

TEACHING GUIDE

EVERYBODY LIES

BIG DATA, NEW DATA,
AND WHAT THE INTERNET
CAN TELL US ABOUT WHO
WE REALLY ARE



SETH STEPHENS-DAVIDOWITZ

FOREWORD BY STEVEN PINKER



About this Book

Much of what we thought of people has been dead wrong—because people lie: to their friends, lovers, doctors, families, and themselves. By analyzing the gold mine that is data from the internet—the traces of information that billions of people leave on Google, social media, and dating sites—we can now learn what people really think, what they really want, and what they really do.

In his foreword to the book Stephen Pinker writes, “This book is about a whole new way of studying the mind . . . an unprecedented peek into people’s psyches. . . . Time and again my preconceptions about my country and my species were turned upside-down by Stephens-Davidowitz’s discoveries.”

About the Author

Seth Stephens-Davidowitz is a contributing op-ed writer for the *New York Times*, a visiting lecturer at the Wharton School at the University of Pennsylvania, and a former Google data scientist. He received a B.A. in philosophy from Stanford, and a Ph.D. in economics from Harvard.

Part 1: Data Big and Small

CHAPTER 1: YOUR FAULTY GUT

- How can data science be helpful to understanding the world around us? What kinds of things can we learn from it?
- Why is conventional wisdom often faulty? What are some examples where trusting your gut instinct goes wrong?
- What is beneficial about intuition with respect to data science? How can it help us reach conclusions?
- How does data allow us to see more precise conditions for an athlete making it into the N.B.A.?

Part 2: The Powers of Big Data

CHAPTER 2: WAS FREUD RIGHT?

- What are the four powers of big data?
- How does data science affect certain Freudian theories?
- Was, as the title of this chapter asks, Freud right? And if not, why not?
- What is regression analysis? How is it used?

CHAPTER 3: DATA REIMAGINED

- How did data science help the owners of the racehorse American Pharoah?
- What are the three ways Stephens-Davidowitz looks at words and language as data?
- Using the newspaper example and your own school newspaper, what can the word data tell you about the publication?
- In what ways can pictures be used as data? How can they help us learn about more social/environmental conditions than meets the eye?

CHAPTER 4: DIGITAL TRUTH SERUM

- Why don't people answer surveys truthfully? What are the best conditions to get them to answer truthfully?
- What do searches reveal about parents' attitudes toward their daughters? Do these types of searches surprise you?
- In what ways does the Google search bar act like a confessional for users?
- How do Google Trends help us see child abuse and abortion trends? Why is this data useful in thinking about social welfare cases?

CHAPTER 5: ZOOMING IN

- What does "zooming in" on data allow us to see? Why is it a useful way to examine data?
- What is a doppelganger search? Why is conducting one helpful?
- When Stephens-Davidowitz discusses the relationship between violent crime and violent movies, were you surprised by the results?

CHAPTER 6: ALL THE WORLD'S A LAB

- What is A/B testing? Why is it useful? What kinds of things can we learn from it?
- In this chapter, Stephens-Davidowitz writes, "In the era of big data, all the world's a lab." Is this cool or creepy? Does big data become intrusive?
- What is regression discontinuity? Why is it significant?
- In this chapter, Stephens-Davidowitz discusses data related to doctors and incentive programs and how that potentially influences medical care. What are the ethical implications of such practices?

Part 3: Big Data: Handle with Care

CHAPTER 7: BIG DATA, BIG SCHMATA? WHAT IT CANNOT DO

- What is the "curse of dimensionality"? How does it affect big data?
- Why can't big data beat the stock markets? Should it ever?
- Are there any other examples of numbers being seductive in the way they were for Zoë Chance and her pedometer study? Other number addictions?
- What role does small data and human expertise play in Jeff Seder's evaluations of horses?

CHAPTER 8: MO DATA, MO PROBLEMS: WHAT WE SHOULDN'T DO

- This chapter raises the following ethical question: "Do corporations have the right to judge our fitness for their services based on abstract but statistically predictive criteria not directly related to those services?" How would you answer this question?
- This chapter talks about creating a "fair fight" between corporations and consumers over what information data provides. How do we create a fair fight?
- Why shouldn't data be used to try and predict crime at an individual level?

Writing Prompts/Activities

- Spend some time exploring Google Trends: <https://trends.google.com/trends/>. What information can a search for “immigration” in Google Trends tell us? What are some interpretations you can make from the data given?
- Using the language analysis and methodology that Stephens-Davidowitz talks about in chapter 3, what can learn about the local paper serving your area? Thinking back to the question for chapter 3 about your college or university’s school newspaper, how does it compare?
- The Jeanne Clery Act, which aims to provide transparency around campus crime policy and statistics, mandates that colleges must post campus crime data on their campus websites. Access the data for your campus. Does the data reveal any trends? For example, are there upticks in certain types of crime during specific days of the week? Create a report that includes graphs to support your theory.
- How serious are the ethical issues surrounding information available from data science? Do you think, in certain circumstances and scenarios, that big data over-reaches and is harmful?
- Read the privacy policy and terms of use sections of a website that you visit on a regular basis. How much of your personal information do you give to have access to the site? Be specific. Do you think this a fair trade off?

Resources

- [“A Hands-on Guide to Google Data”](#) by Seth Stephens-Davidowitz and Hal Varian. Goggle, Inc.
- [PowerPoint](#) for “A Hands-on Guide to Good Data” by Seth Stephens-Davidowitz and Hal Varian. Goggle, Inc.
- [“What can Google search reveal about how we spend, vote, and connect?”](#) by Seth Stephens-Davidowitz
- [“How Uber Uses Psychological Tricks to Push Its Drivers’ Buttons”](#) by Noam Scheiber and graphics by Jon Huang, *New York Times*

Everybody Lies: Big Data, New Data, and What the Internet Can Tell Us About Who We Really Are by Seth Stephens-Davidowitz.
Dey Street: 354 pages; illustrated; notes; index.
2017: hardcover (978-0-06-239085-1)
Also available in ebook and audio editions.

For more teaching guides, please visit HarperAcademic.com.

