



Syllabus

Syllabus

COS422/522 : Computing for Data Science Fall 2022

Professor: Bruce MacLeod (www.cs.usm.maine.edu/~macleod)

Room 222 Science Bldg.

email : macleod at maine.edu

Tel: 780-4285

Office Hours: Mon 9:30-10:45, Wed 12:30-1:30 or by appointment

(Please let me know & we can arrange a meeting) Zoom link :

[https://maine.zoom.us/j/87841802479?](https://maine.zoom.us/j/87841802479?pwd=dEVsQXJUmIDK0ptL2dyK21OTWNqUT09)

[pwd=dEVsQXJUmIDK0ptL2dyK21OTWNqUT09](https://maine.zoom.us/j/87841802479?pwd=dEVsQXJUmIDK0ptL2dyK21OTWNqUT09)

COS 422/522 Grader/Tutor:

Nickolas Littlefield

email: nickolas.littlefield at maine.edu

Hours

Monday: 3 - 4pm

Wednesday: 2-4pm

Thursday: 2 - 4pm

Zoom link : [https://maine.zoom.us/j/88618763015?pwd=](https://maine.zoom.us/j/88618763015?pwd=YjNqcjA5eDRXaVVrbCtwa1lIUeJ0dz09)

[YjNqcjA5eDRXaVVrbCtwa1lIUeJ0dz09](https://maine.zoom.us/j/88618763015?pwd=YjNqcjA5eDRXaVVrbCtwa1lIUeJ0dz09)

Course Objectives:

This class provides a practical introduction to the data science workflow using Python. Students will demonstrate their expertise in:

- using advanced features of python including classes, sequences, maps and lambda functions;
- working with the the numpy library for algebraic computation and pandas libraries for data storage and retrieval;
- retrieving information in data files and prepare it for analysis and visualization;
- constructing meaningful graphs and plots to visualize information;
- identifying ethical issues in Data Science;
- ability to work with data in different forms : numbers, text, images, and sound
- and completing an end to end data science project of their choosing.

All of these outcomes will be evaluated during the course.

Prerequisites: At least one introduction to Computer Science course (for example, COS 160) and at least one Calculus or Statistics course.

Textbook:

Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython, 2nd or 3rd Edition, Wes McKinney (Open access 3rd edition : <https://wesmckinney.com/book>)

Grading

- Class participation, quizzes, readings : (10%)
- Labs (10%)
- Four Programming Assignments (40%)

- Midterm (10%)
- Second Exam (10%)
- Final Project (20%)

Graduate students will have additional exercises in each Programming Assignment and the expectations for the final project will include background reading and a written summary.

Preparation before each class:

The readings for each lecture are listed on the syllabus. You are expected to do the readings before the lecture. Quizzes will evaluate your understanding of the readings.

Teamwork on Programs

It is acceptable to work with other students in the analysis, design, and debugging phases of your programming assignments. However, unless it is a team project, do not write code as a team. Do not copy your code from or for another student. Please list the students with whom you collaborated, and the type of help you gave or received in a comment at the top of your program. Plagiarism will result in disciplinary action that may involve failure of the course.

Good Data science requires focus and clear, logical, problem solving. In this context, multitasking does not work and there is considerable body of research to support this (starting links from a spectrum of sources: Multitasking damages your brain and your career, Multitasking does not work, Modern World is bad for the Brain, The Distracted Mind) When you work on the homework and projects, consider creating an environment with minimal distraction. You may find things get done faster and easier to understand.

Reading, assignment, and exam schedule

Week 1 : (8/29, 8/31)

- Introduction to Python, Jupyter Notebooks
 - Lecture 1 : Intro to course, Data Science
 - Introduction Lab (Wed, 9:30-10:20):
 - **Use the lab zoom link** : Intro to Colab, Data Science Workflow
 - Lecture 2 :
 - Reading : Chapter 2
 - Breakout room exercises
 - post answers to Slack (sign up if you have not done so already)

Week 2: (9/7 wed only)

- Python types & sequences
 - Lecture 3 :
 - reading Chapter 3.1.3.2
 - Overview of Chapter 3, breakout room exercises
- Types and Sequences Lab

Week 3: (9/12, 9/14)

- Python types & sequences
 - Lecture 4 :
 - Finish Chapter 3
 - Overview, breakout room exercises
- Titanic Competition Lab

- Panda Basics : Dataframes, dropping columns, descriptive statistics
 - Lecture 5
 - Homework #1 Due
 - Reading 5.1, 5.2
 - Overview of Pandas, breakout room exercises
- Homework #1 Due

Week 4: (9/19, 9/21)

- Panda Basics : Dataframes, dropping columns, descriptive statistics
 - Lecture 6 :
 - Finish Chapter 5
 - Overview of Pandas, Kaggle Competitions
- Numpy Lab : Vector & Matrix computation
- Numpy Basics : Arrays and Vector/Matrix computation
 - Lecture 7:
 - Reading 4.1-4.3
 - Overview of Numpy, Breakout room exercises

Week 5: (9/26, 9/28)

- Numpy Basics : Arrays and Vector/Matrix computation
 - Lecture 8:
 - Finish Chapter 4
 - Overview, Numpy exercises, GPU acceleration
- Data Loading Lab : Web Api & SQL DB data access
- Data Loading and file formats & Data Cleaning
 - Lecture 9 : Data loading CSV, JSON, SQL DB, image, and Web scrapping

- Reading Chapter 6
- Overview of Reading, JSON, Web API, and SQL breakout room exercise

Week 6: (10/3, 10/5)

- Homework #2 Due
- Data Loading and file formats & Data Cleaning
 - Lecture 10 : Data Cleaning :
 - Reading Chapter 7 (skip 7.3 String manipulation)
 - Overview & Kaggle's House Price Competition
- Combining Data Sources Lab
- Data Wrangling : Join, Combine, and Reshape
 - Lecture 11:
 - Reading Chapter 8.1, 8.2
 - Overview, small breakout room exercises

Week 7: (10/12 ... wed)

- No Lab
- Exam #1

Week 8: (10/17, 10/19)

- Data Wrangling : Join, Combine, and Reshape
 - Lecture 12 :
 - Reading Chapter 8.3 (Reshaping)
 - Overview with examples
- Visualization Lab
- Plotting and Visualization
 - Lecture 13 :
 - Reading : Sections 9.1, 9.2
 - Overview, Students come with good examples and

demo

Week 9: (10/24, 10/26)

- Data Aggregation and Grouping operations
 - Lecture 14 : Groupby and Aggregation
 - Reading 10.1, 10.2
 - Overview, breakout room exercises
- Spatial Data Lab
- Data Aggregation and Grouping operations
 - Lecture 15: General Split-apply-combine & Pivot tables
 - Reading Finish Chapter 10
 - Chapter 14.2, 14.3

Week 10: (10/31, 11/2)

- Homework #3 Due
- Ethical Considerations in Data Science
 - Lecture & breakout out room discussions
- Object Oriented Lab
- Object Oriented Python :
 - Lecture 16 :
 - Reading : <https://realpython.com/python3-object-oriented-programming/>
 - Extended Overview
 - Students come with examples from Kaggle notebooks (including questions about examples)

Week 11: (11/7, 11/9)

- Time Series
 - Lecture 17 :
 - Reading 11.1-11.4

- Overview, Breakout room exercises : :
continue/extend lab exercises
- Time Series Lab
- Lecture 18 :
 - Reading Finish Chapter 11
 - Overview, Breakout room exercises :
continue/extend lab exercises

Week 12: (11/14, 11/16)

- Natural Language Processing
 - NLP, Part 1
 - Reading : <https://realpython.com/nltk-nlp-python/>
up until section on Named Entity Recognition
 - Overview
 - Breakout room exercises : Disaster tweets
 - NLP Lab
 - NLP Part 2
 - Homework #4 Due : refined Project proposals : Time series, NLP
 - Lecture 20:
 - Reading : finish <https://realpython.com/nltk-nlp-python/>
 - Overview, Breakout Room exercises : Disaster tweets

Week 13: 11/21

- Homework # 4 Due
- Deep Learning and Computer vision : classification, identification, and segmentation
 - Lecture 21

- Watch : https://www.youtube.com/watch?v=5tvmMX8r_OM
- Background and applications of Deep Learning

Week 14: (11/28, 11/30)

- Audio Processing
 - Read : Audio Data Analysis using Deep Learning in Python (<https://www.kdnuggets.com/2020/02/audio-data-analysis-deep-learning-python-part-1.html>)
- Deep Learning Lab
- Exam 2

Week 15: (12/5, 12/7)

- Project Presentations
- Project Presentations

Week 16: (12/14)

Submit Final Project

Course Evaluations:

At the end of each semester every student has the opportunity to provide constructive feedback on the course. It is important to me that you take the time to let me know your thoughts about the course. I use your feedback to make improvements in the course materials, assignments, and outcomes.

COURSE POLICIES

ATTENDANCE POLICY

Students must attend all online classes. If, for some reason, you cannot make it to a class please notify the instructor.

ACADEMIC INTEGRITY / PLAGIARISM

Everyone associated with the University of Southern Maine is expected to adhere to the principles of academic integrity central to the academic function of the University. Any breach of academic integrity represents a serious offense. Each student has a responsibility to know the standards of conduct and expectations of academic integrity that apply to academic tasks. Violations of academic integrity include any actions that attempt to promote or enhance the academic standing of any student by dishonest means. Cheating on an examination, stealing the words or ideas of another (i.e., plagiarism), making statements known to be false or misleading, falsifying the results of one's research, improperly using library materials or computer files, or altering or forging academic records are examples of violations of this policy which are contrary to the academic purposes for which the University exists. Acts that violate academic integrity disrupt the educational process and are not acceptable.

Evidence of a violation of the academic integrity policy will normally result in disciplinary action. A copy of the complete policy may be obtained from the office of Community Standards and Mediation, online at usm.maine.edu/community-standards-mediation/academic-integrity or by calling and requesting a copy at (207) 780-5242.

UNIVERSITY POLICIES AND RESOURCES

DISABILITY ACCOMMODATIONS

The university is committed to providing students with disabilities equal access to all university programs and services. If you think you have a disability and would like to request accommodations, please contact the Disability Services Center. Timely notification is essential. The Disability Services Center can be reached by calling 207-780-4706 or by email dsc-usm@maine.edu. If you have already received a faculty accommodation letter from the Disability Services Center and would like to request accommodations for this course, please provide me with that information as soon as possible. Please make a private appointment so that we can review your accommodations together.

TUTORING AND WRITING ASSISTANCE

Tutoring at USM is for *all* students, not just those who are struggling. Tutoring provides active feedback and practice, and is available for writing, math, and many more subjects. Walk-in tutoring is available at the Glickman Library in Portland, the Gorham Library, and the LAC Writing Center. For best service, we recommend making an appointment at <https://usm.maine.edu/learningcommons/schedule-tutoring-appointment>. Questions about tutoring should be directed to Naamah Jarnot at 207-780-4554. Interested in becoming a more effective, efficient learner? Check out <https://usm.maine.edu/agile!>

HEALTH AND COUNSELING

Counseling is available at USM. The best way to schedule an

appointment is by phone at 780-5411. More information is available at <https://usm.maine.edu/uhcs>.

RECOVERY ORIENTED CAMPUS CENTER (ROCC)

A peer support community for students in recovery from substance abuse and other mental health conditions is available at USM. More information may be found online at <https://usm.maine.edu/recovery> or by contacting ROCC at 207-228-8141.

NONDISCRIMINATION POLICY AND BIAS REPORTING

The University of Southern Maine is an EEO/AA employer, and does not discriminate on the grounds of race, color, religion, sex, sexual orientation, transgender status, gender expression, national origin, citizenship status, age, disability, genetic information or veteran's status in employment, education, and all other programs and activities. The following person has been designated to handle inquiries regarding non-discrimination policies: Amie Parker, Interim Director of Equal Opportunity, The Farmhouse, University of Maine Augusta, Augusta, ME 04333, 207.581.1226, TTY 711 (Maine Relay System). Incidents of discrimination or bias at USM should be reported to Associate Vice President for Student Affairs David Roussel at 207-780-5242.

STATEMENT ON RELIGIOUS OBSERVANCE FOR USM STUDENTS

Absence for Religious Holy Days: The University of Southern Maine respects the religious beliefs of all members of the community, affirms their rights to observe significant religious holy days, and will make reasonable accommodations, upon request, for such observances. If a student's religious observance is in conflict with the academic

experience, they should inform their instructor(s) of the class or other school functions that will be affected. It is the student's responsibility to make the necessary arrangements mutually agreed upon with the instructor(s).

TITLE IX STATEMENT

The University of Southern Maine is committed to making our campuses safer places for students. Because of this commitment, and our federal obligations, faculty and other employees are considered mandated reporters when it comes to experiences of interpersonal violence (sexual assault, sexual harassment, dating or domestic violence, and stalking). Disclosures of interpersonal violence must be passed along to the University's Deputy Title IX Coordinator who can help provide support and academic remedies for students who have been impacted. More information can be found online at <http://usm.maine.edu/campus-safety-project> or by contacting Sarah E. Holmes at usm.TitleIX@maine.edu or 207-780-5767.

If students want to speak with someone confidentially, the following resources are available on and off campus: University Counseling Services (207-780-4050); 24 Hour Sexual Assault Hotline (1-800-871-7741); 24 Hour Domestic Violence Hotline (1-866-834-4357).

POLICY ON ACCEPTABLE CONDUCT IN CLASS SETTINGS

If a student substantially disrupts a class, the professor may ask the student to align with this policy on conduct in a class setting. If the student refuses, the professor may, at their discretion, ask the student to leave. If the professor takes this step, they must attempt to communicate with the student and provide informal counsel and advice. The professor

may elect to notify their dean of the situation as well. If the student disrupts the class again, the professor may, at their discretion, provide a written notification to the student, describe the offending behavior, and refer the student's case to the appropriate academic dean and notify the dean of students that an official student conduct code violation has occurred [<https://usm.maine.edu/community-standards-mediation/conduct-process>].

COVID FACE COVERING REQUIREMENT

Please follow State and University guidance.

<https://www.maine.edu/together/community-guidance/everyone/>