Ecological Analytics with R

Location: Harvard Forest, Fisher Museum

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)

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

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

June 2 (Tuesday)

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

- Present the overarching analytics framework
- Basic functions/operations and tasks (e.g., do simple calculations on scalars and vectors)
- Objects = way to keep data in your memory
- Functions = name(arguments)
- Scripting and annotation = save your code and notes about code

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)*

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

Post-Class Challenge

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

June 9 (Tuesday)

9-10AM: Analysis

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

Post-Class Challenge

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

June 16 (Tuesday)

9-10AM: Best Practices and Some Advanced Topics

- Organizing code
- Getting data from the HF archives
- Versioning = github
- Loops and applys
- Simulating data = runif and rnorm
- Data and software ethics
- Data provenance
- 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