

2020

Art, Artfulness, or Artifice?: A Review of *The Art of Statistics: How to Learn From Data*, by David Spiegelhalter

Jason Makansi
jmakansi@gmail.com

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



Part of the [Other Mathematics Commons](#), and the [Statistics and Probability Commons](#)

Recommended Citation

Makansi, Jason. "Art, Artfulness, or Artifice?: A Review of *The Art of Statistics: How to Learn From Data*, by David Spiegelhalter." *Numeracy* 13, Iss. 1 (2020): Article 9. DOI: <https://doi.org/10.5038/1936-4660.13.1.9>

Authors retain copyright of their material under a [Creative Commons Non-Commercial Attribution 4.0 License](#).

Art, Artfulness, or Artifice?: A Review of *The Art of Statistics: How to Learn From Data*, by David Spiegelhalter

Abstract

David Spiegelhalter. 2019. *The Art of Statistics: How to Learn From Data*. (London: The Penguin Group). 444 pp. ISBN 978-1541618510

The author successfully eases the reader away from the *rigor* of statistical methods and calculations and into the *realm* of statistical thinking. Despite an engaging style and attention-grabbing examples, the reader of *The Art of Statistics* will need more than a casual grounding in statistics to get what Spiegelhalter, I believe, intends from his book. It should be viewed as a companion to a more rigorous textbook on statistical methods but not necessarily a book that makes statistics any less complicated for the uninitiated.

Keywords

statistics, quantitative literacy, data, critical thinking

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

Jason Makansi, a member of the National Numeracy Network, is the author of the award-winning *Painting By Numbers: How to Sharpen Your BS Detector and Smoke Out the Experts*. Currently, he is assisting a new high school in Tucson with innovative curricula concepts such as numerical literacy. He earned his BS in Chemical Engineering from Columbia University and runs an independent consulting firm that specializes in technology deployment for the electricity industry. He has published five professional books, hundreds of articles in academic and professional publications, one novel, and numerous short stories.

The drumbeat for better quantitative literacy (QL) among the general population is beginning to sound more like Keith Moon (The Who) than Ringo Starr (The Beatles) – less like just keeping a steady beat in the background and more like an equal musician vying for the attention of the audience.

Partly, this trend can be attributed to a rising concern generally with new forms of literacy — computer literacy, media literacy, government literacy, climate literacy — at a time when literacy is code for citizens not possessing the essential tools for making better decisions about what’s going on around them politically, culturally, and commercially.

In the case of statistical literacy (as a subset of QL), the digitization of our lives is leading to an explosive growth of, interest in, and reliance on, data science. The swirling miasma of data science, in other words, is far outstripping the minimal tools citizens possess to deal with it.

This modest introduction gets at three of Spiegelhalter’s key words in the title of his book – statistics, learn, and data. Later, I will have more to say about the fourth key word (art).

The author apparently wrestled with himself, and his editor, about the title. In the acknowledgments section, he notes that even after the book was finished, “[he and his editor] still could not agree on a title.”

Maybe the first question about the title that needs an answer is: who? Who is Spiegelhalter trying to help “learn from data?” Based on the book’s jacket copy (hardback edition), the answer appears to be people who understand a bare minimum about statistics, but need more guidance in order to “think like a statistician.”

It seems clear from the selection, presentation, and organization of the content that the target readers are those who while already familiar with statistical concepts, aren’t making connections among these concepts and what goes on around them every day. “Improving data literacy means changing the way statistics is taught (12).” Thus, my sense is that Spiegelhalter intends *The Art of Statistics* to be a better and/or supplementary textbook for teaching statistics. If readers without some formal engagement with statistics get something out of it, all the better.

The flap copy notes that “working through a range of both practical and exciting examples (without using mathematics), [Spiegelhalter] shows how to better prepare for life and the challenges of a data-driven society.” My assessment, however, is that the reader most definitely needs to be comfortable with math and statistical terms, even if the reader isn’t required to *solve* math problems as part of the learning experience. In other words, even though the math is in the background, you certainly are having to “use it.”

One of the most appealing features of *The Art of Statistics* is how the author threads the same examples through the various statistical concepts he is trying to explain. For example, the book begins with Harold Shipman, “Britain’s most

prolific convicted murderer (1),” using statistical science to get insight into this question: “What kind of people did Harold Shipman murder, and when did they die” (2)? The Shipman example arises in later chapters, and in Chapter 10, Spiegelhalter delves into whether formal statistical analysis could have helped catch Shipman earlier.

I appreciate that the author immediately captures the limitations of statistical science from the beginning of his text. Statistical science is all about making better judgements and getting better insight, not final answers or definitive conclusions.

The second question the author poses is “How many trees are there on the planet” (7)? Immediately, the author conveys statistical analyses’ reliance on assumptions. As Spiegelhalter notes, before we can answer this question we must first determine “what is a tree?” and how does it differ from a bush or a shrub? Statistics can help us estimate the number of trees on the planet but only if we can agree on terminology. As the author points out, the fifty American States don’t even agree on the “legal definition of death” (9). Given such challenges that transcend statistical methods, Spiegelhalter wonders how it is possible that statistics can be used to determine how happy people were yesterday, as a survey of 150,000 people in the UK tried to determine. “These examples show that statistics are always to some extent constructed on the basis of judgements” (9).

In addition to death and trees, the author explores as examples in later chapters: children having heart surgery (19), cancer risk from bacon sandwiches (31), jelly beans in a jar (41), number of sexual partners (51), global population growth (61), college education and brain tumors (95), prayer (105), passengers surviving the Titanic (148), and many others. You probably get the picture – illustrate exceedingly dry concepts in statistics through non-threatening, emotional, and/or even titillating topics.

The technique is effective to a limited extent. There’s no getting around the fact that statistical concepts are difficult to grasp, even for those who are inclined to math and apply math concepts regularly. To his credit, Spiegelhalter tackles most of the essential statistical concepts – from mean, median, and mode to Bayesian statistics.

From the perspective of pure statistics, the author, I believe, has successfully eased the reader away from the *rigor* of statistical methods and calculations and towards the *goal* of statistical thinking. But no question, the book is still (and does not purport to be anything but) a statistician’s view of the world.

In chapter 2, the author poses the question “Can we trust the wisdom of crowds?” (39). He then describes a “trivial experiment” in which he posted a YouTube video showing a jar filled with jelly beans and asked people to guess the number inside. (He also suggests readers try it themselves). The 915 responses are then subjected to statistical analysis to gain some insight into the wisdom of crowds.

So I took him up on it. Being an engineer, my response was to estimate the number of jelly beans by instead guessing at the dimensions of the container, assuming a shape for it, calculating the volume, estimating the volume of one jellybean, estimating and then determining how many beans would fill that volume, accounting for empty space.

Then I tried a different method; I visually tried to count the number of jelly beans at the bottom of the two sides of the jar which are visible in the photo, estimate the number in a flat layer at the bottom, and multiply this figure by the number of jelly beans from the bottom to the top. My assumption here was that layers of jelly beans - imagine the floors of a tall building - would get me close. Interestingly, the first method grossly overestimated the correct number, and the second method underestimated it by more than half. My point in dwelling on this example is that it is important to recognize that there are multiple ways to think through a problem.

Spiegelhalter reports that of the 915 responses, 18 are the same number, 10,000, a detail that is clear when a histogram is plotted of the responses (42). It's like a skyscraper in the middle of the suburbs outside of the city center. Apparently, 10,000 is a nice large round number people will guess at when staring at a jar of jellybeans, which perhaps says more about human behavior than the wisdom of crowds. For those eager to know, the true value was 1,616, the median group-guess was 1,775, and one person guessed precisely (48).

Following the opening chapters on proportions, percentages, bar and pie charts, icon arrays, bacon, jelly beans, means, medians, and modes, things begin to get challenging in a hurry: standard deviations, inter-quartile range, Pearson's correlation coefficient, Spearman's rank correlation, and others. And we're only to Chapter 3! Later chapters deal with causality, regression analysis, algorithms, pattern recognition, probabilities, curve-fitting, variability, the central limit theorem, confidence intervals, null hypotheses, *p*-values, and Bayesian statistics (Spiegelhalter is a fan!) — with the exception of the last, all the topics one would expect to find in a text on statistics.

The important thing to realize is that the author writes about these complicated concepts in a style that is most accessible. Only in a few passages does he slip into dry academic prose, and it's likely unavoidable. Even at the end of these passages, though, the “takeaway” for the reader is crystal clear. A good example is when the author shows that even though a screening test for doping in sports claims to be 95% accurate, “the majority of people who test positive are in fact innocent” (310).

I wish Spiegelhalter offered more of what he wrote about election surveys: “My personal, rather skeptical heuristic is that any quoted margin of error in a poll should be doubled to allow for systematic errors made in the polling” (245). This is a useful rule of thumb. I assume it is based on his considerable experience as a statistician. Similar rules applied, for example, to clinical trials (such as the ones

referenced in pharmaceutical advertisements), diet and nutrition claims, and economic data would add real value for the true lay reader.

Spiegelhalter misses an opportunity to make his book even more useful to the non-stats person. Each statistical “term” introduced in the text is in boldface type. Then there is a glossary of these terms (381) and an index which also includes these terms (419). I would have loved to see a list of the top twenty terms and concepts used in statistics, with a brief discussion of how the reader should think about each—you know, the terms that could net 95% of all the instances the more casual reader might need to apply statistics in everyday life.

When it comes to doing statistics better, the author rightfully identifies three groups who must all be involved: producers of statistics, communicators, and audiences (362). (The latter two I would term consumers of statistics.) And in a passage probably most useful to the non-stats reader, he offers ten questions to ask when confronted by a claim based on statistical evidence, categorized by how trustworthy are the numbers, how trustworthy is the source, and how trustworthy is the interpretation (369-371).

So let’s turn to the fourth key word in the title: art. The following passage about smoking and lung cancer does a good job summing up what we are usually left with at the end of exhaustive statistical analysis: “...the medical community now agrees that smoking cigarettes causes lung cancer, but it took decades for doctors to come to this conclusion. Why did it take so long? Because most people who smoke do not get lung cancer. And some people who do not smoke do get lung cancer. All we can say is that you are more likely to get lung cancer if you smoke than if you do not smoke...” (98).

Consider all the gnashing of teeth, research conducted, money spent, energy expended, and statistical analyses performed over decades. The medical community does not even associate a percentage number with this risk. But this does illustrate the “art” of statistics. As the author states “...there is unavoidable variability that underlies everything interesting in real life” (98). Statistics helps you ask more and better questions, even if they don’t provide simple black-white answers we often wish for.

Implicit in much of Spiegelhalter’s prose, however, is the “artfulness” of statistics. “Artful” adds a performance dimension, and (according to my on-line Merriam’s dictionary) “an insinuating or indirect means of attaining an end.” Since it is darn near impossible in so many cases to prove a direct cause and effect relationship in “everything interesting in real life,” statistical analysis offers an indirect approach. And for the most part, the expert doing the statistical analysis is “performing” for someone or a group that usually has a purpose.

“The ability to assess the trustworthiness of statistical claims seems a key skill in the modern world, and I hope that this book may help to empower people to question the numbers that they encounter in their daily life” (17). Spiegelhalter also

quotes Nate Silver, *The Signal and the Noise*, at the opening of the book's introductory chapter: "the numbers have no way of speaking for themselves. We speak for them. We imbue them with meaning" (1).

As the author makes clear from these two quotes, sometimes statistics is art, often it is artfulness, even artifice. When evaluating statistical analysis accompanying a study, report, and media account, readers of *The Art of Statistics* should be better equipped to distinguish among the three. But, as much as Spiegelhalter surely hopes that his book helps regular people "think like a statistician" about the real world, he certainly suspects it may not, based on his own conclusion: "To put it bluntly, statistics can be difficult. Although I have tried to tackle underlying issues in this book rather than getting embroiled in technical detail, the narrative has unavoidably had to rely on some challenging concepts. So congratulations for reaching the end" (379).

In sum, I think the reader of *The Art of Statistics* will need more than a casual grounding in statistics to get what Spiegelhalter intends from his book. It should be viewed as a companion to a more rigorous textbook on statistical methods, and therefore of keen value to most readers of *Numeracy*.

Nevertheless, embedded in Spiegelhalter's book is a sequel on the order of "statistics for people in a hurry (paraphrasing Neil DeGrasse Tyson's excellent, brief, *Astrophysics for People in a Hurry*). Perhaps that is the "wild man drummer" that could finally capture the lay public's attention.

Maybe this passage best sums up Spiegelhalter's alter ego on statistics: "Personally, I rather like acting as if all that occurs around us is the result of some random pick from all the possible things that could happen. It is up to us whether we choose to believe it is truly chance, whether it is the will of a god or gods, or any other theory of causation: it makes no difference to the mathematics. This is just one of the mind-stretching requirements for learning from data" (93).

"We imbue [the numbers] with meaning."

References

Silver, Nate. 2012. *The Signal and the Noise – Why So Many Predictions Fail, but Some Don't*. New York, NY: The Penguin Group.