CS 220 / CS319

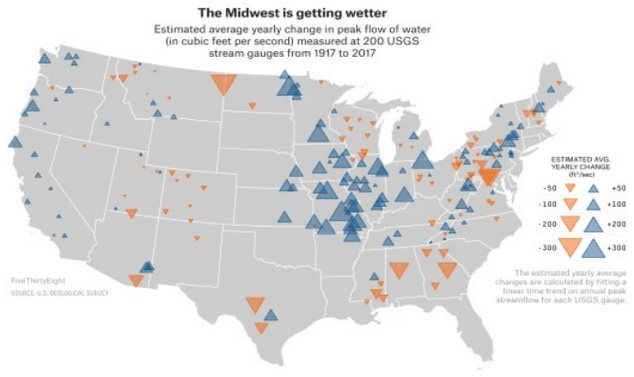
Introduction

Department of Computer Sciences
University of Wisconsin-Madison

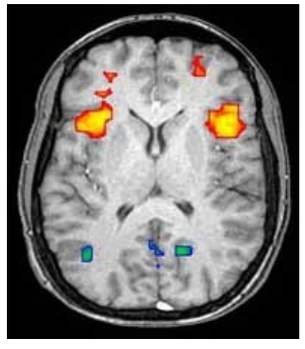
Welcome to Data Science Programming I

Data is now integrated into many fields

- Journalism
- Biology, physics, chemistry
- Psychology, sociology, economics, business
- Engineering (mechanical, biomedical, industrial, etc.)

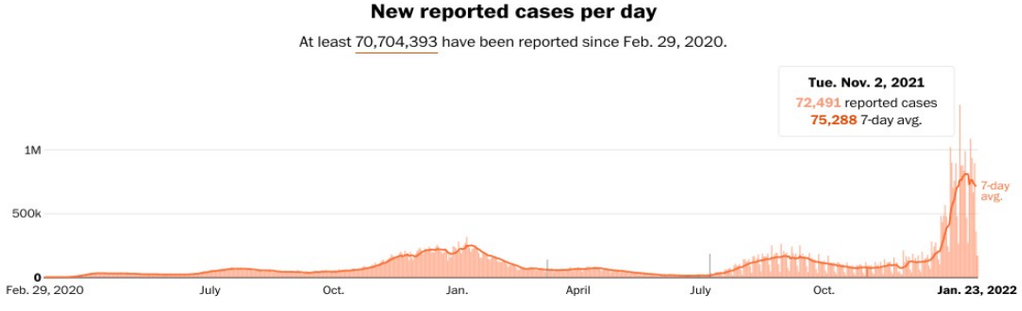


https://fivethirtyeight.com/features/the-midwest-is-getting-drenched-and-its-causing-big-problems/

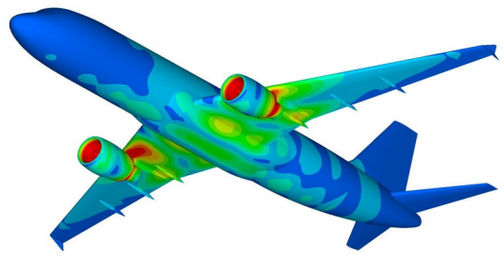


https://en.wikipedia.org/wiki/Neuroimaging

https://science.howstuffworks.com/life/genetic/gattaca-gaptacaz-adding-letters-the-genetic-alphabet.htm



https://www.washingtonpost.com/graphics/2020/national/coronavirus-us-cases-deaths/



http://www.stressebook.com/finite-element-analysis-in-a-nut-shell/

Welcome to Data Science Programming I

Data is now integrated into many fields

- Journalism
- Biology, physics, chemistry
- Psychology, sociology, economics, business
- Engineering (mechanical, electrical, industrial, etc)

How can we gain insights from that data?

- With computation

Approach 1: human computation

https://en.wikipedia.org/wiki/Human_computer

Approach 2: machine computation

http://fortune.com/2015/11/15/intel-super-7/

Welcome to Data Science Programming I

CS 220 is about approach 2

- Faster, more reliable, can churn through more data
- Automate to save human effort

"Find the leverage in the world, so you can be more lazy!"

~ Larry Page

Approach 1: human computation

https://en.wikipedia.org/wiki/Human_computer

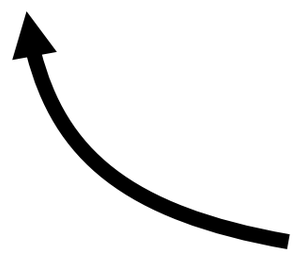
Approach 2: machine computation

http://fortune.com/2015/11/15/intel-super-7/

Welcome to Data Science Programming I

CS 220 is about approach 2

- Faster, more reliable, can churn through more data
- Automate to save human effort
- Requires being able to tell computers what to do!



society needs more domain experts in specific fields who can write code

Goal: become "bilingual"

- Speak the language of X (biology, mech eng, journalism, etc)
- Speak the language of computing

Data Science:

- Combines inquiry, statistics, programming, and communication skills to provide actionable insights from data sets

Why CS 220?

Typical intro CS

- Challenging language (e.g., C++ or Java)
- CS students and other majors together
- Heavy on theory, light on data

vs

CS 220 approach

- Python (powerful but easier to learn)
- Bring more coding into other fields
- Light on theory, heavy on data
- Emphasize questions and communication

Why CS 220?

	Title	Projected Jobs	Median Salary	Education
1	Nurse Practitioner	135,500	$126,260	Masters
2	IT Manager	106,9000	$169,510	Bachelor
3	Physician Assistant	43,700	$130,020	Masters
4	Financial Manager	138,300	$156,100	Bachelor
5	Software Developer	303,700	$132,270	Bachelor
6	Information Security Analyst	59,100	$120,360	Bachelor
7	Med. & Health Services Manager	160,600	$110,680	Bachelor
8	**Data Scientist**	**73,100**	**$108,020**	**Bachelor**
9	Speech-Language Pathologist	33,300	$89,290	Masters
10	Actuary	6,600	$120,000	Bachelor

https://money.usnews.com/careers/best-jobs/rankings/the-100-best-j

Why CS 220?

People use Data to solve the world's problems

https://ourworldindata.org/

Measuring progress towards the Sustainable Development Goals

https://sdg-tracker.org/

Today's Topics

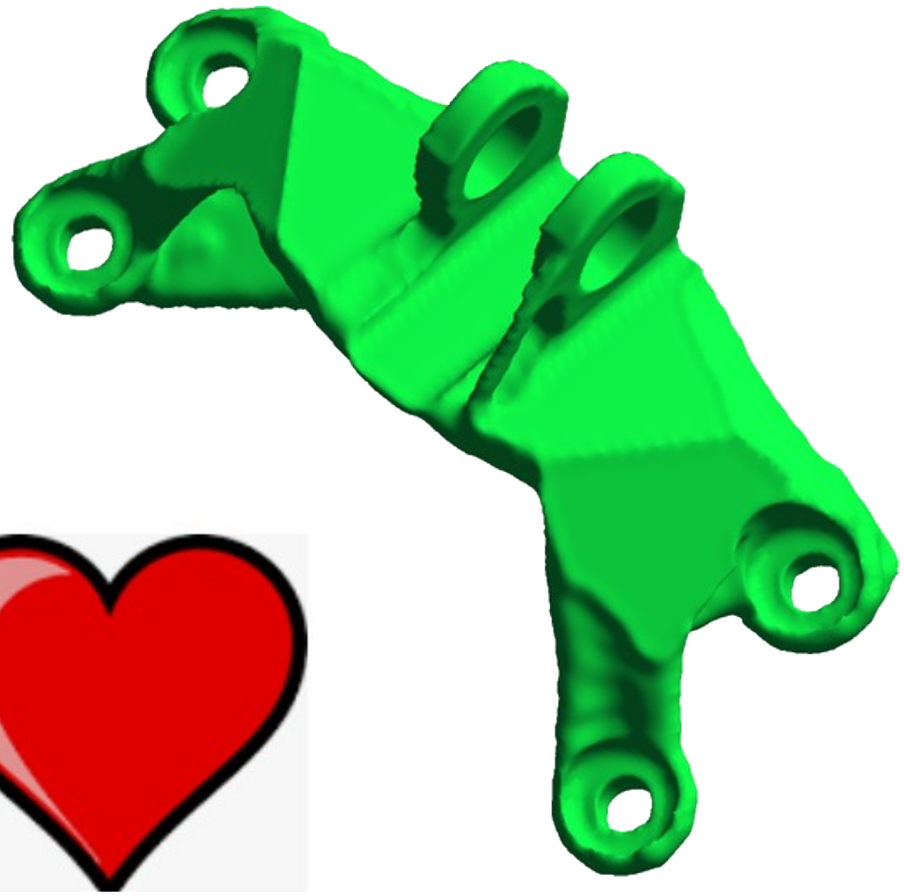
Introductions

- Who are the faculty? Who are you?

Course overview

Website

Who are the faculty?



Mike Doescher
- Email: mdoescher@wisc.edu
- Please call me “Mike”

Industry and Teaching experience
- Naval Research Laboratory
- Benedictine College
- SciArt Software
- UW Madison

Who are the faculty?

Louis Oliphant (i.e. Louis)

Email: ltoliphant@wisc.edu

BA in Mathematics Education @ BYU, 1995

Taught High School for several years

MS and PhD in Computer Science @ UW-Madison

Dissertation in Machine Learning, 2009

Adaptively Finding and Combining First-Order Rules for Large, Skewed Datasets

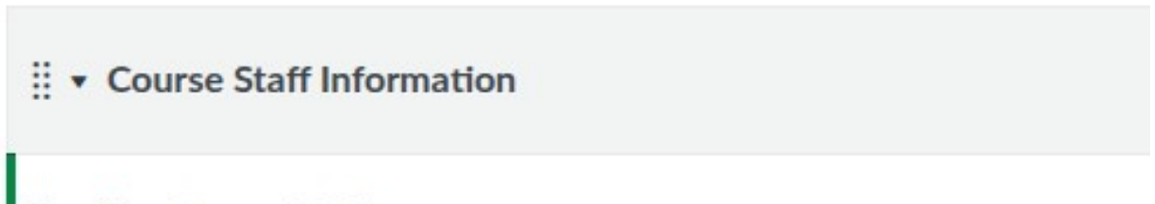
Taught at Hiram College in Ohio from 2009-2023

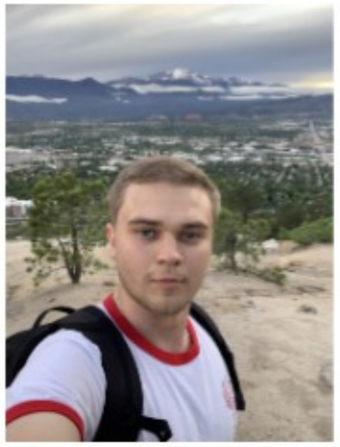
Started at UW-Madison in Fall, 2023

Always looking for chess players

Who are the Staff?

- 2 Head Teaching Assistants
- 14 Teaching Assistants
- 14 Peer mentors



	Hi, I'm Dmitry Pinchuk, a head TA for CS 220, and a second-year master's student in Computer Science with a focus in reinforcement learning. This is my third semester as a CS 220 TA. In my free time I like to play chess and soccer as well as read books and watch movies.
	Hi, I'm Mo Xu, a second year CS PhD student on robotics. This is my third semester TAing CS 220. I like various video games and you can also see me at Bakke.
	Hi! My name is Shuibai. This is my second semester TA at CS220. I'm a second-year CS PhD student specializing in machine learning. I enjoy working out, hiking, and exploring anything interesting in my free time.
	Hi, my name is Yunhao and I am a 2nd year PhD student in Computer Science. I enjoy playing badminton and gaming in my free times.

Who are you?

- 800 students in 4 sections
- Say hi to your neighbor and introduce yourself
  - Name
  - Major (potential major)
  - Year in college
  - Fun fact

https://pixy.org/4356032/

Today's Topics

Introductions

Course overview

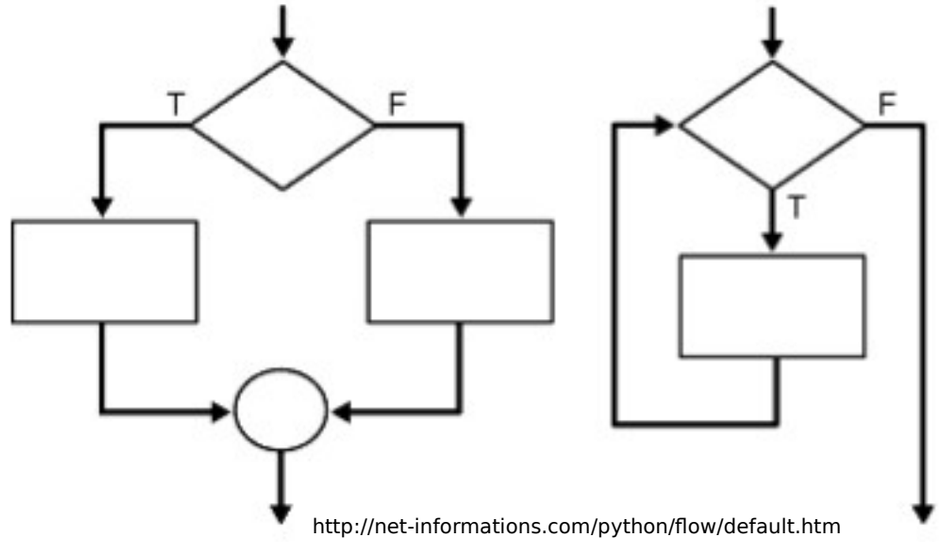
- Topics
- Lectures
- Lab
- Readings
- Course tools
- Grades
- Projects
- Exams & quizzes

Website

220 Topics

Part 1: Control Flow

- What step is currently executing?
- How to write functions?
- How to conditionally do something?
- How to repeat steps?



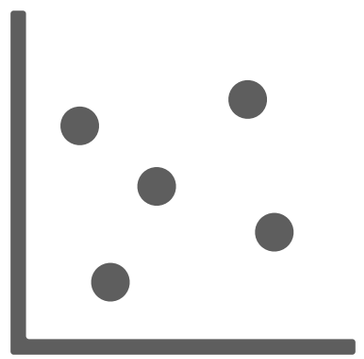
http://net-informations.com/python/flow/default.htm

Part 2: State

- How to structure lots of data?
- How to save data in files?

Part 3: Data Science

- Tabular data
- Internet
- Databases
- Plotting

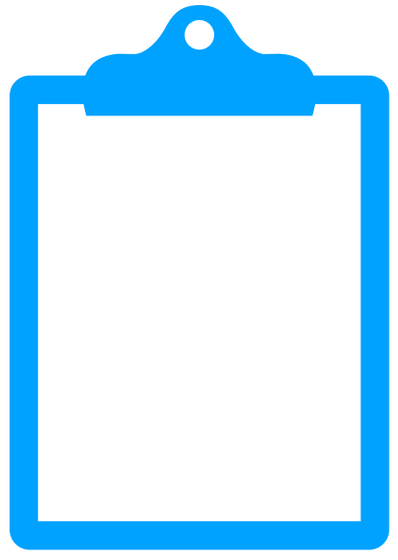


Generative AI: The use of artificial intelligence (AI) tools and applications (including, but not limited to, ChatGPT, DALL-E, and others) for course assignments and assessments does not support the learning objectives of this course and is prohibited. Using them in any way for this course is a violation of the course's expectations and will be addressed through UW-Madison's academic misconduct policy, specifically UWS 14.03(1)b (b) Uses unauthorized materials or fabricated data in any academic exercise.

Lectures

general concepts

live coding



worksheet practice

Your role

- Do readings before or after
- Download the template file and code along in lecture

  `pixi run lecture`
- Ask questions during the lectures + office hours
- Lectures *WON'T* be recorded

Labs – CS220

Format

- 13 labs (3 dropped)
- 75 minutes on Wed, Thu or Fri
- led by Teaching Assistant (TA) and a Peer Mentor (PM)
- Lab attendance is required
- Meant to practice weekly content & help you succeed on projects

Labs – CS319

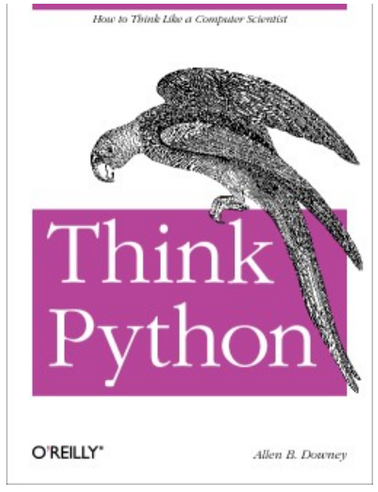
Format

- Attendance is optional
- May attend any lab

we will have labs this first week!!!

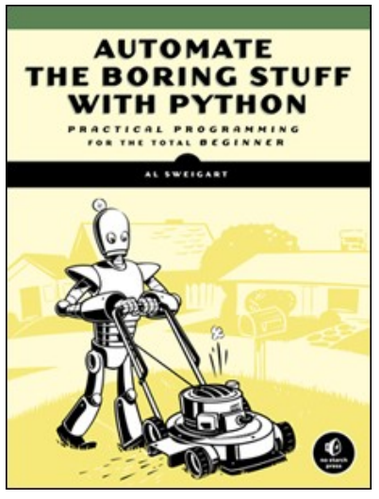
(also, get any help needed installing Python during this one)

Readings (all free!)



Think Python, 2nd Edition

- Allen B. Downey
- Assumes no programming background
- It's very concise
- Get the 2nd edition, which is for Python 3!



Automate the Boring Stuff

- Al Sweigart
- Useful for some more advanced topics related to using data

Python for Everyone – Interactive

- Barb Ericson
- Allows you to practice coding as you learn

CS 220 / CS319 Notes

Course Notes

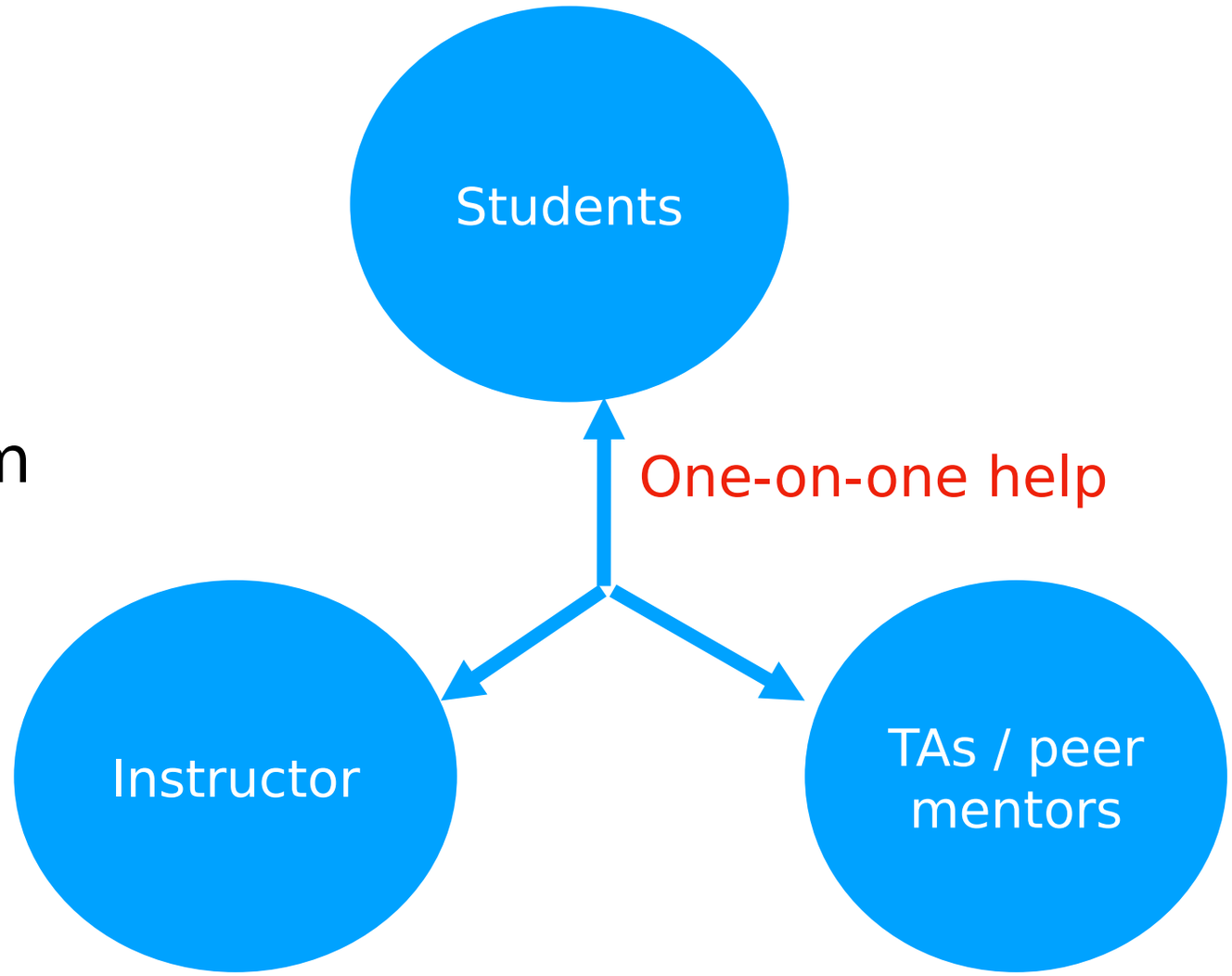
- 220 / 319 instructors
- Mostly for data science part of class

Course tools

Communication tools

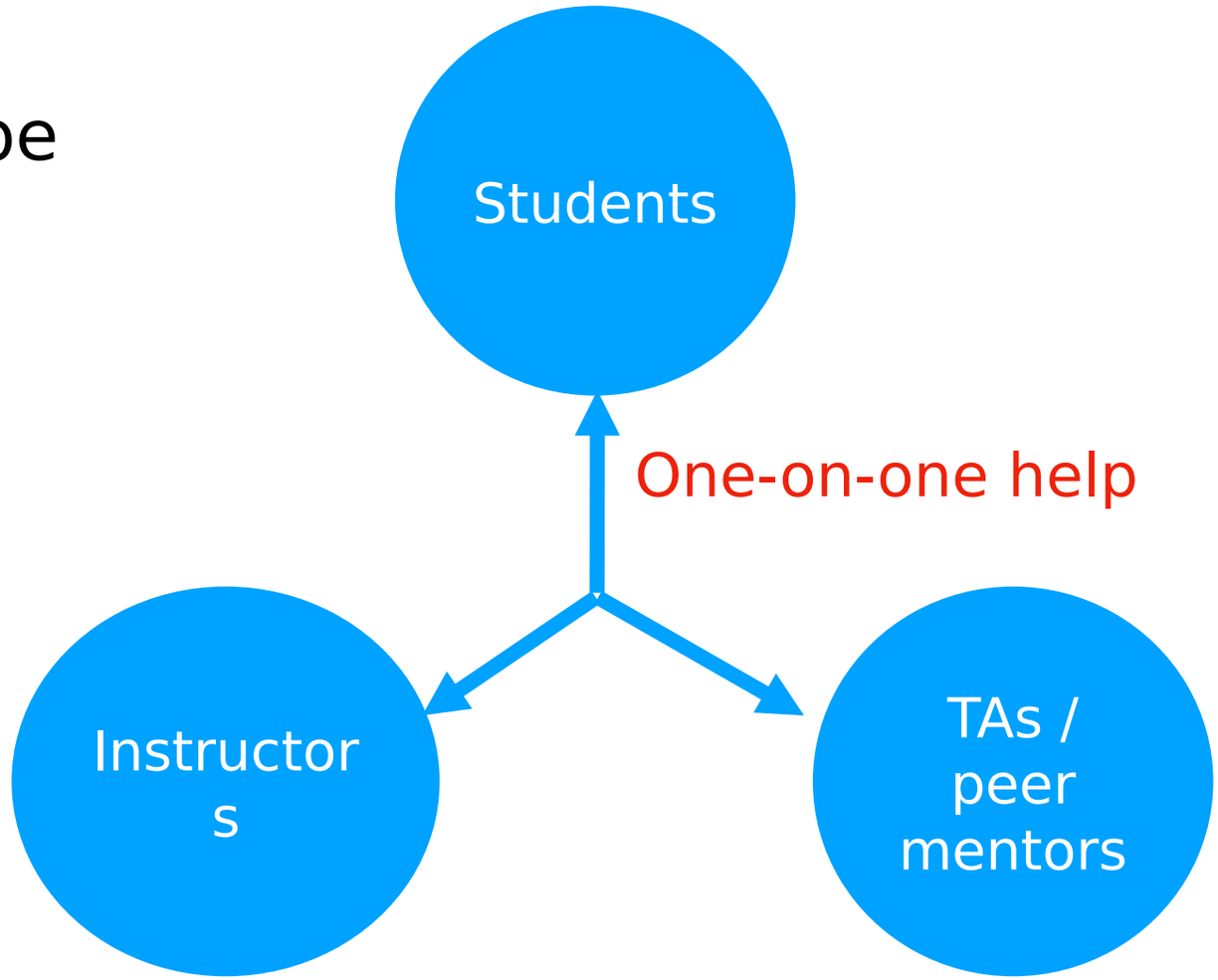
- Office hours (best way to get help):
  - Shared Google calendar: requires login with your wisc email

- Canvas:
  - Quizzes
  - Personal messages (exam room assignment, exam scantron results, etc.,)
  - Grades



Course tools

- Piazza:
  - Rule 1: don’t post more than 5 lines of code
  - Rule 2: check other posts before posting
- Class Forms:
  - Feedback form (anonymous / non-anonymous)
  - Exam conflict forms
  - Thank you form
- Project Submission: GradeScope
- Email (least-preferred):
  - Class mailing lists



Grades - CS220

47% - **programming projects**

- 13 projects
- p1: 1%, p9: 2%, remaining projects 4%

32% - **exams**

- 10% midterm 1
- 10% midterm 2
- 12% final

16% - **quizzes**

- 10 quizzes (drop 2 lowest scores)

4% - **labs**

- 13 labs

1% - **class surveys**

Grades – CS319

31% - **programming projects**

- 9 CS220 projects: p1 to p9
- p1: 1%, p9: 2%, remaining projects 4%

32% - **exams**

- 10% midterm 1
- 10% midterm 2
- 12% final

16% - **quizzes**

- 10 quizzes (drop 2 lowest scores)

20% - **graduate-level project**

- Proposal: 4%
- Data Set: 5%
- Initial Code: 1%
- Final Code: 5%
- Presentation: 5%

1% - **class surveys**

Letter Grades

- Your final grade is based on sum of all points earned
- Your grade does not depend on other students' grade – no curving
- **We will NOT be rounding off scores at the end of the semester**
- **No extra credit**

Grade cut-offs

- 93% - 100%: A
- 88% - 92.99%: AB
- 80% - 87.99%: B
- 75% - 79.99%: BC
- 70% - 79.99%: C
- 60% - 69.99% D

Today's Topics

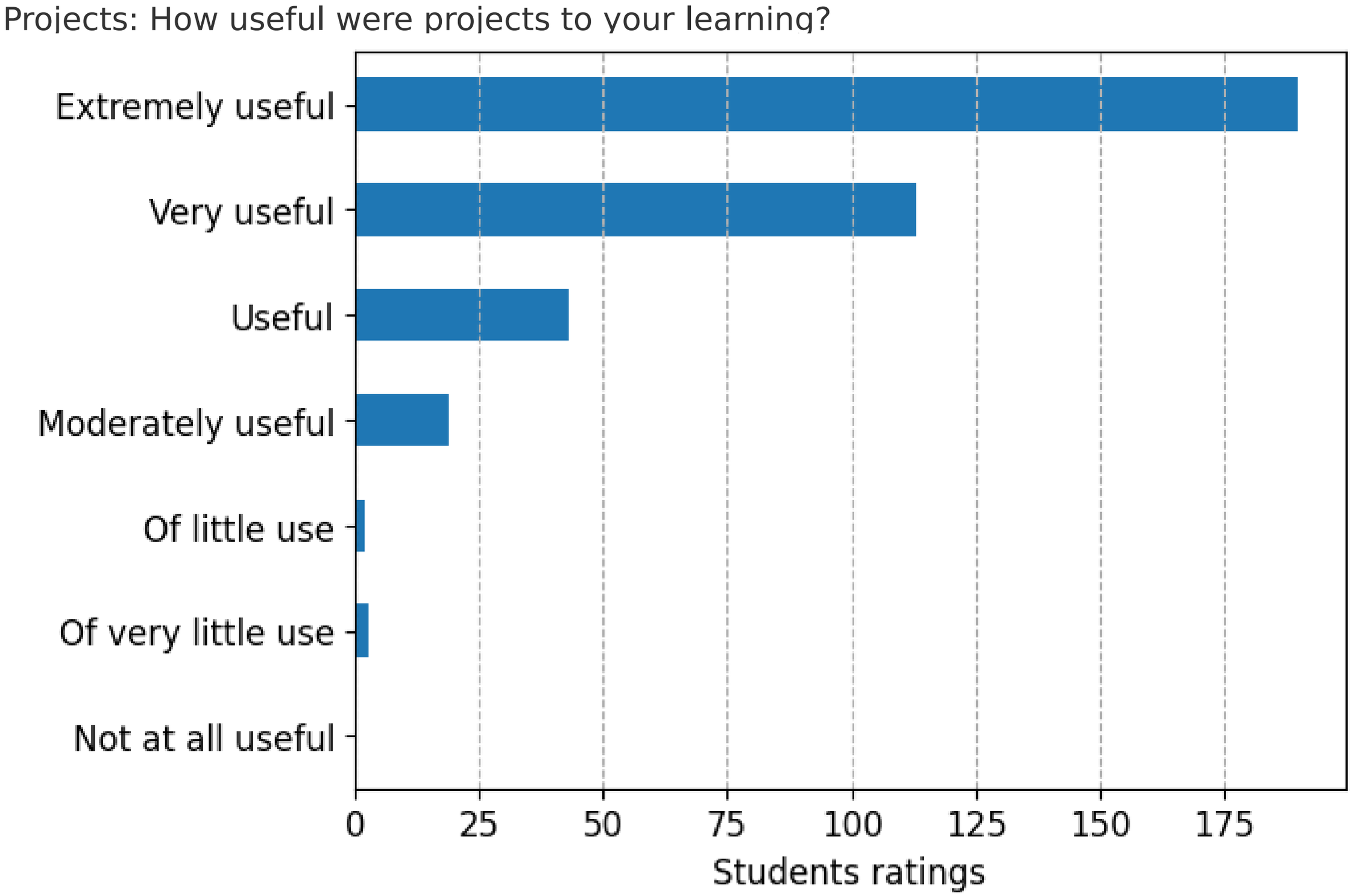
Introductions

Course overview

- Topics
- Lecture
- Lab
- Readings
- Class communication
- Grades
- Projects
- Exams & quizzes

Website

Prior student reaction to projects



Project Overview

Nearly all projects will relate to some dataset

Timeline

- Projects will be due on Fridays at 4:00:00 pm
- You get a bank of 12 late days (automatically applied, may not be used on last project)
- After late days, 5% deduction per day late (up to 7 days late)

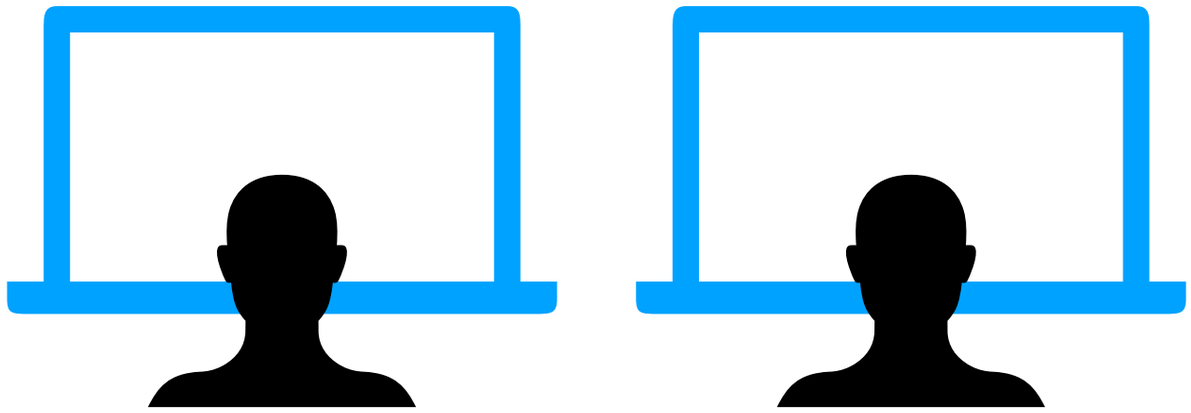
Getting help

- Office hours
- Lab sessions
- Piazza

Pair Programming

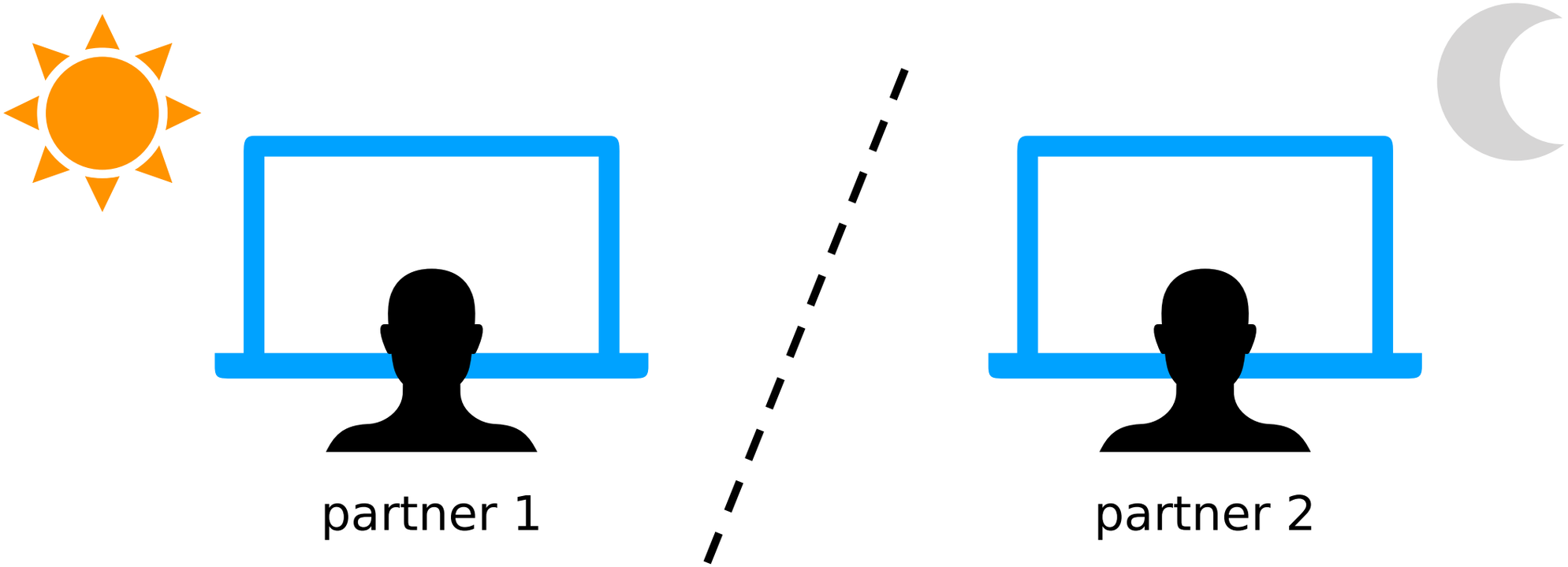
You can optionally work in pairs of two

- CS220 students can partner with any CS220 students, in any section
- CS319 students can partner with any CS319 students
- You can choose to keep the same partner, for multiple projects or choose to switch partners



Best practice: working alongside each other

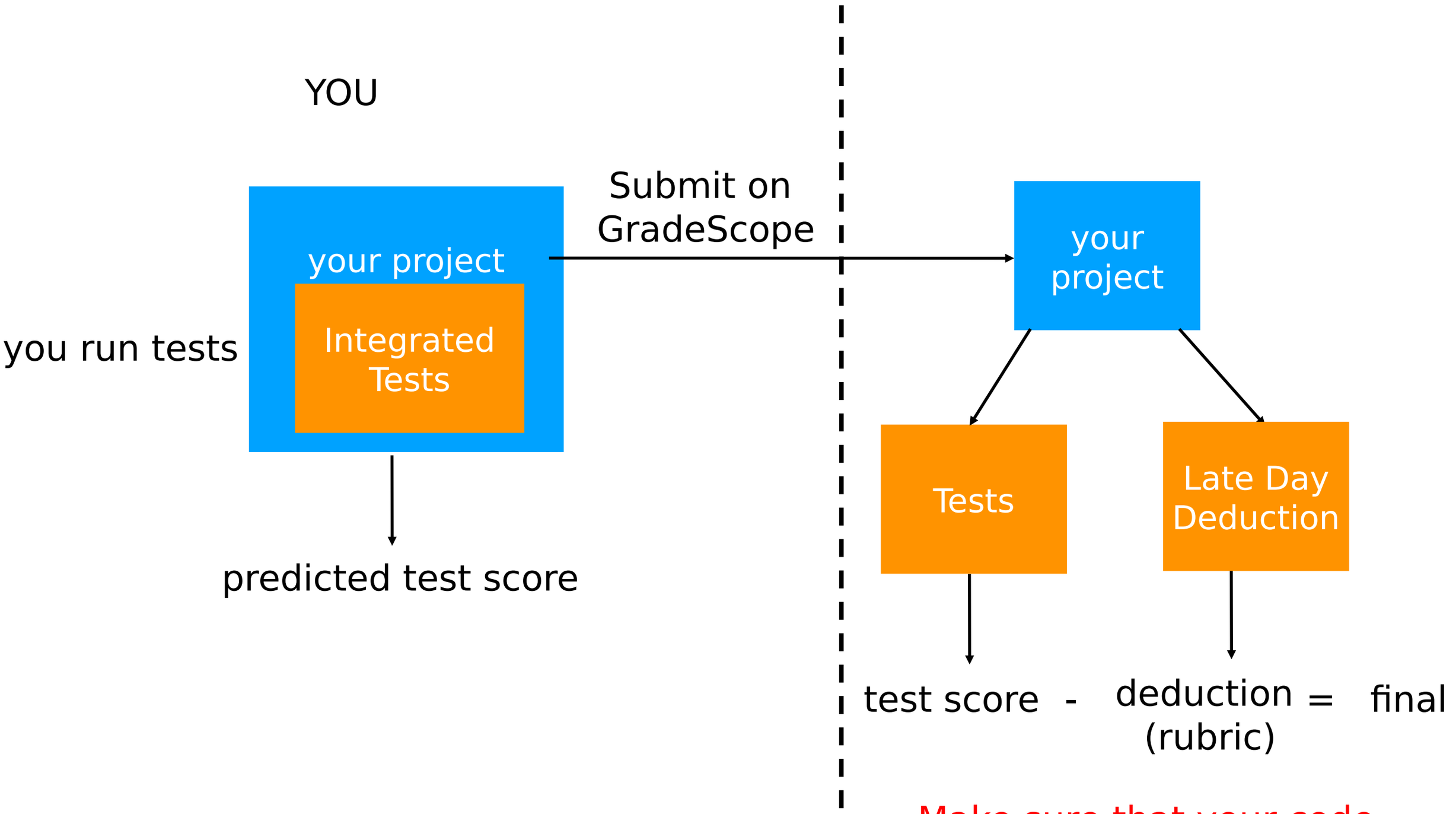
Pair Programming



Breaks syllabus rules: working on different parts at different times

Breaks syllabus rules: working on alternate projects individually

Project Grading



Make sure that your code clears tests on GradeScope!

Today's Topics

Introductions

Course overview

- Topics
- Lecture
- Lab
- Readings
- Class communication
- Grades
- Projects
- Exams & quizzes

Website

Quizzes and Exams

Quizzes

- Will be due most weeks, on Fri, at 11:59pm
- Focus on recent lectures so you stay current and check your knowledge

Exams: two midterms and one final

- Multiple choice
- 1.5 hours midterms
- 2 hours final
- Given in a large lecture hall

projects - writing and testing code with a computer

quizzes - reading and interpreting code with a computer

exams - reading and interpreting code without a computer

Today's Topics

Introductions

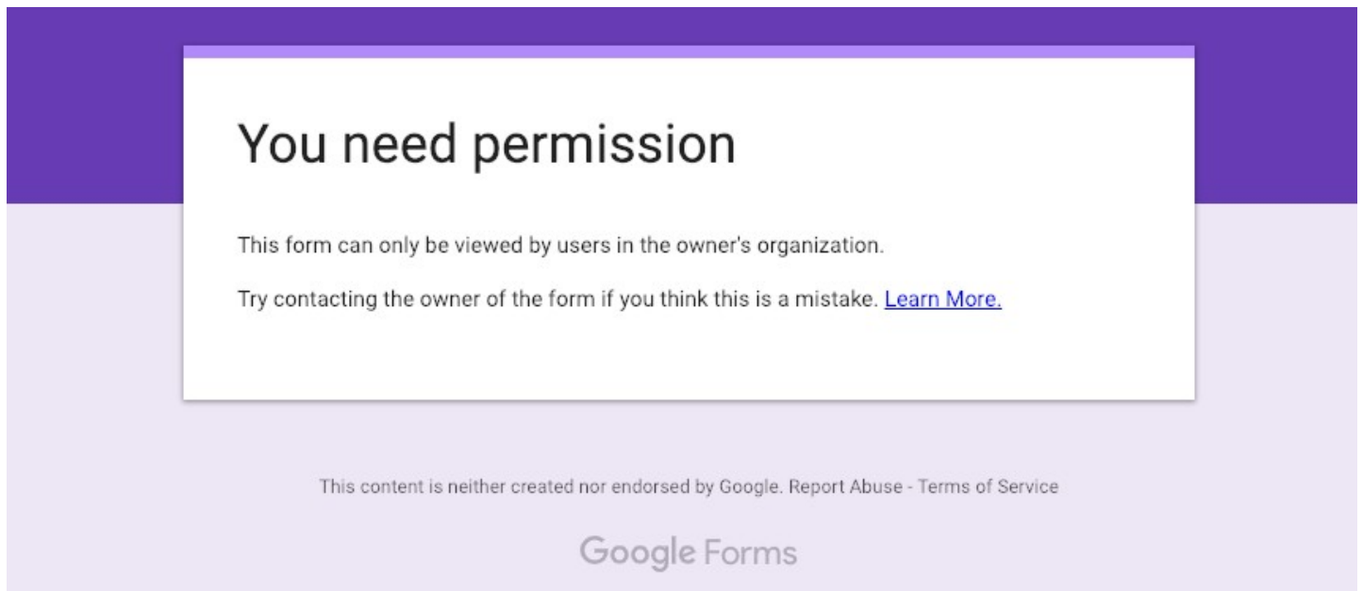
Course overview

Website

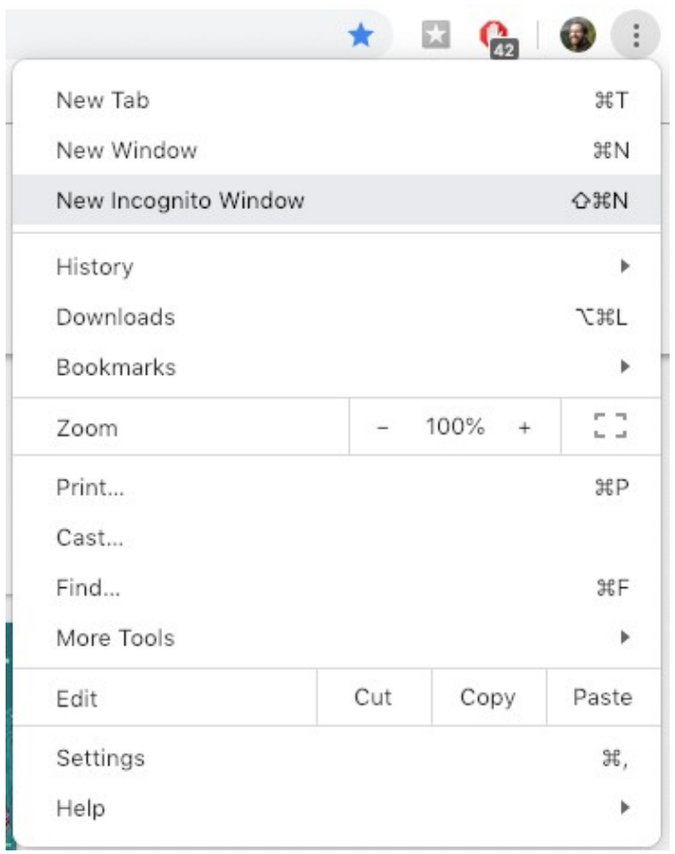
Course Website

https://cs220.cs.wisc.edu/s26/schedule.html

Walk through...



if you see this, it means you're signed in via Gmail instead of your campus email



if you were automatically signed into gmail without being asked, consider clearing cookies or using an Incognito Window (in Chrome)

Next steps...

- Complete the ”Student Information Survey" survey quiz on canvas.
- Read syllabus carefully
- Setup Python on your computer and attend your lab for Lab-P1
- Submit P1 (Project 1) after attending lab: due next Fri
- Sign up for GradeScope