

CS 220

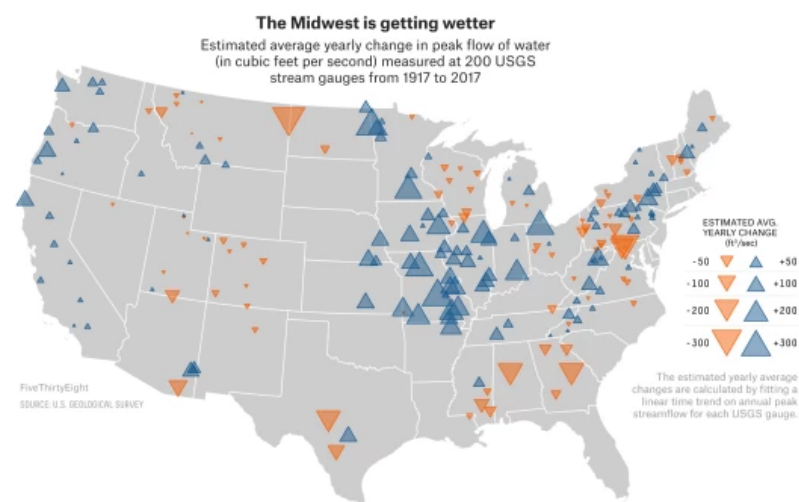
Introduction

Department of Computer Sciences
University of Wisconsin-Madison

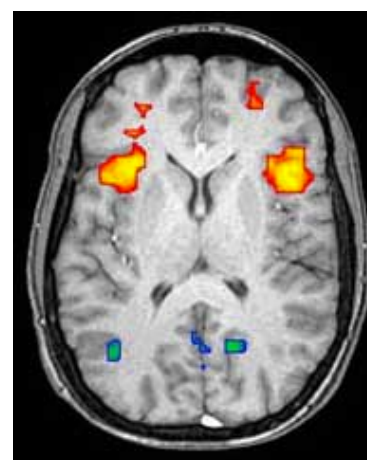
Welcome to Data Science Programming I

Data is now integrated into in many fields

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



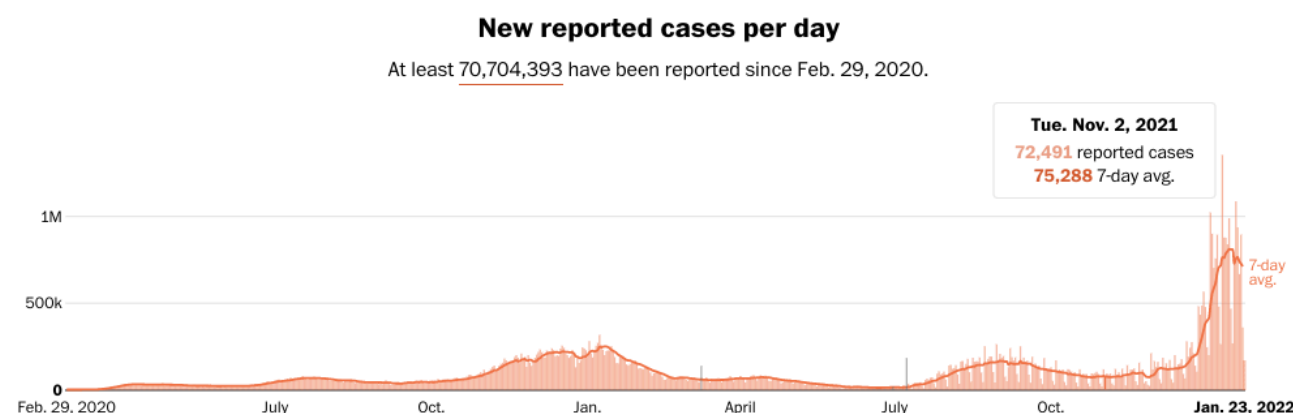
<https://fivethirtyeight.com/features/the-midwest-is-getting-drenched-and-its-causing-big-problems/>



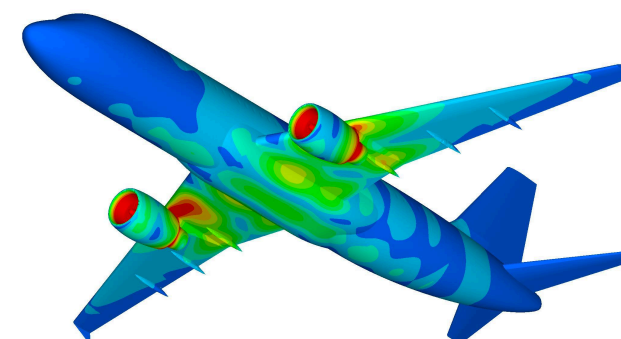
<https://en.wikipedia.org/wiki/Neuroimaging>



<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 exploding in 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!**”*

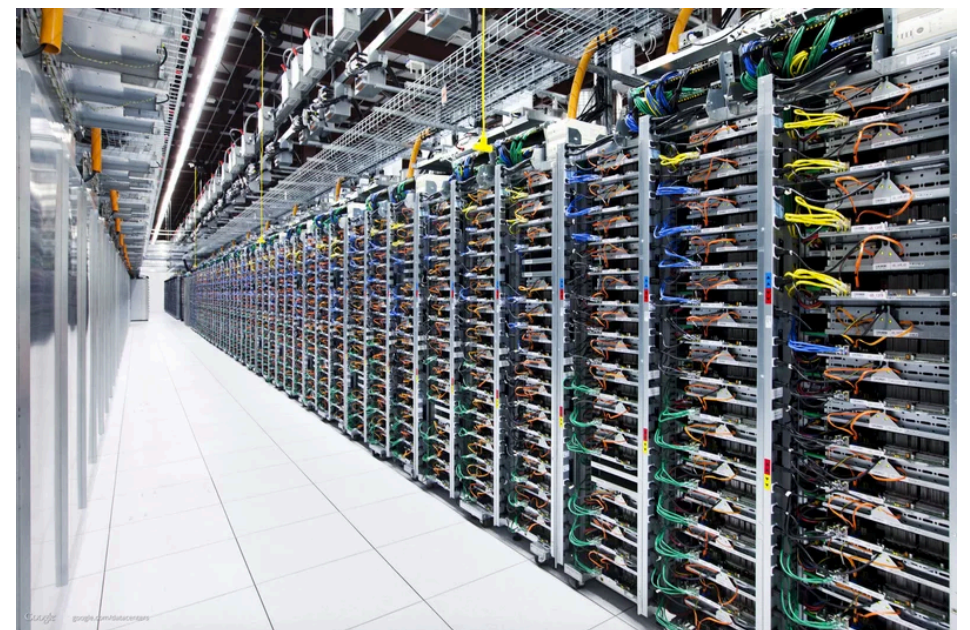
~ Larry Page

Approach 1: human computation



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

Approach 2: machine computation

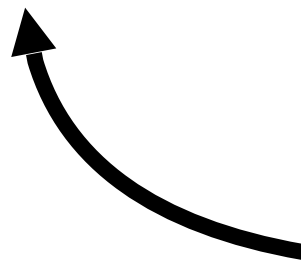


<http://fortune.com/2015/11/15/intel-super-7/>

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?

50 Best Jobs in America for 2022

	Job Title	Median Base Salary	Job Satisfaction	Job Openings
#1	Enterprise Architect	\$144,997	4.1/5	14,021
#2	Full Stack Engineer	\$101,794	4.3/5	11,252
#3	Data Scientist	\$120,000	4.1/5	10,071

https://www.glassdoor.com/List/Best-Jobs-in-America-LST_KQ0,20.htm

Why CS 220?

People use Data to solve the world's problems



Measuring progress towards the Sustainable Development Goals

Today's Topics

Introductions

Course overview part 1

Worksheet: pseudocode

Course overview part 2

Canvas and course materials

Who is the instructor?

Anna Meyer

- Email: apmeyer4@wisc.edu
- Please call me “Anna”

Education & Experience

- BA (Mathematics, Carleton College)
- MS (Computer Science, UW-Madison)
- PhD (Computer Science, UW-Madison) in-progress
- 2 years experience as a software developer

Who are the TA and peer mentors?

TA: Jane Zhang

- Email: zhang2752@wisc.edu
- MS student in Computer Science

Peer Mentor: Adi Tewari

Who are you?

- 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 part 1

- Topics

Worksheet: Pseudocode

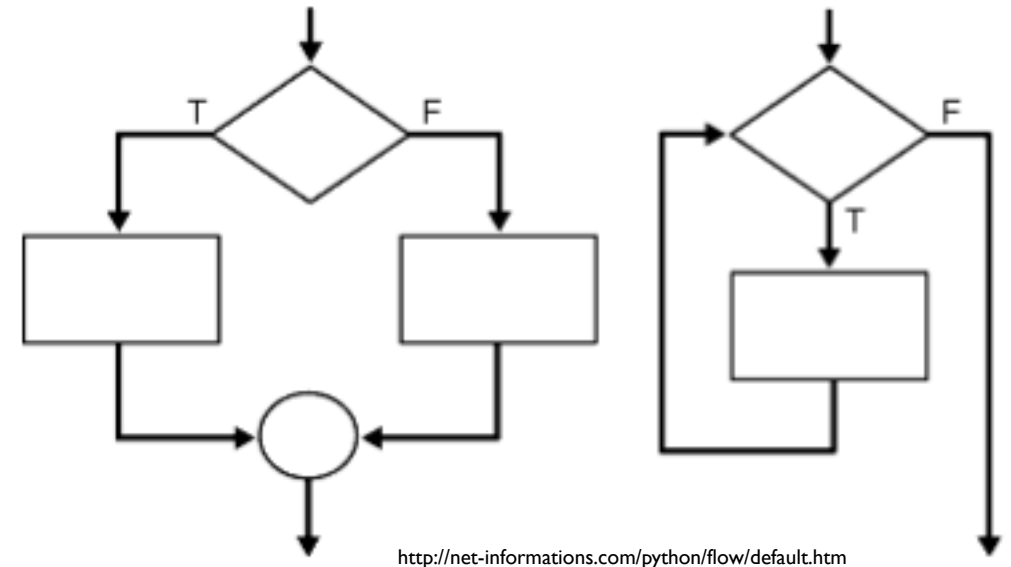
Course overview part 2

Canvas and course materials

220 Topics

Part 1: Control Flow

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



Part 2: State

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



Part 3: Data Science

- Tabular data
- Internet
- Databases
- Plotting



Today's Topics

Introductions

Course overview part 1

Worksheet: Pseudocode

Course overview part 2

Canvas and course materials

Pseudocode

Today's Topics

Introductions

Course overview part 1

Worksheet: Pseudocode

Course overview part 2

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

Canvas and course materials

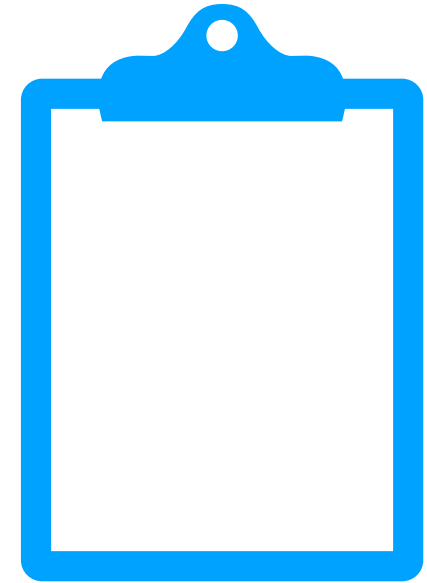
Lectures



general concepts



live coding



worksheet practice

Your role

- Do **readings** before or after
- Download the **template file** and code along in lecture
- Ask **questions** during the lectures + office hours
- Lectures **WON'T** be recorded



Labs – CS220

Format

- Attendance is mandatory – 3 out of 13 labs will be dropped for grading
- 75 minutes on Tuesdays and Thursdays
- Led by Teaching Assistant (TA) and a Peer Mentor (PM)
- Lab document will be posted each week on Wednesday
- Meant to help you succeed on your project

Partnership

- We strongly encourage you to find a lab / project partner
- If you chose to do lab with a partner, make sure they are your project partner

we will have labs starting tomorrow!!!

(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

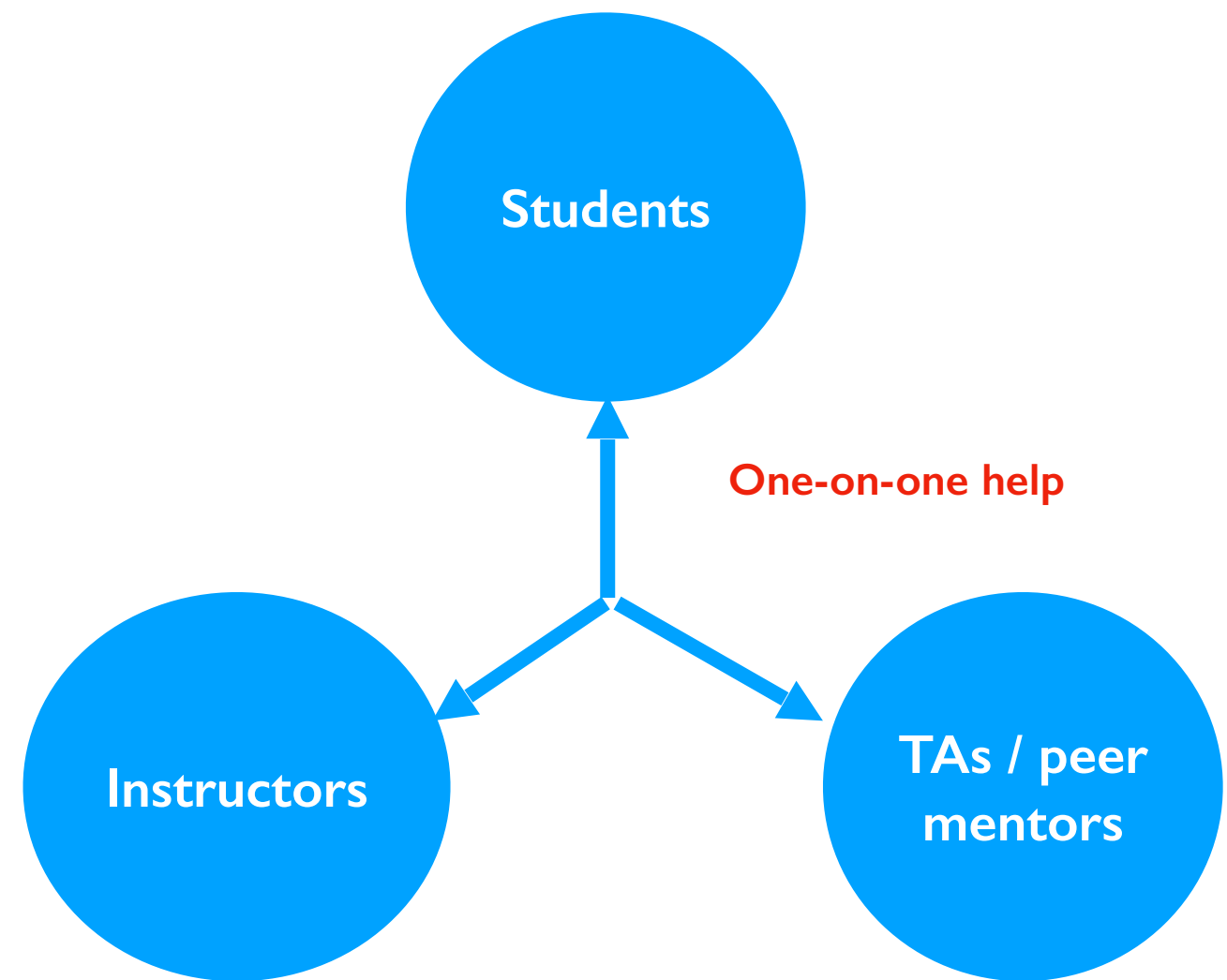
Course Notes

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

Course tools

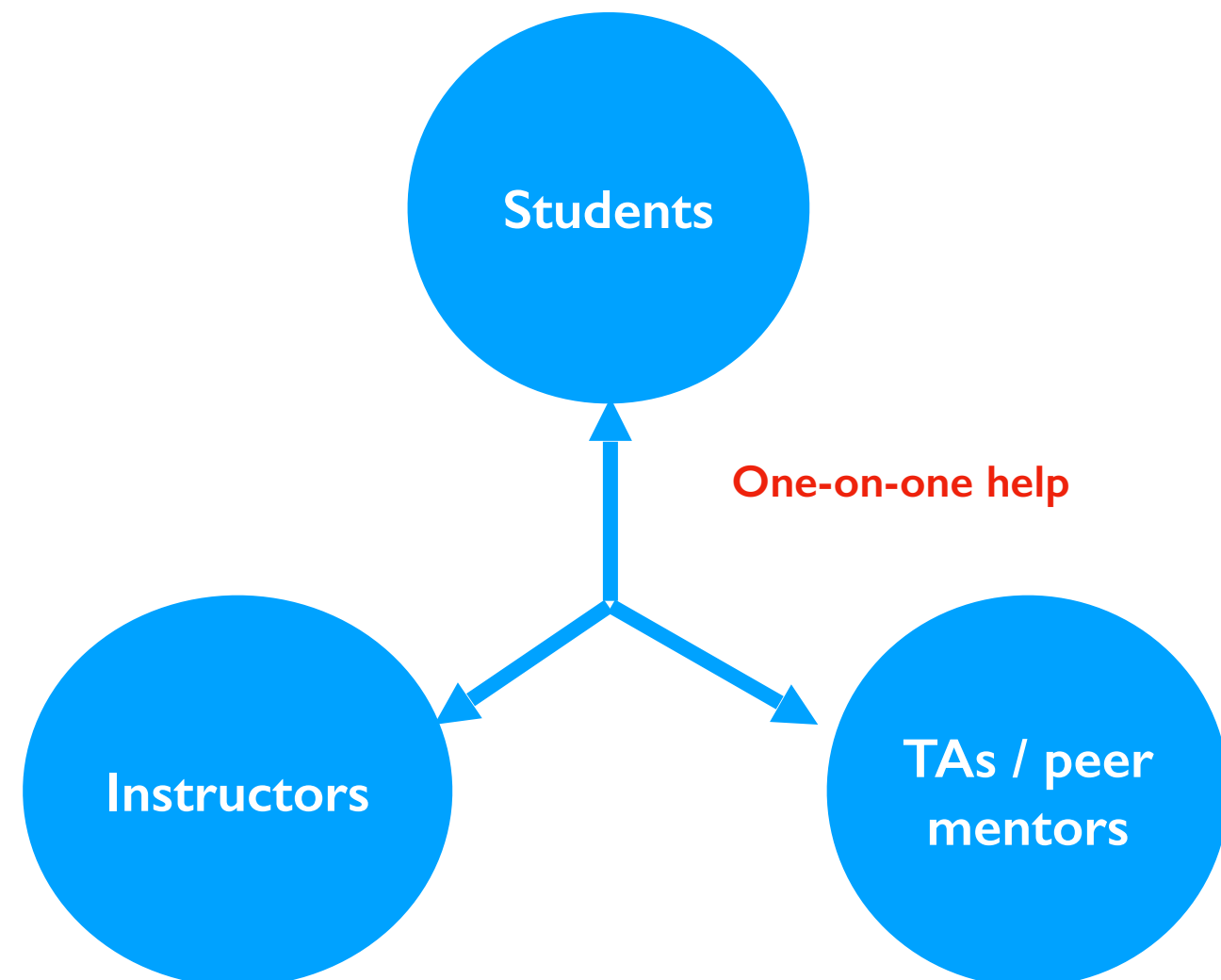
Communication tools

- Office hours (best way to get help):
 - See schedule on Canvas
- Canvas:
 - Announcements
 - Course schedule
 - Quizzes
 - Grades
- Piazza
 - Question asking and answering
 - Communication with course staff (use a private message)



Course tools

- Piazza:
 - Rule 1: don't post more than 5 lines of code
 - Rule 2: check other posts before posting
- Project Submission: GradeScope



Grades - CS220

48% - programming projects

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

30% - exams

- 3 “midterms”
- 10% each

16% - quizzes

- 10 quizzes (drop 2 lowest scores)

5% - lab attendance

- 13 labs (drop 3 lowest scores)

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% - 74.99%: C
- 60% - 69.99%: D

Today's Topics

Introductions

Course overview part 1

Worksheet: pseudocode

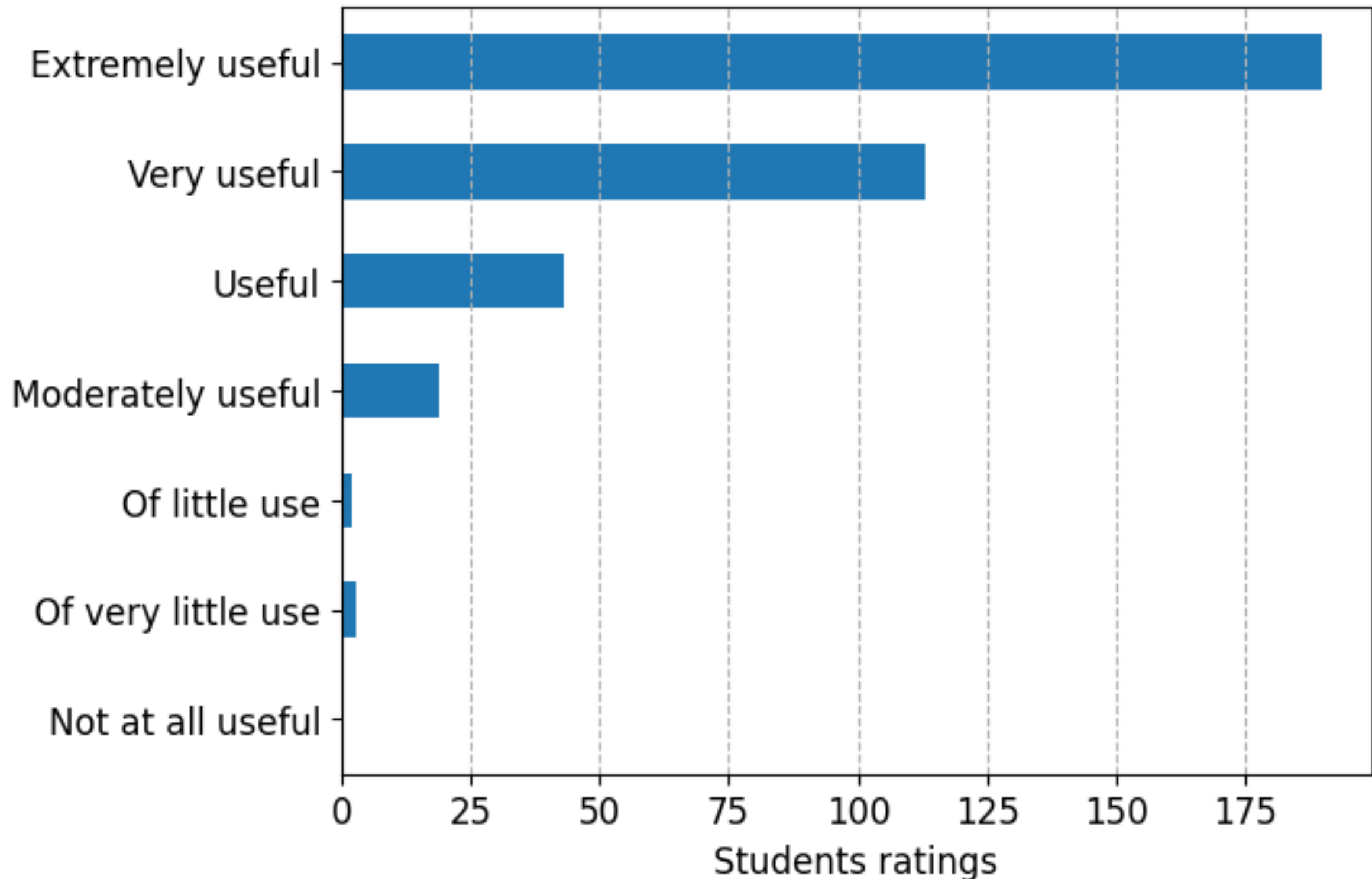
Course overview part 2

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

Canvas and course materials

Prior student reaction to projects

Projects: How useful were projects to your learning?



Project Overview

Nearly all projects will relate to some dataset

Timeline

- Projects will be due on **Tuesdays and Fridays at 11:59:00 pm**
- You get a bank of 8 late days, but can use only 2 on one project
- After late days, 5% deduction per day late
 - 7 days after the project deadline, project submission won't be accepted

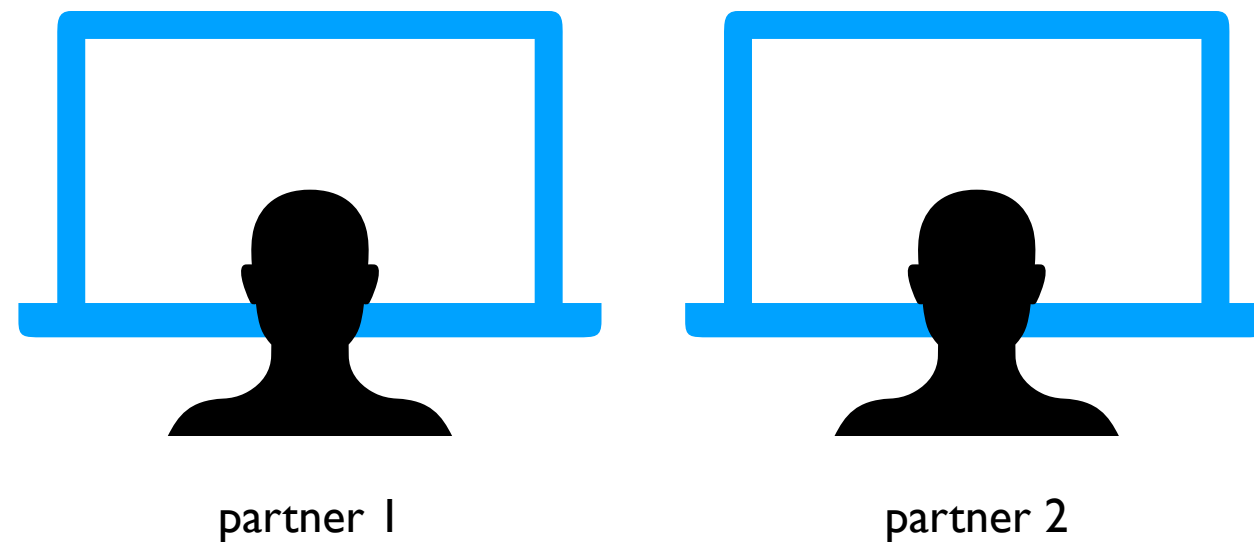
Getting help

- Office hours
- Lab sessions
- Piazza

Pair Programming

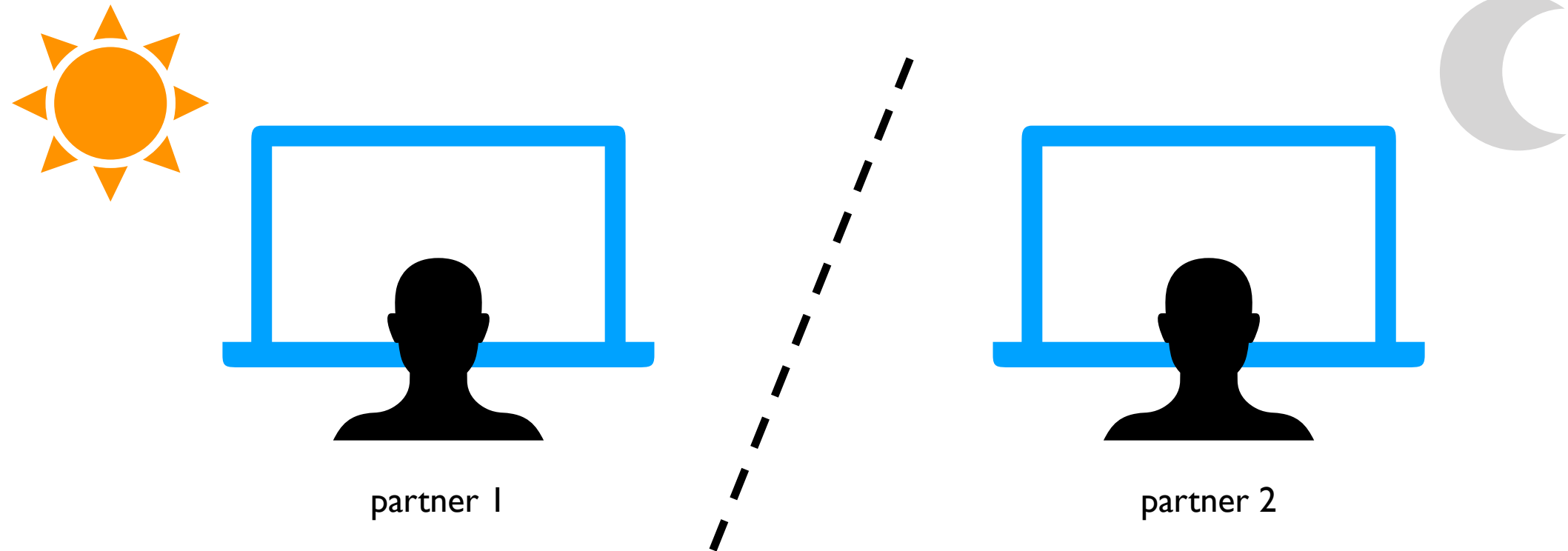
You can optionally work in pairs of two

- Students can partner with students from any lab section
- You can choose to keep the same partner, for multiple projects or choose to switch partners



Best practice: working alongside each other

Pair Programming

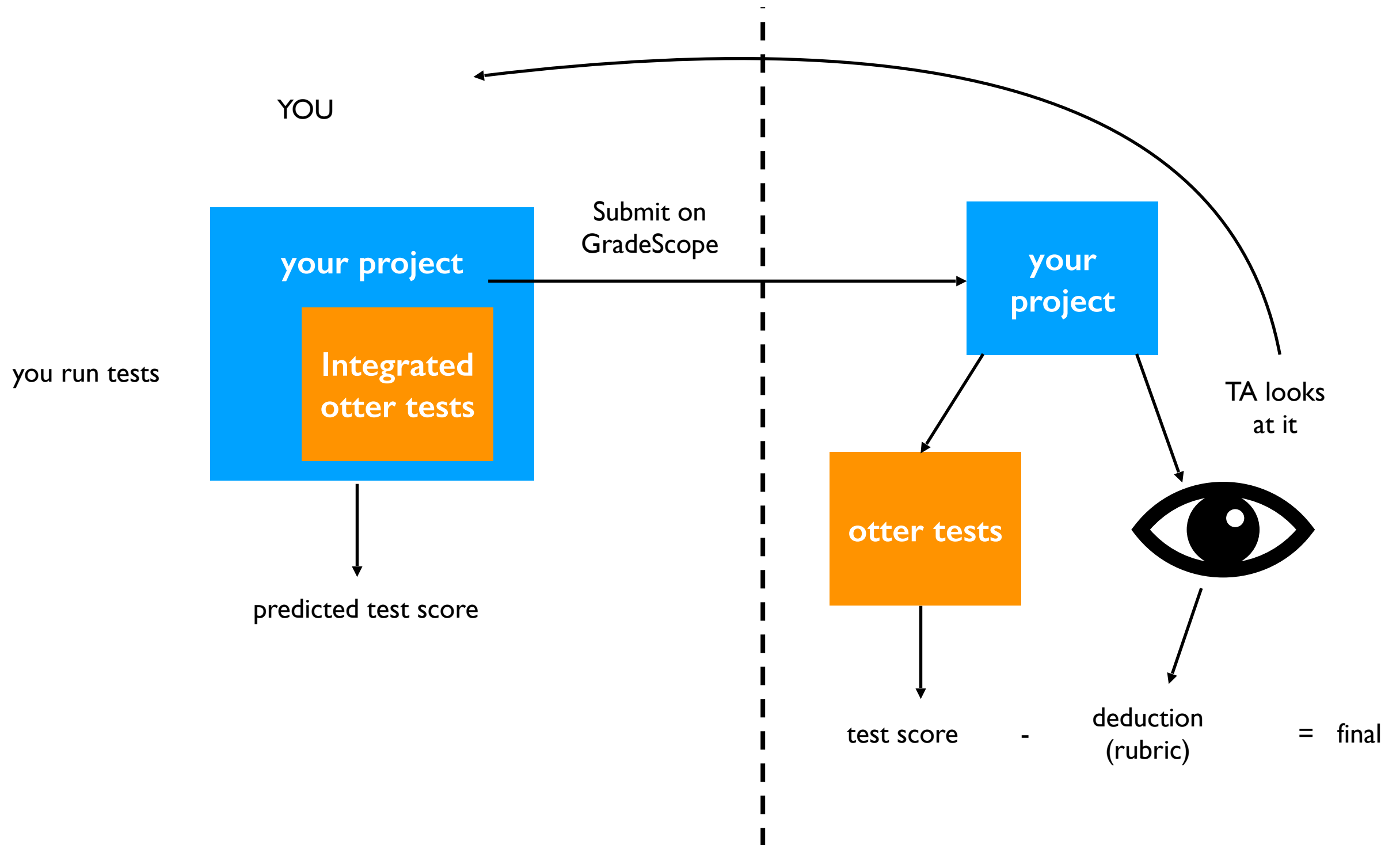


Breaks syllabus rules: working on different parts at different times

Breaks syllabus rules: working on alternate projects individually

Project Grading

feedback is mostly about how
to do things better or more simply
(valuable even if you score 100%)



Make sure that your code
clear tests on GradeScope

Today's Topics

Introductions

Course overview part 1

Worksheet: pseudocode

Course overview part 2

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

Canvas and course materials

Pseudocode

Quizzes and Exams

Quizzes

- 10 quizzes total, due 11:59PM on Mondays, Wednesdays, or Fridays.
 - Typically one quiz every 2 lectures (but there's some variation to work around exams, so pay attention to announcements in lecture)
- Quizzes are on Canvas and you can use any resource you want, except other people or AI
- Each quiz has 2 attempts and the higher score counts
- Focus on recent lectures so you stay current and check your knowledge

Exams: three “midterms”

- Multiple choice, closed-book
- Exams are cumulative but focused on more recent material
- All exams 75 minutes
- Exams will be held these dates
 - July 7 — in class on Friday from **9:50am - 11:05am**
 - July 27 — during lab
 - August 10 — during class from **9:50am - 11:05am**

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 part 1

Worksheet: pseudocode

Course overview part 2

Canvas, Piazza, and online course materials

Canvas

- Syllabus
- Course schedule (readings and due dates)
- Quizzes
- Surveys
- Grades

Piazza

- Need to have cross-site cookies enabled to see Piazza embedded in Canvas, if you prefer, you can also access Piazza directly at <https://piazza.com/class/liabmpaecey579>
- Used to ask and answer questions
- Do NOT post code snippets longer than 5 lines on Piazza (if in doubt, make the post private)

GitLab

Contains all files that you need for this course

- Class notes
- Lab and project materials
- Additional resources: readings and old practice exams

Next steps...

- Complete the "Student Information Survey" survey quiz on canvas.
- Read syllabus carefully
- Setup Anaconda (Python) on your computer and attend your lab for Lab-P I
- Submit P I (Project I) after attending lab: due this Friday
- Sign up for GradeScope and Piazza