



Naked Statistics: Stripping the Dread from the Data

Charles Wheelan

[Download now](#)

[Read Online ➔](#)

Naked Statistics: Stripping the Dread from the Data

Charles Wheelan

Naked Statistics: Stripping the Dread from the Data Charles Wheelan

The best-selling author of *Naked Economics* defies the odds with a book about statistics that you'll welcome and enjoy.

Once considered tedious, the field of statistics is rapidly transforming into a discipline that Hal Varian at Google has actually called sexy. And with good reason from batting averages and political polls to game shows and medical research, the real-world application of statistics is growing by leaps and bounds. In *Naked Statistics*, Charles Wheelan strips away the arcane and technical details to get at the underlying intuition that is so key to understanding the power of statistical concepts. How can we catch schools that cheat on standardized tests? How does Netflix know which movies you'll like? What is causing the rising incidence of autism? Statistics can be used to tackle these questions and others that are incredibly important and relevant to us today. With the trademark wit, accessibility, and sheer fun that turned *Naked Economics* into a bestseller, Wheelan brings another essential discipline to life with that one-in-a-million statistics book that you'll read for pleasure.

Naked Statistics: Stripping the Dread from the Data Details

Date : Published January 7th 2013 by W. W. Norton & Company (first published December 31st 2012)

ISBN :

Author : Charles Wheelan

Format : Kindle Edition 305 pages

Genre : Nonfiction, Science, Business, Economics, Mathematics



[Download Naked Statistics: Stripping the Dread from the Data ...pdf](#)



[Read Online Naked Statistics: Stripping the Dread from the Data ...pdf](#)

Download and Read Free Online Naked Statistics: Stripping the Dread from the Data Charles Wheelan

From Reader Review Naked Statistics: Stripping the Dread from the Data for online ebook

DonkeyPopsicle says

There are many popular science books that try to teach basic statistical concepts, but more often than not they fall into the awful popular science trope of narrative over concepts that Malcolm Gladwell introduced into science writing and then Jonah Lehrer perfected into an awful, horrible art. Take Nate Silver's lauded book 'The Signal and the Noise'. Each chapter is about some specific area of prediction, and along the way some statistical concepts are introduced but rarely elaborated [I will note that Nate Silver only rarely mentions what the expert had for lunch during their interview, unlike much worse science books that presume we are interested in the culinary habits of scientists]. In that book, Silver also tries to make a case for Bayesian statistics over traditional statistics, but because the explanation of the concepts is not very rigorous, we don't get so much an argument as an opinion.

Charles Wheelan's book is a fantastic antidote to modern popular science writing and conceptual hand-waving. In a nutshell, the book is a stats 101 course without the math. Unlike, say, popular physics books where understanding can only be vaguely metaphorical at best without knowing quite a bit of advanced mathematics (giving the illusion of knowledge; yes, you've read "The Elegant Universe", but sorry, you still know bupkis about string theory), statistical concepts can be explained and even employed in a critical fashion without much math at all. Knowing that variation is much more informative than simply the mean doesn't require that you know calculus. Likewise for understanding simple experimental design (and most experimental designs are simple: state a null, apply Student's t-test, and you've got 70% of published scientific papers).

Of course, saying that something can be explained without math is not the same as actually doing it proficiently, but Wheelan has excelled here. The examples are all intuitive, and the writing is clear and easy. Perhaps more importantly, Wheelan spends an entire chapter on the Central Limit Theorem halfway through the book, and then uses that to explain statistical inference, sampling, and regression. Giving the Central Limit Theorem such pride of place is appropriate but is often neglected in basic statistics textbooks (not to mention popular statistics books).

The book is not flawless, but the quibbles are minor. First, Wheelan has a silly sense of humor that intrudes into the book too often (culminating in several pointless footnotes that only serve to extend jokes). Second, although there are a few mathematical appendices for various chapters, they are generally far too short and actually need more math than they have. As it is, they are likely to confuse more than help.

In general, Wheelan's book is a must read for anyone that hasn't taken a basic stats course (so every journalist ever) or can't remember much from when they did take it.

Wen says

A solid five-star. If only I had had Charles Wheelan as my college statistics professor! :)

The synopsis on Goodreads was a good review, so I'll save some ink here.

These were all basic statistic concepts, from probability to regression. While breaking down the basic concepts, Wheelan sought to cultivate intuition around them. And he did a fantastic job. Better yet, he made

me chuckle all the time with those funny, sometimes provocative real-life examples.

Some of the examples, like the route cause of the 2008 financial crisis and the Monty Hall problem, have been widely telegraphed; I'd assume Wheelan's explanations will make them easier to sink in.

I took statistics classes during two phases of my education, but am currently using little of it at work. Given big data is on the rise, and large free data sets are becoming more obtainable, I'm toying the idea of taking data crunching as a pastime.

To challenge myself, I opted for the audiobook and played it at double speed. Not until double-digit chapters when I had to pause and rewind, because keeping four-digit numbers in my head became impossible on the noisy subway. That said, Wheelan did not lose a beat, or his cool, in breaking down more difficult hypothesis testing and multivariate regression for a lay person.

I applaud making audiobook available for math-related subjects; it would benefit visually impaired students who might otherwise find math's too daunting. On that note, Jonathan Davis did a wonderful job with his smooth narration and sense of humor pretty much syncs with the author.

Elizabeth Theiss says

An amusing, clear, and even fun introduction to basic statistics and probability, this gem explains foundational concepts and provides compelling examples to illuminate them. It covers correlation, normal distributions, the central limit theorem, significance, standard error, multiple regression, and so on in a way that math-phobes can likely handle without panic attacks. I wish I had read this before taking grad stats.

The truth is that students of statistics today can use Excel, SPSS, Stata and similar programs to calculate statistics. So the focus has changed from working formulas to gaining a deep understanding of the assumptions, meaning, and limitations of statistics. Wheelan's excellent book provides the background for this understanding in a readily comprehensible way.

Caitlin says

Being a mathematics and statistics teacher, of course I am inclined to enjoy a statistics book. There were times I found myself a bit bored because I was being explained basic statistical concepts of which I already possess a wider understanding.

This book is an excellent recommendation to students just starting statistics as it gives practical and engaging examples of statistics and easy to follow. For those who already have a broad understanding of statistical topics as well as commonly used examples of probability, this book can be repetitive and frustrating.

It also bothered me that at more than one time throughout the book Wheelan assumes that the reader does not wish to know more about these topics and so he has processed information so we don't have to. These are more of the things I would have liked to know more about but then perhaps I should have been reading a different book if that was what I was desiring.

Overall, a well written book about probability and statistics. Useful to anyone who is just starting a journey in statistics as it breaks down various statistical concepts that can be confusing to a beginner.

Sebastian says

This book is brilliant! Although, I'd say that I already have a relatively solid statistics background (currently doing computational biology and machine learning research), I got a lot out of it: Beyond the basic concepts (sampling, central limit theorem, hypothesis testing, linear and multiple linear regression, just to name a few), it is very rich in useful, hands-on examples. I'd even go so far and would recommend it as a first book for someone who is new to the field -- prior to reading any stats textbook since it contains so many illustrative examples and explains the common pitfalls -- often in context of real-world research findings -- very, very well. However, I believe that even the most experienced stats person will be properly entertained by this awesome book!

PS: I got the audio book version and listened to it on my 6-hour drive to Toronto; although, the book contains some diagrams, it works perfectly well as an audio book.

Mook Woramom says

????????????????????????? ????? bias ???
????????????? 5555
??
??
????????????????
????????????????? ?????? ?????? ?????????????????????? ?????????????? ?????????????? ?????????????
????????????????????????? ?????? ?????????????????????? ?????? ?????????????????? ?????? ?????????????????? ????

Herve says

I have already talked about statistics here, and not in good terms. It was mostly related to Nicholas Nassim Taleb's works, The Black Swan and Antifragile. But this does not mean statistics are bad. They may just be dangerous when used stupidly. It is what Charles Wheelan explains among other things in Naked Statistics. Naked Statistics belongs to the group of Popular Science. Americans often have a talent to explain science for a general audience. Wheelan has it too. So if you do not know about or hate the concepts of mean/average, standard deviation, probability, regression analysis, and even central limit theorem, you may change your mind after reading his book. Also you will be explained the Monty Hall problem or equivalent Three Prisoners problem or why it is sometimes better (even if counterintuitive) to change your mind.

Finally Wheelan illustrates why statistics are useless and even dangerous when the data used are badly built or irrelevant (even if the mathematical tools are correctly used!). Just one example in scientific research (which is another topic of concern to me) "This phenomenon can plague even legitimate research. The accepted convention is to reject a hypothesis when we observe something that would happen by chance only 1 in 20 times or less if the hypothesis were true. Of course, if we conduct 20 studies, or if we include 20 junk variables in a single regression equation, then on average, we will get 1 bogus statistically significant finding. The New York Times magazine captured this tension wonderfully in a quotation from Richard Peto, a medical statistician and epidemiologist: "Epidemiology is so beautiful and provides such an important perspective on human life and death, but an incredible amount of rubbish is published".

Even the results of clinical trials, which are usually randomized experiments and therefore the gold standard

of medical research, should be viewed with some skepticism. In 2011, the Wall Street Journal ran a front-page story on what it described as one of the "dirty little secrets" of medical research: "Most results, including those that appear in top-flight peer-reviewed journals, can't be reproduced. [...] If researchers and medical journals pay attention to positive findings and ignore negative findings, then they may well publish the one study that finds a drug effective and ignore the nineteen in which it has no effect. [...] On top of that, researchers may have some conscious or unconscious bias, either because of a strongly held prior belief or because a positive finding would be better for their career. (No one ever gets rich or famous by proving what doesn't cure cancer. [...] Dr. Ioannidis [a Greek doctor and epidemiologist] estimates that roughly half of the scientific papers published will eventually turn out to be wrong." [Pages 222-223]

Jacob says

Another good Wheelan, similar to Naked Economics: Undressing the Dismal Science. There are a few topics that overlap a bit, but the author does a good job of keeping them separate. This has much of the personal anecdotes / history that make the topic more interesting, and the author includes more silly scenarios in this one which keep you engaged, such as the continually missing & crashing buses of marathon runners and sausage festival attendees. Unfortunately, the third quarter of the book gets a bit dry and stultifying when the author discusses Inference, Polling, and Regression Analysis, but the last quarter redeems itself with discussions of regression analysis mistakes and how statistics is being used to address social problems. Those discussions are once again engaging and meaningful, and include issues such as the increase in autism rate, how to tell good teachers from bad and change teacher's pay to reward the good ones, and how to reduce global poverty (this is one of the slightly overlapping bits with Naked Economics).

I also found the numerous discussions of how statistical inference can go wrong to be extremely helpful -- like any tool, you have to know what it's good for and when. Wheelan brings up that even in peer-reviewed medical journals, many if not most of the findings can't be repeated. There's also mention of a paper that finds about half the peer-reviewed papers (remember, these are supposed to be the vetted ones) are wrong, with the irony noted that if the author's right, then he's just as likely to be wrong.

Patrik says

How good is this book? After reading "Naked Statistics" I wanted to teach an introductory statistics course!

I could see myself engaging the students with really cool stories, confuse them with fun probability examples, only to wittily explain it clearly a minute later. I would pursue the connection between probability and inference and they would all clearly understand hypothesis testing. I would give great tales of statistics being misused and the students and I would chuckle together over how famous researchers made mistakes. Then we would promise to never make such silly mistakes ourselves.

I would never show the students "Naked Statistics," always pretending that I came up with the examples by myself... The students would love me. The administration would love me. I would receive awards, promotions, and (finally) significant pay raises. That is how good this book is. Too bad I don't teach statistics...

Jenne says

This is not the most exciting book ever, but it's way more exciting than you would think for a book about statistics.

More importantly, people: YOU NEED TO KNOW THIS STUFF. This is how you separate the lies from the damn lies from the nonsense that TV news shows spew at you. I don't care if you read THIS one, but please just fucking read a book about statistics. THANK you.

Youghourta says

???? ??????. ??? ??????? ??? ??????? ??? ?????????? ??? ?????????? ??? ?????????? ????.
????????? ??? ?????? ??? ?????? ??? ?? ?????? ?? ?????? ?? ?????? (??????? ??? ?? ??? ?????? ????) ???
???? ?? ?????? ?????? ??? ??????? ??? ??????? ?????????? ?? ?????? ??????? ??? ??????? ?????? ???
????? ?????? ?????? ?????? ??? ?????? ?? ?????? ??? ?????? ?????????? ??? ?????? ??????? ?????? ????.

?? ?????? ?????? ??????? ??? ?????? ??? ?? ??? ?????? ?? "???? ??????" ??? ?????? ?????? ?? ???
????? ?????? ??? "??????? ????" ?? ?????? ?????? ??? ??? ??????? ??????? ??? ?????? ??? ?????? ?????????? (??????
????? ??? ???????) ?????? ??? ??????? ??????? ??? ??? ??? ?????? ??? ?????? ??? ?????? ??? ?????? ??????? ???
????? ?????? ?????? ??? ??? ??????? ??? ?? ?????? ??????? ?????? ?? ?????? ?????? ?????? ??? ?????? ??? ???
????? ?????? ?????? ??? ??? ?????? ?????? ?????? ??????? ??? ?? ??? ??? ?????? ??? ?????? ??? ?????? ??? ???
????? ??? ??? ?????? ?????? ?????? ?????? ??????? ??? ?????? ??? ?????? ??? ?????? ??? ?????? ??? ??? ???
????? ?????? ??? ??? ?????? ?????? ?????? ?????? ??????? ??? ?????? ??? ?????? ??? ?????? ??? ?????? ??? ??? ????.

????? ?????? ?????? ?? ??? ?????? (?? ?????? ?? ??? ???) ??? ?????? ?????? ?????? ?????? ?????? ?????? ?????? ?????? ??????.

Diego Eis says

A nota foi 3, porque não é possível colocar 3.5.

Ótimo livro pra quem está começando ou apenas se interessa por estatística. Eu estava buscando exatamente um livro como esse para entender melhor conceitos técnicos. O objetivo não era aprender estatística a fundo, o que você não faria com esse livro, mas entender que estatística, de verdade, é difícil, não pela matemática envolvida, mas porque quem faz estatística precisa ser atenta, sagaz e sensata.

O livro vai ficando cada vez mais técnico da metade pra Frente. O início explica tudo o que você pode encontrar na internet como estatística descritiva, modelos de posições, etc etc. já no final ele apela para as armadilhas das análises de regressão.

Algumas anotações:

- Entender o objetivo do que se quer saber com a estatística. Nem sempre os números puros conseguem te

trazer respostas verdadeiras. Você precisa primeiro entender qual o objetivo que se quiser buscar pra entender se os números disponíveis conseguem te dar as respostas esperadas.

Por exemplo: descobrir se fumar causa câncer não é uma tarefa facil. Não é só separar dois grupos de pessoas que fumam (grupo de pesquisa) e outra que não fumam (grupo controle) e ver no que da. Tem que isolar exatamente a diferença entre os dois grupo. O grupo controle pode ter uma dieta ruim ou outros hábitos que podem estragar a pesquisa.

- Coeficiente de Gini mede a o quanto a riqueza (ou renda) é partilhada equitativamente dentro de um país numa escala de 0 a 1. A estatística pode ser calculada para a riquezas ou para a renda anual a nível individual ou familiar. Quanto mais perto de zero, melhor. Se toda a riqueza está concentrada em uma família, o coeficiente é um e mostra desigualdade. Se toda a riqueza está distribuída de forma igual a todos as famílias do país, o coeficiente é zero.

- "a estatística descritiva existe para simplificar, o que sempre implica alguma perda de nuance ou detalhe."

- "uma função chave da estatística é usar os dados que temos sobre perguntas mais amplas para as quais não temos informação completa. Em suma, podemos usar dados do mundo conhecido para fazer inferências informadas sobre o mundo desconhecido."

- estatística descritiva pode ser totalmente correta, mas pode trazer insights enganosos por causa do "erro das médias".

- o básico da estatística é encontrar o número moderado ou o valor do meio de uma distribuição que pode tentar trazer neutralidade na comparação de dois cenários. A medida mais básica do meio de uma distribuição é a "média".

- para distribuições sem valores atípicos sérios, ou seja, sem valores muito altos ou muito baixos do que os outros valores da distribuição, a mediana e a média serão semelhantes.

- valores absolutos geralmente podem ser interpretados dem qualquer contexto ou informação adicional. Por exemplo: saber que alguém precisou dar 80 tacadas para acertar 18 buracos no golfe. Com essa informação absoluta já é possível avaliar a performance do jogador, sem comparar outro jogador.

- valores relativos do tem significado em comparação com alguma outra coisa, ou num contexto mais amplo. Se o golfista está em nono lugar no ranking, podemos compará-lo com os oito golfistas que tiverem melhor resultado que ele em um mesmo campo, mesmo horário, mesma temperatura, etc.

- Esse resultado comparativo fica muito mais interessante se pegarmos o jogador que precisou das 80 tacadas e compararmos ele numa distribuição de percentil com os outros jogadores, conseguiremos saber se acertar 18 buracos com 80 tacadas é algo bom ou ruim. Dele cair num percentil de 90, quer dizer que ele está acima de 90% dos jogadores daquele grupo.

- A média é influência pelas medidas extremas, pelas dispersões. Um valor muito alto ou muito baixo pode influenciar a média facilmente

- a mediana não tem influência de valores extremos, mas ela não mostra o que está nos extremos

- do ponto de vista da acusação, a questão mediana versus media hora em torno de se os valores extremos numa distribuição distorcem o que está sendo descrito ou não, ao contrario, parte importante da mensagem

- se quiser descrever um grupo de números de uma forma que o faça parecer grande, focalize na média. Se quiser fazer que se pareça menor, focalize a mediana.

- a mediana não pesa observações de distância que os números se situam do ponto médio, apenas se estão acima ou abaixo

- a correlação mede o grau em que dois fenômenos estão relacionados entre si

- a Netflix consegue recomendar filmes para deus usuários porque ela relaciona os filmes que o usuário avaliou bem, com filmes assistidos e avaliações e filmes assistidos de outros clientes.

- probabilidade é o estudo de eventos e resultados envolvendo um elemento de incerteza

- probabilidades não nos dizem o que acontecerá com certeza, dizem o que é provável de acontecer e o que é menos provavelmente de acontecer.

-

Robert Muller says

I couldn't get through this book, mainly because I know too much about statistics and I know too much about the specific examples he gives to illustrate his points. Unfortunately, while at times Wheelan does convey the underlying concepts of probability and statistics in a way that would help you understand them at a basic level, he does so in what I would regard as a patronizingly oversimplified way. If you compare this book to Nate Silver's book on prediction or, indeed, to the book he says motivated him (How to Lie with Statistics), this book simply doesn't deliver the goods. It clothes the concepts of statistics in yet another layer of misunderstanding and half truth. If, for example, he had spent a chapter on "unemployment" and really showed how, as a descriptive statistic, the number is meaningless for all kinds of measurement and theoretical reasons, I would have been impressed. Instead, he used it as an example of a good statistic. If he had cited Savage's "The Flaw of Averages" while making points about averages, dispersion, and distributions (the wrong points, I might add), I would have been impressed. If he had at least *mentioned* Bayes Rule and Bayesian statistics, I would have been impressed. I wasn't impressed.

The Serendipity Aegis ~ ?Misericordia? ?????? ✶✶♥? says

Cool and easy. Even though I would have liked more advanced stuff made easy and cheesy.

Rakibul Islam says