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 statistics from arithmetic numbers. Statistics are numbers in context -- where the context matters. Statistics are socially constructed. They are generated, selected and presented by people with motives, values and goals. Policy makers need to evaluate quantitative evidence using the same skills they use in evaluating non-quantitative evidence. Ask questions! This short 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. Analysing 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. Having had at least 12 years of maths, they may be used to treating numbers as facts. One plus one is always two.

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. 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 two things:

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 the words – market share – really matter.

Statistics are socially constructed. This is most obvious with social statistics. The number of choices involved in selecting and presenting a statistic are bigger than most people realize. This paper examines some of them.

Best (2002) argued that 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.*"

Most traditional statistics courses don't deal with large-group social statistic. They deal with small group random samples. Most policy makers do not need to study traditional statistics: random sampling, sampling error, margin of error, confidence intervals, random assignment, hypothesis tests, statistical significance and p-values.

Policy makers need a course in statistical literacy *Statistical literacy* studies *critical thinking about statistics as evidence in arguments*. Most of these are social statistics. Schield (2010A) provides a brief overview of statistical literacy. This paper provides some of the highlights.

2. Questions

Policy makers typically encounter statistical data used as evidence in arguments. Policy makers typically deal with statistical summaries obtained from large amounts of data for an entire group or time period. Policy makers need to focus on a few simple questions. Here are seven.

1. *How big?*

This simple question 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. "Doing X will increase Y". Another way for others to avoid size is to use 'many' and 'often'.

- "Many scientists said…" "X has happened too many times…"
- "X is often followed by Y". "All too often, doing X leads to Y."

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 in place of a quantity when the size is small. If it were big, they would probably tell you. So, if the size is not given, you have no basis for saying the effect or result is important.

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.

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

3. *Why not a rate?*

We don't need higher math to know that rates can control for the size of a group. We are not surprised if more men are hospitalized than women in the military. But we recognize that the hospitalization rate of women may be higher than that of men.

When given a count or comparison of counts, ask why not a rate or 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%).

Even though a rate may be better than a count, a single rate has no more context than a single count. You need a comparison.

Suppose you are told that New Jersey had a Covid19 death rate of 25.6 deaths per 1,000 cases. You would not know that New Jersey had the highest Covid19 death rate per case in the US.

Suppose you were told that Lithuania has a death rate of 14.6 per 1,000 population. Without a comparison you would not know that Lithuania has one of the highest death rates in the world.

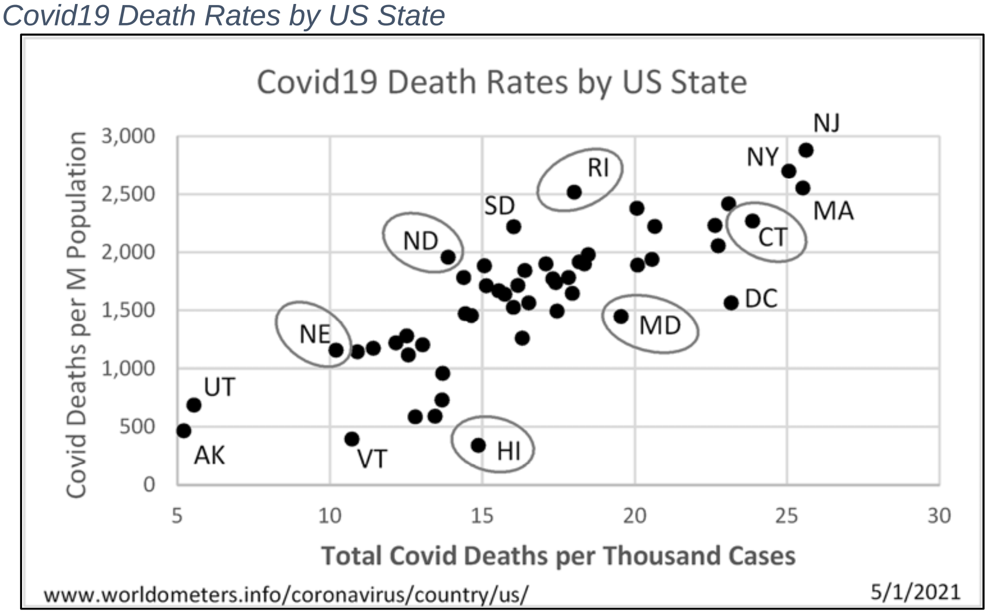
Without a comparison, a single count, rate or percentage stands alone. It has no context.

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.

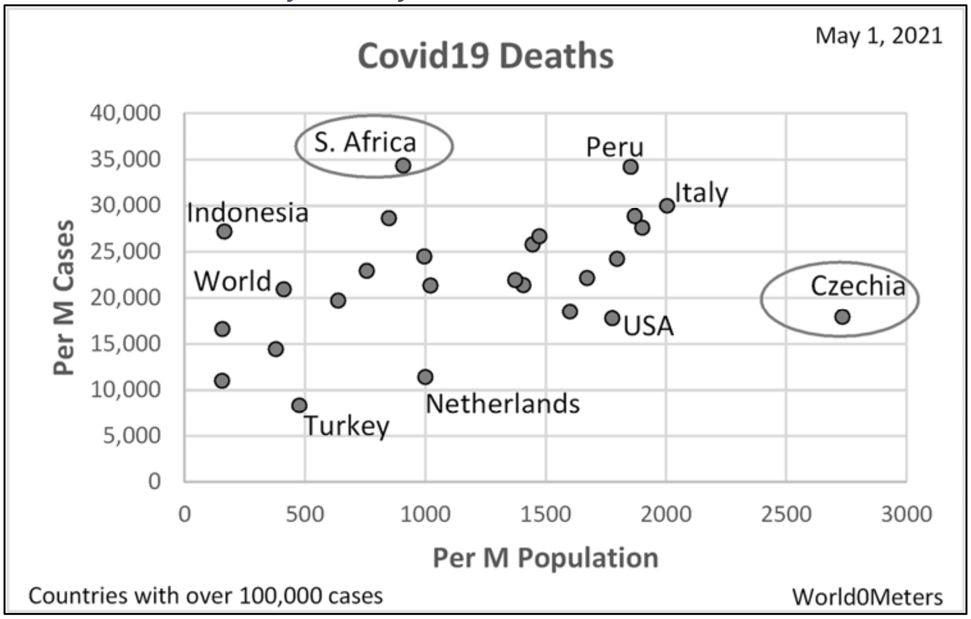
Compare the Covid19 death rates by state using two different denominators.

Figure 1: Covid19 Death Rates by US State



The Covid19 death rate was higher in Rhode Island than in Connecticut *per capita*, but it was lower [in Rhode Island than in Connecticut] *per case*. Higher in North Dakota than in Maryland per person, but lower per case.

Figure 2: Covid19 Death Rates by Country



The Covid19 death rate is higher in South Africa than Czechia per case. But lower [in South African than Czechia] per capita.

Consider another example: How 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? Easy. Women are living longer. Once again, the choice of the denominator can change the direction of a comparison.

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). [2]

One economist found the ratio of late fetal death to neonatal deaths for most countries ranged between one and three. See Figure 3*. Cuba's ratio was over six to one. This implies that Cuban doctors were reclassifying neonatal deaths as late fetal deaths. But, without the underlying data, we can't be sure.*

Figure 3: Classification of Fetal and Infant Deaths.

------- **Fetal Deaths** -------		Birth	-------**Infant deaths** ----------	
Before delivery			First year	
Early fetal	Late fetal	⟵ **Cuba ???**	Neonatal	Post-neonatal
< 22 weeks	> 22 weeks		First 28 days	Rest of year

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

We can estimate how many neonatal deaths would need to be reclassified as late fetal in order to lower in the infant mortality rate per thousand births from six to four.

If Cuba had a population of 10 million with a birth rate of 10 per thousand, then that would be 100,000 births per year. If Cuba's infant mortality rate was really 6 per 1,000 live births, then that would involve 600 infant deaths per year. The infant mortality rate per 1,000 live births could be lowered from 6 to 4 by reclassifying just 200 of the neonatal infant deaths to late fetal deaths. This reclassification might be barely noticeable amidst the 100,000 births per year unless one saw the data on late fetal and neonatal 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 million population is lower in Mexico (5.2) than in the US (8.2). The rates take into account the difference in population 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 million 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 maths. 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,

"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 proposed 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 a policy maker is comfortable with these six 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 else used as evidence): "Good policy makers ask good questions!"

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 (2006). Presenting Confounding Graphically Using Standardization. Copy at www.statlit.org/pdf/2006SchieldSTATS.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. In *Assessment Methods in Statistical Education*, Wiley, P. 133-152. Excerpts at www.statlit.org/pdf/2010SchieldExcerptsAssessingStatisticalLiteracy.pdf