Oregon Health & Science University
Department of Medical Informatics and Clinical Epidemiology
BMI 565/665
Bioinformatics Programming and Scripting
Fall Term 2014
Tuesdays and Thursdays 9:00am-10:25am, BICC 124

Instructor: Michael Mooney, Ph.D. (mooneymi@ohsu.edu)
Teaching Assistant: Ryan Swan (swanr@ohsu.edu)
Office Hours: by appointment
Required Textbook: Python for Bioinformatics by Sebastian Bassi. CRC Press, 2009.
ISBN#9781584889298.
Suggested Resources:
- Python for Data Analysis by Wes McKinney, O'Reilly Media, 2012
- Python Essential Reference by David Beazley, 4th Edition, 2009, ISBN#9780672329784.
- Think Python by Allen B. Downey, http://www.greenteapress.com/thinkpython/

Course Description: The purpose of this course is to equip research scientists with
computational skills necessary to create and automate tools to analyze biological data. The course
is divided into four sub-topics: python programming, scripting in Unix, BioPython library,
bioinformatics workflows. Python will be used to solve simple to sophisticated programming
problems and to review general computational language paradigms such as problem abstraction,
data types, file I/O, iteration, functions, and objects. There will also be an emphasis on writing
Unix operating system shell scripts to automate repetitive tasks and connect disparate
bioinformatics tools using files and pipes. In addition, students will learn to access public
repositories to perform basic bioinformatics tasks such as annotating gene products, sequence
searching, and functional queries. This course is designed to be a first year requirement for
students in the Bioinformatics and Computational Biology graduate program in Biomedical
Informatics.

Course Objectives:
- Working with Problems Algorithmically: Be able to create and automate bioinformatics
  tasks using a high level programming language.
- Working with Pipelines: Using scripting in a Linux operating system to execute and
  connect previously written bioinformatics tools in a workflow
- Working with Files and Public Repositories: Read and write text data from files and
  access data and annotation from common bioinformatics public repositories.
- Working with Bioinformatics Data: Be able to perform basic data processing tasks and
  create a graphical representation of results.

Prerequisites:
Background must include an introductory programming class including concepts such as data
types, loops, I/O, methods, and basic algorithms.

Course Grading Policy:
Grades will be based on scores from examinations, weekly programming assignments, and a final
research project. The point breakdown is as follows:

Programming Assignments	40%
Research Project	20%
Mid-Term Exam	20%

| | Final Exam | 20% |

The final course grade will be based on the following grade breakdown:

93-100	A
87-90	B+
83-87	B
80-83	B-
77-80	C+
73-77	C
70-73	C-
<70	F

Programming assignments will be handed out each Tuesday and will be due the following Monday at 5:00pm. Research projects will involve the analysis of a high throughput dataset over the last 2 weeks of class and will require a 3 page write-up. All late assignments will be penalized 33% for each day late. Assignments should be submitted through Sakai.

Graduate Studies in the OHSU School of Medicine is committed to providing grades to students in a timely manner. Course instructors will provide students with information in writing at the beginning of each course that describes the grading policies and procedures including but not limited to evaluation criteria, expected time needed to grade individual student examinations and type of feedback they will provide.

Class grades are due to the Registrar by the Friday following the week of finals. However, on those occasions when a grade has not been submitted by the deadline, the following procedure shall be followed:

1) The Department[1] /Program Coordinator[2] will immediately contact the Instructor requesting the missing grade, with a copy to the Program Director and Registrar.
2) If the grade is still overdue by the end of next week, the Department[1] /Program Coordinator[2] will email the Department Chair directly, with a copy to the Instructor and Program Director requesting resolution of the missing grade.
3) If, after an additional week the grade is still outstanding, the student or Department[1] /Program Coordinator[2] may petition the Office of Graduate students for final resolution.

> 1   *For courses that are run by a specific department.*
> 2   *For the conjoined courses (course number is preceded by CON_ that are run by Graduate Studies.*

Attendance policy:
If possible…don't miss class! Consider getting notes from a kind fellow student.

Tentative Class Schedule:

Week	Date	Topic
1	30-Sep	Intro, Computing Resources, Linux Basics, Installation Python: Help/Documentation, Coding Best Practices, Data Types In the book (Bassi): Chapters 1, 2, 3, pp. 112, 393-408 Python/Linux Tutorial
1	2-Oct	Python: Operators, String Formatting, Control Structures, File I/O In the book: Chapters 3, 4 and 5, pp. 464-478

2	7-Oct	Python: File I/O continued, Linux/Unix Commands In the book: Chapters 5
2	9-Oct	Python: Bash Scripting
3	14-Oct	Python: Regular Expressions In the book: Chapter 9
3	16-Oct	Python: XML In the book: Chapter 12
4	21-Oct	BioPython: Entrez E-Utils In the book: section 10.2.11 BioPython Tutorial: http://biopython.org/DIST/docs/tutorial/Tutorial.html
4	23-Oct	Python: Functions, Generators, Object-Oriented Programming In the book: Chapters 6, 8
5	28-Oct	Python: Graph Data Structures (NetworkX), Intro to Graph Algorithms
5	30-Oct	Python: Modules, Graph Algorithms, Algorithm Analysis In the book: Chapter 6
6	4-Nov	Python: Error Handling, Exceptions, Code Testing, Exam Review In the book: Chapter 7
6	6-Nov	Mid-term Exam
7	11-Nov	BioPython: seqRecords, seqIO, Alignments, alignIO In the book: Chapter 10 BioPython Tutorial
7	13-Nov	BioPython: BLAST Research Project Assigned In the book: Chapter 10 BioPython Tutorial
8	18-Nov	Python: NumPy, matplotlib
8	20-Nov	Python: Statistics with Pandas, Statsmodels, (Rpy2)
9	25-Nov	Thanksgiving Break
9	27-Nov	Thanksgiving Break
10	2-Dec	Python: Statistics and Machine Learning with SciPy, scikit-learn
10	4-Dec	Python: scikit-learn continued
11	9-Dec	Python: Parallel Processing, Map-Reduce (mrjob), Exam Review
11	11-Dec	Final Exam

Student Access:
OHSU is committed to providing equal access to qualified students with disabilities. Student Access determines and facilitates reasonable accommodations, including academic adjustments and auxiliary aids, for students with documented disabilities. A qualified student with a disability is a person who meets the academic and technical standards requisite to admission or participation in a particular program of study. As defined by the Americans with Disability Act (ADA), a person with a disability has a physical or mental impairment that substantially limits one or more major life activities of the individual. This may include, but is not limited to, physical conditions, chronic health issues, sensory impairments, mental health conditions, learning disabilities and ADHD. Student Access works with students with disabilities from all of OHSU's educational programs and at each campus.

Each school has an assigned Program Accommodation Liaison (PAL), who acts as an "in-house" resource for students and faculty concerning access issues for students with disabilities. The PAL

works in collaboration with Student Access to implement recommended accommodations for students with disabilities.

It is recommended that you contact Student Access to consult about possible accommodations if you a) received disability accommodations in the past, b) begin experiencing academic difficulties, and/or c) are given a new diagnosis from your healthcare provider.

Learn more about Student Access:
Phone: 503 494-0082
Email: studentaccess@ohsu.edu
Website: www.ohsu.edu/student-access

Academic Integrity:
The students will be responsible for following the OHSU guidelines for academic integrity. You may discuss the general concepts and principles behind an assignment with other students. In fact, you are encouraged to do this whenever possible, because it is often a valuable way to reinforce ideas, and to learn new perspectives. However, in doing assignments, each student is expected to develop, write up, and hand in an individual solution and, in doing so, develop a sufficient understanding of the problem and solution so as to be able to explain it adequately to the instructor. Under no circumstances should a student copy or consult the solution of another student, or copy a solution from any other source, including the Internet.

Literature and resources (including Internet resources) employed in fulfilling assignments must be cited. See http://www.ohsu.edu/xd/education/library/research-assistance/plagiarism.cfm?WT_rank=1# for information on code of conduct for OHSU and http://www.ohsu.edu/xd/education/teaching-and-learning-center/for-students/index.cfm for more information on citing sources and recognizing plagiarism.

Copyright Information
Every reasonable effort has been made to protect the copyright requirements of materials used in this course. Class participants are warned not to copy, audio, or videotape in violation of copyright laws. Journal articles will be kept on reserve at the library or online for student access. Copyright law does allow for making one personal copy of each article from the original article. This limit also applies to electronic sources.

To comply with the fair use fair use doctrine of the US copyright law, Sakai course sites close three weeks after grades are posted with the Registrar. Please be sure to download all course material you wish to keep before this time as you will have no further access to your courses.

Use of Sakai
This course will have an online component, which can be accessed through Sakai, OHSU's online course management system. For any technical questions or if you need help logging in, please contact the Sakai Help Desk.

Hours: Sakai Help Desk is available Mon – Fri, 8 am – 10 pm and weekends and holidays 12 pm – 5 pm
Contact Information:
(Local) 503-494-7074 (4-7074 on campus)
(Toll-free) 877-972-5249
(Web) http://atech.ohsu.edu/help
(Email) sakai@ohsu.edu

Inclement Weather Policy

When the weather forecaster is calling for ice or snow, call the **OHSU Alert Line, 503 494-9021**, for information regarding weather conditions that may affect operations at OHSU. This hot line will offer specific recorded messages for road conditions on OHSU's Marquam Hill and West campuses (option 1), and for patients (option 2), students (option 3) and employees (option 4).

If extreme weather conditions present potentially unsafe situations, the provost of the university may choose to delay or cancel classes, or alter office and research activities. If classes are canceled or delayed, residents and students who have patient care responsibilities must meet those obligations.

For more information, please view the website
http://www.ohsu.edu/xd/about/visiting/weather/index.cfm or call the above hotline.

DMICE COMMUNICATION POLICY
1. If the syllabus directs the student to contact the TA before contacting the instructor, the student should do so. Otherwise, the student should contact the instructor and allow 2 business days (not including weekends) for a response.
2. If the student does not receive a response from the instructor within 2 business days, s/he should contact the TA (if there is one). When contacting the TA s/he should cc the instructor and Diane Doctor at doctord@ohsu.edu.
3. If a student does not receive a response from the TA within 1 business day (not including weekends), s/he should contact Diane Doctor at doctord@ohsu.edu and cc the instructor and the TA.
4. If Diane does not reply within 1 business day (not including weekends), the student should contact Andrea Ilg at ilgan@ohsu.edu.
5. Students having difficulties with Sakai should contact the Sakai Help Desk at sakai@ohsu.edu or at (877) 972-5249. Sakai help is available M-F from 8am to 10-pm and weekends from Noon to 5pm. Do not contact the instructor.