EVALUATING STATISTICAL INFORMATION

BRITISH COLUMBIA INSTITUTE OF TECHNOLOGY | BCIT LIBRARY | bcit.ca/library

"We think of statistics as facts that we discover, not as numbers we create."- Joel Best, *Damned Lies and Statistics*

This guide introduces criteria for evaluating statistics gathered from articles, websites, and other sources. It is designed to be useful for researchers with no specialized background in mathematics or statistics. For more in-depth knowledge of how to evaluate the reliability of statistical information, consult any relevant materials from your courses, or see the list of additional resources at the end of this webpage, which contain many instructive examples of bad statistics. Remember that you can always ask a librarian for help with finding reliable sources of statistical information. Here are some of the key questions to ask as you encounter statistics:

WHO COLLECTED OR IS REPORTING THE INFORMATION?

Think critically about the source of the statistics and if necessary research who they are. Is the author a qualified researcher? If statistics are cited in an article, is the article peer-reviewed? Does the organization publishing the numbers have an agenda that may lead them to exclude certain results or spin the numbers in any way?

See our guide to Evaluating Web Sites for related tips on assessing online information sources.

A good first step towards avoiding bad statistics is to use resources linked from the library webpage such as our Statistics Web Links or Statistics + Government Publications databases. Many other BCIT Library databases feature subject-specific statistics, as well as reliable reports and peer-reviewed articles that include statistical information. The catalogue quick search can be used to find print sources of statistics as well.

WHAT IS THE CONTEXT?

Be critical of single statistics mentioned in isolation. In such cases, it is worth seeking out the original source of the number to get more information. Check the author's reference list or footnotes to identify the original source and use the library catalogue, library databases, or the internet to find it and investigate the details behind the number.

IS THAT A LOT?

Statistics reported as a single absolute value or a single percentage may sound dramatically large or small. But depending on the context and what is being counted, a million may not be as big as it sounds; a 200% increase sounds big, but its real significance depends on factors such as the original quantity and the amount of time during which the increase took place. Whenever absolute values or percentages appear without any kind of benchmark, get the complete context of the number, drawing on the full range of resources made available to you by the library.

Continued next page

CONTACT INFO
T 604.432.8371
E ebrarian@bcit.ca
bcit.ca/library

WHAT IS BEING COUNTED? WHAT IS NOT?

Look at the definitions of the study's variables. Try to decide if they are reasonable indicators of what is purportedly being measured or if they exclude too much other important information. For example, does a study of the quality of life rely too heavily on income levels? Sometimes answering such questions requires expert knowledge, but very often it does not. Critical thinking about the topic you are researching can put the numbers in perspective.

WHAT KIND OF AVERAGE IS BEING DISCUSSED?

Statistical and everyday meanings of words are sometimes conflated. If a study refers to averages, be aware that this does not necessarily translate to a study of typical cases. Depending on how "average income" is calculated, the size of the populations above and below the average may be wildly different. Be wary of sources that seem to confuse the everyday and statistical definitions of "average," that do not say whether the average is a median or mean, or that do not discuss effects of extreme, outlying data. Investigating the details behind averages does not necessarily require high-level mathematics! Ultimately, it is worth asking yourself: does the average helpfully simplify an issue, or do I need more detail to properly answer my research question?

DOES THE STUDY COMPARE LIKE WITH LIKE?

Statistics are often used to make comparisons, draw contrasts, and compile rankings. When you see statistics used this way, it is important to read with a critical eye for oversimplification. One of the most rampant abuses of statistics is the declaration without context that a number (e.g. crime, spending on a social program) has gone up or down. Comparison between two time periods is misleading if it does not compensate for other important factors that affect the statistic, such as inflation or population change.

IS IT A CORRELATION OR CAUSATION?

A correlation is an association between two variables; causation is a cause-and-effect relationship between the two. A correlation shows that the two variables are often found together, but it does not prove that one variable causes the other. With correlations, there may be a *confounding factor*: a third variable that really accounts for the relationship. Sometimes specialized knowledge is needed to make sense of the difference, but in many cases it pays to read carefully and use your own knowledge. Does the source call the statistic a "trend," a "link," or an "association" (correlation), or does it refer to a "cause," "effect," or "predictor" (causation)? If it is a correlation, it may still be useful, but brainstorm or read about any confounding factors that could offer alternative explanations for the statistic.

IS IT BASED ON A SCIENTIFIC SAMPLE OR RIGOROUS COUNT?

Some studies (for example, most social science experiments) are based on a sample: they use the analysis of a subset to make generalizations about the larger group. Other studies (for example, the Canadian Census) attempt to count everyone or everything being studied. When done rigorously, both can be valid methods of deriving statistics. However, both methods, done poorly, can produce dubious results. Poor sampling may skew a statistic up or down due to biased or careless selection of the sample. Attempts to count the entire population may be based on poor definitions of what is being counted, human error in counting, or lack of reliable data. With either type of study, sources that obscure their methodology should be viewed critically. As always, it is important to ask questions such as: what are the credentials of the researchers? does the source have a known agenda? if the statistics are reported in a journal, is it peer-reviewed?

IS THE EVIDENCE MERELY ANECDOTAL?

Sometimes statistics are reported as if they were based on a sample when they were actually derived from anecdotal evidence. Evidence is anecdotal if it comes from one case or from an extremely small number of cases. It is better to use statistics based on scientific sampling or rigorous counting of a larger number of cases.

DOES A CHART OR GRAPH CLEARLY COMMUNICATE A STATISTIC, OR DOES IT CHANGE THE MEANING?

Charts and graphs can quickly communicate the significance of statistical information. But they can also be used for exaggeration. Paul Bolton, of the UK House of Commons Library, notes that media often use charts with axes that are adjusted to make a change seem more dramatic or rapid than it actually is. When in doubt, look more closely at the details behind the image.

Adapted from: *The Numbers Game* by Michael Blastland and Andrew Dilnot, and *Statistical literacy guide: How to spot spin and inappropriate use of statistics* by Paul Bolton (UK House of Commons Library).

Additional Resources

WEBSITES

StatLit.org

Statistical literacy guide: How to spot spin and inappropriate use of statistics/ Paul Bolton (UK House of Commons Library)

BOOKS

The numbers game: the commonsense guide to understanding numbers in the news, in politics, and in life / Michael Blastland and Andrew Dilnot

Call #: QA 141.15 B535 2009

Damned lies and statistics: untangling numbers from the media, politicians, and activists / Joel Best

Call #: HM 535 B47 2001

"How to Lie with statistics" / Darrell Huff. In: Strategies for business and technical writing, ed. Kevin J. Harty

Call #: HF 5721 H37 2008

Innumeracy: mathematical illiteracy and its consequences / John Allen Paulos

Call #: QA 93 P38 1988

A mathematician reads the newspaper / John Allen Paulos

Call #: QA 93 P385 1995

Seeing through statistics / Jessica M. Utts

Call #: QA 276.12 U88 1999

Statistics without tears: a primer for non-mathematicians / Derek Rowntree

Call #: QA 276.12 R68 1981

Numbers guide: the essentials of business numeracy / Richard Stutely

e-book available through library catalogue

The visual display of quantitative information / Edward R. Tufte

Call #: QA 276.3 T83 2001

How to lie with maps / Mark Monmonier

Call #: G 108.7 M66 1996

BRITISH COLUMBIA
INSTITUTE OF TECHNOLOGY
3700 WILLINGDON AVENUE
VANCOUVER, BRITISH COLUMBIA
CANADA V5G 3H2

BCIT LIBRARY
bcit.ca/library
June 2010