

DSCI 549: Introduction to Computational Thinking and Data Science

32430D — Fall 2023

Instructors Information

Name: Deborah Khider

Email: khider@usc.edu

Office Hours: By Appointment Only Tu 2-5pm; W 3-5pm; 15min

Graders Information

Name: TBA

Email: TBA

Name: TBA

Email: TBA

Class Information

Dates: 8/24/2023 – 11/30/2023

Time and classroom: Th 3:30pm – 6:50pm PT, SOS B2

All class communications, materials dissemination, and announcements will occur via Blackboard.

Course Description

Introduction to data analysis techniques and associated computing concepts for non-programmers. Topics include foundations for data analysis, visualization, parallel processing, metadata, provenance, and data stewardship.

Expanded Course Description

This course will teach non-programmers to think in computing terms about modern topics, and to approach real-world phenomena through data science. The course is intended for students in disciplines outside of computer science, so no prior experience with computer science is assumed. The course topics will be particularly relevant to students interested in physical, biological, and social sciences. In particular, the course will cover different kinds of data and corresponding approaches to data analysis, including geospatial data, time series, networks, and multimedia data.

Learning Objectives and Outcomes

Students learn to run multi-step analysis through a graphical workflow interface and will experience firsthand complex concepts in data science such as parallel computing, provenance, and visualization. Students also learn to use ontologies and logic representations to capture metadata and other knowledge about complex data. The course includes practical lessons to use workflow and ontology development toolkits, as well as best practices for data stewardship and dissemination.

In particular, the course will enable students to:

- Acquire computational thinking skills that will enable students to represent and reason about complex problems in the digital arena
- Understand different kinds of data in terms of their possibilities and limitations to approach complex problems cast in terms of the emerging field of data science
- Familiarize with data science scholars' best practices in data documentation and dissemination

Prerequisite(s):

None

Co-requisite(s):

None

Recommended preparation:

Mathematics, Statistics, and Logic undergraduate course.

Textbooks and Software

There is no textbook. All required software is freely available for students to install on their personal computers or to access through a web interface.

Lectures

Lecture materials will be pre-recorded and made available at least one week prior to class. Students are expected to listen to the material before the Thursday class and complete a quiz based on these materials. Class time will be used for hands-on practicums.

Quizzes

Students are expected to listen to the materials and take a quiz on Blackboard prior coming to class. The quiz must be completed by **3:30pm on Thursday**. Students will have one week to complete the quiz. **No make-up quizzes will be given with the exception of late registration, providing documentation for the late registration.** There will be nine quizzes throughout the semester.

Examinations

Students will be required to take three, in-class, comprehensive examinations throughout the semester. Each test will be two-hour long and cumulative. The tests will consist of multiple-choice, multiple-answers, matching, fill-in-the-blank, and short answer questions.

Students will not be allowed to use class notes or electronic devices during the tests. However, students will be allowed to bring a 8.5"x11" one-sided, handwritten "cheat sheet" to help them with the test. The cheat sheet will have to be turned in with the test. Missing sheets will result in a **5-point deduction**.

In case of a **documented** emergency (family emergency and illness), students will have **4 business days** to contact the professor to make up the test. Test make-up will be held **in-person at the Information Sciences Institute in Marina Del Rey**. No make-up will be allowed after the answers have been given in class (see schedule), which allows for 2-3 weeks to reschedule the test. The exception to this rule is for the last test, which will coincide with the final examination time for

the course. In case of a **documented** emergency (family emergency and illness), students will be required to receive an Incomplete for the course and schedule their final examination in the first week of January.

In case of a religious conflict with any of the tests, students **must** notify the professor at least two weeks prior to the test and make arrangements to schedule the test **in-person at the Information Sciences Institute in Marina Del Rey**.

Grades will be posted within 1 week of the examination. **However**, answers to individual test will be covered in class after everyone has taken the examination. Scheduling office hours to go over the test before then will result in a **0.5 point penalty**. Grade argumentation may be subjected to a **0.25 point penalty** per instance.

Use of AI Generators

In this course, you may use AI Generators for the quizzes and in-class practicums. In fact, several of these practicums will be using this technology so you can learn to use them as intended and understand the pitfalls associated with them. However, it is also important to learn the basics of data science (to also be able to use AI Generators more effectively). Therefore, the use of AI Generators will not be allowed during in class examinations.

Grading

The course grade is determined by the following components:

Quizzes	10 points (9x1.5 points to a maximum of 10 points)
Exams	90 points (3x30 points)

Grade Scale

Final grades will be assigned according to the following scale:

A	94–100	C+	77–79.9
A–	90 – 94.9	C	74–76.9
B+	87–89.9	C–	70–73.9
B	84–86.9	D	64–69.9
B–	80–83.9	F	0 – 63.9

No rounding up will be applied. No curves will be applied. Extra-credit (if any) will be incorporated as individual questions in exams and will not exceed **4 points**.

Schedule

Lecture materials will be pre-recorded and made available on Blackboard at least one week prior to class. Class time will be dedicated to practicums. Come to class with a laptop. If software needs to be installed, an email will be sent prior to the class.

Please note that the class will neither be streamed on Zoom nor recorded. However, answer keys will be made available on Blackboard at the end of the class.

	Lecture	In class Practicum	Exam Re-views	Quiz
Aug 24		Syllabus - Into to Data Science - Jupyter Notebooks		
Aug 31	Data/Data Privacy/ Software	Finding Data/Data Exploration		Q1
Sep 7	Workflow/Parallel and Distributed Computing	Defining data science problems as computational workflows		Q2
Sep 14	Experimental Design/ Probability/ Inferential Statistics	Obervational vs Experimental Studies/Understanding Probabilities/t-test/ANOVA/Solving with simulations		Q3
Sep 21	A/B Testing	A/B testing in UI/UX	Test 1 Review	Q4
Sep 28		TEST 1	Test 1 Review	
Oct 5	Machine Learning	Training a classifier/clustering algorithm		Q5
Oct 12		<i>Fall Recess</i>		
Oct 19	Natural Language Processing and Multimedia	Creating a spam classifier/Topic Modeling/Text Generation/Object Recognition	Test 1 Answers	Q6
Oct 26	Network and Timeseries	Time series forecasting/Causal network	Test 2 Review	Q7
Nov 2		TEST 2	Test 2 Review	
Nov 9	Semantic Metadata and Ontology	Protege/Designing an Ontology/Working with the Music Ontology		Q8
Nov 16	Provenance and Data Stewardship	Intro to GitHub, Figshare, and Zenodo/Using PROV-O to report a study		Q9
Nov 23		<i>Thanksgiving</i>		
Nov 30			Test 1&2 Answers + Review	

Attendance

Students are expected to attend in-class practicums. In case of absence, students are responsible for obtaining the materials from their classmates.

Please, **DO NOT** attend class if you are feeling sick or have received a positive COVID-19 test as mandated by University policy.

Office Hours

Office hours will be held virtually by appointment only Tu 2-5pm; Th 3-5pm. Students should be using [calendly](#) for scheduling a 15-min session 24hours in advance.

Please come prepared with specific questions about the materials covered in the pre-recorded lectures and in-class exercises. No individual answers for the tests (including the ability to see the test) will be given until after we have covered the answers in the class.

Communication

This content is protected and may not be shared, uploaded, or distributed.

There will be no Slack channel for this class. Emails should be limited to administrative questions and requests for make-up under the conditions explained above. Use the calendly link to schedule office hours.

Academic Integrity

The University of Southern California is foremost a learning community committed to fostering successful scholars and researchers dedicated to the pursuit of knowledge and the transmission of ideas. Academic misconduct is in contrast to the university's mission to educate students through a broad array of first-rank academic, professional, and extracurricular programs and includes any act of dishonesty in the submission of academic work (either in draft or final form).

This course will follow the expectations for academic integrity as stated in the USC Student Handbook. All students are expected to submit assignments that are original work and prepared specifically for the course/section in this academic term. You may not submit work written by others or "recycle" work prepared for other courses without obtaining written permission from the instructor(s). Students suspected of engaging in academic misconduct will be reported to the Office of Academic Integrity.

Other violations of academic misconduct include, but are not limited to, cheating, plagiarism, fabrication (e.g., falsifying data), knowingly assisting others in acts of academic dishonesty, and any act that gains or is intended to gain an unfair academic advantage.

The impact of academic dishonesty is far-reaching and is considered a serious offense against the university and could result in outcomes such as failure on the assignment, failure in the course, suspension, or even expulsion from the university.

For more information about academic integrity see [the student handbook](#) or the [Office of Academic Integrity's website](#), and university policies on [Research and Scholarship Misconduct](#).

If found responsible for an academic violation, students may be assigned university outcomes, such as suspension or expulsion from the university, and grade penalties, such as an "F" grade on the assignment, exam, and/or in the course.

Course Content Distribution

USC has policies that prohibit recording and distribution of any synchronous and asynchronous course content outside of the learning environment.

Recording a university class without the express permission of the instructor and announcement to the class, or unless conducted pursuant to an Office of Student Accessibility Services (OSAS) accommodation. Recording can inhibit free discussion in the future, and thus infringe on the academic freedom of other students as well as the instructor. ([Living our Unifying Values: The USC Student Handbook](#), page 13).

Distribution or use of notes, recordings, exams, or other intellectual property, based on university classes or lectures without the express permission of the instructor for purposes other than individual or group study. This includes but is not limited to providing materials for distribution by services publishing course materials. This restriction on unauthorized use also applies to all information, which had been distributed to students or in any way had been displayed for use in relationship to the class, whether obtained in class, via email, on the internet, or via any other

media. ([Living our Unifying Values: The USC Student Handbook, page 13](#)).

Students and Disability Accommodations

USC welcomes students with disabilities into all of the University's educational programs. The [Office of Student Accessibility Services \(OSAS\)](#) is responsible for the determination of appropriate accommodations for students who encounter disability-related barriers. Once a student has completed the OSAS process (registration, initial appointment, and submitted documentation) and accommodations are determined to be reasonable and appropriate, a Letter of Accommodation (LOA) will be available to generate for each course. The LOA must be given to each course instructor by the student and followed up with a discussion. This should be done as early in the semester as possible as accommodations are not retroactive. More information can be found at osas.usc.edu. You may contact OSAS at (213) 740-0776 or via email at osasfrontdesk@usc.edu.

Support Systems

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

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

[988 Suicide and Crisis Lifeline](#) - 988 for both calls and text messages – 24/7 on call

The 988 Suicide and Crisis Lifeline (formerly known as the National Suicide Prevention Lifeline) provides free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week, across the United States. The Lifeline is comprised of a national network of over 200 local crisis centers, combining custom local care and resources with national standards and best practices. The new, shorter phone number makes it easier for people to remember and access mental health crisis services (though the previous 1 (800) 273-8255 number will continue to function indefinitely) and represents a continued commitment to those in crisis.

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

Free and confidential therapy services, workshops, and training for situations related to gender- and power-based harm (including sexual assault, intimate partner violence, and stalking).

[Office for Equity, Equal Opportunity, and Title IX \(EEO-TIX\)](#) - (213) 740-5086

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.

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

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

[The Office of Student Accessibility Services \(OSAS\)](#) - (213) 740-0776

OSAS ensures equal access for students with disabilities through providing academic accommoda-

tions and auxiliary aids in accordance with federal laws and university policy.

[USC Campus Support and Intervention](#) - (213) 740-0411

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

[Diversity, Equity and Inclusion](#) - (213) 740-2101

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

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

Non-emergency assistance or information.

[Office of the Ombuds](#) - (213) 821-9556 (UPC) / (323-442-0382 (HSC)

A safe and confidential place to share your USC-related issues with a University Ombuds who will work with you to explore options or paths to manage your concern.

[Occupational Therapy Faculty Practice](#) - (323) 442-2850 or otfp@med.usc.edu

Confidential Lifestyle Redesign services for USC students to support health promoting habits and routines that enhance quality of life and academic performance.