

Syllabus

Data Literacy in the Age of Big Data

EMPA Workshop
November 18, 2018

Outline

Data play an ever-increasing role in evidence-based policymaking. But often, they are poorly collected, analyzed, and interpreted. In this course, participant will be enabled to evaluate data-based research and recognize pitfalls and imposture that are often hidden in quantitative reports, in particular when it comes to the use of new or “big” data.

Course goals

Having attended this workshop, you will have...

1. acquired basic knowledge on what’s possible (and not possible) with statistics in the age of Big Data,
2. improved a sense to spot design flaws in quantitative research, and
3. some guidance on hand to get better at consuming/producing quantitative information.

Session Overview

Session	Topic
1	Spotting statistical bullshit
2	Consuming quantitative research
3a	Understanding the promise and perils of Big Data
3b	Getting a feel for hands-on statistical coding

Further Reading

Bergstrom, C, and J West, 2016: Calling Bullshit in the Age of Big Data. callingbullshit.org .

Conway, Drew, and John White, 2012: Machine learning for hackers. O’Reilly.

Healy, Kieran, 2018: Data Visualization: A Practical Introduction. Princeton University Press.

Huck, Schuyler W, and Howard M Sandler, 1979: Rival hypotheses: Alternative interpretations of data based conclusions. HarperCollins.

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani, 2013: An Introduction to Statistical Learning. Springer.

Kahneman, Daniel, and Patrick Egan, 2011: Thinking, fast and slow. Farrar, Straus and Giroux.

Kampen, Jarl K, and Peter Tamás, 2014: Should I take this seriously? A simple checklist for calling bullshit on policy supporting research. Quality & Quantity 48:1213–1223.

Pearl, Judea, and Dana Mackenzie, 2018: The Book of Why: The New Science of Cause and Effect. Basic Books.

Salganik, Matthew J, 2017: Bit by bit: social research in the digital age. Princeton University Press.

Silver, Nate, 2012: The signal and the noise: why so many predictions fail—but some don't. Penguin.

Stephens-Davidowitz, Seth, 2017: Everybody lies: big data, new data, and what the internet can tell Us about who we really are. HarperCollins.

Watts, Duncan J, 2011: Everything is obvious: Once you know the answer. Crown Business.*

Wickham, Hadley, and Garrett Grolemund, 2016: R for data science: import, tidy, transform, visualize, and model data. O'Reilly.