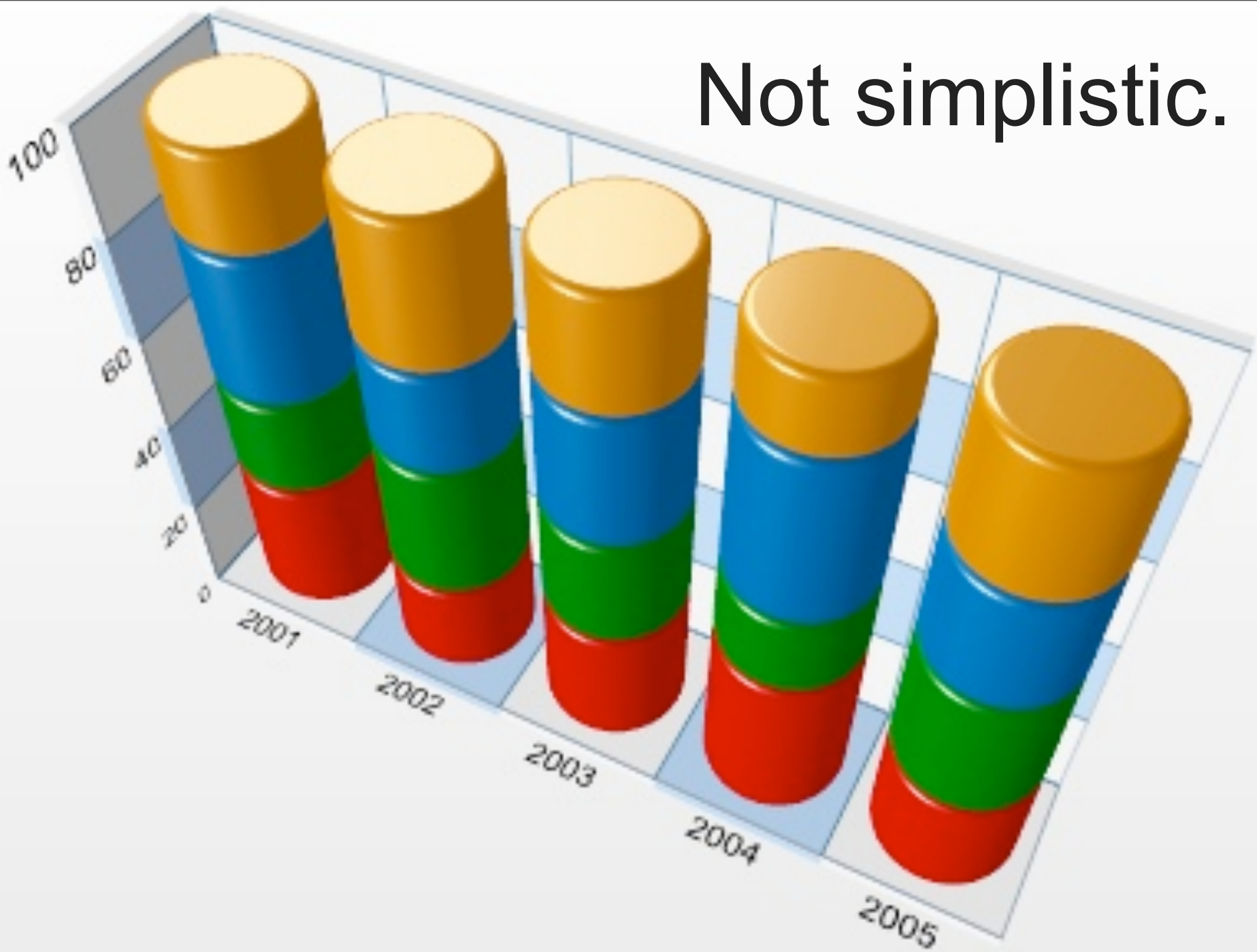
I want to talk about five characteristics of visual analysis that are needed to reach analytical enlightenment.



Simple

It's amazing how a few simple well-placed such as these lines can communicate so much. For me, simplicity isn't a design preference, it is a fundamental principle of usability. Thinking and communicating are dramatically improved through simplification by abstracting information in a way that eliminates all but what's essential for the purpose at hand.

Not simplistic.



By simple, I don't mean simplistic. Nothing is gained by decorating data displays or dumbing them down.

Information

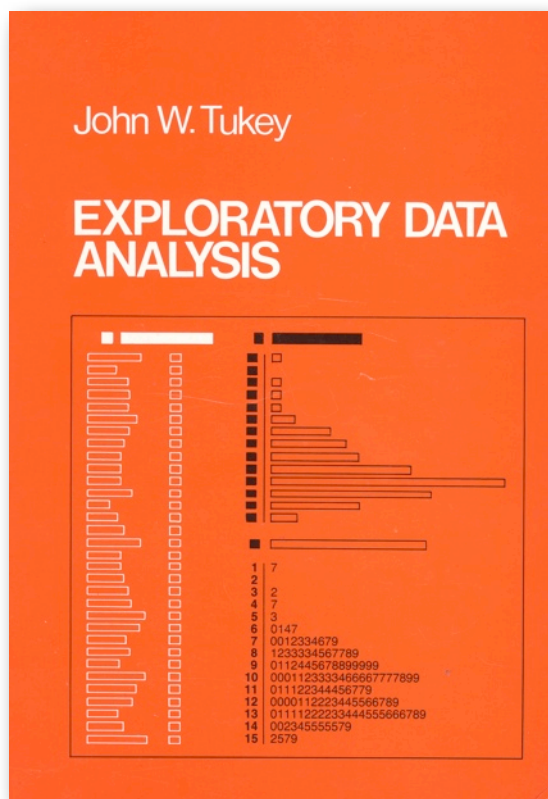
Its essence



Not overly simplified

Not overly complicated

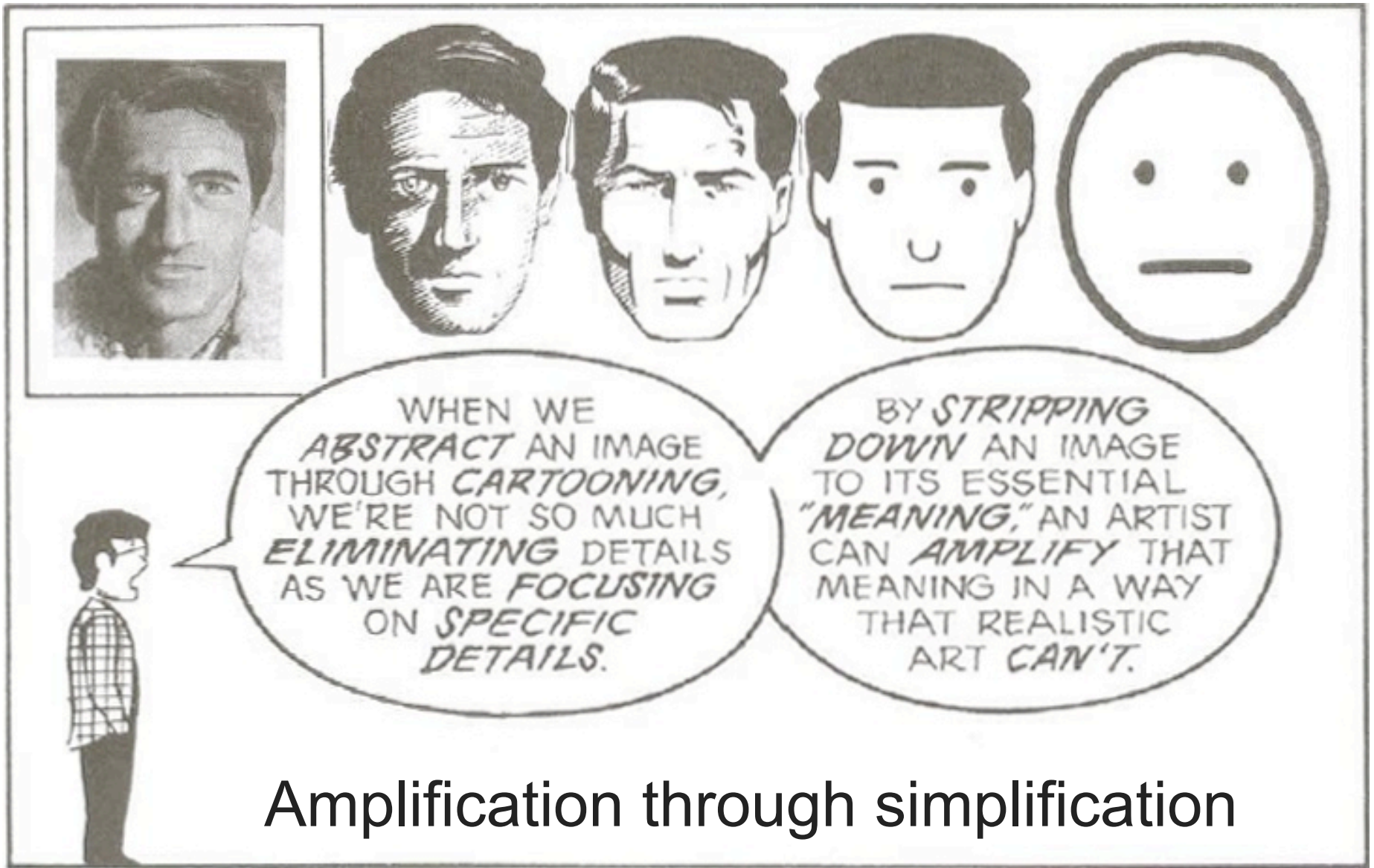
Simplicity doesn't only mean keeping things from becoming unnecessarily complicated, but also keeping them from becoming simplified to the point where anything useful is lost. Simplicity involves a balanced approach to information that allows us to see it for what it is.



“When is one expression better than another for analysis? Basically, when the data are more simply described, since this implies easier and more familiar manipulations during analysis and, even more to the point, easier and more thorough understanding of the results.” John Tukey

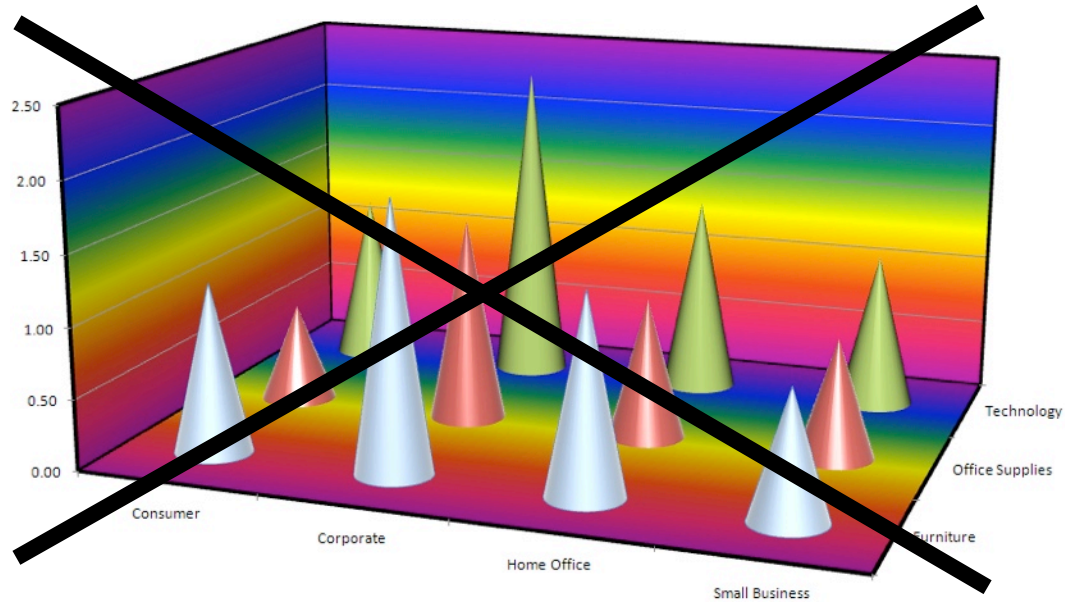
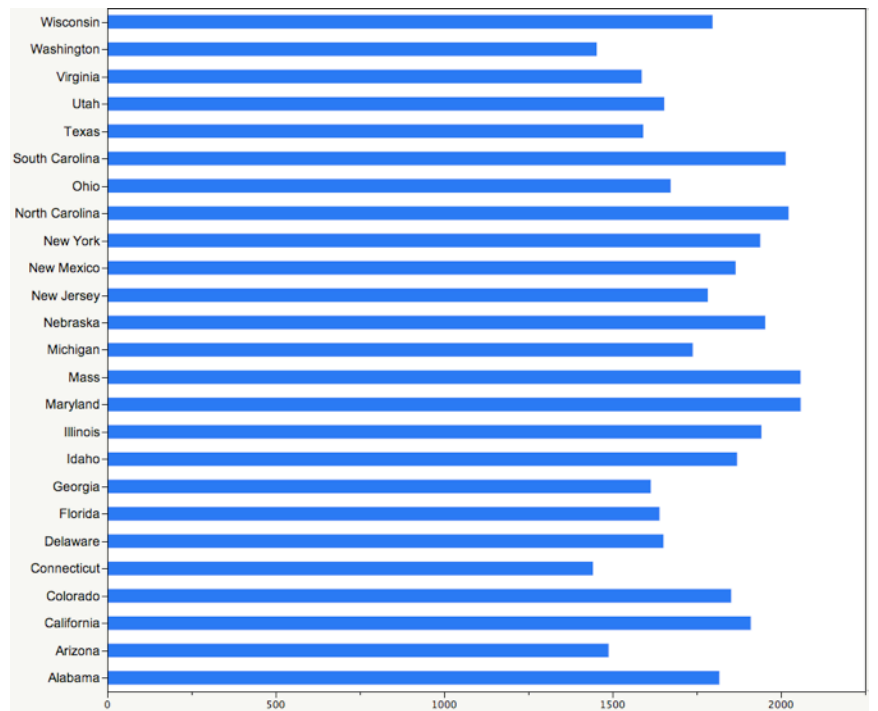
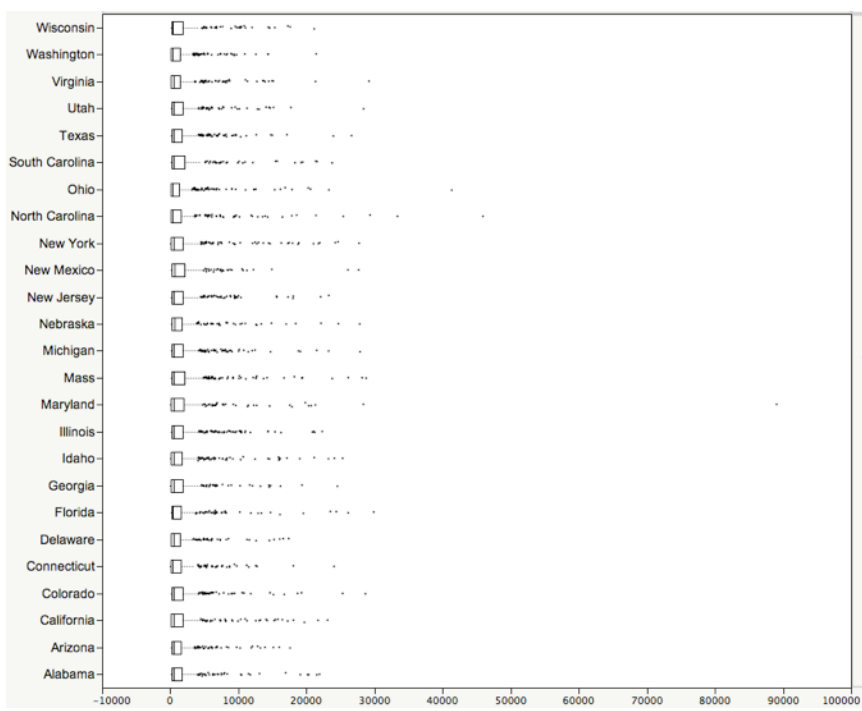
Princeton statistician John Tukey, the author of the milestone work *Exploratory Data Analysis*, once posed the question and responded with the answer above. (*The Collected Works of John W. Tukey*, John W. Tukey, Wadsworth, Inc.: Belmont, CA, 1988, page 12)

Scott McCloud, *Understanding Comics*



Comic book artist Scott McCloud describes his work as the amplification of meaning through simplification of visual representation. To do this well, you must be able to recognize what's essential to the data's meaning and strip away all else.

Antoine de Saint-Exupery wrote: "Perfection is achieved, not when there is nothing more to add, but when there is nothing left to take away."



When we're ready to visualize our data, we want to choose from a small set of effective graphs. More choices are not always better. More than we need complicates our lives. Ideally, we'd like the software to make an intelligent guess about the type of graph that works best with the data, then make it easy to switch to something else if it guesses wrong. Once we've chosen the graph that we want, don't want it to appear with a bunch of visual fluff decorating it, which we'll have to work to ignore.



Simplicity is the ultimate sophistication.

Leonardo da Vinci

Simplicity isn't always easy; it takes skill.

Simple data analysis tools are easy to build; they're built on an understanding of how we see and think, so the display can be matched with our visual and cognitive strengths. There is no room for fluff; anything that might distract from a clear view of the data.

Contextual

Compared
to what?

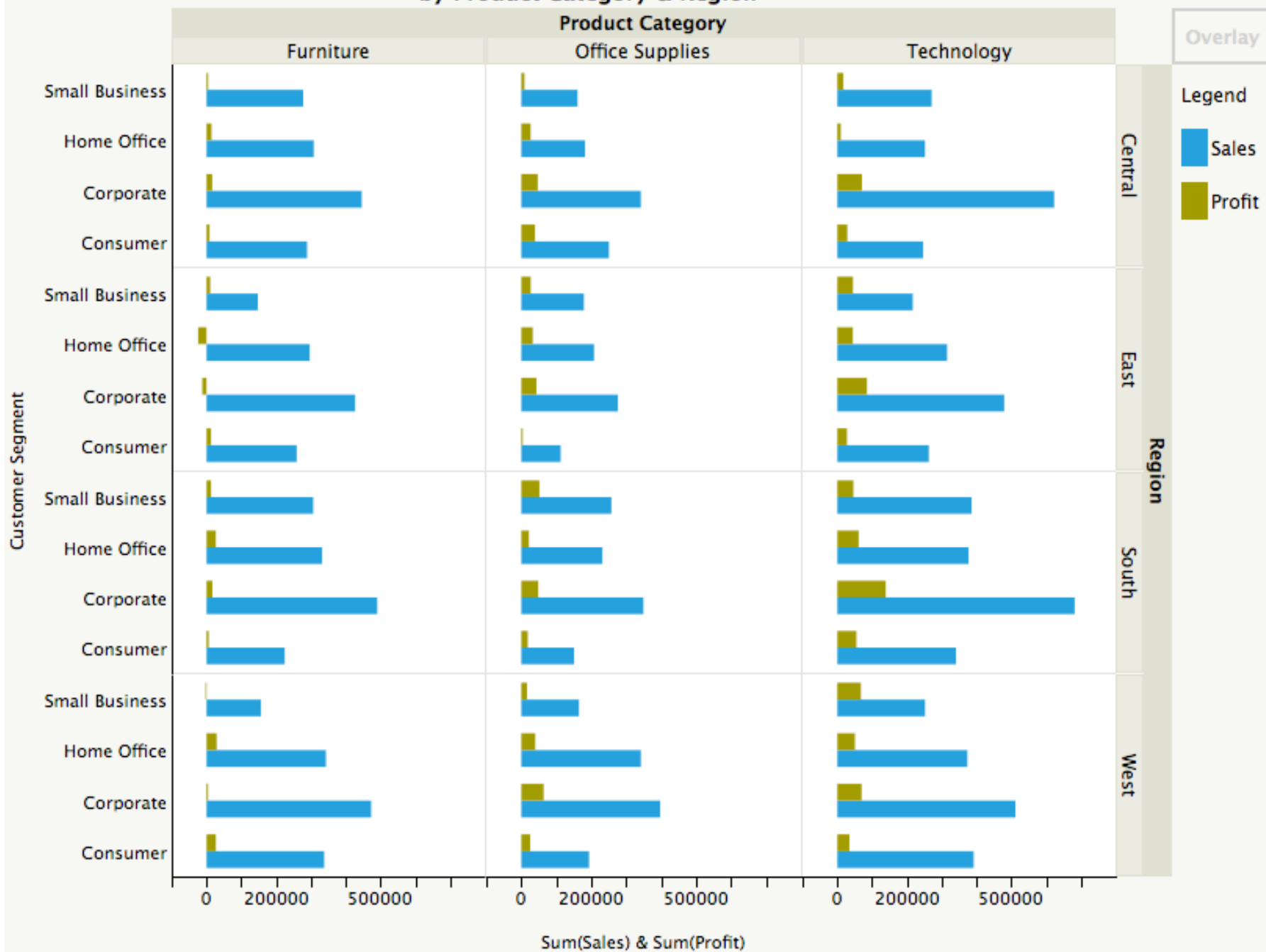


Edward Tufte says that the primary question that we ask when striving to understand quantitative information is “Compared to what?”

Quantitative stories cannot be told effectively just by throwing numbers at people. Numbers alone—even those that measure something perceived as important—are meaningless unless we present them in context. In part, this means that we should reveal the pedigree of the numbers (that is, where they came from and how we might have adjusted them). Even more importantly, however, this means that we must provide additional information to which the numbers can be compared. We discover the meanings in numbers primarily by comparing them to other numbers and by comparing the patterns, trends, and exceptions that live within them to those formed by other relevant sets of numbers. Appropriate comparisons are the context that make numbers meaningful in a way that allows us to form judgments, make decisions, and take action.

For quantitative information, comparisons are vital to the story.

Customer Segment vs. Sum(Sales) & Sum(Profit) by Product Category & Region



Displays like this visual crosstab, constructed using JMP's Graph Builder, provide means of comparison.



Focused

Effective data analysis tools help us attend without distraction to those aspects of data that are important to understanding.

We're not aware of everything we see.



“Ever saw a magic show and wondered just how the magician took your watch without you even noticing? Ever wonder why is it that you can search for a set of misplaced keys for a long time, only to later find them sitting in the exact place where you were looking? Research has shown that we don't always see everything we're looking at, and that attention plays a big part in what consciously registers to us. The effect where we're blind to things we don't attend to is known as 'Inattentional Blindness'.”

(Source: This demonstration of inattentional blindness and the explanation above was prepared under the direction of Ronald A. Rensink of the University of British Columbia. Several other examples of this visual phenomenon can be found at <http://psychlab1.psych.ubc.ca/~viscoglab/>.)

Data Filter for profit by product

Data Filter

Clear

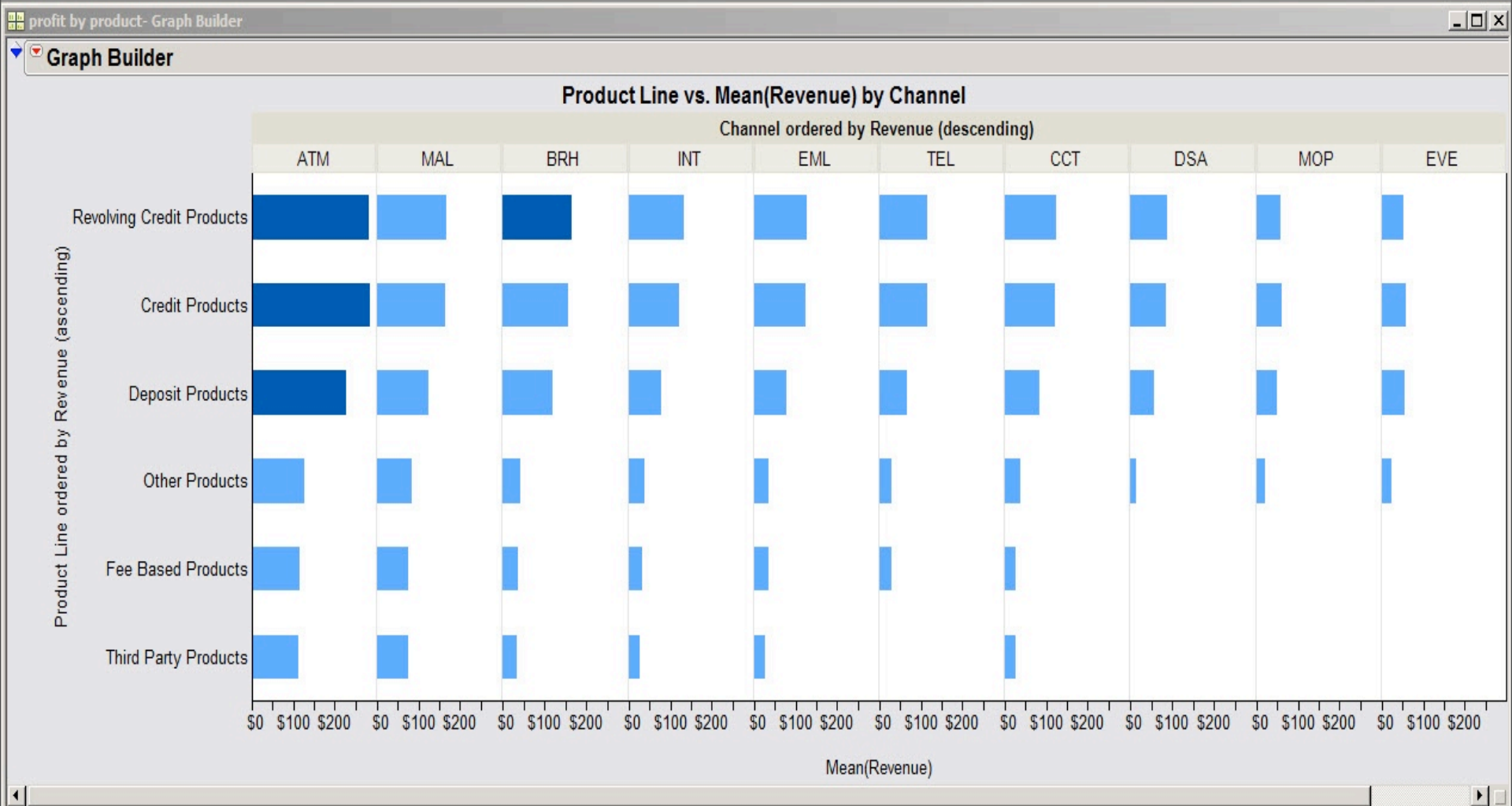
39 matching rows

Select Show Include

$4660.2 \leq \text{Cost of Sales} \leq 9256.4$

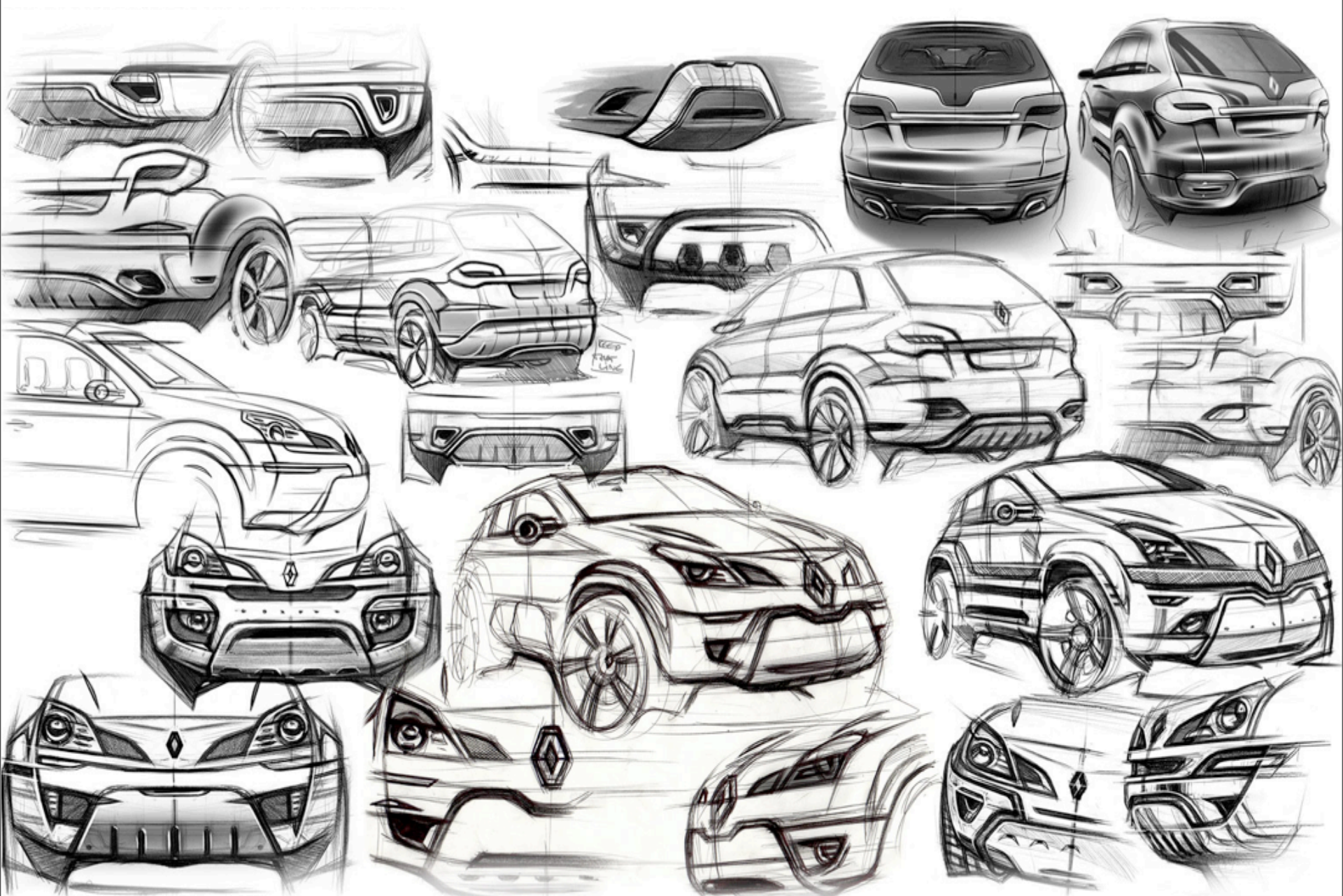
+

Start Over

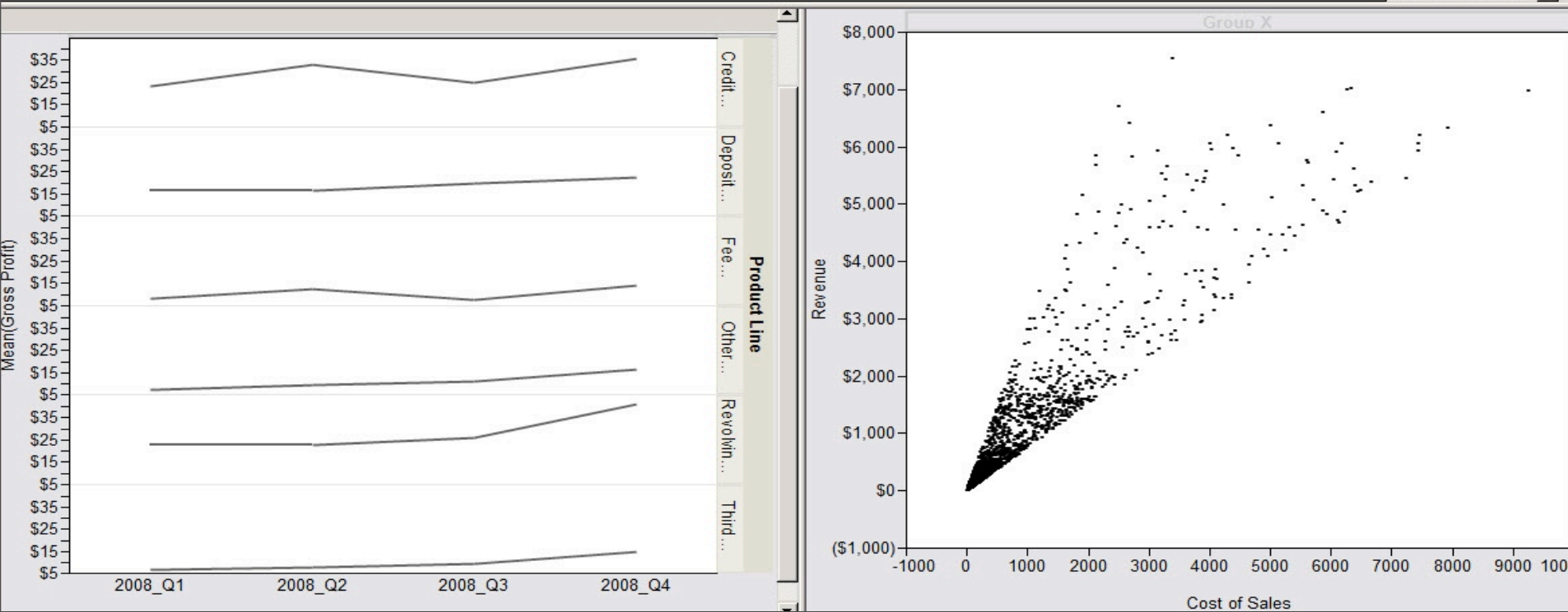
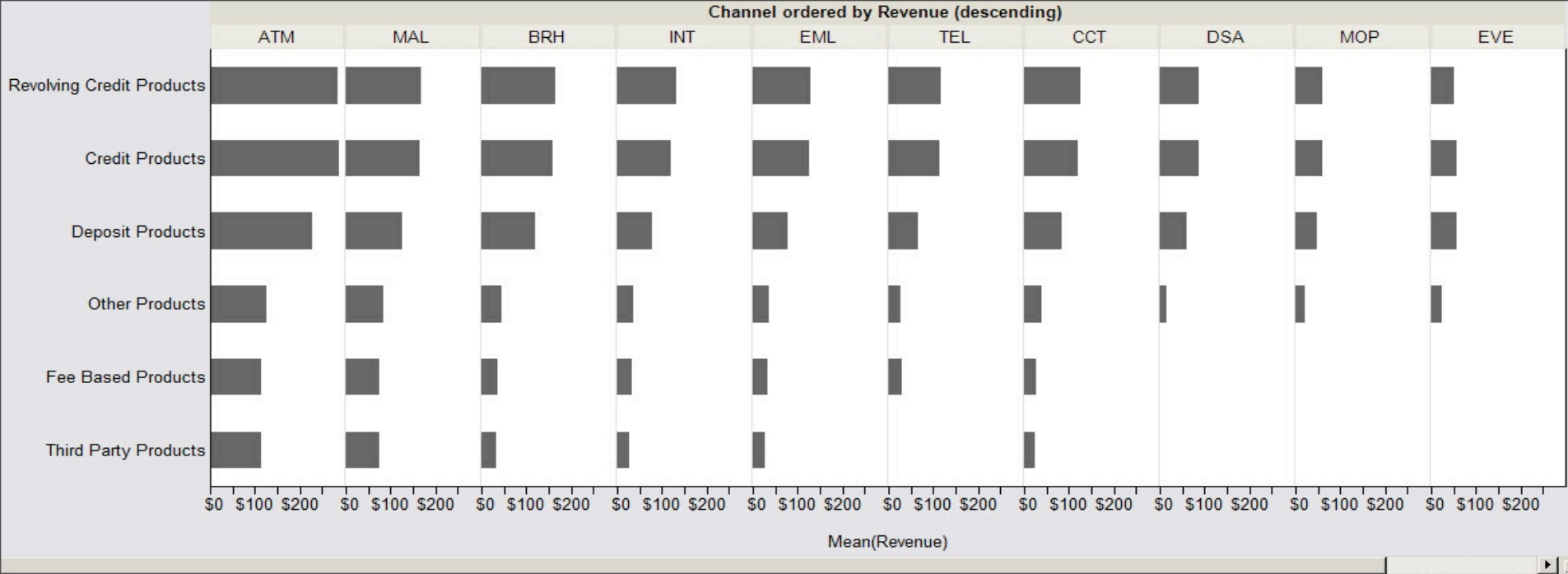


To attend to what's important, we need the ability to easily and immediately remove distractions. Dynamic data filters like the slider shown above are one convenient means to remove data that doesn't interest us at the moment.

Multifaceted



The next step along the path to enlightenment involves being multifaceted--that is the practice of looking at information from every perspective: from a distance as well as up close and from every possible angle. No single view of our data will ever tell us the whole story. Any single view without being seen in the context of the whole has the potential of misleading us entirely.



Having the ability to place several views of a data set in front of our eyes at the same time makes it possible to see connections that would remain invisible if we were forced to look at each of the views independently.

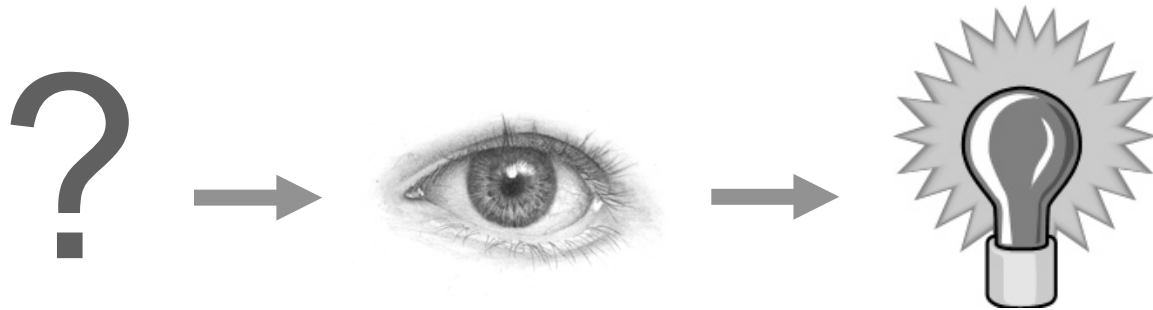


Flexible

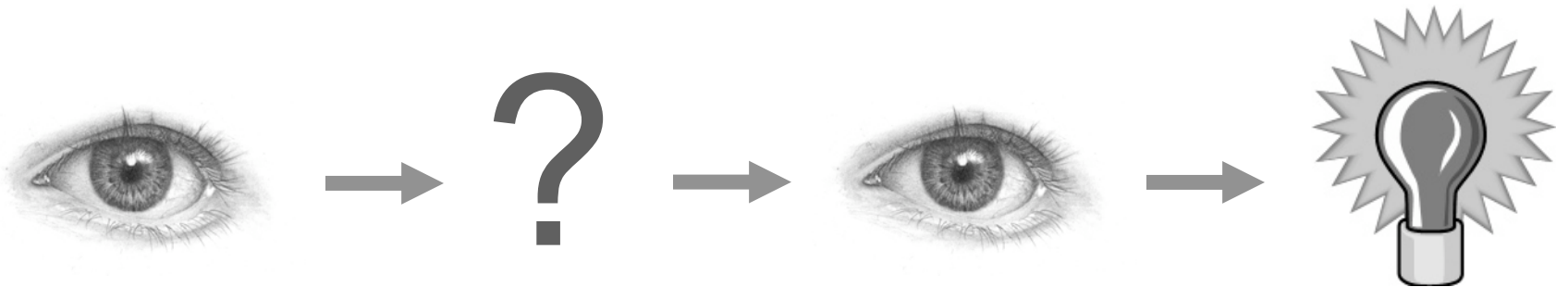
Effective data analysis tools allow us to explore the landscape freely, without getting in the way or forcing us down narrow paths.

Directed vs. exploratory navigation

Directed



Exploratory



Visual data analysis consists of many steps and many roads to get us from where we start, knowing little, to where we need to be, understanding much. Some ways of navigating your path from step to step are more effective than others. “Data analysis, like experimentation, must be considered as an open-minded, highly interactive, iterative process, whose actual steps are selected segments of a stubbornly branching, tree-like pattern of possible actions.” (*The Collected Works of John W. Tukey*, John W. Tukey, Wadsworth, Inc.: Belmont, CA, 1988, pages 5 and 6)

At the most fundamental level, analytical navigation can be divided between two approaches: directed and exploratory.

Directed analysis begins with a specific question that you wish to answer, proceeds to a search specifically for the data that will answer that question, such as a particular pattern, and hopefully results in finding the answer.

Exploratory analysis begins by looking at the data without predetermining what you expect to find, proceeds to noticing things in the data that are interesting and asking a question about it, then proceeding in a directed fashion in search of an answer to that question.

Both approaches are vital to data analysis. The main point I wish to make here is that comprehensive analysis requires that you sometimes start with a blank slate and let the data itself direct you to items worth examining.

Contained within the data of any investigation is information that can yield conclusions to questions not even originally asked. That is, there can be surprises in the data...To regularly miss surprises by failing to probe thoroughly with visualization tools is terribly inefficient because the cost of intensive data analysis is typically very small compared with the cost of data collection.

(*The Elements of Graphing Data*, William S. Cleveland, Hobart Press, 1994, pages 8 and 9)

Shneiderman's mantra

“Overview first,
zoom and filter,
then details-on-demand.”

An overview is almost always a good place to start when you begin to explore a new data set for the first time. It makes sense to find some high point from which to view the landscape of data around you before setting off in any particular direction. One way to do this is to look at the frequency distribution of all the variables--both dimensions and measures--to see where values fall in each. Here I've arranged the major dimensions across the top row and the major measures along the bottom. This view can really come alive when enriched through linked highlighting. For example, if we want to see how the West region relates to the whole across each variable, let's select in in the Market view to cause the West's portion of values to be highlighted in all the views at once.

When new recruits by intelligence organizations are trained in spy craft, they are taught a method of observation that begins by getting an overview of the scene around them while being sensitive to things that appear abnormal, not quite right, which they should then focus in on for close observation and analysis.

A visual information-seeking mantra for designers: 'Overview first, zoom and filter, then details-on-demand.'
(*Readings in Information Visualization: Using Vision to Think*, Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman, Academic Press, San Diego, California, 1999, page 625)

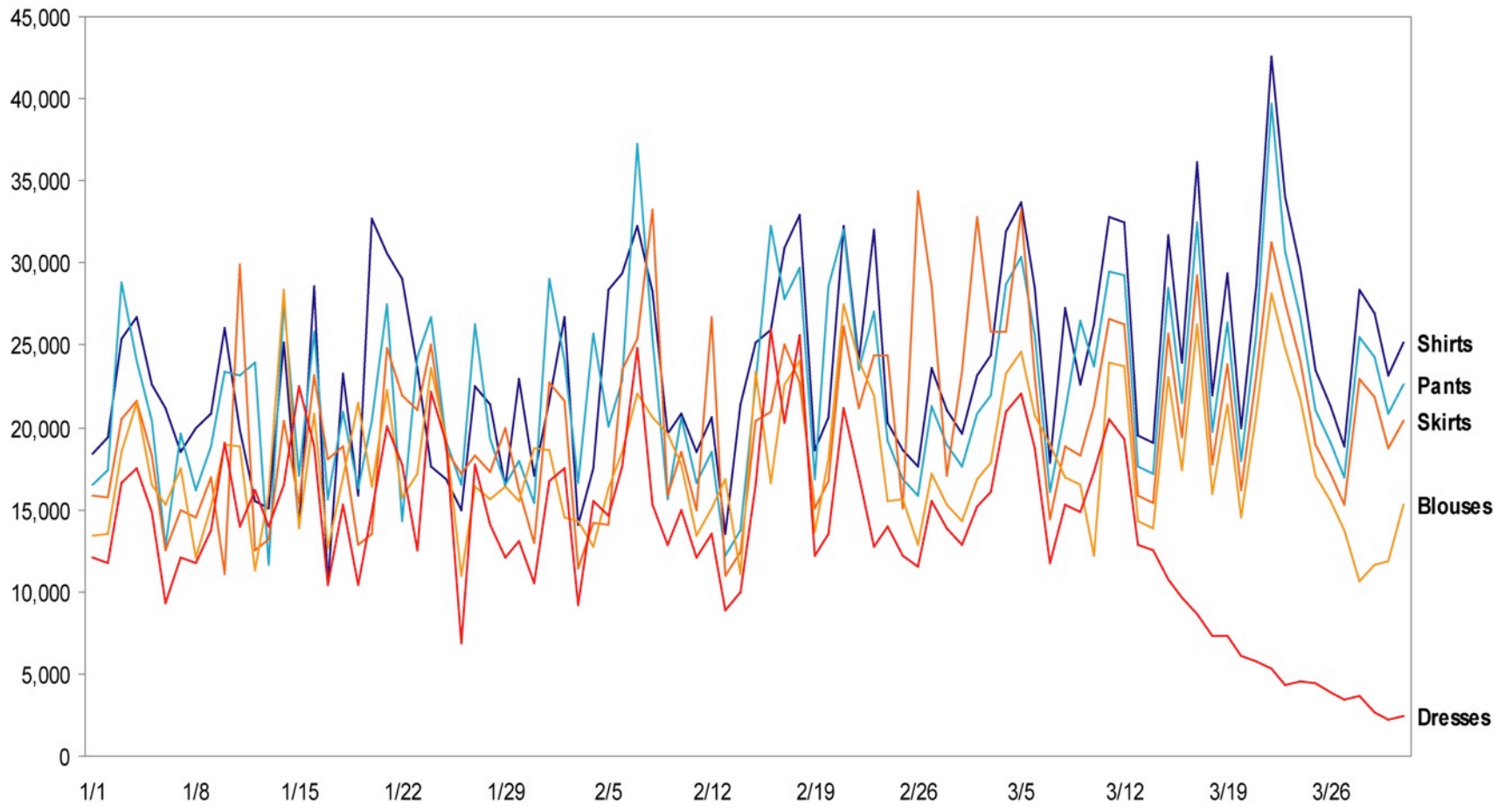
Having an overview is very important. It reduces search, allows the detection of overall patterns, and aids the user in choosing the next move. A general heuristic of visualization design, therefore, is to start with an overview. But it is also necessary for the user to access details rapidly. One solution is overview + detail: to provide multiple views, an overview for orientation, and a detailed view for further work.

(*Ibid.*, page 285)

Users often try to make a 'good' choice by deciding first what they do not want, i.e. they first try to reduce the data set to a smaller, more manageable size. After some iterations, it is easier to make the final selection(s) from the reduced data set. This iterative refinement or progressive querying of data sets is sometimes known as hierarchical decision-making.

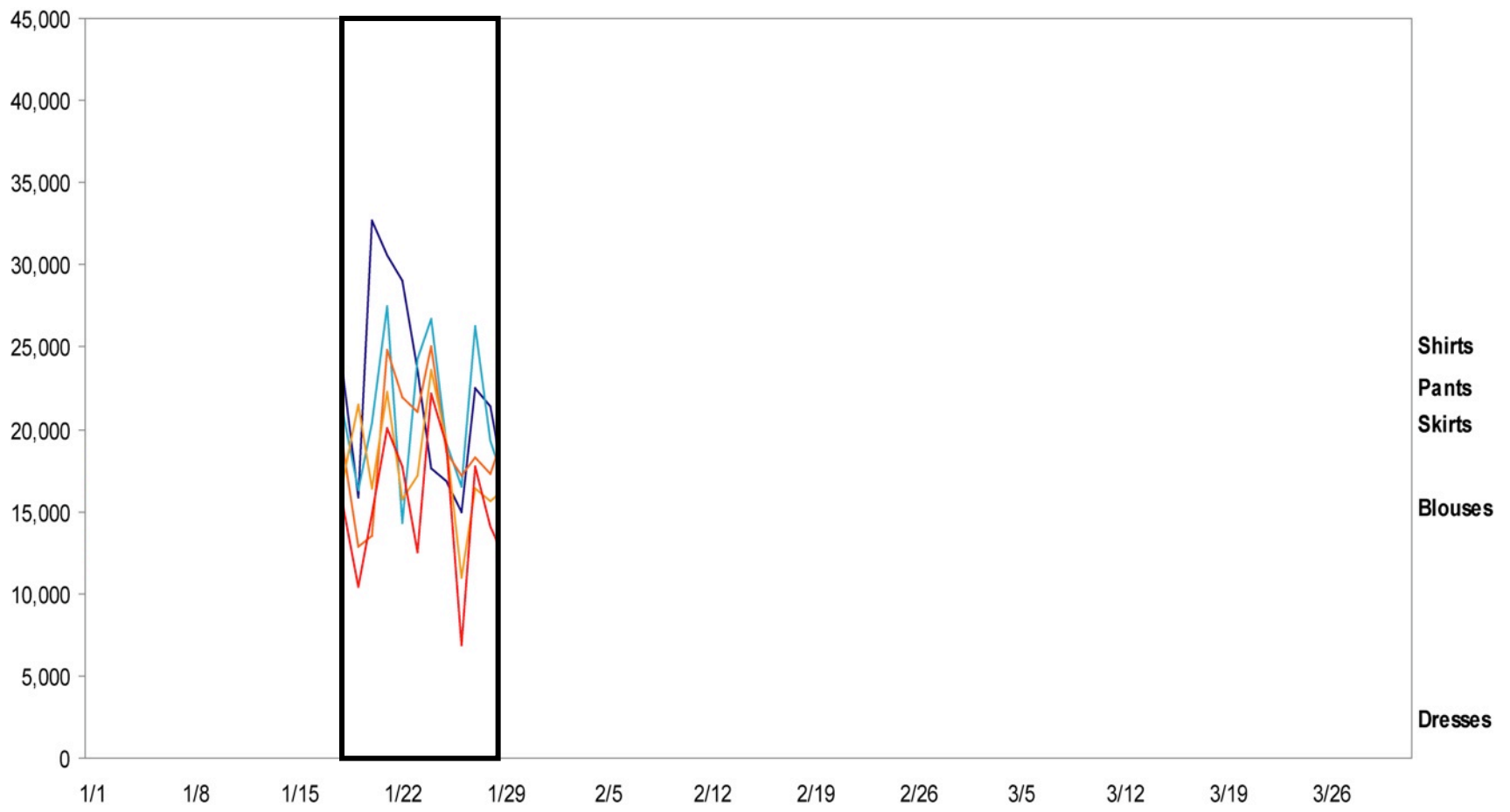
(*Ibid.*, page 295)

Overview first, zoom and filter, then details-on-demand



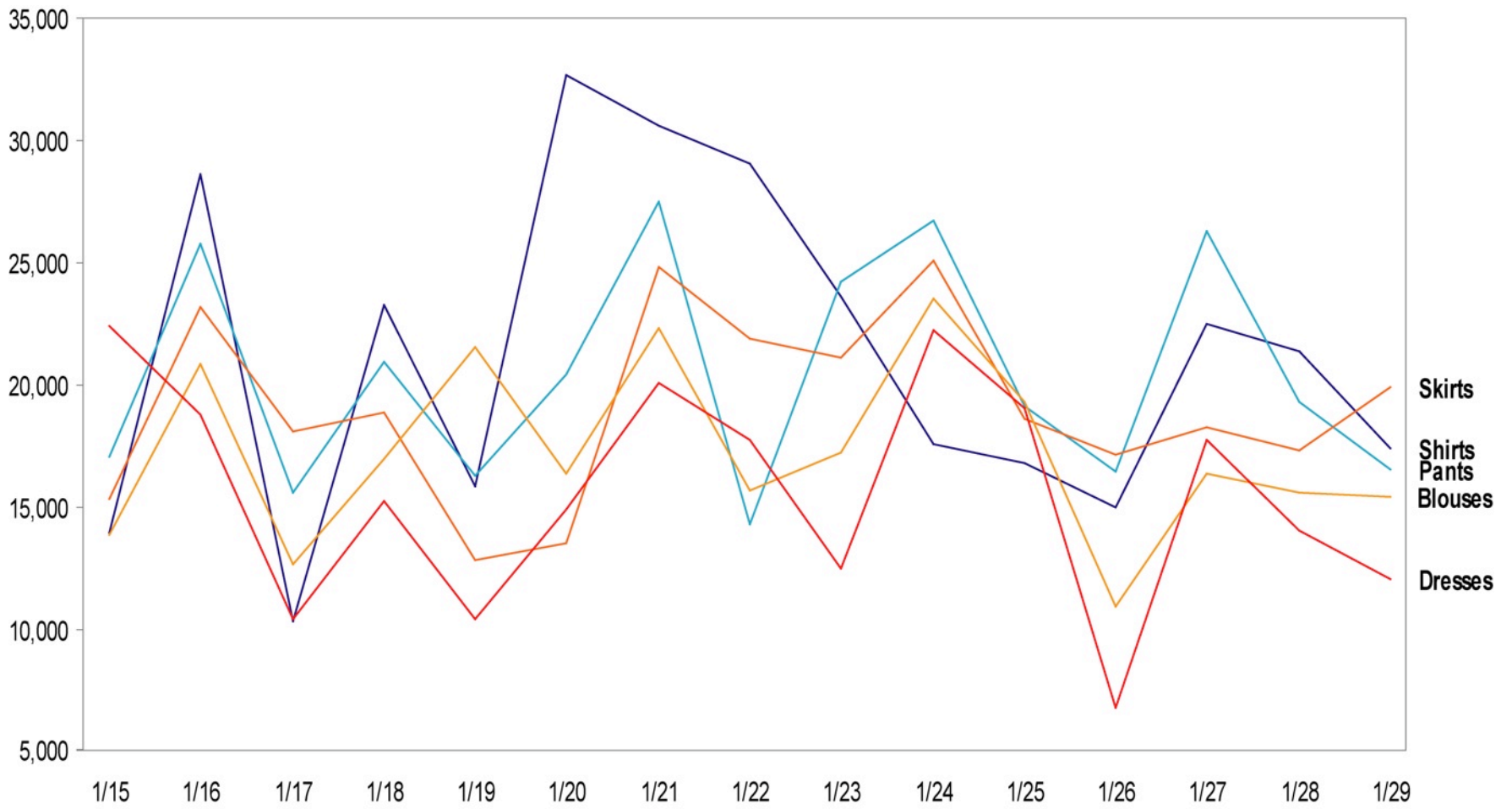
Shneiderman's technique begins with an overview of the data – the big picture. Let your eyes search for particular points of interest in the whole.

Overview first, **zoom** and filter, then details-on-demand



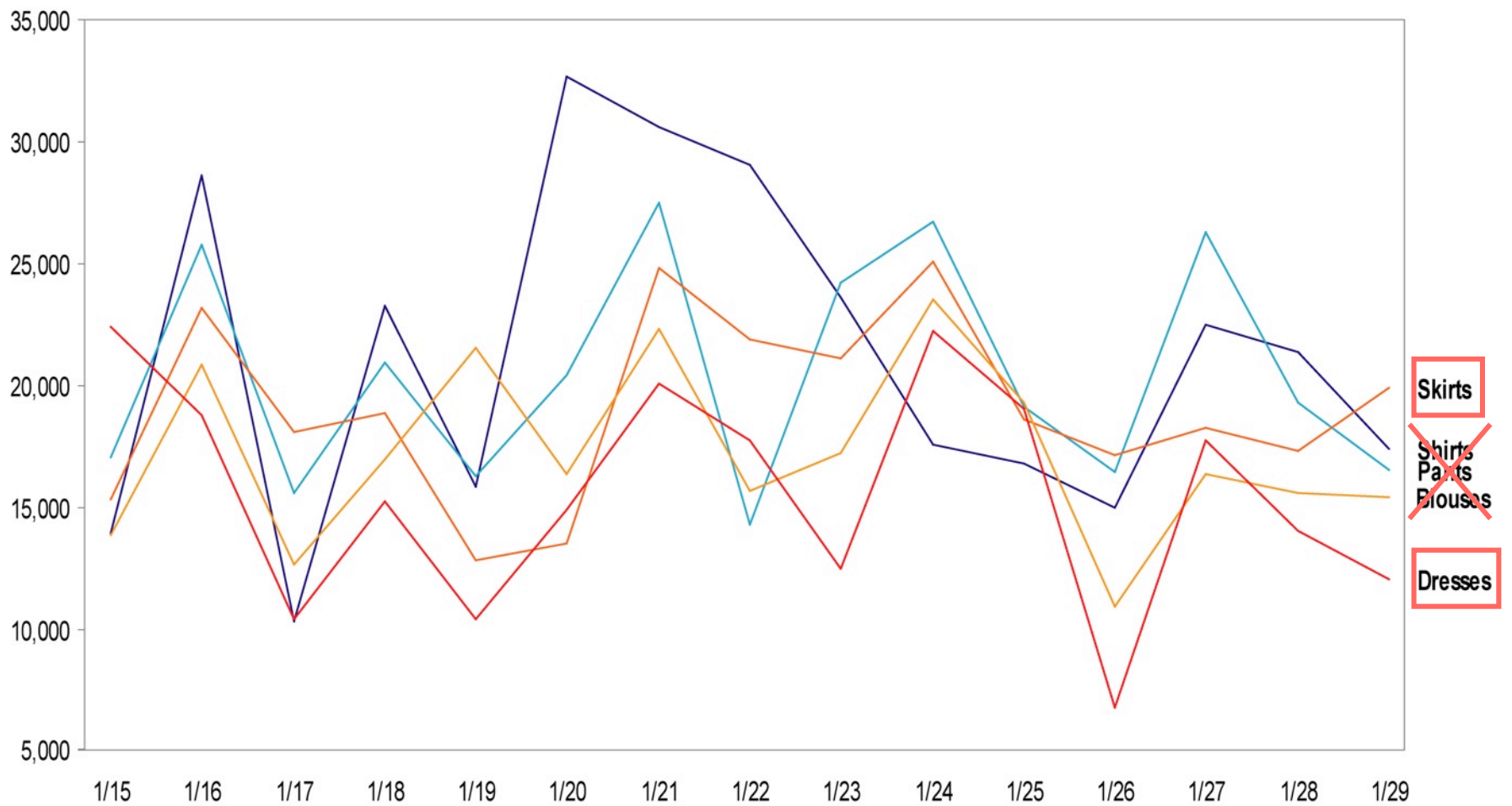
When you see a particular point of interest, then zoom in on it.

Overview first, **zoom** and filter, then details-on-demand



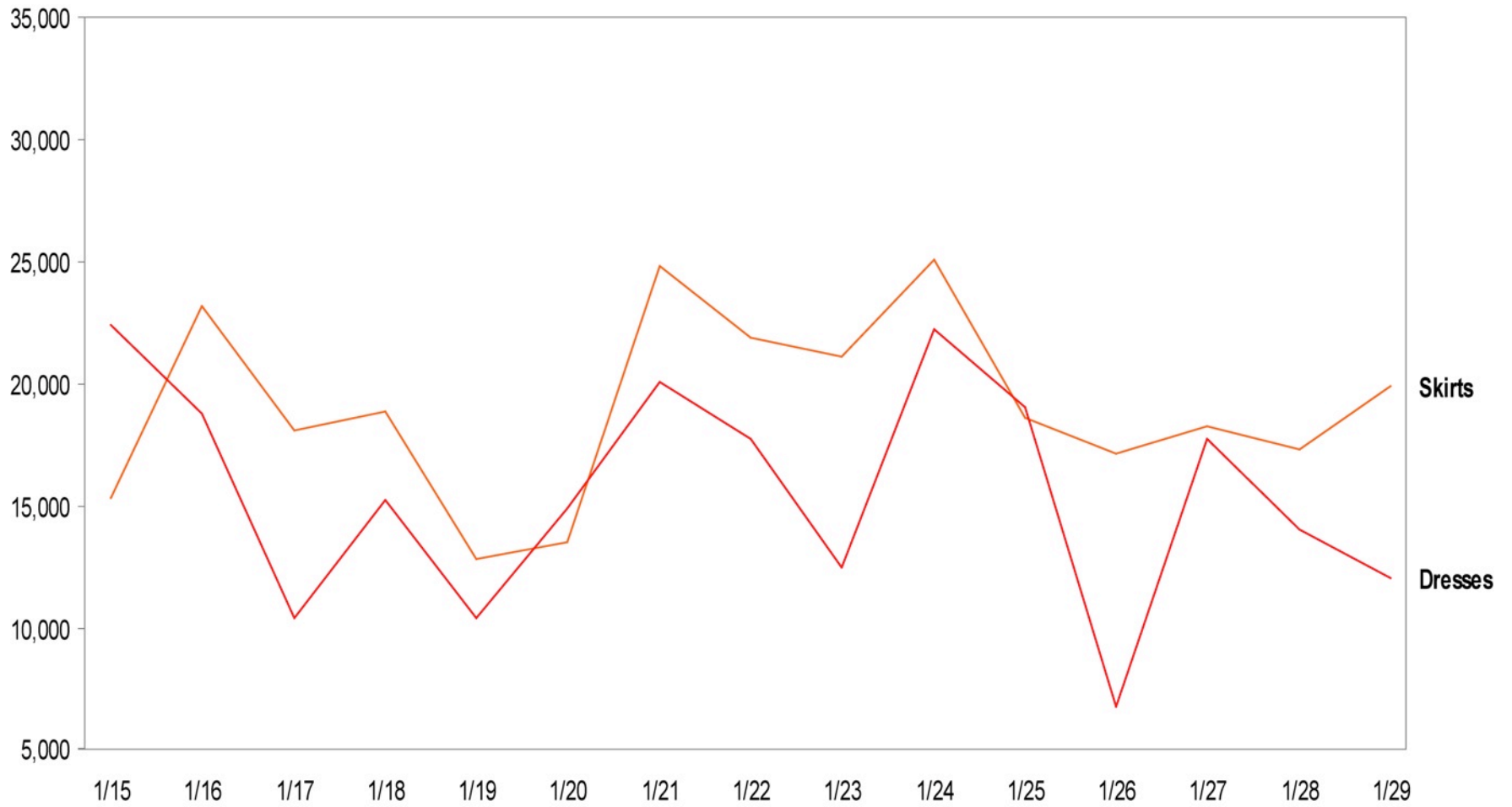
Once you've zoomed in on it, you can examine it more closely and in greater detail.

Overview first, zoom and **filter**, then details-on-demand



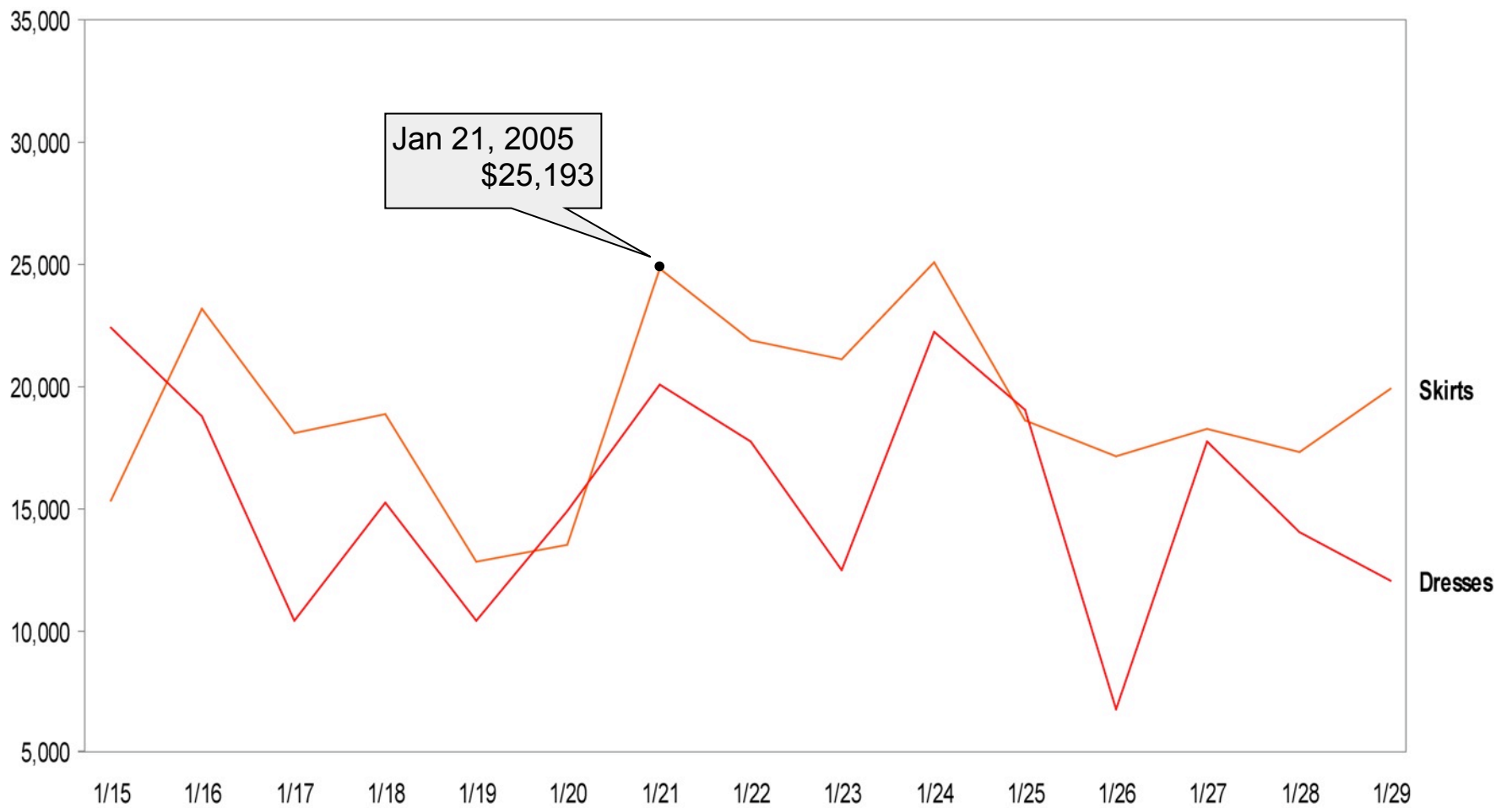
Often you must remove data that is extraneous to your investigation to better focus on the relevant data.

Overview first, zoom and **filter**, then details-on-demand

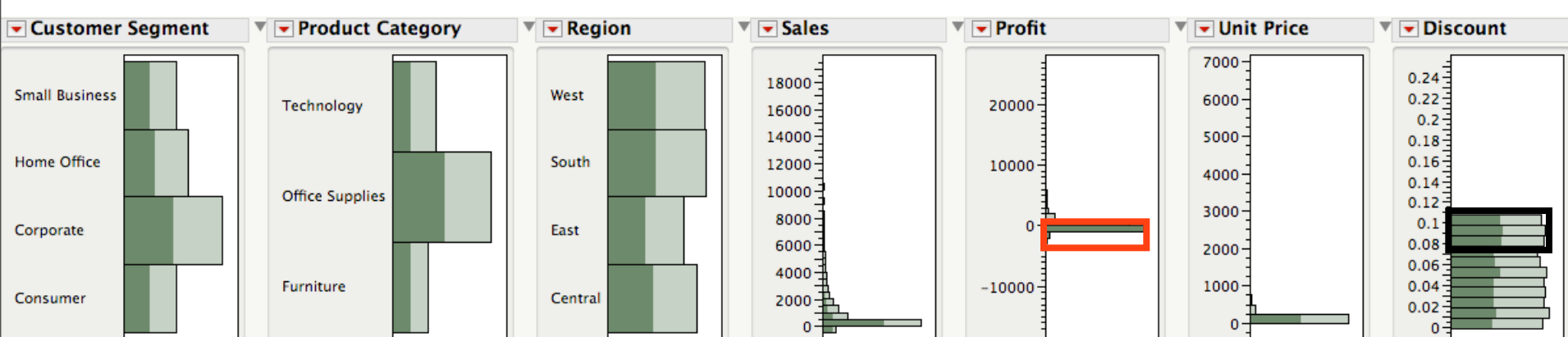
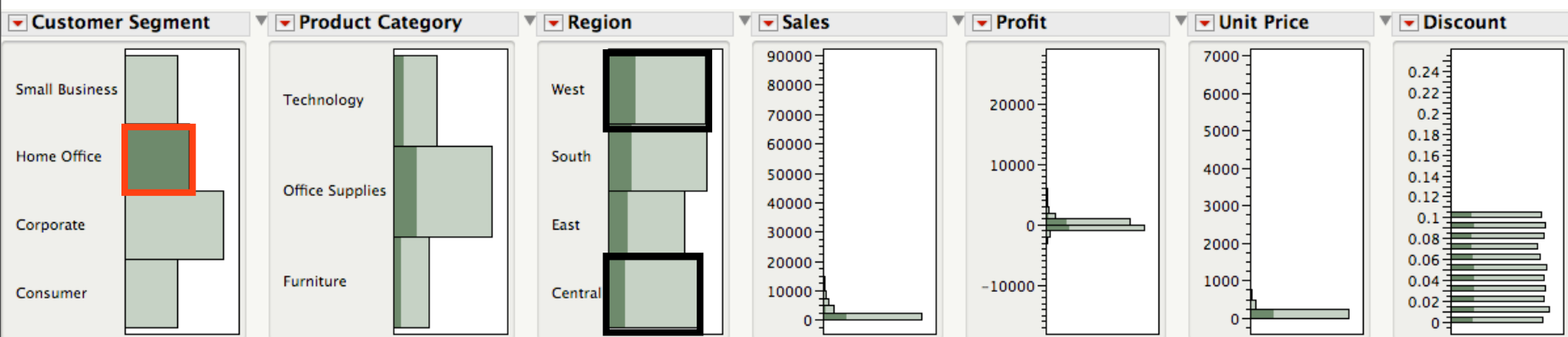
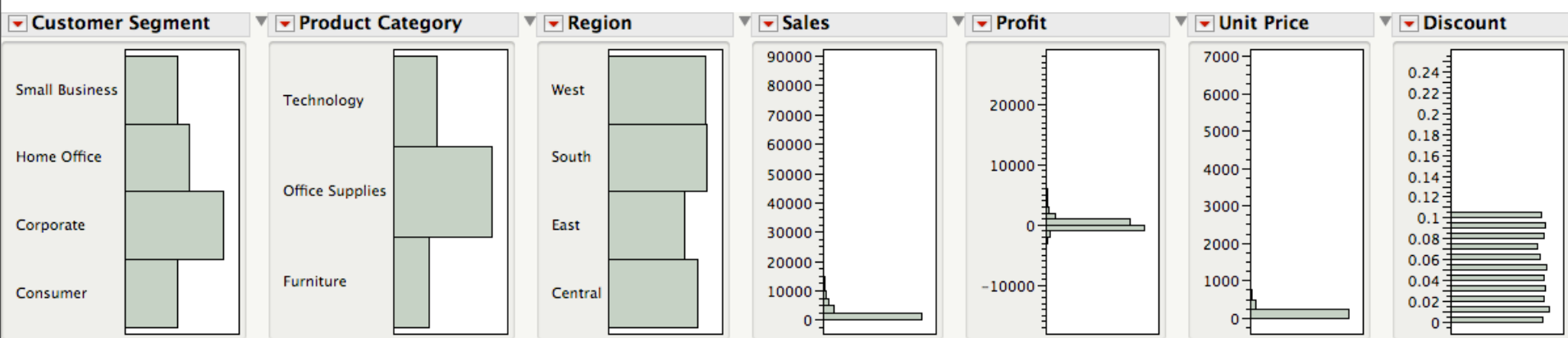


Filtering out extraneous data removes distractions from the data under investigation.

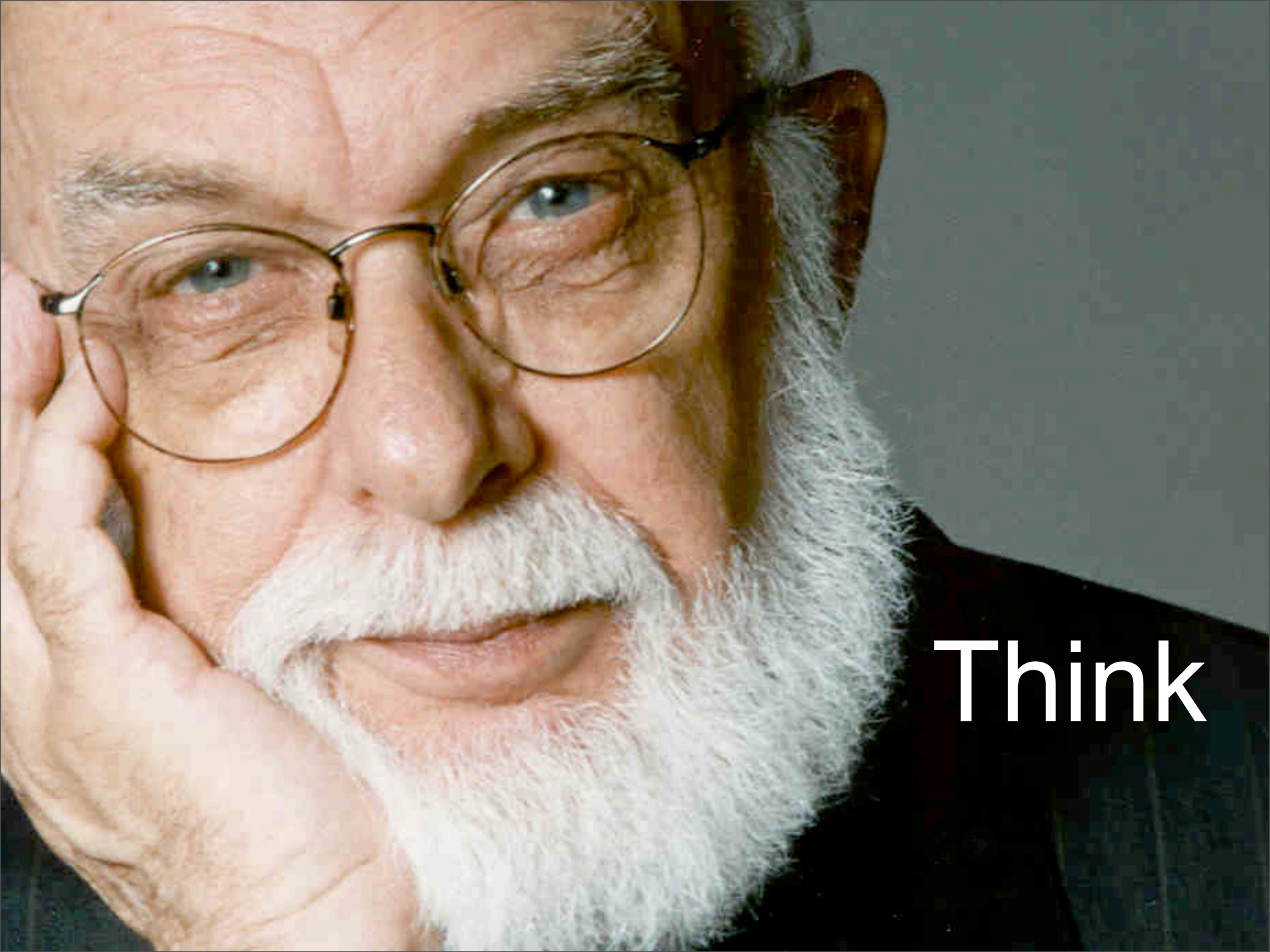
Overview first, zoom and filter, then **details-on-demand**



Visual data analysis relies mostly on the shape of the data to provide needed insights, but there are still times when you need to see the details behind the shape of the data. Having a means to easily see the details when you need them, without having them in the way when you don't works best.



When beginning to examine a new data set, I often begin with a set of frequency distributions for each of the primary variable, all arranged as histograms on the screen at once. This gives me a good overview of the data set. With JMP, I can select and highlight individual items, such as “home office” sales (see middle example), and see where home office sales fit in relation to each of the other variables because they are automatically highlighted in every histogram. This allows me to see that home office sales appear to be distributed in even proportions among the variables, except that there is a slightly greater proportion associated with sales in the West and a slightly smaller proportion in the Central region. If I wanted to see where losses (profits under zero) fit among the other variables, by selecting and highlighting them (see the bottom example), I can see them highlighted in all the other histograms, which reveals that losses are associated to a greater than normal degree with high discounts.

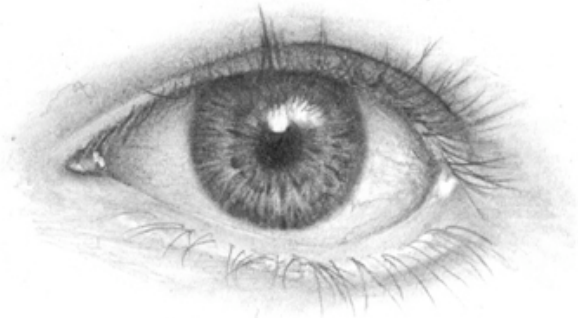


Think

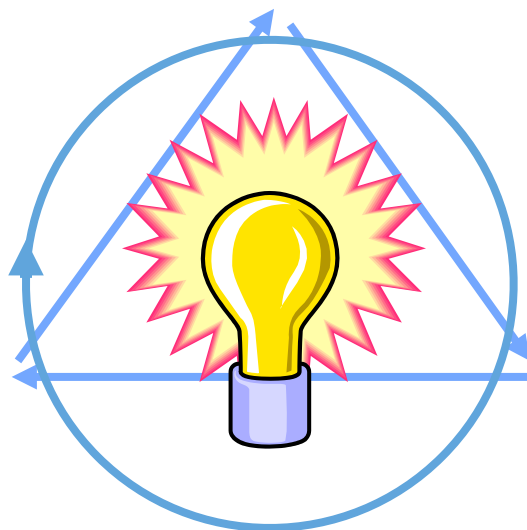
Good tools allow us to think about data, remaining immersed in thought without being distracted by the mechanics of using the software.



Think



See



Modify

Direct dynamic interaction with the properly visualized data allows us to see discover meaningful patterns, trends, and exceptions in the display and to interact with it directly to filter out what we don't need, drill into details, combine multiple variables for comparison, etc., in ways that promote a smooth flow between seeing something, thinking about it, and manipulating it, with no distracting lags in between. This is what I call "visual analysis at the speed of thought."

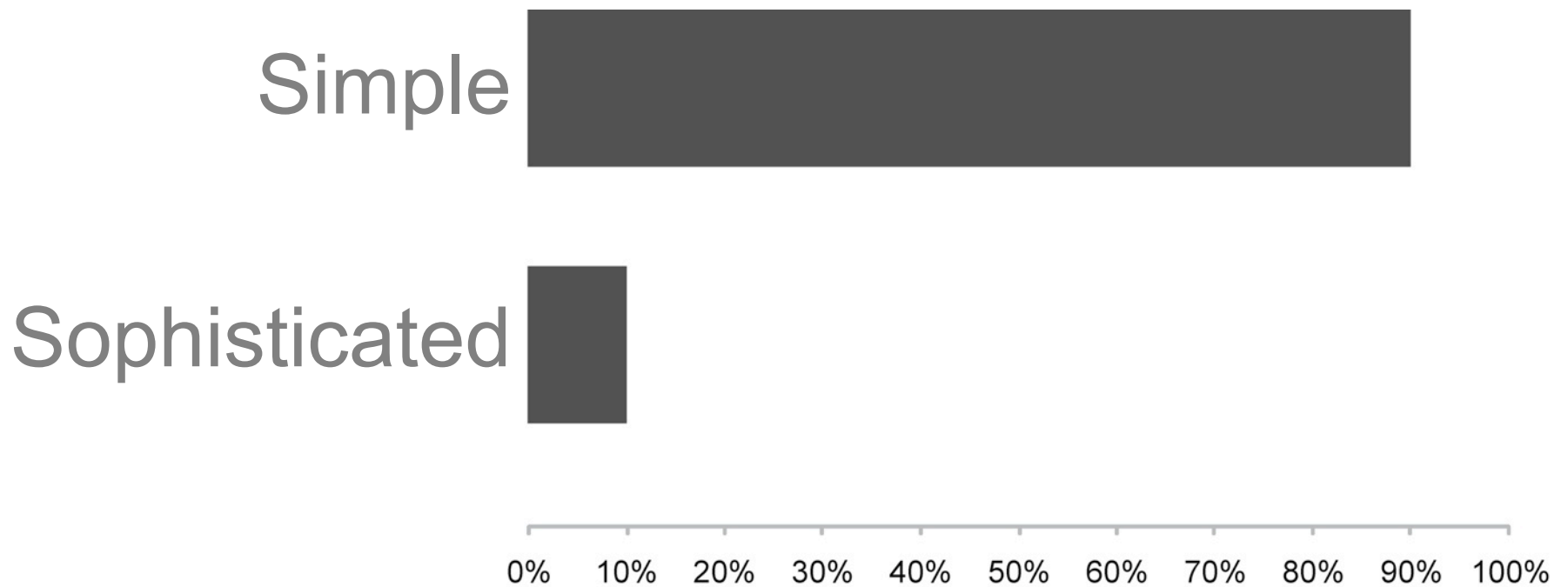
When our tools support this kind of analytical agility, allowing us to remain in the flow of analysis, we're given permission to be skeptical, because asking "Is this really what's going on?" is no longer like standing at the base of Mt. Everest and trying to decide if we feel like taking a hike.

Share the work

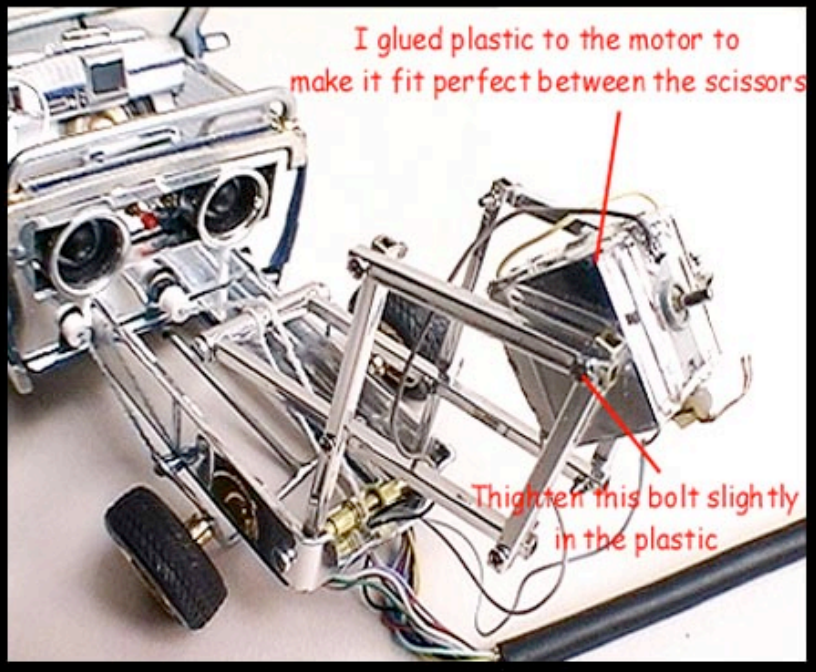
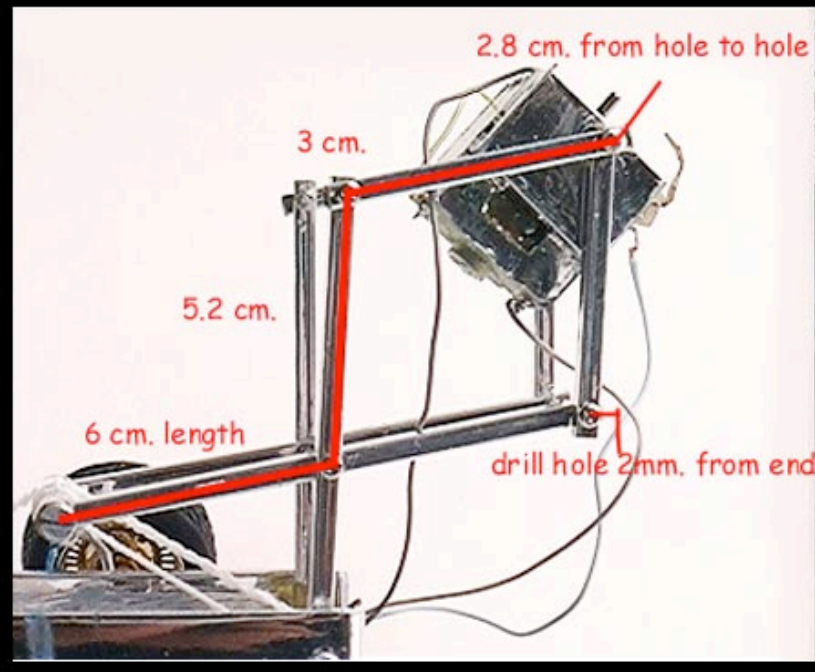
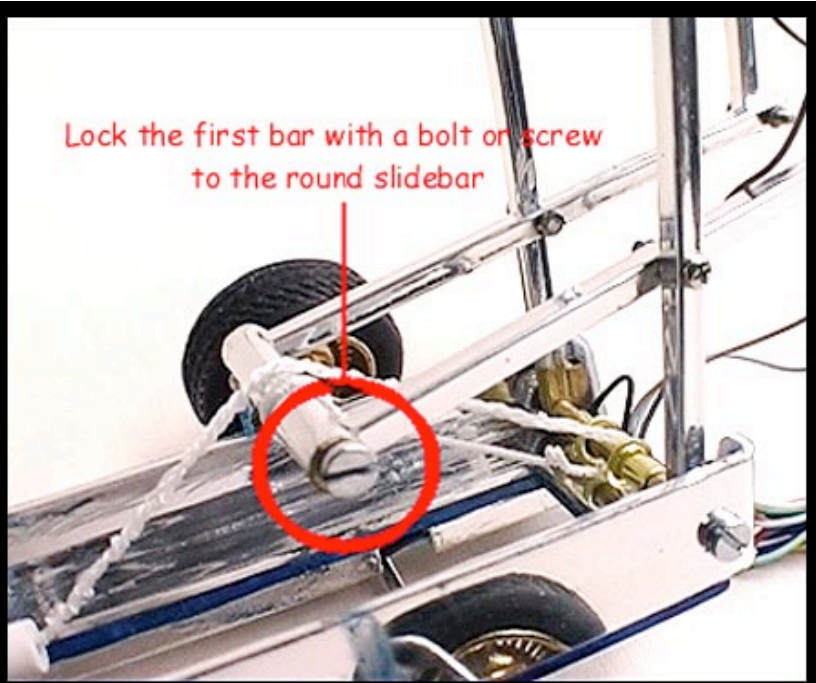


We can support the dawning of the analytics age by sharing the work with people throughout our organizations, including non-statisticians. We all have something to contribute to the analytical quilt. Others might not be able to contribute to the quilt in a way that is as sophisticated as your contribution, but everything that others can do for themselves gives you more time to do the hard stuff.

Statistics required for analytical tasks



In most organizations, whether business, government, academic, or non-profit, 90% of the data analysis that's needed can be done by a broad population of people using relatively simple statistics; only a small portion requires sophisticated statistical analysis. Statisticians' time is wasted if they do what others they easily do for themselves. Statisticians ought to be spending their time doing the complicated stuff.

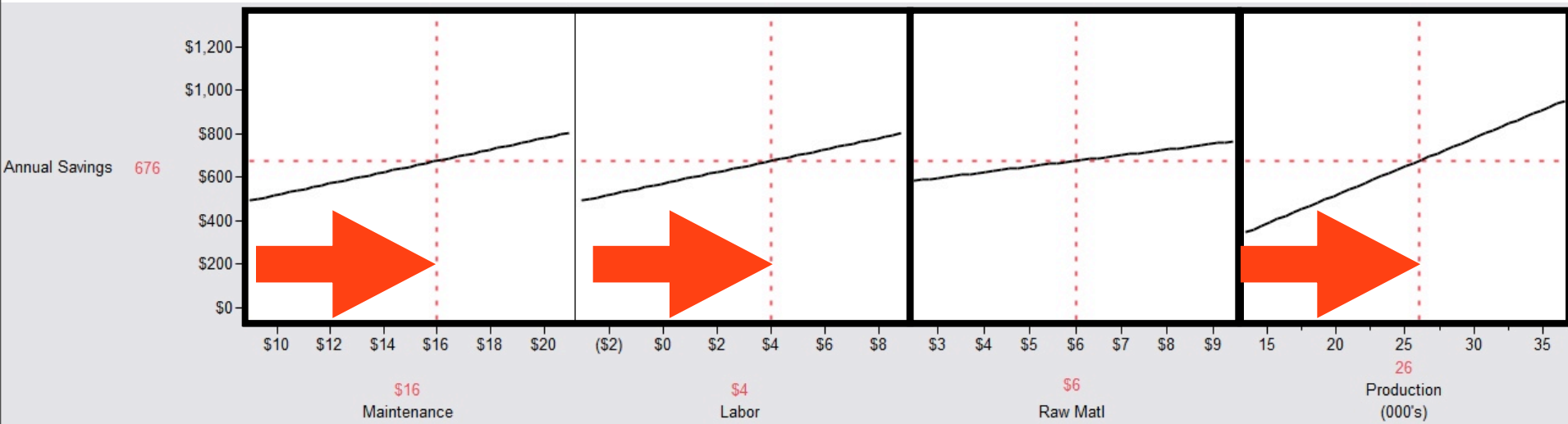


Not everyone can build a statistical model,...



...but almost everyone who is familiar with the data domain can use a model if you build a good one for them.

JMP Profiler

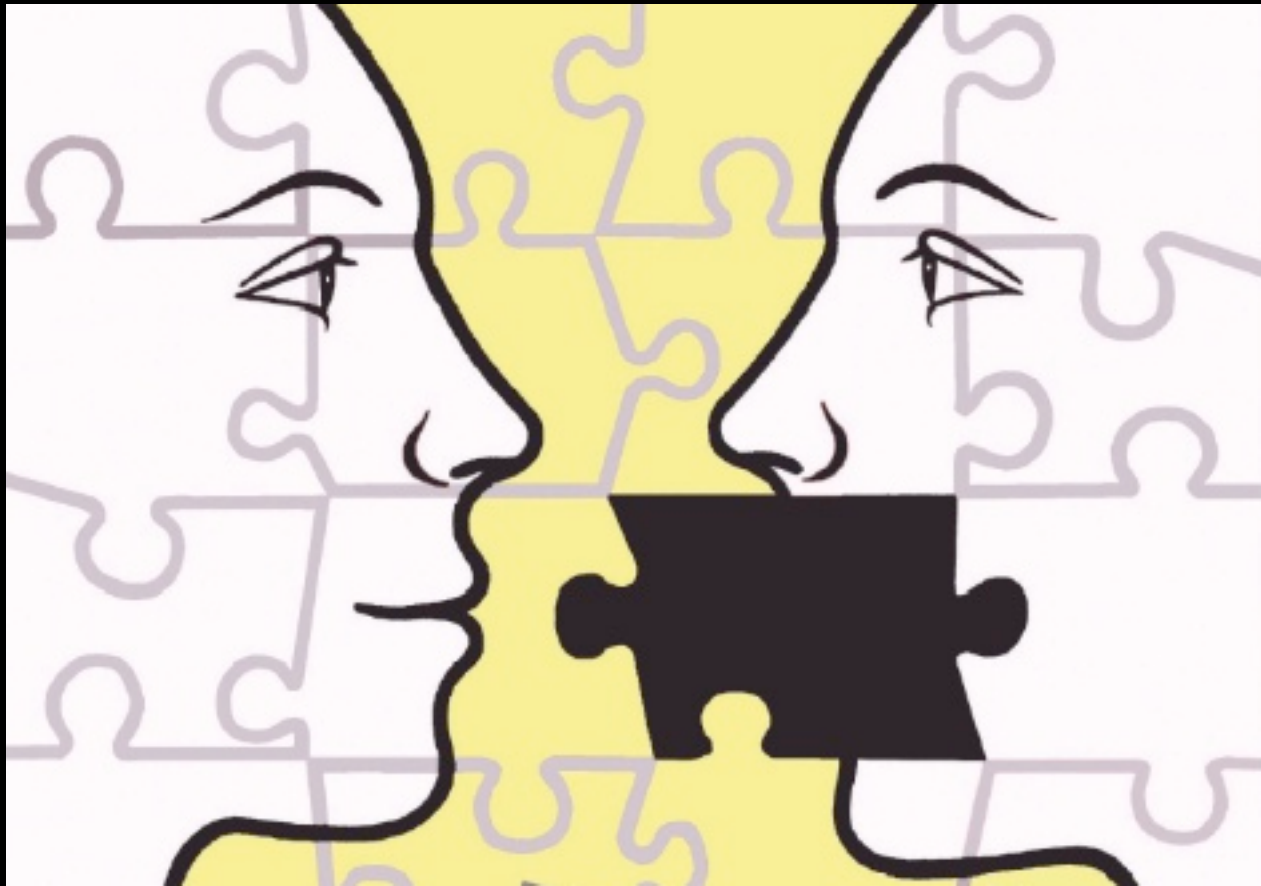


Projected Savings

Total	Maintenance	Labor	Raw Materials	Production
\$600,000	\$15	\$3	\$6	25,000
\$650,000	\$16	\$4	\$6	25,000
\$676,000	\$16	\$4	\$6	26,000

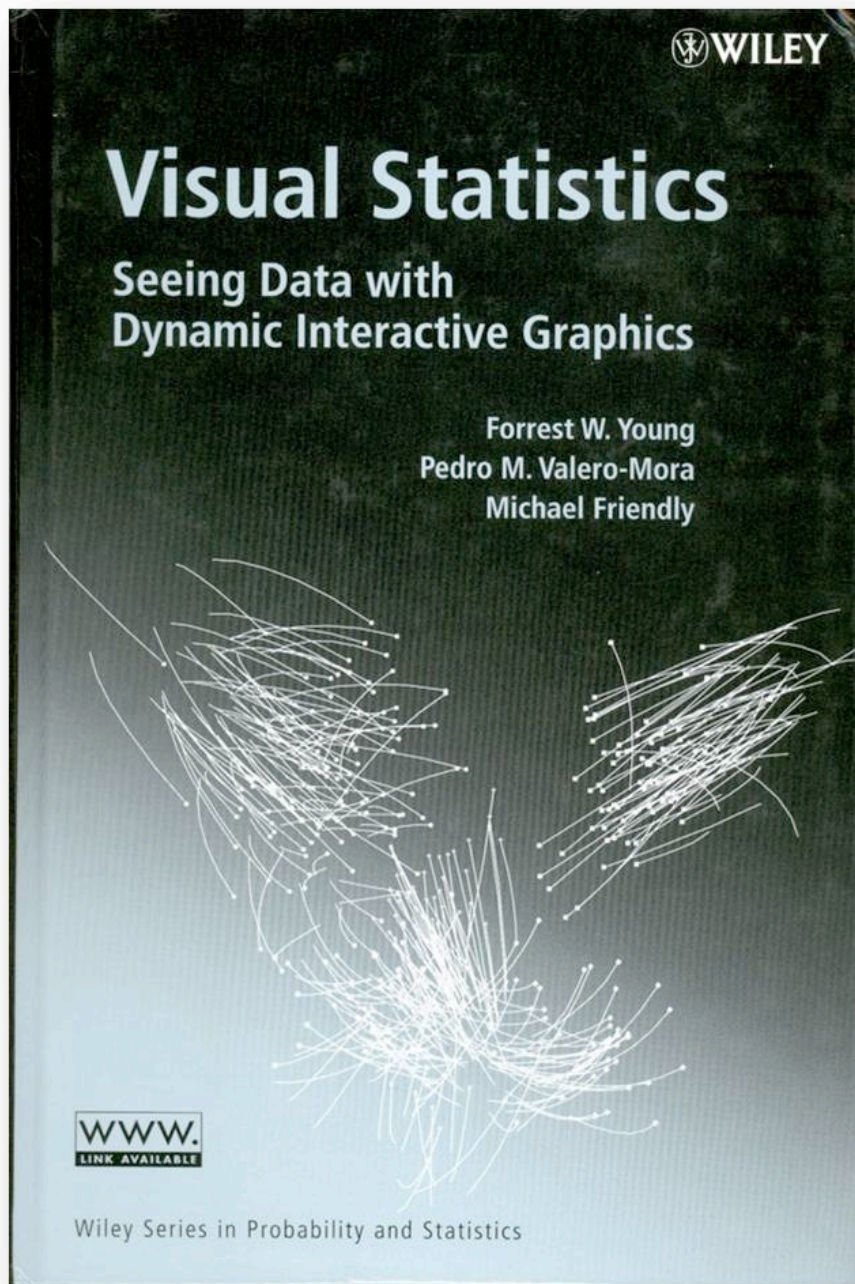
I appreciate how easy it is to build good predictive models using the Profiler in JMP. These models are good partly because they allow people to see the effects among variables as changes to one result in specific changes to others, rather than doing what many models do, which is to hide these interactions in a black box. People who use statistical models need to see these relationships so they can become engaged in thinking about the data, rather than relying on the computer to spit out answers that they don't understand.

Communicate



If you're an expert in any field, you use a vocabulary that is particular to that field; one that works well for interacting with other experts, but not with others. It takes a special effort to express statistical concepts in terms that are familiar to the uninitiated, but it must be made. Your analytical findings might be brilliant, but if you can't explain them to the people who need the information to make decisions, they're worthless.

Analysts must resist the urge, after spending days working to make sense of data, to take those who need the information painfully through the entire journey. You want people to appreciate how hard you worked, but this approach will always backfire. They'll appreciate your work much more if you can tell them what they need to know, and nothing more, in terms that they can understand. If you fully understand your findings, it is in your power to express them in simple terms.



“The spreadplot...for the initial model, (GPE)(M) is shown...This model fits very poorly, of course ($G^2 = 107$, $df = 7$, $p < 0.001$). The G^2 measure is a badness-of-fit measure. Low values are good, high values are bad. The empty model, reported here, has a very large value of G^2 , meaning the fit is very poor, which, of course, it must be, since it has no terms. The hypothesis test, when rejected, as is the case here, indicates that the model does not fit the data.”

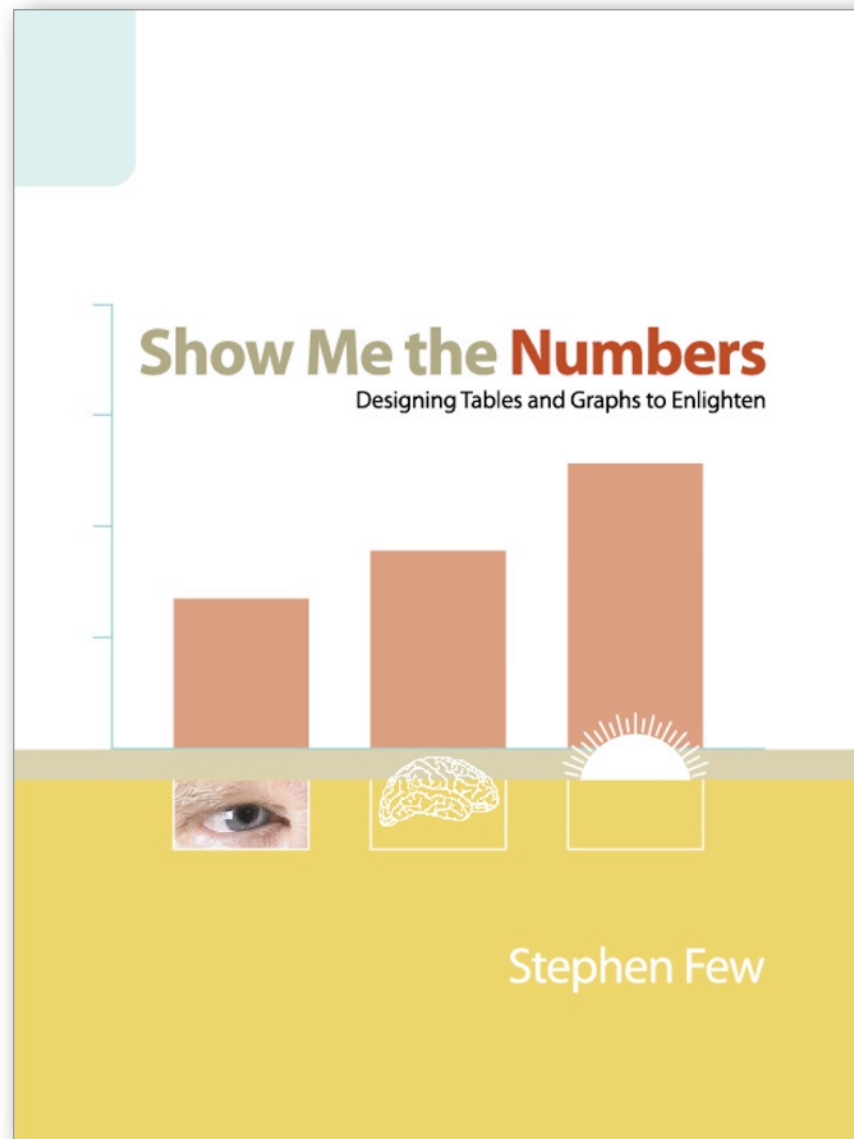
(Excerpt from Chapter 2)

Although this book is quite good for statisticians, the authors thought they wrote it in a way that non-statisticians could understand.

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 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 were too immersed in an academic statistical mindset to recognize when they were 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.

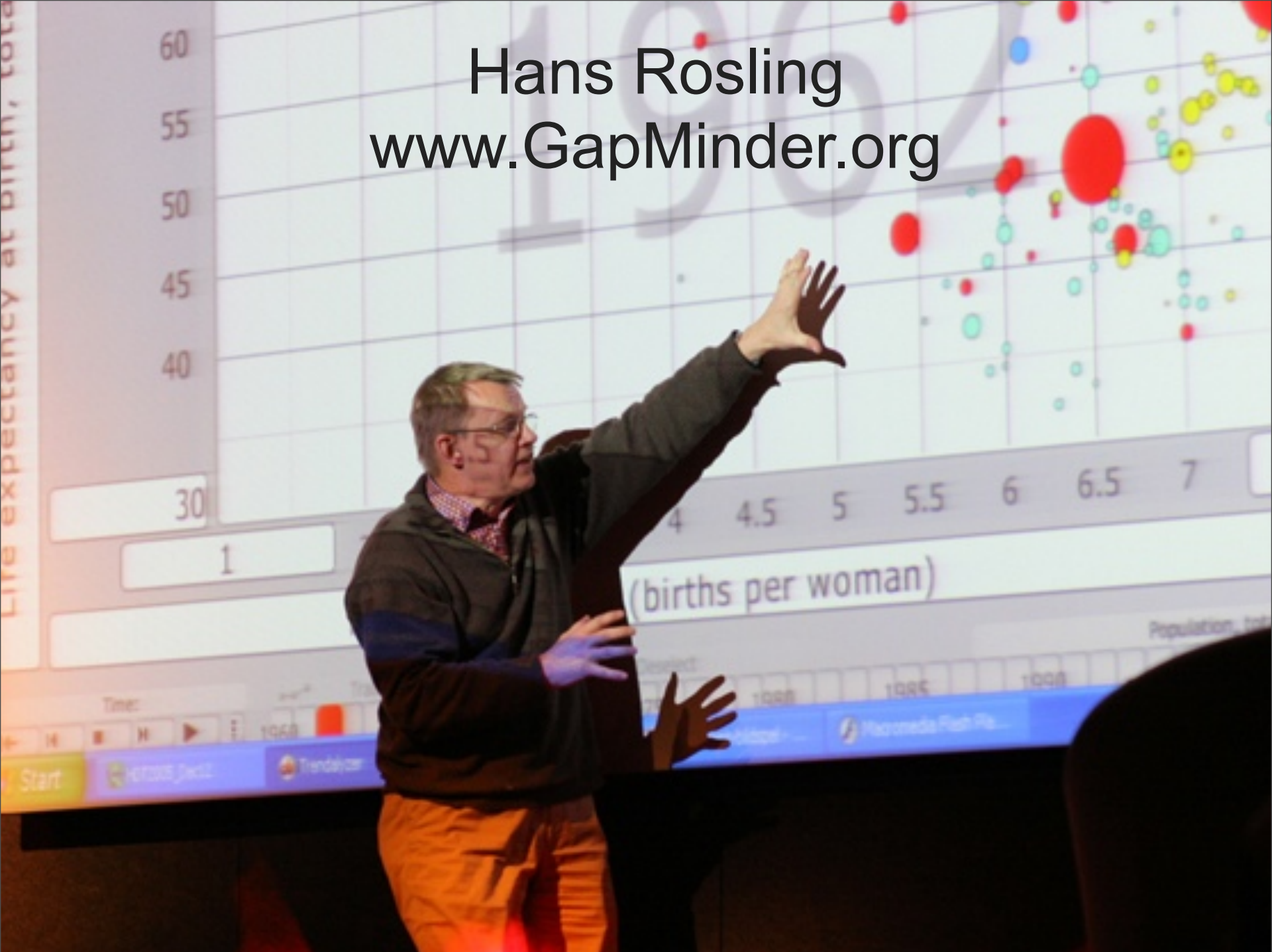
Communicate with pictures.



Pictures of data often serve as an effective way to communicate stories that would be difficult to tell with words. My book *Show Me the Numbers* teaches the use of simple tables and graphs to communicate quantitative data simply and clearly, in a way that everyone can understand.

Hans Rosling

www.GapMinder.org



Few people tell quantitative stories as compellingly as Hans Rosling. When he spoke at the TED Conference (Technology, Entertainment, and Design) for the first time in 2006, it was perhaps the first time in history that a large audience of people found themselves on the edge of their seats watching a bubble plot.

Wisdom

Knowledge

Information

Information cannot speak for itself. It needs our help. It relies on us to give it a voice. When we do, information can tell its story, and will thus become knowledge. The ultimate goal, however, isn't knowledge; it is wisdom. Knowledge becomes wisdom when it is used to do something good. Only when we use what we know to make the world a better place has information served its purpose and we have done our job.

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.



*O perpetual revolution of configured stars,
O perpetual recurrence of determined seasons,
O world of spring and autumn, birth and dying!
The endless cycle of idea and action,
Endless invention, endless experiment,
Brings knowledge of motion, but not of stillness;
Knowledge of speech, but not of silence;
Knowledge of words, and ignorance of The Word.
All our knowledge brings us nearer to our ignorance,
All our ignorance brings us nearer to death,
But nearness to death no nearer to God.
Where is the Life we have lost in living?
Where is the wisdom we have lost in knowledge?
Where is the knowledge we have lost in information?*

Excerpt from *The Rock*, 1930, T.S. Elliot

[Image source: www.irishastronomy.org]



*O perpetual revolution of configured stars,
O perpetual recurrence of determined seasons,
O world of spring and autumn, birth and dying!
The endless cycle of idea and action,
Endless invention, endless experiment,
Brings knowledge of motion, but not of stillness;
Knowledge of speech, but not of silence;
Knowledge of words, and ignorance of The Word.
All our knowledge brings us nearer to our ignorance,
All our ignorance brings us nearer to death,
But nearness to death no nearer to God.
Where is the Life we have lost in living?
Where is the wisdom we have lost in knowledge?
Where is the knowledge we have lost in information?*

Excerpt from *The Rock*, 1930, T.S. Elliot

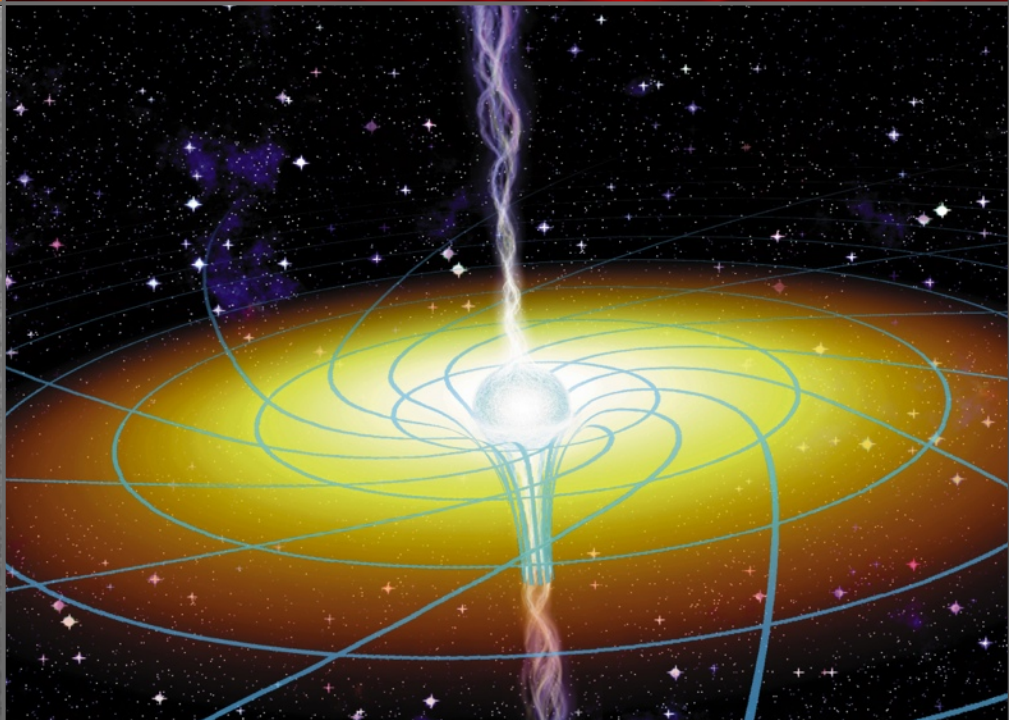
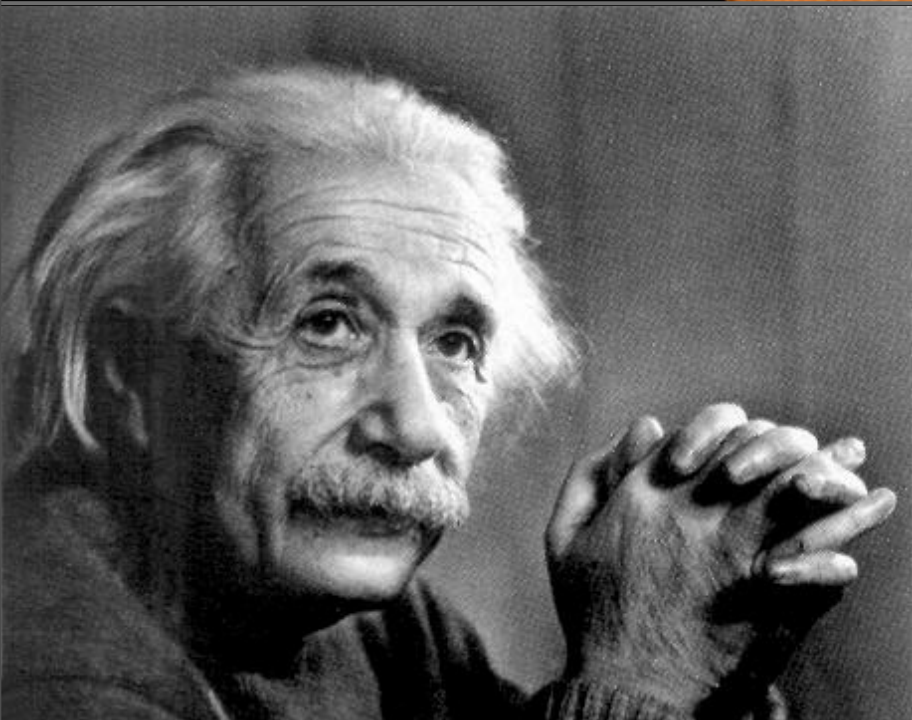
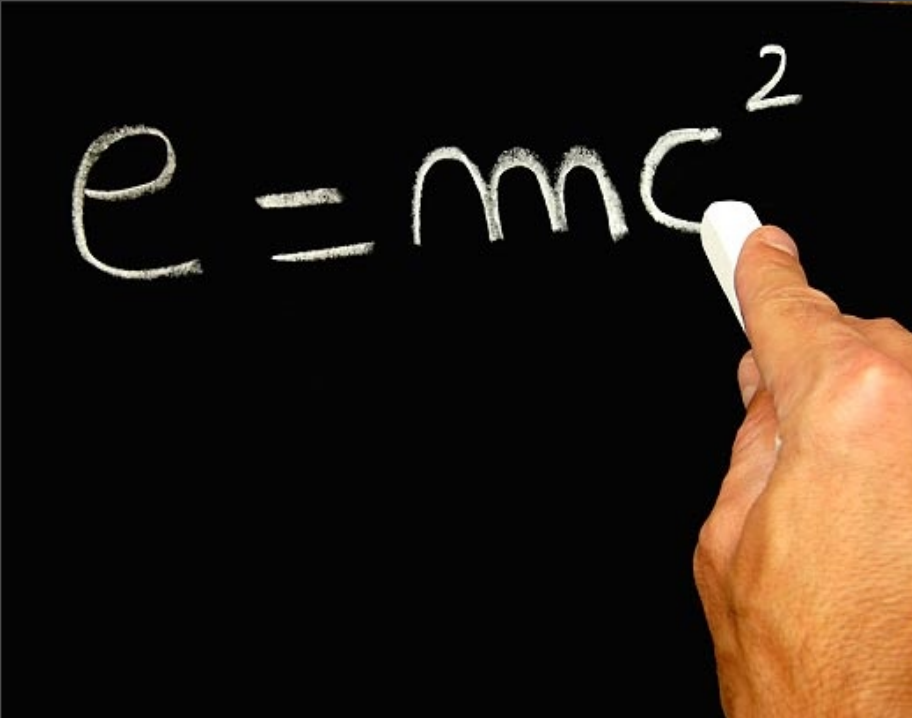
[Image source: www.trekvisual.com]



*O perpetual revolution of configured stars,
O perpetual recurrence of determined seasons,
O world of spring and autumn, birth and dying!
The endless cycle of idea and action,
Endless invention, endless experiment,
Brings knowledge of motion, but not of stillness;
Knowledge of speech, but not of silence;
Knowledge of words, and ignorance of The Word.
All our knowledge brings us nearer to our ignorance,
All our ignorance brings us nearer to death,
But nearness to death no nearer to God.
Where is the Life we have lost in living?
Where is the wisdom we have lost in knowledge?
Where is the knowledge we have lost in information?*

Excerpt from *The Rock*, 1930, T.S. Elliot

[Image source: www.i.pbase.com]



*O perpetual revolution of configured stars,
O perpetual recurrence of determined seasons,
O world of spring and autumn, birth and dying!
**The endless cycle of idea and action,
Endless invention, endless experiment,**
Brings knowledge of motion, but not of stillness;
Knowledge of speech, but not of silence;
Knowledge of words, and ignorance of The Word.
All our knowledge brings us nearer to our ignorance,
All our ignorance brings us nearer to death,
But nearness to death no nearer to God.
Where is the Life we have lost in living?
Where is the wisdom we have lost in knowledge?
Where is the knowledge we have lost in information?*

Excerpt from *The Rock*, 1930, T.S. Elliot

[Image source: www.]

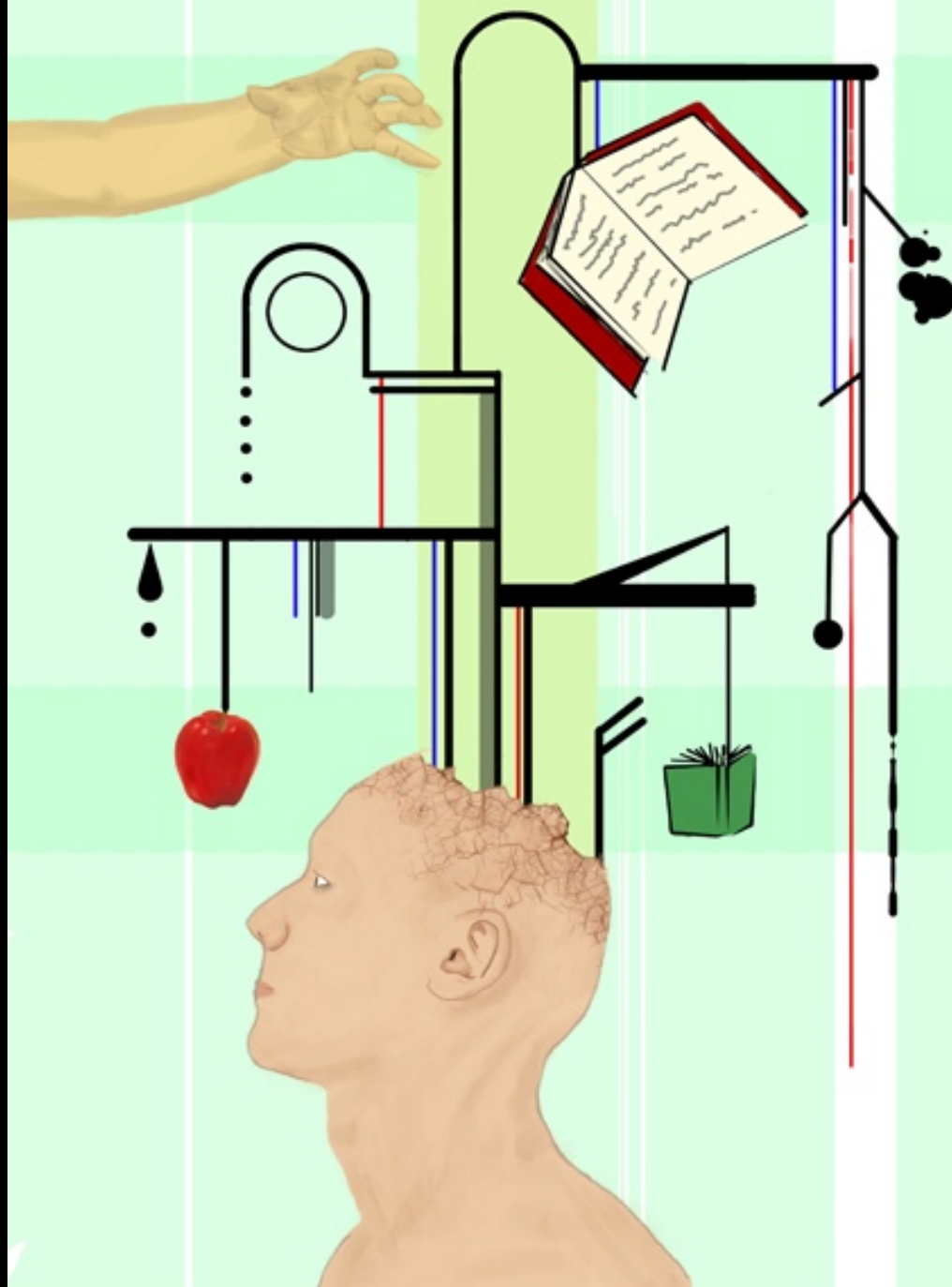


Truth

*O perpetual revolution of configured stars,
O perpetual recurrence of determined seasons,
O world of spring and autumn, birth and dying!
The endless cycle of idea and action,
Endless invention, endless experiment,
**Brings knowledge of motion, but not of stillness;
Knowledge of speech, but not of silence;
Knowledge of words, and ignorance of The Word.**
All our knowledge brings us nearer to our ignorance,
All our ignorance brings us nearer to death,
But nearness to death no nearer to God.
Where is the Life we have lost in living?
Where is the wisdom we have lost in knowledge?
Where is the knowledge we have lost in information?*

Excerpt from *The Rock*, 1930, T.S. Elliot

[Image source: www.shepherdpics.com]



*O perpetual revolution of configured stars,
O perpetual recurrence of determined seasons,
O world of spring and autumn, birth and dying!
The endless cycle of idea and action,
Endless invention, endless experiment,
Brings knowledge of motion, but not of stillness;
Knowledge of speech, but not of silence;
Knowledge of words, and ignorance of The Word.
**All our knowledge brings us nearer to our ignorance,
All our ignorance brings us nearer to death,
But nearness to death no nearer to God.**
Where is the Life we have lost in living?
Where is the wisdom we have lost in knowledge?
Where is the knowledge we have lost in information?*

Excerpt from *The Rock*, 1930, T.S. Elliot

[Image source: www.i163.photobucket.com]

The value of information depends on how it's used.



Use it wisely.

*O perpetual revolution of configured stars,
O perpetual recurrence of determined seasons,
O world of spring and autumn, birth and dying!
The endless cycle of idea and action,
Endless invention, endless experiment,
Brings knowledge of motion, but not of stillness;
Knowledge of speech, but not of silence;
Knowledge of words, and ignorance of The Word.
All our knowledge brings us nearer to our ignorance,
All our ignorance brings us nearer to death,
But nearness to death no nearer to God.
**Where is the Life we have lost in living?
Where is the wisdom we have lost in knowledge?
Where is the knowledge we have lost in information?***

Excerpt from *The Rock*, 1930, T.S. Elliot

[Image source: www.jamin.org]