BILD 62 | Introduction to Python for Biologists

Fall 2022

Instructor:
Johnatan (Yonatan) Aljadeff, aljadeff@ucsd.edu

Instructional Assistants:
Sriram Shreedharan, sshreedharan@ucsd.edu
Alexei Martsinkovskiy, amartsinkovskiy@ucsd.edu

Office hours
Monday at 4pm by zoom (link provided on Canvas)

Class Schedule
Lectures: Tuesday/Thursday, 9:30-10:50am in TATA 3201

Discussion Sections:
Wednesday, 2-2:50pm in HSS 2321
Wednesday, 3-3:50pm in HSS 2321
Wednesday, 4-4:50pm in HSS 2321

Course GitHub: http://github.com/BILD62

Course Description: Introductory class for biology students interested in using Python for data analysis and visualization. Course covers the basics of programming in Python and introduces students to various implementations of Python analyses for biological data such as time series and images.

By the end of this course, you will be able to:

- Read and run basic Python programs, recognizing the structures used (i.e. variables, conditionals, loops, functions) and explaining how they work
- Manipulate and create objects in Python, including data structures and classes
- Write, edit, and execute Python code in Jupyter Notebooks as well as the command line
- Visualize and run hypothesis-testing on simple datasets in Python
- Implement common algorithms for analyzing biological data (e.g., time series, images) and determine when such computations are appropriate

Grading

- **Assignments (50%)**: Weekly take-home coding assignments will support your progression through the course topics. Assignments will be submitted through the DataHub (http://datahub.ucsd.edu) and graded automatically using a tool called NBGrader.
    - All assignments are released by Monday at 8am, due **Friday at 8am**, and are worth 2.5-10% each.
    - These assignments should be completed individually and should take you about 1-2 hours.

- **Final Projects (30%)** Includes the project proposal, code, and deliverables. During finals week, we'll have a final project roundtable where you'll share your project with classmates and visitors.

- **Midterm Exam (20%)** About two thirds through our course, we will have an open note, open Python exam in which you will be asked to practice some of the fundamentals of Python and apply your knowledge of how to work with biological datasets.

Additional notes about grading:

- We will be using Canvas (http://canvas.ucsd.edu) to manage grades and assignments.
- **Late policy**: Assignments and projects will lose -10% for each day they are late.
- **Grading Scheme:** Final scores will be converted to letter grades, where A=100-90%, B=89-80%, C=79-70%, D=69-60%, and F=59-0%. For positive and minus grades, A+ = 97-100, A = 93-96.99, A- = 90-92.99, B+ = 87-89.99, B = 83-86.99, and so on.
- Final scores are as you see them on Canvas, once all of your assignments are graded. There is no rounding up to the closest score.

Course Resources

Computing Resources

You will need access to a computer and an internet connection for our course. It will also significantly help if you have a laptop (or an iPad + keyboard) for both lectures and discussions. If you need a laptop for the quarter, you can request a loaner laptop by filling out this form: https://eforms.ucsd.edu/view.php?id=490887.

Textbook

There is no official textbook for this course. However, we'll be relying on several online resources:

- Mansuri & Moshiri, **Intro to Computer Science & Python Stepik Course**
- VanderPlas, **Whirlwind Tour of Python**
- Software Carpentry, **Plotting and Programming in Python**
- Libeskind-Hadas & Bush, **Computing for Biologists**

Course Philosophy

A note on our course's environment

We'll be working together to create an equitable and inclusive environment of mutual respect, in which we all feel comfortable to share our moments of confusion, ask questions, and challenge our understanding. Everyone should be able to succeed in this course. If you do not feel that is the case please let the instructor know.

Course accommodations

If you need accommodations for this course due to a disability, please contact the Office for Students with Disabilities (osd@ucsd.edu) for an Authorization for Accommodation letter. Please speak with me in the first week of class if you intend to apply for accommodations. For more information, visit http://disabilities.ucsd.edu.

This course, and the work it entails, is for *you*

You won't benefit if others do your work. If you're unclear about what constitutes cheating in this course, please ask. Cases of academic dishonesty or cheating will be first handled by the course instructor, and then by the Academic Integrity Office. If you become aware of cheating in this class, you can anonymously report it.

We'll be relying a lot on other people's code as we learn. Here are some guidelines as to how you should use other code in the process of writing your own, as well as how you can talk to your classmates about the code we're working with in class:

Do explain the thought process behind your code.
Do share the general steps you took to solve a problem.
Do describe your code to others, either verbally or in writing.
Do use examples on the internet to inform your code.

Do not screenshot someone else's code.
Do not directly share your code with others, either in text or image format.
Do not directly copy 5+ lines of code from examples on the internet.
Do not share the values of variables that are explicitly asked for in the validation of the question.

Syllabus **(subject to change!)**

Date	Topic	Assignments
Week 0	Setting up and motivation for learning coding as a biology student	
Sep 22	**Welcome to BILD 62!** Introduction to the course, people & tools	
Week 1	To set the foundation for **this** course, we'll introduce the approaches and tools that we'll use throughout.	
Sep 27	**#2: Where Python lives, and how to talk to it**	Take the computing attitudes survey
Sep 29	**#3: Syntax & variable types**	*Due Friday 8am*: **a0** (2.5%)
Week 2	Fundamental coding skills in Python	
Oct 4	**#4: Data structures:** lists, tuples, and dictionaries	
Oct 6	**#5: Functions, booleans & conditionals**	*Due Friday 8am*: **a1** (5%)
Week 3	Fundamental coding skills in Python (continued)	
Oct 11	**#6: For Loops**	
Oct 13	**#7:** Object-oriented programming	*Due Friday 8am*: **a2** (7.5%)
Week 4	Scientific Computing	
Oct 18	**#8** NumPy introduction	
Oct 20	**#9** Using NumPy to load and analyze data	*Due Friday 8am*: **a3** (7.5%)
Week 5		
Oct 25	**#10:** Visualizing data	

Oct 27	**#11:** Data Analysis	*Due Friday 8am*: **a4** (7.5%)
Week 6		
Nov 1	**#12:** Data Analysis (continued)	
Nov 3	**#13:** Pandas	*Due Friday 8am*: **a5** (7.5%)
Week 7		
Nov 8	Review for midterm and Information for final projects	
Nov 10	*No Class This Day — Take Home Midterm*	**Note:** Take home midterm due at 6pm
Week 8		
Nov 15	**# 14:** Time series & signal processing Group formation for final projects	
Nov 17	**#15:** What counts as a cell? Image processing & cell detection in Python	*Due Friday 8am*: **a6** (7.5%)
Week 9		
Nov 22	**#16: Using Python to solve equations**	*Due Wednesday at 6pm*: **Project proposal**
Nov 24	Thanksgiving break	
Week 10	Wrapping up	
Nov 29	Documentation, version control, and collaborating on code	*Due Monday 8am*: **a7** (5%)
Dec 1	Next steps in bioinformatics, biological data science & computational approaches to big data (and time to work on final projects)	

Final Project Showcase: Dec 8, 8-11am