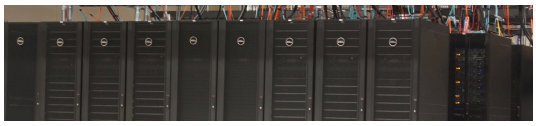
UNIVERSITY OF FLORIDA

COMPUTATIONAL TOOLS FOR RESEARCH

BSC 4926/6927
3 Credits
Fall 2018
Room: Rogers 110
M,W,F 8:30-9:20 (2nd Period)



INSTRUCTOR:

Matt Gitzendanner
magitz@ufl.edu
Dickinson Hall room 301c
273-1960
Office Hours: Tuesdays 12:30-1:30; Fridays 9:30-10:30, or by appointment

CATALOG DESCRIPTION

Information technology has dramatically transformed how research across many disciplines is conducted. This is particularly true in the biological sciences where researchers frequently find themselves faced with massive amounts of diverse data to analyze. As data types and volumes continue to grow, knowledge of scripting, database management, and advanced computing skills are critical for researchers.

Topics will address a gap in how research has advanced—*and become increasingly computational*— while student training in the use of computational tools has lagged. The course will cover basic concepts that will provide the ability for students to apply new technologies to a wide array of research questions. A foundation in information management concepts opens doors for well-trained scientists and allows them to work in multidisciplinary research.

This course will survey areas where high performance computing, large-scale data access and integration, informatics tools and software, and multi-disciplinary collaboration have had high impact on research as a foundation for computationally-enabled research.

THE COURSE IS ROUGHLY DIVIDED INTO THREE SECTIONS:

SECTION 1: Linux command line, Bash scripting, version control and using high-performance computing resources
SECTION 2: Python scripting
SECTION 3: SQL database introduction and integration with Python

PRE-REQUISITES AND CO-REQUISITES

Junior or senior standing.

COURSE OBJECTIVES

By the end of the course, the student will:

- Demonstrate how technology infrastructure can improve research and open new avenues of investigation.
- Competently navigate the Unix/Linux command line interface.
- Effectively and efficiently manipulate text files, performing complex regular expression replacements, reformatting and merging files in various ways.
- Raise and address current issues through class participation and discussion.
- Use High Performance Computing resources such as the UF Research Computing for cluster-based analyses. Including batch scripting and running multi-processor applications (threaded and MPI).
- Explain the basic anatomy of computer scripts/programs, with particular focus on Python scripting.
- Construct analytical pipelines to accomplish complex tasks.
- Describe basic database design, creation and manipulation. Perform scripted database operations for information discovery, data exploration and research data curation.
- Have a basic understanding of research graphics formats, preparation and manipulation

COURSE WEBSITE AND COMMUNICATIONS

Course materials and related information will be posted on the course E-Learning (Canvas) website at http://lss.at.ufl.edu. You are responsible for all announcements made in class and/or posted on the course website for this course.

Students are encouraged to use the course discussion board in Canvas for questions and discussion, but should not hesitate to email the instructor.

REQUIRED MATERIALS

FEES

None

TEXTBOOKS OR OTHER READINGS

This course will use various free texts and web sites.The main texts for the course are:

- The Linux Command Line: http://linuxcommand.org
- Python For Everyone: https://www.py4e.com

Each of these is available as a **free PDF download** or for purchase in print.

SOFTWARE AND HARDWARE

PARTICIPANTS WILL BE REQUIRED TO HAVE A PORTABLE COMPUTER WITH ABILITY TO CONNECT TO THE INTERNET IN CLASS EACH PERIOD. As access to power during course time may be

limited, students should ensure their computer is charged and able to function for the 50 minute class period.

Several free/open source software packages will be used throughout the course, and students will be required to install some of these.

Students will be required to apply for a (free) Research Computing account to access HiPerGator for coursework.

Students will be required to apply for a (free) Github.com account for coursework.

FOR ISSUES WITH TECHNICAL DIFFICULTIES FOR CANVAS, PLEASE CONTACT THE UF HELP DESK AT:

- http://helpdesk.ufl.edu
- (352) 392-HELP (4357)
- Walk-in: HUB 132

Any requests for make-ups due to technical issues MUST be accompanied by the ticket number received from the Help Desk when the problem was reported to them. The ticket number will document the time and date of the problem. You MUST e-mail your instructor within 24 hours of the technical difficulty if you wish to request a make-up.

All faculty, staff and student of the University are required and expected to obey the laws and legal agreements governing software use. Failure to do so can lead to monetary damages and/or criminal penalties for the individual violator. Because such violations are also against University policies and rules, disciplinary action will be taken as appropriate.

COURSE OUTLINE

Week	Date	Topic
1	8/22/18 8/24/18	Introduction and course objectives. UF Research Computing Intro & getting started
2	8/27/18 8/29/18 8/31/18	Linux Basics: Command line Linux Basics: Pipes and Redirects Linux Basics: Text files, regular expressions
3	9/3/18 9/5/18 9/7/18	**Labor Day Holiday—No Class** Shell Scripts Shell Scripts, Version Control: git and GitHub
4	9/10/18 9/12/18 9/14/18	Version Control: git and GitHub Using UF Research Computing resources Running batch jobs, Compiling source code
5	9/17/18 9/19/18 9/21/18	Singularity/Docker Data Management/Data Curation Wrap-up
6	9/24/18 9/26/18 9/28/18	Introduction to Python Python data types Working in Python, File I/O
7	10/1/18	Working in Python, Loops, Conditional Statements

	10/3/18 10/5/18	Working in Python: try/except, RegEx, imports SciPy, NumPy, Pandas
8	10/8/18 10/10/18 10/12/18	Class project, begin Scripting data acquisition **UF Homecoming—No Class**
9	10/15/18 10/17/18 10/19/18	Python modules and libraries Working with dictionaries Class project, continue
10	10/22/18 10/24/18 10/26/18	Writing functions Class project, complete Matplotlib and data visualization
11	10/29/18 10/31/18 11/2/18	Overview of databases Database design MySQL
12	11/5/18 11/7/18 11/9/18	Setting up a database SQL Queries, joins, etc. SQL Queries
13	11/12/18 11/14/18 11/16/18	**Veteran's Day Holiday–No class** Python scripting of database actions Uses of databases
14	11/19/18 11/21/18 11/23/18	Class projects begin **Thanksgiving break—No class** **Thanksgiving break—No class**
15	11/26/18 11/28/18 11/30/18	Security and data management Class projects, continue Graphics (Ch. 17 of text)
16	12/3/18 12/5/18	Class projects, complete Wrap-up—Last day of classes.

GRADING

- QUIZZES: 5 @ 20 points each (25% of final grade)
- PROBLEM SETS: 5 @ 20 points each (25% of final grade)
- GITHUB ASSIGNMENTS: 5 @ 10 points each (25% of final grade)
- CLASS PROJECTS: 2 @ 40 points each (20% of final grade)
- CLASS PARTICIPATION: 20 points (5% of final grade)

GRADE DISPUTES:

Should a student wish to dispute any grade received in this class (other than simple addition errors), the dispute must be in writing and be submitted to the instructor within a week of receiving the grade. The dispute should set our very clearly, the grade that the student believes the assignment should have received as well as why he or she believes that he or she should have received such a grade.

GRADING SCALE (& GPA EQUIVALENT):

A	A-	B+	B	B-	C+	C	C-	D+	D	D-	E
100-93	92-90	89-87	86-83	82-80	79-77	76-73	72-70	69-67	63-66	62-60	59-
(4.0)	(3.67)	(3.33)	(3.0)	(2.67)	(2.33)	(2.0)	(1.67)	(1.33)	(1.0)	(0.67)	(0)

Note: A grade of C- is not a qualifying grade for major, minor, Gen Ed, or College Basic distribution credit. For further information on UF's Grading Policy, see:
https://catalog.ufl.edu/ugrad/current/regulations/info/grades.aspx#hgrades
http://www.isis.ufl.edu/minusgrades.html

COURSE POLICIES:

QUIZ AND ASSIGNMENT DATES/POLICIES:

Quiz and assignment dates will be announced at least one week in advance and students will have at least one week to complete the quiz or assignment. Each quiz or assignment will clearly state if it is an individual or group assignment. Individual assignments must be the student's own work, completed without the assistance of others. **ALL QUIZZES AND ASSIGNMENTS ARE "OPEN BOOK, OPEN INTERNET"**, you may use whatever resources you desire to complete the quiz/assignment.

MAKE-UP AND LATE POLICY:

Please notify the instructor of circumstances that lead to late work or missed classes. I will generally work with you and accept late work. Without prior notification, late work will be penalized 5% per week after the due date.

ATTENDANCE & MAKEUP POLICY:

Requirements for class attendance and make-up exams, assignments, and other work in this course are consistent with university policies that can be found in the online catalog at:
https://catalog.ufl.edu/ugrad/current/regulations/info/attendance.aspx.

CELL PHONE AND TEXTING POLICY:

Students must turn cell phones to vibrate before coming to class. Each time a student's cell phone rings or each time that a student texts during class, 1% will be deducted from that student's final grade for each instance.

ACADEMIC HONESTY:

All students registered at the University of Florida have agreed to comply with the following statement: "I understand that the University of Florida expects its students to be honest in all their academic work. I agree to adhere to this commitment to academic honesty and understand that my failure to comply with this commitment may result in disciplinary action up to and including expulsion from the University."
In addition, on all work submitted for credit the following pledge is either required or implied: "*On my honor I have neither given nor received unauthorized aid in doing this assignment*."
If you witness any instances of academic dishonesty in this class, please notify the instructor or contact the Student Honor Court (392-1631) or Cheating Hotline (392-6999). For additional information on Academic Honesty, please refer to the University of Florida Academic Honesty Guidelines at:
http://www.dso.ufl.edu/judicial/procedures/academicguide.html.

ACCOMMODATIONS FOR STUDENTS WITH DISABILITIES:

Students who will require a classroom accommodation for a disability must contact the Dean of Students Office of Disability Resources, in Peabody 202 (phone: 352-392-1261). Please see the University of Florida Disability Resources website for more information at: http://www.dso.ufl.edu/drp/services/.
It is the policy of the University of Florida that the student, not the instructor, is responsible for arranging accommodations when needed. Once notification is complete, the Dean of Students Office of Disability Resources will work with the instructor to accommodate the student.

COUNSELING AND MENTAL HEALTH RESOURCES:

Resources are available on-campus for students having personal problems or lacking clear career and academic goals. The resources include:

- UF Counseling & Wellness Center, 3190 Radio Rd, 392-1575, psychological and psychiatric services.
- Career Resource Center, Reitz Union, 392-1601, career and job search services.

Many students experience test anxiety and other stress related problems. "A Self Help Guide for Students" is available through the Counseling Center (301 Peabody Hall, 392-1575) and at their web site: http://www.counsel.ufl.edu/.

COURSE EVALUATION PROCESS:

Students are expected to provide feedback on the quality of instruction in this course based on 10 criteria. These evaluations are conducted online at https://evaluations.ufl.edu. Evaluations are typically open during the last two or three weeks of the semester, but students will be given specific times when they are open. Summary results of these assessments are available to students at https://evaluations.ufl.edu/results

CLASS DEMEANOR OR NETIQUETTE:

Students are expected to arrive to class on time and behave in a manner that is respectful to the instructor and to fellow students. Please avoid the use of cell phones and restrict eating to outside of the classroom. Opinions held by other students should be respected in discussion, and conversations that do not contribute to the discussion should be held at minimum, if at all.