Introduction to R for Bench Biologists

Nolan Ung
nolan.ung@cammlab.org

8/23/2019

Workshop description

This 6 hr six week course is a brief introduction to the R statistical computing language designed specifically with bench biologists with no coding experience in mind so that they can get more done in less time and with less pain. This course is intended to be very hands-on and will include a final project of each individuals choosing, demonstrating the skills they have learned. There will be a brief introduction to programming principals, R and R studio, version control to make sure the students have the tools to use R effectively. We will cover importing data and common manipulations of data and data processing. The end of the course will cover data visualization both theory and practice.

---

learning objectives

- Understand the rational for using R and a scripting language over a GUI based workflow.
- Become familiar with the various capabilities of the Rstudio IDE.
- Understand how version control works and be able to use Rstudios git interface.
- Basics of programming: assignment, data types, functions, arguments, help
- How to install and load packages.
- Grasp how to import and combine data sets.
- Basic data manipulation using `tidyr` and `dplyr` including pipes, mutate, group_by and summarize.
- How to export data
- Explain the gestalt principals of data visualization and the goals of data viz.
- Be able to build simple plots using `ggplot2` and a simple heat map using `pheatmap`.
- Build plots interactively and customize aesthetics and labels, facet plots and save figures.
- Have the ability to take a simple experimental analysis from beginning to end.
- Advance to the intermediate R course.

Office Hours

I will holding Office hours in an effort to make sure that I am available for all questions and practice. I am still always available on Slack or and outside of office hours.

Office Hours: Fridays from 2pm-3pm location TBD

Prepartion for workshop

Please have the following installed on your laptop computer:

- R version 3.6.1 (Action of the Toes)
- Rstudio version 1.2.1335
- git
- Git for Windows
- jupyter notebook for Mac
- jupyter notebook for Windows
- Anaconda

Also please create a GitHub account so that you can access the materials that we will use for the workshop. If you need help on how to create an account have a look here.

downloaded these on your mac or windows lap top and bring it to the workshop. If you do not have a laptop please contact Nolan and/or Tom ahead of time and they can get you a loaner lap top. If you need assistance with downloading these two tools please feel free to contact Nolan for assistance.

For detailed instructions on how to install and download everything, check out this video that Ken and I made!!

Final Project

I find that R is learned best when you use it on your own data and when you need to use it. To help you learn and really interalize this content I'm "assigning" a final project. Its very open ended and I want you to use R in one of your analyses for your project. you can do a new analysis or reproduce one that you did in excel. Other than that there is no requirement. I don't want to you to spend a lot of time on this, rather the goal is to see how you can use R to meet your everyday analysis needs. I would love to have everyone give a 10min or less walkthrough of their code.

Workshop Calendar

Each session will last between 1 and 1.5 hrs.

Date	Week #	Topic
September 2nd	Week 1	Before we start: Getting to know R, Rstudio and Git
September 9th	Week 2	Introduction to programming and R basics
September 16th	Week 3	Starting with data: importing and formatting
September 23rd	Week 4	Manipulating data: wrangling and transforming data
September 30th	Week 5	Data visualization: basic plots and heat maps
October 7th	Week 6	Beginning to end real world examples