



Introduction to R for Data Analysis

COURSE OBJECTIVE: By the end of the course, you will be familiar with navigating the *RStudio* environment and reading in different data types. You will also learn the fundamentals of cleaning, analyzing, and visualizing real data. Finally, you will be introduced to, and have the chance to practice using, built-in R functions to explore data and apply basic statistical methods to different datasets.

INSTRUCTOR: Ben Brew | Data Scientist and Bioinformatician at SickKids' Genetics and Genome Biology Lab | ben@databrew.cc

Course Outline

- **Introduction (15 minutes)**
- **Session 1: Getting familiar with RStudio (60 minutes)**
 - RStudio environment
 - R for basic calculations
 - R as an object oriented programming language
 - Saving objects
- **Session 2: Reading in and exploring data (90 minutes)**
 - Flat files (csv, xlsx, txt, etc.)
 - Using R libraries to get data
 - Reading in data from your laptop
 - Data structures and variable types
 - R syntax
 - Basic R functions for summarizing data
 - Basic visualization (plots, ggplot)
- **Session 3: Manipulating data (90 minutes)**
 - Changing variable names
 - Conditional statements
 - Subsetting data
 - Regular expressions, cleaning data
 - Manipulating data with dplyr package
- **Wrapping up (15 minutes)**
 - Useful resources to continue learning
- **Additional Support (*Optional*)**
 - Your chance to ask questions about and get support with your own data