EEB698: Data management and analysis in R for Ecologists & Evolutionary Biologists

Instructor: Dr. Haldre Rogers
334 Bessey
Monday lecture 11:00-12:30, Wednesday optional coding session 11:00-12:30
1 credit

Course description:
In this course, students will learn how to follow best practices for conducting reproducible research. Using the open-source software, R, we will practice manipulating, analyzing, and graphing the types of data commonly collected in ecology and experimental evolutionary biology. Mondays will start off with a presentation to introduce the topic, followed by a group exercise in R where students will apply this method to an example dataset. Students will be asked to apply the skills learned in class to their own datasets for homework, which they can work on during an optional group work session on Wednesdays. In doing the homework, students will work through the problems that inevitably arise when analyzing real, messy datasets plagued by missing values, small sample sizes, over-dispersion etc. For the final project, each student will be expected to analyze their own dataset, and turn in the script, methods & results sections, and manuscript-quality figures/tables. We will finish the semester with a symposium where each student will give a short presentation about his or her research.

Requirements: Students should have a basic understanding of R (e.g. how to load data, make a simple graph). A basic background in statistics (e.g. STAT401(now587)/402, EEOB 590 - Advanced Biostatistics) would also be helpful. Students should bring a dataset, ready for analysis. If you are not sure whether you are ready for this course, email the instructor (haldre@iastate.edu).

Course Objectives:
By the end of the semester, students will:
1. Be comfortable using R to import, explore, and graph data.
2. Know how to follow best practices for sharing data and code.
3. Be comfortable discussing statistical analyses and sharing code.
4. Have analyzed their own dataset, written the analysis section and the results section of a paper, and have produced manuscript-quality graphics.
5. Become part of an active community of R-users in Ecology & Evolutionary Biology at Iowa State.

Topics we will cover include:
- Intro to R, R Studio, and GitHub
- Data management plans
- Designing datasheets and databases
- Data munging/wrangling
- Data exploration
- Data visualization (base graphics and ggplot)
- Intro to linear models
- Other topics to be determined by the course participants & instructor

Resources and other courses/opportunities for building R skills:
- ***LunchinatoRs*** - Every Friday from 12-1. Informal brown bag seminar where people take turns sharing code and getting feedback or teaching others new skills. 334 Bessey.
- ***Iowa State Statistics Consulting*** is a fantastic resource. They can help with (from their website) "research design, sample size calculations, choosing statistical methods, use of statistical computing packages (R, SAS, and some JMP) to analyze data, and interpretation of results. Just make an appointment here: http://stat.iastate.edu/statistical-consulting
- **EEOB 590: Spring, Odd Years: Advanced Biostatistics with Dean Adams**
    - Review of the basic univariate and multivariate statistics commonly used in evolutionary and ecological research. The goal of the course is to give students a general idea of what statistical methods are

commonly used in evolutionary ecology, which methods are appropriate for which types of data, and to provide a general knowledge of how the methods work.

- R-ladies-Ames - local hub of a global organization for women using R.
  - rladies.org
  - meetup.com/R-Ladies-Ames
  - twitter.com/RLadiesAmes
- **Various Statistics courses useful for ecology/evolutionary biology, with sections that use R:**
  - STAT579 - Learning how to code in R;
  - STAT587- Intro to statistics
  - STAT402- Experimental design
  - STAT406- Spatial data, typically taught by Dr. Phil Dixon
  - STAT407- Multivariate analyses
  - STAT444- Bayesian data analysis
  - STAT457 - Categorical data, linear models

Reference books

- Bolker, B. M. 2008. Ecological models and data in R. Princeton University Press, Princeton, N.J.
- Clark, J. S. 2007. Models for ecological data : an introduction. Princeton University Press, Princeton, N.J.
- Crawley- The R Book. Available online at: http://www.kharms.biology.lsu.edu/CrawleyMJ_TheRBook.pdf
- Kery, M. 2010. Introduction to WinBUGS for Ecologists: Bayesian approach to regression, ANOVA, mixed models and related analyses. Academic Press.
- Zuur, A. F., E. N. Ieno, and G. M. Smith. 2007. Analysing ecological data. Springer, New York ; London. http://link.springer.com/book/10.1007%2F978-0-387-45972-1 (free download)
- Zuur, Ieno, Walker, Savelieve, and Smith. 2009. Mixed Effects Models and Extensions in Ecology with R. http://link.springer.com/book/10.1007%2F978-0-387-87458-6 (free download)

Online learning resources

https://www.coursera.org/learn/r-programming
http://ecology.msu.montana.edu/labdsv/R/labs/R_ecology.html
http://swirlstats.com/students.html
GLMM wiki: http://glmm.wikidot.com/faq
http://ropensci.github.io/reproducibility-guide/sections/introduction/

Class Schedule

Note that this is subject to change, depending on the pace of the class and what is working/not working.

Date	Topic	Homework (due following Monday)
20-Aug-18	Introduction to course, GitHub	Get setup with GitHub, Fill out Analysis_Outline
22-Aug-18	Coding Wednesday: Getting comfortable with GitHub	
27-Aug-18	Introduction to R with Erica Baken	TBD
29-Aug-18	Coding Wednesday: getting comfortable with R	
3-Sep-18	Labor Day - NO CLASS	None
5-Sep-18	Coding Wednesday: Open topic	
10-Sep-18	Introduction to Reproducible Research	Data Management Plan
12-Sep-18	Coding Wednesday: Data management plan	
17-Sep-18	Introduction to Data Wrangling, intro to tidyverse	Data wrangling exercise, your own repository
19-Sep-18	Coding Wednesday: Data wrangling	
24-Sep-18	Data Wrangling continued	Data wrangling script for your own dataset
26-Sep-18	Coding Wednesday: Data wrangling	
1-Oct-18	Exploring and Visualizing datasets, intro to ggplot2	Exploratory graphics from your own dataset
3-Oct-18	Coding Wednesday: Exploring and visualizing datasets	
8-Oct-18	Data simulation	Simulated data set & accompanying script
10-Oct-18	Coding Wednesday: Data simulation	
15-Oct-18	Choosing your statistical approach	Proposed statistical approach for your dataset
17-Oct-18	Coding Wednesday: Choosing your statistical approach	
22-Oct-18	Data analysis part 1	First draft of data analysis
24-Oct-18	Coding Wednesday: Data analysis	
29-Oct-18	Data analysis part 2	Final data analysis
31-Oct-18	Coding Wednesday: Data analysis	
5-Nov-18	Graphing part 1	First draft of graphs
7-Nov-18	Coding Wednesday: Graphing	
12-Nov-18	Graphing part 2	Final draft of graphs
14-Nov-18	Coding Wednesday: Graphing	
19-Nov-18	No class- Thanksgiving	
21-Nov-18	No class- Thanksgiving	
26-Nov-18	What to include in a paper	Written methods & results section
28-Nov-18	Coding Wednesday: Preparation for presentations	
3-Dec-18	Final presentations & peer review	
5-Dec-18	No class	

Academic Dishonesty
The class will follow Iowa State University's policy on academic dishonesty. Anyone suspected of academic dishonesty will be reported to the Dean of Students Office.

Accessibility Statement
Iowa State University is committed to assuring that all educational activities are free from discrimination and harassment based on disability status. Students requesting accommodations for a documented disability are required to meet with staff in Student Accessibility Services (SAS) to establish eligibility and learn about related processes. Eligible students will be provided with a Notification Letter for each course and reasonable accommodations will be arranged after timely delivery of the Notification Letter to the instructor. Students are encouraged to deliver Notification Letters as early in the semester as possible. SAS, a unit in the Dean of Students Office, is located in room 1076, Student Services Building or online at www.sas.dso.iastate.edu. Contact SAS by email at accessibility@iastate.edu or by phone at 515-294-7220 for additional information.

Dead Week
This class follows the Iowa State University Dead Week policy as noted in section 10.6.4 of the Faculty Handbook.

Discrimination and Harassment
Iowa State University does not discriminate on the basis of race, color, age, ethnicity, religion, national origin, pregnancy, sexual orientation, gender identity, genetic information, sex, marital status, disability, or status as a U.S. Veteran. Inquiries regarding non-discrimination policies may be directed to Office of Equal Opportunity, 3410 Beardshear Hall, 515 Morrill Road, Ames, Iowa 50011, Tel. 515-294-7612, Hotline 515-294-1222, email eooffice@iastate.edu

Religious Accommodation
Iowa State University welcomes diversity of religious beliefs and practices, recognizing the contributions differing experiences and viewpoints can bring to the community. There may be times when an academic requirement conflicts with religious observances and practices. If that happens, students may request reasonable accommodation for religious practices. In all cases, you must put your request in writing. The instructor will review the situation in an effort to provide a reasonable accommodation when possible to do so without fundamentally altering a course. For students, you should first discuss the conflict and your requested accommodation with your professor at the earliest possible time. You or your instructor may also seek assistance from the Dean of Students Office at 515-294-1020 or the Office of Equal Opportunity at 515-294-7612.

Contact Information
If you are experiencing, or have experienced, a problem with any of the above issues, email academicissues@iastate.edu