





## Course Description and Learning Objectives

The intention of this intensive short course is to strengthen the methodological skills of the research community. At the end of the week, participants will have:

- “Mastered” the tidyverse, a set of principled tools for data science. The tidyverse is a friendly, readable, and fast set of packages intended to work well together, improve code readability, and make analyses more reproducible.
- Written dynamic documents using best practices for reproducible research, including using Quarto. Quarto intertwines code and text so that reports and articles are fully reproducible and exportable to PDF, HTML, Word, and more. Quarto also has excellent support for citation management and formatting for journals.
- Modeled simple statistical problems using linear and logistic regression.

### Audience

The course is for researchers, public health professionals, epidemiologists, and clinicians who want to improve their R coding skills, who want to learn modern tools in the R ecosystem, like the tidyverse and Shiny, or who want to get started writing software in R.

### Course Pre-requisites

Basic epidemiology and statistics. Some experience with R programming will be helpful, but those new to R can take a free introductory course on Data-Camp: <https://www.datacamp.com/courses/free-intro-duction-to-r>

### Course materials

R for Data Science (<https://r4ds.had.co.nz/>) and Advanced R, ed. 2 (<https://adv-r.hadley.nz/>), both free online

## Program

14 - 18 August 2023 | 9am - 5pm

### Monday, 14 August

am Reading and wrangling data  
pm Data visualization with ggplot2

### Tuesday, 15 August

am Assignment 1: Plot recreation  
pm Tidying data

### Wednesday, 16 August

am Assignment 2: Modeling in R  
pm Types, vectors, functions; Functional programming with purrr

### Thursday, 17 August

am Assignment 3: Data science whole game  
pm Reproducible research with Quarto

### Friday, 18 August (ends 13:00)

am Assignment 4: Tidy Tuesday