Introduction to R/RStudio

Experiential Data science for Undergraduate Cross-disciplinary Education

Dr. Kim Dill-McFarland, U. of British Columbia

Contents

Introduction to R/RStudio

Learning objectives

- Describe general uses for R
- List several advantages to using R and RStudio

What are R and RStudio?

At its heart, R is statistical software. Base R is used to manipulate, analyze, and visualize tabular data, *i.e.* data in tables. However, since its original inception, thousands of add-on R packages have been created to build on base R. Packages allow the import of diverse data types as well add functionalities to all aspects of data manipulation, analysis, and visualization. Thus, there is likely a package to suit each of your data needs!

RStudio is an integrated development environment (IDE) for R that provides a more user-friendly environment and integrates R with additional features. These include a customizable visual interface, integrated file, package, and plot management, and links to other data science resources like Git/GitHub, Shiny apps, etc.

Advantages of R/RStudio

1. Free and open-source. There is no fear of losing access when you graduate, change jobs, get a new computer, etc.

2. Reproducible methods. Unlike graphical user interface (GUI) software like Microsoft Excel, R relies on scripts (*i.e* text-based commands) that can be saved and shared. RStudio also offers additional tools like Rmarkdown and RPubs to create reproducible documents like reports, slides, websites, etc.

3. Community. While R was originally for statistics, it has be adopted in many diverse fields. Custom R packages and individual developers around the world provide accessible resources for all users, from novice to expert. This has built an international community of support.