

Lies, Damn Lies, and Statistics

Accounting is one of the oldest information systems. Its purpose is to provide meaningful data to stakeholders and decision makers, but do we defeat our purpose through the use of specialized jargon, obscure concepts, and a focus on the mechanics of bookkeeping? Rather than address that issue directly, I will give an example from a sister science, statistics.

There is a legend about how President Eisenhower was shocked to learn that fully half of all public school students were below average. This tale is supposed to ridicule the intelligence of the late President, but it really tells us more about how ridiculous statistics can be. Obviously, half of all students will fall below the median, but what does that really mean?

In standard usage, an "average" student is not just one who exactly falls on the median. We couldn't identify such a student, anyway. Say all students were ranked by a standard IQ test. If the median score is normalized to 100 and the standard deviation is 22, then we can expect 1.8% of all students taking the exam to score 100. Half of these students fall below the median and half are above, but how can we tell one from another? Even the most devoted statisticians will acknowledge that the same student will score differently each time they take the test, based on a number of environmental factors or sheer luck. Someone who takes the test three times might get a score of 98, 102, and 101. But the law of large numbers provides assurance we won't get a large score spread, like 89, 91, and 132--that is, unless the student cheats or intentionally answers questions wrong on some exams and not on others. So students who score 98 and 102 can be of similar intelligence (as far as the exam is concerned), even though, technically, one is below average and the other is above.

How far a spread do we allow to call someone "average"? That's where our old friend the standard deviation comes in. Statistical language can get much closer to standard usage if we say that someone is average if they score within one standard deviation of the mean. Using a truncated normal distribution, about 68% of all students will be average, so only 16% or almost one-sixth will actually be below average. President Eisenhower would have been right to be concerned if it were one half!

Of course, it would be far more meaningful to discuss what skill level qualifies as average. If the median high school graduate reads at a sixth grade level, or can't write a short essay, then that is a greater cause for concern than arbitrary statements about who is below average. In government accounting, many agencies had to work hard to get expenditures to equal expenses plus purchases, or appropriated capital to equal funding available plus commitments and undelivered orders. These are necessary steps, but they are only the first steps to properly track financial activity and generate measures of

outputs and outcomes to produce meaningful data that can be properly used for making decisions.

Comments, suggestions, and critiques are welcome. Send them to Simcha.Kuritzky@cgi-ams.com, and not to the AGA.