**Programming for Biologists**

BSC 4932

3 credits, Fall 2019

**Course Instructor**

Dr. Eric Goolsby
Office: BL 402D
Email: eric.goolsby@ucf.edu
Office Hours: Tuesdays 12pm-2pm or by appointment

**Course Description**

This course covers basic concepts in software applications for biologists with little or no programming background. We will be working in a UNIX environment with various biological programs, as well as the programming languages Python and R. Topics will include operating in a command line interface, basic coding techniques, and finding and using external software libraries for applications in biology. Specifically, we will cover applied topics in phylogenomics, selection analysis, statistical analysis of experimental data, toxicology, and image and morphometric analysis.

**Course Goals and Learning Objectives**

Upon completion of the course, students will be able to:

* Perform basic tasks using a command line (UNIX-based) interface.
* Automate repetitive tasks by writing custom scripts and programs.
* Understand basic programming syntax.
* Identify and implement functions from external software packages for specific biological applications.
* Solve unfamiliar biological computing problems using a combination of covered course materials and external resources (e.g., Stack Overflow, etc.).
* Self-teach more advanced programming methods.

**Prerequisites**

Statistical Methods (STA 2023), Biology II (BSC 2011C), Calculus with Analytical Geometry (MAC 2311C), and Genetics (PCB 3063).

**Class Meetings**

Combined lecture and computer laboratory meetings twice per week (Tu/Th).

**Webcourses Site**

There is a course website available through Webcourses ([https://webcourses.ucf.edu](https://webcourses.ucf.edu/)) that will be used to post materials for the course, including the syllabus, lecture slides, and grades.

**Course Text**

Recommended: “Computing for Biologists: Python Programming and Principles.” 1st Edition by Ran Libeskind-Hadas and Eliot Bush.

**Methods of Evaluation and Grading System**

Grades will be assigned on the following scale without rounding or the use of plus/minus grades:

A: 90-100% B: 80-89% C: 70-79% D: 60-69% F<60%

The grade for this course will be based on the following components:

(1) Three **take-home exams** (10% each, 30% total). All exams will be cumulative, building on programming concepts covered throughout the semester.

(2) **In-class assignments and homework** (40% total). The lowest two grades will be dropped.

(3) Leading three **group projects** (30% total).

**Course Schedule**

*Course schedule is an approximation and subject to change at any time by the instructor.*

|  |  |  |
| --- | --- | --- |
| **Meeting** | **Unit** | **Description** |
| 1 | **Unit I: UNIX** | Working in a command line interface |
| 2 | File management |
| 3 | Vi text editing |
| 4 | Manipulating genetic data (regular expressions I) |
| 5 | **Survey of biologically relevant software on Stokes / Group project I** |
| 6 | Making a single-gene phylogeny (sequence alignment, ML tree search) |
| 7 | Writing a simple guessing game (user interface, random number generation, loops) |
| 8 | Scripting a miniature phylogenetics pipeline (functions, scripts, submitting jobs) |
| 9 | Regular expressions II |
| 10 | Writing a Tic-Tac-Toe game |
| 11 | **Group project I presentations / Take-home Exam I** |
| 12 | **Unit II: Python** | Basic Python syntax I (data types, conditional statements) |
| 13 | Basic Python syntax II (loops and functions) |
| 14 | **Survey of biologically relevant Python modules / Group project II** |
| 15 | Handling genomic data |
| 16 | Performing BLAST searches |
| 17 | Accessing NCBI via Entrez |
| 18 | Selection analysis |
| 19 | **Group project II presentations / Take-home Exam II** |
| 20 | **Unit III: R** | Basic R syntax I (data types, conditional statements, loops, and functions) |
| 21 | **Survey of biologically relevant R packages / Group project III** |
| 22 | Analysis of experimental data I: t-tests, ANOVAs, multiple comparisons |
| 23 | Analysis of experimental data II: linear regression, multiple regression, and ANCOVA |
| 24 | Analysis of experimental data III: mixed models |
| 25 | Analysis of experimental data IV: generalized linear models |
| 26 | Parameter optimization: mixture toxicity |
| 27 | Simulation |
| 28 | Image and morphometric analysis |
| 29 | **Group project III presentations / Take-home Exam III (final exam)** |
|  |  | **Exam III due on scheduled day of final** |

**Course Policies**

1. Attendance is extremely important for success in this course. Additionally, most class meetings will involve graded in-class activities, which cannot be made up without a valid excuse.
2. Make-up exams and in-class assignments will be given only with valid documentation of illness, mandatory religious observance, or required participation at an authorized university activity.
3. Students are highly encouraged to discuss any and all portions of this course with me. If you are struggling, please do not wait until you fall behind to meet with me. I am available during my weekly office hours or by appointment and will always be happy to discuss the course.
4. Written communication with the instructor should be sent via UCF email. Note that I will not be able to respond to course inquiries sent from third-party email addresses (e.g. Gmail) where student identity cannot be confirmed, in order to comply with FERPA regulations.
5. Academic dishonesty (e.g. plagiarism or cheating) will not be tolerated. Students found to have committed academic dishonesty will receive an automatic grade of zero for the assignment in question. Additionally, at the discretion of the instructor and based on the severity of the violation, an overall course grade of “F” may be assigned with referral to the Office of Student Conduct.

**Accessibility**

It is my goal that this class be an accessible and welcoming experience for all students, including those with disabilities that may impact learning in this class. If anyone believes the design of this course poses barriers to effectively participating and/or demonstrating learning in this course, please meet with me to discuss reasonable options or adjustments. You may also contact SAS (Ferrell Commons 185; 407-823-2371; sas@ucf.edu) to talk about academic accommodations.