

2015

Review of *Naked Statistics: Stripping the Dread from Data* by Charles Wheelan

Michael T. Catalano

Dakota Wesleyan University, micatala@dwu.edu

Follow this and additional works at: <http://scholarcommons.usf.edu/numeracy>



Part of the [Statistical Methodology Commons](#)

Recommended Citation

Catalano, Michael T. (2015) "Review of *Naked Statistics: Stripping the Dread from Data* by Charles Wheelan," *Numeracy*: Vol. 8 : Iss. 1 , Article 13.

DOI: <http://dx.doi.org/10.5038/1936-4660.8.1.13>

Available at: <http://scholarcommons.usf.edu/numeracy/vol8/iss1/art13>

Review of *Naked Statistics: Stripping the Dread from Data* by Charles Wheelan

Abstract

Wheelan, Charles. *Naked Statistics: Stripping the Dread from Data* (New York, NY, W. W. Norton & Company, 2014). 282 pp. ISBN 978-0-393-07195-5

In his review of *What Numbers Say* and *The Numbers Game*, Rob Root (*Numeracy* 3(1): 9) writes “Popular books on quantitative literacy need to be easy to read, reasonably comprehensive in scope, and include examples that are thought-provoking and memorable.” Wheelan’s book certainly meets this description, and should be of interest to both the general public and those with a professional interest in numeracy. A moderately diligent learner can get a decent understanding of basic statistics from the book. Teachers of statistics and quantitative literacy will find a wealth of well-related examples and stories to use in their classes.

Keywords

Review, Numeracy, Statistics, Data

Creative Commons License



This work is licensed under a [Creative Commons Attribution-Noncommercial 4.0 License](https://creativecommons.org/licenses/by-nc/4.0/)

Cover Page Footnote

Mike Catalano is the Book Review Editor for this journal. He is a professor of mathematics and chair of the department at Dakota Wesleyan University. He is a former member of the Board of Directors of NNN and has a particular interest in incorporating numeracy into college algebra through the use of examples related to social issues.

Introduction

Naked Statistics, by Charles Wheelan (2014), is an interesting, example-packed, inviting, and occasionally facetious discussion of concepts covered in most introductory statistics courses. The traditional list of topics includes (not in this order) sampling and data collection techniques, experimental design, descriptive statistics, correlation, regression analysis, basic probability, inferential statistics, polling, confidence intervals and margin of error, and the Central Limit Theorem. The treatment is far from traditional, however, and Wheelan does a fine job finding easily understood and interesting examples to illustrate fundamental concepts. This is not a textbook; there are no exercises, and the calculations are generally rudimentary. Still, even a moderately diligent learner can get a decent understanding of basic statistics from the book. Teachers of statistics and quantitative literacy will find a wealth of well-related examples and stories to use in their classes.

Wheelan teaches economics and public policy at Dartmouth, where he earned his B.A. and previously authored *Naked Economics: Undressing the Dismal Science* (Wheelan 2003). For five years, he served as a correspondent for *The Economist*. He has a Master's in Public Affairs from Princeton, and a Ph.D. in Public Policy from the University of Chicago's Harris School.

As an economist, Wheelan certainly knows his way around statistics, but he also brings a practical approach to it. Chapter One is entitled "What's the Point?" Wheelan wastes no time in making his case that statistics are ubiquitous, interesting, and extremely useful, even if the information they provide is imperfect and approximate. The focus is on the questions that statistical methods can help answer, rather than the techniques used – certainly an approach promoters of numeracy can identify with.

The Book

The contents of the book might be considered fairly standard, with a couple of exceptions, and give a reader familiar with the subject of statistics a good idea of what is coming. They also indicate that real-world examples will be front and center.

- 1) What's the Point?
- 2) Descriptive Statistics: *Who was the best baseball player of all time?*
- 3) Deceptive Description: *"He's got a great personality!"*
- 4) Correlation: *How does Netflix know what movies I like?*
- 5) Basic Probability: *Don't buy the extended warranty.*
- 5 ½) The Monty Hall Problem

- 6) Problems with Probability: *How overconfident math geeks nearly destroyed the global financial system.*
 - 7) The importance of Data: *"Garbage in, garbage out."*
 - 8) The Central Limit Theorem: *The Lebron James of statistics*
 - 9) Inference: *Why my statistics professor thought I might have cheated*
 - 10) Polling: *How we know that 64 percent of Americans support the death penalty (a sampling error of ± 3 percent)*
 - 11) Regression Analysis: *The miracle elixir*
 - 12) Common Regression Mistakes: *The mandatory warning label*
 - 13) Program Evaluation: *Will going to Harvard change your life?*
- Conclusion: *Five questions that statistics can help answer*
Appendix: *Statistical Software*

Chapters 2, 4, 9, 10, and 11 have appendices which get into the gory mathematical details, for those who are interested. The Appendix at the end of the book discusses the capabilities and usefulness of the software packages Excel, Stata, SAS, R, and IBM SPSS, a nice feature for instructors and those who may actually be taking a statistics class.

Without going through every chapter in detail, I will highlight some of the more interesting topics and examples, more or less in order of coverage.

Wheelan answers the question posed in the title of Chapter One with the following list at the end of the chapter.

- To summarize huge quantities of data.
- To make better decisions.
- To answer important social questions.
- To recognize patterns that can refine how we do everything from selling diapers to catching criminals.
- To catch cheaters and prosecute criminals.
- To evaluate the effectiveness of policies, programs, drugs, medical procedures, and other innovations.
- And to spot the scoundrels who use these very same powerful tools for nefarious ends.

These general points are illustrated by some of the questions posed in the chapter titles, as well as the following questions from within this first chapter:

- How can we catch schools that are cheating on their standardized tests?

- How can we figure out what substances or behaviors cause cancer, given that we cannot conduct cancer-causing experiments on humans?
- Does praying for surgical patients improve their outcomes?
- Is there really an economic benefit to getting a degree from a highly selective college or university?
- What is causing the rising incidence of autism?

Familiar statistics related to sports, criminal investigations, and economics are included in the examples, but I also appreciated Wheelan's discussion of the Gini Index (see Catalano et al. 2009 for another treatment of this index), a measure of inequality that has been receiving occasional mention in the popular press over the last several years.

Income inequality is discussed again in Chapter Two, which includes graphs for percentiles of income in the U.S. as one example of the general theme of the chapter: What do the standard summary statistics tell us, and what are their shortcomings? The shape and skew of data distributions are discussed, with Bill Gates serving as the iconic outlier. Wheelan does include discussion of the normal distribution here, example histograms which are skewed in various ways, and even trots out the formula for standard deviation. Various ways in which summary statistics and graphical representations can be deceptive are covered in Chapter Three, which includes a fairly thorough and cogent discussion of inflation and real versus nominal currencies.

Chapter Five begins with the story of a 1981 Schlitz Brewing Company ad campaign culminating in a blind taste test during the Super Bowl between Schlitz and Michelob refereed (literally) by an NFL official. Wheelan explains how Schlitz' marketing ploy exploited the (apparent) fact that most beer drinkers cannot really tell the difference between similar brands. Schlitz conducted the test using samples of 100 self-identified Michelob drinkers. The "apparent fact" and basic binomial probability meant that Schlitz could be highly confident the live test would support the desired headline: "Half of all Michelob drinkers like Schlitz better!" (I suppose this example could have gone in Chapter Three as well as Chapter Five.)

Other examples in this chapter include rare causes of death, lotteries, medical testing, and DNA analysis. I found Wheelan's explanation of the latter rather more intuitive and straightforward than that given in Schneps and Colmez's *Math on Trial* (2013), although that book is also quite good in general (Catalano 2014).

I appreciated the organization of Chapter Six, which provides section headings that give a list of common problems in thinking about probability.

- Assuming events are independent when they are not. This is illustrated by the same miscarriage of justice regarding SIDS deaths in Great Britain discussed in *Math on Trial*.
- Not understanding when events ARE independent, illustrated by sporting streaks of various kinds.
- Clusters happen. Here, the example is cancer clusters, and how seemingly significant findings can occur simply by chance.
- The prosecutor's fallacy, illustrated via DNA testing, again as in *Math on Trial*.
- Reversion to the mean (or regression to the mean). Examples from sports as well as investing. No mutual fund manager outperforms the Dow for long.
- Statistical discrimination. Is it fair to practice racial profiling or have gender-based differences in insurance premiums?

The latter example illustrates that statistics can only describe and possibly explain. They do not tell us what to do on the basis of those descriptions, nor which possible actions are fair. This is not part of numeracy per se, but is certainly relevant to the larger issue of critical thinking.

I like the analogy Wheelan makes in discussing the importance of data-gathering techniques in Chapter Seven. "Data are to statistics what a good offensive line is to a star quarterback." No amount of fancy analysis can make up for bad data. This chapter also includes sections which provide a nice list, in this case of common problems in data collection. These include selection bias, publication bias, recall bias, survivorship bias, and healthy user bias, all of which will be familiar to statistics and numeracy instructors. One story here related how women who were diagnosed with breast cancer changed their memories regarding what they were eating decades earlier, in effect creating a fictional but plausible explanation for the condition they found themselves in. In both Chapters Six and Seven, the sections are short and to the point, typically about one or two pages.

Chapter Eight provides an intuitive explanation of the power of the Central Limit Theorem, including graphs to illustrate the point. There are some technical points here, including a discussion of standard error. One could skip over this chapter without doing harm to one's understanding of the rest of the book.

Some of Wheelan's examples are both hypothetical and, by his own admission, implausible. In this chapter, he picks up one such example introduced earlier which relates how a group of marathoners and a second group of attendees of a sausage festival were randomly assigned to busses as

they traveled around Boston. The groups are kept separate, and none of them speak English. When a bus crashes, statistics is used to determine which group the riders belong to by considering the average weight and standard deviation of the 60-odd riders on each bus. In Chapter Nine, which covers statistical inference, one of the busses is hijacked by a pro-obesity terrorist group, and statistics is employed to determine if the commando group (led by Angelina Jolie in the movie version of this hypothetical event) is rescuing the right bus. All ends well when the terrorists are enthralled by the sausage festival, and agree to lay aside violence to promote such events throughout the world.

The subtitle of Chapter Nine stems from the author's personal experience of getting an A on his statistics final after, by his own admission, barely passing the course up to that point. The professor's implicit inference was reasonable from a statistical standpoint, but incorrect, illustrating both the power and shortcomings of such inferences. Statistics can never prove anything for certain, but can tell us about the likelihood of certain claims and explanations being correct, and how certain we can be they are correct.

Wheelan does use the standard language of hypothesis testing, null hypothesis, standard deviation, etc. (this is true here and with respect to other topics). Examples include deciding whether providing drug treatment to inmates reduces recidivism, testing the effectiveness of medical treatments, and the Atlanta standardized test cheating scandal of 2009. Several pages are devoted to a 2011 study in the *Archives of General Psychiatry* (Hazlett et al. 2011) on autism spectrum disorder, and how a study that includes only 97 subjects can actually produce results we can be statistically confident of.

A study in the *Journal of Personality and Social Psychology*¹ which rejected the null hypothesis, that ESP does not exist, was used to illustrate the pitfalls of hypothesis testing. Wheelan goes on to discuss Type I and Type II errors, using the analogy of convicting innocent defendants and acquitting guilty ones, a context that most readers would find very intuitive. He explains the trade-off between avoiding Type I Errors versus avoiding Type II errors by comparing spam filters to cancer screening. In the first case, a Type II error lets spam into your inbox, but that is probably preferable to losing an important e-mail to the filter, a Type I error. For cancer screening, however, a false-positive Type I error is clearly preferable to the alternative. Preventing terrorism is a situation where there is more tension between picking which error is more allowable.

¹ Wheelan cites a discussion of this study in a *New York Times* article (Zeleny and Thee-Brenan 2011).

Wheelan's interest in public policy and politics comes even further to the fore in Chapter Ten's discussion of polling. We also learn that people tend to have sex with others who are like themselves a few times a month in the "old-fashioned" way, and 80% of adults had sex with zero or one partner during the previous year. The appendix to this chapter explains why a sample proportion of 50% gives the largest standard error for a given sample size.

Chapters Eleven and Twelve cover regression. Perhaps referring to it as the miracle elixir is hyperbole, especially as Wheelan spends a good deal of time discussing confounding variables and other pitfalls of regression. He does get a bit more technical in these two chapters, giving examples of linear regression equations along with a residuals term e . For example:

$$\text{Weight} = a + b(\text{HEIGHT}) + e$$

There is also some discussion of the t-distribution, although the more-technical details and some relevant graphs are left to the appendix.

Chapter Eleven does get into multivariable regression as well. Wheelan does a nice job of intuitively explaining what the coefficients signify, through an example relating weight of individuals to not only height, but age, gender, race, education, and eligibility for food stamps. Yes, food stamp recipients are on average 5.6 pounds heavier than U.S. citizens overall.

Another perhaps surprising finding is related to the gender gap in pay equity. Women do tend to fall further behind men as they age, but this gap is not likely due to discrimination because of factors like time spent out of the work force by women, differences in what college classes men and women take, etc. (Bertrand, et al. 2009). The authors explain all but 1% of the wage difference by factors other than discrimination.

The warning label referred to in the title of Chapter Twelve reads "Do not kill people with your research." Wheelan goes on to discuss a Nurses' Health Study² of 122,000 women which showed how estrogen (hormone replacement therapy) was associated with lower rates of heart disease and osteoporosis. Millions of women underwent this therapy as a result of the study. When clinical trials were later done, the *New York Times Magazine* (Taubs 2007) pointedly asked how many women died because of hormone replacement therapy, since the trials showed an increased risk of other negative outcomes, including strokes and breast cancer.

Wheelan lists the top seven abuses of this 'otherwise extraordinary tool' as follows:

² See <http://www.channing.harvard.edu/nhs/> for information about The Nurses' Health Studies.

- Using regression to analyze a nonlinear relationship.
- Correlation does not equal causation.
- Reverse causality.
- Omitted variable bias
- Highly correlated explanatory variables (multicollinearity)
- Extrapolating beyond the data.
- Data mining (too many variables)

The last issue received mention at a plenary session of the National Numeracy Network's Annual Conference this past October at Carleton College. Jake Porway of DataKind was the main speaker. One audience member noted that if you consider enough variables, you will find lots of pairs giving you strong correlations just by chance (see Schield 2014). Porway allowed that this is one potential pitfall of 'big data.'

Concluding Comments

Wheelan's book is certainly a worthy addition to the genre of popular books on statistics. Examples from this genre include Best (2008, reviewed by Swingle 2009), Blastland (2009, reviewed by Root 2009), Donnelly (2007), and the classic *How to Lie with Statistics* (Huff 1954). This reviewer found *Naked Statistics* not only to be an enjoyable read, but also a book with the potential to promote numeracy among the general public, and a resource for instructors of numeracy and statistics. Examples from throughout the book, both real and hypothetical, serve to illustrate the usefulness of statistical techniques and understanding. Along with Wheelan's generally intuitive and insightful explanations, these examples can also serve to demystify statistics.

References

- Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz. 2009. Dynamics of the gender gap for young professionals in the corporate and financial sectors. *NBER Working Paper* 1468.
- Best, Joel. 2008. *Stat-Spotting: Field Guide to Identifying Dubious Data*. Berkeley: University of California Press.
- Blastland, Michael, and Andrew Dilnot. 2009. *The Numbers Game: The Commonsense Guide to Understanding Numbers in the News in Politics, and in Life*. New York, NY: Gotham Books/Penguin.
- Catalano, M. T, 2014. A sampling of books for *Numeracy* readers. *Numeracy* 7(1): Article 7.

- , T. L. Leise, and T. J. Pfaff. 2009. Measuring resource inequality: The Gini Coefficient. *Numeracy* 2(2): Article 4.
- Donnelly, Jr. 2007. *The Complete Idiot's Guide to Statistics*. New York: Alpha Books.
- Hazlett, H. C., M. D. Poe, G. Gerig, M. Sstyner, C. Chappell, R. G. Smith, C. Vachet, and J. Piven . 2011. Early brain overgrowth in autism associated with an increase in cortical surface area before age 2 years. *Archives of General Psychiatry* 68(5): 467–476.
<http://dx.doi.org/10.1001/archgenpsychiatry.2011.39>
- Huff, Darrell. 1954. *How to Lie with Statistics*. New York: W. W. Norton & Company.
- Levitt, Steven D., and Stephen J. Dubner. 2009. *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything*. New York: William Morrow Paperbacks.
- Root, Robert G. 2009. Two popular books for quantitative literacy: *What the Numbers Say*, and *The Numbers Game*. *Numeracy* 3(1): Article 9.
- Schild, Milo. 2014. Two big ideas for teaching big data: Coincidence and confounding. ECOTS invited paper. Statistical Literacy (www.statlit.org);
<http://statlit.org/pdf/2014-Schild-ECOTS.pdf>
- Schneps, Leila and Coralie Colmez. 2013. *Math on trial: How numbers get used and abused in the courtroom*. Basic Books.
- Taub, Gary. 2007. Do we really know what makes us healthy? *New York Times Magazine*, September 16.
http://www.nytimes.com/2007/09/16/magazine/16epidemiology-t.html?pagewanted=all&_r=0
- Swingle, Joe. 2009. Review of *Stat-Spotting: A Field Guide to Identifying Dubious Data* by Joel Best. *Numeracy* 2(2): Article 7.
- Wheelan, Charles. 2003. *Naked Economics: Undressing the Dismal Science*. New York: W. W. Norton & Company.
- . 2014. *Naked Statistics: Stripping the Dread from Data*. New York: W. W. Norton & Company.
- Zeleny, J. and M. Thee-Brenan. 2011. New poll finds a deep distrust of Government. *The New York Times*, October 2011.
<http://www.nytimes.com/2011/10/26/us/politics/poll-finds-anxiety-on-the-economy-fuels-volatility-in-the-2012-race.html?hp>