

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.)



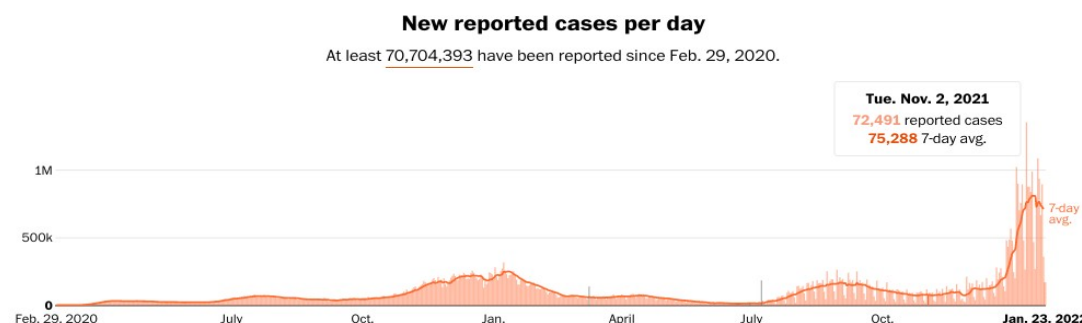
<https://fivethirtyeight.com/features/the-midwest-is-getting-drenched-and-its-causing-big-problems/>



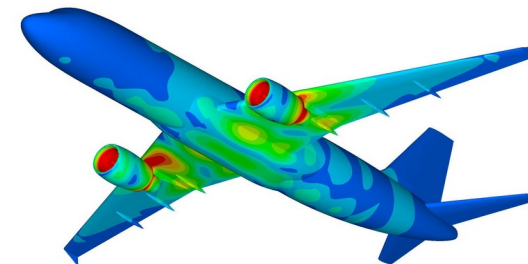
<https://en.wikipedia.org/wiki/Neuroimaging>



<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 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



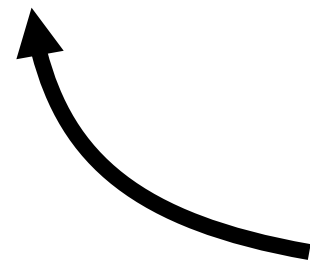
<http://fortune.com/2015/11/15/intel-super-7/>

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

<https://ourworldindata.org/>

<https://sdg-tracker.org/>

Today's Topics

Introductions

- Who are the faculty? Who are you?

Course overview

Website

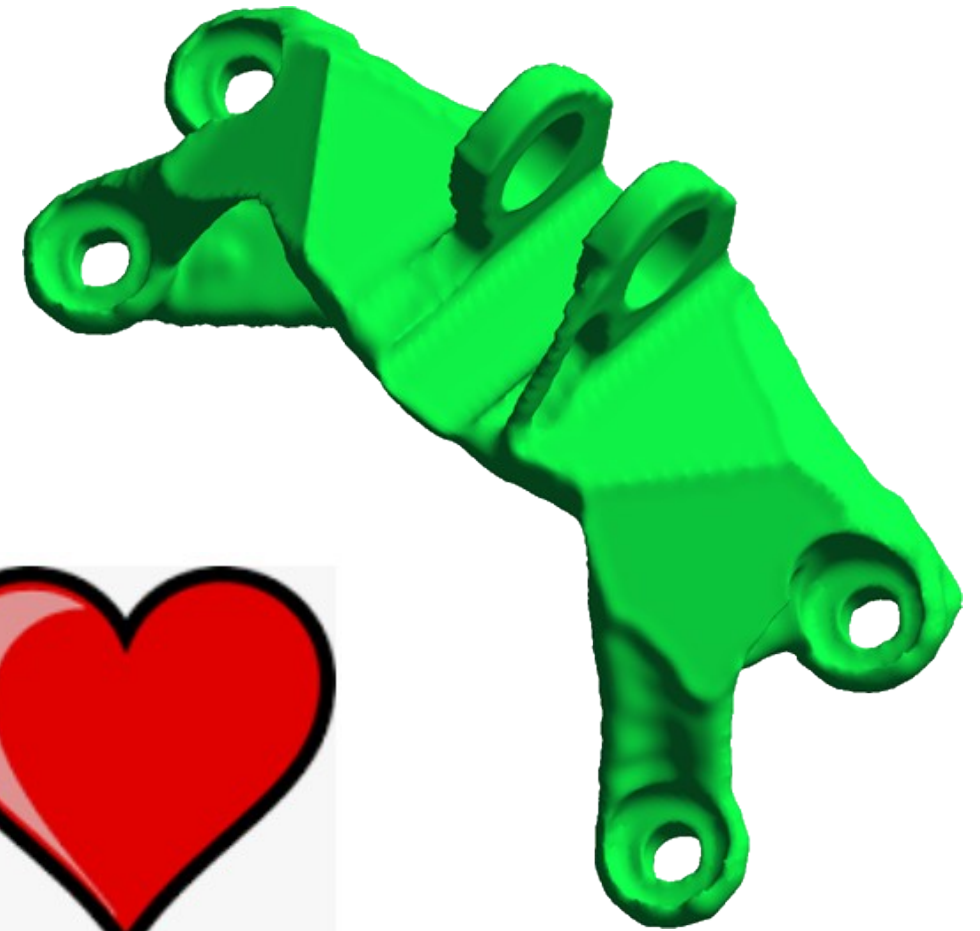
Who are the faculty?

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

Email: ltoliphant@wisc.edu

BA in Mathematics Education @ BYU, 1995

Taught High School for several years

MS and PhD in Computer Science @ UW-Madison

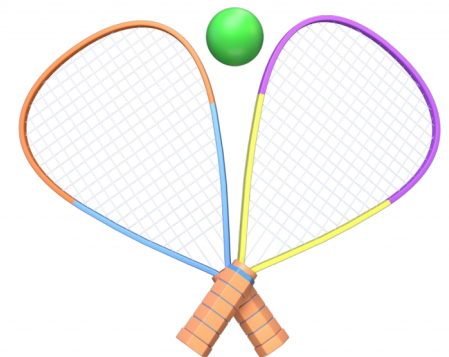
Dissertation in Machine Learning, 2009

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

Taught at Hiram College in Ohio from 2009-2013

Started at UW-Madison in Fall, 2023

Always looking for racquetball partners



Who are you?

- 1200 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

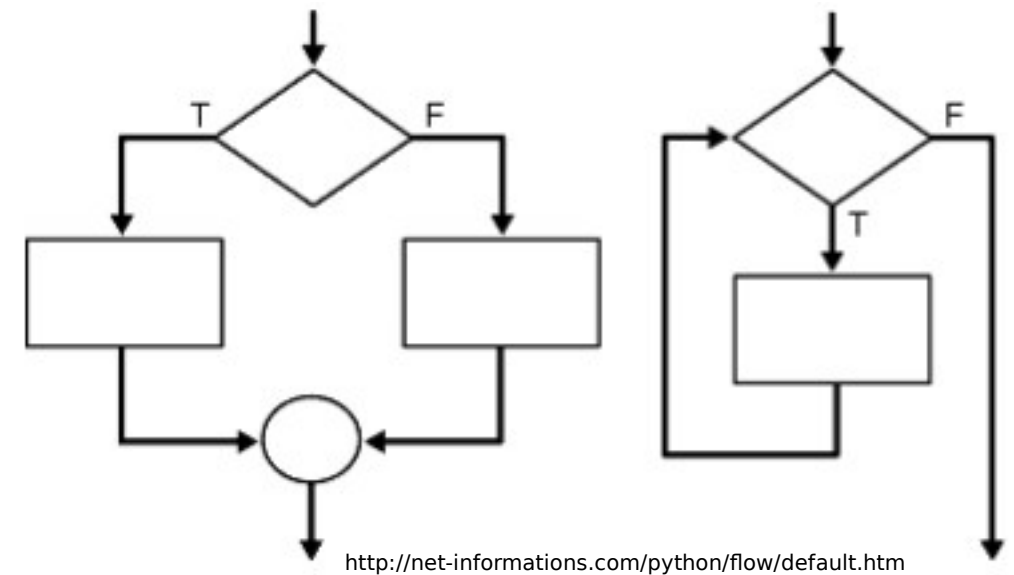
- Topics
- Lectures
- Lab
- Readings
- Course tools
- Grades
- Projects
- Exams & quizzes

Website

220 Topics

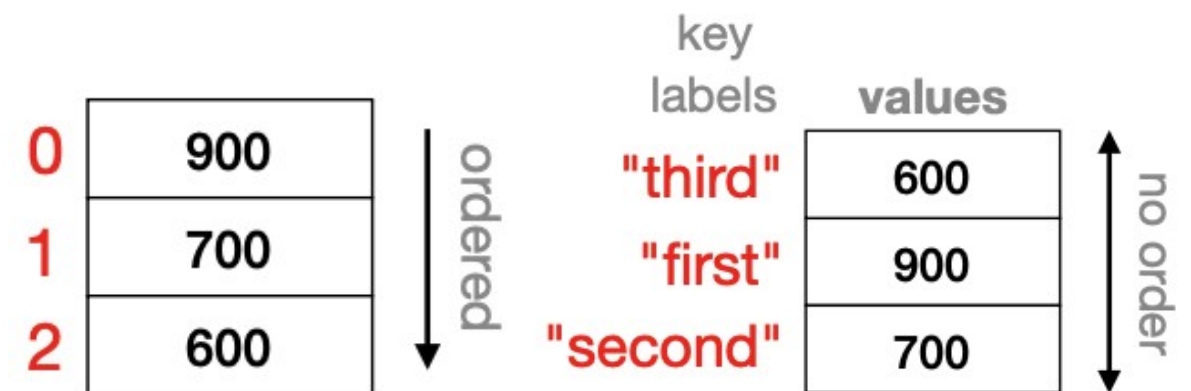
Part 1: Control Flow

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



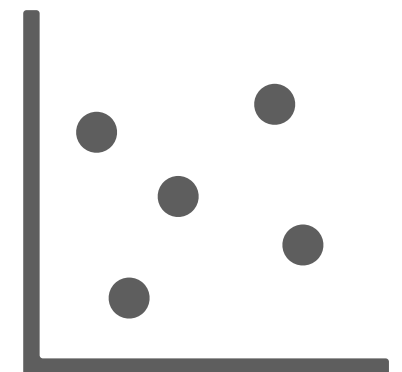
Part 2: State

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



Part 3: Data Science

- Tabular data
- Internet
- Databases
- Plotting



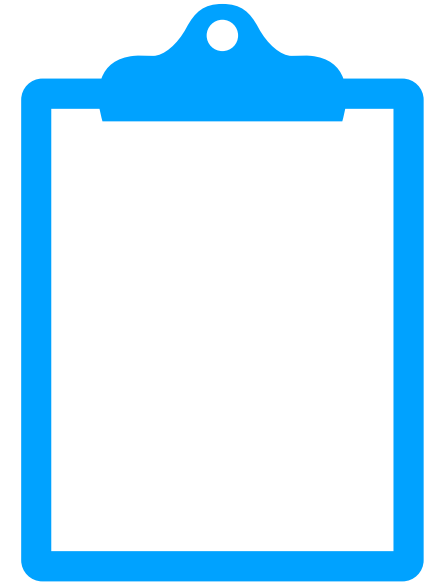
Lectures



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

- 13 labs
- 75 minutes on Wed, Thu or Fri
- 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
- CS220 students can partner with CS220 students
- If you chose to do lab with a partner, make sure they are your project partner

we will have labs this first week!!!

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

Labs – CS319

Format

- Self-guided
- Utilize Office Hours to get questions answered

Partnership

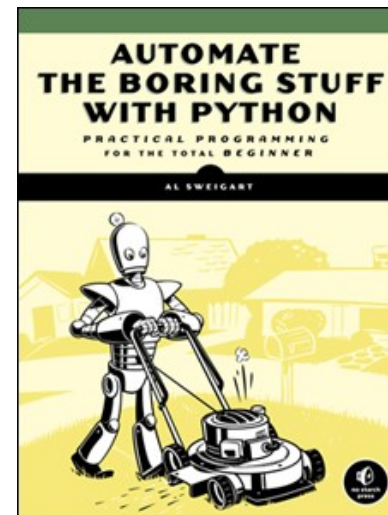
- CS319 students can partner with CS319 students
- If you chose to do lab with a partner, make sure they are your project partner

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

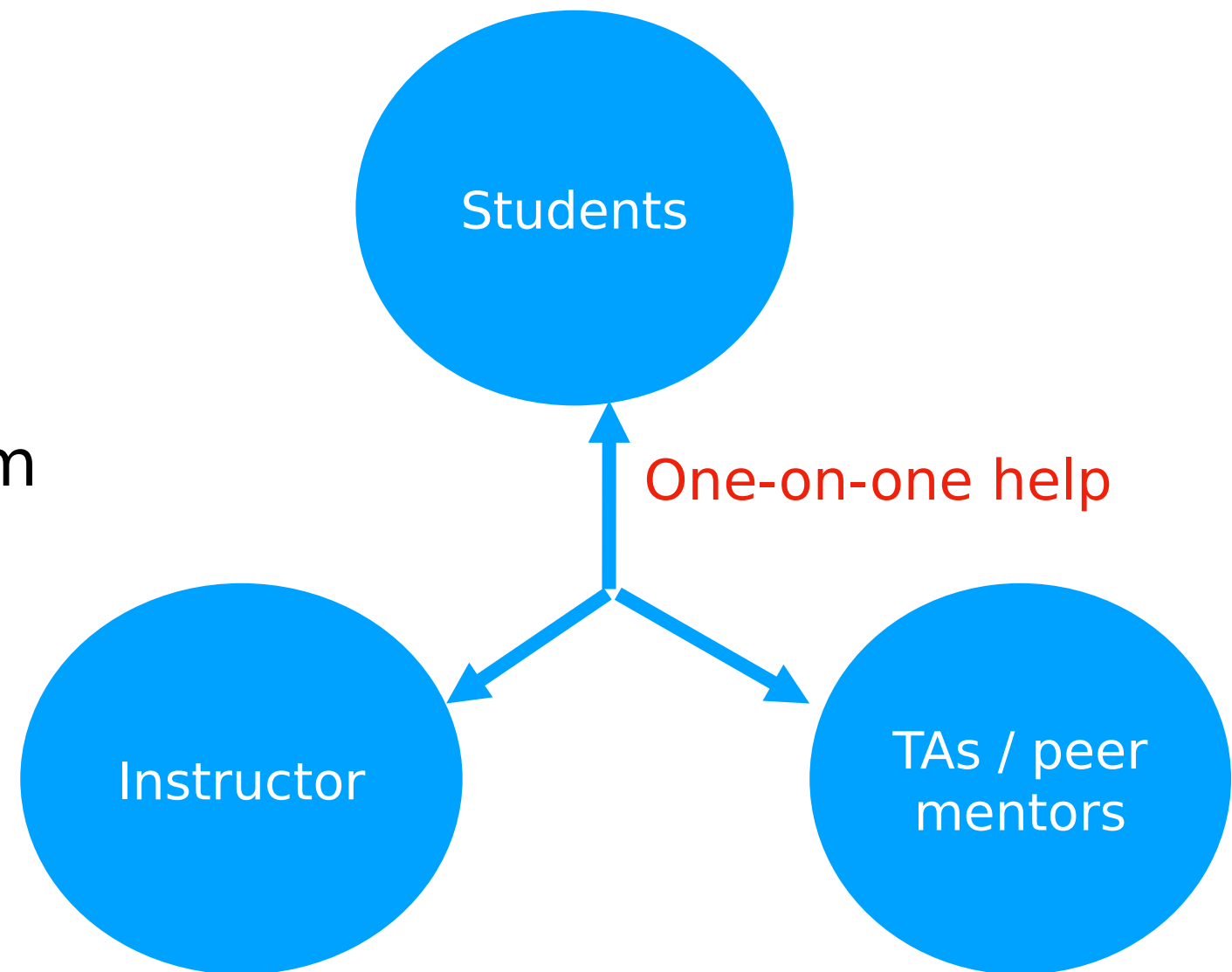
Course Notes

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

Course tools

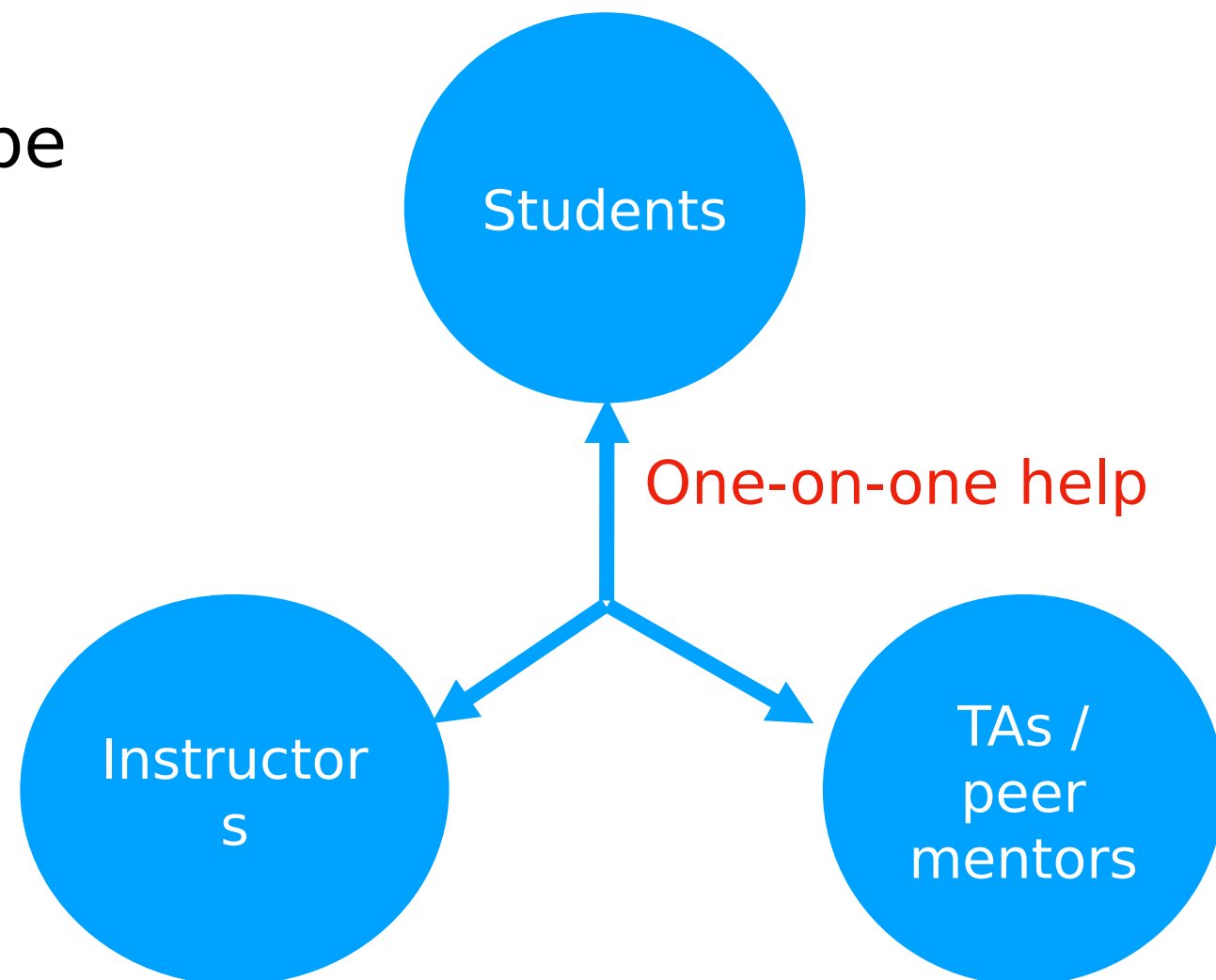
Communication tools

- 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)

2% - labs

- P1 – P9
- Self-guided

20% - graduate-level project

- Proposal: 4%
- Data Set: 4%
- Initial Code: 4%
- Final Code: 4%
- Presentation: 2%

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

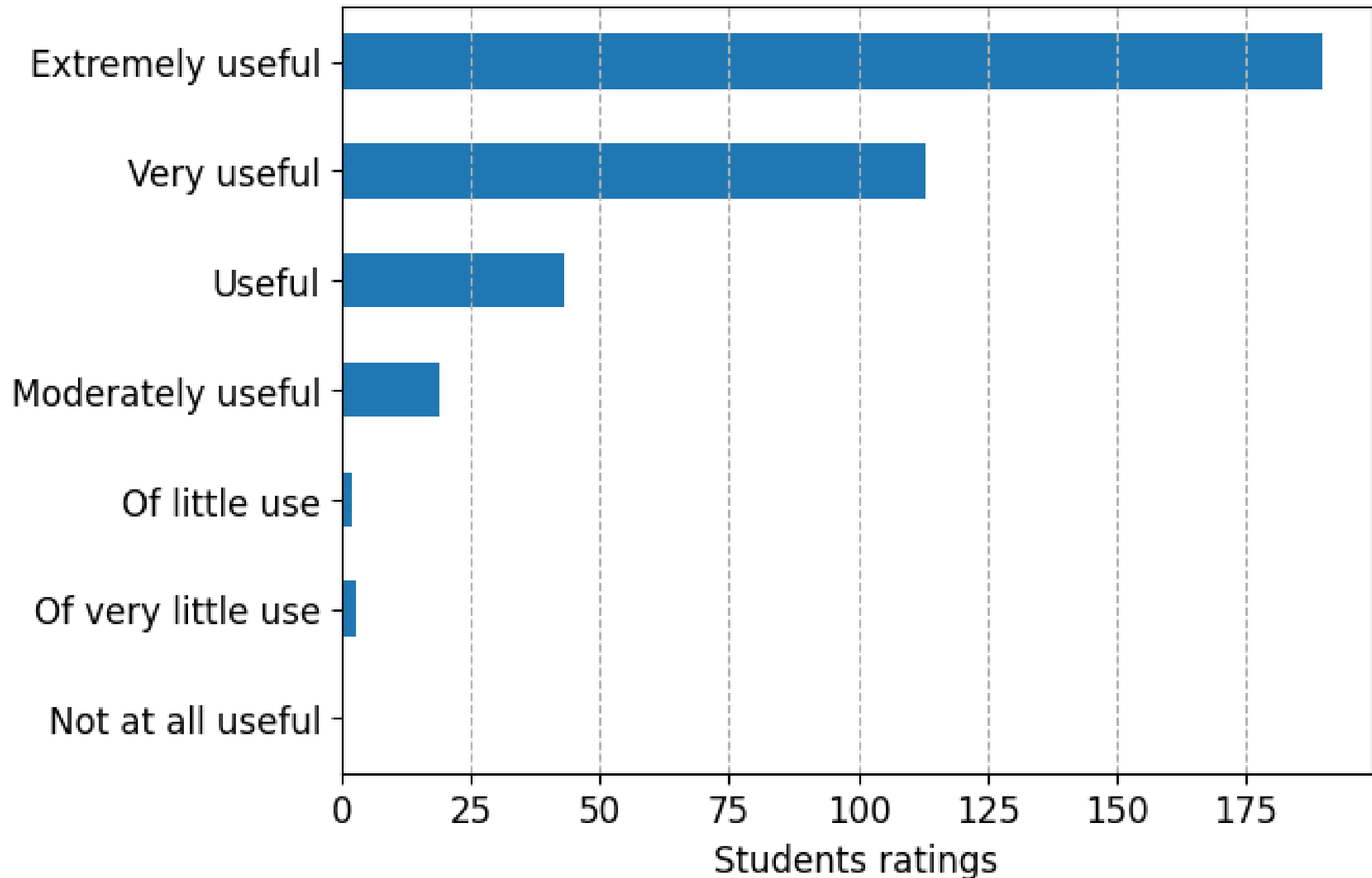
Course overview

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

Website

Prior student reaction to projects

Projects: How useful were projects to your learning?



Project Overview

Nearly all projects will relate to some dataset

Timeline

- Projects will be due on **Wednesdays at 11:59:00 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)

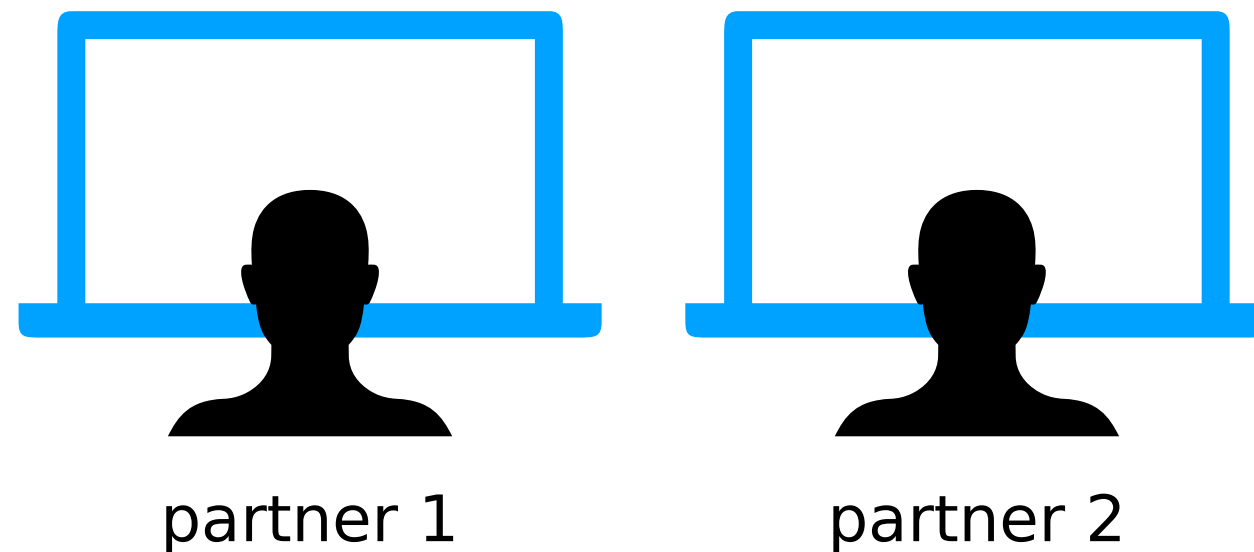
Getting help

- Office hours
- Lab sessions
- Piazza

Pair Programming

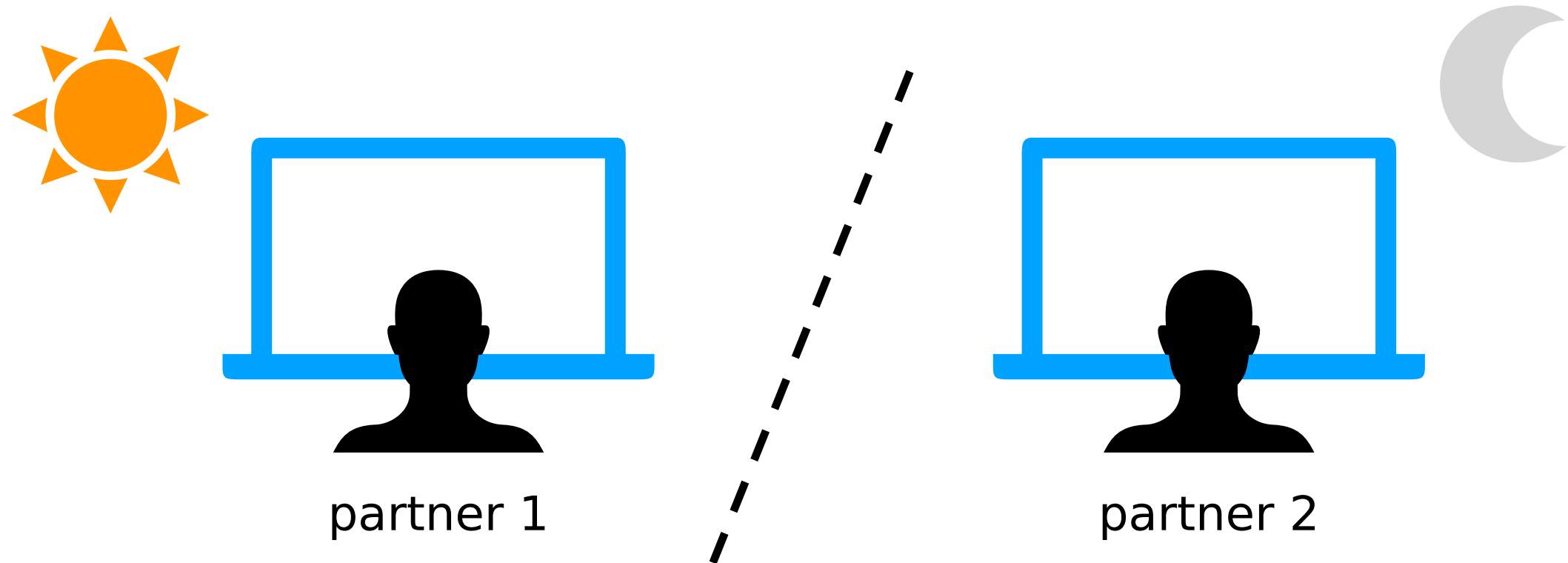
You can optionally work in pairs of two

- 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

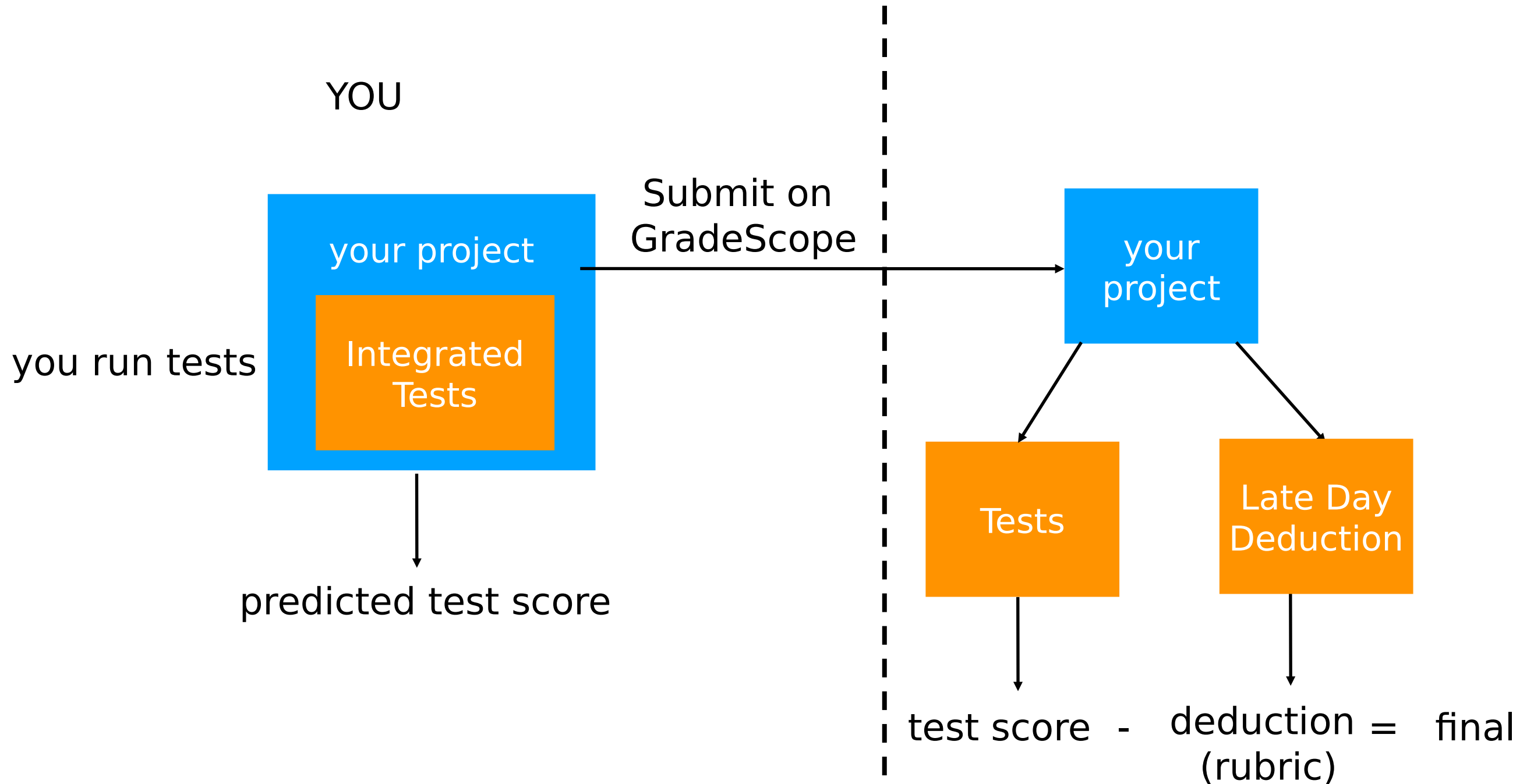
Pair Programming



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

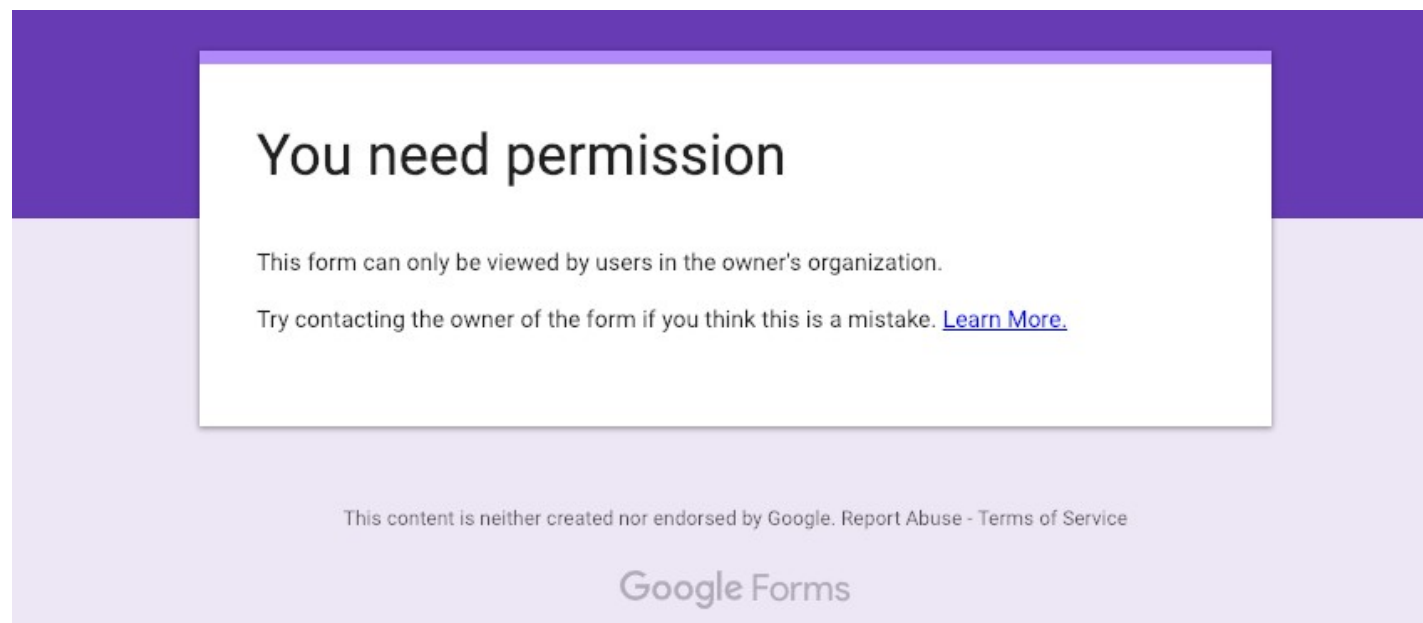
Website

Course Website

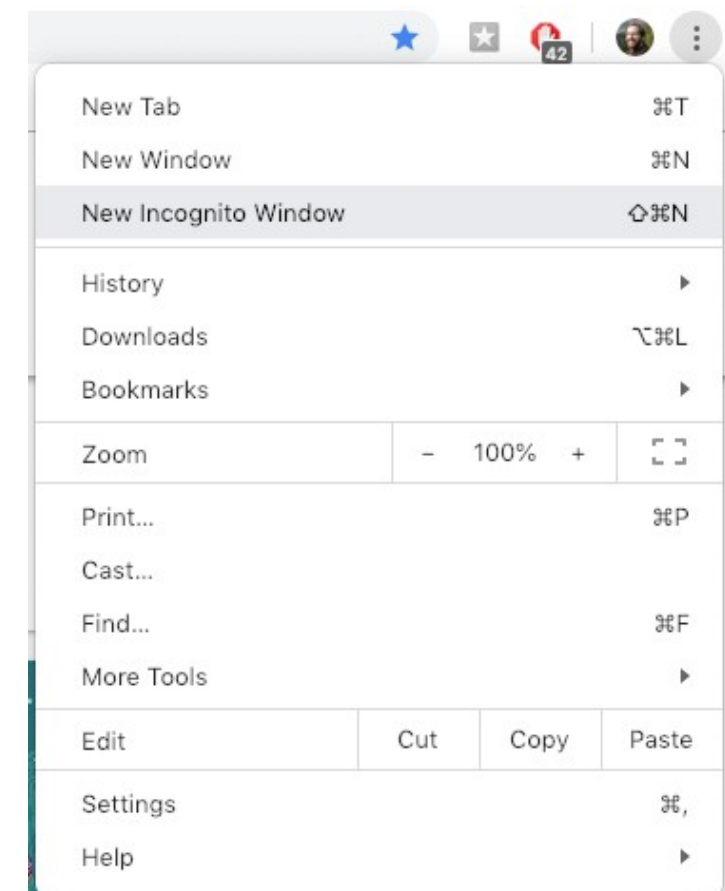
<https://cs220.cs.wisc.edu/f24/schedule.html>

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 Anaconda (Python) on your computer and attend your lab for Lab-P1
- Submit P1 (Project 1) after attending lab: due next Wed
- Sign up for GradeScope