

**R and Programming Skills Development: A SoLS Grad Student  
Workshop  
May 4 – 5, 2018  
Building/Room: LSC active learning labs (tentative)  
Tempe, AZ**

**Workshop Details**

The R and Programming Skills Development Workshop 2018 is a free spring workshop featuring an introduction to R, \*nix commands, and scientific programming. This two - day course will provide tools for reading in, manipulating and analyzing data in R, as well as for using other open-source tools for visualizing and manipulating data. It will provide a basic introduction to scientific programming, creating publication quality graphics in R, as well as wrangling messy data and files with \*nix commands in bash. Although the course is aimed at graduate students, you need not be a graduate student to attend. The modular design of the course should allow participants to attend sections of the course that are useful to them and skip the parts that are not. We encourage you to attend the first day if you are not already an avid R user.

**Instructor:** Rebecca Clark

**Requirements**

No previous knowledge of R is required. The course is open to all students, postdocs, and faculty. **A laptop is required.** If you do not have a laptop, please contact Rebecca (rmclark1@asu.edu) and we will work to set you up with a loaner. The software needed for this course is all open-source. Installation instructions will be provided in a separate document and will vary slightly for different operating systems (Mac, Windows, Linux).

**Schedule**

The workshop will start at 1:00 pm May 4<sup>th</sup> and wrap up at 3 pm on May 5<sup>th</sup>. We are hoping to provide snacks and coffee during afternoon and morning breaks.

**Day 1 – May 4<sup>th</sup>**

**1:00 – 2:30 pm**

**Session 1 – Introduction to the R environment (Rebecca Clark)**

1. Overview - (Presentation)
2. Basic interface (Handout)
  - a. Simple calculations
  - b. Variables and assignments
  - c. Functions
3. Loading data (Handout)
  - a. Directories, paths, and finding files
  - b. .csv and .txt

- c. Headers
- d. Formats (vectors and data frames)
- e. Group activity (“fixing” an input file for R)
- 4. Manipulating data (Handout) (R File)
  - a. Subsetting
  - b. Indexing
  - c. Handling missing values

**2:30-2:45 coffee break**

**2:45 – 4:15 pm**

**Session 2 – Introduction to the R environment continued (Rebecca Clark)**  
(presentation and handout)

- 1. Project organization
  - a. File hierarchy
  - b. Reference cards
  - c. R style guide
  - d. Well-documented scripts
- 2. Packages and CRAN
- 3. Where to go for help
- 4. R studio

**Day 2 – May 5<sup>th</sup>**

**9:00 – 10:30 am**

**Session 3 – Publication-Quality Graphics (Rebecca Clark)**

- 1. base graphics details
- 2. Plotting with lattice
- 3. Plotting with ggplot2

**10:30 – 10:45 coffee break**

**10:45-12:00 pm**

**Session 4 - \*nix commands in bash and R (Rebecca Clark and Ofir Levy)**

- 1. \*nix commands
- 2. Regular expressions
- 3. Invoking R from command line

**12:00 – 1:00 pm – Lunch**

**1:00-2:30 pm**

**Session 5 –Scientific Programming (Rome)**  
**-Tentative-**

- 1. Introduction to programming
  - a. Control structures

- b. Vector notation
- c. How to write instructions as a script
- 2. Random walk example
- 3. Randomization with Monte Carlo and bootstrapping

**2:30 – 3:00 pm**

**Wrap-up: Discussion of special topics and organization of working groups (Rebecca Clark)**

**Initial Websites**

<http://cran.r-project.org>

Quick-R: Online quick-reference index for many common statistical procedures in R:

<http://www.statmethods.net>

<http://stackoverflow.com/questions/tagged/r>