Data Science for Public Policy

Aaron R. Williams - Georgetown University

PPOL 670 | Reading List

Week 01

- R for Data Science – Chapters 1, 2, & 4
- What all policy analysts need to know about data science
- What is Data Science @ the Urban Institute?

Week 02

- R for Data Science – Chapters 5, 6, 8, & 12
- The Statistical Crisis in Science
  - This article, by Andrew Gelman and Eric Loken, discusses the importance of data manipulation choices in statistical inference.
- OPTIONAL: Tidy Data
  - Please ignore the R code as it is out-of-date.

Week 03

- R for Data Science – Chapter 3 – Data Visualization
- Fundamentals of Data Visualization – Chapter 2 – Mapping data onto aesthetics
- R for Data Science – Chapters 26, 27, 29, & 30
- RStudio R Markdown introduction
- Mastering Markdown

Week 04

- Git Handbook
- Mastering Markdown
- Understanding the GitHub Flow
- Documenting Your Projects on GitHub
- After class: Git Tutorial

Week 05

- Read more about custom functions in R4DS
- Read more about `library(purrr)` in R4DS.
- Marble Science: Monte Carlo Simulation
- Forecasting a Presidential Election with Monte Carlo Simulation

Week 06

- Why we built the Education Data Portal
- Why we built an API for Urban's Education Data Portal
- Analyzing US Census Data by Kyle Walker – Chapter 5
- Simple Features for R

Week 07

- Bit By Bit section 3.6.2
- Hands on Machine Learning with R sections 1.1-1.2 and chapter 2
- R2D3 introduction to the bias-variance tradeoff

Week 08

- Get Started with Tidymodels
- R2D3 introduction to decision trees

Week 09

- Hands on Machine Learning with R Chapter 17
- Reenvisioning Rural America

Week 10

- Hands on Machine Learning with R Chapter 17
- Reenvisioning Rural America
- Hands on Machine Learning with R Chapter 20
- Catalyzing Policing Reform with Data