Lying with Data

By

Jordan Ellenberg

June 26, 2015 10:16 a.m. ET

A number has a way of ending an argument. What can you say to it? There’s no nuance, no room for interpretation—it is what it is.

Unfortunately, numbers turn out to be a lot like words: powerful and illuminating but capable of being deployed to bad ends. Here’s a little manual of some of the most common ways that data, for all its precision, can take you down a wrong path.

**Failure to compare.** Last fall, New York Gov. [Andrew Cuomo](http://topics.wsj.com/person/C/Andrew-Cuomo/5961) crowed, “The news that our unemployment rate has dropped to its lowest since 2008 is proof that New York is on the move.” What Mr. Cuomo said about the unemployment rate was true: [Only 6.2% of New Yorkers were unemployed](https://www.governor.ny.gov/news/statement-governor-cuomo-new-yorks-unemployment-rate-falling-62-lowest-october-2008) in September 2014. But he didn’t mention another number: the overall U.S. unemployment rate, which stood at 5.9%—also the lowest since 2008. If New York is on the move, it is moving at the same speed as the country as a whole. A number by itself is often meaningless; it is the *comparison* between numbers that carries the force. (Gov. Cuomo’s office didn’t respond to a request to comment.)

My favorite example of this lapse came from the blogger Vani Hari (aka “Food Babe”), [who warned her air-traveling readers in 2011](http://www.freezepage.com/1415667665TBMRBWICKU), “The air that is pumped in [to an airplane cabin] isn’t pure oxygen either, it’s mixed with nitrogen, sometimes almost at 50%.” Almost 50% adulteration sounds terrible—until you remember that the natural proportion of nitrogen in Earth’s atmosphere is 78%. (Yvette d’Entremont wrote about the mistake on Gawker; Ms. Hari has pulled down [the offending post](http://foodbabe.com/2014/12/06/food-babe-critics/).)

**Unrepresentative representative.** Suppose you give college students around the world a values questionnaire, asking them (for instance) to agree or disagree with the statement, “Making a lot of money is a high priority for me.” Now suppose that 35% of American students strongly agree—the highest proportion in the developed world. Does this mean that American capitalism has soured our youth into nasty greedheads?

Maybe, maybe not. Questionnaires have lots of questions. Maybe this one also included items like, “Material comfort is more important than personal fulfillment” or “I would sedate children and sell their kidneys if it got me into a higher tax bracket,” on which the U.S. was in the middle of the pack.

When you’re telling a story, it’s natural to pick the most vivid and persuasive detail. In this case, it was the question on which the answers of U.S. students represented an extreme. But providing the impressive number without conceding the existence of the unimpressive ones is a kind of numerical malpractice.

**Needle in the haystack.** A closely related trick is to pull out the most exciting finding—the needle—from a scientific study that is mostly a big heap of hay. [A 1998 study in New Zealand](http://www.ncbi.nlm.nih.gov/pubmed/9648329) asked: Did a serious fall on the playground make children more fearful of heights later?

A fall before age 5 had no effect on fear at age 11 or fear at age 18. A fall between ages 5 and 9 also had no effect on fear at age 11. But a fall between 5 and 9 was associated with a *reduced* fear of heights at age 18! The story got reported like this: Make sure to expose your kids to dangerous playground equipment because if they never break a bone, they might grow up to be cringing wimps. In this telling, the hay is gone, leaving only a neat moral lesson with a needle’s sting.

**More is more.** The U.S. now has about as many bank tellers as it did in 1980. Does this mean that such work is immune to tech-driven obsolescence? No: The U.S. population has increased by 40% over this period, which suggests that tellers are less in demand and make up a smaller share of the workforce.

It doesn’t make sense to directly compare the number of tellers then to the number of tellers now, any more than it makes sense to compare the box-office take of “The Sound of Music” (1965) and “The Croods” (2013) by unadjusted dollar value. “The Croods,” by this measure, was the bigger hit, taking in some $187 million to the von Trapp family’s $159 million. (No, I don’t know what “The Croods” is about either.)

All these mistakes have one thing in common: They don’t involve any actual falsehoods. Still, despite their literal truth, they manage to mislead. It is as if you said, “Geraldo Rivera has been married twice.” Yes—but this statistic leaves out 60% of his wives.

In the era of data journalism, truth is not enough. We need people in the newsroom who can check not only a number’s value but also its meaning. Unless we can ensure that, we’re going to be reading a lot of data-driven stories that are true in every particular—but still wrong.

*—Dr. Ellenberg is a professor of mathematics at the University of Wisconsin, Madison, and the author of “How Not to Be Wrong: The Power of Mathematical Thinking,” recently released in paperback by the Penguin Press.*