EEB R Bootcamp
Intro to class and basic stats and plotting

Noa Pinter-Wollman

Welcome!

- **Goals**: Introduce graduate students to the R environment and provide a foundation for skills in scientific programming.
- Course webpage with materials: http://michaelalfaro.github.io/eeb201/

Who R we?

Instructors

Noa Pinter-Wollman

Kirk Lohmueller

Mike Alfaro

Jamie Lloyd-Smith

Graduate students

Marvin Browne

Lauren Smith

Mark Juhn

Shawn Schwartz

Riley Mummah

Ana Gomez

Liz Karan

Christiane Jacquemetton

Maddi Cowen

Kenji Hayashi

Learning objectives

By the end of the course students will:

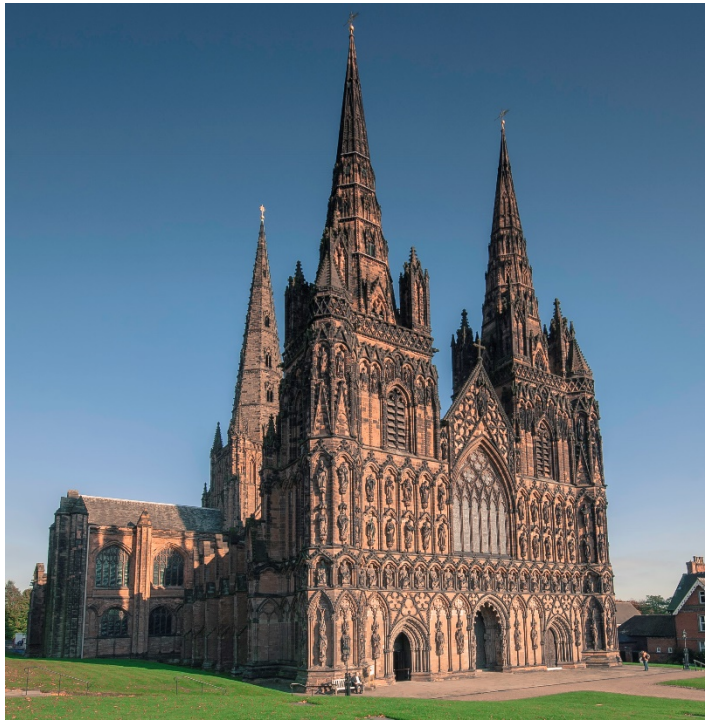
- Be comfortable executing basic commands in the R environment.
- Be able to load packages in R and make use of their added functionality.
- Be able to read in data files, manipulate data, and perform simple analyses in R.
- Be capable of plotting curves, scatter-plots, histograms, and other graphic outputs in R.
- Be able to write their own computer programs to simulate population models in discrete or continuous time.

What is R?

- “R is a language and environment for statistical computing and graphics”
- A free, open-source, community-built software package.
- Powerful, flexible computational tool with a broad array of add-on ‘packages’ that perform advanced analyses.
- A high-level programming language for scientific computing.
- A powerful platform for statistics.
- A major research tool in ecology and other fields.

What is R not?

- A commercial package
- A drop-down menu environment
- Excel
- JMP, SPSS, SAS, etc.
- Matlab



Why R?

- Free
- Widely used in ecology and evolution
- Helpful, active, and large user community
- Lots of resources (formal packages and informal community)
- Versatile
- Flexible
- Data manipulation capabilities
- Reproducible
- Scalable
- Transferable programming skills

R Bootcamp schedule

- Wednesday September 18th
  - 9-10:30 intro to the bootcamp and Basic stats and plotting – Pinter-Wollman
  - 10:30-12 Advanced plotting - Lohmueller
  - 12-1:30 Lunch
  - 1:30-5PM Work on assignments

- Thursday September 19th
  - 9-10:30 Flow control, scripts, functions –Alfaro
  - 10:30-12 Working with data and population dyanmics - Lloyd Smith
  - 12:30-1:30 Lunch
  - 1:30-5PM Work on assignments

What will you be learning in the next two days?

- Command line, scripts, and functions
- Getting help in R
- Packages/libraries
- Reading data
- Variables
- Subsetting and finding specific data
- Control flow (for, if, while…)
- Functions
- Plotting
- Basic modelling

R resources

- https://cran.r-project.org/
- R studio (https://www.rstudio.com/)
- R for biologists (https://cran.r-project.org/doc/contrib/Martinez-RforBiologistv1.1.pdf)
- Springer series 'Use R!' is available free through UCLA library
- R for Data Science http://r4ds.had.co.nz/
- Quick R: http://www.statmethods.net/
- R reference card: ftp://cran.r-project.org/pub/R/doc/contrib/Short-refcard.pdf
- R colors: http://research.stowers-institute.org/efg/R/Color/Chart/ColorChart.pdf
- Other grad students
- Stats consulting on campus
- Google!

What is the plan for the next 1.5 hours?

- Running statistical tests
- Plotting

Basic statistics

- T-tests
- ANOVA
- Correlation
- Other linear models
- More advances statistics – EEB202C

Some advice on script layout

1. Setup (e.g., clean workspace, load packages)

```
rm(list=ls()) # clean workspace
graphics.off() #close all figures
#load packages
library('igraph')
```

2. Load data and prepare it for analysis (e.g., subset, assign data type…)
3. Perform analysis
4. Display results (plot, save output etc…)

Comparing means of groups

Sepal length
Sepal width
Petal length
Petal width

Figure 1: Three species of irises in the Anderson/Fisher data set: setosa (left), versicolor (center), and virginica (right). *Source*: The photographs are respectively by Radomil Binek, Danielle Langlois, and Frank Mayfield, and are distributed under the Creative Commons Attribution-Share Alike 3.0 Unported license (first and second images) or 2.0 Creative Commons Attribution-Share Alike Generic license (third image); they were obtained from the Wikimedia Commons.