

CSS 2 Syllabus: Winter 2024

(Note that this is a static, fixed version of the CSS 2 syllabus; students can access a website with additional information at: <https://ucsd-css2.github.io/ucsd-css2-website/course/syllabus.html>).

Course Overview

This course will teach how you to think about **data** and **modeling** for **computational social science**. This will involve raising and examining questions like:

- How is a dataset **formatted**, and is this the appropriate format for what I want to do?
- Is this dataset **representative** or does it reflect a **biased sample**?
- What **ethical considerations** should I take into account when obtaining and analyzing data?
- What kind of **model** is the most appropriate for these data?
- How do I **design and implement** these models—ranging in complexity from [linear regression](#) to [support vector machines](#)?

These topics will be discussed in the context of **hands-on practice** with real-world datasets in a **Python programming environment**.

Key Learning Outcomes

This course is designed to give students a range of conceptual and technical tools. My goal is that by the end, you are able to:

- **Identify** and **explain** ethical issues that arise in CSS, along with analytical issues more broadly.
- **Design** and **implement** clear, concise, and accurate visualizations of data.
- **Propose** and **test** hypotheses about data using statistical models.
- **Construct** statistical models in Python and **interpret** the results.
- **Weigh** the pros and cons of different model evaluation metrics.

Course Logistics

Teaching Team OH

Who?	When?	Where?
Sean Trott	Monday 11-12	Zoom (link on Canvas)

Grading

Grade Components

Your grade will be determined by three kinds of assessments: [coding labs](#), [problem sets](#), and a [final project](#).

Grade Component	Percentage of Final Grade
8 Coding Labs	50% (6.25% each)
4 Problem Sets	32% (8% each)

Grade Component	Percentage of Final Grade
1 Final Project	18%

Letter Grades

If you're taking the course for a letter grade, your grade will be determined according to the scale below.

Note that the number on the **right-hand** side of the range is *not included* in that range: that is, an "A-" ranges from 90% all the way to 91.99% but does not include 93% (93% is an A).

Percentage	Letter Grade
97%+	A+
93-97%	A
90-93%	A-
87-90%	B+
83-87%	B
80-83%	B-
77-80%	C+
73-77%	C
70-73%	C-
60-69%	D
<60%	F

On Rounding

Note that my policy is *not* to round up grades for two reasons:

1. If rounding is applied selectively (i.e., only to students who ask), it is unfair to other students.
2. If rounding is applied across the board, it simply redefines the boundary between two letter grades (e.g., making an 89% the cut-off for an A-).

Late submissions

Students may submit late assignments up to 48 hours after the submission deadline, for 75% credit of what you would've received (i.e., if you scored 90%, you'd get 67.5% with the late penalty).

Questions, feedback, and communication

Instructors can be reached in the following ways:

- Office hours.
- Public question on Piazza.
- Private message over Piazza.
- Email.

The course Piazza can be found here: <https://piazza.com/ucsd/winter2023/602>

Note that in general, we prefer communication over Piazza as opposed to email.

Other Information

Academic Integrity

Please turn in your own work. While you are encouraged to work together on some assignments (e.g., on [labs](#)), you should still understand the code you've submitted. Problem sets and final project should be completed independently.

Please review academic integrity policies [here](#). Cheating and plagiarism are unfair to other students and ultimately to yourself, and you will be penalized if caught. Instead, if you're struggling with something, please come to office hours and ask for help!

Class Conduct

All of us (instructors + students) should treat others with respect and follow the UC San Diego [Principles of Community](#). This class should be a welcoming and inclusive environment for everyone, regardless of gender, gender identity and expression, sexual orientation, physical appearance, disability, race, or ethnicity.

Please be considerate and respectful of your fellow classmates (and instructors), refrain from discriminatory language and harassment, and interact with good faith.

Schedule of Course Content

See the schedule [here](#).

CSS 2 Schedule

Note that this schedule is **aspirational** and is therefore **subject to change**.

Date	Topics	Assignment due	Notes
Week 1	Introduction, Ethics		
M-01-08-2024	Introduction		
W-01-10-2024	Ethics: Bias and Fairness		
F-01-12-2024	Ethics: Privacy, CSS for Social Good		
Week 2	Python Review		
M-01-15-2024	Holiday (MLK, Jr. Day)		

W-01-17-2024	Python review: Fundamentals		
F-01-19-2024	Python review: pandas	Lab 1 due	
Week 3	Data Visualization		
M-01-22-2024	Data visualization (matplotlib.pyplot)		
W-01-24-2024	Data visualization (seaborn)	PS 1 due	
F-01-26-2024	Data visualization (seaborn)	Lab 2 due	
Week 4	Working with Data		
M-01-29-2024	Data visualization (principles) (Remote, pre-recorded lecture)		
W-01-31-2024	Data wrangling, "tidy" data		
F-02-02-2024	Data wrangling, "tidy" data"	Lab 3 due	
Week 5	Describing Data		
M-02-05-2024	Descriptive statistics, data transformations		
W-02-07-2024	Descriptive statistics, data transformations	PS 2 due	
F-02-09-2024	Introduction to statistical modeling	Lab 4 due	
Week 6	Linear regression		
M-02-12-2024	Linear regression: basics		
W-02-14-2024	Linear regression: heteroscedasticity, residuals		
F-02-16-2024	Linear regression: heteroscedasticity, residuals	Lab 5 due	
Week 7	Multiple regression		
M-02-19-2024	Holiday (President's Day)		
W-02-21-2024	No class	PS 3 due	
F-02-23-2024	Multiple regression	Project proposal due	
Week 8	Regression, generalized		
M-02-26-2024	Multiple regression, continued		
W-02-28-2024	Non-linear regression		
F-03-01-2024	Overfitting; cross-validation	Lab 6 due	
Week 9	Classification		
M-03-04-2024	Logistic regression; Type 1 vs. Type 2		
W-03-06-2024	Evaluation using ROC Curve, AIC, Confusion Matrix	PS 4 due	
F-03-08-2024	Evaluation, continued	Lab 7 due	
Week 10	<i>*Unsupervised approaches</i>		

M-03-11-2024	K-means clustering		
W-03-13-2024	K-means clustering		
F-03-15-2024	Wrap-up, final project work	Lab 8 due	
Finals week			
M-03-18-2024			
W-03-20-2024		Final project due	
F-03-22-2024			