Learning Outcomes

1. Attendees will have a firm grasp of managing and manipulating data within R.
2. Learn to code in a cleaner, less error-prone, fashion to facilitate sharing R-scripts and working collaboratively.
3. Keep in step with open science practices in terms of sharing code that other users can more easily understand.

Target Audience:

* Researchers and graduate students

Abstract (250 word limit):

* Data cleaning is a time consuming and often error-prone process which every researcher will experience. It is said that 80% of data analysis time is spent on the process of cleaning and preparing the data (Dasu & Johnson, 2003), but this time spent cleaning and preparing data can be dramatically reduced by using the right tools for the job. The workshop will begin with a foundational discussion on data structures within R and general coding practices. The focus will then shift to the more practical topics of using the *tidyverse* collection of packages to import/export data, inspect and manipulate specific types of data (including categorical, dates, and character type), reformat data, and produce quality graphics to showcase your data. A publicly available dataset will be used as a running example throughout the workshop to provide hands-on experience using each of the techniques discussed. Periodic interactive exercises will be presented to allow attendees an opportunity to work collaboratively to solve common data cleaning dilemmas. To help tie the whole workshop into a cohesive process, the concept of a data “pipeline” will be introduced as way to conceptualize the data cleaning process. By teaching better coding practices throughout the workshop, our resulting data cleaning pipelines will be coded in a more human-readable format making it both easier to debug and share our code when working collaboratively on projects. It is recommended that attendees bring personal laptops and pre-install R, RStudio, and the *tidyverse* package.

Key words/phrases :

1. Data-cleaning
2. R
3. tidyverse
4. Better coding practices