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 with:

- Managing analytical aspects of a project
- Inputing, manipulating and exporting data
- Statistical functions
- Exploring 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

Pre-Class Meeting

Location: Fisher Museum

Time: TBD

Brief intro to ecological analyses.

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

First Meeting

Location: Fisher Museum

Time: TBD

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
- Analytics project framework
- Operations
- Objects

- Functions
- Scripting and annotation (Save Our Source!)
- Setting the working directory

Post-Class Challenge

- Explore the HF data archive and find a dataset relevant to your project.
- Make sure to read the data ownership information!

Second Meeting

Location: Fisher Museum

Time: TBD

- Entering data by hand
- Manipulating vectors (sorting, ordering)
- Manipulating matrices (sorting, appending)
- Inputting data (read.csv,read.table)
- Advanced data = lists

Post-Class Challenge

- Write a function that will import your HF dataset into R.

Third Meeting

Location: Fisher Museum

Time: TBD

- Overview of data visualization
- 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

- Write a script that will import your HF dataset and conduct analyses on the data.

Fourth/Last Meeting

Location: Fisher Museum

Time: TBD

- Organizing code
- Getting data from the HF archives
- Loops and applys
- Versioning with github
- Simulating data (runif and rnorm)

- Data provenance

Post-Class Challenge

- Think of an ecological question/challenge that might be addressed with software
- Summer Hackathon?
- Check out Code for America

Readings and Resources

- R Cheat Sheet – http://cran.r-project.org/doc/contrib/Short-refcard.pdf
- Plots with ggplots – https://www.rstudio.com/wp-content/uploads/2015/03/ggplot2-cheatsheet.pdf
- Code School – http://tryr.codeschool.com
- 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