The Good, the Bad, and the Biased:

Five Ways Visualizations Can Mislead (and How to Fix Them)

Danielle Albers Szafir, University of Colorado Boulder

Insights

→ Visualizations allow people to readily analyze and communicate data. However, many common visualization designs lead to engaging imagery but false conclusions.

→ By understanding what people see when they look at a visualization, we can design visualizations that support more accurate data analysis and avoid unnecessary biases.

Data visualizations allow people to readily explore and communicate knowledge drawn from data. Visualization methods range from standard scatterplots and line graphs to intricate interactive systems for analyzing large data volumes at a glance. But how can we craft visualizations that effectively communicate the right information from our data? What aspects of data and design need to come together to develop accurate insights? The answer lies in the way we see the world: People use their visual and cognitive systems (i.e., our eyes and brain) to extract meaning from visualized data. However, flashy visualizations are not always optimized to help people see what matters. This article reviews common visualization practices that may inhibit effective analysis, why these designs are problematic, and how to avoid them. The discussion illustrates a need to better understand how visualizations can support flexible and accurate data analysis while mitigating potential sources of bias.

Glancing at the bar chart in Figure 1 will likely convince you that one method performs twice as well as the other. However, this visualization is misleading: The true difference between methods is only 5 percent. Talks and articles frequently feature flashy visualizations like this—visualizations

that, despite the data's simplicity, break several rules for honest and effective data visualization, exaggerating the differences between methods and calling into question the statistical conclusions drawn from the results. Are these violations nefarious? No. Are they done with the intention of making a cool graph? Probably. Do they lie with that data? Yes.

The mistakes made in this visualization—unnecessary use of 3D, a lack of uncertainty information, axes starting above zero—are common throughout the scientific world. People often justify these designs with comments like "I have learned to read these charts correctly" or "If I label my axes, no one will make that mistake." While there are small individual differences in how we interpret visualizations, everyone has the same visual system, is subject to the same visual biases, and can be fooled by the same visual illusions. And we are only fooling ourselves if we assume differently.

The choice to use flashy rather than accurate data visualizations is growing increasingly problematic. Data provides a crucial foundation for the decisions on which our society operates. It allows us to characterize the world in new ways and drive innovative discoveries. While algorithms and computational tools provide powerful mechanisms for harnessing data, interpretation and decision making are ultimately done by people. People bring context, expertise, and situational awareness to analyses that are not easily integrated into databases but that are critical to disentangling the signal from the noise. How can we as developers and data scientists enable access to the right information to support effective data analysis and communication?

The answer lies in understanding what people actually see in data visualizations. Our sense of sight provides us with a well-tuned pattern-recognition system. Centuries of evolution have refined our visual abilities to rapidly process large

amounts of complex information. We can find a tiger in long grass or ripe red berries in a bush. We can detect whether people are approaching or moving away. Visualizations leverage the high-throughput processing capabilities offered by our sense of sight to help people make sense of data. If we understand the patterns and information people extract from a visualization, we can enable people to draw informed conclusions from data at a glance.

Visualizations must be crafted with care, as we are easily tricked into seeing patterns in data that are not actually present, such as the 50 percent difference in Figure 1. While some visual biases and illusions are difficult to avoid, by understanding how information is transformed between the visualization and the knowledge it creates, we can encourage designs that help people better communicate, and ultimately understand, data.
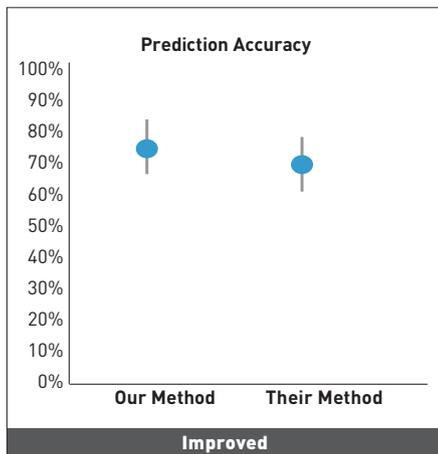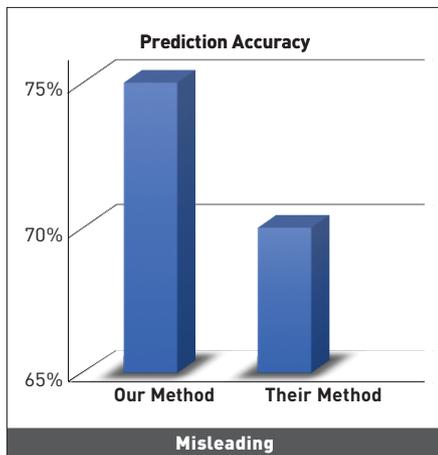
Figure 1. 3D marks, truncated axes, and other design choices create stylish visualizations; however, these visualizations are at best difficult to read and at worst lead to incorrect conclusions. Avoiding known bad practices leads to more honest and accurate data communication.

Here, we identify several (sometimes controversial) visualization design choices that can lead to potentially erroneous conclusions and offer solutions to overcome them, focusing on color choice, animation, axis scales, unnecessary 3D, and privileging statistics over data.

A PRIMER IN VISUALIZATION: WHEN, WHY, AND HOW

Visualizations are powerful tools for discovering and communicating insights in data. However, visualizations are not always necessary—people are not optimized to compute precise statistical quantities from abstract images. Many analysis problems can be solved with direct queries and algorithmic methods. For example, statistical models allow companies to optimize shipping procedures. Purely computational approaches scale further and more accurately estimate precise quantities than people. If you can distill what you need to know about your data into one computable value, you likely do not need a visualization.

However, visualizations often prove robust where statistics fall short. Visualizations take advantage of the universality of visual structure: We can see the shapes these data points make even when we cannot directly enumerate them. Take, for example, Anscombe's Quartet: four datasets with identical means, variance, correlation, and regressions (Figure 2). While these datasets appear statistically identical, visualizing them shows substantial qualitative differences in their structure. Our sight detects these high-level structures within 100 milliseconds of looking at a graph [1], far faster than the blink of an eye.

How do you decide when to visualize and when to compute? Factors such as uncertainty (how well do statistics represent the data?), transparency (what does the underlying data look like?), context (what additional knowledge could inform analysis and decision making?), scale (how many distinct quantities do we need to evaluate?), exposition (what story must the data tell?), and purpose (do we know what we are looking for?) all help determine when visualizations are valuable. For example, if you cannot readily quantify (or even know) what data properties matter, you can use visualizations to synthesize a diverse set of conclusions.
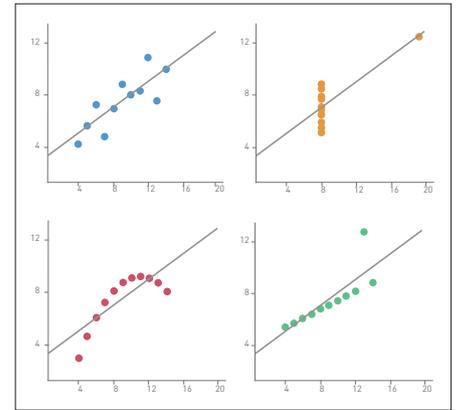
Figure 2. The four datasets of Anscombe's Quartet share the same basic descriptive statistics, but visualizing these datasets reveals four qualitatively different structures.

This trade-off between flexibility and precision is often the primary deciding factor for determining when a visualization is necessary: If access to the data underlying a statistic or prediction might change our decisions about that data, we should use a visualization.

Crafting visualizations generally follows a systematic process: clean the data, precompute relevant information, map that information to different visual channels (e.g., position, size, color), and integrate interaction and other details where appropriate. By combining a small number of channels, visualization designers can create intricate interactive systems that reveal patterns in large data collections at a glance. Choosing among these channels, while simple in concept, is where most visualizations go wrong. While many combinations create flashy and engaging graphics, these approaches may inadvertently obscure or even misrepresent data in ways that lead to flawed and biased interpretations. Misleading visualizations appear in our news reports, creating public mistrust in data, in scientific results, leading to incorrect theories, and even in Congress, where policymakers find themselves in conflict over data. So how do we avoid faulty visualizations? Science still cannot fully answer that question, but we can start by avoiding well-studied design pitfalls.

GETTING OVER THE RAINBOW

Many visualizations, such as geographic choropleth maps, eye-tracking heatmaps, and scalar field visualizations, represent data using

a familiar red-yellow-green-blue scheme referred to as the rainbow colormap. A longtime default of tools like MatLab, this colormap creates bright and engaging imagery that has led to incorrect conclusions and even retracted papers in top scientific venues. Many insist that the rainbow colormap allows them to interpret more variations in their data, as they have "learned to read the colormap correctly." However, a number of studies have proven that rainbow colormaps distort data even for people who use them daily. For example, researchers at Harvard worked with cardiologists who used rainbow colormaps to diagnose arterial disease [2]. Despite experts' insistence that they could accurately interpret rainbows, switching from rainbows to more mundane colors increased experts' abilities to correctly identify cardiac issues from 50 percent to 81 percent.

While in most cases using a rainbow colormap is not life-or-death, getting over the rainbow can improve data interpretation. Rainbows trick people into seeing false patterns in data. Color changes over rainbows are not uniform in their magnitude or direction, causing mismatches between perceived color differences and actual data differences. These mismatches distort value relationships and lead people to see data differences as being artificially smaller or larger than they actually are. For example, in Figure 3 the yellows appear far more similar to the oranges than to the equidistant greens.

Rainbows also cause people to visually group colors sharing the same name, such as shades of blue. In practice, this grouping makes rainbows useful for visualizing categorical data (e.g., apples and oranges). However, using rainbows for continuous values introduces artificial divisions in smoothly varying data. These divisions create false associations within grouped colors and dissociations between colors that bias what we see as same and different data. In Figure 3, we see clear bands of blues, greens, yellows, and reds, even though the data varies smoothly across the entire dataset. More appropriate colormaps overcome these biases by visually preserving relative data magnitudes.

Even if you consider yourself robust to the rainbow, consider that nearly 1 in 12 men are colorblind [3]. Colorblind individuals see the rainbow differently: They cannot discriminate between certain hues. This lack of discrimination does not just cause people to see reds and greens as the same but also shifts the perception of all hues by removing individual color components from each color in the rainbow. This shift further skews the mapping between color and data, leading to significant misperceptions and inaccessible data.

Tools such as ColorBrewer, Colorgorical, and Adobe Kuler offer principled alternatives to rainbows and allow you to tailor colormaps to best represent the visualized data types. If your data is categorical (e.g., dogs and cats), rainbows are fair game. However, ordered or continuous data should use either sequential or diverging colormaps. To choose between them, determine if there is a meaningful middle point in your data (e.g., differences from a baseline or natural zero value). If so, diverging colormaps (those that extend continuously from a neutral middle color) allow easy comparisons to that middle point. If not, sequential colormaps intuitively represent data magnitudes (Figure 3). By matching color to data, visualizations can avoid needless distortions that so often lead to false conclusions.

DATA ON THE MOVE

Many visualizations use animations. For example, a data point's velocity may represent its value. We visualize data at different time points in sequence to show change over time. Animated visualizations are flashy and engaging; however, they also blind people to important changes in data.

While we can use motion direction and velocity to encode data, people can distinguish only a handful of different speeds and motion directions [4] and can trace the specific movement of only three to four data points at a time [5]. Our limited abilities to track moving objects imply that representing data using motion may help us identify only a few high-level patterns with little sense of what those patterns mean.

These limitations are especially problematic for showing values changing over time. For example, Hans Rosling's GapMinder TED Talk [6] leverages animation to narrate changes in the global economy. Much of the power in this story lies in Rosling's ability to direct your attention to important changes in the data. However, our attention is a
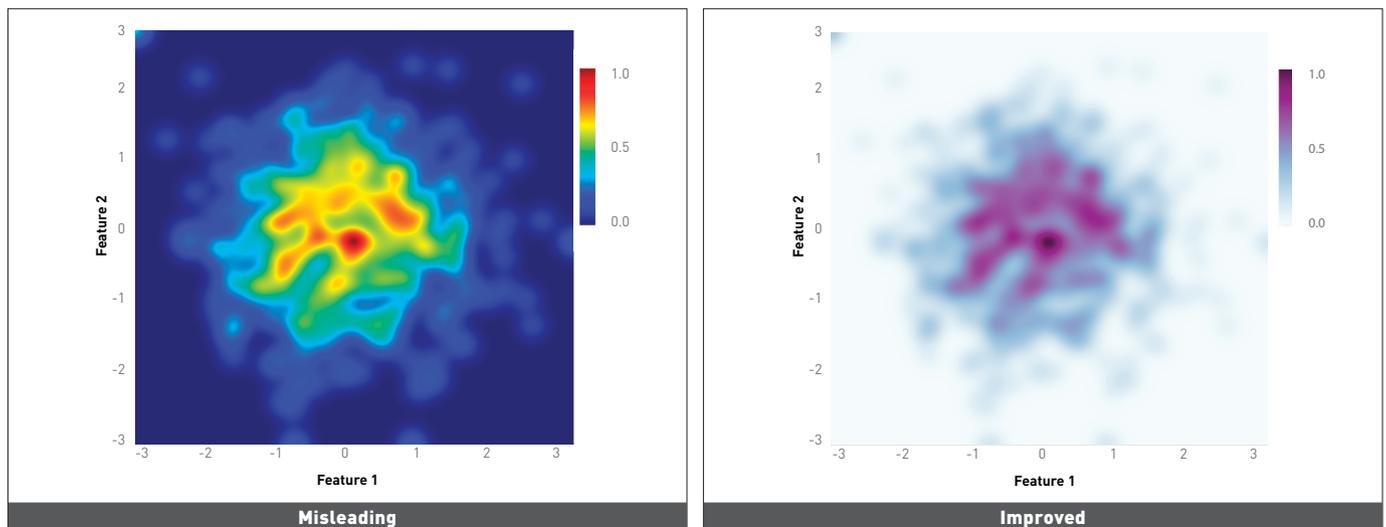
Figure 3. Rainbow colormaps make engaging figures but also create artificial divisions and skew value differences in ways that have caused innumerable false conclusions. Using a sequential colormap supports more accurate insights into smoothly varying datasets.
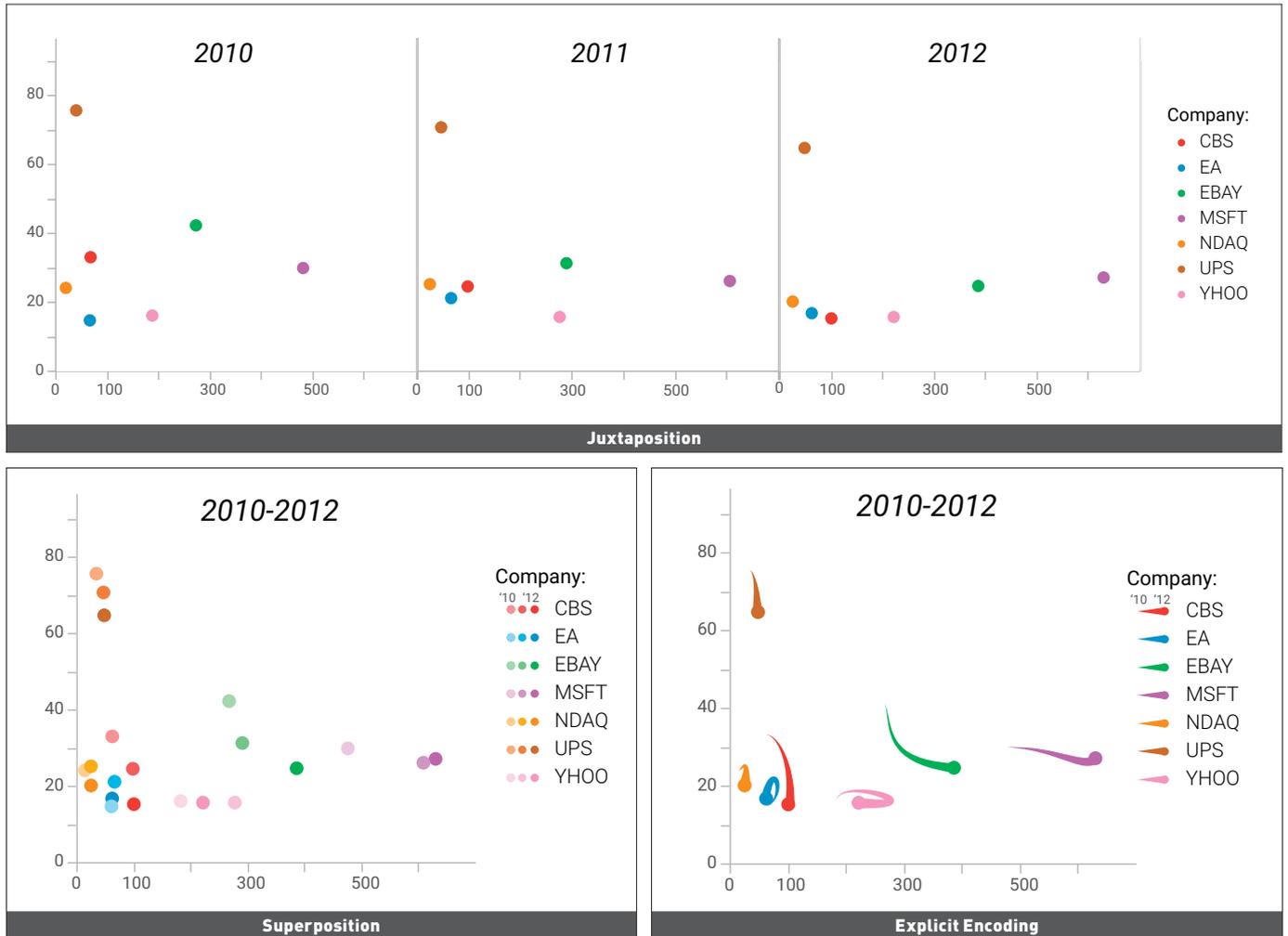
Figure 4. Animated data can leave people blind to important changes. Instead, consider methods for directly supporting comparison across time points.

scarce resource: We can allot a limited amount to any given set of data points. By directing our attention to one set of values, we effectively ignore changes in the rest of the dataset. As a result, animating your data over time may cause people to lose sight of most of the data.

This information loss is in large part due to change blindness, a phenomenon where attending to one change leaves us blind to others. For example, counting the number of times a basketball is passed causes us to miss a gorilla dancing through the passers

[7]. We can replace a conversant mid-discussion without notice [8]. In data visualizations, change blindness means that if we don't tell the analyst what aspects of an animated visualization to pay attention to, they may never see important changes in their data. Even if they see these changes, our limited memory prevents us from recalling precise differences over time.

We can overcome these limitations by directly visualizing how data changes over time. Methods for such temporal comparison fall into three categories (Figure 4): juxtaposition

(visualizing multiple time points side by side), superposition (arranging data from multiple time points on the same axes), and explicit encoding (directly visualizing the differences between time points). We choose between these different techniques by focusing on what aspects of change we want to highlight in our visualization and how many time points we need to see at any one time. Superposition facilitates precise and immediate comparison across a small number of time points; however, layering too many time points causes data points to occlude one another. Juxtaposition scales comparisons across larger datasets; however, it is difficult to precisely compare visualizations that are far apart. Explicit encoding can extract and represent salient information about changes over time, such as the trajectory a point follows on a scatterplot; however, these techniques require determining what differences matter for the analysis. By considering data scale and relevant questions, we

Our attention is a scarce resource: We can allot a limited amount to any given set of data points. By directing our attention to one set of values, we effectively ignore changes in the rest of the dataset.

can use these visualizations to compare changes over time without blinding people to critical changes in their data.

A MATTER OF SCALES

When we represent data on a standard Cartesian plane, many systems by default fit axis ranges to natural data scales, such as the minimum and maximum value. This choice maximizes the space in a graph dedicated to data. However, it also may cause people to see differences in the data that simply do not exist.

This issue is most problematic when visualizations begin their y-axes above zero. In many common visualizations, we interpret visualized values by measuring the distance between the x-axis and our marks (e.g., the top of a bar, the position of a point). Non-zero y-axes distort the difference between values, causing small differences to appear much larger than they truly are. Consider the example shown in Figure 1: The data difference is only 5 percent, yet the left bar appears twice as large as the right bar. Many argue that labeling axes counteracts the biasing effects of truncating the y-axis. However, people seldom read axis labels: The ratios people see at a glance often reflect the conclusions they will draw from data [9].

The same is true of normalized axes. If you have multiple consecutive plots showing the same variables, the axes should map to the same data ranges. Consider the infamous Planned Parenthood comparison chart [10]. The y-axis corresponds to the number of services provided; however, these axes are normalized to two different ranges, creating a false crossing in the data. Renormalizing these axes to the same scale tells a different story: At no point does the dominant provided service change. The most salient feature of the original graph led to a false conclusion because of improper normalization.

The distortion caused by poor axis scaling is a by-product of the way we read visualizations. Axis labels require conscious attention to interpret: We have to actively read these numbers to make sense of them. However, when we look at a visualization, we form the gist of a visual scene unconsciously. We get a sense of the data's shape and distributional properties without actively reading anything. If we use different axes to represent different facets of our data, the resulting shapes

and structures distort our perceptions of the data.

The one place where starting y-axes at values greater than zero is still a matter of debate is in communicating variation. For line graphs, small variations become less noticeable as the amount of space dedicated to those varying elements grows smaller; the magnitude of these small-scale variations becomes distorted by truncated axes. But if an analyst cares about variation rather than magnitude, many argue that the loss of fidelity from non-zero y-axes may be mitigated: The distortions created by the axis may not matter.

Instead of truncating your axes, consider the story the visualization is supposed to tell. What are the important differences in the data? For example, if you want to visualize change in a value over time, instead of communicating the raw magnitudes, you may wish to compute change relative to some baseline and visualize that computed value instead. To tell a story about growth or decline, visualize the rate of growth rather than the full population. By visualizing metrics more closely tied to the actual quantity of interest using honest axes, visualizations can focus on data that matters without introducing unnecessary bias.

THREE PROBLEMS WITH 3D

Three-dimensional visualizations create graphics that appear to pop out of the page. They are seen as engaging, futuristic, and sophisticated. And removing the ability to generate them is one of the best things presentation tools could do for honest data communication.

3D visualizations in two-dimensional media like slideshows and papers suffer from three primary issues that bias analysis: occlusion, projection, and perceptual ambiguity. Occlusion occurs when some marks make it difficult (or even impossible) to view others. Consider Figure 5: Center bars are occluded by outer values, complicating analysis. In the real world, people can move around objects to resolve occlusions. For example, we peek around a wall to see what lies behind it. In 2D, people generally cannot change their viewpoint to see occluded data. Occluded data is effectively lost. Even if people can move their viewpoint,

occlusion may prevent us from knowing where to look.

Our ability to resolve 3D objects stems from both monocular cues (e.g., one object being partially occluded by another) and binocular cues (e.g., information coming from each eye fused into a single picture). When we project 3D data onto a 2D image, we lose binocular cues. For example, we cannot engage motion parallax—the same depth cue that cats use when bobbing back and forth to judge how far to leap—or vergence—our brain's ability to resolve 3D position using the angles between an object and our eyes. As a result, 2D projections are inherently imperfect approximations of 3D space and are often difficult to resolve. For
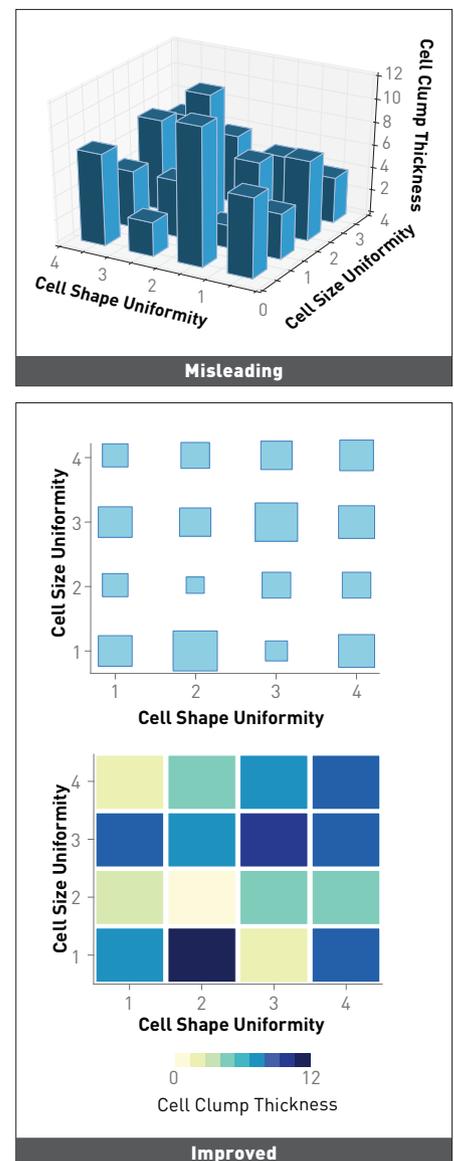
Figure 5. 3D bar charts can occlude data and distort values. Leveraging a third visual variable, such as size or color, supports more accurate comparisons over multiple dimensions.

example, when we tilt a pie chart in 3D, we distort the angles between slices of the pie (Figure 6) [11]. This distortion at best makes it harder to read the data and at worst causes incorrect analysis by distorting mark shape and size (and consequently perceived values) at different depths. These distortions worsen when we map data to size: As objects get farther away, they also appear smaller. In 3D visualizations, a small object may either have a small value or be far away. We cannot visually resolve the two possibilities.

To avoid occlusion and ambiguity in visualizations, use 3D only when absolutely necessary. Instead of representing the third dimension of your data using depth, try using alternative visual variables like color or size (Figure 5). Some kinds of data, like molecular surfaces or architectural structures, have inherent 3D shapes. In these cases, 3D can provide important contextual information. However, 3D is

still often imperfect for these scenarios. For example, we can see only half of any 3D volume from a single viewpoint. Pairing 2D summary representations with 3D structures can help overcome these limitations, even for complex geometries and inherently spatial data.

SHOW, DON'T TELL

As algorithms improve, it is tempting to rely on statistical processing for most data analysis. Visualizations increasingly represent the outputs of these processes rather than the original data. People often see algorithms as less error prone and unbiased; however, like people, algorithms are subject to bias and make mistakes. Electing to visualize algorithmic outputs without the context of the underlying data deprives people of the information necessary to evaluate the output's meaning and validity.

In collaboration with Microsoft and the University of Wisconsin, we surveyed the ways in which people visualize large collections of data. The majority of systems (74 percent) computed and directly visualized representative statistics [12]. While such statistical aggregation allows people to make precise claims about target quantities, it comes at the expense of context and flexibility. Consider a scatterplot comparing two clusters, A and B. If we choose to show the means of A and B, we have precise information about these means but have no data about other statistics of each cluster, such as the variance or density.

People can efficiently estimate aspects of a statistical distribution at a glance [13]. They can use visualizations to estimate properties of a distribution like means, variance, and even higher-order statistics like correlation quickly and accurately [14]. For example, within a half second of looking at a bubbleplot, we already have an approximate sense of the mean size of the collection of bubbles. Our abilities to visually compute these values relates to the concept of *ensemble coding*—a process our brain uses to compactly represent large quantities of visual information by recalling the data's distributional parameters.

When reasonable, visualizations should err toward providing more data rather than less. This design choice sacrifices precise statistical comparison in order to enrich analysis. There

are two primary cases where we may choose to explicitly aggregate data: when aggregate statistics are sufficient for our analysis and when we have too much data to visualize at once. In some cases, we may not need much data to address the question at hand. However, such visualizations should use caution when communicating statistics. For example, analysts often compare sample populations using bar charts with error bars. This method, despite its popularity, causes people to interpret values inside of a bar as statistically more likely than those outside of the bar, a phenomenon known as *within-the-bar bias* [15,16]. We can avoid this bias by using representations that provide more transparent insight into the data distribution. A violin plot (Figure 7) visualizes data distributions alongside means to help avoid within-the-bar bias; it also surfaces aspects of the data distribution that enrich analysis, such as the normal, bimodal, and skewed distributions in the figure's three samples.

Showing the full dataset is not always an option. Modern datasets may simply have too much data to visualize. Trying to show all available data can lead to clutter—we have so much visual information, we cannot find the data that matters. For example, network visualizations may gain so many connections that they become a "hairball": It is impossible to disentangle the individual relationships between entities in the graph. We can overcome clutter by carefully coupling statistics and visualization to construct visual summaries—visualizations that reduce the amount of data shown while preserving important properties of the distribution. For example, we can compute representative statistics for relevant subsets such as clusters or connected components. Alternatively, we can filter out irrelevant information to focus on relevant elements of the dataset. We can even randomly subsample our data, preserving the underlying data distribution while reducing the overall amount of information shown.

Balancing showing and telling in visualization is more of an art than a science, as we need to allow accurate and flexible analysis while not overwhelming people with too much information. Ideal visualizations should be transparent: People should understand how the data
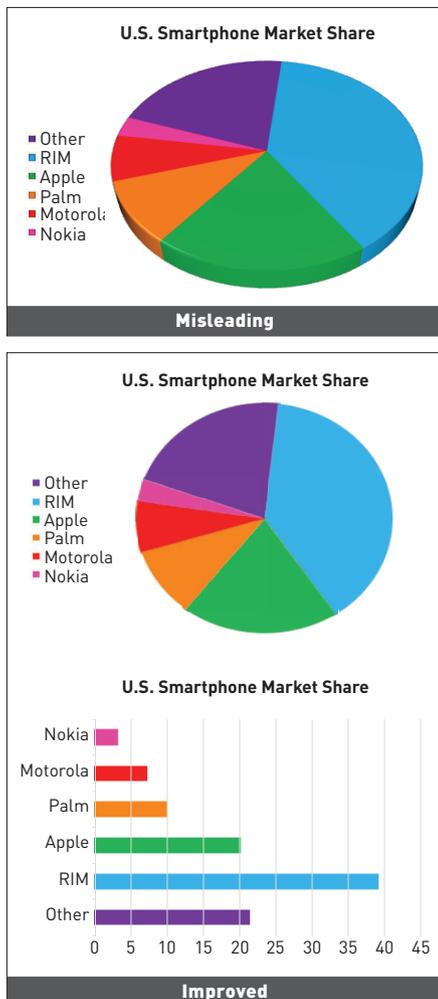
Figure 6. Distortion due to projection in the 3D pie chart causes the green wedge to represent a far larger market share than the data supports.
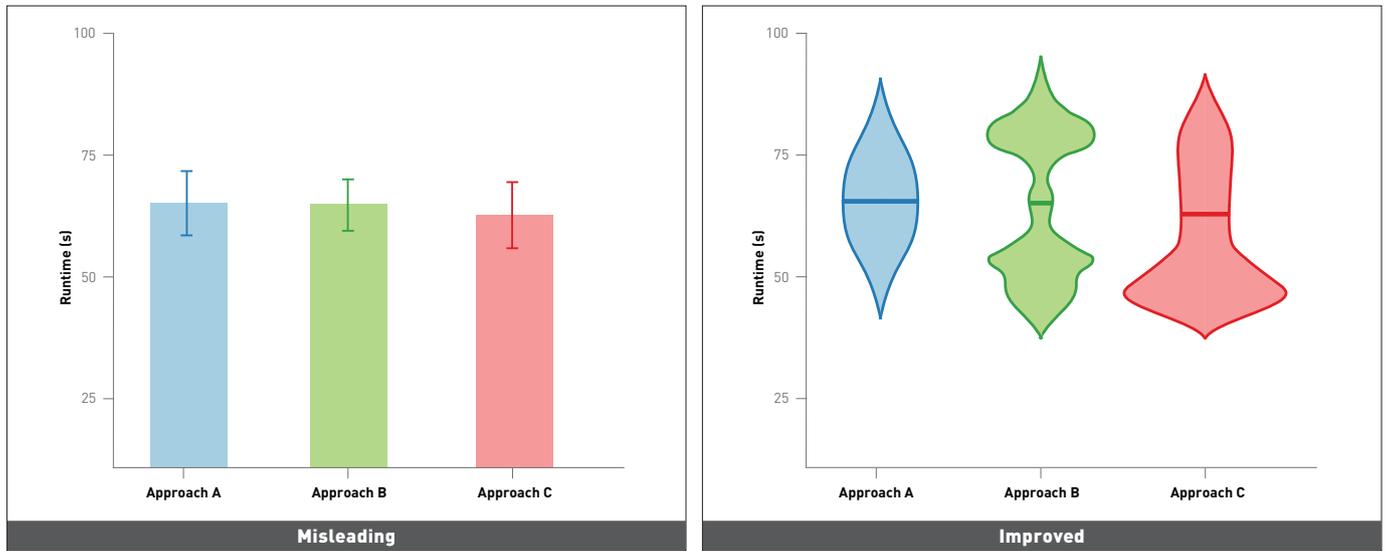
Figure 7. Traditional aggregation methods, such as bar charts encoding means, replace data with statistics, obscuring important patterns in the underlying data distribution.

changed between the raw, unprocessed file and visualized marks, and how the patterns they see reflect the underlying data. What statistics are used? What was filtered for? What happened to outliers? By being transparent with visualizations, we can help people better understand the available data and intuitively generate informed insights and decisions, even with large data collections.

TOWARD BETTER PRACTICES
This article focuses on common mistakes in visualizations that bias data analysis. These guidelines are deeply grounded in empirical studies and decades of observation and practice. Vision science and visualization offer some explanation for why these phenomena occur and allow us to design alternative representations that more faithfully depict data.

However, we are far from understanding all of the mechanisms at play when people interpret data. For example, how might visualizations account for illusions that occur naturally in data? Can we rescale or renormalize visualizations to account for biases introduced by the ways we see the world? How do we intuitively navigate high-dimensional data? How do we effectively pair visualization and computation to help people better leverage petabyte datasets?

A principled and quantified understanding of the way we see data can empower people to better leverage the many benefits offered by data. Crafting optimal visualizations is still an unsolved and wicked problem. Deeper collaboration between data science, cognitive science, and vision science is necessary to move us toward algorithmic and visual solutions that can scaffold an informed and inclusive data-driven society.

ENDNOTES
1. Larson, A.M., Freeman, T.E., Ringer, R.V., and Loschky, L.C. The spatiotemporal dynamics of scene gist recognition. *Journal of Experimental Psychology: Human Perception and Performance 40*, 2 (2014), 471.
2. Borkin, M.A., Gajos, K.Z., Peters, A., Mitsouras, D., Melchionna, S., Rybicki, F.J., Feldman, C.L., and Pfister, H. Evaluation of artery visualizations for heart disease diagnosis. *IEEE Trans. on Visualization and Computer Graphics 17*, 12 (2011), 2479–2488.
3. Wong, B. Points of view: Color blindness. *Nature Methods 8*, 441 (2011).
4. Ball, K., and Sekuler, R. A specific and enduring improvement in visual motion discrimination. *Science 218*, 4573 (1982), 697–698.
5. Franconeri, S.L., Jonathan, S.V., and Scimeca, J.M. Tracking multiple objects is limited only by object spacing, not by speed, time, or capacity. *Psychological Science 21*, 7 (2010), 920–925.
6. https://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen
7. Neisser U. The control of information pickup in selective looking. In *Perception and Its Development: A Tribute to Eleanor J. Gibson*. A.D. Pick, ed. Erlbaum, New York, 1979, 201–219.
8. Simons, D.J. and Levin, D.T. Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review 5*, 4 (1998), 644–649.
9. Pandey, A.V., Rall, K., Satterthwaite, M.L., Nov, O., and Bertini, E. How deceptive are deceptive visualizations?: An empirical analysis of common distortion techniques. *Proc. of the ACM Conference on Human Factors in Computing Systems*. ACM, New York, 2015, 1469–1478.
10. http://www.msnbc.com/msnbc/congressman-chaffetz-misleading-graph-smear-planned-parenthood
11. https://www.wired.com/2008/02/macworlds-iphon/
12. Sarikaya, A., Gleicher, M., and Szafir, D.A. Design factors for summary visualization in visual analytics. *Computer Graphics Forum 37*, 3 (2018).
13. Ariely, D. Seeing sets: Representation by statistical properties. *Psychological Science 12*, 2 (2001), 157–162.
14. Szafir, D.A., Haroz, S., Gleicher, M., and Franconeri, S. Four types of ensemble coding in data visualizations. *Journal of Vision 16*, 5 (2016), 1–19.
15. Correll, M. and Gleicher, M. Error bars considered harmful: Exploring alternate encodings for mean and error. *IEEE Trans. on Visualization and Computer Graphics 20*, 12 (2014), 2142–2151.
16. Newman, G.E. and Scholl, B.J. Bar graphs depicting averages are perceptually misinterpreted: The within-the-bar bias. *Psychonomic Bulletin & Review 19*, 4 (2012), 601–607.

⬤ **Danielle Albers Szafir** is an assistant professor in the Department of Information Science at the University of Colorado Boulder. Her research bridges data science and vision science to develop interactive visualization systems, guidelines, and techniques for exploratory data analysis.
→ danielle.szafir@colorado.edu