

Lies, Damned Lies, and Statistics

The Use and Abuse of Numbers

by DR. JAYME A. SOKOLOW

Benjamin Disraeli, Victorian England's most famous prime minister, once said there were three kinds of lies: "lies, damned lies, and statistics." There are actually countless forms of mathematical chicanery all around us, especially when we begin calculating the number of creative ways we use statistics to win arguments, sell products, or just plain bamboozle people. The potential for abuse may even exist with proposals, though there are no studies to prove or disprove this point. Some examples may be instructive and cautionary signposts of what to avoid.

As Darrel Huff has argued in his hilarious classic, *How to Lie with Statistics* (1954), now back in print, the "secret language of statistics, so appealing in a fact-minded culture, is employed to sensationalize, inflate, confuse, and oversimplify." Sometimes statistical methods and terms are unwittingly misused, especially in the media. On other occasions, however, statistics are consciously used to baffle, deceive, legitimize decisions, and bolster authority and power.

more...

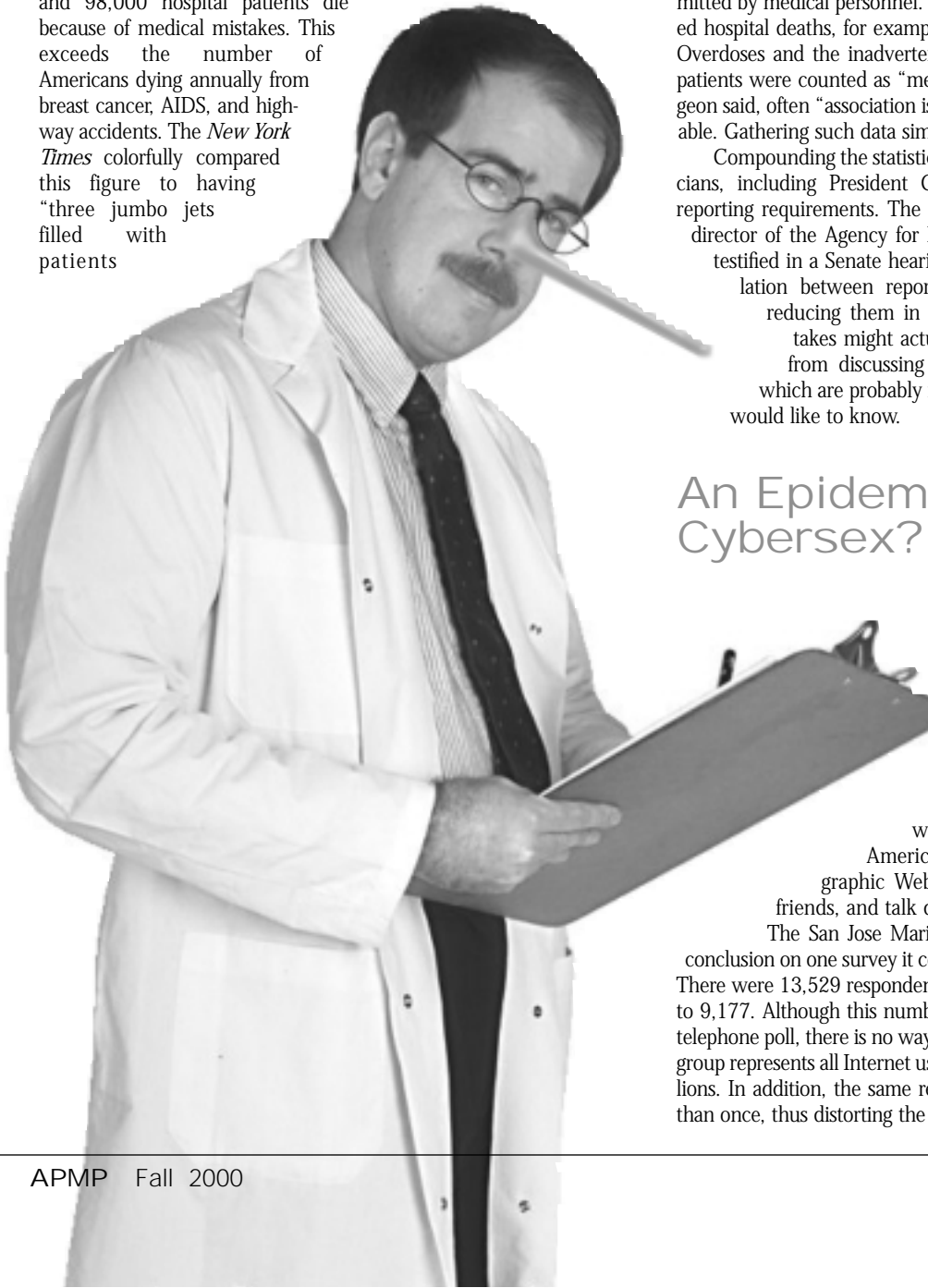


According to STATS, statistical confusion and inaccuracy are thriving in the United States.

Doctoring Statistics

To study the lively art of statistical misuse and manipulation, I recommend that you visit the Web site of the Statistical Assessment Service (www.stats.org), a nonprofit organization that examines the ways quantitative research is used by the media. According to STATS, statistical confusion and inaccuracy are thriving in the United States.

Two highly publicized recent reports highlight the misuse of statistics. According to the Institute of Medicine, every year between 44,000 and 98,000 hospital patients die because of medical mistakes. This exceeds the number of Americans dying annually from breast cancer, AIDS, and highway accidents. The *New York Times* colorfully compared this figure to having "three jumbo jets filled with patients



crash every two days." These numbers, however, are very unreliable because the Institute of Medicine's report is riddled with questionable assumptions and dubious calculations.

The Institute of Medicine based its conclusions on two studies: a 1984 study of hospital discharges in New York with 129 fatalities in 30,000 cases and a 1992 study that covered Utah and Colorado with 59 deaths in 15,000 cases. The Institute of Medicine extrapolated these figures to the 1997 national hospital admissions figure of 33.6 million and arrived at their expansive numerical range.

One problem is the states that were used in the study. Are hospitals in New York, Utah, and Colorado representative of the entire country? Another problem is the use of hospital admissions. If estimates had been made based on hospital discharges, the 44,000 to 98,000 range would have decreased, becoming 39,650 to 88,450.

Another flaw with this study was its loose definition of medical error. Because errors were defined as "inappropriate decisions...when an appropriate alternative could have been chosen," it is very difficult to separate patient errors from those committed by medical personnel. More than 7,000 of the extrapolated hospital deaths, for example, were medication-related errors. Overdoses and the inadvertent use of the wrong medicines by patients were counted as "medical errors." As one skeptical surgeon said, often "association is indirect, hard to make, and debatable. Gathering such data simply isn't an exact science."

Compounding the statistical errors was the response of politicians, including President Clinton, who called for statutory reporting requirements. The problem with this solution, as the director of the Agency for Health Care Research and Quality testified in a Senate hearing, is that there is no direct correlation between reporting medical errors and actually reducing them in hospitals. In fact, publicizing mistakes might actually discourage hospital personnel from discussing real examples of medical errors, which are probably more widespread than most people would like to know.

An Epidemic of Cybersex?

In another highly publicized report, the San Jose Marital and Sexuality Center recently claimed that eight percent of all Internet users are cybersex addicts. Almost five million more people could be at risk, the report darkly warned. Is it possible that so many Americans feel compelled to visit pornographic Web sites, send lewd e-mail to their friends, and talk dirty in chat rooms?

The San Jose Marital and Sexuality Center based its conclusion on one survey it conducted on MSNBC.com last year. There were 13,529 respondents, which the Center filtered down to 9,177. Although this number is 13 times larger than a typical telephone poll, there is no way of ascertaining how accurately this group represents all Internet users, who number in the tens of millions. In addition, the same respondents could have voted more than once, thus distorting the results of the survey.

No research has yet established that there is a disorder of Internet addiction that is separable from problems such as loneliness ... or that a passion for using the Internet is long-lasting.

Another problem is the definition of cybersex addiction. The word addiction usually refers to activities that are compulsive, that involve withdrawal symptoms, and that physically alter the brain. As one scholar argued after the report was issued, it "seems misleading to characterize behaviors as 'addictions' on the basis that people say they do too much of them. No research has yet established that there is a disorder of Internet addiction that is separable from problems such as loneliness ... or that a passion for using the Internet is long-lasting."

Perhaps cybersex is a problem with a growing number of Internet users. But using a self-selecting group of 9,177 people who happened to learn about a survey on MSNBC.com to represent the nation's millions of Internet users is very questionable from a statistical standpoint. The sample is certainly not representative and probably too small.

Join the Navy!

My third example comes from Huff's *How to Lie with Statistics* and concerns the US Navy. During the Spanish-American War (1898), the death rate for Navy personnel was nine per thousand. In the same year, New York City's mortality rate was 16 per thousand. Navy recruiters later used these figures to argue that a career in the Navy was far safer than living in the Big Apple.

But these are not comparable population samples. Most Navy personnel are young men who have passed a stringent physical exam and are in excellent health. New York City's population, however, is very different. You do not need a physical exam to become a resident of the five boroughs, and the population includes the elderly and large numbers of people with serious illnesses.

Perhaps it was safer to be in the Navy than to live in the Bronx. Perhaps it was not. But the Navy's argument about its favorable differential mortality compared to New York City was simply statistical nonsense.



Statistical Miscalculations

These three examples are just the tip of the statistical iceberg. Newspapers, magazines, television, and radio are full of numerical misinformation. Too many statistical statements are based on bad mathematics; samples that are not randomized; samples so small that differences produced by chance are likely to be large; samples with low levels of statistical significance; and statistical conclusions that confuse correlation with cause.

The single most accurate single predictor of the S&P 500 stock index was Bangladesh's butter production!

Huff argues that non-randomized and small samples are the two most common causes for statistical inaccuracy, especially in the slippery world of advertising. In large data sets, mistaking correlation for cause may be a frequent error. As *Business Week* reported, one fund manager humorously claimed, based on his study of a United Nations CD-ROM, that the single most accurate single predictor of the S&P 500 stock index was Bangladesh's butter production!

more...



Another problem is the pictorial representation of numbers. As Edward Tufte demonstrated in his *The Visual Display of Quantitative Information* (1983), too many displays of statistics—charts, graphs, tables, and other representations of quantity—do not depict numbers and numerical trends accurately.

The Zeal for Quantification

Many cultures throughout history have been fascinated with numbers, but in the modern world quantification has a prestige and power unparalleled in ancient India or medieval Europe. One reason undoubtedly is its many successes in the physical and life sciences, technology, engineering, government, and the social sciences.

But quantification serves another important function. As we move beyond our own localities, we need to find ways to transact business from afar and deal with strangers. Because numbers convey information in a familiar, standardized, and reassuring form, they are superbly adapted to long-distance commerce and communications. In a large heterogeneous world of strangers, quantification functions as a seemingly neutral, objective, and value-free discourse that promotes interaction across time and space.

Until recently, few people were concerned about the lack of numerical standards. Before the nineteenth century, for example, many European towns had their own particular weights and measures, which they proudly defended as a symbol of their sovereignty and independence. In pre-Revolutionary France, every province coined its own money and had its own methods for calculating a bolt of silk or a bushel of wheat.

In 1860, when it was noon in Chicago it was 11:50 AM in St. Louis, 11:27 AM in Omaha, and 12:18 PM in Detroit.

In the United States, despite uniform coinage, weights and measures, and the absence of internal tariffs, the time of day was locally determined until after the Civil War. In 1860, when it was noon in Chicago it was 11:50 AM in St. Louis, 11:27 AM in Omaha, and 12:18 PM in Detroit. Railroad companies got so tired of setting their clocks to 53 different standards that on November 18, 1883, they created four time zones, which encouraged communities to switch from local to railroad time.

This imprecision still survives informal conversations. Who has not heard a parent say, “I’ve told you 1,000 times” or been baffled by a manager’s hope that “every team member gives us 110 percent.” The only time precise statistics are used in everyday speech may be when males discuss baseball.

But as national and international commerce began to link disparate communities and as central governments became more powerful, quantification emerged as a substitute for local knowledge and personal trust. Quantification became an effective form of communication because it transcended local boundaries to produce credible information while bolstering the authority and expertise of those who created the numbers.

The City of the Big Shoulders

The rise of nineteenth-century Chicago provides a vibrant example of how numbers enable strangers to transact business over great distances. Grain made Chicago the most powerful city in the Midwest by the Civil War, and statistics played a vital role in turning crops into commodities.

Before the 1850s, in the absence of railroads and decent roads, farmers in the Midwest sent their wheat in personally marked sacks on river flatboats to Chicago, St. Louis, or New Orleans.

Downstream, a miller or merchant would closely inspect each bushel sack with his eyes and hands and then offer the farmer a price.

There were no uniform prices for a bushel of wheat or barley, and no standard definitions of what constituted high- or low-quality grain.

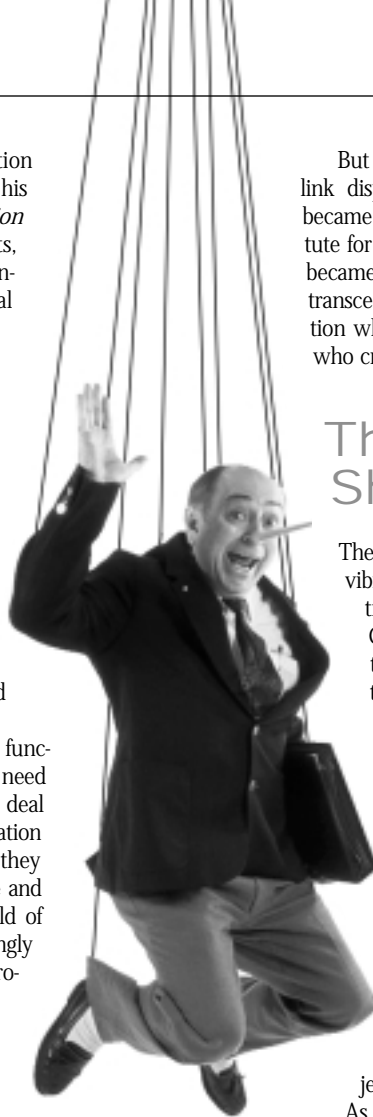
Instead, millers and merchants used their personal experience to decide how much each bushel was worth. Knowledge was local, subjective, and imprecise.

As Chicago grew, the local world of Midwestern farmers and merchants dramatically changed. By 1860, thousands of miles of railroad lines brought wheat to Chicago from Chicago’s hinterland. After it arrived, steam-powered conveyor belts moved a farmer’s wheat sacks to the top of a grain elevator where they were weighed and then dumped into a bin. By 1857, the city had 12 grain elevators with a capacity of 4 million bushels.

Grain elevator operators, however, faced a major problem. It was not cost-effective to keep individual sacks of grain in separate bins. A new organization, the Chicago Board of Trade, solved this problem and unknowingly helped transform Midwestern agriculture. Founded in 1848, the same year as Chicago’s first railroad, grain elevator, telegraph, canal, and stockyard, the Board established a standard weight for a bushel of grain. When farmers learned that Chicago businesses would pay them the same price for any bushel, they started adding dirt, chaff, and much worse to their wheat and barley.

In 1856, the Board responded by classifying grain into grades based on its quality. Now elevator operators could mix the grains of different farmers and give farmers a receipt for their produce. This made grains interchangeable between elevator bins, cities, and even continents. Now No. 3 spring wheat could be sold in New York City, London, and Moscow on the basis of prices quoted over the telegraph.

The next year, the Board appointed its own city grain inspector and assistants to certify the proper grades for all grain traded on the Chicago Exchange. In 1859, the Illinois state legislature authorized the Board to create standardized grades and inspection codes for its members. By the Civil War, Chicago dominated the Midwestern grain market because of its extensive railroads and



elevator warehouses and the grading and marketing systems established by the Board of Trade. At the same time, merchants and speculators began trading elevator receipts on the floor of the Exchange. The futures market had been born.

Grain prices were no longer established by local farmers, millers, and merchants as rural production grew more remote from the economic point of processing and consumption. Now, grain was bought and sold on the floor of the Chicago Exchange by businessmen who never touched or saw any natural produce. They could even speculate on grain that had yet to be harvested.

Gone were the days when merchants talked to farmers and personally knew their crops. Grain elevators and grading systems had transformed cereals from a crop into a numerical abstraction.

The futures market completed this process by freeing the market from the literal exchange of cereals. In 1875, Chicago's grain business was approximately \$200 million. The volume of futures, in contrast, was \$2 billion, ten times greater than the buying and selling of actual grain.

As one bemused visitor noted in 1880, "in the business centre of Chicago you see not even one 'original package' of the great cereals." Moving produce from farmers' sacks into grain elevators unintentionally started the revolutionary process of turning crops into statistics—elevator receipts, national and international prices, production data, railroad and shipping schedules, and the value of commodity futures. Chicago may have been the "City of the Big Shoulders," in Carl Sandberg's memorable phrase, but its power depended on numbers as well as muscle in its dominance of Midwestern agriculture.

A Senseless Census?

The controversy over the 2000 Census illustrates an important point about the use of numbers by the government. Although they may be statistically sound, numbers are never neutral, value-free, or objective. Since quantification is always embedded in a social and political context, government numbers are often the subject of heated analysis and dispute.

According to Article I, Section II of the Constitution, every ten years an "actual enumeration" must be conducted to determine the number of members each state is entitled to have in the House of Representatives. The first census in 1790 recorded 3.9 million inhabitants.

As the nation grew, so did the US census. In 1810, the census asked questions about manufacturing and the amount and value of products. In 1850, new questions covered taxation, religion, the indigent, crime, and insanity. There were so many new questions in the censuses of 1880 and 1890 that it took the government almost a full decade to publish the results.

Over the past three decades, the Census Bureau has experienced increasing difficulty counting everyone. From 1970 to 1990, the percentage of people in houses mailing back census forms dropped from 78 to 65 percent. From 1980 to 1990, the census undercount also increased from 1.2 to 1.8 percent of the population, or almost 4 million Americans. Most of the undercounted were poor, Black, or Hispanic.

Several years ago, Congress directed the Bureau to devise plans for the 2000 Census that would reduce the undercount and also limit costs, which had sharply increased even after allowing for inflation and population growth. The Bureau responded by proposing to use statistical sampling once again because it seemed a scientific and non-partisan solution to the twin problems of undercounting and rising costs. Sampling had been used in previous censuses without much comment.

The Bureau expected little controversy over statistical sampling, which is widely used in medicine, industry, accounting, and other fields that demand mathematical rigor.

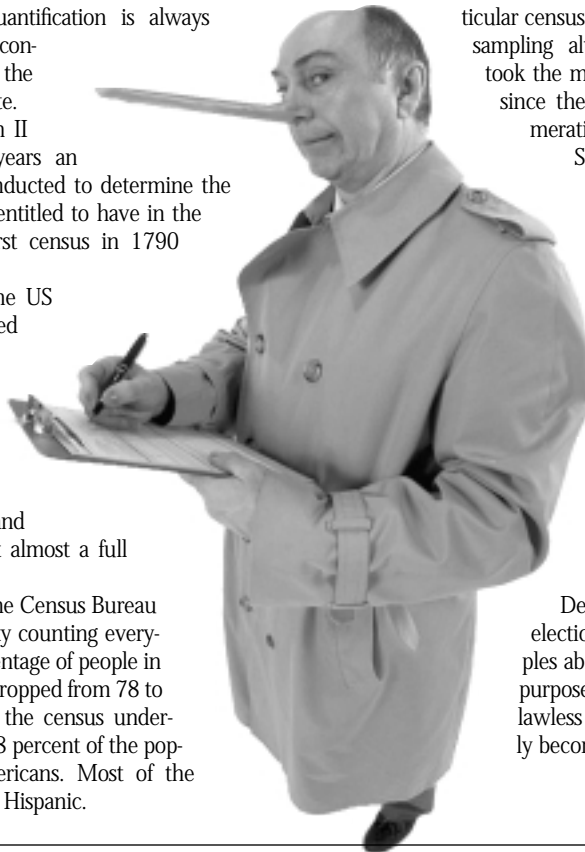
The Census Bureau worked closely with the American Statistical Science Association to develop an accurate sampling method for the 2000 Census. The Bureau expected little controversy over statistical sampling, which is widely used in medicine, industry, accounting, and other fields that demand mathematical rigor. The Bureau believed that it could develop carefully designed sampling techniques that would generate population data with a high degree of accuracy.

But Republicans became outraged over the Bureau's proposed use of statistical sampling. One Republican Congresswoman introduced a bill to use sampling only after direct contact had been made with 90 percent of households in a particular census tract. Another bill would have prohibited sampling altogether. Finally, dyspeptic Republicans took the matter to the Supreme Court, arguing that since the Constitution stipulated an "actual enumeration," sampling was unconstitutional. The Supreme Court agreed.

The battle over the 2000 Census was not really about the accuracy of statistical sampling. It was about two highly partisan political issues that had become entwined with discussions of census numbers.

First, many Republicans objected to any statistical sampling with President Clinton in the Oval Office. As one critic fulminated, "this is a White House that had no scruples about getting the Immigration and Naturalization Service to drop criminal checks on applicants for citizenship so that more Democrats could be naturalized for the 1996 election; why would it suddenly develop scruples about adjusting census numbers for political purposes?" Since the president was viewed as a lawless person, statistical sampling would logically become his latest form of political abuse.

more...



Their second objection was rarely voiced publicly. In a House of Representatives where Republicans have a slim majority, there are powerful political reasons to attack statistical sampling. Those counted in statistical sampling—primarily the poor and minorities—are overwhelmingly Democratic voters. Statistical sampling might help lead to the creation of new Congressional districts with potential Democratic majorities.

Even after the Bureau began the 2000 Census, Republicans continued attacking it. Trent Lott (R-Mississippi), the Senate Majority Leader, condemned the census as being too “intrusive” and urged his constituents not to return their census forms. After critics pointed out that the Senate had approved every question and category in the 2000 Census, Senator Lott hastily beat an ignominious retreat from the statistical battlefield. His press secretary lamely argued that the senator was actually “agnostic” about the census, a strange word to use from a politician strongly supported by the Christian Coalition.

Meanwhile, there is good news about the 2000 Census. Following the Supreme Court decision, the Bureau has undertaken a concerted media campaign to encourage all Americans and especially minorities to complete their census forms. Nationwide, as of June 2000, 65 percent of the 2000 Census questionnaires have been returned, a rate equal to the previous census.

Implications for Proposal Professionals

There is a great irony in our eagerness to use statistics and believe them. Quantification makes knowledge more open, understandable, and uniform. Numbers enable people to communicate across languages, cultures, and continents. Numerical standards promote interdependence by enabling strangers to transact business over vast distances. Quantification has fundamentally altered the way we understand the world.

As the articles in this issue demonstrate, some companies are using proposal metrics to become more efficient, which they associate with a higher win rate. In today’s business environment, numbers can do more than measure success. They may also help us better understand what elements of the proposal development process work well and what elements need to be changed.

At the same time, the very power and persuasion of quantification obscures the fact that numbers are social and historical artifacts. They are never abstract, neutral, or value-free. As the grain market in mid-nineteenth-century Chicago and the 2000 Census demonstrate, statistics are not timeless, objective entities that exist outside society. When we use statistics in proposals, we are doing much more than merely counting or displaying numerical trends.

Usually, numbers in a proposal serve one purpose—to help convince reviewers that we are best qualified to be awarded a contract. In other words, proposal statistics primarily function as part of a persuasive argument to demonstrate that we are highly experienced and qualified, regardless of what the numbers may actually mean. Numbers augment our authority and expertise by making us appear scientific, rigorous, and credible, whether they are

real, false, or misleading.

Let Mark Twain have the last word about the use and abuse of numbers. In *Life on the Mississippi* (1883), Twain entertained his readers with an explanation of the changing length of the Lower Mississippi River that I believe has rarely been equaled for its quantitative power:

“In the space of one hundred and seventy-six years the Lower Mississippi has shortened itself two hundred and forty-two miles.

This is an average of a trifle over one mile and a third per year. Therefore, any calm person, who is not blind or idiotic, can see that in the Old Oölitic Silurian Period, just a million years ago next November, the Lower Mississippi River was upward of one million three hundred thousand miles long, and stuck out over the Gulf of Mexico like a fishing-rod. And by the same token any person can see that seven hundred and

forty-two years from now the Lower Mississippi will be only a mile and three-quarters long, and Cairo and New Orleans will have joined their streets together, and be plodding comfortably along under a single mayor and a mutual board of aldermen. There is something fascinating about science. One gets such wholesale returns of conjecture out of such a trifling investment of fact.”

The same might be said of many statistics, even those that appear in proposals.

“And by the same token any person can see that seven hundred and forty-two years from now the Lower Mississippi will be only a mile and three-quarters long.”

References

- Cronon, William. *Nature’s Metropolis: Chicago and the Great West*. New York: W.W. Norton, 1991.
- Huff, Darrell. *How to Lie with Statistics*. 1954, New York: W.W. Norton, 1993.
- Mickle, Lee. “Lies, Damned Lies, and Statistics.” *The Editorial Eye*, 1998 (21): 1-3.
- Porter, Theodore M. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University, 1995.
- Sandberg, Carl. *The Complete Poems of Carl Sandberg*. New York: Harcourt Brace Jovanovich, 1970.
- Tufte, Edward. *The Visual Display of Quantitative Information*. Cheshire, CONN: Graphics Press, 1983.
- Twain, Mark. *Life on the Mississippi*. 1883, New York: Penguin Books, 1984.
- Urton, Gary. *The Social Life of Numbers: A Quechua Ontology of Numbers and Philosophy of Arithmetic*. Austin, TX: University of Texas, 1997.

Web Sites

Statistical Assessment Service: <http://www.stats.org>
2000 Census: <http://www.census.gov>; <http://reporternews.com>

Jayne A. Sokolow, Ph.D., is founder and president of The Development Source, Inc., a proposal services company located in Silver Spring, Maryland, that works with both businesses and nonprofit organizations. He is also Chair of the Editorial Advisory Board of *Proposal Management*. He can be reached at JSoko12481@aol.com.