Client: Penguin Press
Source: Financial Times
Date: 06/05/2019
Page: 22
Reach: 180915
Value: 19339.32

A statistically significant data advocate

Book review

Alan Smith

There is just one problem with famous statisticians: they tend not to be very famous. One notable exception is the late Hans Rosling, who made Time Magazine's list of the world's most 100 influential people in 2012. Perhaps his closest living equivalent is now Sir David Spiegelhalter, former president of the UK's Royal Statistical Society and current Winton professor of the public understanding of risk at the University of Cambridge.

The comparison with Rosling is easy to make, not least because Spiegelhalter is humorously critical of his own field which, by his reckoning, has spent too much time arguing with itself over "the mechanical application of a bag of statistical tools, many named after eccentric and argumentative statisticians".

His latest book, its title, *The Art of Statistics*, laced with oxymoronic appeal, is ambitiously intended for students of statistics and general readers. To his credit, there is indeed something in here for everyone.

Introductions to topics such as correlation, regression and probability successfully use real world stories to make their point without resorting to mathematical formulas. You will learn the identity of the luckiest passenger on the Titanic, whether serial killer Harold Shipman could have been caught earlier, and if it really was Richard III's body discovered under a Leicester car park.

There's a good balance of depth and pace: useful summaries at the end of each chapter reinforce key points without being overbearing. There is an accessible technical glossary. And although the book's chapters become progressively more challenging for general readers, there's never too far to jump.

The author likes charts — another parallel with Rosling. The book is sprinkled with examples of how to read and present data in visual form. But for Spiegelhalter, this is only part of the secret: "The first rule of communication is to shut up and listen." Trying to make an audience smile using data alone is difficult, but Spiegelhalter achieves it repeatedly. Describing the tendency of respondents to a sexual attitudes survey to round the number of sexual partners they have had to the nearest five or 10, the author highlights "the rather pedantic man, possibly a statistician, who said precisely '47'."

As engaging as such examples are, it is the book's conceptual and philosophical focus that provides greatest value, especially to the serious student. It arrives at a time of existential crisis — for both the science of statistics itself as well as the broader use of data and evidence in public debate and decision-making.

Mired for decades in an old-fashioned system of rote learning, statistics has threatened to collapse under the weight of its own reliance on p-values, its standard (if often misunderstood) unit of currency for the past half century.

Meanwhile, the emergence of big data and "data science" threatens to overtake it with a nimbler, multidisciplinary approach to real-world problem solving. It is clear the direction Spiegelhalter would like to see the field take. He is one of more than 850 signatories to an article in Nature magazine calling for an end to "statistical significance" and its reliance on p-values.

The book's closing chapter focuses on the future of statistics. Spiegelhalter makes it clear that responsibility lies with three distinct groups: producers of statistics, communicators and audiences (educators could be added to the list). And therein lies part of the broader appeal of the book: it's a call to arms for greater societal data literacy.

Readers expecting software tutorials will be disappointed — there are none. On balance, this seems a wise move by the author, ensuring the book has a shelf life that is not threatened by rapid changes in the software landscape. Nevertheless, it does mean the book stops slightly short of the full "what, why and how" experience.

I once teased a statistician I worked with that the term data science was only coined to usurp statistics, an anachronistic discipline rooted in the past, paralysed by a refusal to evolve. Spiegelhalter's work serves as a reminder that there are passionate, self-aware statisticians who can argue eloquently that their discipline is needed now more than ever.

The reviewer is the FT's data visualisation editor

Client: Penguin Press
Source: Financial Times
Date: 06/05/2019
Page: 22
Reach: 180915
Value: 19339.32

The Art of Statistics
David Spiegelhalter
Pelican Books, £16.99