Biology 456 – Computer Skills for Biologists
Course Syllabus for Fall 2014

Professor: Dr. James A. Foster
885-7062 (but email is better)
foster _AT_ uidaho _DOT_ edu

Guest instructor: Dr. Celeste Brown

TA: Daniel Caetano da Silva, BCB PhD student

Summary: This course develops skills to manage and analyze complicated datasets such as those in molecular evolution, systematics, (meta)genomics, and transcriptomics. Datasets in biology are growing explosively, so computational skills are vital for graduate studies and technical careers in the life sciences. This course will use demonstrations, exercises, and student projects to learn advanced Unix skills, Python programming, and data management. This course is explicitly designed to prepare students for independent research in computational biology and biological sciences and will be helpful for those students taking CS 515, Computational Biology, Sequence Analysis; Biol 421, Advanced Evolutionary Biology; Biol 444, Genomics; or Biol 545, Principles of Systematic Biology.

Requirements: Stat 251 and Biology 210, or permission of instructor, class is strictly limited to the first 24 students

Class Hours: Tuesday 12.30pm to 2.00pm and Fridays 1.30pm to 3pm, LSS 440 (IBEST Classroom), or as posted

Website: Course website: http://computerskillsforbiologists.wordpress.com. You are responsible for checking regularly for reading materials, exercises, and exams.

Format: Lectures, demonstrations, and hands-on exercises on computer workstations, and student discussion. There will be regular homework assignments–usually online tutorials. Graduate students will prepare and present a final project, which will demonstrate facility with course materials. Undergraduates will have regular homework.

Grading: Grades will be determined by the instructor based on demonstrated mastery of the material as determined by in-class

activities, exercises, and the final project. *My philosophy is to give you the highest grade that won't embarrass either of us.* My decisions are final.

Course Materials: All course materials will be free and online. The textbook will be *Python for Informatics: Exploring Information*, Charles Severance. The book is available in several formats at *http://www.pythonlearn.com/book.php*. We will also use online tutorials, including *http://wiki.scipy.org/Tentative_NumPy_Tutorial, http://scipy-lectures.github.io/intro/matplotlib/matplotlib.html,* and *http://biopython.org/DIST/docs/tutorial/Tutorial.html*.

Makeup policy: There is none. All assignments and exams must be completed and turned in on time. Plan ahead.

Academic honesty: Anything you turn in must be your own work. I will be very unforgiving of plagiarism. If in doubt, ask me. Use available resources, but cite your sources. Feel free to discuss things online. Discuss and help each other learn the material. But any copying of turned in work from other students or elsewhere will be punished harshly, as governed by Article II of the University if Idaho's Student Code of Conduct (http://www.webs.uidaho.edu/fsh/2300.html). All incidents of academic dishonesty will be reported to the dean of students. Individuals guilty of academic dishonesty will receive a failing grade for the course and may face further disciplinary action. **Don't be evil.**

Civility: Discussions and interactions in class must be kept civil. Offenders will be asked to leave.

Changes: *This course is under development, so details (including the course schedule) may change periodically.*

Syllabus: After a serious discussion in class of our goals and objectives for the course, we agree to change the syllabus structure. On Fridays we will walk through each others' code, which can be either Rosalind solutions, answers to textbook exercises, or anything else. Rosalind homework will be due on **Tuesday Morning at 10:00** (see schedule below). You should look at each others' code and choose which solutions they want to go over in class. You could also send me code to go over in class, but please do it before the end of the day Tuesday. The objective is for everyone to have a copy of the code we are discussing before class on Friday. On Tuesdays I will lecture about technical aspects of Python programming and software engineering in general. We will cover the topics in the syllabus in order until we run out of time, rather than feeling compelled to go over everything.

Homework and Reading

Day	Date	Homework
F	12-Sep	INI1
F	19-Sep	INI2
F	26-Sep	INI3, INI4, DNA
T	7-Oct	RNA, REVC, HAMM
T	17-Oct	GC, SUBS, FIB
T	28-Oct	INI5, PROT, INI6
T	4-Nov	ORF, IPRB, CONS
T	11-Nov	GRPH, PERM
T	18-Nov	TREE, PROB
T	2-Dec	KMP
T	9-Dec	PDST, LEXF

Day	Date	Topic	Chap.
F	12-Sep	Variables, Expressions, Statements	Ch 2
F	19-Sep	if/then/else; subroutines	Ch 3
T	23-Sep	Functions	Ch 4
F	3-Oct	Iteration: for, loops; while	Ch 5
F	10-Oct	Strings	Ch 6
T	21-Oct	Files, dictionaries	Ch 7, 9
T	28-Oct	Lists	Ch 8
T	4-Nov	Tuples, list comprehension	Ch 10
T	11-Nov	Regular expressions	Ch 11
T	18-Nov	BioPython	online
T	9-Dec	Presentations	
F	12-Dec	Presentations	

Topics

Course intro; Basics of bash and unix

Really helpful commands

ADH as an example

Scripting

git, directory structure, self-documentation, Rosalind, resources

Variables, Expressions, Statements

Intro computing, CS, performance, architecture, OS, pipelines/workflows

if/then/else; subroutines

Functions

diff langs, hardware, lang types, ...

Programming style, debugging tips, regression testing

Iteration: for, loops; while

Strings

Files

Lists

sets, dictionaries

Tuples

RE package

pdb and online help

Biopython package

Matplotlib package