

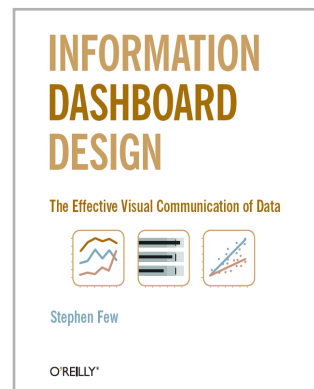
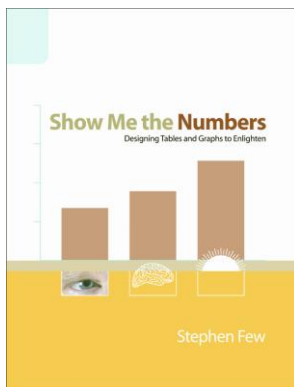
Telling Compelling Stories with Numbers



Stephen Few
of
Perceptual Edge

Data visualization for enlightening communication.

Stephen Few, Principal, Perceptual Edge
sfew@perceptualedge.com
(510) 558-7400



So much data; so little understanding

*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

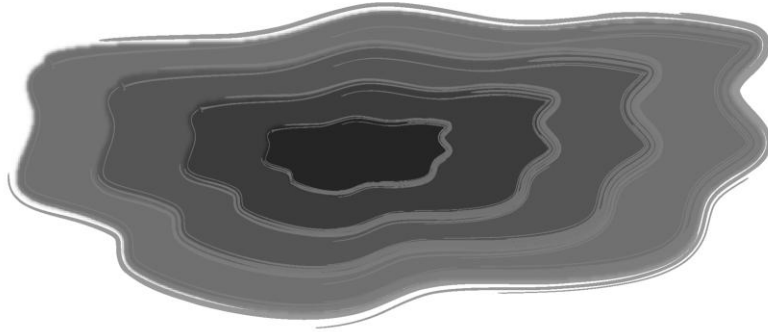
The amount of information that is available to businesses has increased dramatically in the last few years, but the ability to make use of it has increased little, if any.

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.

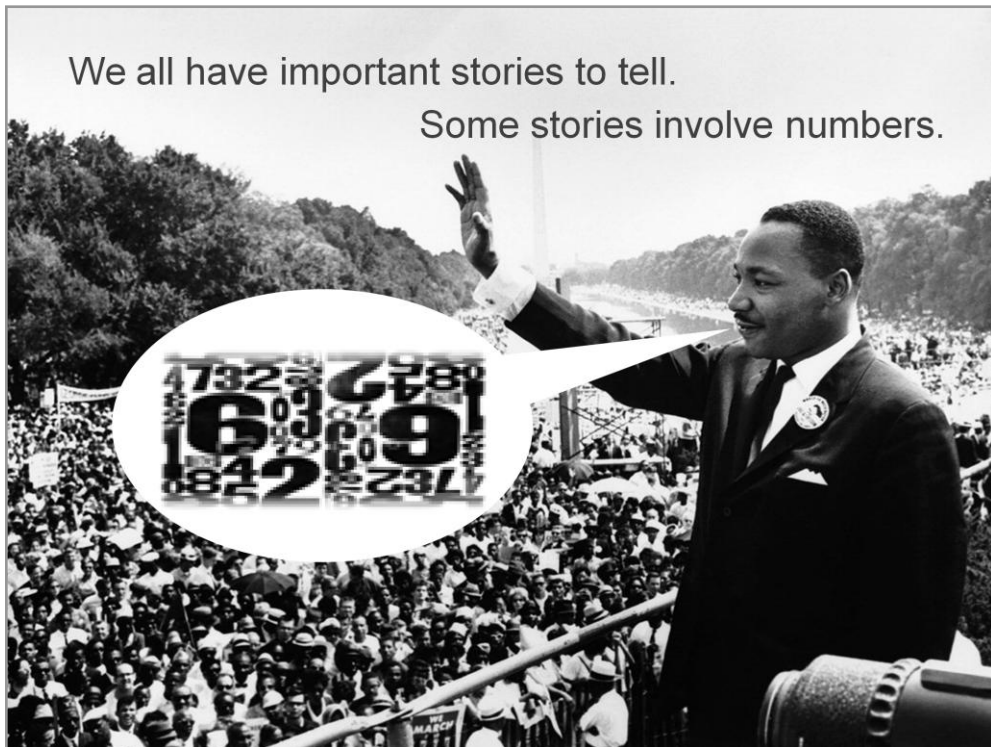
(*Turning Numbers into Knowledge*, Jonathan G. Koomey, 2001, Analytics Press: Oakland, CA page 5, quoting Clifford Stoll)

Most of us 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 analyze it.

Floods of data collect in ever-expanding,
stagnant pools.

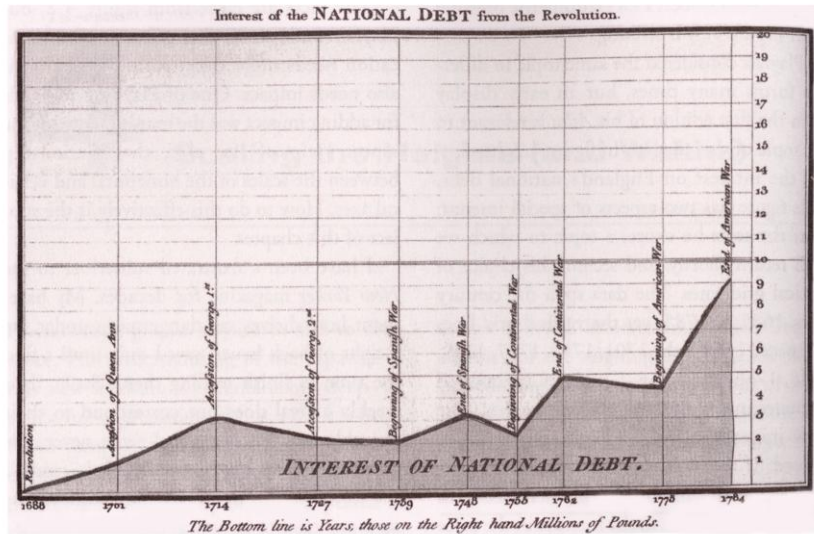


We live in the so-called information age. We have far more data than we've ever had, but our ability to make good use of it hasn't caught up. Information is useless until we understand what it means and can clearly communicate that meaning to those who need it, those whose decisions affect our world.



I work with people and organizations of all types. If they've requested my help, they have one thing in common—they have important quantitative information to understand and present, and realize they could be doing it much better.

We've been telling them with graphs for quite awhile.



William Playfair – 1786

In 1786, a roguish Scot – William Playfair – published a small atlas that introduced or greatly improved most of the quantitative graphs that we use today. Prior to this, graphs of quantitative data were little known.

But most graphs today communicate poorly.

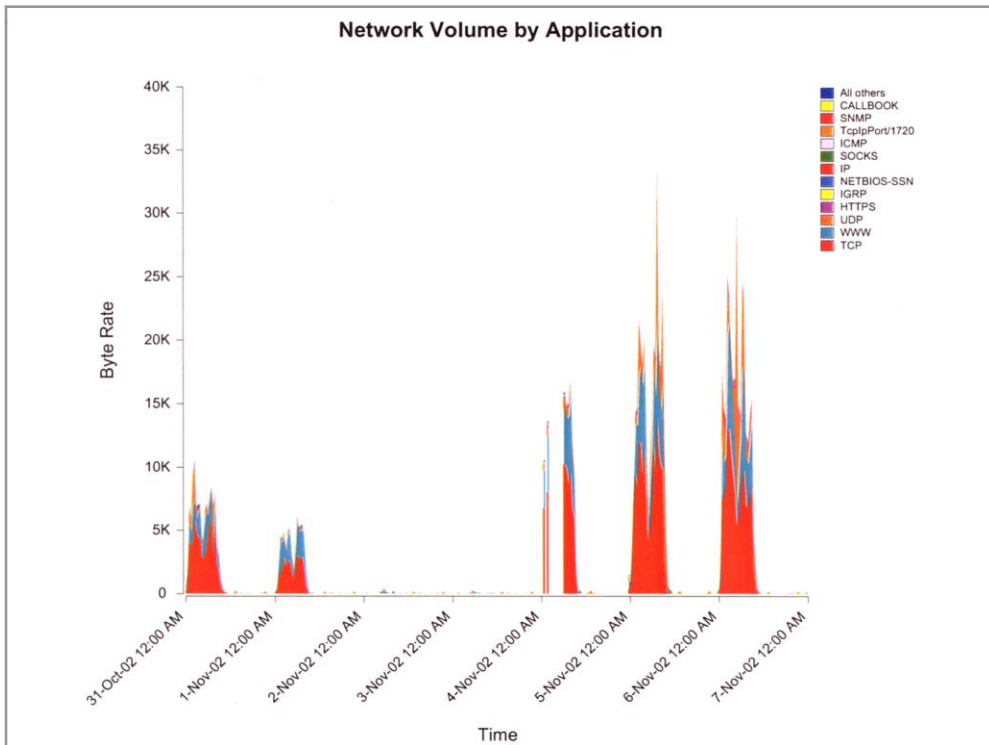


Finally...

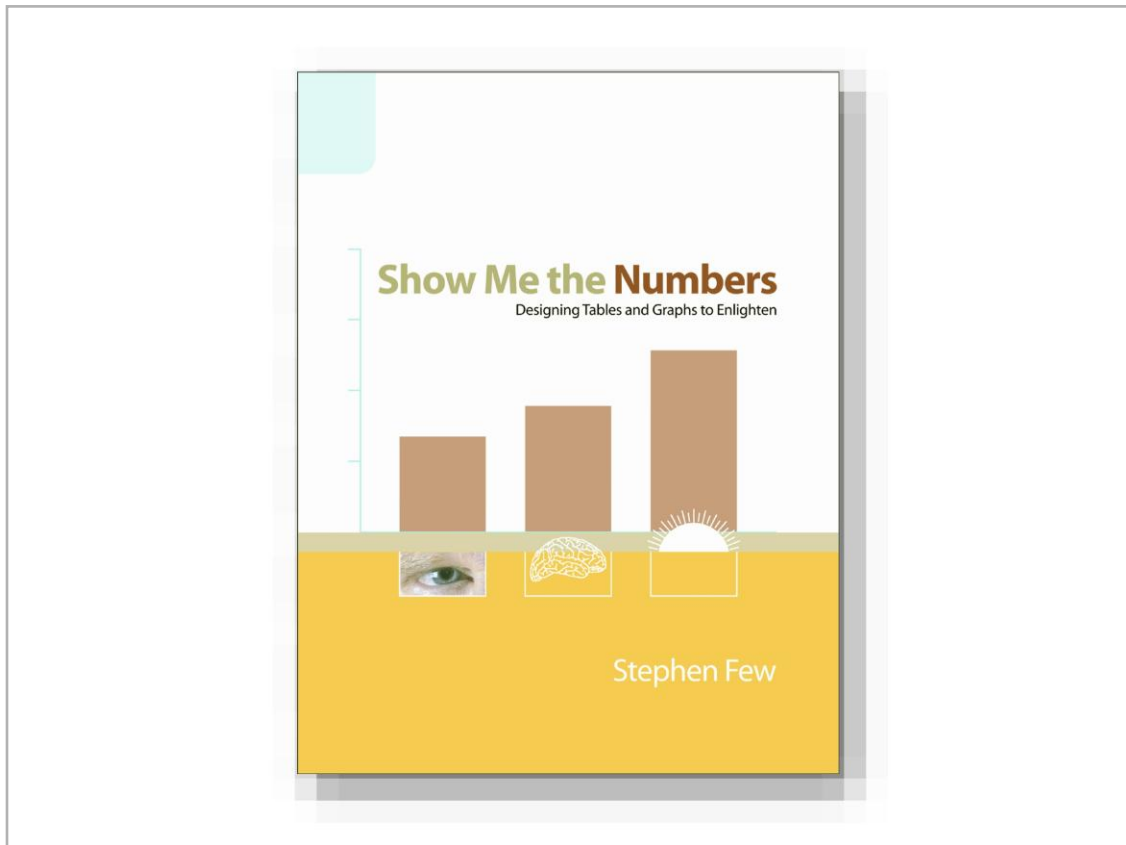
**Effective network
monitoring has arrived!**

- **Near real-time**
- **Phenomenally user-friendly**
- **Instant insight → effective response**

You've been invited to another of the many meetings that you're required to attend. You're one of many managers in the Information Technology department. Like most meetings, this one begins with the light of a projector suddenly illuminating a screen. Bursting with excitement, the speaker announces that you and everyone else in the room will now receive a daily report that will inform you how the network is being utilized, and then the graph on the next slide appears.



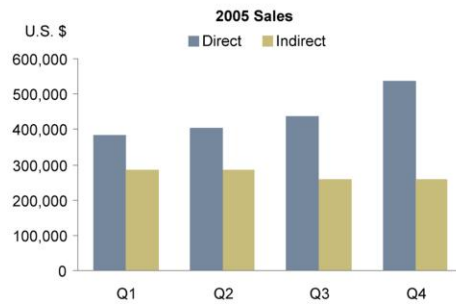
You stare at this graph very intently, trying your best to keep any hint of confusion from crossing your face. From your peripheral vision you can see that the CIO (Chief Information Officer) is smiling broadly and nodding with obvious understanding. You and everyone else in the room begins to nod enthusiastically as well. You feel very dumb. What you don't realize is that you are not alone.



I wrote the book, ***Show Me the Numbers: Designing Tables and Graphs to Enlighten*** in 2004 to help people like you respond to the challenges that you face every day when presenting quantitative information.

Quantitative information is usually communicated through tables and graphs.

Department	Expenses			
	Jan	Feb	Variance	Change %
Sales	9,933	9,293	-640	-6%
Marketing	5,385	5,832	+447	+8%
Operations	8,375	7,937	-438	-5%
Total	\$23,693	\$23,062	-\$1,327	-3%



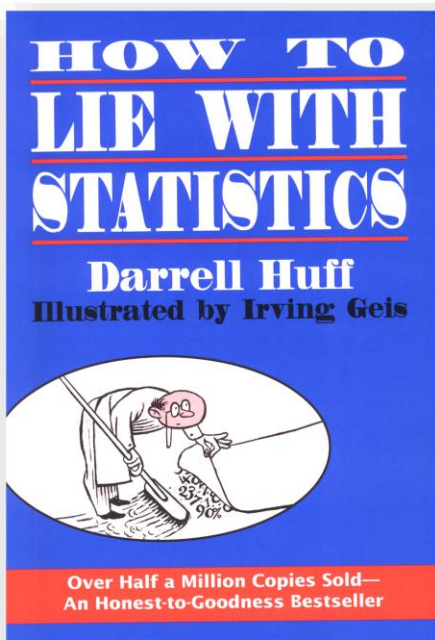
But few communicate effectively. **Why?**

Why? Few people are trained.

Why? Few people recognize the need to be trained.

Why? Few examples of good design exist to expose the problem.

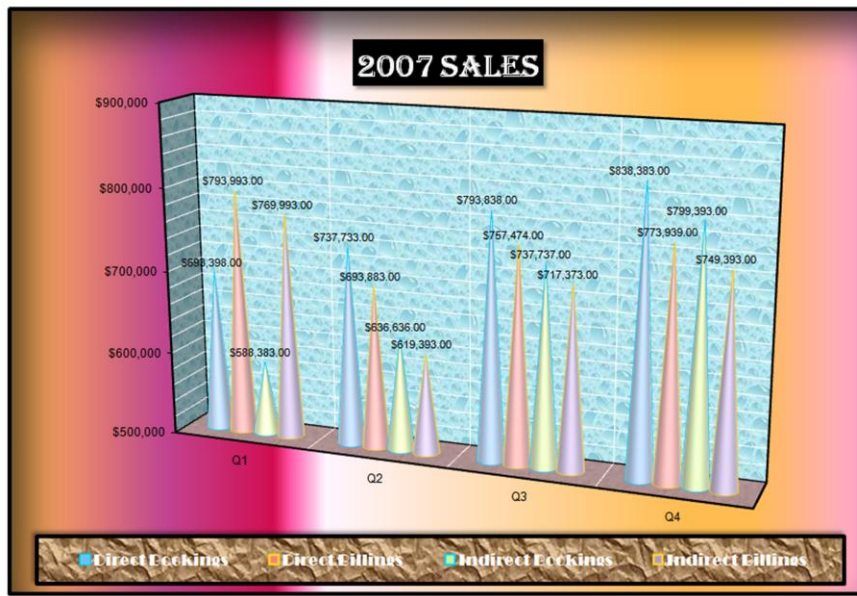
Intentional deceit is no longer our biggest problem.



In 1954, Darrell Huff wrote his best-selling book about how people often intentionally use graphs to spread misinformation, especially in favor of their own products or causes. Today, vastly more misinformation is disseminated unintentionally because people don't know how to use charts to communicate what they intend.

Numbers are commonly obscured, then dressed up to look *sexy*.

Why?

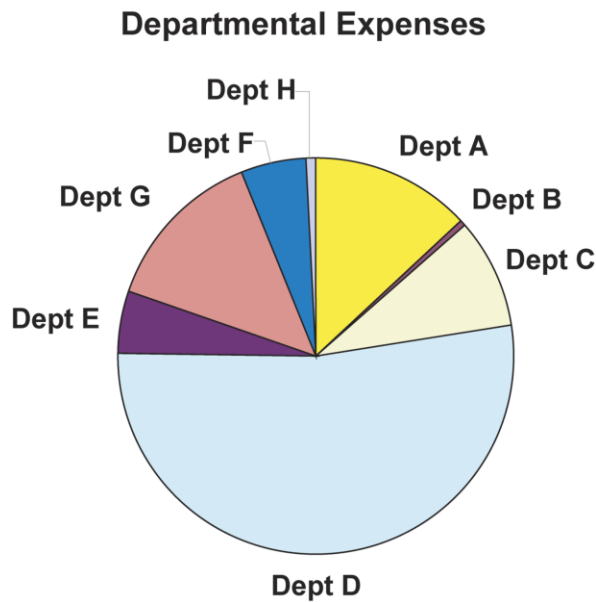


When the PC was introduced, software soon made the arduous task of table and graph creation as easy as 1-2-3 (literally “Lotus 1-2-3”, the software that was the first to legitimize the PC as a viable tool for business). Unfortunately, this improvement in ease and efficiency was not accompanied by instruction in visual design for communication. People today think that if they know how to click with the mouse to create a table or graph, they know how to present data effectively.

“In the two centuries since [the invention of the first graphs], ...charts have become commonplace. With the advent of modern computer tools, creating graphs from data involves trivial effort. In fact, it has probably become too easy. Graphs are often produced without thought for their main purpose: to enlighten and inform the reader.” Jonathan G. Koomey, *Turning Numbers into Knowledge*, Analytics Press, 2001

I can talk about this all day, but the best way to make my point convincingly is to show you.

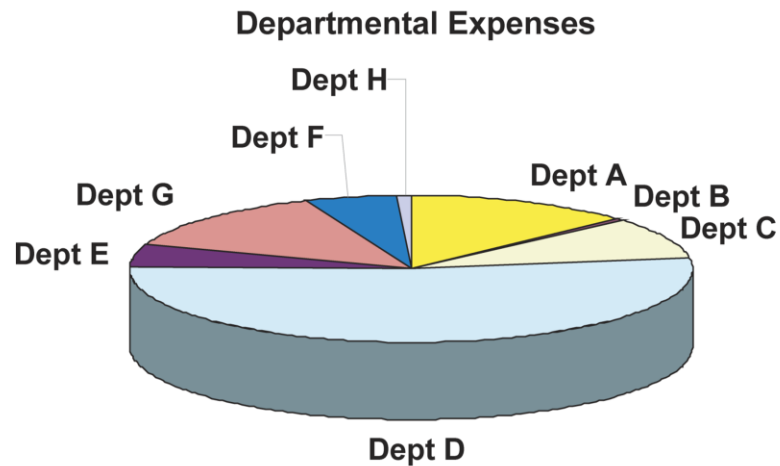
Example #1 – Poor



The purpose of this graph is to display how *Department G* is doing regarding expenses compared to the other departments. Is the message clear?

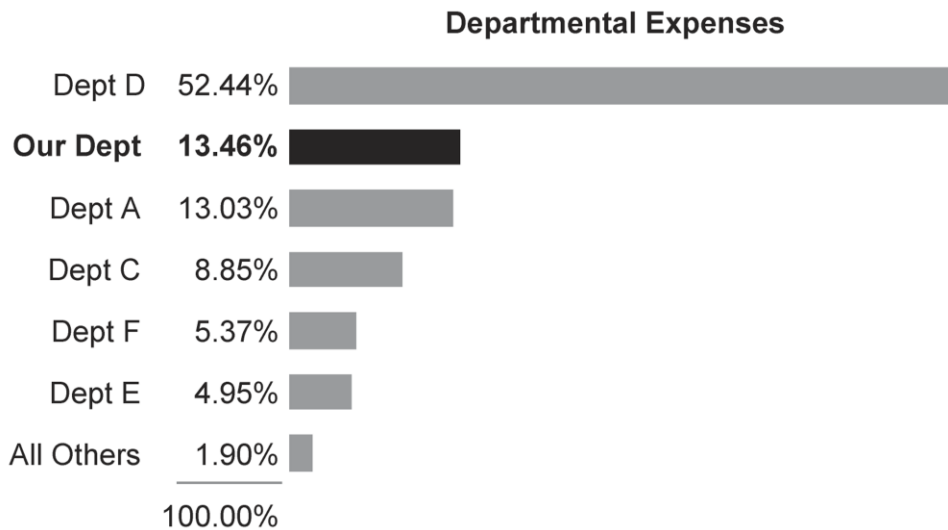
Often, when someone creates a graph that appears inadequate somehow, they try to fix it with sizzle, as in the next slide.

Example #1 – Worse



Does the addition of 3D improve this pie chart? Definitely not. In fact, it actually makes it harder to read.

Example #1 – Good



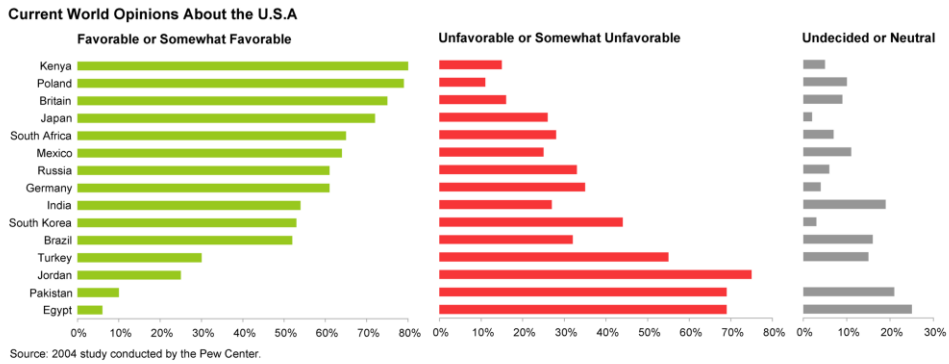
Though it lacks flash and dazzle, this simple bar graph tells the story elegantly.

Example #2 - Poor

Favorable or Unfavorable View of the U.S.	
Brazil: % with somewhat or very favorable opinion of the U.S.:	52%
Brazil: % with somewhat or very unfavorable opinion of the U.S.:	32%
Mexico: % with somewhat or very favorable opinion of the U.S.:	64%
Mexico: % with somewhat or very unfavorable opinion of the U.S.:	25%
Britain: % with somewhat or very favorable opinion of the U.S.:	75%
Britain: % with somewhat or very unfavorable opinion of the U.S.:	16%
Germany: % with somewhat or very favorable opinion of the U.S.:	61%
Germany: % with somewhat or very unfavorable opinion of the U.S.:	35%
Russia: % with somewhat or very favorable opinion of the U.S.:	61%
Russia: % with somewhat or very unfavorable opinion of the U.S.:	33%
Poland: % with somewhat or very favorable opinion of the U.S.:	79%
Poland: % with somewhat or very unfavorable opinion of the U.S.:	11%
South Africa: % with somewhat or very favorable opinion of the U.S.:	65%
South Africa: % with somewhat or very unfavorable opinion of the U.S.:	28%
Kenya: % with somewhat or very favorable opinion of the U.S.:	80%
Kenya: % with somewhat or very unfavorable opinion of the U.S.:	15%
India: % with somewhat or very favorable opinion of the U.S.:	54%
India: % with somewhat or very unfavorable opinion of the U.S.:	27%
Japan: % with somewhat or very favorable opinion of the U.S.:	72%
Japan: % with somewhat or very unfavorable opinion of the U.S.:	26%
South Korea: % with somewhat or very favorable opinion of the U.S.:	53%
South Korea: % with somewhat or very unfavorable opinion of the U.S.:	44%
Egypt: % with somewhat or very favorable opinion of the U.S.:	6%
Egypt: % with somewhat or very unfavorable opinion of the U.S.:	69%
Pakistan: % with somewhat or very favorable opinion of the U.S.:	10%
Pakistan: % with somewhat or very unfavorable opinion of the U.S.:	69%
Turkey: % with somewhat or very favorable opinion of the U.S.:	30%
Turkey: % with somewhat or very unfavorable opinion of the U.S.:	55%
Jordan: % with somewhat or very favorable opinion of the U.S.:	25%
Jordan: % with somewhat or very unfavorable opinion of the U.S.:	75%

I found this table on the Web site for Bill Moyers' public television show "Now". I felt that it provided important information that deserved a better form of presentation. In this case the story could be told much better in visual form.

Example #2 - Good



This series of related graphs tells the story in vivid terms and brings facts to light that might not ever be noticed in the table.

What do tables and graphs help us do?

Think

and

Communicate

Grice's conversational maxims

4 categories:

- Quantity
- Quality
- Relevance
- Manner

Paul Grice was a 20th century philosopher whose work ventured into the realm of linguistics. He is well known for his conversational maxims, which attempt to describe the characteristics of polite conversation.

Every one of these maxims of conversation apply equally well to the communication of quantitative information in the workplace. We'll strive in this workshop to translate these maxims into effective and polite communication via tables and graphs.

Grice's conversational maxims: Quantity

1. Make your contribution to the conversation as informative as necessary.
2. Do not make your contribution to the conversation more informative than necessary.

Grice's conversational maxims: Quality

1. Do not say what you believe to be false.
2. Do not say that for which you lack adequate evidence.

Grice's conversational maxims: Relevance

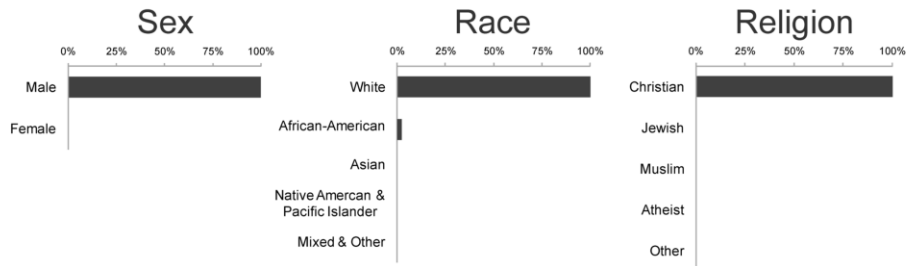
Be relevant (that is, say things related to the current topic of conversation).

Grice's conversational maxims: Manner

1. Avoid obscurity of expression.
2. Avoid ambiguity.
3. Be brief (avoid unnecessary wordiness).
4. Be orderly.

The best displays are often the simplest.

Statistics about U.S. Presidents



Often, the simplest form of display is the most powerful.

Which would you rather look at?



What if you are Jessica's dermatologist?

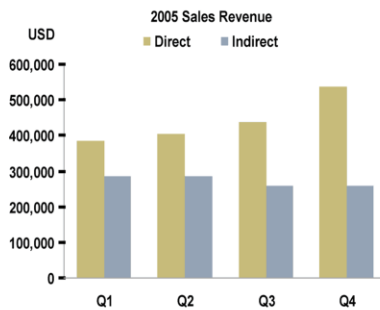
Dressing things up is appropriate for advertising, because the illusion pleases and sells. When you're responsible for discovering the truth and understanding it, makeup only gets in the way.

Data-Ink Ratio

Data Ink

2005 Sales Revenue (USD)

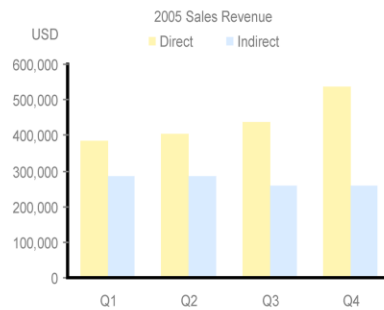
Sales Channel	Q1	Q2	Q3	Q4
Direct	383,383	403,939	437,373	538,583
Indirect	283,733	283,833	257,474	258,474
Total	667,116	687,772	694,847	797,057



Non-Data Ink

2005 Sales Revenue (USD)

Sales Channel	Q1	Q2	Q3	Q4
Direct	383,383	403,939	437,373	538,583
Indirect	283,733	283,833	257,474	258,474
Total	667,116	687,772	694,847	797,057



But besides the data, what else is there? According to Edward Tufte, tables and graphs consist of two types of ink: data ink and non-data ink. He introduced the concept of the “data-ink ratio” in his 1983 classic *The Visual Display of Quantitative Data*. He argued that the ratio of ink used to display data to the total ink should be high. In other words, ink that is used to display anything that isn’t data should be reduced to a minimum.

Steps in the design process

1. Reduce the non-data ink.

- Remove unnecessary non-data ink.
- De-emphasize and regularize the remaining non-data ink.

2. Enhance the data ink.

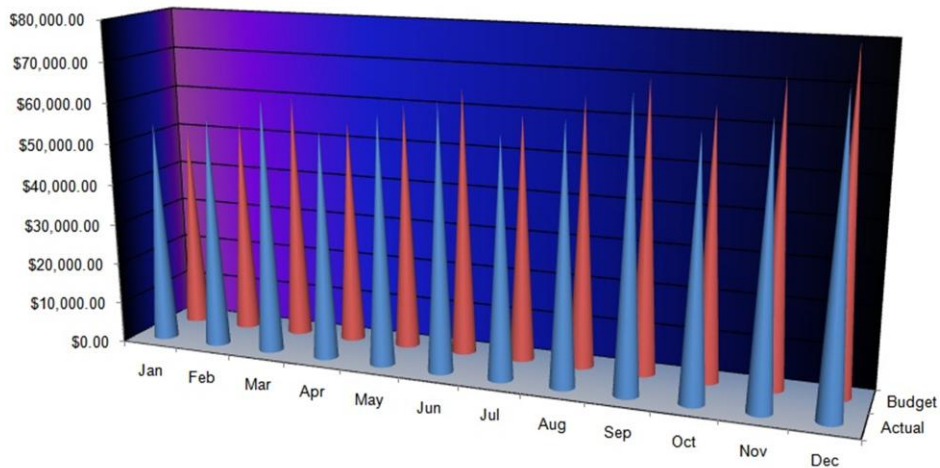
- Remove unnecessary data ink.
- Emphasize the most important data ink.

“In anything at all, perfection is finally attained not when there is no longer anything to add, but when there is no longer anything to take away.” Antoine de St. Exupery

John Maeda, in *The Laws of Simplicity*, offers a maxim about design simplicity, which I have massaged into the following statement:

Simplicity is about eliminating the obvious (and everything else that doesn't support your purpose), and enhancing the meaningful.

Poor graphs can be transformed.

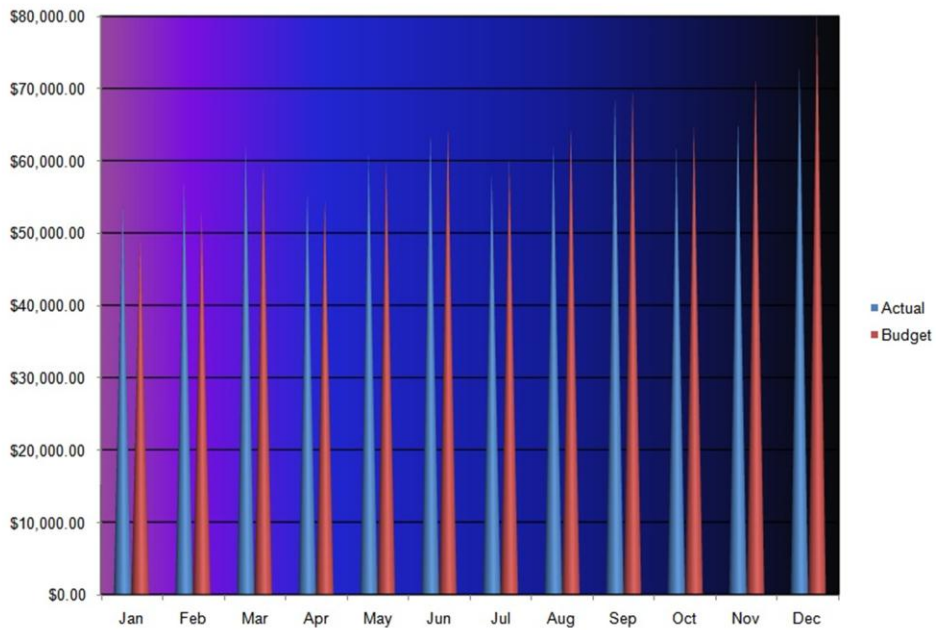


This is the kind of graph that software products, including Excel, encourage us to create. They give us an infinite selection of poorly-designed graphs from which to choose. What we really need, however, is a small selection of graphs that really work.

Using this graph, try to see the pattern of change across the months in actual expenses. Try to determine one of the actual values. Try to compare actual expenses to the budget across time.

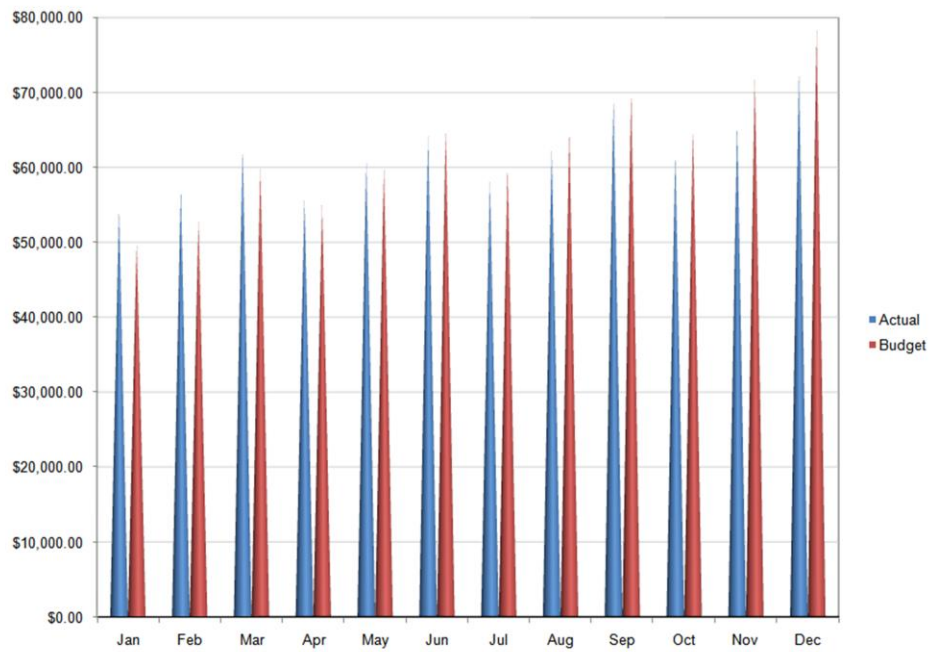
Let's transform this graph into one that communicates.

Poor graphs can be transformed.



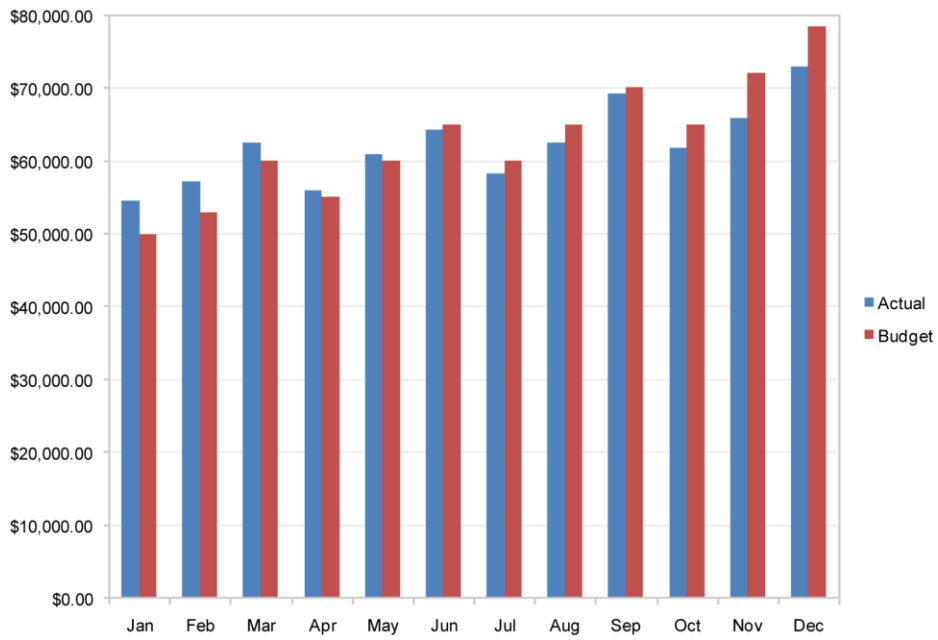
We have now removed the useless 3-D effects and angle, which makes the data easier to read.

Poor graphs can be transformed.

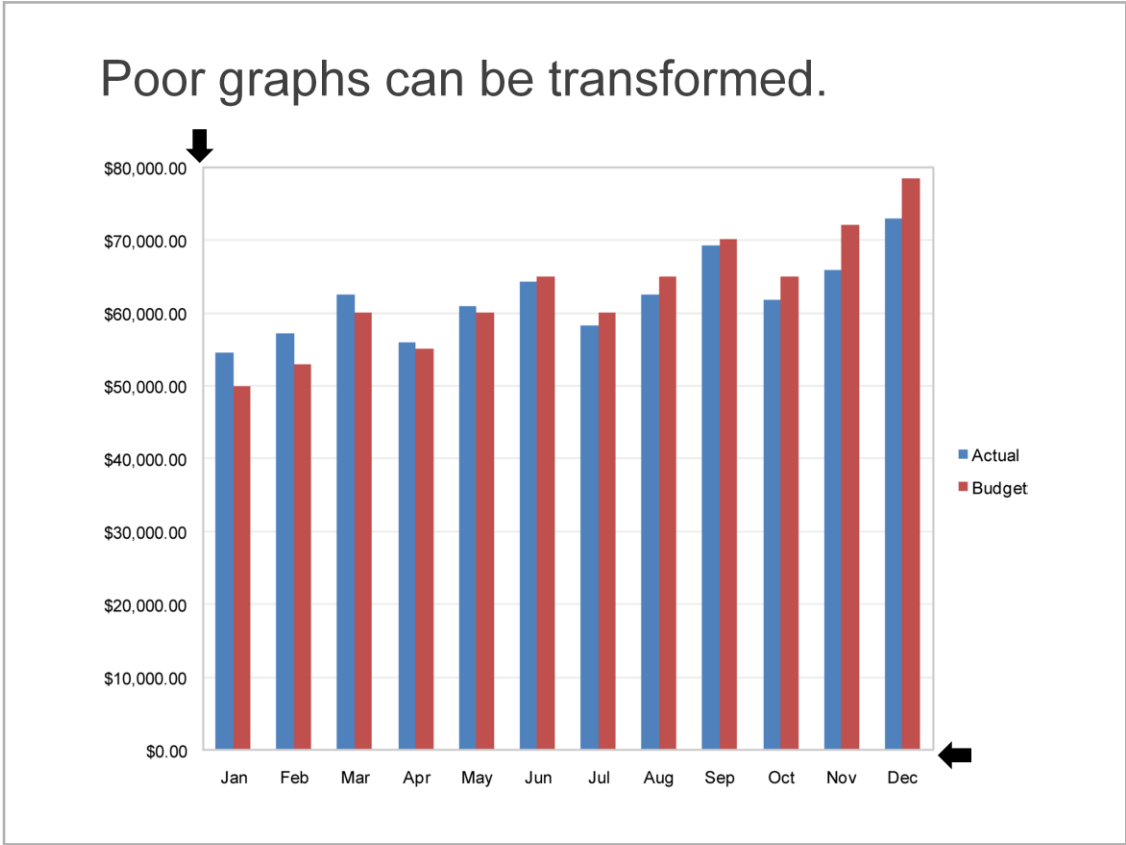


We have now removed the background fill color.

Poor graphs can be transformed.

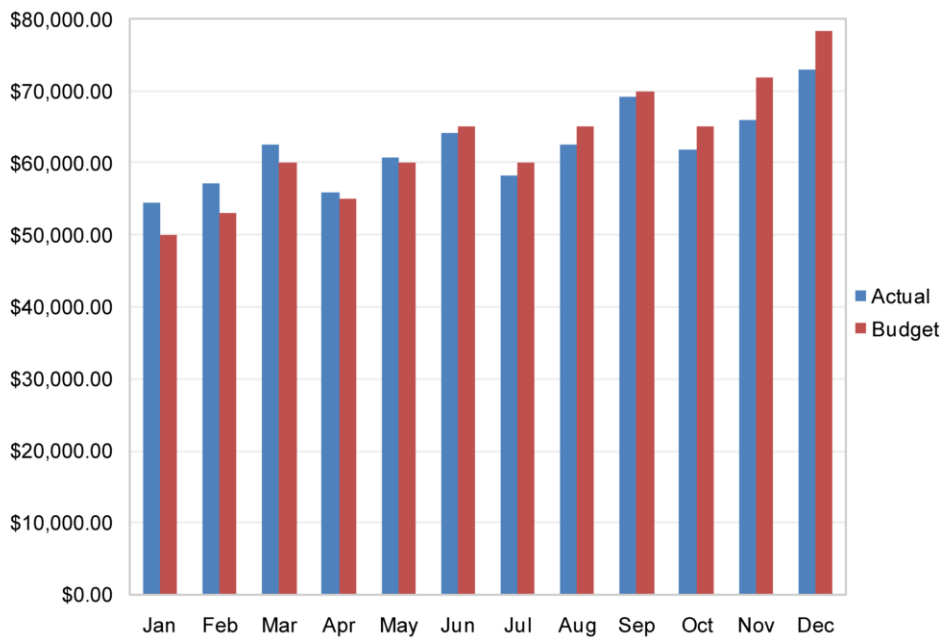


We have now replaced the silly cones with regular bars.



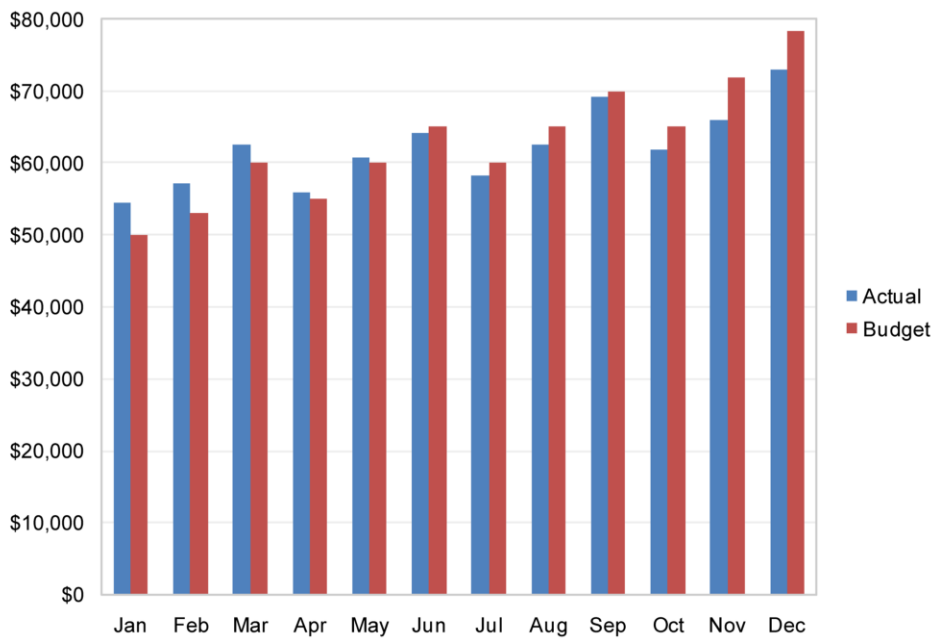
We have now removed the tick marks, which aren't necessary. Tick marks are not needed to separate the months along the X-axis and because horizontal grid lines are being displayed, there is no need for tick marks on the Y-axis either.

Poor graphs can be transformed.

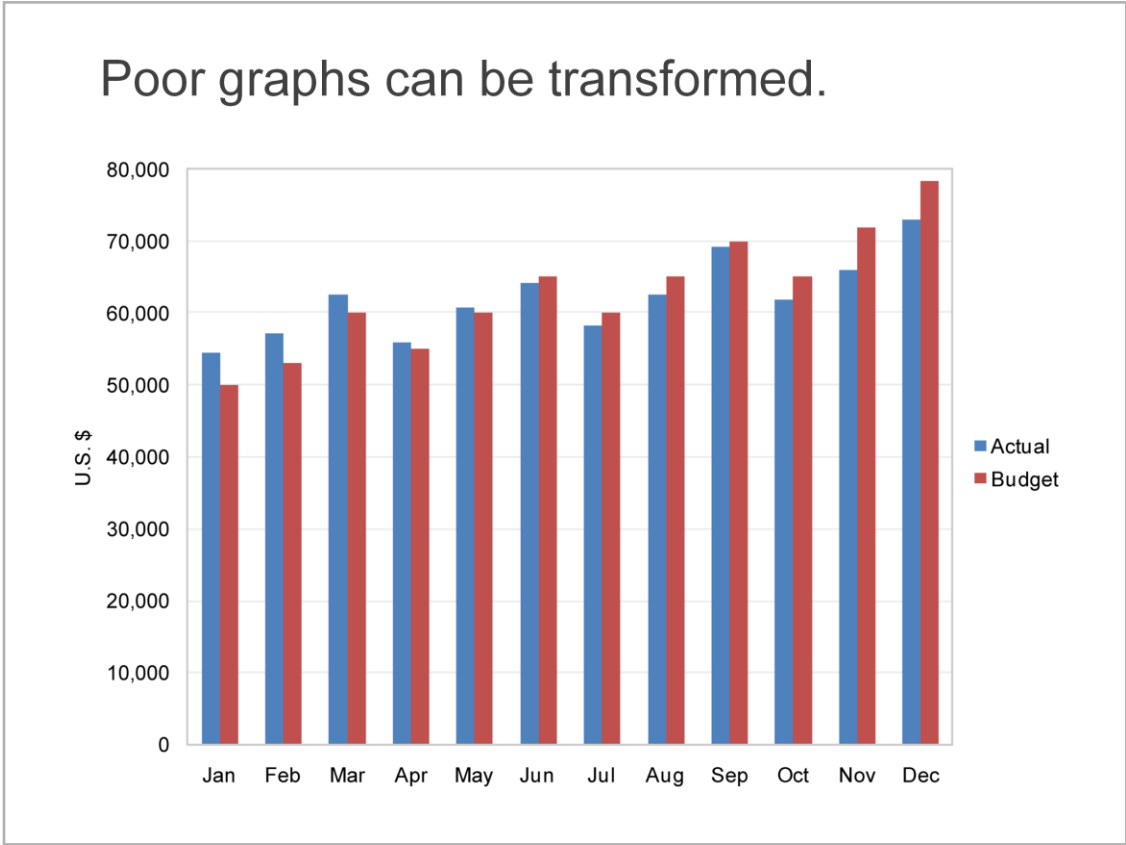


We have now enlarged the text, making it easier to read.

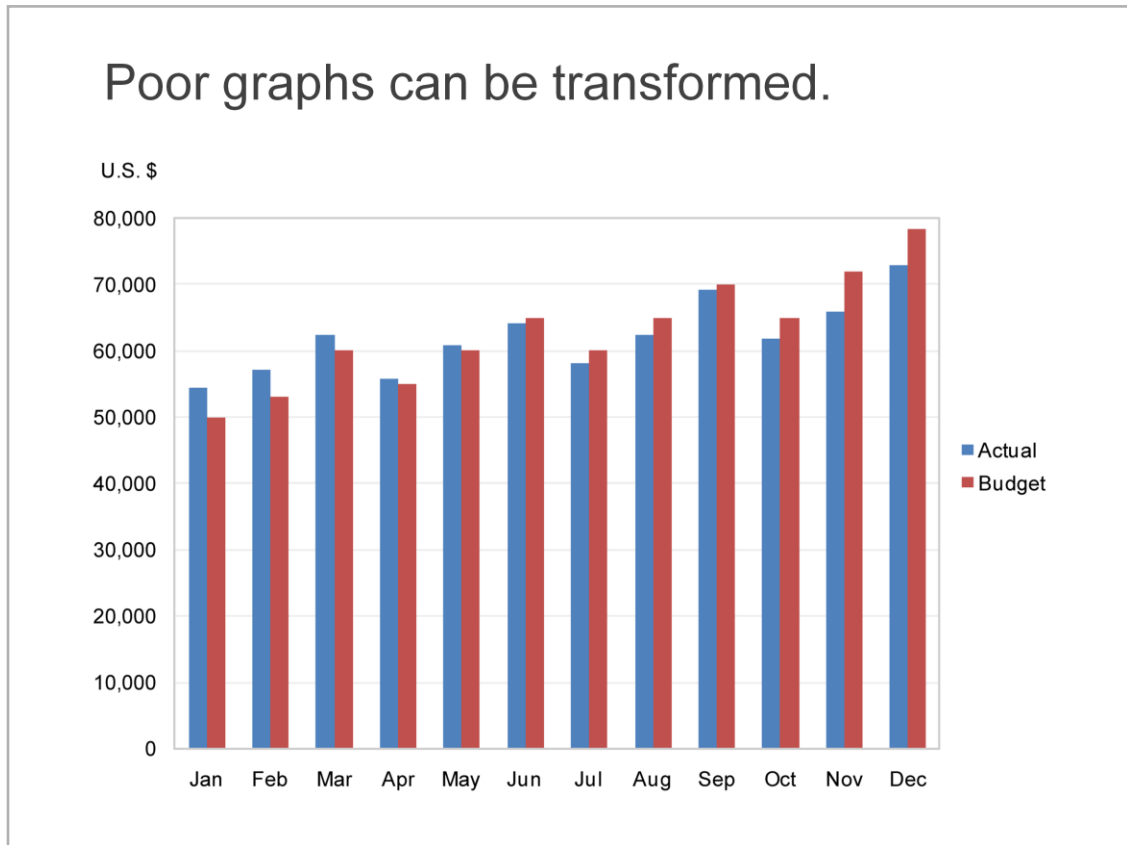
Poor graphs can be transformed.



We have now removed the unnecessary decimal places in the dollar amounts along the Y-axis.

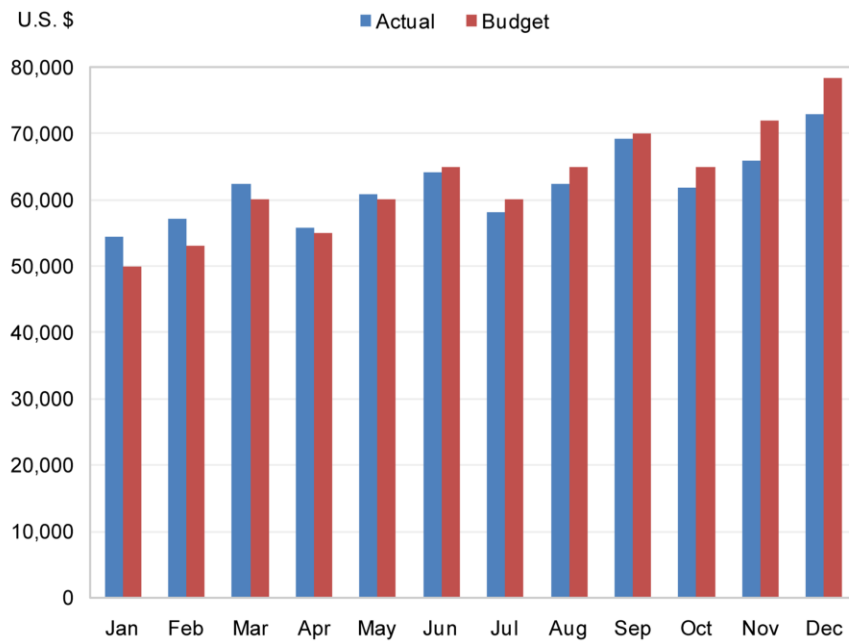


We have now removed the redundant dollar signs and labeled the unit of measure (U.S. \$) clearly.



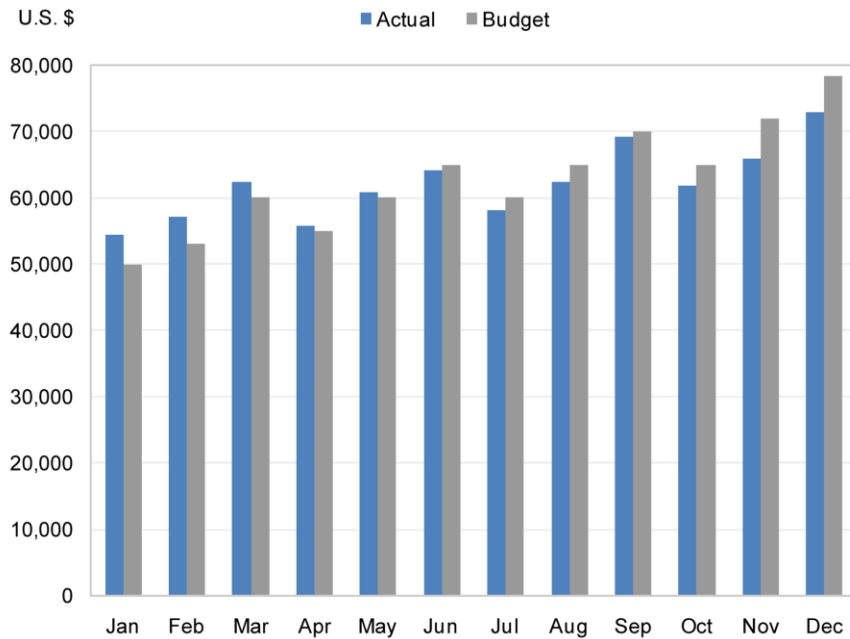
We have now reoriented the Y-axis label to the horizontal and placed it above the axis to make it easier to read.

Poor graphs can be transformed.



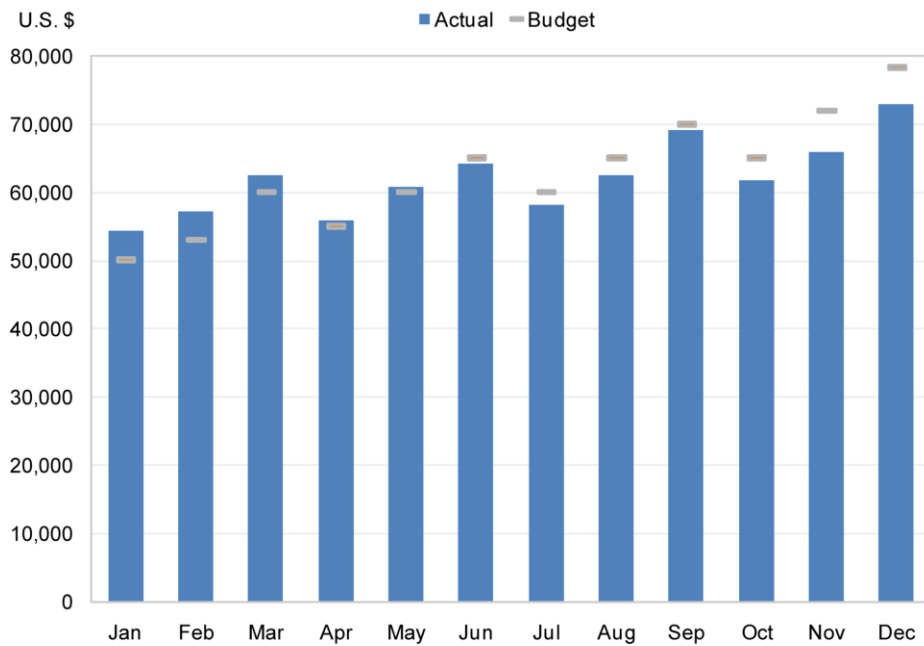
We have now reoriented and repositioned the legend to make it easier to associate it with the data bars.

Poor graphs can be transformed.



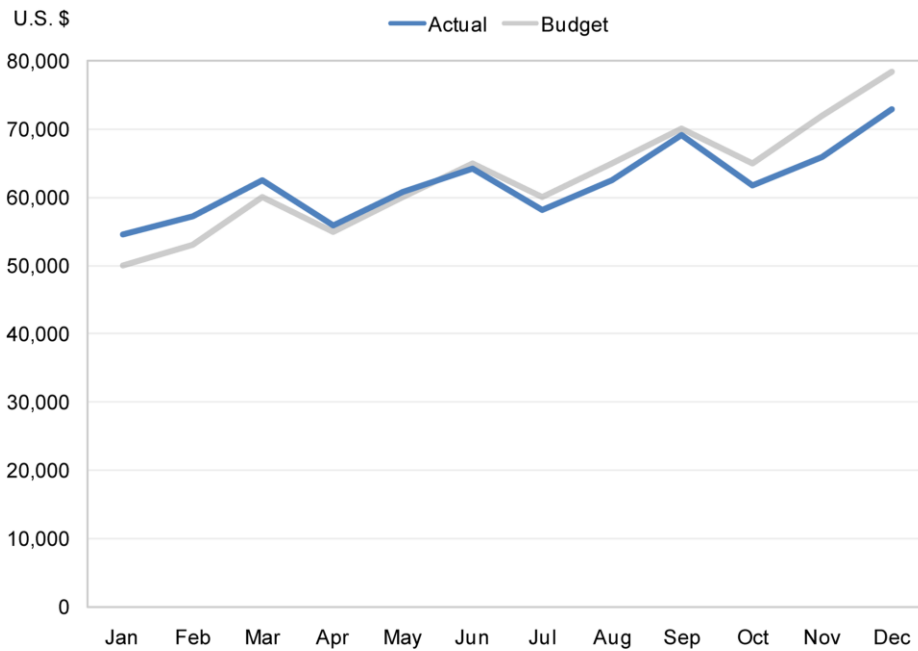
We have now changed the color of the Budget bar to be more visually pleasing in relation to the blue Actual bars. Changing from the color red also removed the possibility people interpreting the data as something bad or a warning, which red is often used to represent.

Poor graphs can be transformed.

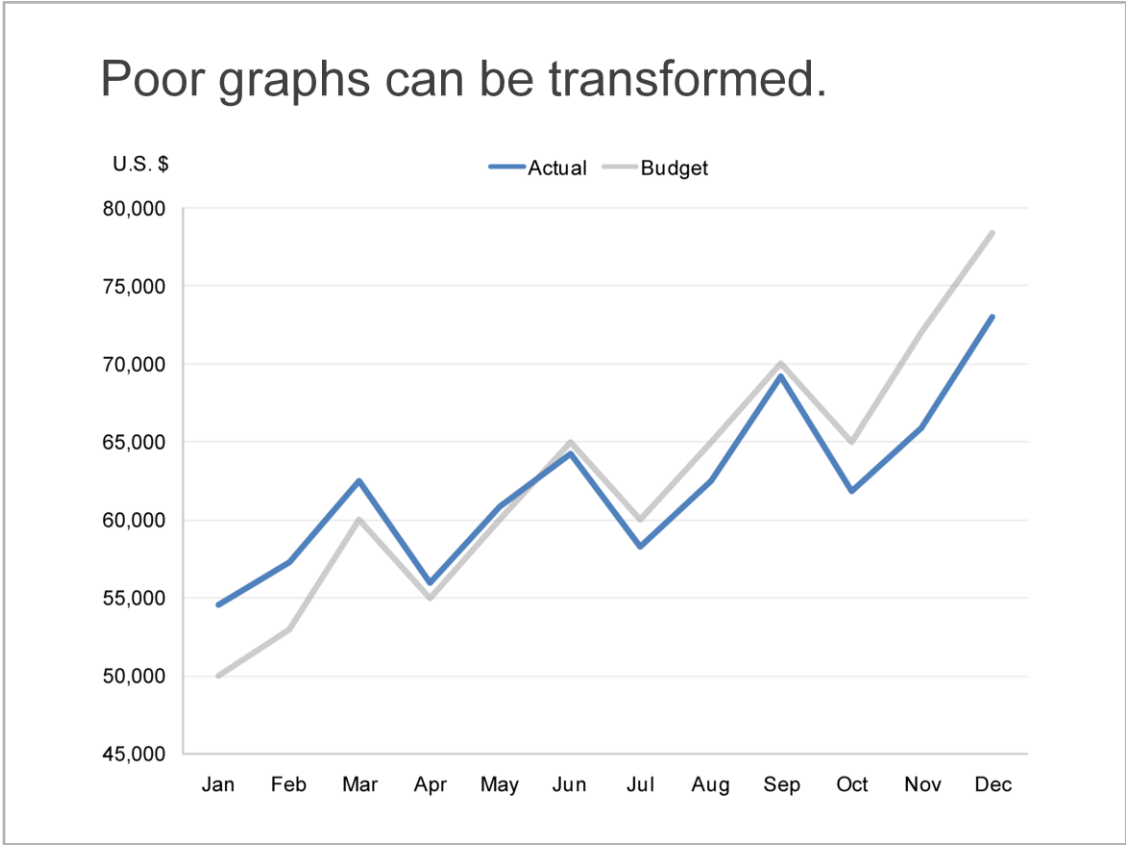


We have now reduced the visual salience of the Budget values, because they are less important than the Actual values, and have done so in a way that reduced clutter.

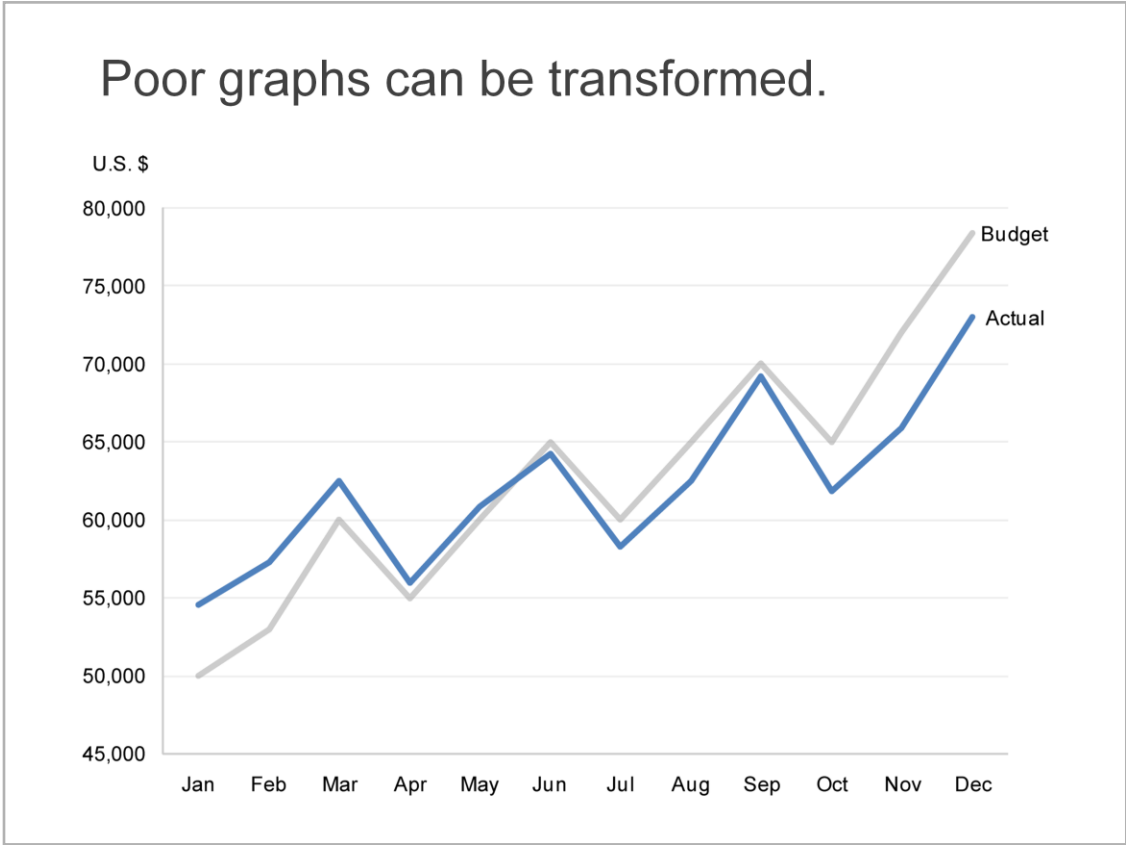
Poor graphs can be transformed.



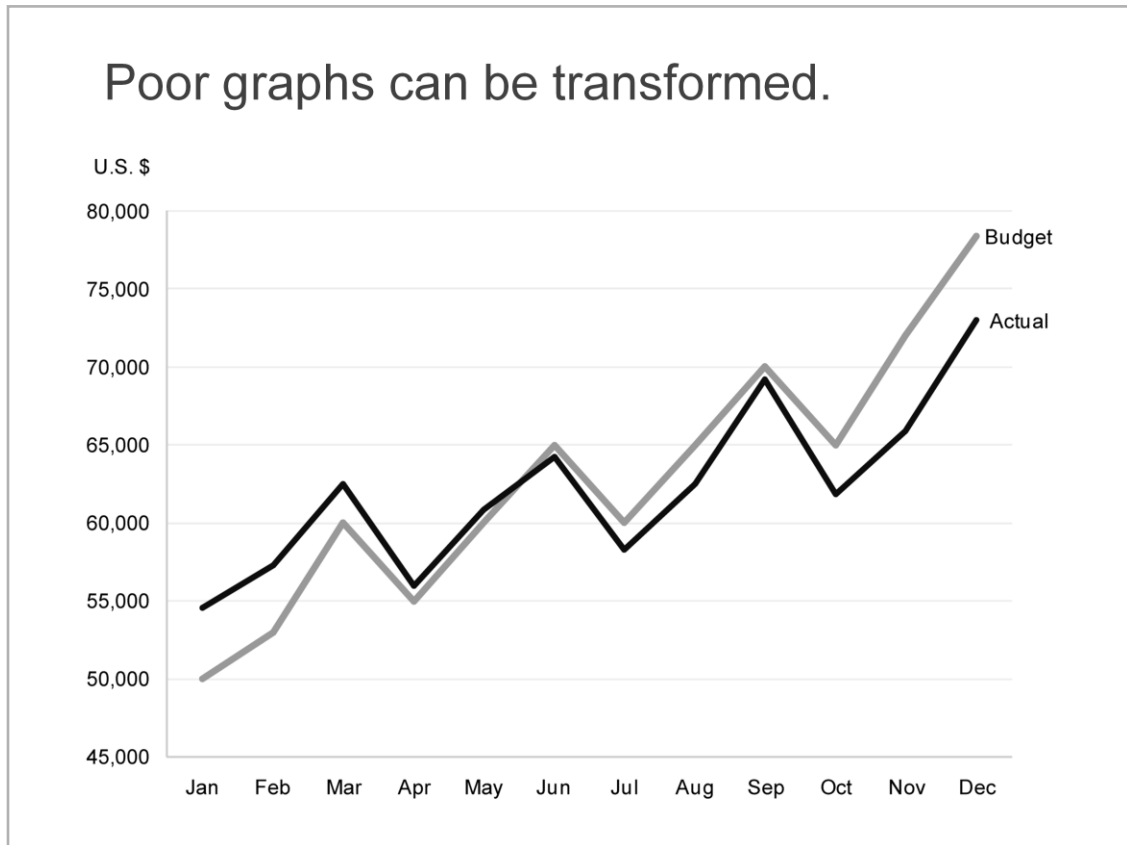
We have now made it much easier to see the pattern of change through time by using lines rather than bars to represent the data.



We have now made it much easier to examine the differences between actual expenses and the budget by spreading them across more space.

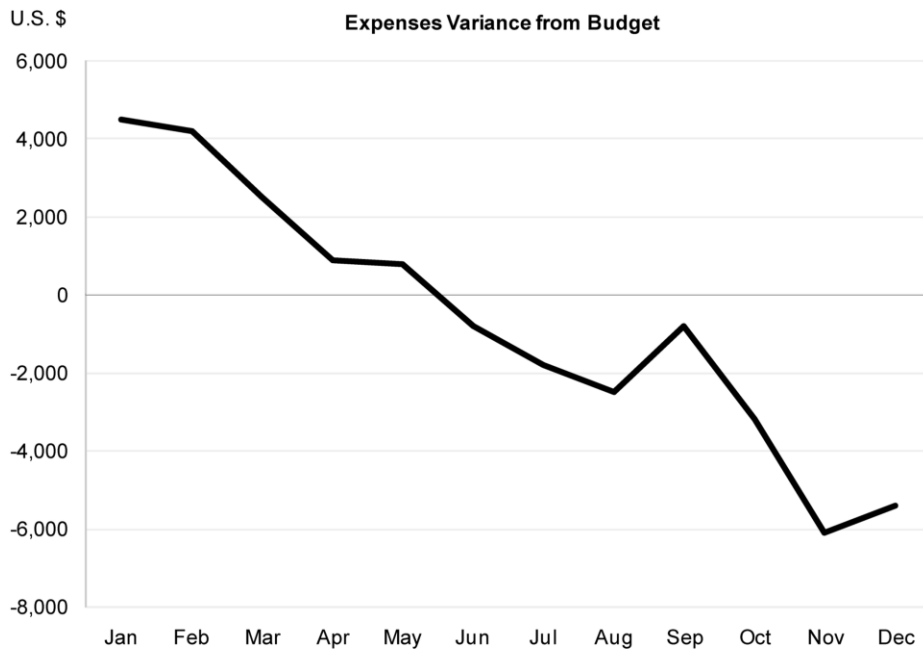


We have now labeled the lines directly, removing the need for a legend.



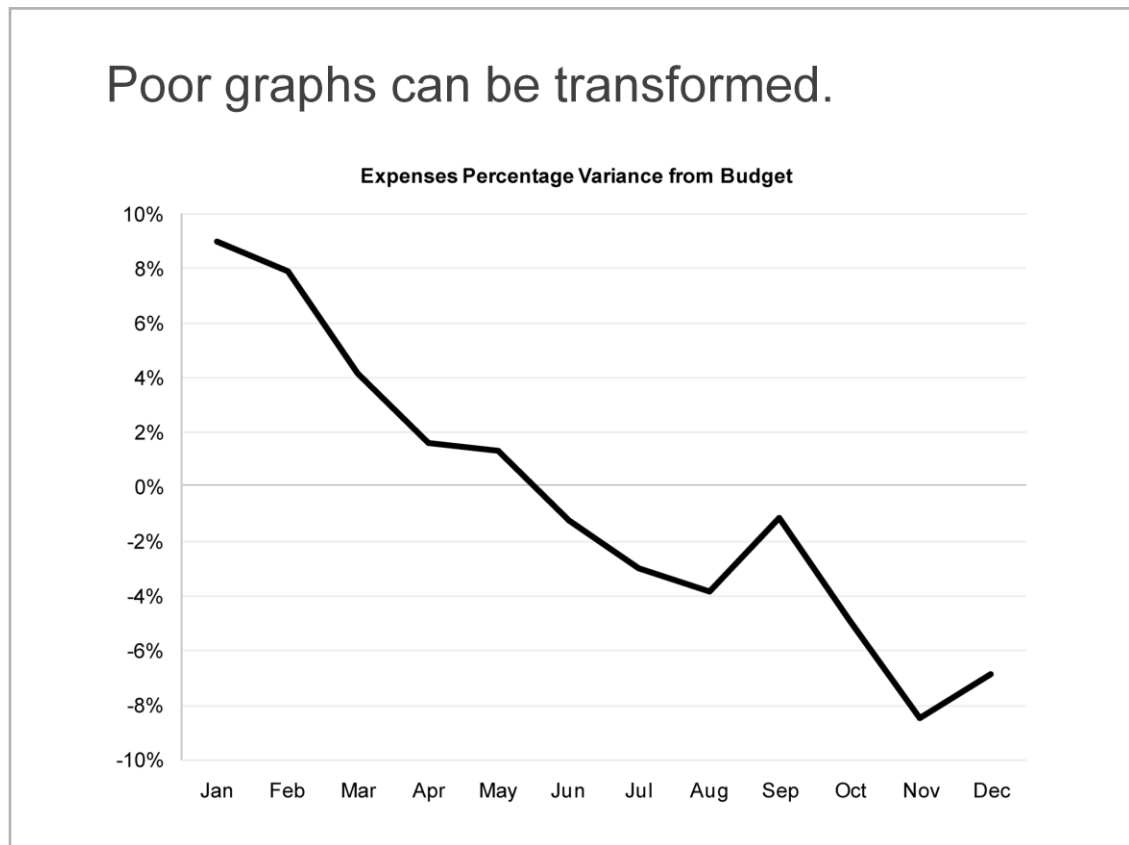
We have now changed the lines to two shades of gray to guaranty that even if the graph is printed on a black-and-white printer or photocopier, they will still look distinctly different from one another.

Poor graphs can be transformed.



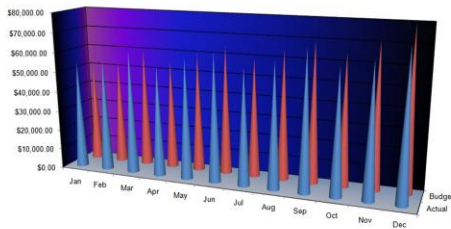
We have now represented the variance of actual expenses from the budget directly, as a single line.

Poor graphs can be transformed.

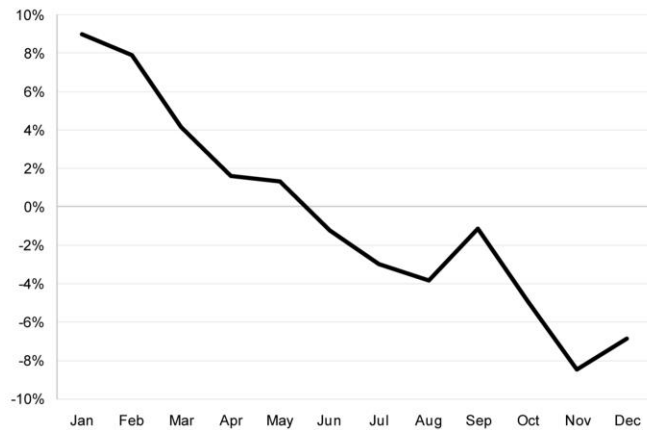


As our final step, we have expressed variance as a percentage, to provide a better measure of performance.

From incomprehensible to clear and simple



Expenses Percentage Variance from Budget



Our final solution, which we produced in sixteen steps, could have easily been our original solution. It usually takes no longer to design effective graphs than those that communicate poorly, if at all.

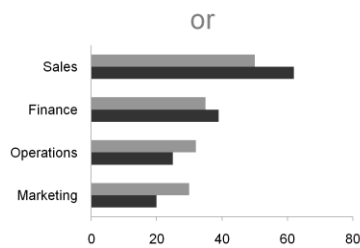
Fundamental challenges of data presentation

1. Determining the medium that tells the story best

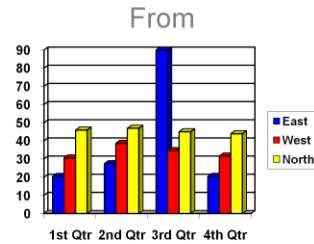
2. Designing the visual components to tell the story clearly

Either

Product	Units Sold	Revenue
Food	34,837	746,383
Beverage	42,374	845,984
Total	77,211	1,592,367



and which kind?



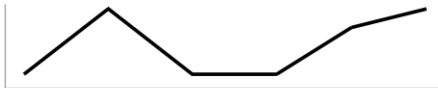
1. You begin by determining the best medium for your data and the message you wish to emphasize. Does it require a table or a graph? Which kind of table or graph?
2. Once you've decided, you must then design the individual components of that display to present the data and your message as clearly and efficiently as possible.

Tables work best when...

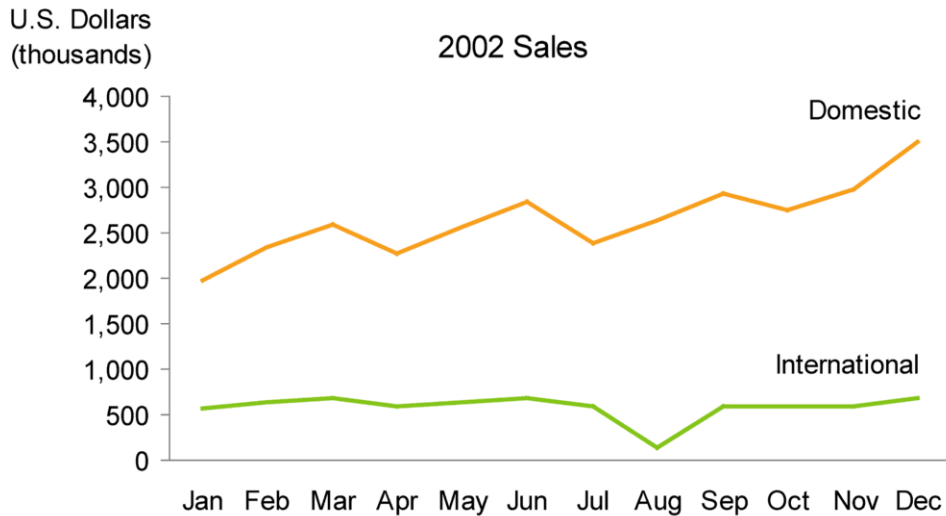
Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
1990	127.4	128.0	128.7	128.9	129.2	129.9	130.4	131.6	132.7	133.5	133.8	133.8	130.7
1991	134.6	134.8	135.0	135.2	135.6	136.0	136.2	136.6	137.2	137.4	137.8	137.9	136.2
1992	138.1	138.6	139.3	139.5	139.7	140.2	140.5	140.9	141.3	141.8	142.0	141.9	140.3
1993	142.6	143.1	143.6	144.0	144.2	144.4	144.4	144.8	145.1	145.7	145.8	145.8	144.5
1994	146.2	146.7	147.2	147.4	147.5	148.0	148.4	149.0	149.4	149.5	149.7	149.7	148.2
1995	150.3	150.9	151.4	151.9	152.2	152.5	152.5	152.9	153.2	153.7	153.6	153.5	152.4
1996	154.4	154.9	155.7	156.3	156.6	156.7	157.0	157.3	157.8	158.3	158.6	158.6	156.9
1997	159.1	159.6	160.0	160.2	160.1	160.3	160.5	160.8	161.2	161.6	161.5	161.3	160.5
1998	161.6	161.9	162.2	162.5	162.8	163.0	163.2	163.4	163.6	164.0	164.0	163.9	163.0
1999	164.3	164.5	165.0	166.2	166.2	166.2	166.7	167.1	167.9	168.2	168.3	168.3	166.6
2000	168.8	169.8	171.2	171.3	171.5	172.4	172.8	172.8	173.7	174.0	174.1	174.0	172.2
2001	175.1	175.8	176.2	176.9	177.7	178.0	177.5	177.5	178.3	177.7	177.4	176.7	177.1
2002	177.1	177.8	178.8	179.8	179.8	179.9	180.1	180.7	181.0	181.3	181.3	180.9	179.9

- Used to look up individual values
- Data must be precise

What do graphs do well?



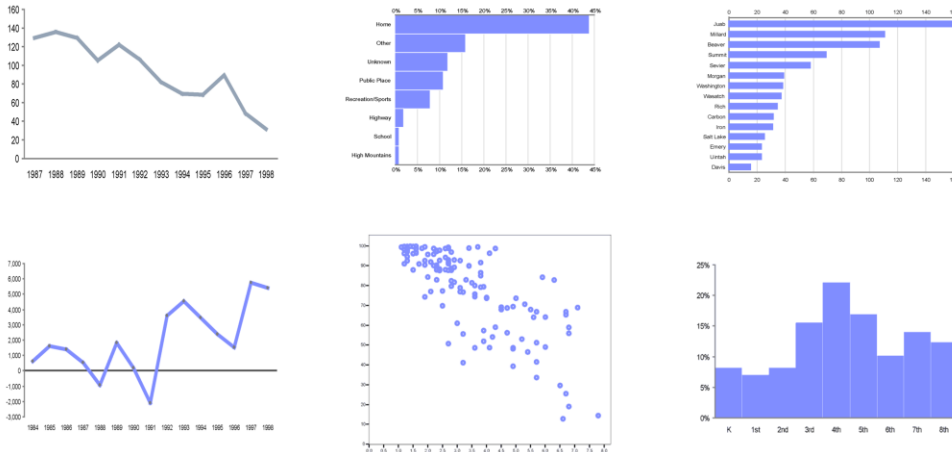
Feature patterns, trends, and exceptions



The saying, “A picture is worth a thousand words,” applies quite literally to quantitative graphs. By displaying quantitative information in visual form, graphs efficiently reveal information that would otherwise require a thousand words or more to adequately describe.

Joseph Berkson once stated what happens quite powerfully: When we visualize the data effectively and suddenly, we experience “interocular traumatic impact”—a conclusion that hits us between the eyes.

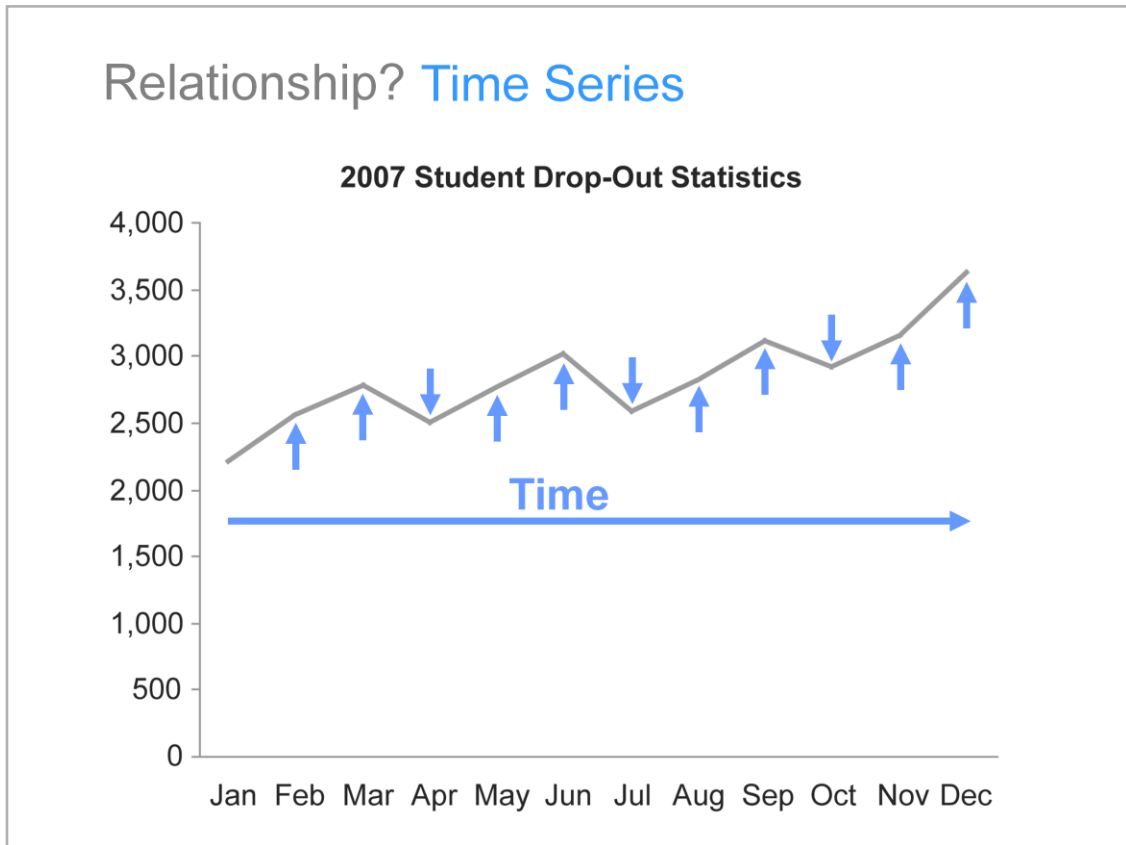
Quantitative messages always feature relationships and make comparisons.



The stories contained in numbers all involve relationships. Numbers that measure public health, in fact, involve six fundamental types of relationships.

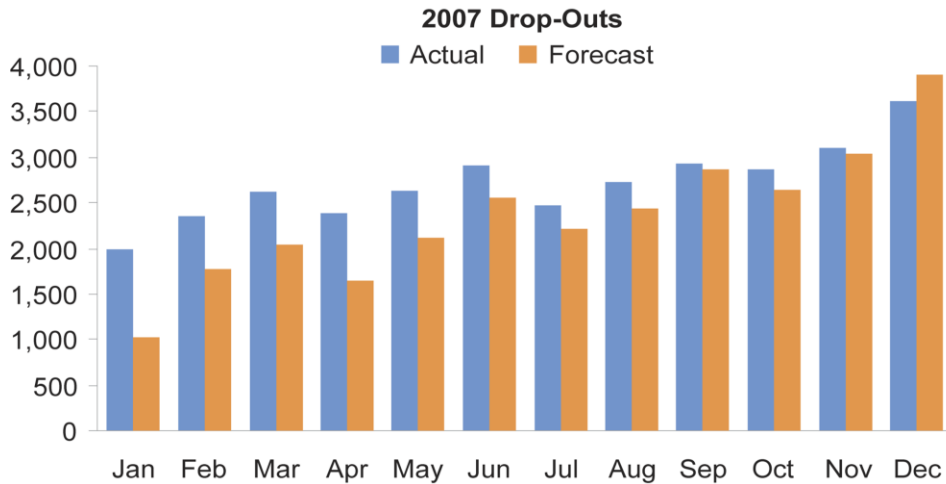
If you know the relationship you're trying to communicate graphically and you know the best way to graphically encode that relationship, you possess a simple vocabulary for telling quantitative stories.

Allow me to introduce the six relationships that you should get to know.



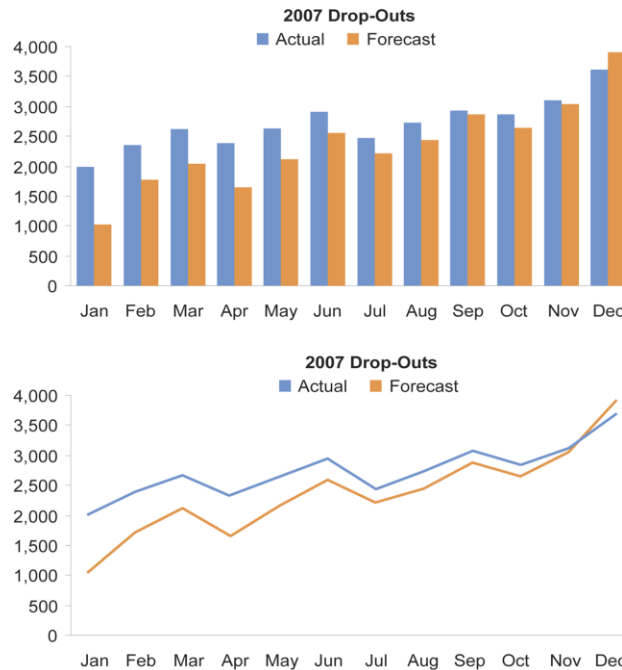
This graph features the relationships between values as they change through time, which is perhaps the most common quantitative relationship that you need to communicate.

Could this pattern of change be displayed more clearly?



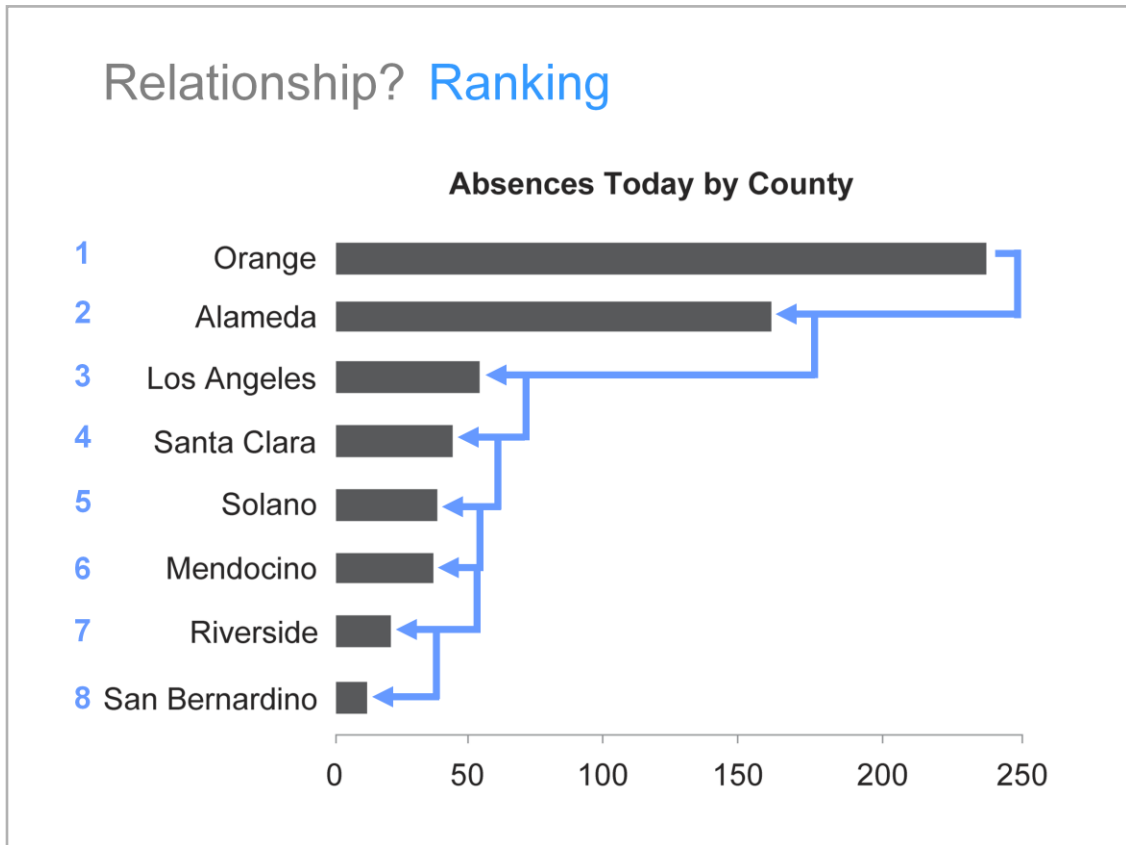
Here's an attempt to display a time-series relationship regarding actual vs. budgeted sales, which works, but the trend and patterns could be much more clearly displayed.

Bars and lines tell time differently.



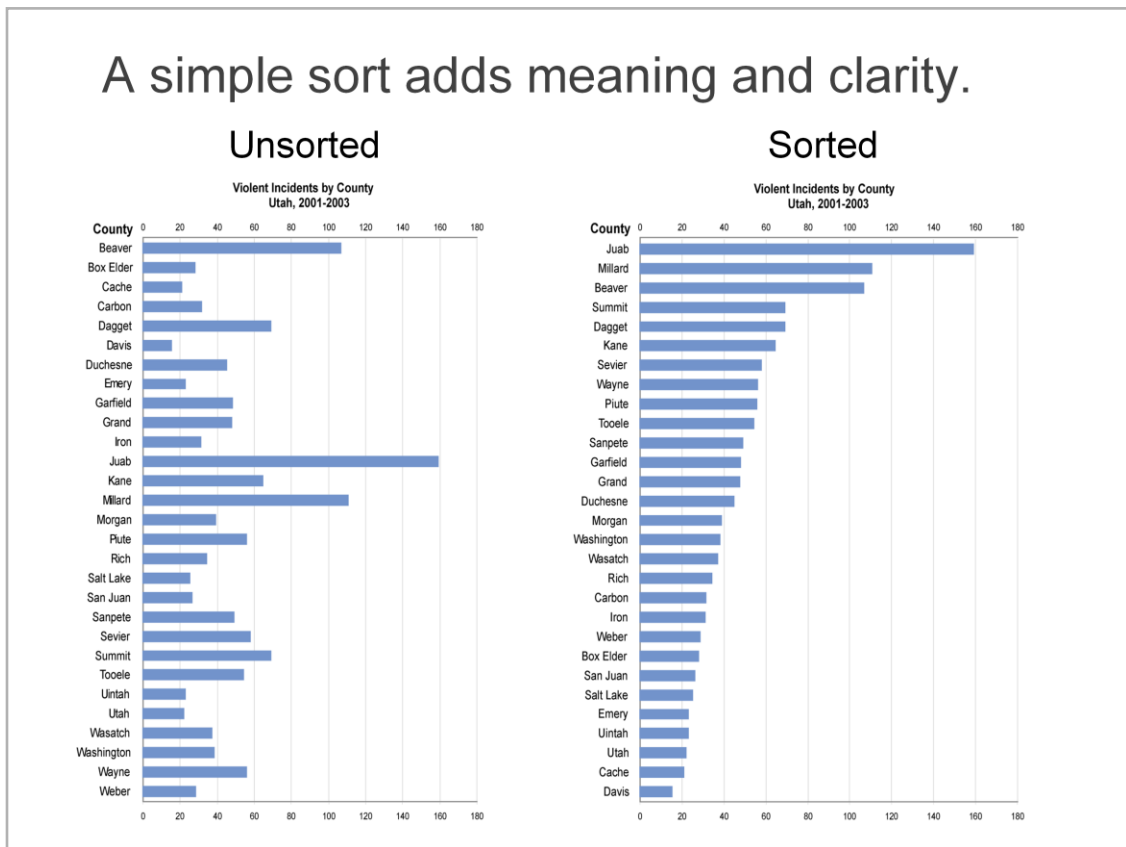
Here’s the same exact data presented in two ways: the top graph uses bars and the bottom graph uses a line. Which displays the shape of change through time more clearly?

Bars work well for comparing individual values to one another, but lines show the shape of change through time much more clearly.



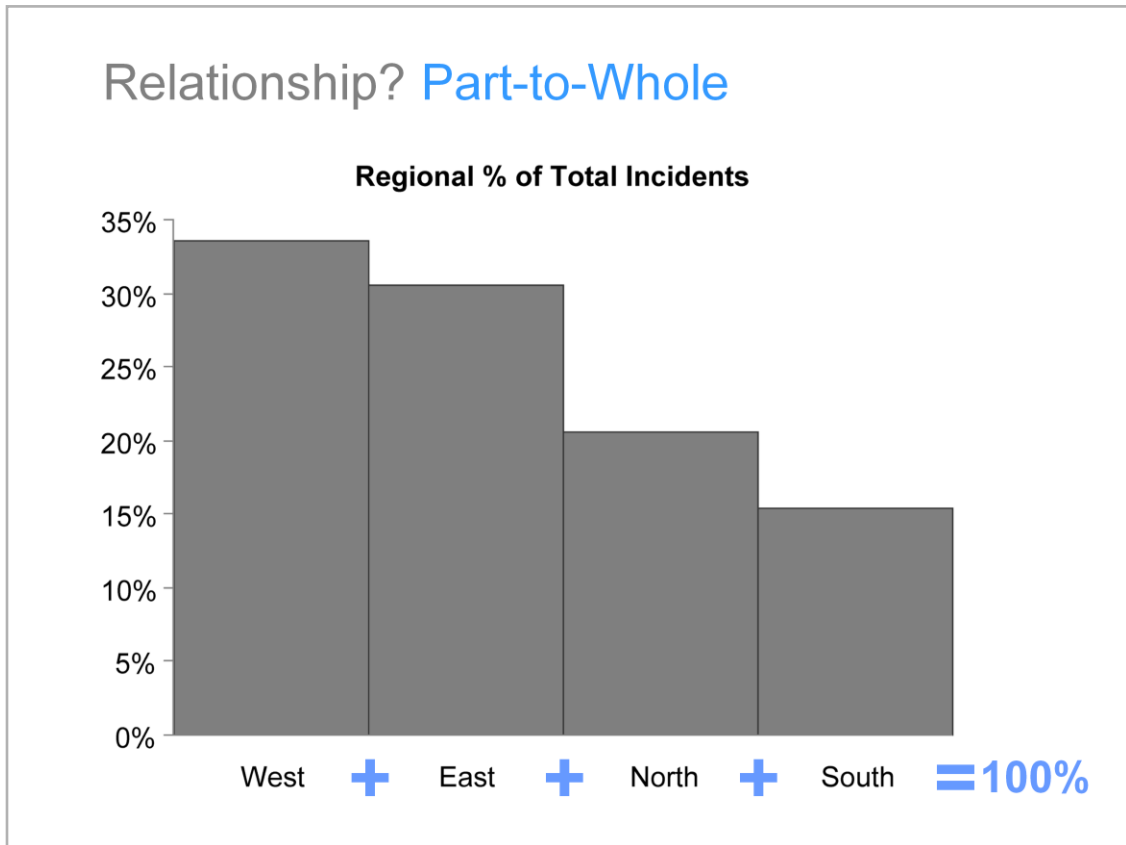
We've been talking about time-series relationships, but this graph features a different relationship between the values. The values are arranged in order of size, in this case from big to small. Arranging values sequentially makes them easier to compare to one another and directly communicates the relationship of rank.

A simple sort adds meaning and clarity.



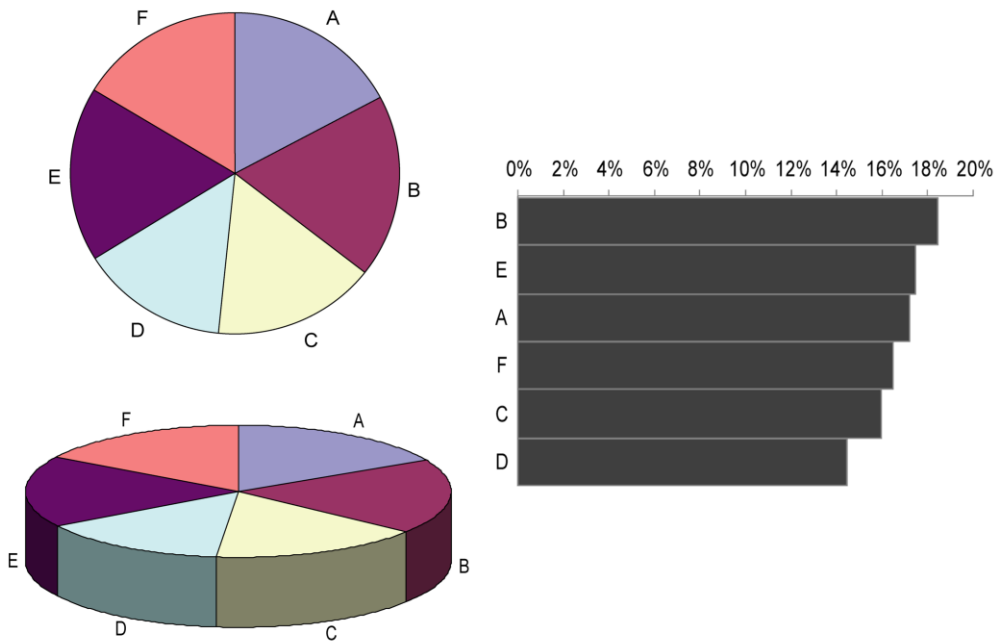
In the display of trauma registry injuries by county on the left, notice how difficult it is to compare the values and to get a sense of rank when they aren't sequenced by size.

The same information is displayed on the right, this time with the counties arranged by the number of injuries. If the purpose of the display is to look up individual values, which is the only thing that alphabetical order supports, a table would work much better. The ranking display on the right, however, tells a useful story.



This graph features another relationship that is commonly displayed. Notice that if you add the values they total 100%. This is what I call a part-to-whole relationship, which shows how the individual values that make up some whole relate to one another and to the whole.

Save the pies for dessert!



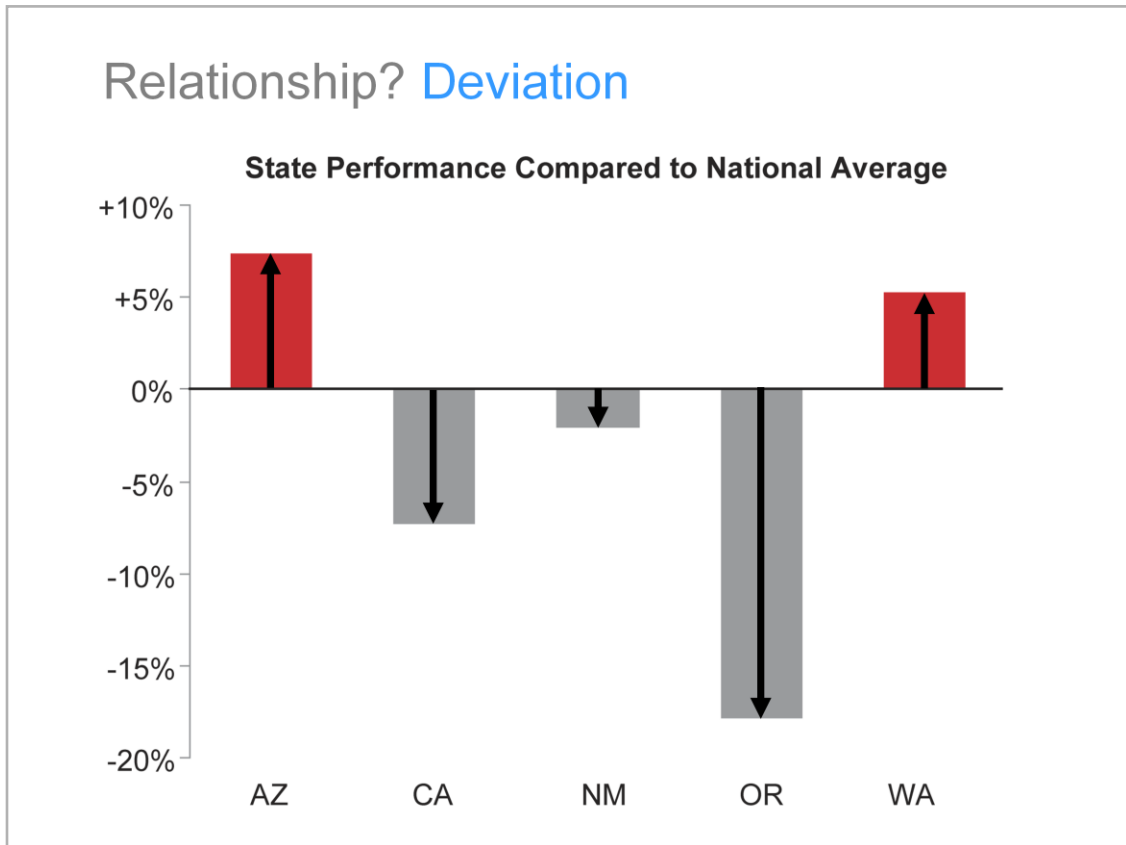
Pie charts use 2-D areas and the angles formed by slices to encode quantitative values. Unfortunately, our perception of 2-D areas and angles as measures of quantity is poor.

Since all graphs have one or more axes with scales, there must be one on a pie chart, but where is it? The circumference of the circle is where its quantitative scale would appear, but it is rarely shown.

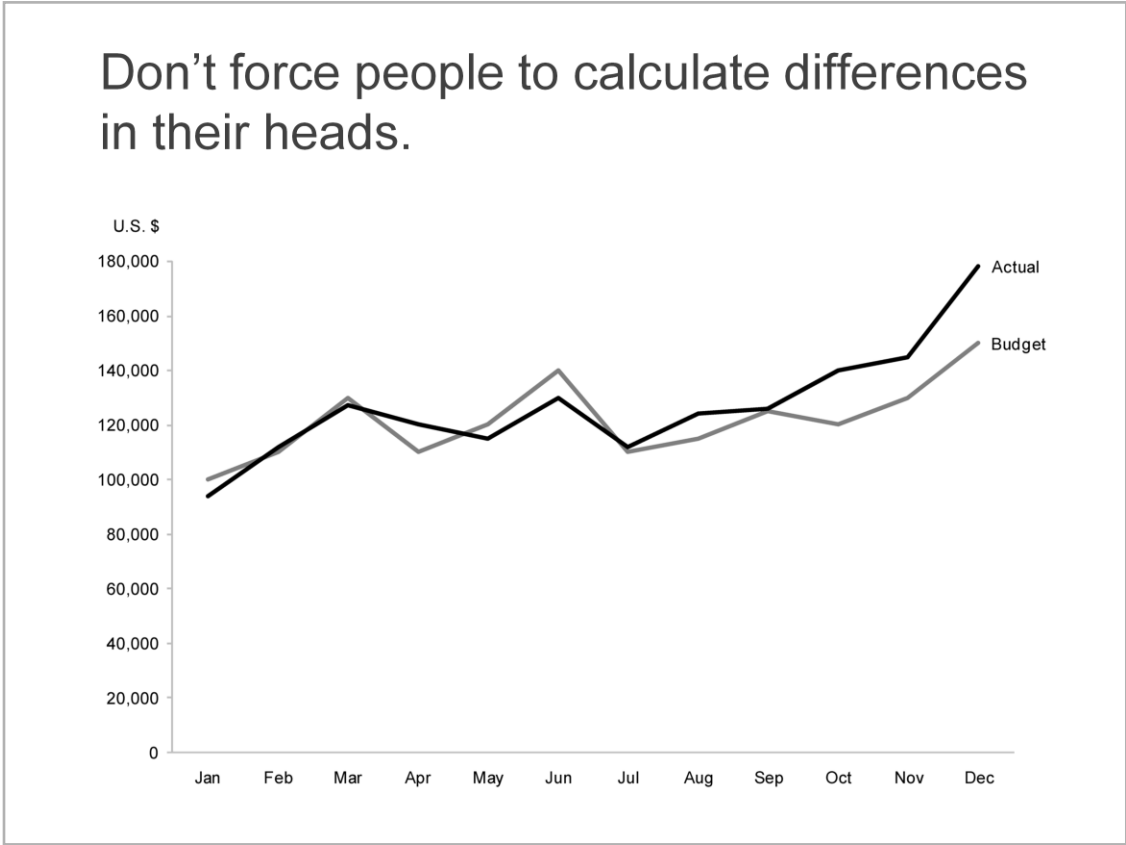
Try using either one of the pie graphs to put the slices in order by size. Can't do it, can you? Now see how easy this is to do when the same data is encoded in a bar graph.

Coda Hale once expressed his opinion of pie charts quite colorfully:

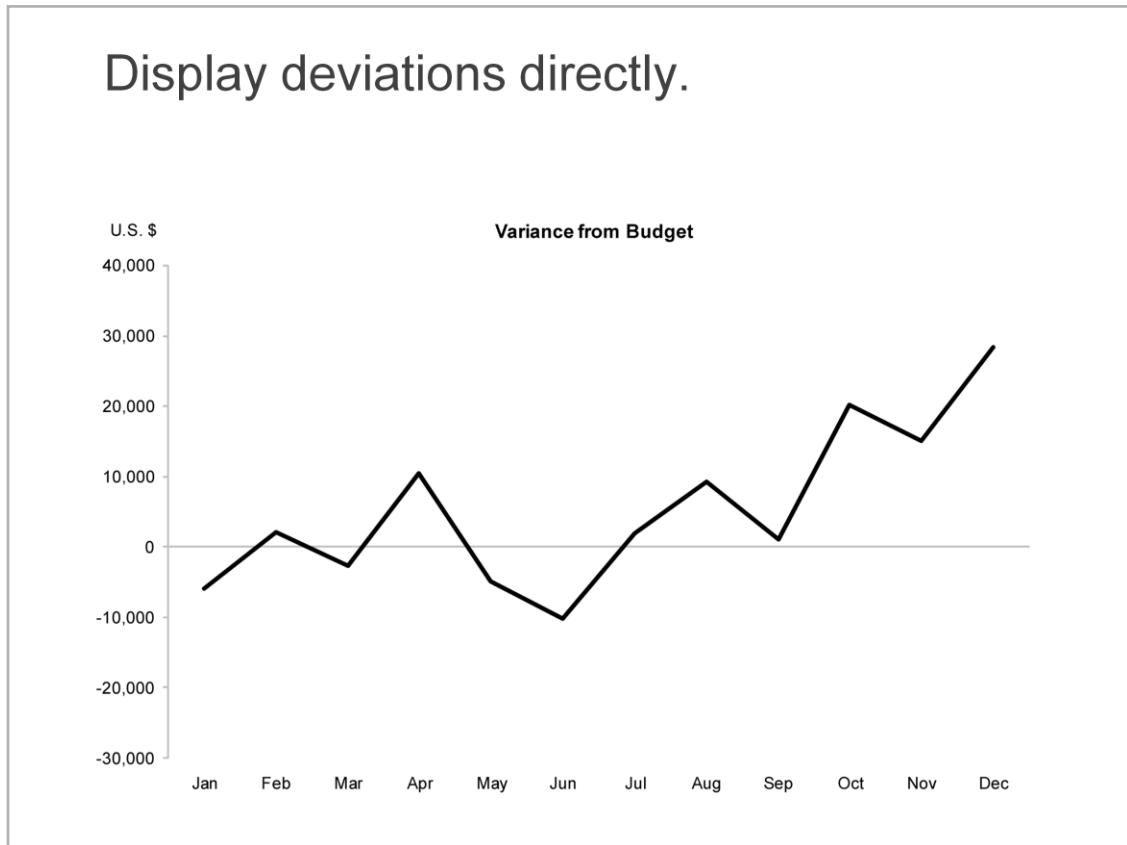
Pie charts are the information visualization equivalent of a roofing hammer to the frontal lobe...[Piecharts] have no place in the world of grownups, and occupy the same semiotic space as short pants, a runny nose, and chocolate smeared on one's face. They are as professional as a pair of assless chaps. Anyone who suggests their use should be instinctively slapped.



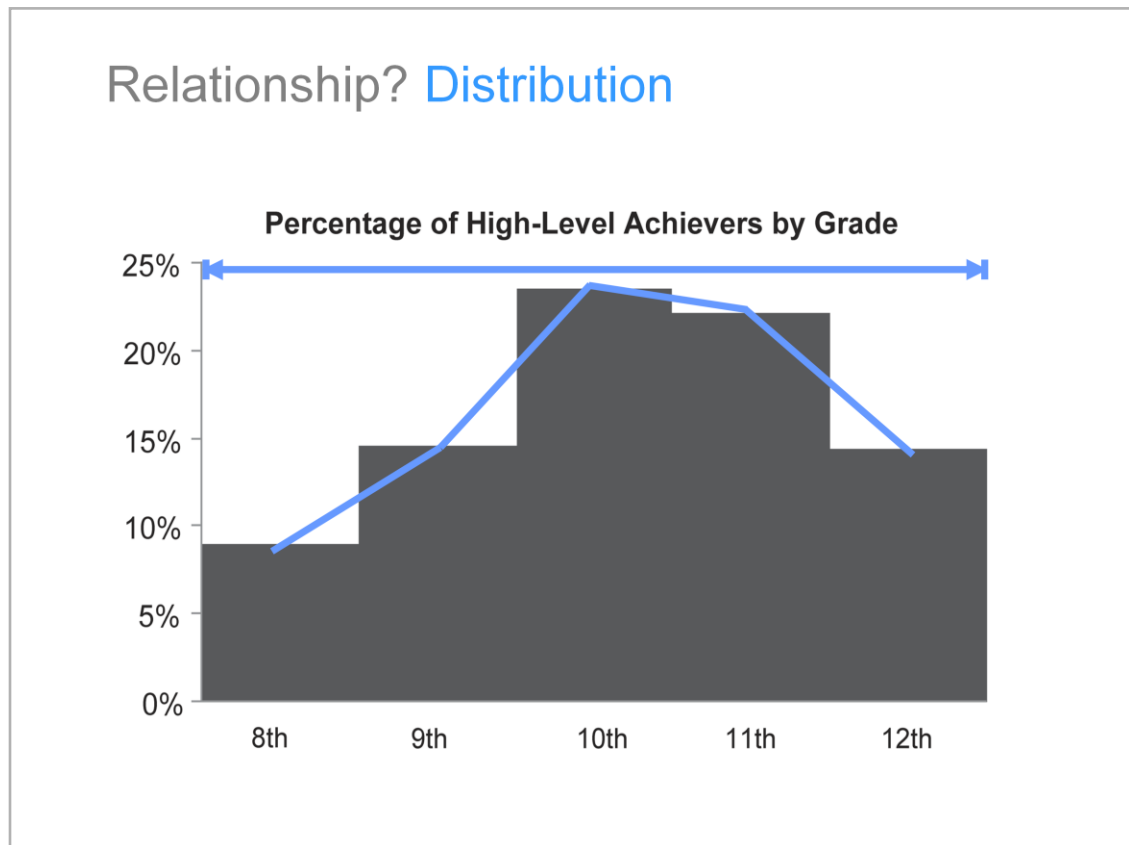
This graph features the next common relationship between values. This approach is used when you want to feature how one set of values differs from another reference set of values. I call this a deviation relationship.



When people primarily need to see the differences between things, show them the difference directly, rather than showing them the two sets of values and forcing them to construct a new picture in their heads of how they differ.



This graph directly expresses how actual sales differed from the budget through the year.

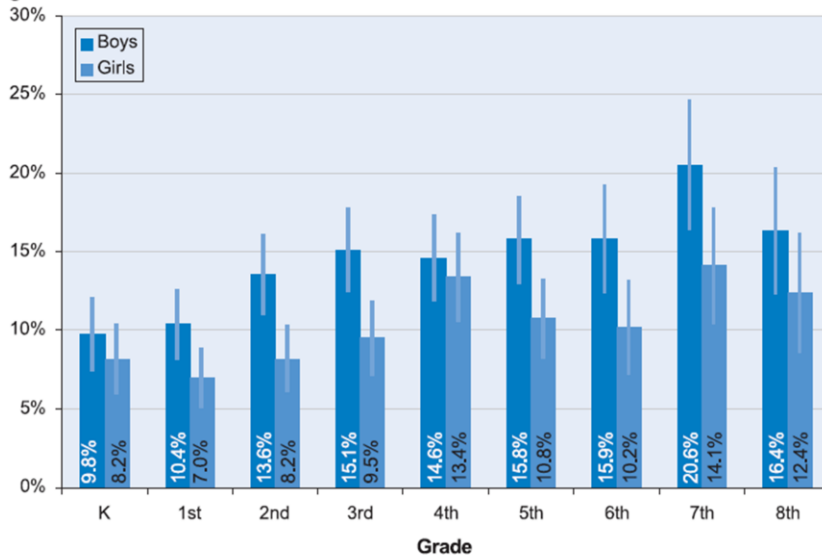


The relationship that this graph features the spread of a set of values from lowest to highest and the shape of their distribution across that range.

Histograms with multiple data sets can look cluttered.

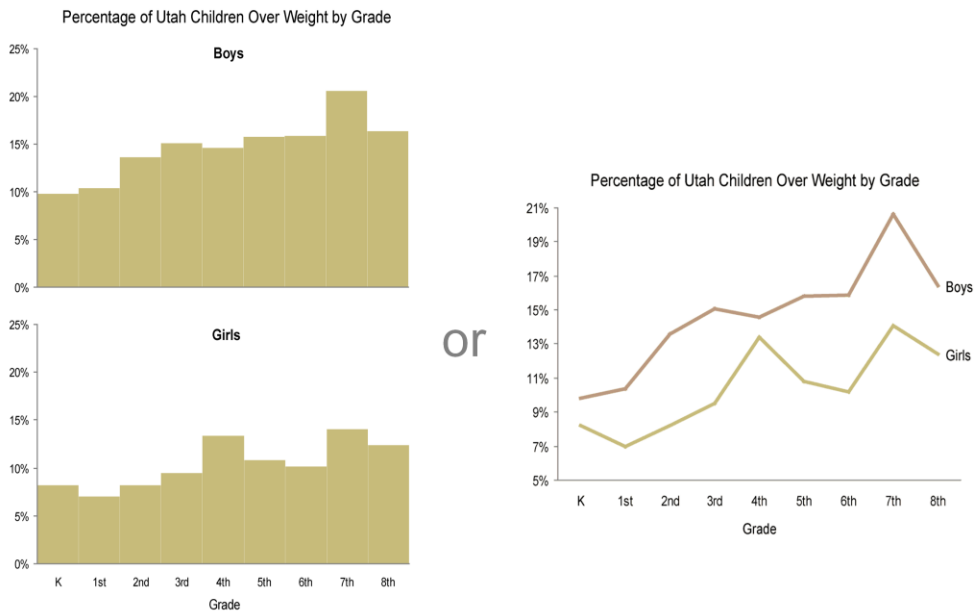
Overweight* by Grade and Gender

Figure 1. Percentage of Utah children who were overweight by grade and gender, Utah, 2002



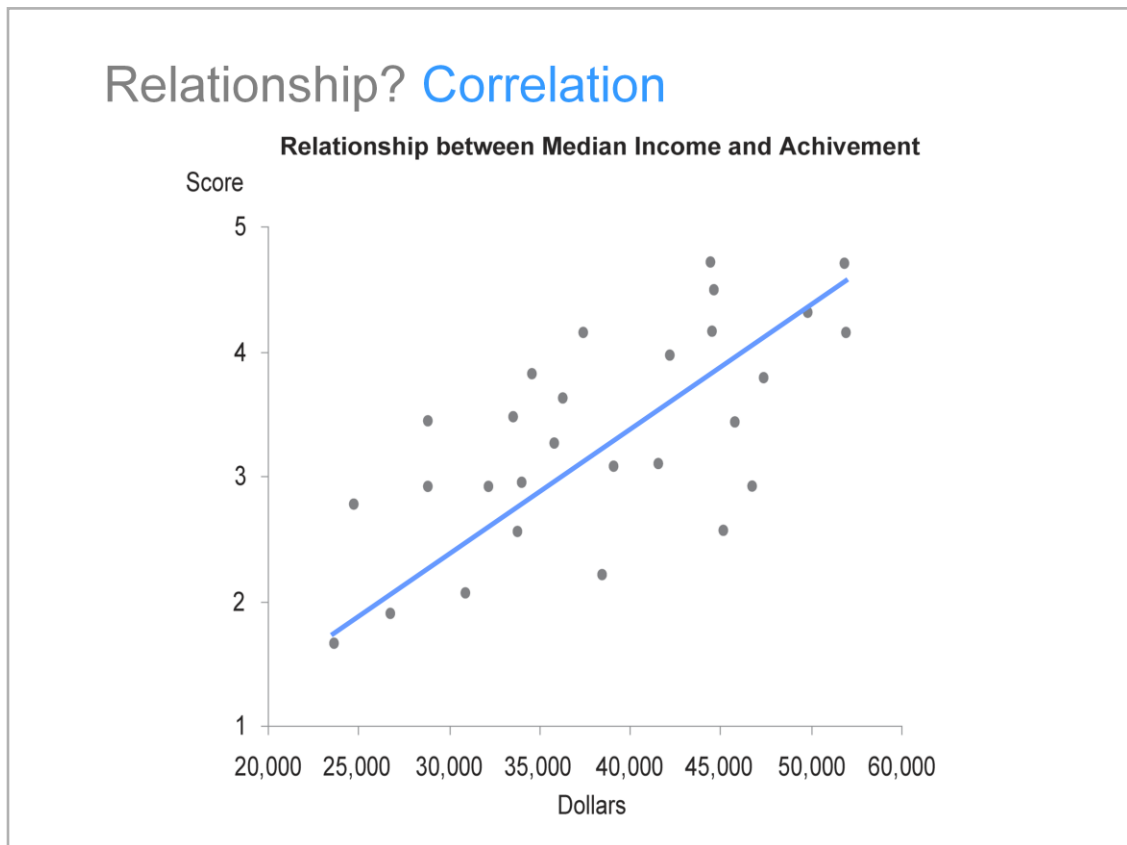
Here's a graph that attempts to show the distribution of overweight children by grade separately for boys and girls, but doing it in this way results in clutter that makes the patterns difficult to segregate and compare.

Clutter can be easily eliminated.



This pair of histograms—one for boys and one for girls—however, are arranged in a way that makes the patterns of each easy to see, yet still easy to compare.

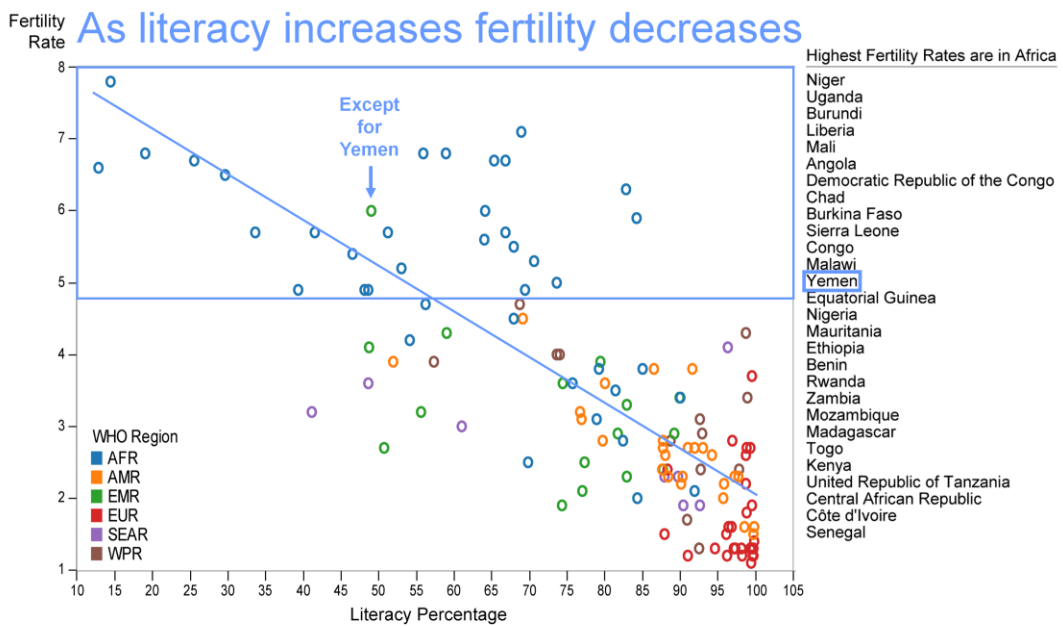
Even better, by using lines rather than bars, the separate patterns can be shown in the same graph in a way that features the shape of the patterns and how they differ.



This graph illustrates the last of the six relationships. Graphs such as this feature co-relations between two paired sets of values, so this relationship is called a correlation.

Correlations show whether two paired sets of measures, such as these purely fictional sets of people's heights and salaries, vary in relation to one another, and if so, in which direction (positive or negative) and to what degree (strong or weak).

Interesting correlations abound.



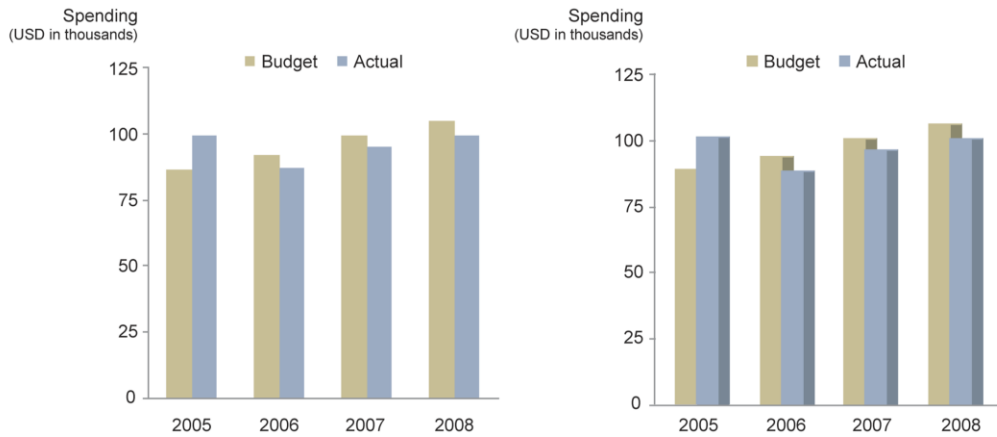
This example, based on WHO data, explores the correlation between adult literacy and fertility rate by country. A correlation clearly exists: higher literacy corresponds to lower rates of fertility. It is also clear from this display that the highest rates of fertility all occur in Africa (the blue circles), which the one exception of Yemen (the one green circle at the high end of fertility).

Six common relationships in graphs

- Time-series
- Ranking
- Part-to-whole
- Deviation
- Distribution
- Correlation

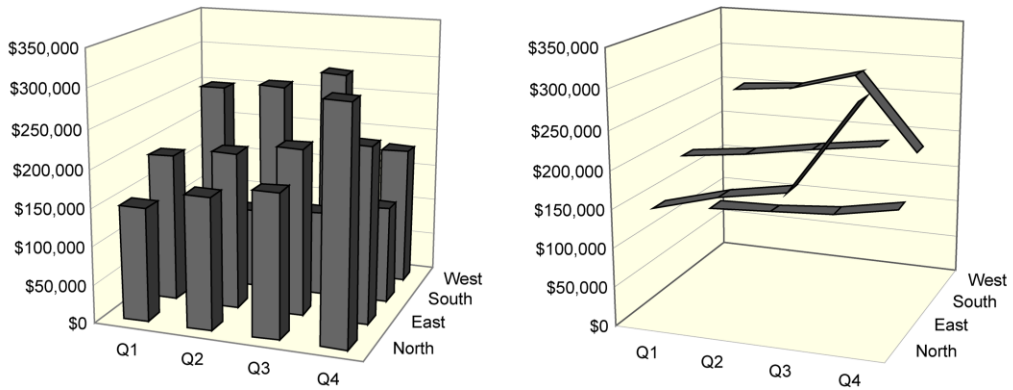
Almost every quantitative story involves one or more of these fundamental relationships, these comparisons between numbers. The skills required to tell these stories clearly are easy to learn.

What about 3-D? It's so cute!!!



Adding a third dimension of depth to the bars on the right without adding a corresponding third variable is not only meaningless, it makes it more difficult to decode the data.

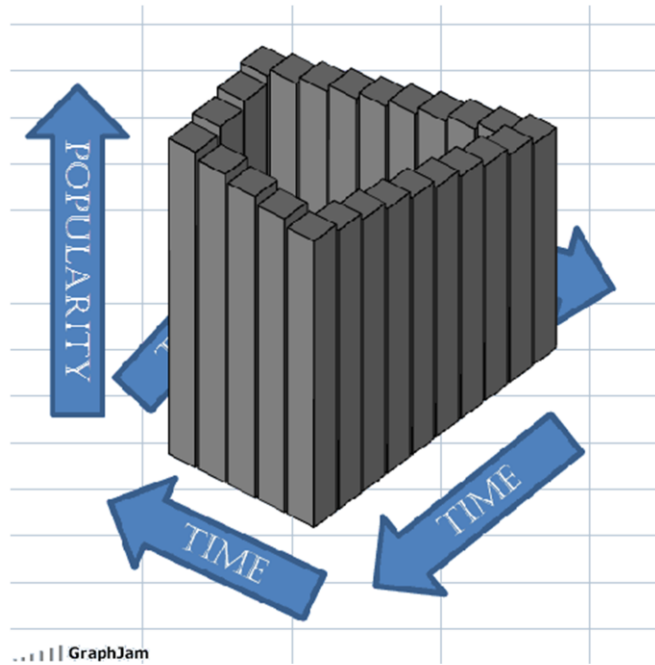
But what if there's a 3rd variable?



Can you determine which of the lines in the graph on the right represents the East region? Are you sure?

A third dimension with a corresponding variable is too hard to read.

Don't make people solve a 3-D puzzle.

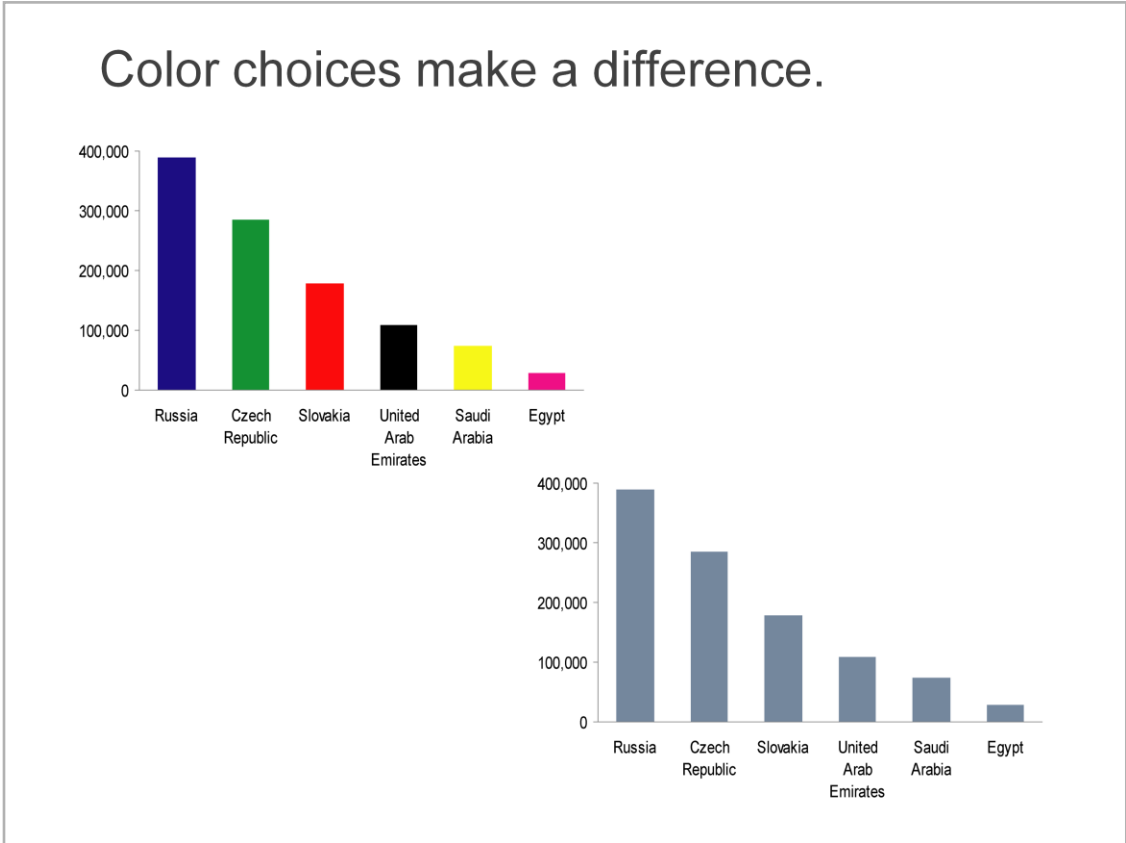


This chart of Escher's changing popularity through time was created by B. Brucker. I found it at www.GraphJam.com.

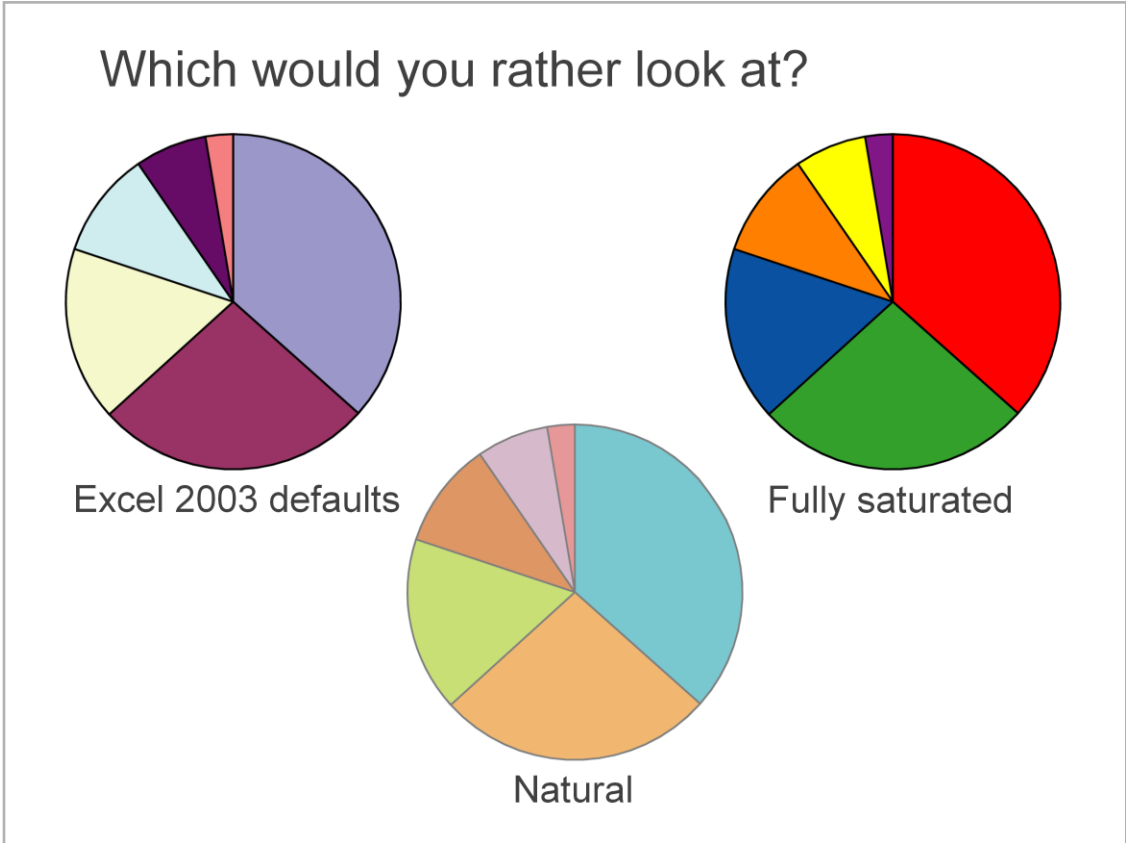
Is more color always better?



Lots of bright color is great if you're a preschooler. For adults, frequent use of bright colors accosts visual perception. Pastels and earth tones are much easier to look at. Use bright colors only to make particular data stand out above the rest.

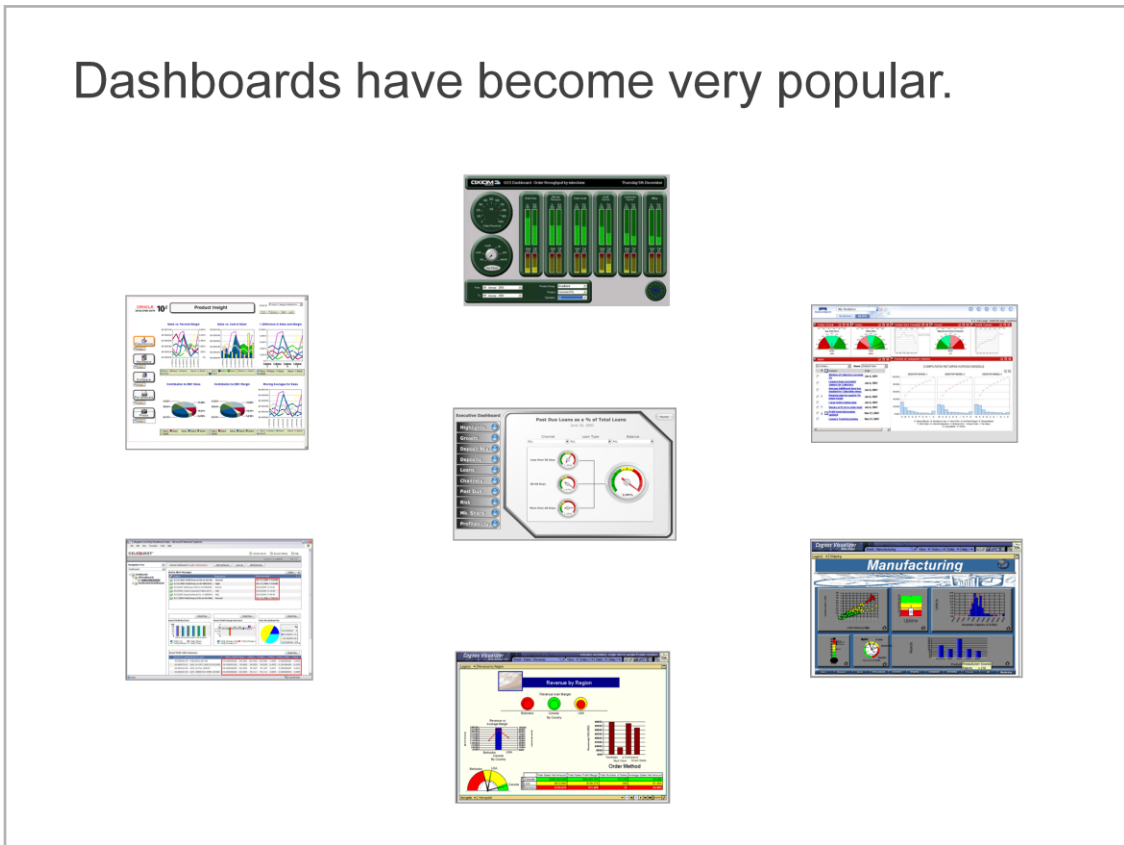


The top graph varies the colors of the bars unnecessarily. We already know that the individual bars represent different countries. Varying the colors visually separates the bars by making them look different from one another, but we want them to look alike to encourage people to compare them and to see the ranking pattern that they form as a whole.



Notice that, despite the softness of the colors in the example of natural colors, they still do the job of separating the sections of these pies just as well as the other examples, but do so in a manner that is much more pleasant to look at.

Dashboards have become very popular.



Dashboards are extremely popular today. Everyone assumes that dashboards are useful. Like all aspects of business intelligence, however, they can be done well or they can be done poorly.

Dashboard defined

A visual display
of
the most important information needed
to achieve one or more objectives
that has been
consolidated on a single computer screen
so it can be
monitored and understood at a glance

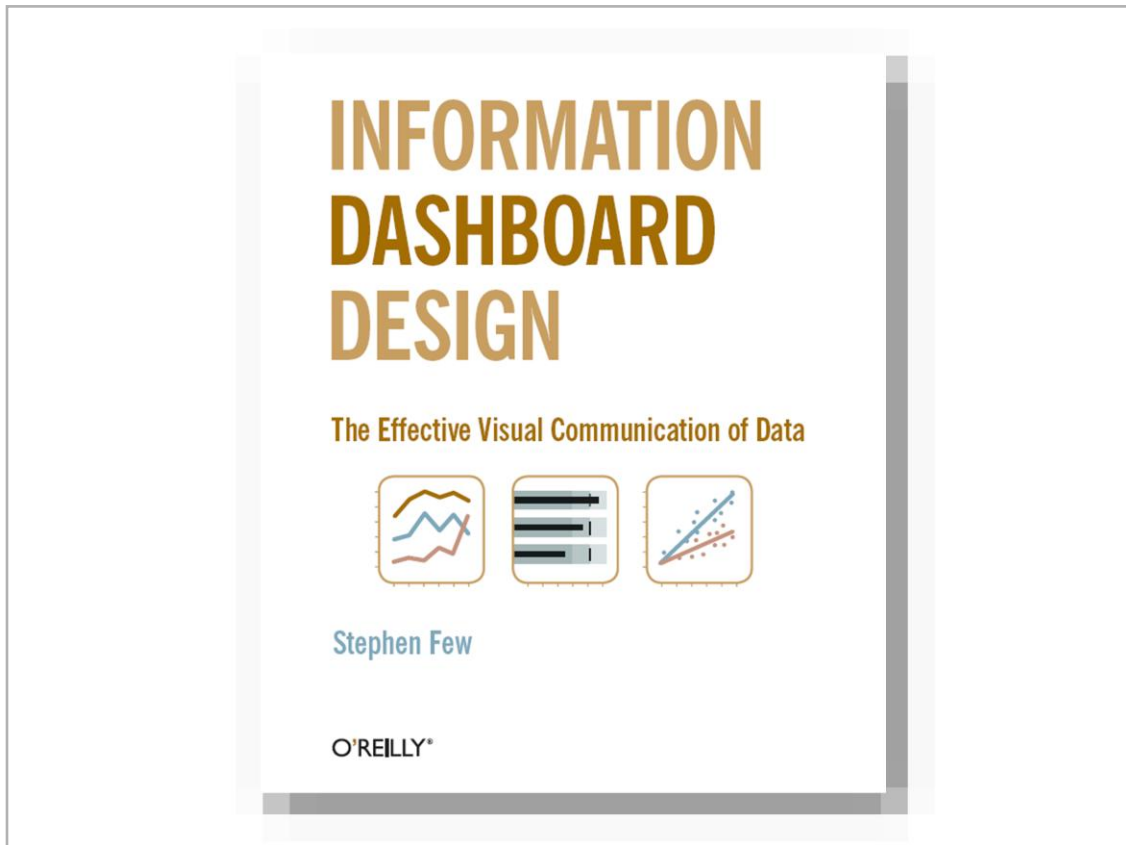
A dashboard is a visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at a glance.

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

But most dashboards fail.

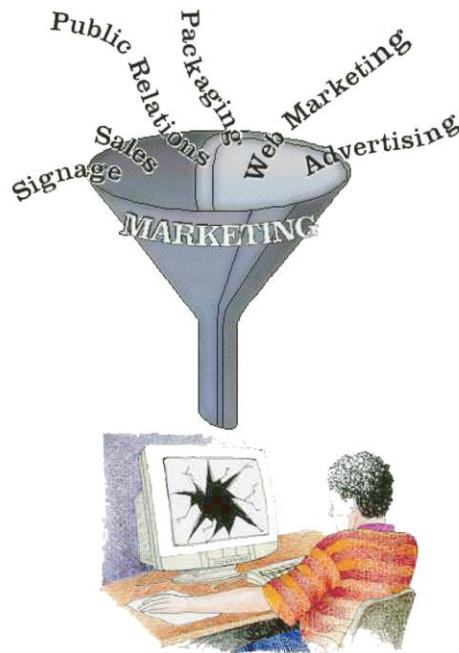


Dashboards like this are typical. You can find them on the websites of most business intelligence vendors. In their attempt to dazzle us visually, they fail entirely to present information in a way that can be understood at a glance.



It is because dashboards are rarely designed effectively that I wrote the book *Information Dashboard Design*. The principles are not difficult to learn, but they aren't obvious until someone points them out.

The fundamental design challenge



The fundamental challenge of dashboard design is to display all the required information on a single screen:

- clearly and without distraction
- in a manner that can be quickly examined and understood

Think about the cockpit of a commercial jet. Years of effort went into its design to enable the pilot to see what's going on at a glance, even though there is much information to monitor. Every time I board a plane, I'm grateful that knowledgeable designers worked hard to present this information effectively. Similar care is needed for the design of our dashboards. This is a science that few of those responsible for creating dashboards have studied.

Dashboards should be designed to support visual monitoring.

1. Scan the big picture



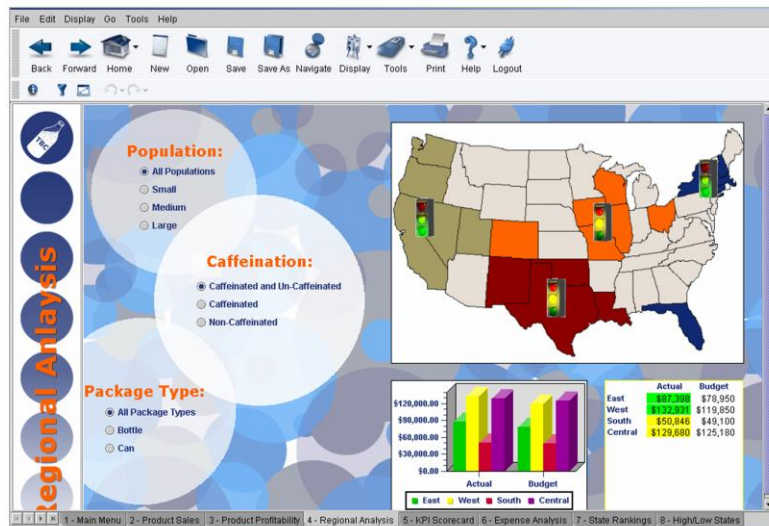
3. Link to supporting details



2. Zoom in on important specifics

The process of visual monitoring involves a series of sequential steps that the dashboard should be designed to support. The user should begin by getting an overview of what's going on and quickly identifying what needs attention. Next, the user should look more closely at each of those areas that need attention to be able to understand them well enough to determine if something should be done about them. Lastly, if additional details are needed to complete the user's understanding before deciding how to respond, the dashboard should serve as a seamless launch pad to that information, and perhaps even provide the means to initiate automated responses, such as sending emails to those who should take action.

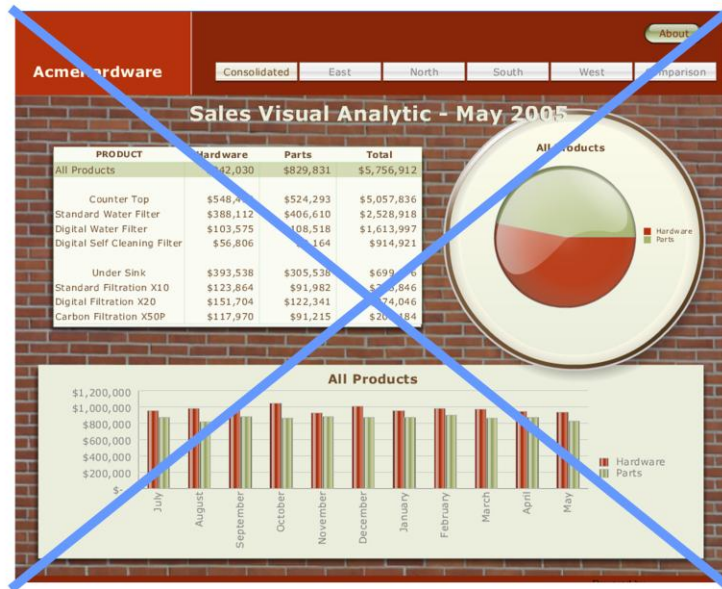
The problem with dashboards today



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

This failure is not rooted in technology; it is a failure of communication, rooted in poor design. The main problem that plagues dashboards today is that they don't say enough and what they do say they don't say very well. In other words, they are not very informative. When designed properly, a dashboard provides an overview of what's going on, clearly and rapidly.

Dashboards are not for show.



They're for communication.

Dashboards are not for show. No amount of cuteness and technical wizardry can substitute for clear communication. My purpose today is to present the design practices you cannot ignore if you wish to build dashboards that set the stage for informed response.

Typical dashboard graphs lack context.



YTD Units

YTD Units: 7,822

Measures of what's going on in the business rarely do well as solo acts; they need a good supporting cast to get their message across. To state that quarter-to-date sales total \$736,502 without any context means little. Compared to what? Is this good or bad? How good or bad? Are we on track? Is this better than before? The right context for the key measures makes the difference between numbers that just sit there on the screen and those that enlighten and inspire action.

The gauges above could have easily incorporated useful context, but they fall short of their potential. The center gauge tells us only that 7,822 units have sold year to date and that this number is good, indicated by the green arrow.

Quantitative scales on a graphic, such as those suggested by the tick marks around these gauges, are meant to help us interpret the measures, but they can only do so when scales are labeled with numbers, which these gauges lack. A great deal of the space that is used by these gauges tells us nothing whatsoever.

(Source of gauge image: Website of Informmersion Incorporated.)

Dashboards should tell us...

- Are we doing well or poorly?
- How well or poorly?
- What has led to what's happening today?

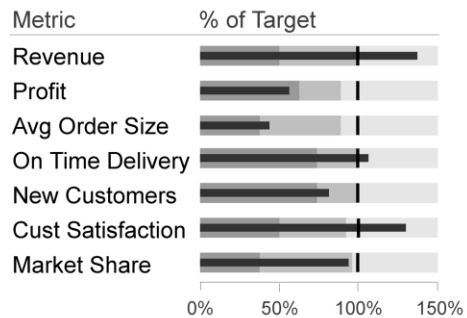
Many dashboards—perhaps most—are data poor. They report numbers, but don't express them meaningfully enough for people to gain the level of understanding needed to respond. To be effective, dashboards must be data rich. This is achieved by comparing measures of what's going on to related measures (for example, to a target). When monitoring what's going on, people usually need answers to these questions:

- Are we doing well or poorly?
- How well or how poorly?
- What has led to what's happening today?

The advantage of graphical displays

Metric	Variance to Target
Revenue	-0.68%
Profit	-17.90%
Avg Order Size	-5.90%
On Time Delivery	2.35%
New Customers	58.38%
Cust Satisfaction	-12.22%
Market Share	-13.33%

Text must be read,
processed serially



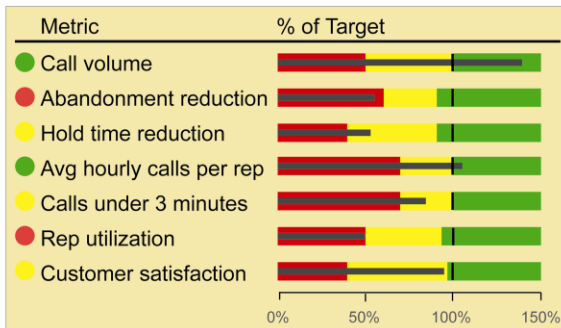
Graphics can be
perceived at a glance

The visual orientation of dashboards is important due to the speed of perception that's usually required to monitor information. The faster you must assess what's going on, the more you should rely on graphical means to display the information.

Text must be read, which involves a relatively slow, serial process.

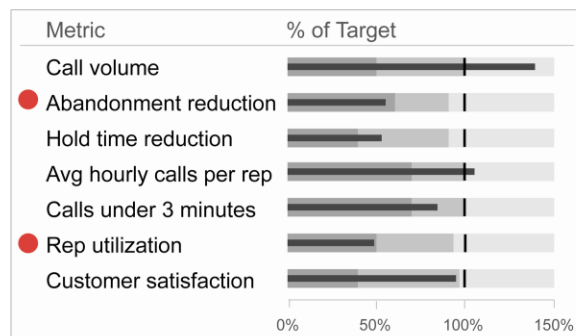
Certain visual properties, however, can be perceived at a glance, without conscious thought. With the graphical display on the right, it's quick and easy to see which bars exceed target, marked by the short vertical line, and which fall short.

Using too much color undermines its power



Avoid bright, saturated colors, except to highlight

Highlight sparingly



People tend to overuse color on dashboards, assuming that the brighter and more colorful they are, the better they work, but the opposite is true. People who understand color and how it can be used to display information use it sparingly.

Too much color is visually overwhelming; it tires our eyes. Also, if you use color gratuitously, you undermine its ability to be used to make things stand out. Notice how the red alerts clearly stand out in the bottom display in contrast to the neutral grays and blacks that been used elsewhere, rather than being lost in the meaninglessly colorful display above.

The key to effective data presentation

Eloquence through simplicity

To clearly present everything on a single screen, even the slightest lack of organization will result in a confusing mess. You must condense the information, you must include only what you absolutely need, and you must use display media that can be easily read and understood even when they are small, which is often necessary.

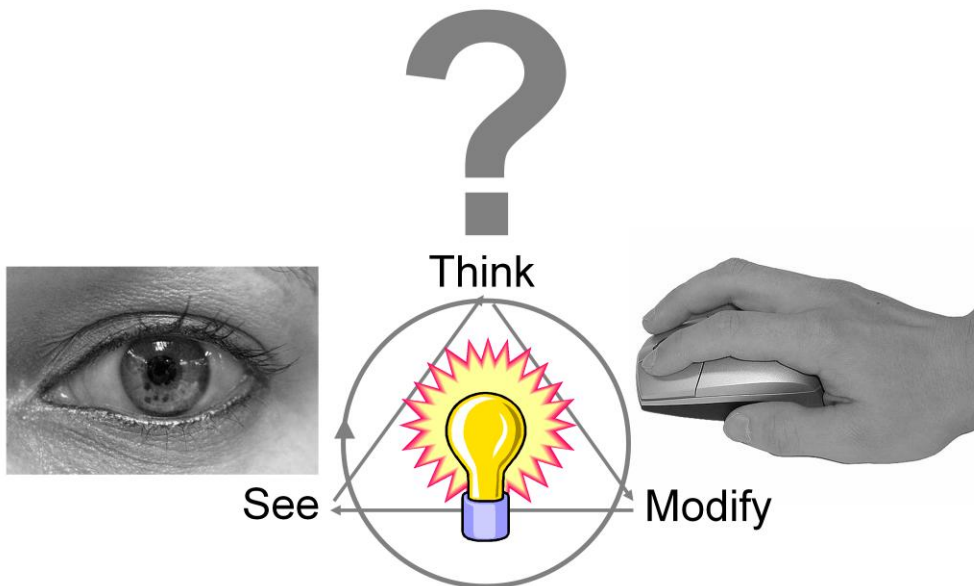
Elegance in communication can be achieved through simplicity of design. Too often we smear a thick layer of gaudy makeup on top of the data in an effort to impress, rather than to communicate the truth in the clearest possible way.

“Simplify, simplify, simplify.”

Henry David Thoreau

Think and Communicate

Analysis at the speed of thought



Direct dynamic interaction with the visualized data allows you to see something in the data visualization and interact with it directly to filter out what you don't need, drill into details, combine multiple variables for comparison, etc., in a way that promotes a smooth flow between seeing something, thinking about it, and manipulating it, with no distracting lags in between.

Demo: Effective visual analysis



Analytical navigation



Visual data analysis is a process that consists of many steps and many paths 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 stubbily 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)

Shneiderman's mantra

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

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.'

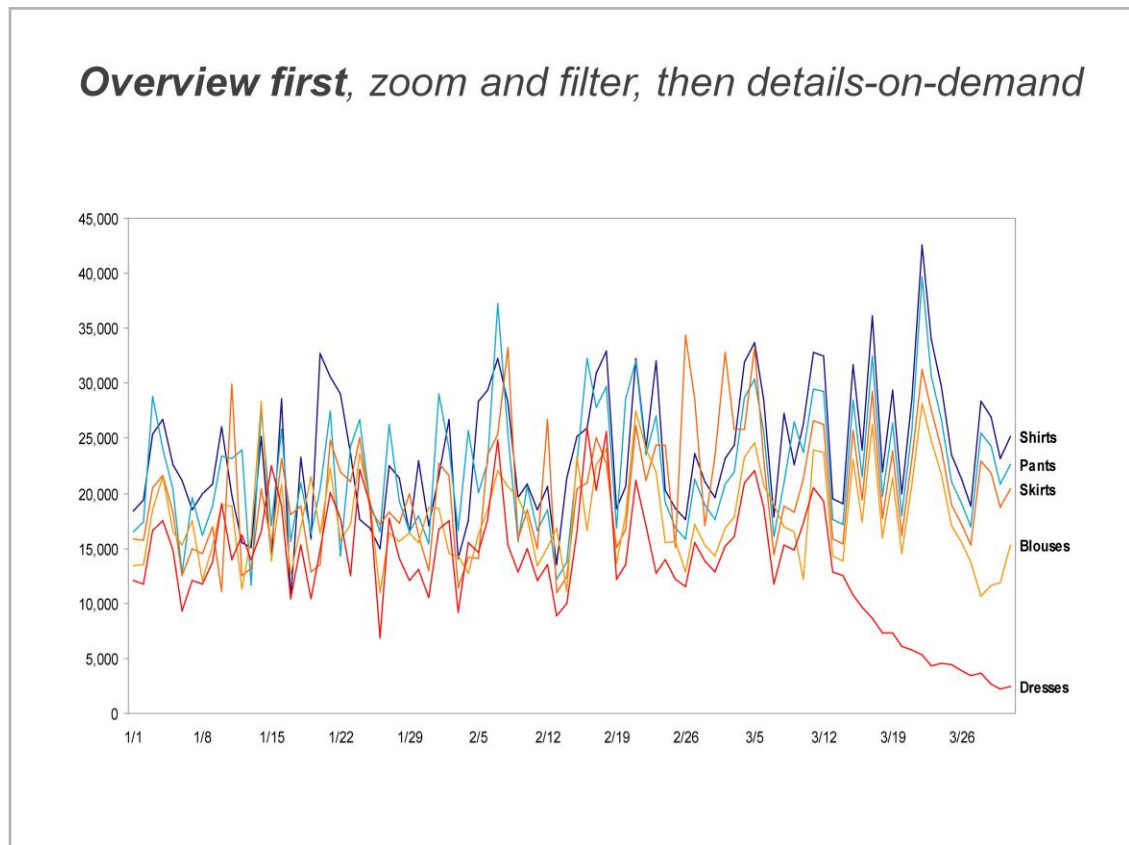
(Readings in Information Visualization: Using Vision to Think, Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman, Academic Press, San Diego, California, 1999, page 625)

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.

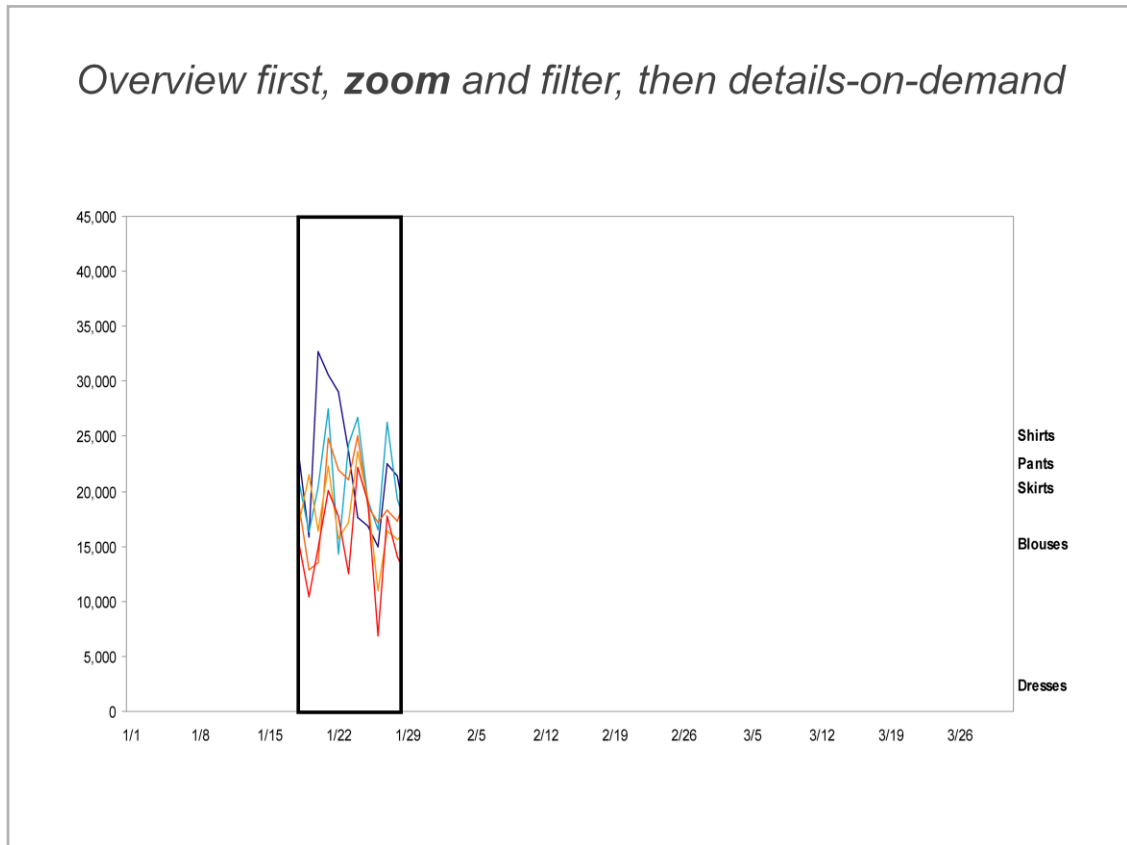
(Ibid., page 285)

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.

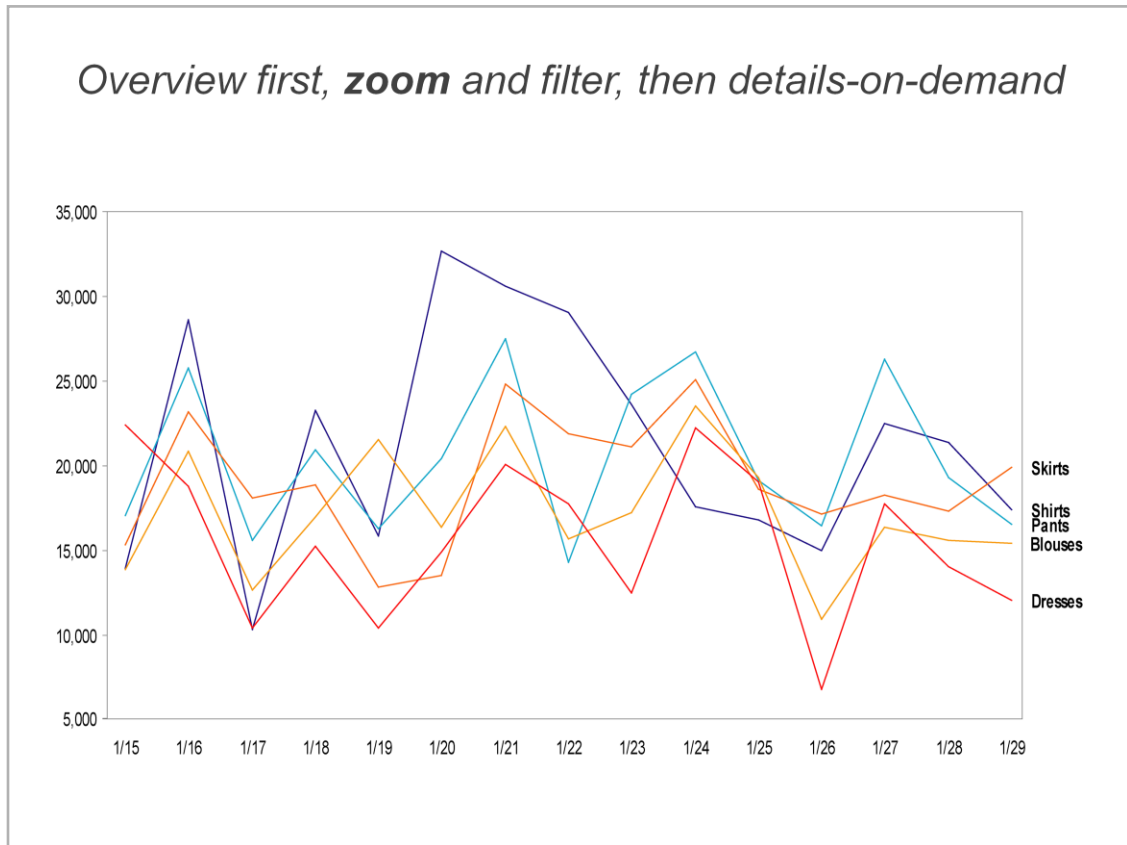
(Ibid., page 295)



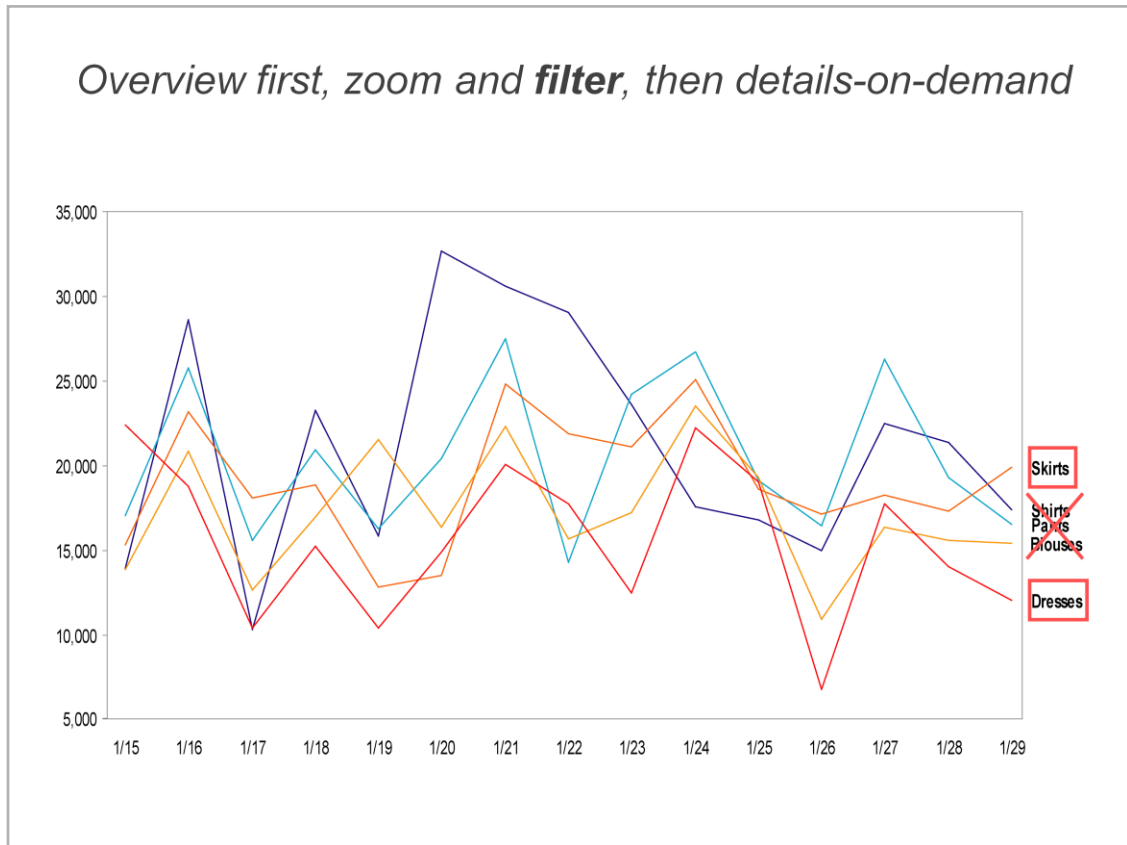
Shneiderman's technique begins with an overview of the data – the big picture. Let your eyes search for particular points of interest in the whole.



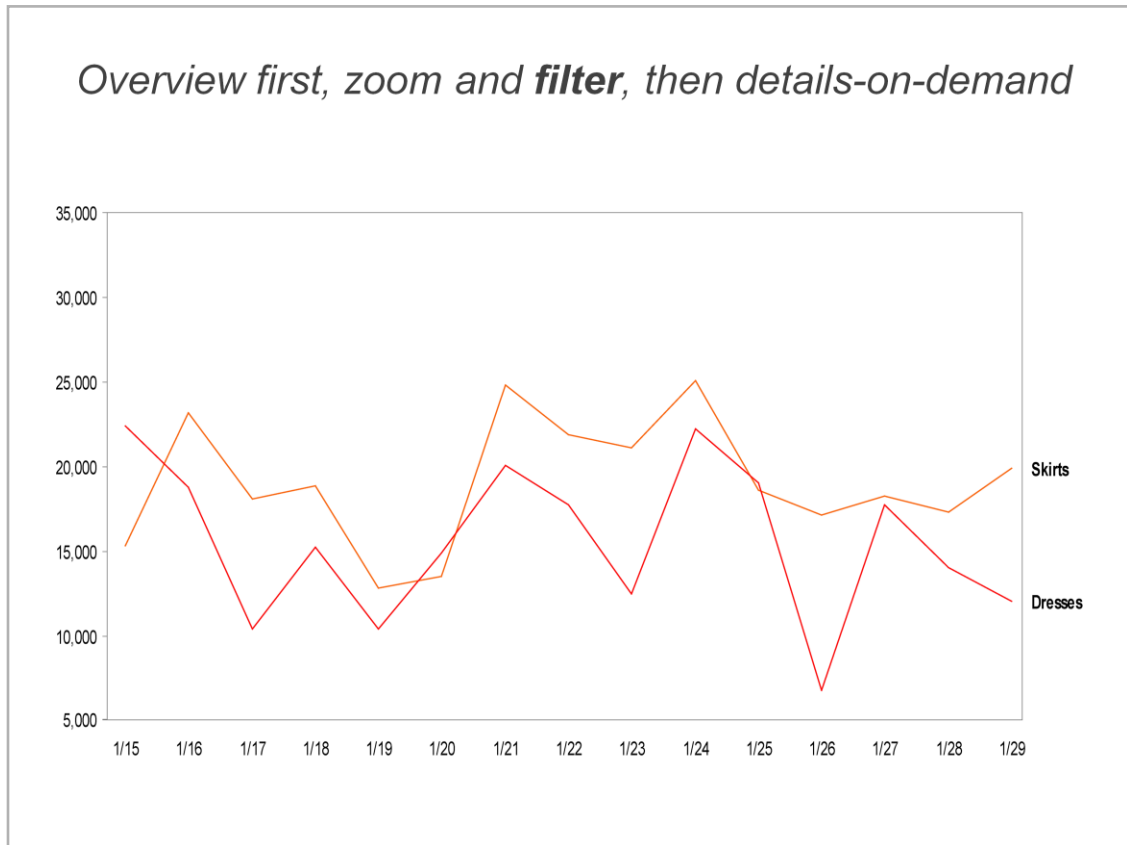
When you see a particular point of interest, then zoom in on it.



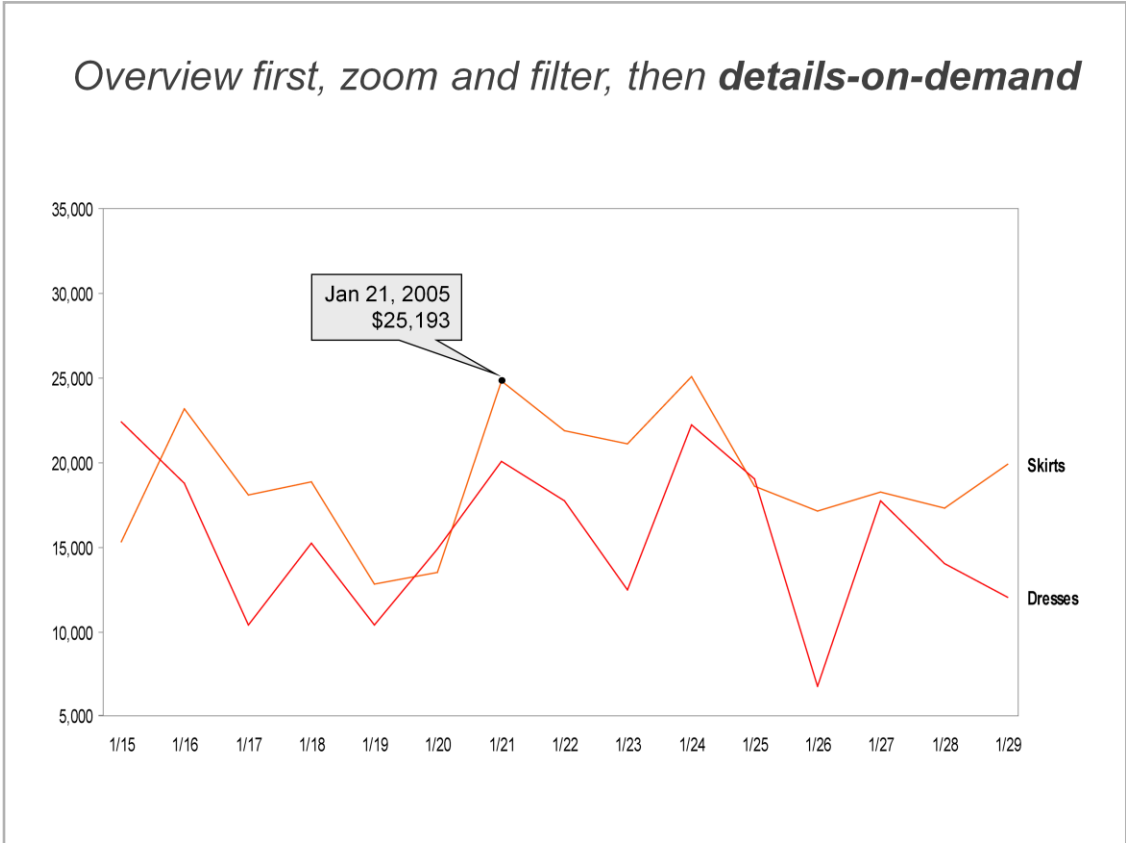
Once you've zoomed in on it, you can examine it more closely and in greater detail.



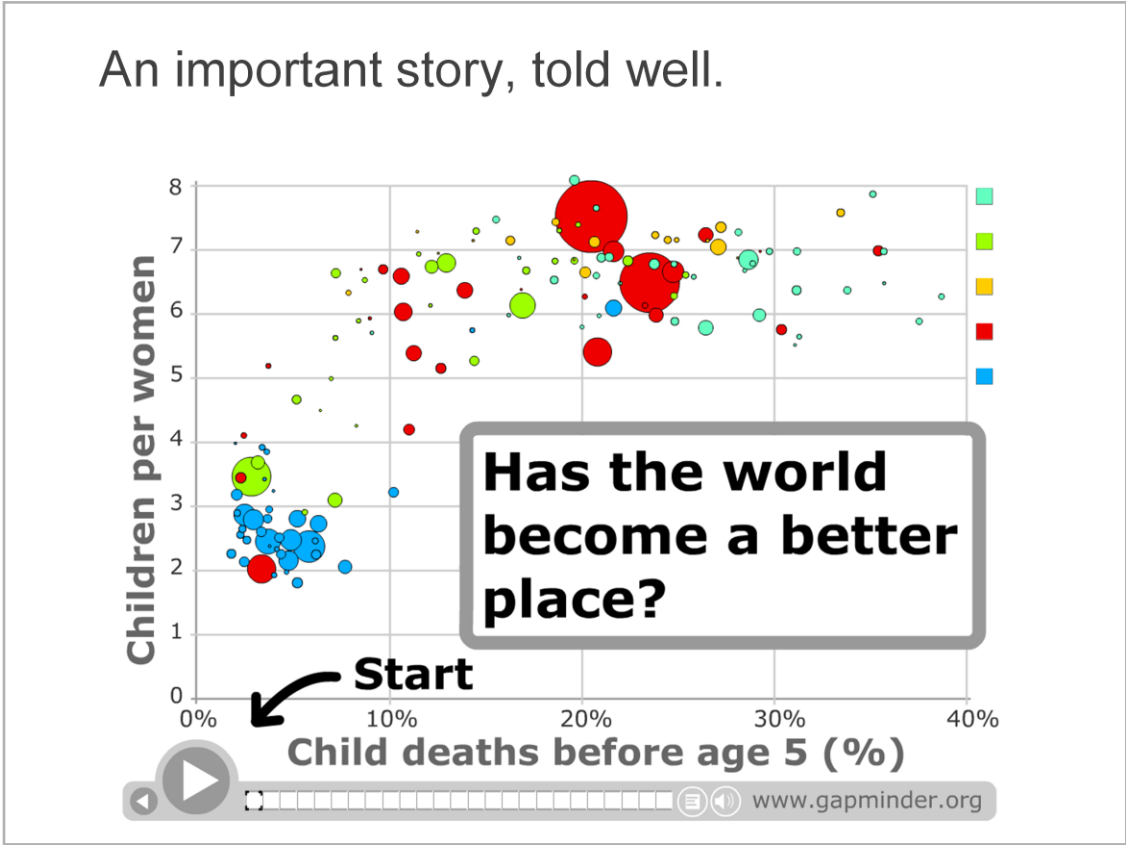
Often you must remove data that is extraneous to your investigation to better focus on the relevant data.



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



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.



Once you have the basics down, you can begin to tell more complex stories using more advanced techniques, such as this example from GapMinder.org, which uses an animated display to tell the story of how the correlation between the number of births per woman and mortality among young children throughout the world has changed through time.



The good news is, although the skills required to present data effectively are not all intuitive, they are easy to learn. The resources are available, such as my books *Show Me the Numbers* and *Information Dashboard Design*, but it won't happen unless you recognize the seriousness of the problem and commit yourself to solving it. It is up to you.