Statistical Literacy for Policy Makers

Milo Schield
Augsburg University

Abstract: Information literacy, data literacy and statistical literacy overlap when they deal with data as evidence in arguments. All three require analysis and evaluation. To effectively evaluate data as evidence, policy makers need to untangle data and social statistics from arithmetic numbers. Social statistics are numbers in context -- where the context matters. Social statistics are socially constructed. They are generated, selected and presented by people with motives, values and goals. Social statistics can be influenced. Policy makers need to evaluate quantitative evidence using the same skills they use in evaluating non-quantitative evidence. Ask questions! This paper presents seven simple questions.

Keywords: information literacy; data literacy, decisions

1. Introduction:

Data literacy, information literacy and statistical literacy are new areas for many policy makers. Schield (2004A) argued that there is a substantial overlap between information literacy, data literacy and statistical literacy. The results of all three require analysis and evaluation. Analyzing and evaluating are things that policy makers do. Yet these may seem like new situations when they involve statistical data.

Most policy makers are used to facing new situations: situations for which they have little prior experience. To handle such situations, they ask questions. Sometimes, the questions may seem stupid, but this is how policy makers get the information they need. Unfortunately policy makers are not used to asking questions about data. In order to ask good questions about statistics, policy makers need to know something about statistics.

Policy makers need to recognize the vast difference between statistics and numbers. Numbers are more like book-keeping: arithmetic operations that don't involve assumptions or choices. Statistics are different – very different. Statistics deal with reality. It is easier to lie, to mislead or to prevaricate with statistics. Statistics are more like finance or economics: Statistics – certainly social statistics – involve assumptions and choices. To understand statistics, you need to recognize seven things:

1. *Statistics are numbers in context – where the context matters.* In arithmetic, 1 + 1 = 2. But in statistics, the reality involved makes a difference. In bunny statistics, adding one bunny and one bunny can yield more than two bunnies. In ice-cube statistics, adding one ice-cube and one ice-cube can result in no ice-cubes (in a hot cup of coffee). A company may have a 60% market share in the Eastern US and a 70% market share in the Western US. What is their market share in the entire US? 130%? Hardly. Here, along with geography, the words – market share – really matter.

2. *Statistics are socially constructed.* This is most obvious with social statistics. Most people don't realize how many choices are involved in constructing a social statistic. Best (2002) argued that social statistics are socially constructed:

   Statistics "*are the products of social activities. There's a tendency in our culture to believe that statistics – that numbers – are little nuggets of truth, that we can come upon them and pick them up very much the way a rock collector picks up stones. A better metaphor would be to suggest that statistics are like jewels; that is, they have* *to be selected, they have to be cut, they have to be polished, and they have to be placed in settings so that they can be viewed from particular angles. All of that is work that people do. Any number that we come across is evidence that somebody has gone about counting something. That means that it was a social process, that there were people involved, that somebody, for some reason, in some way, counted something and produced a number. This may seem obvious, but that obvious process tends to get ignored most of the time when we talk about statistics.*"

3. *Statistics can be influenced – just like words.* Just because a statistic is true, doesn't mean that's the end of the story. All-too-often, there's a story behind the statistic. The influences can be grouped under the four letters in CARE.
   C *Confounders* can influence statistics. Did you know that people who read home and fashion magazines are 10 times more likely to get pregnant than people who read car and sport magazines? This statistic is confusing. You know pregnancy has nothing to do with magazines. But you guess that women are more likely to read home and fashion than to read car and sport. Not taking into account gender confuses the connection between magazines and pregnancy. Statisticians would say that gender was a confounder
   A *Assembly* can influence statistics. They can be manipulated by how they are assembled: how they are defined, counted, measured, compared and presented. In 1998, the National Institutes of Health (NIH) lowered the overweight BMI score for overweight. Overnight 30 million Americans suddenly became overweight.
   R *Randomness* can influence statistics – even in big data. The bigger the data, the more unlikely the coincidences that emerge. C.f., the numerology associated with 9/11 and the unlikely similarities between the Lincoln and Kennedy assassinations.
   E *Error* or bias can influence statistics. Shoppers outside Costco are more likely to have a favourable opinion of Costco than shoppers outside Walmart: selection bias.
   The best advice in dealing with statistics – specifically social statistics – is to "Take CARE". The four letters of care indicate the four kinds of influence on social statistics: C for Confounding, A for Assembly, R for Randomness and E for Error or bias.

4. *Association is not [necessarily] causation.* People who shave their face are generally taller than people who shave their legs. The truth of this comparison does not prove that you should shave your face if you want to get taller.

5. *Disparity is not [necessarily] discrimination*. Men are about 90% of prison inmates – but only 50% of the population. The disparity provides evidence of sexual discrimination against men, but it does not prove discrimination – even though this is a big disparity.

6. *Ratios and average may be confounded.* The best research hospital may have the highest patient death rate. One explanation is that the people, procedures and facilities at the research hospital are the cause. Another is that the best research hospital handles the sickest patients.

7. *The bigger the effect size, the more resistant an association is to being nullified or reversed by a confounder*. There was no way to conduct a scientific experiment to determine if smoking caused lung cancer. But smokers were 10 times as likely to develop lung cancer as non-smokers. This large effect size effectively neutralized all known confounders and allowed decision-makers to conclude that smoking caused cancer. The converse is this: the smaller the effect size, the more likely an association is to being nullified or reversed by a confounder.

Statistical literacy studies *critical thinking about statistics as evidence in arguments*. Most of these are social statistics. Schield (2010A) provides an overview. This paper provides some highlights.

2. Questions

Policy makers typically deal with statistical summaries obtained from large amounts of data for a group or time period. Policy makers need to focus on a few simple questions. Here are seven.

1. *How big? How much? How many?*

These simple questions can reveal a lot. Sometimes, the actual effect size is never mentioned. The simplest way for others to avoid size is to indicate a direction. "The more soda consumed, the greater the person's weight." Or "Accounting majors who took a review course were more likely to pass the CPA than those who did not." Another way for others to avoid size is to use 'many' or 'often': "Many scientists believe that much of global warming is man-made" or "All too often, those in prison are school dropouts (leavers)."

If possible, ask a follow-up question: "By how much?" If asking a question is impossible, you have little reason to consider the argument. Think of why the size wasn't given. Direction or quality words are convenient distractors in place of a quantity when the size is small. If it were big, they would tell you. Without a size or quantity, there is no basis for saying a statistical difference or change is important. Without numbers, statistical comparisons may be suggestive, but unimportant. Without numbers, statistical comparisons are almost meaningless.

2. *Compared to what?*

A given statistic may seem small or big. But without comparing it to something relevant, it is difficult to analyse or evaluate it. The simplest case involves a count or total. Suppose you are told that Florida had 487,000 unemployed.[1] By itself, one has no idea of whether that is big or small. But if we know that New Jersey just had 333,000 unemployed, now we have some basis of comparison. But why was New Jersey chosen? Was it to make Florida's count look big or small?

Whenever you are given a comparison, think about what other comparison might have been given. Why was this comparison given rather than others? Everyone has motives, values and interests. Everyone has an agenda. Statistics are no different than words. People choose the words and the statistics that best support their agenda.

3. *Why not a rate?*

We don't need higher math to know that rates can control for the size of a group. When given a count, ask "Why not a rate?" In April, 2020, the number of unemployed workers was 1,218,000 in Florida. Is this big or small? Rates provide an internal comparison. The unemployment rate in Florida in April 2020 was 12.9%. Now we have some context.

Even though rates may be better than counts, a rate has just slightly more context than a count.

---

[1] https://www.bls.gov/news.release/laus.t01.htm

New Jersey had a Covid19 death rate of 25.6 deaths per 1,000 cases as of May, 2021. You would not know that New Jersey had the highest Covid19 death rate per case of any state in the US. Lithuania has a death rate of 14.6 per 1,000 population as of May, 2021. Without a comparison you would not know that Lithuania has one of the highest Covid death rates in the world.

We can combine "Why not a rate?" with "Compared to what?" We can compare rates of two groups at the same time. We can compare rates at two times for the same group;

- In April, 2020, Florida had a 12.9% unemployment; New Jersey had 15.3%
- Florida's April unemployment rate was lower in 2021 (4.8%) than in 2020 (12.9%).

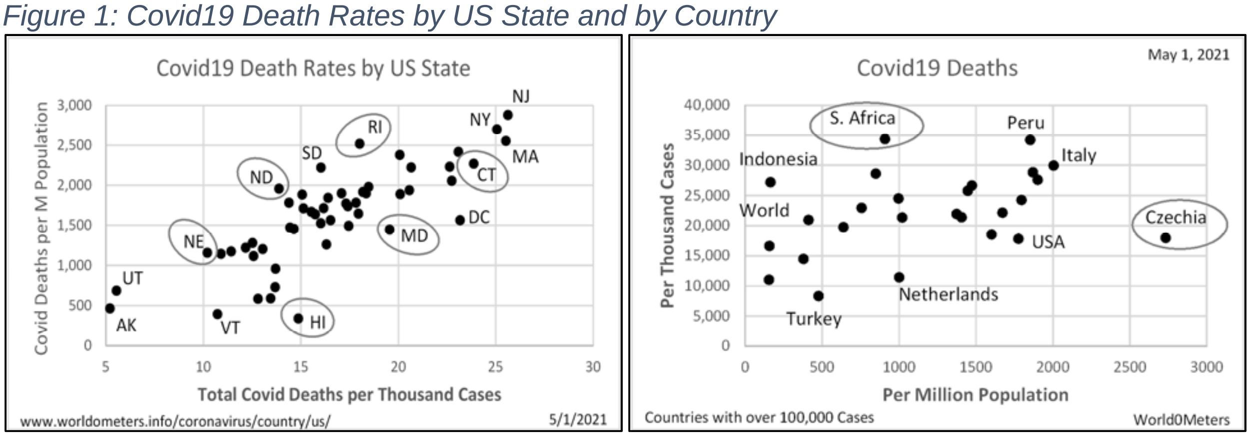
When given a comparison of counts, ask why not a comparison of rates. In April, 2021, the number of unemployed workers was bigger in Florida (487,000) than in New Jersey (333,000). But, the unemployment rate was smaller in Florida (4.8%) than in New Jersey (7.5%).

4. *Per what? The diabolical denominator.*

Sometimes there is choice in selecting the denominator in a ratio. The choice of the denominator can influence the size of a comparison and even change its direction. Compare the Covid19 death rates by state using two different denominators.

On the left side of Figure 1, the Covid19 death rate was higher in Rhode Island than in Connecticut *per capita* or person (on the vertical axis), but it was lower *per case* (on the horizontal axis).

Figure 1: Covid19 Death Rates by US State and by Country



On the right side of Figure 1, the Covid19 death rate is higher in South Africa than in Czechia per case (on the vertical axis), but lower per capita or person (on the horizontal axis). Why the difference? The relationship between the denominators (cases and population) differed for the two groups. See Schield (2021) for details.

A second example: Can the birth rate per 1,000 women ages 15-44 be going up over time while the birth rate per 1,000 women (all ages) be going down – for the same population over the same time period? Yes! The birth rate among potentially fertile women (15-44) might go after a depressions or war. The birth rate among all women could go down if women are living longer. The choice of the denominator can change the direction of a comparison.

Here's a third: The 1996 auto death rate was 7 times as high in Hawaii (35) as in Arkansas (7) per mile of road, but twice as high in Arkansas (38) as in Hawaii (19) per vehicle.[2]

The choice of the denominator is important in comparing crime statistics. Consider comparing convictions by race or gender. Are we comparing convictions per person, per victimization, per reported crime or per arrest? Four different choices can give four different comparisons.

5. *How were things defined, counted or measured?*

For years, Cuba was touted as having a very low infant mortality rate: infant deaths per 1,000 births during the first year after birth.

In 2017, the infant mortality rate (IMR) per 1,000 live births was supposedly lower in Cuba (4.1) than in Canada (4.5) or the US (5.7). [3]

A simple way to arrange this is to count a few of the deaths in the first month after birth (neonatal) as late pregnancy deaths (late fetal). With around 100,000 births per year, the stated difference of less than 2 infant deaths per 1,000 live births (5.7 minus 4.1) would involve reclassifying less than 200 early-post pregnancy infant deaths as late pre-pregnancy deaths.

6. *What was taken into account (what was controlled for)?* Is this a crude association?

Suppose that a decision maker is told that Mexico has a better health care system than the US. You might ask, "What is the evidence?" The death rate per thousand population is lower in Mexico (5.2) than in the US (8.2). The rates take into account the difference in population size between Mexico and the US. But these rates are still *crude statistics*; their association is a crude association. These rates didn't take anything else into account that is relevant.

What else would be relevant in comparing death rates? Age! Older people are much more likely to die than younger people. Mexico has a much younger population than the US. This effect is clearly seen when you find that the death rate per thousand population is lower in Gaza (3.5) than in Mexico (5.2). The idea that Gaza has better health care than the US is all but unthinkable. For more on taking things into account, see Schield (2004B).

7. *What else should have been taken into account (controlled for)?*

This question involves hypothetical thinking. Hypothetical thinking does not require a high IQ or knowledge of advanced math. Just asking some very simple – but practical – questions, can open up some very helpful discussion. Hypothetical thinking does require some knowledge of the situation. In the case of the death rates for the US, Mexico and Gaza, you need to know that older people are more likely to die than younger people.

Suppose you asked those comparing the death rates of countries if they had taken into account age. Suppose they said "No." Do you have to argue that they were wrong? No! The burden of proof lies with those making the claim. Besides, policy makers are seldom experts. But they can ask questions. The simplest way to take something into account for a rate or percentage is selection. In the case of the death rates, a policy maker can say,

---

[2] 1998 US Statistical Abstract, Table 1029. See Schield (2010): The Social Construction of Rankings.

[3] https://fee.org/articles/why-cubas-infant-mortality-rate-is-so-low/

"I'm not convinced. I'd be more convinced if you had taken age into account. For example, just show me the death-rate comparison for seniors and the comparison for all others (the non-seniors). If Mexico has lower death rates for both groups, then I'd be more persuaded by your claim that Mexico has better health care than the US.

So, if you think a crude association of rates or percentages may be influenced by a third factor, and this third factor has just a few values, then ask for the size of the comparison for each value of that third factor. To repeat, selection is the simplest way to control for the influence of a related factor on a comparison of rates or percentages.

3. Conclusions:

This paper identified seven basic elements of a statistic or statistical association. It proposed that policy makers consider seven simple questions: How big? Compared to what? Why not rates? Per what? Defined, counted or measured how? What was controlled for? What should have been controlled for?

These questions are simple and straightforward. The main thing is for policy makers to treat statistics the same way they treat people. People have motives, values and agendas. So do statistics – because they were selected, assembled and presented by people who have motives, values and agendas. Statistics are closer to words than to numbers. Yes, statistics involve numbers, but statistics are numbers in context and the words give the context.

For more examples, see Schield (1999, 2004B and 2005). For more details on study design, selection, ratios and standardization, see Schield (2010B).

Once policy makers are comfortable with these seven questions, they are ready to ask more complex questions. How was the data generated? What kind of study design was involved? The moral for statistics (as for anything involving evidence): "Good policy makers ask good questions!"

Acknowledgements:
Thanks to Cynthia Schield for reading a draft and to Marc Isaacson for suggestions.

References:

Best, Joel (2002). People Count: The Social Construction of Statistics. Copy at www.statlit.org/pdf/2002BestAugsburg.pdf

Schield, Milo (1999). Statistical Literacy: Thinking Critically About Statistics. Copy at www.statlit.org/pdf/2019SchieldIASE.pdf

Schield, Milo (2004A). Information Literacy, Statistical Literacy and Data Literacy. Copy at www.statlit.org/pdf/2004-Schield-IASSIST.pdf

Schield, Milo (2004B). Statistical Literacy and Liberal Education at Augsburg College. Copy at www.statlit.org/pdf/2004SchieldAACU.pdf

Schield, Milo (2005). Statistical Prevarication: Telling Half Truths Using Statistics. Copy at www.statlit.org/pdf/2005SchieldIASE.pdf

Schield, Milo (2010A). Statistical Literacy: A Short Introduction. Copy at this URL: www.statlit.org/pdf/2010Schield-StatLit-Intro4p.pdf

Schield, Milo (2010B). Assessing Statistical Literacy: Take CARE. Excerpts at www.statlit.org/pdf/2010SchieldExcerptsAssessingStatisticalLiteracy.pdf

Schield, Milo (2021). Statistical Literacy: Teaching Confounding. Copy at www.StatLit.org/pdf/2021-Schield-USCOTS.pdf