

Book Review

Title: Standard Deviations: Flawed assumptions, Tortured Data, And Other Ways to Lie With Statistics

Author: Gary Smith

Publisher: Duckworth Overloo, Pages: 326

Gary Smith, the Fletcher Jones Professor of Economics at Pomona College, Claremont, California, has written a deeply studied and thought-provoking book on how researchers torture data to produce outcomes that prove their assumptions which themselves could be flawed. People are taken for a ride sometimes even for prolonged periods. They lead people to draw false inferences and making wrong decisions. Even informed people are misled by biased or irrelevant data or by the flawed approaches to research. One reason could be the assumption that computers are infallible and that 'no matter what kind of garbage we put in, computers will spit out gospel.' This, according to the author, is because in the past data were scarce and computers non-existent, researchers had to work hard to collect good data and used to spend long time for making the necessary calculations. But today, with data so plentiful, researchers don't spend time to distinguish good data from rubbish, and sound analysis from junk science. And most people assume that since we are able to handle large amounts of data, nothing can go wrong. And all these lead us to make decisions based on the nonsense that the computers spit out.

Prof. Gary Smith gives a vivid narration of how data get distorted and mangled by graphs, how selection bias or survivor bias leads to wrong decisions, how regression towards the mean get used for making wrong conclusions, clustering effect of certain phenomena lead to wrong theories, how certain outlier data get consciously avoided in arriving at conclusions, how superstitions get into those events which occur at random, how distant prayers could heal deadly diseases like aids or cancer etc etc. While all these are related to evolving theories

with tortured data, Prof. Smith also narrates a number of examples connected with arriving at data without any theory.

A few illustrations by Prof. Smith in support of his arguments can highlight how even the so-called experts make mistakes or lead to wrong conclusions and decisions. During World War II, the British Royal Air Force planned to fix heavy plates to the fighter planes to protect them from German fighter planes and anti-aircraft guns. If protective plates were used to cover the entire plane, the weight will increase to prohibitive levels. So the RAF collected data on the location of bullet and shrapnel holes on planes that returned from bombing missions. They found that most holes were on the wings and the rear of the plane, and very few on the cockpit, engines or fuel tanks, suggesting that protective plates be put on the wings and radar. The data actually suffered from survivor bias. Returning planes seldom had holes in the cockpit and the fuel tank because those planes that were hit at these locations did not survive to return to England. Returning planes were more likely to have holes in the wings because these holes did little damage. Abraham Wald, a Hungarian Jew, provided the insight that these data suffered from survivor bias and advised that instead of reinforcing the location with the most holes, they should reinforce the locations with no holes.

In another example, Prof. Smith explains how selective omissions of data by two Harvard professors, Carmen Reinhart and Kenneth Rogoff, lead to a conclusion that a Debt/GDP ratio above 90% leads countries to slip into recession. Many European governments even tried to reduce their budget deficits by cutting spend and raising taxes while other researchers working with the same data and including the

Standard Deviations: Flawed assumptions, Tortured Data, And Other Ways to Lie With Statistics

data omitted by Reinhart and Rogoff, found that there is not a simple, close relationship between debt and growth.

Prof Smith has also narrated a number of examples to show how treacherous is data without theory. For example, many people compare stock prices with Consumer Price Index (CPI) to predict stock market bubbles. The comparison between prices of Berkshire Hathaway's (the investment vehicle of the legendary investor Warren Buffet) prices and the CPI between 1995 and 2005. The Berkshire Hathaway prices increased by 269 percent while the CPI increased by 27 percent. Was it a reason for predicting bubble in Berkshire Hathaway? No way. In 2005, each share of Berkshire Hathaway was priced at \$90,000. In 2007, it increased to \$141,600 and in 2003, it quoted \$170,000. The gap between the Berkshire Hathaway share price and the CPI actually increased after 2005. Actually, the reason for its higher price is that the number of shares outstanding is fewer than 2 million (IBM has one billion and Exxon Mobil has six billion) and the company never went for any stock split, while many other companies resorted to stock splits specifically to make their shares find more trades in the market.

The book contains a number of such examples of flawed analysis or tortured data and those too attributable at times to well-known personalities and among them even Nobel-laureates. Many of them were published in difficult-to-get-entry journals, only to be contradicted later in the same journals.

A very interesting book which describes how some researchers lead us to make wrong conclusions or decisions based on flawed assumptions and tortured data.

About the Author

Prof. Gary Smith is the Fletcher Jones Professor at Pomona College in Claremont, California. He taught at Yale University for seven years. He also taught at University of Houston and Rice University. He teaches courses like Economic Statistics, Economic Modeling, Security

Valuation and Portfolio Theory, among others. He has published a large number of papers in reputed international journals. He has also published a number of textbooks on Economics, Banking, Investments, and statistics.

Standard Deviations: Flawed assumptions, Tortured Data, And Other Ways to Lie With Statistics