Parametric Statistics

Week 1 - Introduction

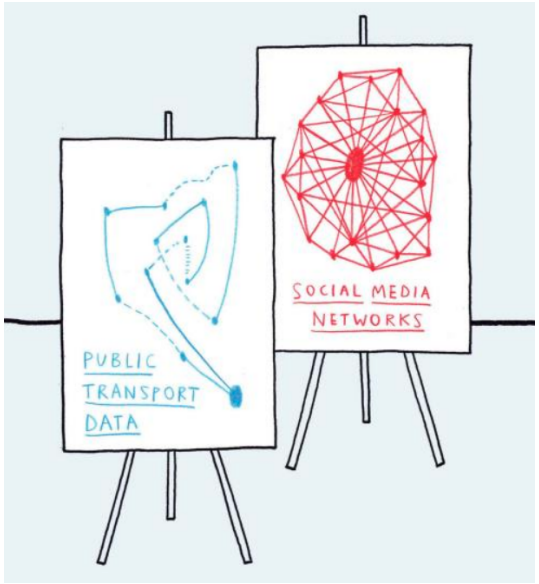
Juan Carlos Medina Serrano

Technische Universität München
Hochschule für Politik
Political Data Science

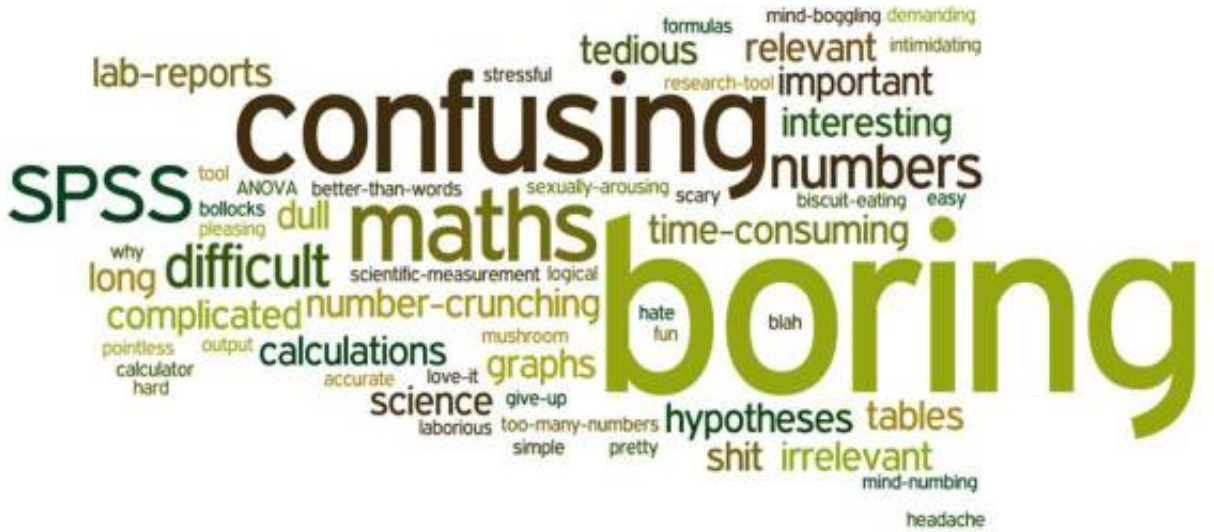
Munich, 16. October 2019

political data science

https://politicaldatascience.blogspot.de



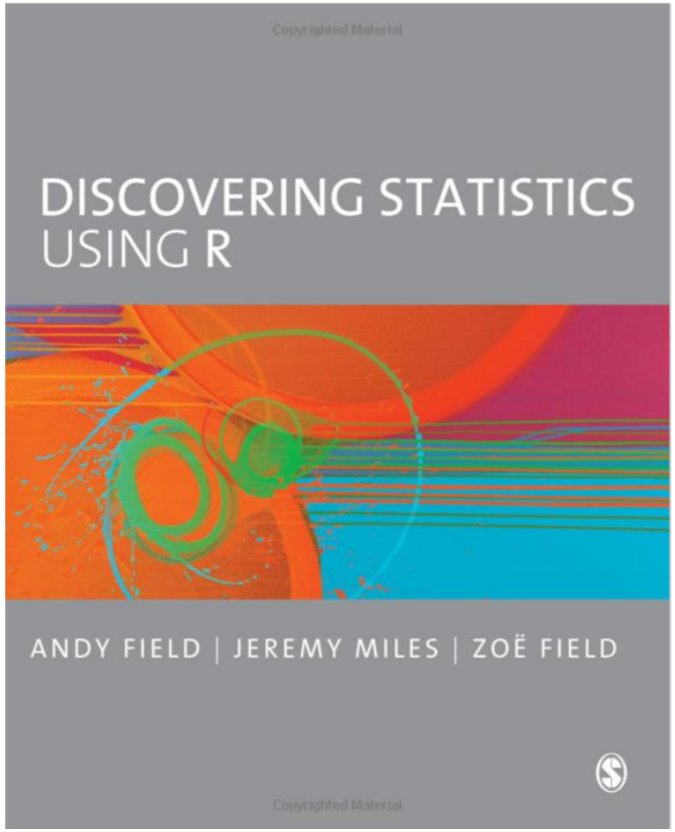
Introduction

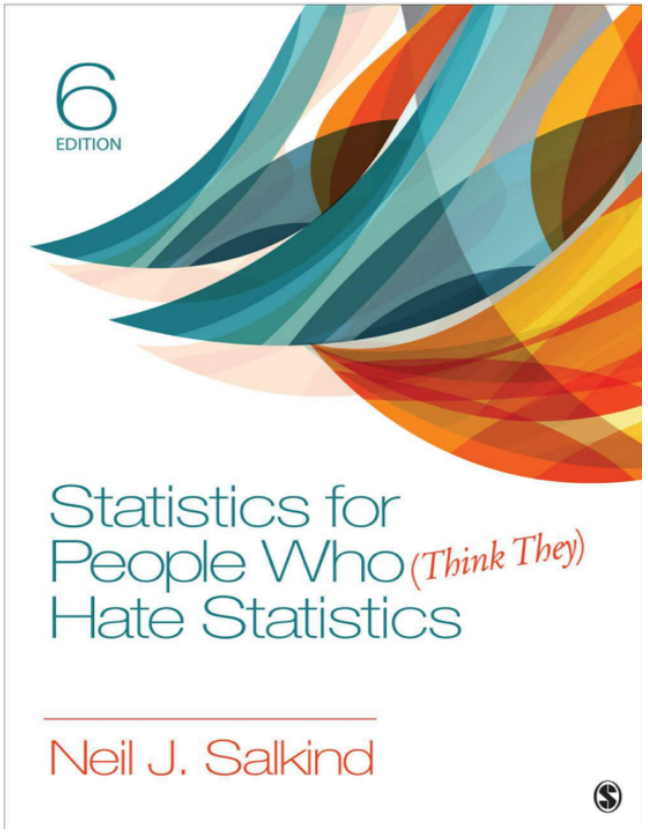


Content

1. Descriptive Statistics and Probability
2. Research Design
3. Statistical Modeling and Sample Distributions
4. Hypothesis Testing
5. Tests of Significance
   - Z-tests, t-tests, ANOVA tests
6. Correlation
   - Pearson, Spearman and Kendall coefficients
7. Statistical Assumptions
8. Non-normality
   - KS, Wilcoxon, Kruskal Wallis tests, Chi-squared
9. Fitting the Data
10. Statistics done wrong
11. The linear model
    - Linear Regression, ANOVA revisited, Multiple Linear Regression
12. Basics of Statistical Learning
    - Predicting the future, Logistic Regression, Clustering

Introduction





Further Literature

- Statistics in a Nutshell - Sarah Boslaugh & Paul Andrew Watters
- Practical Statistics for Data Scientists (Chapters 1-3) - Peter Bruce & Andrew Bruce
- Parametrische Statistik - Carsten Dormann
- Statistics done wrong - Alex Reinhart
- R Cookbook - Paul Teetor

Organization

- Course : Wednesday 09:45 - 11:15
- R Tutorial : Thursday 15:00 - 16:30
- R Tutorial : Bring your own laptop!
- Exam: 20th February 11.00 am - 12.30 pm.

Cool Statistics = Applied Statistics!

REAL EXAMPLES from political data science (Social bots, fake news, online political campaigns, microtargeting)
INTERESTING QUESTIONS:

- How can I analyze demographic data to make a better political campaign?
- How to create a poll that is significant of the whole German electorate?
- Which political party had the best online advertising campaign?
- How to predict if Trump will be elected in 2020?

Homework 1

Install Anaconda in your laptop: `https://www.anaconda.com/distribution/`
Python 3.7 (We will not use Python, but is nice to have for other projects)