

# Ecological Analytics with R

*Location:* Harvard Forest

*Instructor:* Matthew Lau (<http://people.fas.harvard.edu/~matthewklau>)

*Goal:* Students will learn how to use the R programming language for ecological analyses and gain experience:

- Managing analytical aspects of a project
- Inputting, manipulating and exporting data
- Statistical functions
- Plotting data patterns
- Coding and software best practices
- Getting more help, experience and practice

**Although statistics will be introduced very briefly, this will not be a statistics class.**

*Pre-requisites:* experience with using basic computer software

*Required Materials:* laptop or access to some computing device

*No laptop?* You can borrow a Harvard Forest laptop (contact Manisha Patel)

## Class Schedule

### June 1 (Monday)

**Location: Fisher Museum**

7:45-8PM: brief intro to ecological analyses.

- Why program?
- Brief introduction to project management
- Connect to syllabus and course materials

### June 2 (Tuesday)

**Location: Fisher Museum**

*Before Class*

- Connect to Harvard Forest wireless
- Install R on your computer: <http://lib.stat.cmu.edu/R/CRAN/>\*
- Download the example project: <https://github.com/HarvardForest/myProject/archive/master.zip>

*9-10AM: Analytics and programming basics*

- Analytics project framework
- Operations
- Objects
- Functions
- Scripting and annotation (Save Our Source!)
- Setting the working directory

### *Post-Class Challenge*

- What's the area of a square with a side = 10 cm?
- What's the circumference of a circle with diameter 10 cm?

### **June 4 (Thursday)**

#### **Location: Seminar Room**

*9-10AM: Data, Data and more Data!* - Entering data by hand - Manipulating vectors (sorting, ordering) - Manipulating matrices (sorting, appending) - Inputting data (read.csv, read.table) - Advanced data = lists

### *Post-Class Challenge*

- How does a circle's circumference change as its diameter changes?
- Find a cool equation and play with it in R (e.g.  $E=mc^2$ )!

### **June 9 (Tuesday)**

#### **Location: Seminar Room**

*9-10AM: Analysis*

- Overview of data visualization
- Calculating basic statistics (mean and variance)
- Writing your own functions (se: input, process, output)
- What are packages? (e.g., *ggplot*)
- Barplot with error bars

### *Post-Class Challenge*

- Create a plot of how a circle's circumference changes as its diameter changes.

### **June 16 (Tuesday)**

#### **Location: Seminar Room**

*9-10AM: Best Practices and Some Advanced Topics*

- Organizing code
- Getting data from the HF archives
- Loops and applies
- Versioning with github
- Simulating data (runif and rnorm)
- Data and software ethics
- Data provenance
- The Hackathon

### *Post-Class Challenge*

- Find and analyze an HF dataset

## Readings and Resources

- R Cheat Sheet – <http://cran.r-project.org/doc/contrib/Short-refcard.pdf>
- Version Control – <https://help.github.com/articles/good-resources-for-learning-git-and-github>
- Code for America – <https://www.codeforamerica.org>
- Learning Statistics
- *Primer of Ecological Statistics* by Ellison and Gotelli
- *The Ecological Detective* by Hillborn and Mangel