

Data Science for Everyone – Welcome!

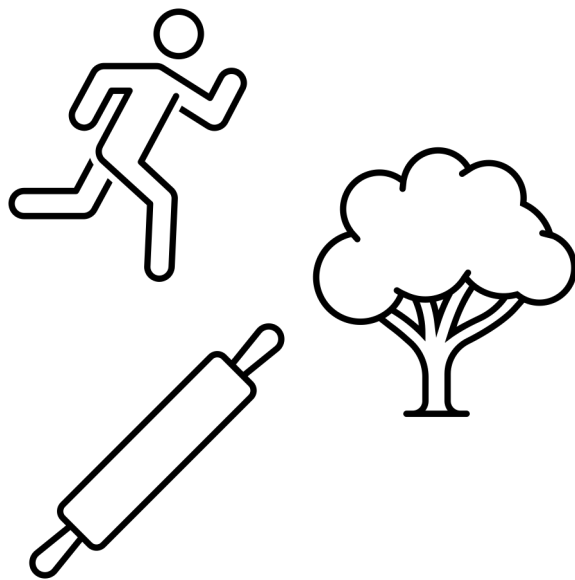
Dr. Ab Mosca (they/them)

Slides based off slides courtesy of Jordan Crouser (<https://jcrouser.github.io/>)

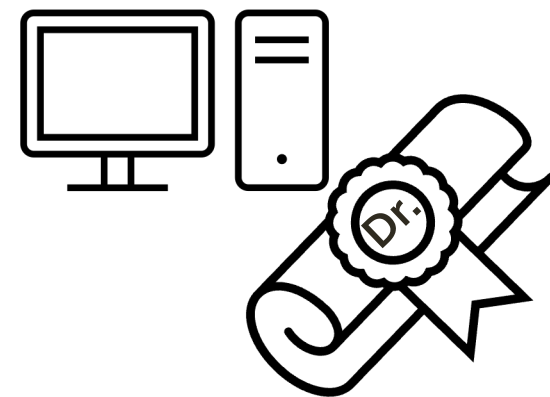
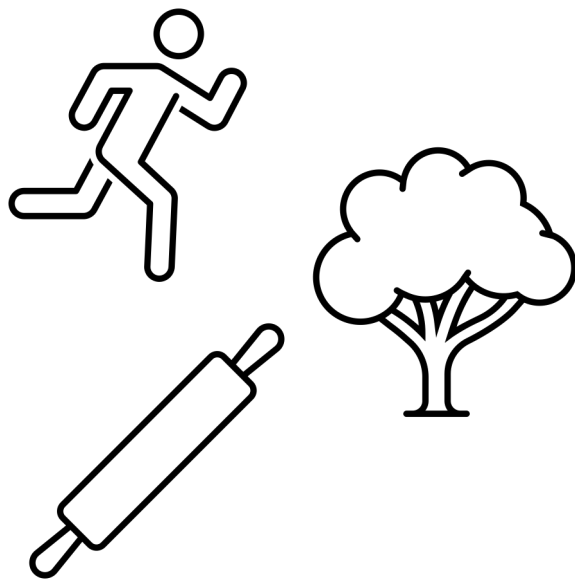
Plan for Today

- Who am I?
- Who are you?
- What will we do in this class?
- What is Data Science?

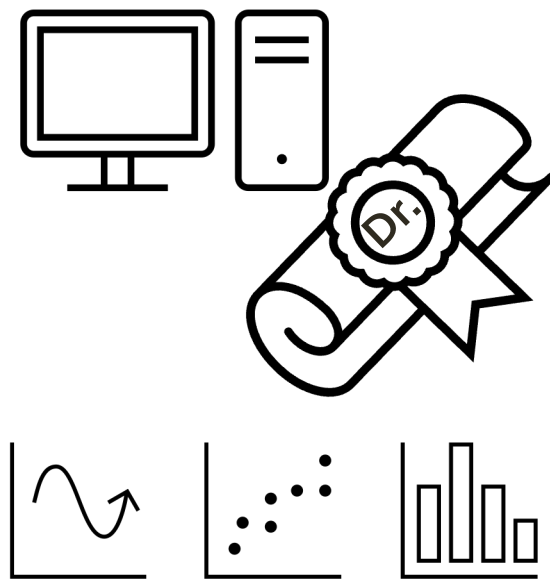
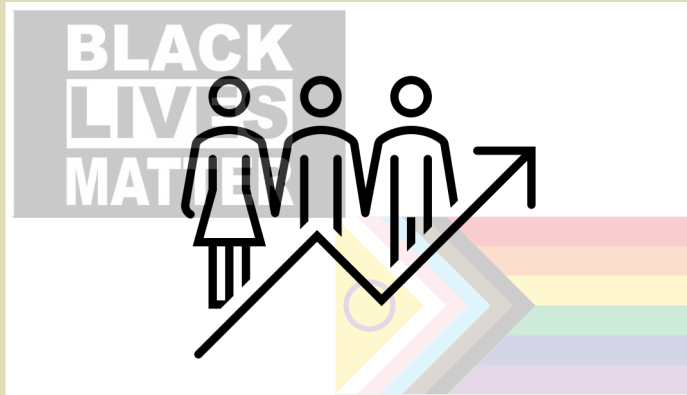
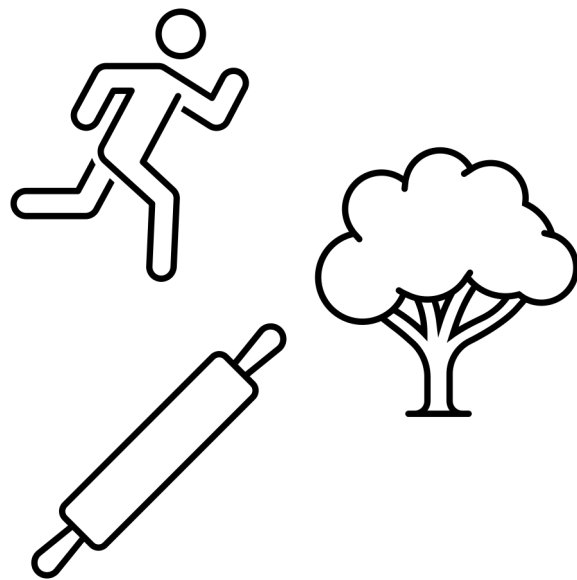
Who Am I?



Who Am I?



Who Am I?



Who Are You?

- Form groups of 3
- Introduce yourselves (name, pronouns)
- Share:
 - A highlight of your summer
- Find 1 thing that your entire group has in common (favorite color? hometown? left-handed? Be creative!)
- After about 5 minutes we will go around, introduce ourselves, and share what each group has in common

Who Are You?

- Form **new groups** of 3 (move around!)
- Introduce yourselves (name, pronouns)
- Share:
 - Would you rather be able to read any book in seconds OR remember everything you ever read perfectly?
- Find 1 thing that your entire group has in common (favorite color? hometown? left-handed? Be creative!)
- After about 5 minutes we will go around, introduce ourselves, and share what each group has in common

Who Are You?

- Form **new new groups** of 3 (move around!)
- Introduce yourselves (name, pronouns)
- Share:
 - Would you rather have school 7 days a week for 6 months + 6 months off OR have school 3 days a week year-round?
- Find 1 thing that your entire group has in common (favorite color? hometown? left-handed? Be creative!)
- After about 5 minutes we will go around, introduce ourselves, and share what each group has in common



What You Will Learn & Logistics

What Is This Class?

- An introduction to the field of Data Science assuming no prior knowledge of the subject
- You will learn...
 - How to do data science (basic techniques and foundations)
 - How data science shows up in different fields (public health, city planning, etc.)
 - How to be a critical consumer of data science

****Important Info****

- Course website (**write this down!**):
<https://amoscao1.github.io/Intro-DS-F23/>
- PLATO: please use for all course related communication
- OH's: T/R 11:30 – 13:30 (stop by and say hi!)

****Important Info****

- Textbook: *Modern Data Science with R*, 2nd Edition
 - Available free of charge here: <https://mdsr-book.github.io/mdsr2e/>
- Assignments:
 - Turn in on PLATO
 - Homeworks – largely effort based
 - Quizzes – open book / notes / google
 - Projects – group based, application of skills
- Due Dates: As listed on course schedule.
 - 24hr grace period; no late submissions
 - Lowest homework and quiz dropped
 - No regrades; see syllabus for revise and resubmit instructions

****Important
Info****

- I'm here to help you succeed
- Please come to office hours or reach out on PLATO if you need any additional support

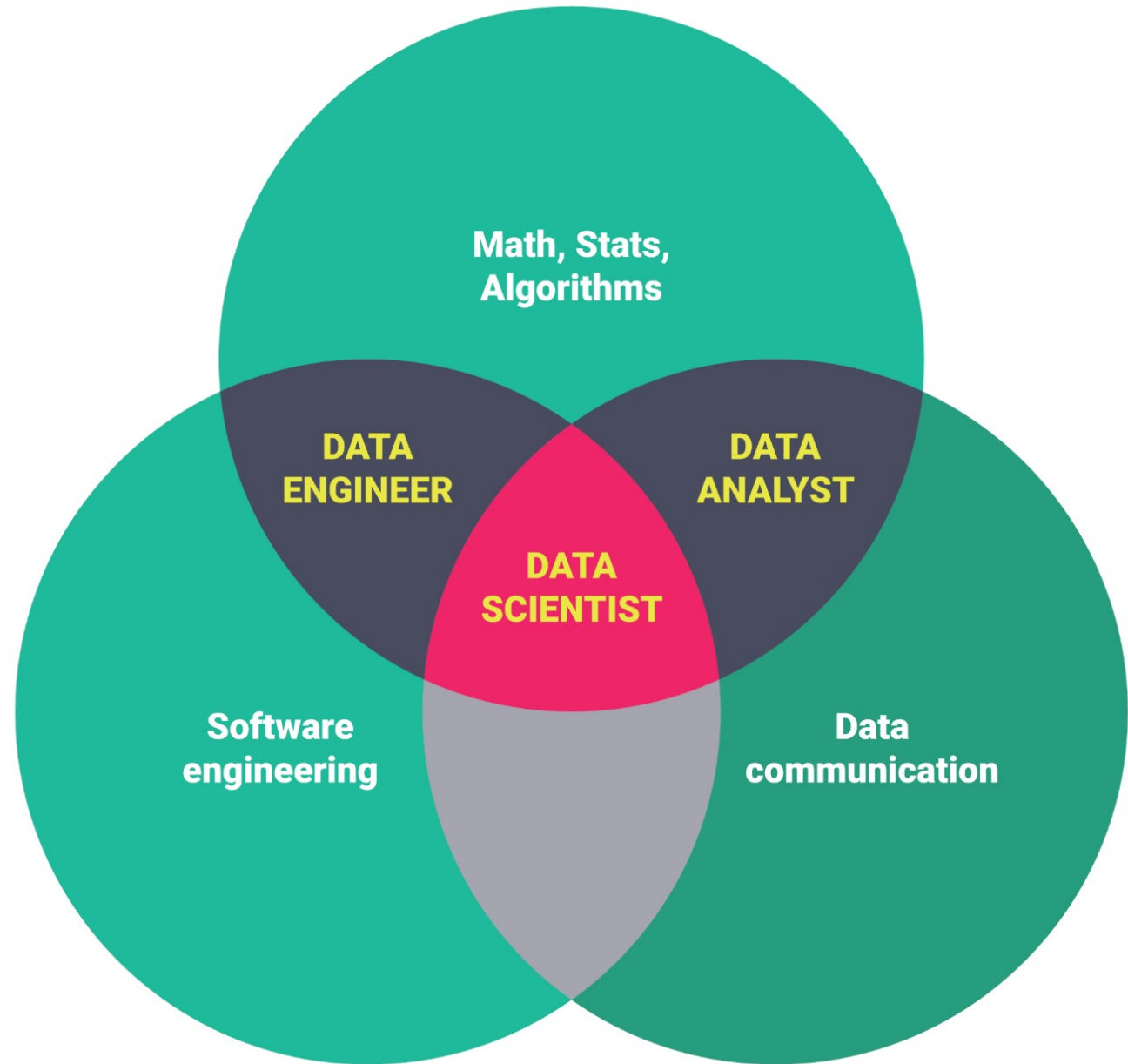


Now the good stuff

Defining Data Science

- How would you **define data science**?
- Talk it out with the person sitting next to you then add your definition to the board

Definition: Data Science




Defining Ethics

- How would you **define ethics**?
- Talk it out with the person sitting next to you then add your definition to the board

Definition: Ethics

Definition of *ethic*

- 1 a : a set of moral principles : a theory or system of moral values
// the present-day materialistic *ethic*
// an old-fashioned work *ethic*
—often used in plural but singular or plural in construction
// an elaborate *ethics*
// Christian *ethics*
- b **ethics** \ 'e-thiks  \ plural in form but singular or plural in construction : the principles of conduct governing an individual or a group
// professional *ethics*
- c : a consciousness of moral importance
// forge a conservation *ethic*
- d : a guiding philosophy

<https://www.merriam-webster.com/dictionary/ethic>

Defining Ethical Data Science

- How would you **define ethical data science**?

Definition: Ethical Data Science

- Applying theories and systems of moral values (ethics) to our data science work

Ethical Data Science

- Applying theories and systems of moral values (ethics) to our data science work
- Why is this important?

Ethical Data Science

- Applying theories and systems of moral values (ethics) to our data science work
 - Why is this important?
-



Ethical Data Science

- Applying theories and systems of moral values (ethics) to our data science work
- Why is this important?

Can you think of other examples of things that can go wrong with data science?

Ethical Data Science

- Applying theories and systems of moral values (ethics) to our data science work
- Why is this important?
 - To combat algorithmic bias
 - To protect personal, identifying information
 - To increase reproducibility and replicability
- Systems of power are woven into our society, and so are misunderstandings of what data science is and how it works