R Course Overview

Course Goals

Course designed to demonstrate basic use of R for reading, manipulating, and plotting data via R Studio. By the end of this course students will be able to:

- read and write basic R programs
- import well formatted data into R
- do basic data manipulation in R
- produce common numerical and graphical summaries in R
- describe a use case of an analysis done in R

Module Overviews

Module 0: Introduction to the Course

- Information about the course structure

- Installing R and R Studio

Module 1: R Programming

- Introduction to R
  - R Studio
  - Console
  - Scripts
  - Objects & Functions
- Common Data Structures
  - Vectors, Matrices, Data Frames, & Lists
  - `help()` Function
  - Accessing & Basic Subsetting of Common Data Structures

Module 2 - Importing Data

- Basics & Delimited Data
  - Common data formats
  - Asides: R projects, `factors`, and R packages
  - Read 'clean' delimited data with `readr`
- Excel Data, Databases, & More
  - Read Excel data with `readxl`
  - Read SAS & SPSS data with `haven`
  - Resources for JSON, XML, databases, and APIs

Module 3 - Manipulating Data

- Documenting with Markdown
  - Data manipulation idea
  - Documenting with Markdown

- Logicals & `dplyr`
  - Logical statements
  - `dplyr` package
- Creating New Variables & Reshaping Data
  - Conditional execution (if then)
  - For loops
  - Vectorized functions
  - Reshaping data with `tidyr`

Module 4 - Summarizing Data

- Descriptive Statistics
  - Variable types
  - Contingency tables
  - Measures of center and spread
  - Summaries by subgroups of data
- Graphical Displays via Base R
  - `barplot`, `hist`, `boxplot`, `plot`
  - `lines`, `abline`, `points`, `text`
  - `lty`, `lwd`, `pch`, `cex`, `color`
- Graphical Displays via `ggplot2`
  - `ggplot`, `aes`, common `geom_*`
  - Labels, scales, title
  - `stat` vs `geom` layer
  - `position` arguments
  - Faceting

Module 5 - Example Program

- Recreation of River Water Analysis