Computational Tools for Research in Biology

Summary: An introduction to computational tools for research

 Edit me

> **Note:** This page is under construction.

About this site

This site is the being developed as the main repository for Matt Gitzendanner's Computational Tools for Research in Biology. This course is typically offered every-other-year at the University of Florida . My hope is that this site can serve both as the primary resource when teaching the course, and as a resource for others teaching similar courses or learning on their own.

The course was designed to be a one semester, in-person, 3-credit course for upper-level undergraduate and starting graduate students. During Summer 2020, I am offering it as an online course with some modifications.

Please use the feedback button to send feedback.

Course 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 covered in the course 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 assumes no prior coding or command line skills, and covers concepts that will provide the ability for students to apply new technologies to a wide array of research questions. A foundation in data management and analysis 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.

Course Organization

The course is divided into three main sections:

- Section 1
    - Linux command line
    - Bash scripting
    - Version control
    - Using high-performance computing resources
- Section 2
    - Python scripting
- Section 3
    - SQL database introduction and integration with Python

Course Objectives

- 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

Meeting Times

Summer A runs May 11, 2020 through June 19, 2020. May 25th is Memorial Day holiday.

Synchronous Meetings: Mon, Wed, Fri 11:00am - 12:15pm

- I will be online and presenting material and/or helping students: **Mon, Wed, Fri during 3rd period (11:00am - 12:15pm)**

> ⚠ **Important:** You should make every effort to attend these sessions. While there will be note pages and hopefully a video recording of these sessions, **this is the best opportunity to ask questions and get help from me and others in the class.**

- I understand that with all that is going on, some students will need to miss classes sometimes. That is fine and I will do my best to help you catch up, but regular attendance is the best way to learn.

Asynchronous Work: Tue, Thurs

- On Tuesday and Thursday, students will work on projects either independently or in small, virtual groups.

- These sessions can be more flexible in timing and will give you and your classmates times to get more hands-on experience and challenge yourselves to learn on your own.

- There will still be content that needs to be covered, and students should plan on working on Tuesday and Thursday, but can generally do the work whenever is convenient for them.

Office Hours: Thursday 11:00 to 2:00 and by Appointment

- I have scheduled office hours for **Thursday between 11:00am and 2:00pm**. I am also available at other times by appointment.

> ☑ **Tip:** Coding is not always easy. Simple solutions are not always obvious. There will be some frustration. I expect that you will need help. You should expect that you will need help. I want to help you! I cannot always help if you don't ask. Please ask for help.

Course Textbooks

The main texts for the course are:

- The Linux Command Line: http://linuxcommand.org ⬈ (my notes and referenced page numbers will be based on the 19.01A PDF, Fifth internet edition from Jan 28, 2019)

- Python For Everyone: https://www.py4e.com/book.php ⬈

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

Course Calendar

This is subject to change, please check back frequently.

For readings, there may be links to pages with my notes and additional explanations on the content from the texts. The texts are abbreviated as TLCL = The Linux Command Line ⬈; Py4E = Python for Everyone ⬈.

Week	Date	Reading/Assignment	Topic
1	Mon, May 11	Read TLCL Introduction & Ch 1-6 (/TLCL_1.html) Download Software (/software.html)	Introduction and course objectives ⬈ Getting started: Computers ⬈ The Linux command line
1	Tues, May 12	Watch UFRC Training Video ⬈ Read TLCL Ch 7 & 8 (/TLCL_2.html)	UF Research Computing Intro & getting started
1	Wed, May 13	Read TLCL Ch 19 & 20 through p.293 (/TLCL_3.html)	Regular Expressions Handout (/images/RegEx_CheetSheet.pdf)
1	Thurs, May 14	Watch Learn the Linux Command Line (/LinkedInLearningLinux.html) Register for github.com account (/github_account.html) (Includes Quiz 1: Due by start of class Friday, May 15)	Keep working on Linux command line skills
1	Fri, May 15	Read TLCL Ch 24-26 (/TLCL_4.html) Problem Set 1: Due Tues, May 19 (In Canvas)	Shell Scripts and version control with git and GitHub
2	Mon, May 18	Read TLCL Ch 27, 29 & 33 (/TLCL_5.html)	Flow control: `if`, `while`, `until`, `for` etc.

Week	Date	Reading/Assignment	Topic
2	Tues, May 19	Read TLCL Ch 23 (/TLCL_6.html) Problem Set 1 due	Start Problem Set 2 (details in Canvas)
2	Wed, May 20	Work on Problem Set 2	Using UF Research Computing resources Running batch jobs, Compiling source code
2	Thurs, May 21	Work on Problem Set 2	Catching up
2	Fri, May 22	Problem Set 2 Due Read Py4E Ch 1 (/py4e_1.html) Read Py4E Ch 2 ☑	Introduction to Python Python data types
3	Mon, May 25		*Memorial Day Holiday, No Class*
3	Tues, May 26	Read Py4E Ch 3 & 5	
3	Wed, May 27		Working in Python, Flow Control
3	Thurs, May 28	Read Py4E Ch 6 & 7	Working in Python: try/except, Strings, File I/O
3	Fri, May 29	Read Py4E Ch 4	Working in Python, Functions
4	Mon, June 1	Read Py4E Ch 8-10	Lists, Dictionaries, Tuples

Week	Date	Reading/Assignment	Topic
4	Tues, June 2	Read Py4E Ch 12 & 13	SciPy, NumPy, Pandas
4	Wed, June 3		Scripting data acquisition
4	Thurs, June 4		Class project
4	Fri, June 5	P3 Due, Quiz 2	Matplotlib and data visualization
5	Mon, June 8	Read Py4E Ch 16	Writing Functions
5	Tues, June 9		Class project, continue
5	Wed, June 10		Finishing up
5	Thurs, June 11		Class project, continue
5	Fri, June 12	Programming Foundations Databases	Overview of databases Database design
6	Mon, June 15		Setting up a database SQLAlchemy Introduction

Week	Date	Reading/Assignment	Topic
6	Tues, June 16		SQLAlchemy continued
6	Wed, June 17		Graphics
6	Thurs, Jun 18		Class project, complete
6	Friday, June 19		Last day of class

Software and Hardware

Participants will need a computer with internet connection, webcam and microphone for all classes.

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.

If you have technical difficulties with Cancas, please contact the UF Helpdesk at:

- http://helpdesk.ufl.edu

- (352) 392-HELP (4357)

- Walk-in: HUB 132

Any requests for make-ups due to technical issues should 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. Please e-mail the 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.

Grading

Coming soon

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 out 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 and GPA Equivalent

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

Course Policies

Class Attendance and Makeup Policy

Requirements for class attendance and makeup 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/UGRD/academic-regulations/attendance-policies/ ⬀

Quiz and Assignment Policy

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. Though only assignments specifically noted as group assignments should be worked on with other people.

Makeup 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 day after the due date.

Students Requiring Accommodations

Students with disabilities requesting accommodations should first register with the Disability Resource Center (352-392-8565, https://disability.ufl.edu/students/get-started/ ⧉) by providing appropriate documentation. Once registered, students will receive an accommodation letter which must be presented to the instructor when requesting accommodation. Students with disabilities should follow this procedure as early as possible in the semester.

Course Evaluation

Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at gatorevals.aa.ufl.edu/students/ ⧉. Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via ufl.bluera.com/ufl/ ⧉. Summaries of course evaluation results are available to students at gatorevals.aa.ufl.edu/public-results/ ⧉.

Class Demeanor and 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.

Cell Phone and Texting Policy

Students should turn cell phones to vibrate before coming to class. Cell phone rings and notifications disrupt fellow students and the instructor. Please be courteous to all and silence your phone.

Computer Use

This is a hands-on computational class and students are expected to have their computer at every class. Students should be working on course content while in class. Students are expected to be respectful of others with the content visible on their computers. Sometimes videos are a helpful way to learn new content. If that is the case, please use headphones so that the audio is not audible to others in class.

Discussion Boards

The Canvas discussion boards can be used to ask for and provide help by all. Students should be supportive and considerate of others at all times. Rude or inappropriate comments will be removed and the poster will be warned.

University Honesty Policy

UF students are bound by The Honor Pledge which states:

> We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: "On my honor, I have neither given nor received unauthorized aid in doing this assignment."

The Honor Code ⧉ specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor or TAs in this class

Health and Wellness

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: https://counseling.ufl.edu/ ⧉.

- **U Matter, We Care**: If you or a friend is in distress, please contact umatter@ufl.edu or 352 392-1575 so that a team member can reach out to the student.

- **Counseling and Wellness Center**: https://counseling.ufl.edu/ ⧉, 392-1575; and the

- **University Police Department**: 392-1111 or 9-1-1 for emergencies. http://www.police.ufl.edu/ ⧉

- **Sexual Assault Recovery Services (SARS)**: Student Health Care Center, 392-1161.

- **Student Health Care Center**: Call 352-392-1161 for 24/7 information to help you find the care you need, or visit shcc.ufl.edu ⧉

- **Food insecurity**: The Pantry is a resource on the University of Florida campus committed to supporting students, staff, and faculty who are experiencing food insecurity. These individuals do not have reliable access to nutritious foods for themselves and their families. If you, or anyone you know, is experiencing food insecurity, the Pantry is a resource to visit. We offer non-perishable food, toiletries and fresh produce grown at the Field and Fork Gardens during certain times of the year.

There is no proof of need required in order to shop at the Pantry; you must only bring in your valid UFID card. At the Pantry, we know that a good meal makes for a good student, and we work to support all Gators who are experiencing food insecurity. Field & Fork Food Pantry ⎘

Inclusive Learning Environment

This course embraces the University of Florida's Non-Discrimination Policy, which reads:

> The University shall actively promote equal opportunity policies and practices conforming to laws against discrimination. The University is committed to nondiscrimination with respect to race, creed, color, religion, age, disability, sex, sexual orientation, gender identity and expression, marital status, national origin, political opinions or affiliations, genetic information and veteran status as protected under the Vietnam Era Veterans' Readjustment Assistance Act.

If you have questions or concerns about your rights and responsibilities for inclusive learning environment, please see the instructor or refer to the Office of Multicultural & Diversity Affairs website: http://multicultural.ufl.edu ⎘.

Privacy

There are federal laws protecting your privacy with regards to grades earned in courses and on individual assignments. For more information, please see: https://registrar.ufl.edu/ferpa.html ⎘

[]:

[]:

[]:

[]: