Doing Data Science: Straight Talk from the Frontline

By Cathy O'Neil, Rachel Schutt

Doing Data Science: Straight Talk from the Frontline By Cathy O'Neil, Rachel Schutt
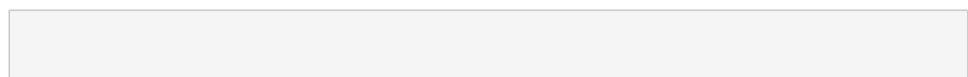
Now that people are aware that data can make the difference in an election or a business model, data science as an occupation is gaining ground. But how can you get started working in a wide-ranging, interdisciplinary field that's so clouded in hype? This insightful book, based on Columbia University's Introduction to Data Science class, tells you what you need to know.

In many of these chapter-long lectures, data scientists from companies such as Google, Microsoft, and eBay share new algorithms, methods, and models by presenting case studies and the code they use. If you're familiar with linear algebra, probability, and statistics, and have programming experience, this book is an ideal introduction to data science.

Topics include:

- Statistical inference, exploratory data analysis, and the data science process
- Algorithms
- Spam filters, Naive Bayes, and data wrangling
- Logistic regression
- Financial modeling
- Recommendation engines and causality
- Data visualization
- Social networks and data journalism
- Data engineering, MapReduce, Pregel, and Hadoop

Doing Data Science is collaboration between course instructor Rachel Schutt, Senior VP of Data Science at News Corp, and data science consultant Cathy O'Neil, a senior data scientist at Johnson Research Labs, who attended and blogged about the course.

Doing Data Science: Straight Talk from the Frontline

By Cathy O'Neil, Rachel Schutt

Doing Data Science: Straight Talk from the Frontline By Cathy O'Neil, Rachel Schutt

Now that people are aware that data can make the difference in an election or a business model, data science as an occupation is gaining ground. But how can you get started working in a wide-ranging, interdisciplinary field that's so clouded in hype? This insightful book, based on Columbia University's Introduction to Data Science class, tells you what you need to know.

In many of these chapter-long lectures, data scientists from companies such as Google, Microsoft, and eBay share new algorithms, methods, and models by presenting case studies and the code they use. If you're familiar with linear algebra, probability, and statistics, and have programming experience, this book is an ideal introduction to data science.
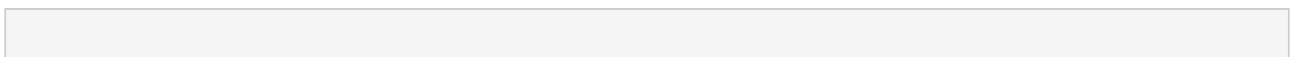
Topics include:

- Statistical inference, exploratory data analysis, and the data science process
- Algorithms
- Spam filters, Naive Bayes, and data wrangling
- Logistic regression
- Financial modeling
- Recommendation engines and causality
- Data visualization
- Social networks and data journalism
- Data engineering, MapReduce, Pregel, and Hadoop

Doing Data Science is collaboration between course instructor Rachel Schutt, Senior VP of Data Science at News Corp, and data science consultant Cathy O'Neil, a senior data scientist at Johnson Research Labs, who attended and blogged about the course.

Doing Data Science: Straight Talk from the Frontline By Cathy O'Neil, Rachel Schutt Bibliography

- Sales Rank: #24248 in Books
- Brand: Brand: O'Reilly Media
- Published on: 2013-11-03
- Original language: English
- Number of items: 1
- Dimensions: 9.00" h x .78" w x 6.00" l, 1.13 pounds
- Binding: Paperback
- 408 pages

Download and Read Free Online Doing Data Science: Straight Talk from the Frontline By Cathy O'Neil, Rachel Schutt

Editorial Review

Review
"Every once in a while a single book comes to crystallize a new discipline. If books still have this power in the era of electronic media, "Doing Data Science: Straight Talk from the Frontline" by Rachel Schutt and Cathy O'Neil: O'Reilly, 2013 might just be the book that defines data science."

 -- Joseph Rickert
Revolutions Blog

"I enjoyed Rachel and Cathy's book, it's readable, informative, and like no other book I've read on the topic of statistics or data science."
—Andrew Gelman
Professor of statistics and political science, and director of the Applied Statistics Center at Columbia University

"I got a lot out of Doing Data Science, finding the chapter organization on business problem specification, analytics formulation, data access/wrangling, and computer code to be very helpful in understanding DS solutions."
—Steve Miller
Co-founder, OpenBI, LLC, a Chicago-based business intelligence services firm

About the Author

Cathy O'Neil earned a Ph.D. in math from Harvard, was postdoc at the MIT math department, and a professor at Barnard College where she published a number of research papers in arithmetic algebraic geometry. She then chucked it and switched over to the private sector. She worked as a quant for the hedge fund D.E. Shaw in the middle of the credit crisis, and then for RiskMetrics, a risk software company that assesses risk for the holdings of hedge funds and banks. She is currently a data scientist on the New York start-up scene, writes a blog at mathbabe.org, and is involved with Occupy Wall Street.

Rachel Schutt is the Senior Vice President for Data Science at News Corp. She earned a PhD in Statistics from Columbia University, and was a statistician at Google Research for several years. She is an adjunct professor in Columbia's Department of Statistics and a founding member of the Education Committee for the Institute for Data Sciences and Engineering at Columbia. She holds several pending patents based on her work at Google, where she helped build user-facing products by prototyping algorithms and building models to understand user behavior. She has a master's degree in mathematics from NYU, and a master's degree in Engineering-Economic Systems and Operations Research from Stanford University. Her undergraduate degree is in Honors Mathematics from the University of Michigan.

Users Review

From reader reviews:

Aline Moran:

Does one one of the book lovers? If so, do you ever feeling doubt when you are in the book store? Try to pick one book that you find out the inside because don't evaluate book by its deal with may doesn't work is difficult job because you are frightened that the inside maybe not as fantastic as in the outside appear likes. Maybe you answer is usually Doing Data Science: Straight Talk from the Frontline why because the fantastic cover that make you consider concerning the content will not disappoint you. The inside or content will be fantastic as the outside or maybe cover. Your reading sixth sense will directly assist you to pick up this book.

Gerald Hackler:

The book untitled Doing Data Science: Straight Talk from the Frontline contain a lot of information on the item. The writer explains your ex idea with easy technique. The language is very straightforward all the people, so do not worry, you can easy to read the item. The book was compiled by famous author. The author will take you in the new era of literary works. It is easy to read this book because you can continue reading your smart phone, or device, so you can read the book within anywhere and anytime. If you want to buy the e-book, you can wide open their official web-site in addition to order it. Have a nice study.

Marian Storie:

Don't be worry should you be afraid that this book will probably filled the space in your house, you will get it in e-book technique, more simple and reachable. This specific Doing Data Science: Straight Talk from the Frontline can give you a lot of close friends because by you looking at this one book you have matter that they don't and make anyone more like an interesting person. This particular book can be one of one step for you to get success. This guide offer you information that perhaps your friend doesn't learn, by knowing more than other make you to be great men and women. So , why hesitate? We need to have Doing Data Science: Straight Talk from the Frontline.

Tamara Evans:

What is your hobby? Have you heard this question when you got pupils? We believe that that query was given by teacher for their students. Many kinds of hobby, Everybody has different hobby. Therefore you know that little person similar to reading or as reading through become their hobby. You need to understand that reading is very important and also book as to be the matter. Book is important thing to provide you knowledge, except your teacher or lecturer. You will find good news or update regarding something by book. Different categories of books that can you take to be your object. One of them is actually Doing Data Science: Straight Talk from the Frontline.

Download and Read Online Doing Data Science: Straight Talk from the Frontline By Cathy O'Neil, Rachel Schutt #D03SUWNKT46

Read Doing Data Science: Straight Talk from the Frontline By Cathy O'Neil, Rachel Schutt for online ebook

Doing Data Science: Straight Talk from the Frontline By Cathy O'Neil, Rachel Schutt Free PDF d0wnl0ad, audio books, books to read, good books to read, cheap books, good books, online books, books online, book reviews epub, read books online, books to read online, online library, greatbooks to read, PDF best books to read, top books to read Doing Data Science: Straight Talk from the Frontline By Cathy O'Neil, Rachel Schutt books to read online.

Online Doing Data Science: Straight Talk from the Frontline By Cathy O'Neil, Rachel Schutt ebook PDF download

Doing Data Science: Straight Talk from the Frontline By Cathy O'Neil, Rachel Schutt Doc

Doing Data Science: Straight Talk from the Frontline By Cathy O'Neil, Rachel Schutt Mobipocket

Doing Data Science: Straight Talk from the Frontline By Cathy O'Neil, Rachel Schutt EPub

D03SUWNKT46: Doing Data Science: Straight Talk from the Frontline By Cathy O'Neil, Rachel Schutt