



CHAPIN HALL
AT THE UNIVERSITY OF CHICAGO

A Hands-On, Pain-Free Introduction to R for Practical Data Work

Nick Mader, PhD

IL CQI Conference – Nov 28th, 2018

nmader@chapinhall.org

Link to this presentation and all materials: <https://goo.gl/xcLCvA>

Goals for this Session

Describe R's features and range of capabilities

Describe when--and when not--to use R

Previous ideas for how to learn more

Try out R!

A Little About R



is--

- **free** -- \$0
- **open source** -- many diverse folks continually adding features
- **broadly supported** -- wide range of classes, forums, and support tools
- **a programming language** -- high learning curve, high capabilities
- **increasingly popular** -- decreases the learning curve, increases value

What R can do

- **Read and writes** virtually all file formats, connects to databases and APIs
- Full range of **data cleaning/transformation**s (e.g. the tidyverse)
- State-of-the-art **statistical analysis** + Big Data methods
- Generate interactive visualizations, maps, web applications, and reports
- Once set up, can perform/update all of the above **automatically**, for any size of data, and any number of operations and users
- **Integrate** with other software, e.g. Power BI

R is like a universal remote formula bar

When to Use R

- When a combination of staff expertise and size of challenges means R can save time and money
- When your organization has a wide range of needs/potential uses, especially if there is a lot of manual work and repetition
- If you have advanced analysis needs, R is cheaper and has broader functionality than SPSS, SAS, and Stata

When to Use Another Tool

- When needs can be met with more straightforward, more interactive tools
- When creating and maintaining programming code is impractical given staff expertise and availability

Resources for Learning More

- **R for Data Science** (<http://r4ds.had.co.nz/>) -- free web book covering a range of common data needs
- **DataCamp** (<http://www.datacamp.com>) is a highly recommended online platform:
 - Short courses: 4-6 hours
 - Instructors are legit experts
 - Every courses mixes short videos and progressions of helpful videos
- **R-Bloggers** (<https://www.r-bloggers.com/>) as an aggregator of R blogs, which show the code as worked examples
- **Google** -- especially with “vignette” as keyword for learning specific packages
- Read my detailed contrast of R with Power BI [here](#)

Let's Try R!

- We're going to do this in the style of Mob Programming
 - One screen and keyboard
 - The whole group sets direction and pace
 - Every 5 minutes, we pass the keyboard
 - Questions and discussion are required throughout. Nick will help steer.
- Sample data sets to use
 - Wage rates for public employees in IL, by trade and county (link)
 - American Community Survey data from the Census (list of tables, “vignette” for pulling Census data with R)