

## 2021 Fall Special Topics: Introduction to R for Biologists

Profs. Andersen and Blythe

Mondays and Thursdays, 9:00-11:00 AM

### Goals and learning objectives:

This course focuses on basic training in data science for biologists using the versatile programming language R. Because the course will be both didactic and project-driven, we will limit enrollment to 10 IBiS students. The course will be divided into three sections: (1) R basics, (2) advanced R for biologists, and (3) personal projects working one-on-one with the instructors. During the course students will bring their own dataset, use a published dataset, or choose a dataset donated by an IBiS preceptor to move through these different sections and create a final presentation. Beyond the final presentation, students will learn to be comfortable using R in their own graduate studies and future careers.

This course is appropriate for students ranging from those interested in an introduction to using R, to those who maybe were once introduced to R, but now need a refresher so that they can effectively incorporate it into their dissertation work. It will be less useful for students with extensive experience with R but could be good for students familiar with other languages who are curious to see what R has to offer.

### Additional Learning objectives:

- Organization of a project, maintenance of records
- Management of diverse data types, use of NIH FAIR data principles (Findable, Accessible, Interoperable, Reusable)
- Efficient R programming
- Development of reproducible analysis pipelines for the “future you” and your lab

Date	Topic	Prof
Sept. 23	Preliminaries, computer setup, data types in R, and basic command entry	Shelby
Sept. 27	Reading in data (other people's data), packages, saving data, Markdown	Erik
Sept. 30	Data structures: long, wide, ragged; how to deal with NAs, spaces, etc.	Erik
Pou5F1	Automating: If/Then/Else, For Loops, Custom Functions	Shelby
Oct. 7	Data processing, Good data habits, tidy data, data bigger than Excel allows	Erik
Oct. 11	Making Nice Plots, ggplot2 and the grammar of graphics	Shelby
Oct. 14, 18, 21	Personal project time	NA
Oct. 25	Advanced R: Statistical tests, significance, multiple hypothesis testing	TBD
Oct. 28	Advanced R: linear modeling, mixed models	TBD
Nov. 1	Advanced R: Clustering, PCA, dimensional reduction	TBD
Nov. 4	Advanced R: GO term enrichment, Fisher's exact test	TBD
Nov. 8	Advanced R: Integrating R with command line and python, R on an HPC	TBD
Nov. 11, 15, 18	Personal project time	NA
Nov. 29, Dec. 2	Student presentations	NA