With the advent of big data across a number of sub-disciplines within biology, there is an increasing need for biologists to be comfortable and competent with command-line data processing and analysis. This course will develop basic skills necessary for biologists working with big data sets. Topics will include UNIX commands, scripting in R and Python, version control using Git and GitHub, use of high performance computing clusters, and writing effective data-management plans. These topics will be taught using a combination of lectures and computational exercises.

Instructors: Matt Hufford, Tracy Heath, & Dennis Lavrov
Fall 2019, Wednesday & Friday, 1:10-2:30 PM