Biosystems Analytics 434/534 (3 units)

Shantz 440; Monday and Wednesday 9:30-10:45AM

Description of Course

This course provides a comprehensive introduction to Python for data analytics. This course teaches students Python's basic data types and structures like strings, numbers, lists, tuples, dictionaries, and sets. In addition, the course focuses on testing code by writing small functions to develop larger, testable programs. As programs become complex, this "test-driven development strategy" ensures that code is robust and can be delivered to the community as cutting-edge open source packages. The course focuses on hands-on coding exercises to develop practical coding skills for interpreting and analyzing high-dimensional biological data. Students work in a collaborative learning classroom to participate in live coding exercises and gain skills in:

- basic Python data structures
- writing and testing functions and programs,
- using Python to read/write/parse files, and
- using regular expressions.

These skills are taught by implementing real-world coding examples to manipulate and process data in Python and test out your new skills through writing games in Python.

Course Prerequisites or Corequisites

An Apple, Linux, or Windows computer. Some experience programming Python is useful but not required.

Instructor and Contact Information

- Instructor: Ken Youens-Clark, MS, Adjunct Lecturer
- Office: None
- E-mail: kyclark@arizona.edu
- Office Hours: 8:30-9:30AM M/W in Shantz 440
- Course Websites: https://github.com/kyclark/be434-fall-2021

Course Objectives

The course objective is to teach formal, structured methods to write Python well. This course introduces techniques seasoned programmers use daily, from how to start writing a new program, to how to verify that it actually works. This course teaches students regular expressions, algorithm design, and random events using simple puzzles and games. Once students have mastered these coding concepts and practices, they can be applied broadly to advanced concepts in scientific computing and biological data. Expected Learning Outcomes This course is designed to provide students with a foundational understanding of manipulating data in Python. Upon completing this course students will be able to:

- Write command-line Python programs;
- Process a variety of command-line arguments, options, and flags;
- Write and run tests for your programs and functions;
- Manipulate of Python data structures including strings, lists, tuples, sets, dictionaries;
- Use higher-order functions like map and filter;
- Write and use regular expressions;
- Read, parse, and write various text formats;
- Use and control randomness;
- Use these skills to write challenging games in Python.

Graduate students will go beyond these basic learning outcomes to (1) apply basic coding principles to solve more advanced challenges, (2) develop optimized solutions to working with large-scale datasets and (3) evaluate when it is important to move from a single server solution to high-performance computing.

Course Format and Teaching Methods

The course is taught using a collaborative learning approach consisting of live-coding examples that demonstrate core concepts, an interactive group assignment, a collective discussion on the material, and an introduction to the homework. Importantly, because the development of scripting skills is essential to this class, programming homework is assigned weekly to reinforce the concepts taught in class. Each homework assignment is due a week after it was assigned. Students are quizzed on the material once a week on Tuesdays. At the end of the semester, students will implement code for a novel program via the class project. Undergraduates may choose from a set of simple games, whereas graduate students will program more complex games. Graduate students are also required to take a final examination comprised of small coding assignments based on concepts taught throughout the semester.

Absence and Class Participation Policy

Attending lectures and participating are vital to the learning process given the hands-on nature of the class. As such, attendance is required at all lectures. Students who miss class due to illness or emergency are required to bring documentation from their health-care provider or other relevant, professional third parties. Failure to submit third-party documentation will result in unexcused absences.

The UA's policy concerning Class Attendance, Participation, and Administrative Drops is available at http://catalog.arizona.edu/policy/class-attendance-participation-and-administrative-drop

The UA policy regarding absences for any sincerely held religious belief, observance or practice will be accommodated where reasonable, http://policy.arizona.edu/human-resources/religious-accommodation-policy.

Absences pre-approved by the UA Dean of Students (or Dean Designee) will be honored. See: https://deanofstudents.arizona.edu/absences

Makeup Policy for Students Who Register Late Students who register after the first class meeting may make up missed assignments/quizzes within the first two weeks of class.

Course Communications

Online communication, quizzes, and grading will be conducted through D2L. Course materials (reading, homework assignments, homework answers, and lectures) will be available on D2L and the course GitHub site (https://github.com/hurwitzlab/biosystems-analytics-2020) The GitHub privacy statement is located here: https://help.github.com/articles/github-privacy-statement.

Required Texts or Readings

Tiny Python Projects (http://tinypythonprojects.com/), Ken Youens-Clark, Manning, 2020 All texts are available free of charge, required sections are listed on the course schedule in D2L.

Optional Texts or Readings

- The Python Quick Book (3rd ed), Manning, 2018.
- Beyond the Basic Stuff with Python, Sweigart, No Starch, 2021
- Mastering Python for Bioinformatics, Youens-Clark, O'Reilly, 2021
- Beginning Python, Hetland, Apress, 2017.
- Python for Data Analysis 2nd edition, O'Reilly, 2017.

- Additional required readings will be listed weekly on the course schedule and D2L.

Required or Special Materials

Students will be required to create a (free) account on GitHub to download and turn in assignments. Students must install Python on their laptops for completing and turning in assignments. Windows users will need to install Windows Subsystem for Linux (WSL) in order to have a functional command-line interface. Students will also need to install an editor. In this class, we use Visual Studio Code (https://code.visualstudio.com/), but students are welcome to use a different editor of their choice. Information on accessing these resources can be found under Getting Started in D2L. We walk-through these resources in-class during the first week of the course.

Required Extracurricular, Activities

Optional software carpentry workshops and events related to Python data analytics will be announced in D2L.

Grading Scale and Policies

The final letter grades for the class are based on the TOTAL NUMBER OF POINTS that each student accumulates for the following assessments (see details for each section below):

Undergraduate Students:

Category	# Assessments	Points per assessment	Points total
Homework (48%)	12	40	480
Quizzes (36%)	12	30	360
Project (16%)	1	160	160
Total			1000

Graduate Students:

Category	# Assessments	Points per assessment	Points total

Homework (36%)	12	30	360
Quizzes (30%)	12	25	300
Project (14%)	1	140	140
Final exam (20%)	1	200	200
Total			1000

Grading Scale:

Grade	A	B	C	D	F
Points	900 or more	800–899	700–799	600–699	less than 600

University policy regarding grades and grading systems is available at http://catalog.arizona.edu/policy/grades-and-grading-system.

Requests for incomplete (I) or withdrawal (W) must be made in accordance with University policies, which are available at http://catalog.arizona.edu/policy/grades-and-grading-system#incomplete and http://catalog.arizona.edu/policy/grades-and-grading-system#Withdrawal respectively.

Dispute of Grade Policy: Disputes on a grade for an assignment, quiz, or exam must be made within three days of when the grade is posted.

Assignments and Examinations

All assignments, quizzes, and examinations are listed on the course GitHub site https://github.com/hurwitzlab/biosystems-analytics-2020 and linked in D2L.

Homework is assigned every Thursday and is due the following Thursday. The homework provides you with practice in scripting and analytics. Homework is turned in by committing your code to your GitHub repository that is shared with the instructor. At the beginning of class on Thursday, I will download your code from GitHub. Your grade will be the percentage of tests you pass in the provided test suite.. We encourage group work but be sure the code you turn in has been independently written (see Code of Academic Integrity below).

Quizzes

Each week you will be tested on the reading for the week ahead (see the schedule in D2L). This will prepare you for the lectures and in-class work. Each of the quizzes will be available on D2L during the first 10 minutes of class on Tuesdays. Answers for each quiz are reviewed in class and posted on the class D2L website.

Final Project

At the end of the semester, students will implement a novel program in Python via the class project that demonstrates the core programming concepts for the class. The project will include documentation (a README), a program, a test suite, and any required input/data to run the program. The program may be a game, a solution to a problem from something like Rosalind.info, or some program needed by your lab or PI to further your research. Graduate students will be expected to create more detailed and involved programs than undergraduates. Final Examination The final exam (for graduate students only) will be drawn from a sample of small coding assignments to be completed during the final using programming concepts in Python discussed throughout the semester. The core concepts will be highlighted each week as expected learning outcomes and given as practice in homework assignments. The final is an open note and open book exam.

Final Exam Schedule: http://www.registrar.arizona.edu/schedules/finals.htm

Final Exam Regulations: https://www.registrar.arizona.edu/courses/final-examination-regulations-and-information

Scheduled Topics/Activities

TBD

Honors Credit

Students wishing to contract this course for Honors Credit should email me to set up an appointment to discuss the terms of the contract. Information on Honors Contracts can be found at http://www.honors.arizona.edu/faculty-and-advisors/contracts.

Classroom Behavior Policy

To foster a positive learning environment, students and instructors have a shared responsibility. We want a safe, welcoming, and inclusive environment where all of us feel comfortable with each other and where we can challenge ourselves to succeed. To that

end, our focus is on the tasks at hand and not on extraneous activities (e.g., texting, chatting, reading a newspaper, making phone calls, web surfing, etc.). Threatening Behavior Policy The UA Threatening Behavior by Students Policy prohibits threats of physical harm to any member of the University community, including to oneself. See http://policy.arizona.edu/education-and-student-affairs/threatening-behavior-students.

Accessibility and Accommodations

At the University of Arizona, we strive to make learning experiences as accessible as possible. If you anticipate or experience physical or academic barriers based on disability or pregnancy, you are welcome to let me know so that we can discuss options. You are also encouraged to contact Disability Resources (520) 621-3268 to explore reasonable accommodation.

If our class meets at a campus location: Please be aware that the accessible table and chairs in this room should remain available for students who find that standard classroom seating is not usable.

Code of Academic Integrity

Students are encouraged to share intellectual views and discuss freely the principles and applications of course materials. However, graded work/exercises must be the product of independent effort unless otherwise instructed. Students are expected to adhere to the UA Code of Academic Integrity as described in the UA General Catalog. See http://deanofstudents.arizona.edu/academic-integrity/students/academic-integrity.

The University Libraries have some excellent tips for avoiding plagiarism, available at http://new.library.arizona.edu/research/citing/plagiarism. Selling class notes and/or other course materials to other students or to a third party for resale is not permitted without the instructor's express written consent. Violations to this and other course rules are subject to the Code of Academic Integrity and may result in course sanctions. Additionally, students who use D2L or UA e-mail to sell or buy these copyrighted materials are subject to Code of Conduct Violations for misuse of student e-mail addresses. This conduct may also constitute copyright infringement.

UA Nondiscrimination and Anti-harassment Policy

The University is committed to creating and maintaining an environment free of discrimination; see http://policy.arizona.edu/human-resources/nondiscrimination-and-anti-harassment-policy Our classroom is a place where everyone is encouraged to express well-formed opinions and their reasons for those opinions. We also want to create a tolerant and open

environment where such opinions can be expressed without resorting to bullying or discrimination of others.

Additional Resources for Students

UA Academic policies and procedures are available at http://catalog.arizona.edu/policies Student Assistance and Advocacy information is available at http://deanofstudents.arizona.edu/student-assistance/students/student-assistance

Confidentiality of Student Records

http://www.registrar.arizona.edu/personal-information/family-educational-rights-and-privacy-act-1974-ferpa?topic=ferpa

Subject to Change Statement

The information contained in the course syllabus, other than the grade and absence policy, may be subject to change with advance notice, as deemed appropriate by the instructor.