BILD 62 │ Introduction to Python for Biologists

SPRING 2022

Instructor Ashley Juavinett

Instructional Assistant Jiawei Shen

Office hours
Mondays at 1 pm (link on Canvas)

Class Schedule
Lectures: T/Th, 2-3:20 pm in TATA 2501
No official podcasting, but I will record screen + audio & share

Discussion Sections:
   Monday, 3-3:50 pm in U301 122
   Wednesday, 9-9:50 am in WLH 2206

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
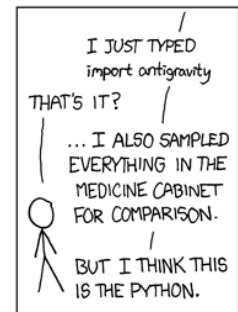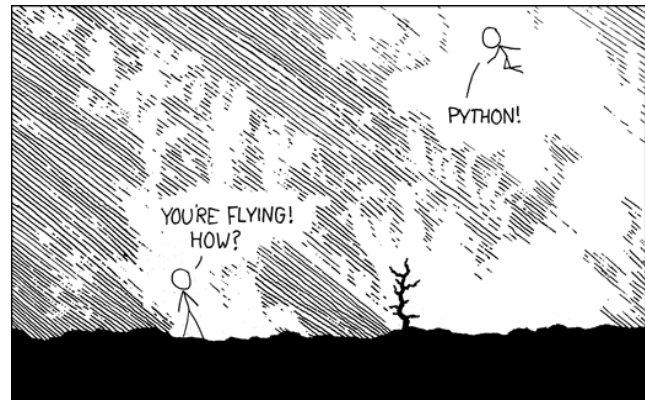
Grading

- **In-class work & participation (10%)**:
  - <u>Attendance</u> (5%): If you attend >75% of discussion sections (6 out of 8), you'll receive all 50 points.
  - <u>In-class work</u> (5%): We will complete several activities in class for credit. If you miss these, you'll need to contact the IA and make them up before Friday at 5 pm.

- **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 Thursday at 5 pm, due **Monday at 5 pm**, and are worth 2.5-10% each.
  - These assignments should be completed individually and should take you about 1-2 hours.

- **Final Projects (25%)** 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 (15%)** 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.
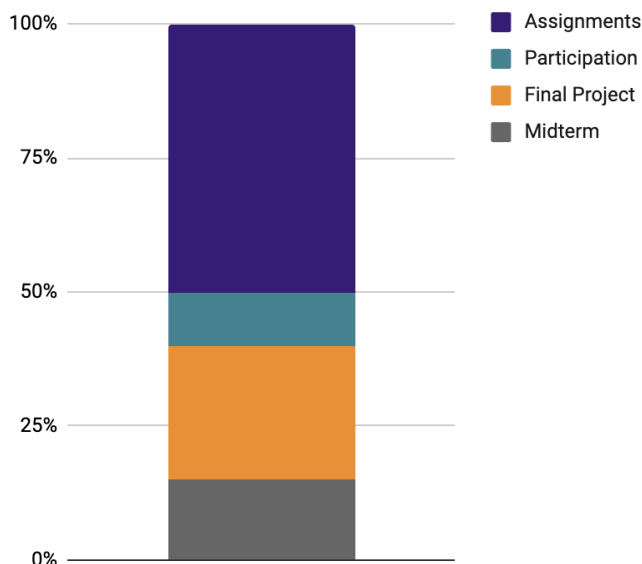
Additional notes about grading:

- 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. We have extra laptops in TATA 2501 if you need one — please come to class a few minutes early. We'll ask for your ID (student or drivers) in exchange for borrowing a laptop for the class period. 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**

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 me 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 me, 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	Before class
Week 1	To set the foundation for this course, we'll introduce the approaches and tools that we'll use throughout, as well as the motivation for learning how to code as a biology student.	
	Note: No discussion sections this week.	
Mar 29	**Welcome to BILD 62!** Introduction to the course, people & tools	Take the entry survey
Mar 31	**#2: Where Python lives, and how to talk to it**	Take the computing attitudes survey
Week 2	Fundamental coding skills in Python	
Apr 5	**#3: Syntax & variable types** *Note*: In-Class Activity, due Friday @ 5 pm	*Due Monday at 5 pm*: **a0**
Apr 7 **VIRTUAL**	**#4: Data structures:** lists, tuples, and dictionaries	
Week 3	Fundamental coding skills in Python (continued)	
	Discussion sections: Guessing Game Challenge	

Apr 12	**#5: Functions, booleans & conditionals**	*Due Monday at 5pm*: **a1**
Apr 14 **VIRTUAL**	**#6: Loops**	

Week 4

	Discussion sections: Fibonacci Sequence Task	
Apr 19	**#7:** Object-oriented programming	*Due Monday at 5pm*: **a2**
Apr 21 **VIRTUAL**	**#8:** Using NumPy to load and analyze data	

Week 5   Scientific Computing

Apr 26	**#9:** Pandas	*Due Monday at 5pm*: **a3**
Apr 28 **VIRTUAL**	**#10:** Simple Statistics in Python Descriptive, one-sample, and two-sample statistics	

Week 6   Statistics & Visualization

May 3	**#11:** Visualizing data Best practices for data visualization & implementation using Matplotlib	*Due Monday at 5pm*: a**4**
May 5 **VIRTUAL**	Review for midterm	

Week 7

May 10	**Midterm**	~~*Due Monday at 5pm*: **a5**~~
May 12 **VIRTUAL**	**Are those brain areas correlated?** Signal processing & correlations	

Week 8

May 17	Information for final projects	*Due Monday at 5pm*: **a6**
May 19 **VIRTUAL**	**What counts as a cell?** Image processing & cell detection in Python	

Week 9   _____ *- informatics*

May 24	Neuroinformatics: Using Python in literature searches	*Due Monday at 5pm*: **Project proposal**
May 26 **VIRTUAL**	**How many copies of a gene?** Bioinformatics in Python	

Week 10   Wrapping up

	Note: No discussion sections this week (Memorial Day)	
May 31	Documentation, version control, and catching bugs	*Due Monday at 5pm*: **a7**
June 2	Next steps in bioinformatics, biological data science & computational approaches to big data	^

Final Project Showcase: Tuesday, June 7th, 3-6 pm