

USC VITERBI SCHOOL OF ENGINEERING DATA SCIENCE PROGRAM

DSCI 510: Principles of Programming for Data Science (4 units) Spring 2024

Instructor:

Prof. Jose-Luis Ambite ambite@isi.edu

Teaching Assistant /Course Producers:

Jehil Yogesh Vora jehilyog@usc.edu (TA)

Masooma Naqvi syedamas@usc.edu

Omkar Metri metri@usc.edu

Karthik Kancharla kkanchar@usc.edu

When emailing, please put “DSCI510” in the subject line

Lecture

Day: Wednesday

Time: 4:00pm–5:50pm

Room: RTH 109

Lab

Day: Friday

Time: 4:00pm–5:50pm

Room: RTH 109

Platform: DEN D2L (not blackboard!): <https://courses.uscden.net/d2l/home/<forthcoming>>
Zoom links available in D2L class page.

Textbook and other resources

Required: [Python for Everybody: Exploring Data in Python 3](#), by Charles R. Severance, plus other materials. Free online.

Recommended: You may also wish to consult [Think Python \(2nd edition\)](#), by Allen B. Downey. Free online. For more advanced Python, you could consult [Fluent Python, 2nd Edition](#), by Luciano Ramalho. Free at the USC Libraries. There are many other free materials on Python on the USC Libraries and on the web, for example, the many O’Reilly Books, and pythontutor.com

Instructor’s office hours: After class, or by appointment.

Although this is not required, students are advised to make appointments with the instructor ahead of time in any event and be specific with the subject matter to be discussed. Students should also be prepared for their appointment by bringing all applicable materials and information.

Catalogue Description:

Introductory programming course for non-Computer Science majors. Programming in Python for retrieving, searching, and analyzing data from the Web. Learning to manipulate large data sets.

Expanded Course Description:

This course is designed to serve as an introduction to computer science concepts and basic programming skills that are specifically geared toward Informatics and Data Science. The main objective of this course is to introduce the fundamental concepts behind general computer science and programming, and to give students practical hands-on experience reading and writing computer programs, in order to give them the tools to manipulate large data sets.

The course is designed to be accessible to non-Computer Science major students with little or no programming experience, and emphasizes writing programs that are capable of retrieving and manipulating large amount of data. The first half of the course focuses on Python as a first programming language, while the second half of the course covers selected advanced topics including web scraping, database access/SQL, data manipulation, data visualization, and more.

The course will combine lectures, labs, in-class discussion and problem-solving, readings, programming assignments, a mid-term exam, and a final project.

Course Objectives:

The objective of this course is to train students to write computer code capable of manipulating large data sets. Specifically, students successfully completing this course will achieve two main objectives:

1. Acquire basic concepts in computer science and programming.
2. Develop sufficient proficiency in Python to write applications capable of retrieving, searching, manipulating, analyzing, and displaying data.

Methods of Teaching:

The primary teaching methods will be discussion, programming assignments, and lectures. Students are expected to perform directed self-learning outside of class which encompasses, among other things, a considerable amount of programming practice.

There will be a midterm exam. There will **not** be a final exam, but there will be a final project. There will be weekly lab meetings and short programming assignments.

Students are expected to have access to a computer that can run the Python programming language and a web browser, and to bring it to class. Beyond that, no special computing facility, hardware or software will be necessary for this course.

Grading Scheme (still subject to change):

- Weekly Lab Assignments: 39% (13 * 3%)
- Mid-Term: 25%
- Final Project 36% (1st Deliverable: 4%; 2nd Deliverable: 7%; 3rd Deliverable: 25%)

Grades will range from A through F. The following is the breakdown for grading:

94 - 100 = A	74 - 76 = C
90 - 93 = A -	70 - 73 = C-
87 - 89 = B+	67 - 69 = D+
84 - 86 = B	64 - 66 = D
80 - 83 = B-	60 - 63 = D-
77 - 79 = C+	Below 60 is an F

Grading for the course will be based on three major components:

1. LAB ASSIGNMENTS – The labs will reinforce the content discussed in the lecture. During the labs some programming problems will be solved. Then, some similar problems will be assigned to be solved either during the lab or in the next few days. There will be laboratory assignments given every week. These assignments are to be completed individually.

2. MID-TERM EXAM – A written exam, roughly halfway through the term, covering material taught up to that point, which is expected to be core Python.

3. FINAL PROJECT – The final project will be a programming assignment that addresses a data science task using the techniques taught in class. The project should integrate and analyze three related data sources from the web. There will be three deliverables: 1) a brief description of the proposed project; 2) code and data scraped from one source; and 3) final code and a short paper on the project and its analysis. Further details on the project will be provided.

Class Communication:

USC DEN Desire2Learn (D2L) platform and Zoom. Classes will be recorded.

Class Structure & Schedule:

Class sequence, dates, topics and guest speakers are subject to change as the semester proceeds. Any revisions will be noted and announced in class in advance.

Week	Topics/Activities	Book Chapters
1 (1/10)	Intro to programs, variables, types, expressions.	1, 2
2 (1/17)	Conditional execution. Functions.	3, 4
3 (1/24)	Iteration. Strings.	5, 6
4 (1/31)	Files. Exceptions. Lists.	7,8
5 (2/7)	Dictionaries, Tuples, Sets.	9, 10
6 (2/14)	Object-Oriented Programming. Python objects, inheritance, operator overloading, creating multiple/dynamic objects.	14
7 (2/21)	Function Params: *args, **kargs. Scope, Modules, Packages, Libraries.	14
8 (2/28)	MIDTERM.	
9 (3/6)	Networked Programs. HTML DOM, requests library, web scraping with BeautifulSoup.	12
10	SPRING BREAK	
11 (3/20)	XML, JSON, web APIs. <i>Project Proposal due March 25</i>	13
12 (3/27)	Regular Expressions.	15
13 (4/3)	Intro to databases, DB design and normalization, SQL queries.	15
14 (4/10)	Data manipulation with Pandas and Numpy. <i>Project Submission 2, due April 16</i>	
15 (4/17)	Data visualization with Matplotlib and Seaborn.	
16 (4/24)	More advanced Python, Databases, SQL, OR-Mapping <i>Final Project due April 30</i>	11

Statement on Academic Conduct and Support Systems

Academic Conduct

Plagiarism

Presenting someone else’s ideas as your own, either verbatim or recast in your own words – is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in *SCampus* in Section 11, *Behavior Violating University Standards* <https://scampus.usc.edu/1100-behavior-violating-university-standards-and-appropriate-sanctions>. Other forms of academic dishonesty are equally unacceptable. See additional

information in *SCampus* and university policies on scientific misconduct, <http://policy.usc.edu/scientific-misconduct>.

Use of AI/LLMs

Use of AI and specifically Large Language Models (LLMs) is allowed. However, it is only allowed as a tool to assist in learning. That is, you may use AI models such as ChatGPT or Claude 2 to help understand the assignments, to ask generic questions about programming in Python and help you by generating code samples that could be of use to explain how certain programming constructs work. Submitting assignments completely generated by AI is strictly prohibited, and when discovered, you will be awarded 0 points for the assignment. We will be utilizing additional software to check for code generated by an AI.

Support Systems

Counseling and Mental Health - (213) 740-9355 – 24/7 on call

<https://sites.usc.edu/counselingandmentalhealth/>

Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

National Suicide Prevention Lifeline - 1 (800) 273-8255 – 24/7 on call

<https://988lifeline.org/>

Free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week.

Relationship and Sexual Violence Prevention and Services (RSVP) - (213) 740-9355 (WELL) – 24/7 on call

<https://sites.usc.edu/clientservices/>

Free and confidential therapy services, workshops, and training for situations related to gender-based harm.

Office of Equity and Diversity (OED)- (213) 740-5086 | Title IX – (213) 821-8298

<https://eeotix.usc.edu/>

Information about how to get help or help someone affected by harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants. The university prohibits discrimination or harassment based on the following protected characteristics: race, color, national origin, ancestry, religion, sex, gender, gender identity, gender expression, sexual orientation, age, physical disability, medical condition, mental disability, marital status, pregnancy, veteran status, genetic information, and any other characteristic which may be specified in applicable laws and governmental regulations. The university also prohibits sexual assault, non-consensual sexual contact, sexual misconduct, intimate partner violence, stalking, malicious dissuasion, retaliation, and violation of interim measures.

Reporting Incidents of Bias or Harassment - (213) 740-5086 or (213) 821-8298

<https://campussupport.usc.edu/trojans-care-4-trojans/>

Avenue to report incidents of bias, hate crimes, and microaggressions to the Office of Equity and Diversity | Title IX for appropriate investigation, supportive measures, and response.

Office of Student Accessibility Services (OSAS) - (213) 740-0776

<https://osas.usc.edu/>

Support and accommodations for students with disabilities. Services include assistance in providing readers/note-takers/interpreters, special accommodations for test taking needs, assistance with architectural barriers, assistive technology, and support for individual needs.

USC Support and Advocacy - (213) 821-4710

<https://campussupport.usc.edu/>

Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

Diversity at USC - (213) 740-2101

<https://diversity.usc.edu/>

Information on events, programs and training, the Provost's Diversity and Inclusion Council, Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

USC Emergency - UPC: (213) 740-4321, HSC: (323) 442-1000 – 24/7 on call

<https://www.usc.edu/emergency/>

Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

USC Department of Public Safety - UPC: (213) 740-6000; HSC: (323) 442-1200 – 24/7 on call

<https://dps.usc.edu/>

Non-emergency assistance and safety information.