

Course Content

Linux/Bash/HPC

Python

SQL


Week-by-Week Pages

Additional Resources



Computational Tools for Research in Biology

Summary: An introduction to computational tools for research-- Linux command line, HPC, Bash scripting, Python, SQL, and Artificial Intelligence

 Edit me [↗](#)

BSC4452 and BSC6451 Registration information for Fall 2024

- **BSC 4452:** Class 19032
- **BSC 6451:** Class 22340

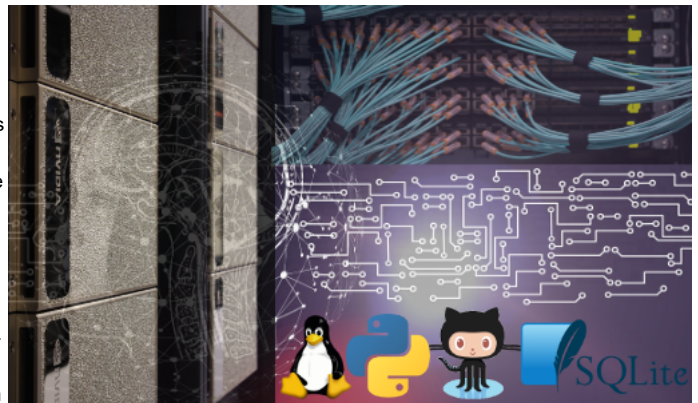
Course Description

The early 2000's were characterized as the era of *Big Data*, with researchers across disciplines finding research transformed by large volumes of diverse data. In the past decade, this has been again transformed by the re-emergence of artificial intelligence and machine learning systems aiding in interpretation of Big Data. As data types and volumes continue to grow, knowledge of scripting, database management, and advanced computing skills, including AI fundamentals, are critical for researchers regardless of discipline.

This course introduces students to the tools needed to be proficient, computationally-enabled researchers, providing a foundation in Linux and Bash scripting for data management, Python coding, basic SQL database fundamentals, and an introduction to artificial intelligence methods.

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 researchers and allows them to work in multidisciplinary fields.

Note: While the title of the course includes '*in Biology*' there is little that is actually focussed on biology. This course strives to be broadly applicable to many areas of research. Bureaucracies being what they are, the title is what it is...



Course Organization

The course is divided into four main sections:

- **Section 1:** Linux command line and Bash scripting; version control using Git and GitHub; using high-performance computing resources
- **Section 2:** Python scripting
- **Section 3:** SQL database introduction and integration with Python
- **Section 4:** A brief introduction to Artificial Intelligence

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 AI-assisted tools to generate and debug code.
- Use High Performance Computing resources such as the UFIT 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
- Have a basic understanding of artificial intelligence and gain hands-on experience with computer vision.

Meeting Times

Synchronous Meetings

- Mon, Wed, Fri, 9:35 AM - 10:25 AM (Period 3)
- Bartram Hall, room 211 and on Zoom

⚠ Important: You should make every effort to attend these sessions synchronously. While I will post most lectures ahead of time and will record these session, **this is the best opportunity to ask questions and get help from me and others in the class.**



- I understand that 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

- Students will work on projects either independently or in small, groups.
- This work 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.

Student Help Time

- Mondays from 11:00 am to noon, Physics (NPB 2334) or on Zoom
- Thursdays from 3:00 pm to 4:00 pm, Physics (NPB 2334) or on Zoom
- Fridays from 10:30 am to 11:30 am, UFII/UFBI Conference Room (ECE Building)
- By Appointment—please Email me (mailto:magitiz@ufl.edu?)



subject=Comp%20Tools%20Res%20office%20hours%20request) to setup a different time if needed.

☑ 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 do not ask for the help you need. 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. Because there are no textbook costs for this course, it has been recognized by Affordable UF as an affordable course.



Tip: Also check out the [additional resources page \(/resources.html\)](#).

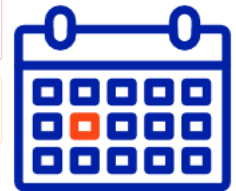
Course videos

Note: Most of the video content for this course was made in the Summer and Fall of 2020. The course was fully remote during the pandemic and these were mostly recorded in my living room. The quality is decent, but they are older and definately home made...That said, I think they may be useful to some, so I have opted to keep them as part of this site. Students should not feel that they need to watch all the videos, but may want to check them out for sections they miss or find more challenging.

Course Calendar

Warning: I am still working on updating things for Fall 2024. While the information below is a good guide, many pages are still being worked on.

Important: 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 (Week_01.html)	Fri, Aug 23	Download Software (/software.html)	Introduction and course objectives
2 (Week_02.html)	Mon, Aug 26	Take UFRC New User Training GitHub Account assignment due tomorrow (/github_account.html) Quiz 1 available, due Friday, September 06	Getting started: Computers Getting a GitHub.com account
2 (Week_02.html)	Wed, Aug 28	Read TLCL Introduction & Ch 1-4 (/TLCL_1.html)	Building shell skills
2 (Week_02.html)	Fri, Aug 30	Read TLCL Ch 5-8 (/TLCL_2.html) Problem Set 1 available, due Friday, September 13	Here are some exercises to work on
3 (Week_03.html)	Mon, Sep 02		Labor Day, no class
3 (Week_03.html)	Wed, Sep 04	Read Notes on Regular Expressions and TLCL Ch 19 (/TLCL_3.html) Quiz 2 available, due Friday, September 13	Text manipulation Regular Expressions Handout (/pdf/Regular_Expressions_Cheat_Sheet.pdf)
3 (Week_03.html)	Fri, Sep 06	Quiz 1 due Read TLCL Ch 20 (/TLCL_3.5.html) Optional: Watch Learn the Linux Command Line (/LinkedInLearningLinux.html) Read TLCL Ch 24-26 (/TLCL_4.html)	Shell Scripts and version control with git and GitHub

Week	Date	Reading/Assignment	Topic
4 (Week_04.html)	Mon, Sep 09		Continue working on shell scripts and git/github, Github Branching exercise (/github_branches.html)
4 (Week_04.html)	Wed, Sep 11	Problem Set 2 available, due Friday, September 20	Git and GitHub.com
4 (Week_04.html)	Fri, Sep 13	Quiz 2 due Problem Set 1 due Watch the HiPerGator: SLURM Submission Scripts 🔗 training	
5 (Week_05.html)	Mon, Sep 16	Read TLCL Ch 23 (/TLCL_6.html)	Compiling source code Google and Documentation Using UFIT Research Computing resources Running batch jobs
5 (Week_05.html)	Wed, Sep 18	Read Py4E Ch 1 (/py4e_1.html)	Introduction to Python
5 (Week_05.html)	Fri, Sep 20	Problem Set 2 due Quiz 3 available, due Friday, September 27 Read Py4E Ch 2 🔗	Python data types
6 (Week_06.html)	Mon, Sep 23	Problem Set 3 available, due Friday, October 04 Read Py4E Ch 3 🔗	Python: Flow Control
6 (Week_06.html)	Wed, Sep 25	Read Py4E Ch 4 🔗	Python: Functions
6 (Week_06.html)	Fri, Sep 27	Quiz 3 due Read Py4E Ch 5 🔗	Python: Iteration
7 (Week_07.html)	Mon, Sep 30	Read Py4E Ch 6 🔗 & Ch 7 🔗	Python: try/except, Strings, File I/O
7 (Week_07.html)	Wed, Oct 02	Read Ch 11 🔗	RegEx in Python
7 (Week_07.html)	Fri, Oct 04	Problem Set 3 due [Read Py4E Ch 12 🔗 & Ch 13 🔗	Scripting data acquisition
8 (Week_08.html)	Mon, Oct 07	SciPy, NumPy, Pandas 🔗	SciPy, NumPy, Pandas
8 (Week_08.html)	Wed, Oct 09		Pandas with Messy Data 🔗 Data visualization with Pandas 🔗
8 (Week_08.html)	Fri, Oct 11	Matplotlib and data visualization] (https://github.com/comptoolsres/Jupyter_content/blob/main/Data_visualization.ipynb)	Visualization

Week	Date	Reading/Assignment	Topic
9 (Week_09.html)	Mon, Oct 14	Problem Set 4 available, due Friday, November 01 Project 1, due Wednesday, October 30	More data visualization
9 (Week_09.html)	Wed, Oct 16		Matt at Conference, no class work on: TBD
9 (Week_09.html)	Fri, Oct 18		Homecoming, no class
10 (Week_10.html)	Mon, Oct 21	Quiz 4 available, due Monday, October 28 Scan Py4E Ch 16 Py4E Ch 14: Object oriented Programming	Py4E Ch 14: Object oriented programming ↗
10 (Week_10.html)	Wed, Oct 23		Work on Project 1
10 (Week_10.html)	Fri, Oct 25	Programming Foundations Databases (/LinkedInLearningDatabases.html)	Work on Project 1
11 (Week_11.html)	Mon, Oct 28	Quiz 4 due	Database intro (/Database_Introduction.html) Flight DB Example (/Convert_Flights_to_DB.html)
11 (Week_11.html)	Wed, Oct 30	Project 1 Due	Overview of databases Database design
11 (Week_11.html)	Fri, Nov 01	Read Py4E Ch. 15, through 15.5 and my notes ↗ Problem Set 4 Due Problem set 5 available, due Friday, November 15	Py4E Ch. 15, through 15.5 and my notes ↗ Databases, SQL and sqlite (/SQL_Introduction.html)
12 (Week_12.html)	Mon, Nov 04		Py4E Ch. 15, through 15.5 and my notes ↗ Databases, SQL and sqlite (/SQL_Introduction.html)
12 (Week_12.html)	Wed, Nov 06	Quiz 5 available, due Wednesday, November 13	More on databases and Joins ↗
12 (Week_12.html)	Fri, Nov 08	Project 2 available, due Wednesday, December 04	SQLAlchemy ↗
13 (Week_13.html)	Mon, Nov 11		Veteran's Day, no class
13 (Week_13.html)	Wed, Nov 13		SQLAlchemy and Pandas ↗
13 (Week_13.html)	Fri, Nov 15	Problem Set 5 due	Argparse (/Argparse.html)

Week	Date	Reading/Assignment	Topic
14 (Week_14.html)	Mon, Nov 18		Graphics (Graphics.html)
14 (Week_14.html)	Wed, Nov 20		Work on Project 2
14 (Week_14.html)	Fri, Nov 22	Quiz 5 due Quiz 6 available, due 2024-12-04	Intro to AI
15 (Week_15.html)	Mon, Nov 25		Thanksgiving, no class
15 (Week_15.html)	Wed, Nov 27		Thanksgiving, no class
15 (Week_15.html)	Fri, Nov 29		Thanksgiving, no class
16 (Week_16.html)	Mon, Dec 02		Intro to AI
16 (Week_16.html)	Wed, Dec 04	Project 2 due Quiz 6 due	Intro to AI

Software and Hardware

Participants will need a computer with internet connection 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.

A (free) Research Computing account will be created for students 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 Canvas, 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

See also the List of Graded Work page (Grading.html).

Assignment Values

- **Quizzes:** 6 @ 20 points each: 120 points (37%)

- **Problem Sets:** 5 @ 20 points each: 100 points (30%)
- **Class Projects:** 2 @ 40 points each: 80 points (24%)
- **Class Participation** 30 points: 30 points (9%)
 - 5 points for GitHub account assignment
 - 10 points for github commits
 - 15 points for code peer review

Grading in this class is consistent with UF policies available at: <https://catalog.ufl.edu/UGRD/academic-regulations/grades-grading-policies/>

Grade Disputes

Grading Scale and GPA Equivalent

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/>

In general, I do not take attendance. You are all adults and I assume you are taking the class to learn. The best way to learn is to regularly attend class. I am sure students will miss class for various reasons. I am happy to help you catch up. If you regularly miss class and fall behind, I may ask that you hold questions on content you have missed until after class, or ask that you coordinate a time to go over the content. I will make every effort to record and post all classes to help those that miss classes.

Quiz and Assignment Policy

Quiz and assignment dates will be announced at least one week in advance and students will have at least three days 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 one point 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.bluer.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 behave in a manner that is respectful to the instructor and to fellow students. Opinions held by other students should be respected in discussion, and conversations that do not contribute to the discussion should be held at a minimum, if at all.

Students should be working on course content during class.

GitHub Discussion Board

Canvas can be a challenge when working with code. As such we will use the GitHub discussion board [🔗](#) 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

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> [🔗](#)

Statement Regarding Course Recording

Our class sessions may be audio visually recorded for students in the class to refer back to and for use of enrolled students who are unable to attend live. Students who participate with their camera engaged or utilize a profile image are agreeing to have their video or image recorded. If you are unwilling to consent to have your profile or video image recorded, keep your camera off and do not use a profile image. Likewise, students who un-mute during class and participate verbally are agreeing to have their voices recorded. If you are unwilling to consent to have your voice recorded during class, you will need to keep your mute button activated and communicate exclusively using the "chat" feature, which allows students to type questions and comments live. The chat will not be recorded or shared. As in all courses, unauthorized recording and unauthorized sharing of recorded materials is prohibited.

Disclaimer on Free and Open Discussion

Students are encouraged to employ critical thinking and to rely on data and verifiable sources to interrogate all assigned readings and subject matter in this course as a way of determining whether they agree with their classmates and/or their instructor. No lesson is intended to espouse, promote, advance, inculcate, or compel a particular feeling, perception, viewpoint or belief.

Additional UF Policies and Resources

